Top-Down Network Design Third Edition A systems analysis approach to enterprise network design Identifying Your Customer's Needs and Goals Chapter 1 Analyzing Business Goals and Constraints This chapter serves as an introduction to the rest of the book by describing top-down network design. The first section explains how to use a systematic, top-down process when designing computer networks for your customers. Depending on your job, your customers might consist of other departments within your company, those to whom you are trying to sell products, or clients of your consulting business. After describing the methodology, this chapter focuses on the first step in top-down network design: analyzing your customer's business goals. Business goals include the capability to run network applications to meet corporate business objectives, and the need to	Thiết kế mạng theo trình tự từ trên xuống Tái bản lần ba Phương pháp phân tích hệ thống trong quá trình thiết kế mạng doanh nghiệp Xác định nhu cầu và mục tiêu của khách hàng Chương 1 Phân tích các mục tiêu kinh doanh và những giới hạn (ràng buộc) Chương này sẽ trình bày kiến thức nền tảng cho phần còn lại của sách, đó là khái niệm thiết kế mạng theo trình tự từ trên xuống. Trong phần đầu, chúng tôi sẽ nói rõ cách dùng một quy trình có hệ thống, theo trình tự từ trên xuống để thiết kế các mạng máy tính cho khách hàng của bạn. Tùy thuộc vào nhiệm vụ của bạn, khách hàng có thể là những người thuộc những bộ phận khác trong công ty bạn, những người mà bạn đang cố bán sản phẩm, hoặc những khách hàng cần tư vấn. Sau khi mô tả cơ sở lý thuyết của phương pháp, chương này tập trung vào bước đầu tiên trong phương pháp thiết kế mạng từ trên xuống: phân tích mục tiêu kinh doanh của khách hàng. Các mục tiêu kinh doanh bao gồm khả năng chạy các ứng dụng mạng để đáp ứng các mục tiêu kinh doanh

work within business constraints, such as budgets, limited networking personnel, and tight timeframes.	của công ty, và nhu cầu làm việc trong những điều kiện kinh doanh (những ràng buộc kinh doanh), chẳng hạn như ngân sách, nguồn nhân lực mạng hạn chế, và khung thời gian chặt chẽ.
This chapter also covers an important business constraint that some people call the eighth layer of the Open System Interconnection (OSI) reference model: workplace politics. To ensure the success of your network design project, you should gain an understanding of any corporate politics and policies at your customer's site that could affect your project.	Chương này cũng đề cập đến một ràng buộc kinh doanh quan trọng mà một số người gọi là lớp thứ tám của mô hình tham chiếu Kết Nối Các Hệ Thống Mở (OSI): các chính sách ở nơi làm việc. Để đảm bảo sự thành công trong dự án thiết kế mạng của bạn, bạn cần phải có những hiểu biết về chính sách và chính kiến ở nơi khách hàng của bạn đang làm việc, những yếu tố này có thể ảnh hưởng đến dự án của bạn.
The chapter concludes with a checklist to help you determine if you have addressed the business issues in a network design project.	Ở cuối chương, chúng tôi cũng đưa ra một danh sách kiểm tra để giúp bạn xác định được mức độ nắm vững các vấn đề liên quan đến kinh doanh trong dự án thiết kế mạng của bạn.
Using a Top-Down Network Design Methodology According to Albert Einstein: "The world we've made as a result of the level of thinking we have done thus far creates problems that we cannot solve at the same level at which we created them."	Dùng phương pháp thiết kế mạng từ trên xuống Theo Albert Einstein: "Thế giới mà chúng ta tạo ra chính là kết quả của mức độ trừu tượng hóa trong suy nghĩ của chúng ta, cho đến nay, thế giới đó tạo ra những vấn đề mà chúng ta chưa thể giải quyết được nếu chúng ta đang

	ở mức bằng với mức mà chúng ta đã tạo ra nó.”
To paraphrase Einstein, networking professionals have the ability to create networks that are so complex that when problems arise they can’t be solved using the same sort of thinking that was used to create the networks. Add to this the fact that each upgrade, patch, and modification to a network can also be created using complex and sometimes convoluted thinking, and you soon realize that the result is a network that is hard to understand and troubleshoot.	Trong trường hợp cụ thể ở đây, chúng ta có thể hiểu như sau: những chuyên gia mạng có thể tạo ra những mạng quá phức tạp đến nỗi có những lúc sẽ có quá nhiều vấn đề phát sinh, và họ không thể giải quyết được những vấn đề đó nếu dùng cách suy nghĩ giống như cách suy nghĩ lúc thiết kế mạng. Ngoài ra, mỗi lần nâng cấp, vá lỗi và hiệu chỉnh mạng cũng có thể vô cùng khó khăn, và bạn sẽ sớm thấy rằng mạng khó kiểm soát và khắc phục sự cố.
A network created with this complexity often doesn’t perform as well as expected, doesn’t scale as the need for growth arises (as it almost always does), and doesn’t match a customer’s requirements. A solution to this problem is to use a streamlined, systematic methodology in which the network or upgrade is designed in a top-down fashion.	Mạng được tạo ra trong điều kiện phức tạp như vậy không thể hoạt động tốt như mong đợi, không thể mở rộng khi cần thiết (mà nhu cầu này lại rất phổ biến), và không phù hợp với các yêu cầu của khách hàng. Một giải pháp cho vấn đề này là dùng một phương pháp tổ chức hợp lý, có hệ thống trong đó việc thiết kế mạng hoặc nâng cấp được thực hiện theo kiểu từ trên xuống.
Many network design tools and methodologies in use today resemble the “connect-the- dots” game that some of us played as children. These	Nhiều công cụ và phương pháp luận thiết kế mạng đang dùng ngày nay giống với trò chơi “nối các điểm” mà một vài người trong chúng ta đã

tools let you place internetworking devices on a palette and connect them with LAN or WAN media. The problem with this methodology is that it skips the steps of analyzing a customer's requirements and selecting devices and media based on those requirements.	chơi từ thời thơ ấu. Những công cụ này cho phép bạn đặt các thiết bị liên mạng trên một bảng màu và kết nối chúng với môi trường LAN hoặc WAN. Nhược điểm của phương pháp này là nó bỏ qua các bước phân tích nhu cầu khách hàng và chọn lựa các thiết bị và môi trường dựa trên những yêu cầu đó.
Good network design must recognize that a customer's requirements embody many business and technical goals, including requirements for availability, scalability, affordability, security, and manageability. Many customers also want to specify a required level of network performance, often called a service level. To meet these needs, difficult network design choices and tradeoffs must be made when designing the logical network before any physical devices or media are selected.	Một nhà thiết kế mạng tốt phải nhận ra rằng yêu cầu của khách hàng là hiện thân của các mục tiêu kinh doanh và kỹ thuật, bao gồm các yêu cầu về tính sẵn có, khả năng mở rộng, tính hợp lý về giá cả, bảo mật và khả năng dễ quản lý. Nhiều khách hàng cũng muốn xác định mức hiệu suất mạng cần thiết, thường được gọi là mức dịch vụ. Để đáp ứng những nhu cầu này, chúng ta phải chọn lựa và cân bằng thiết kế mạng một cách nghiêm túc trong quá trình thiết kế mạng trước khi chọn bất kỳ thiết bị và môi trường vật lý nào.
When a customer expects a quick response to a network design request, a bottom-up (connect-the-dots) network design methodology can be used, if the customer's applications and goals are well known. However, network designers often think they understand a	Khi khách hàng muốn đáp ứng nhanh chóng với một yêu cầu thiết kế mạng, chúng ta nên dùng thiết kế mạng từ dưới lên (kết nối các dấu chấm), nếu đã biết các ứng dụng và mục tiêu của khách hàng. Tuy nhiên, các nhà thiết kế mạng thường nghĩ họ tìm hiểu các ứng dụng và yêu cầu của khách hàng chỉ

customer's applications and requirements only to discover, after a network is installed, that they did not capture the customer's most important needs. Unexpected scalability and performance problems appear as the number of network users increases. These problems can be avoided if the network designer uses top-down methods that perform requirements analysis before technology selection.	nhằm mục đích khám phá, sau khi mạng được lắp đặt, họ không nắm bắt được những nhu cầu quan trọng nhất của khách hàng. Khả năng mở rộng và hiệu suất mạng có thể không như mong đợi khi số lượng người dùng mạng tăng. Những nhà thiết kế mạng có thể tránh được vấn đề này bằng cách dùng các phương pháp từ trên xuống, thực hiện phân tích yêu cầu trước khi lựa chọn công nghệ.
Top-down network design is a methodology for designing networks that begins at the upper layers of the OSI reference model before moving to the lower layers. The top-down methodology focuses on applications, sessions, and data transport before the selection of routers, switches, and media that operate at the lower layers.	Thiết kế mạng từ trên xuống là một phương pháp thiết kế mạng bắt đầu ở các lớp trên của mô hình tham chiếu OSI trước khi chuyển sang các lớp thấp hơn. Phương pháp thiết kế từ trên xuống tập trung vào các ứng dụng, các phiên, và vận chuyển dữ liệu trước khi chọn lựa các bộ định tuyến, các chuyển mạch và môi trường hoạt động ở các lớp dưới.
The top-down network design process includes exploring organizational and group structures to find the people for whom the network will provide services and from whom the designer should get valuable information to make the design succeed.	Quá trình thiết kế mạng từ trên xuống bao gồm khám phá cơ cấu tổ chức và nhóm để tìm đối tượng mà mạng sẽ cung cấp dịch vụ và nhà thiết kế nên thu thập những thông tin có giá trị từ họ để làm cho thiết kế thành công.
Top-down network design is also iterative. To avoid	Thiết kế mạng từ trên xuống cũng có tính chất lặp. Để tránh

getting bogged down in details too quickly, it is important to first get an overall view of a customer's requirements. Later, more detail can be gathered on protocol behavior, scalability requirements, technology preferences, and so on. Top-down network design recognizes that the logical model and the physical design can change as more information is gathered. Because top-down methodology is iterative, some topics are covered more than once in this book. For example, this chapter discusses network applications. Chapter 4, "Characterizing Network Traffic," covers network applications in detail, with emphasis on network traffic caused by application- and protocol-usage patterns. A top-down approach enables a network designer to get "the big picture" first before spiraling downward into detailed technical requirements and specifications.	bị sa lầy vào các chi tiết quá nhanh chóng, điều quan trọng trước mắt là chúng ta cần có cái nhìn tổng quan về các yêu cầu của khách hàng. Sau đó, chúng ta cần thu thập những thông tin chi tiết về đặc điểm của giao thức, các yêu cầu về khả năng mở rộng, sở thích công nghệ, và v.v…Thiết kế mạng từ trên xuống cho rằng mô hình logic và thiết kế thực có thể thay đổi khi nhiều thông tin hơn được thu thập. Bởi vì phương pháp thiết kế từ trên xuống có tính chất lặp, một số chủ đề được đề cập đến nhiều lần trong sách này. Chẳng hạn, chương này thảo luận về các ứng dụng mạng. Chương 4, "Nghiên cứu lưu lượng mạng", đề cập đến các ứng dụng mạng một cách chi tiết, nhấn mạnh vấn đề lưu lượng mạng do việc dùng các pattern sử dụng ứng dụng và giao thức. Cách tiếp cận từ trên xuống giúp nhà thiết kế mạng có thể nhận được "bức tranh tổng thể" đầu tiên trước khi đi vào các yêu cầu và đặc tả kỹ thuật chi tiết.
Using a Structured Network Design Process	Sử dụng một quá trình thiết kế mạng có cấu trúc
Top-down network design is a discipline that grew out of	Thiết kế mạng từ trên xuống là

the success of structured software programming and structured systems analysis. The main goal of structured systems analysis is to more accurately represent users' needs, which unfortunately often are ignored or misrepresented. Another goal is to make the project manageable by dividing it into modules that can be more easily maintained and changed.	một lĩnh vực phát triển từ sự thành công của phương pháp lập trình phần mềm có cấu trúc và phân tích hệ thống có cấu trúc. Mục tiêu chính của việc phân tích các hệ thống có cấu trúc là biểu diễn chính xác các nhu cầu của người dùng, những yếu tố thường bị bỏ qua hoặc hiểu sai. Một mục tiêu khác là giúp chúng ta dễ quản lý dự án bằng cách chia nó thành các mô-đun có thể dễ dàng duy trì nguyên trạng cũng như dễ thay đổi.
Structured systems analysis has the following characteristics:	Phân tích các hệ thống cấu trúc có những đặc điểm sau đây:
■ The system is designed in a top-down sequence.	■ Hệ thống được thiết kế theo trình tự từ trên xuống.
■ During the design project, several techniques and models can be used to characterize the existing system, determine new user requirements, and propose a structure for the future system.	■ Trong một dự án thiết kế, người ta có thể sử dụng một số kỹ thuật và mô hình để xác định các tính chất của một hệ thống hiện tại, và đề suất cấu trúc cho hệ thống tương lai.
■ A focus is placed on data flow, data types, and processes that access or change the data.	■ Chúng ta tập trung vào một số vấn đề như luồng dữ liệu, kiểu dữ liệu, và các quá trình truy cập hoặc thay đổi dữ liệu.
■ A focus is placed on understanding the location	■ Chúng ta cũng tập trung vào việc tìm hiểu vị trí và các nhu

and needs of user communities that access or change data and processes.	cầu của cộng đồng người dùng truy cập hoặc thay đổi dữ liệu và các quá trình.
■ A logical model is developed before the physical model. The logical model represents the basic building blocks, divided by function, and the structure of the system. The physical model represents devices and specific technologies and implementations.	■ Chúng ta sẽ xây dựng mô hình logic trước mô hình vật lý. Mô hình logic thể hiện các yếu tố cấu thành cơ bản, được phân chia theo chức năng, và cấu trúc của hệ thống. Mô hình vật lý biểu diễn các thiết bị và các công nghệ đặc trưng cũng như việc thực thi chúng.
■ Specifications are derived from the requirements gathered at the beginning of the top-down sequence. With large network design projects, modularity is essential. The design should be split functionally to make the project more manageable. For example, the functions carried out in campus LANs can be analyzed separately from the functions carried out in remote-access networks, virtual private networks (VPN), and WANs.	■ Các đặc tả kỹ thuật được rút ra từ những yêu cầu thu thập được tại lúc bắt đầu trình tự từ trên xuống. Với những dự án mạng lớn, việc phân chia thành các mô đun là điều rất cần thiết. Thiết kế được phân chia theo chức năng để dễ quản lý dự án hơn. Ví dụ, các chức năng được thực hiện trong các LAN campus có thể được phân tích một cách tách biệt so với các chức năng được thực hiện trong các mạng truy cập từ xa, các mạng riêng ảo (VPN), và WAN.
Cisco recommends a modular approach with its three-layer hierarchical model. This model divides	Cisco đề nghị phương pháp tiếp cận mô-đun với mô hình phân cấp ba lớp của nó. Mô

networks into core, distribution, and access layers. The Cisco SAFE architecture, which is discussed in Part II of this book, “Logical Network Design,” is another modular approach to network design.	hình này chia mạng thành các lớp trung tâm, phân phối, truy cập. Chúng tôi sẽ đề cập đến kiến trúc Cisco SAFE trong Phần II của cuốn sách này, "Thiết kế mạng logic" cũng là một phương pháp tiếp cận mô-đun trong thiết kế mạng.
With a structured approach to network design, each module is designed separately, yet in relation to other modules. All the modules are designed using a top-down approach that focuses on requirements, applications, and a logical structure before the selection of physical devices and products to implement the design.	Với phương pháp tiếp cận cấu trúc trong vấn đề thiết kế mạng, mỗi mô-đun được thiết kế riêng biệt, nhưng có mối liên hệ với các mô-đun khác. Tất cả các mô-đun được thiết kế theo phương pháp từ trên xuống tập trung vào các yêu cầu, các ứng dụng, và một cấu trúc logic trước khi lựa chọn các thiết bị vật lý và các sản phẩm để thực thi thiết kế.
Systems Development Life Cycles Systems analysis students are familiar with the concept that typical systems are developed and continue to exist over a period of time, often called a systems development life cycle. Many systems analysis books use the acronym SDLC to refer to the system’s life cycle, which might sound strange to older networking students who know SDLC as Synchronous	[redacted]

Data Link Control, a bit-oriented, full-duplex protocol used on synchronous serial links, often found in a legacy Systems Network Architecture (SNA) environment. Nevertheless, it's important to realize that most systems, including network systems, follow a cyclical set of phases, where the system is planned, created, tested, and optimized. Feedback from the users of the system causes the system to then be redesigned or modified, tested, and optimized again. New requirements arise as the network opens the door to new uses. As people get used to the new network and take advantage of the services it offers, they soon take it for granted and expect it to do more. In this book, network design is divided into four major phases that are carried out in a cyclical fashion:	[redacted]

■ Analyze requirements: In this phase, the network analyst interviews users and technical personnel to gain an understanding of the business and technical goals for a new or enhanced system. The task of characterizing the existing network, including the logical and physical topology and network performance, follows. The last step in this phase is to analyze current and future network traffic, including traffic flow and load, protocol behavior, and quality of service (QoS) requirements. ■ Develop the logical design: This phase deals with a logical topology for the new or enhanced network, network layer addressing, naming, and switching and routing protocols. Logical design also includes security planning, network management design, and the initial investigation into which service providers can meet WAN and remote access requirements. ■ Develop the physical design: During the physical design phase, specific	[redacted]

technologies and products that realize the logical design are selected. Also, the investigation into service providers, which began during the logical design phase, must be completed during this phase.

■ Test, optimize, and document the design: The final steps in top-down network design are to write and implement a test plan, build a prototype or pilot, optimize the network design, and document your work with a network design proposal.

These major phases of network design repeat themselves as user feedback and network monitoring suggest enhancements or the need for new applications. Figure 1-1 shows the network design and implementation cycle.

Figure 1-1 Network Design and Implementation Cycle

Plan Design Implement Operate Optimize (PDIOO) Network Life Cycle

Cisco documentation refers to the Plan Design

Implement Operate Optimize (PDIOO) set of phases for the life cycle of a network. It doesn't matter which life cycle you use, as long as you realize that network design should be accomplished in a structured, planned, modular fashion, and that feedback from the users of the operational network should be fed back into new network projects to enhance or redesign the network. The PDIOO life cycle includes the following steps:

- Plan: Network requirements are identified in this phase. This phase also includes an analysis of areas where the network will be installed and an identification of users who will require network services.
- Design: In this phase, the network designers accomplish the bulk of the logical and physical design, according to requirements gathered during the plan phase.
- Implement: After the design has been approved, implementation begins. The network is built according to the design specifications. Implementation also serves

to verify the design.

- Operate: Operation is the final test of the effectiveness of the design. The network is monitored during this phase for performance problems and any faults to provide input into the optimize phase of the network life cycle.
- Optimize: The optimize phase is based on proactive network management that identifies and resolves problems before network disruptions arise. The optimize phase may lead to a network redesign if too many problems arise because of design errors or as network performance degrades over time as actual use and capabilities diverge. Redesign can also be required when requirements change significantly.
- Retire: When the network, or a part of the network, is out-of-date, it might be taken out of production. Although Retire is not incorporated into the name of the life cycle (PDIOO), it is nonetheless an important phase. The retire phase wraps around to the plan phase. The PDIOO life cycle repeats as network requirements evolve.

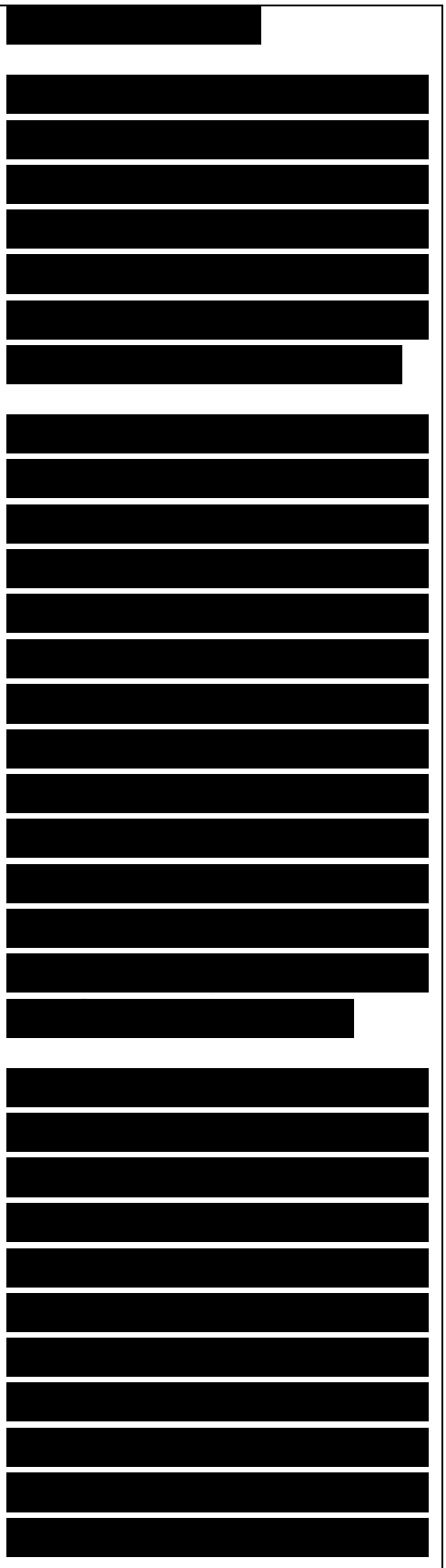


Figure 1-2 shows a graphical representation of the Cisco PDIOO network life cycle.

Analyzing Business Goals

Understanding your customer's business goals and constraints is a critical aspect of network design. Armed with a thorough analysis of your customer's business objectives, you can propose a network design that will meet with your customer's approval.

It is tempting to overlook the step of analyzing business goals, because analyzing such technical goals as capacity, performance, security, and so on is more interesting to many network engineers. Chapter 2, "Analyzing Technical Goals and Tradeoffs," covers analyzing technical goals. In this chapter, you learn the importance of analyzing business goals, and you pick up some techniques for matching a network design proposal to a customer's business objectives.

Working with Your Client Before meeting with your customer to discuss business goals for the network design project, it is a good idea to research your client's business. Find out what industry the client is in. Learn something about the client's market, suppliers, products, services, and competitive advantages. With the knowledge of your customer's business and its external relations, you can position technologies and products to help strengthen the customer's status in the customer's own industry. In your first meeting with your customers, ask them to explain the organizational structure of the company. Your final internetwork design will probably reflect the corporate structure, so it is a good idea to gain an understanding of how the company is structured in departments, lines of	[redacted]

business, vendors, partners, and field or remote offices. Understanding the corporate structure can help you locate major user communities and characterize traffic flow. Chapter 4 covers traffic flow in more detail. Understanding the corporate structure can also help you understand the corporate culture, which can affect the network design. For example, a company with a centralized management structure might require that products and vendors be chosen by headquarters management. A decentralized company might let branch offices have more say.

Note Understanding the corporate structure can also help you recognize the management hierarchy. One of your primary goals in the early stages of a network design project should be to determine who the decision makers are. Who will have the authority to accept or reject your network design proposal? Sometimes, this can be a rather complicated issue, as discussed in the section "Politics and Policies," later in this

chapter.	[redacted]
Ask your customer to state an overall goal of the network design project. Explain that you want a short, business-oriented statement that highlights the business purpose of the new network. Why is the customer embarking on this new network design project? For what will the new network be used? How will the new network help the customer be more successful in the customer's business?	[redacted]
After discussing the overall business goals of the network design project, ask your customer to help you understand the customer's criteria for success. What goals must be met for the customer to be satisfied? Sometimes success is based on operational savings because the new network allows employees to be more productive. Sometimes success is based on the ability to increase revenue or build partnerships with other companies. Make sure you know upfront how "success" is defined by	[redacted]

executives, managers, end users, network engineers, and any other stakeholders. Also, determine whether the customer's definition of success will change as yearly fiscal goals change.	[redacted]
In addition to determining the criteria for success, you should ascertain the consequences of failure: ■ What will happen if the network design project fails or if the network, when installed, does not perform to specification? ■ How visible is the project to upper-level management? ■ Will the success (or possible failure) of the project be visible to executives? ■ To what extent could unforeseen behavior of the new network disrupt business operations? In general, gather enough information to feel comfortable that you understand the extent and visibility of the network	[redacted]

design project. You should try to get an overall view of whether the new network is critical to the business's mission. Investigate the ramifications of the network failing or experiencing problems. Chapter 2 discusses the details of performance and reliability analysis, but at this point in the design process, you should start addressing these issues. (Remember that top-down network design is iterative. Many network design requirements are addressed more than once.) Changes in Enterprise Networks Enterprise networks at many corporations have been undergoing major changes. The value of making vast amounts of data available to employees, customers, and business partners has been recognized. Corporate employees, field employees, contract employees, and telecommuters need access to sales, marketing, engineering, and financial data, regardless of whether the data is stored on	[redacted]

centralized or distributed servers or mainframes. Suppliers, vendors, and customers also need access to many types of data. A network that is used by only internal users is no longer the norm at many companies. Companies are seeking ways to build networks that more closely resemble modern organizations. Many modern organizations are based on an open, collaborative environment that provides access to information and services for many different constituents, including customers, prospective customers, vendors, suppliers, and employees. To remain competitive, companies need ways to reduce product development time and take advantage of just-in-time manufacturing principles. A lot of companies achieve these goals by partnering with suppliers and by fostering an online, interactive relationship with their suppliers. An example is automobile manufacturing.	[redacted]

Instead of producing every automobile component in-house, many manufacturers contract with partners who specialize in specific components and technologies. For example, one partner might produce the engine while another produces the body. If all the partners can access data and services on the manufacturer's network, production costs are reduced, just-in-time manufacturing can be accomplished, and it is easier to plan around component shortages. The ability to share information saves time and money for the automobile manufacturer and for its partners.	[redacted]
A network designer must carefully consider requirements for extending the network to outside users. For security reasons, external access should not mean full network access. Using a modular approach to network design is important here so that a clear boundary exists between the enterprise's private networks and the portions of the internetwork that partners	[redacted]

can access. Networks Must Make Business Sense Although in the past many companies made "technology for technology's sake" choices, this is no longer the case. Business leaders are more involved in Information Technology (IT) decisions than they once were, and IT managers rely on business managers to help them prioritize and fund IT projects. Network upgrades are made not because some new technology sounds interesting to the engineers, but because it will help an enterprise increase profits, productivity, market share, and cash flow. Network designers must choose solutions that address the business dilemmas faced by business managers. Network applications have become mission critical. Despite this trend, large budgets for networking and telecommunications operations have been reduced at some companies. Many companies have gone	[redacted]

through difficult reengineering projects to reduce operational costs and are still looking for ways to manage networks with fewer resources and to reduce the recurring costs of WAN circuits.	[redacted]
Companies are researching ways to make their data centers more efficient in their usage of power, cabling, racks, storage, and WAN circuits. Companies seek to reduce data center costs and to make data centers more "green" (whereby energy usage is reduced). Data center managers have discovered that many of their servers' CPUs are underutilized. A major trend in enterprise network design is server virtualization, where one hardware platform supports multiple virtual servers. Instead of many underutilized hardware boxes, there are now just a few hardware boxes, each of which supports multiple virtual servers. Each virtual server looks and acts just like a physical server, including a fully functional operating system and one or more applications.	[redacted]

[redacted]

Streamlining processes and protocols has also led to an increased use of IP telephony and to the continued convergence of voice and data networks. To save money and to reduce the need for specialized data or voice engineers, companies continue to adopt IP telephony technologies. In previous network designs, telecommunications and voice networks were separate.Telecommunications engineers knew little about data networks, and data communications engineers didn't know the difference between a time-division multiplexer (TDM) and a tandem switching system (TSS). In today's environment, voice, data, and video networks are merged.

[redacted]

Networks Offer a Service

[redacted]

Modern IT departments are more service-oriented than they used to be. To meet the needs of their customers, IT departments are spending

more time analyzing and documenting their processes for delivering services. A focus on processes helps to ensure effective service delivery and to avoid wasted expenditures on technology that doesn't provide a needed service.	[redacted]
As a network designer, you might find yourself working with IT architects who adhere to the IT Service Management (ITSM) discipline. ITSM defines frameworks and processes that can help an organization match the delivery of IT services with the business needs of the organization. ITSM focuses on processes rather than technology and helps an IT organization think of its users as valued customers rather than problem-generating adversaries. A version of ITSM is documented in the Information Technology Infrastructure Library (ITIL), a series of books published by the United Kingdom Office of Government Commerce (OGC), each of which covers an IT management topic. The	[redacted]

details of ITSM and ITIL are outside the scope of this book, but it is worth noting that both ITSM and top-down network design address the need to align the delivery of IT services to the business needs of an organization. This book will help you design networks that comply with ITSM practices. Other trends in IT management that affect network design are related to governance and compliance. Governance refers to a focus on consistent, cohesive decisions, policies, and processes that protect an organization from mismanagement and illegal activities of users of IT services. Compliance refers to adherence to regulations that protect against fraud and inadvertent disclosure of private customer data. For example, in the United States, retail organizations must comply with the Payment Card Industry Data Security Standard (PCI DSS) and healthcare organizations must comply with the Health Insurance Portability and Accountability Act (HIPAA).	[redacted]

[redacted]

The Need to Support Mobile Users

[redacted]

Notebook computers have finally become small enough to carry around, and workers now expect to get work done at home, on the train, in hotels, in meeting rooms, at customer sites, and even while having their morning latte at the local coffee shop. Notebook computers ship with wireless networking built in to facilitate users getting work done outside the office.

[redacted]

It shouldn't matter (to the user anyway) where data is

[redacted]

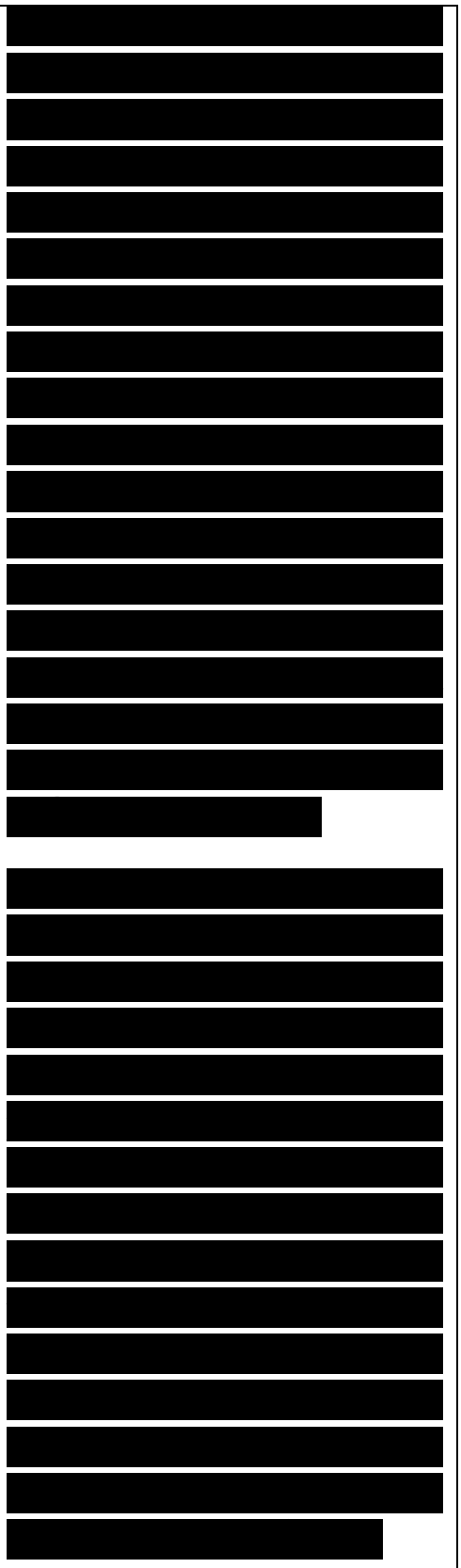
and in what format. Network users expect network performance to be uniform, regardless of where the user or data resides. A user should be able to read email on a cell phone, for example, and read voice mail from a web browser while sipping coffee in an Internet cafe. Users should have secure and reliable access to tools and data wherever they are. The challenge for network designers is to build networks that allow data to travel in and out of the enterprise network from various wired and wireless portals without picking up any viruses and without being read by parties for whom it was not intended. One of the biggest trends in network design is virtual private networking (VPN), where private networks make use of the Internet to reach remote locations or possibly other organizations. Customers getting involved in VPN projects have concerns about security, reliable and predictable performance, and data throughput requirements.	[redacted]

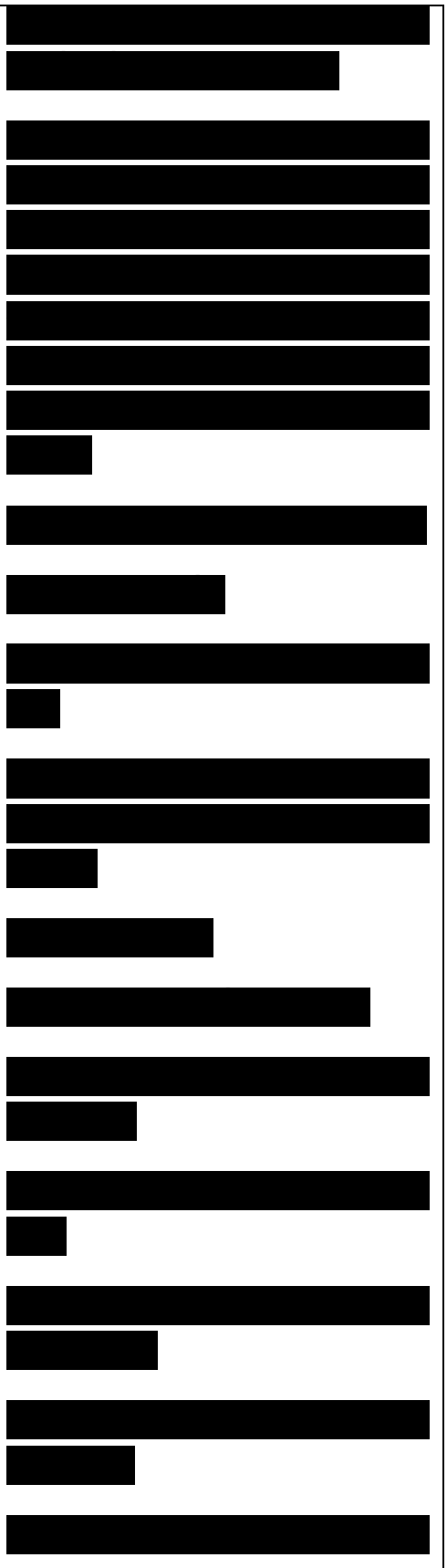
Chapter 5, “Designing a Network Topology,” covers VPNs in greater detail.	[redacted]
Network architectures are taking on a virtual and ubiquitous form for users, while remaining highly structured and managed from the network engineers’ point of view. The designer is challenged to develop secure, resilient, and manageable solutions that enable users to work efficiently and securely wherever they are physically located.	[redacted]
The Importance of Network Security and Resiliency	[redacted]
Network security has filtered to the top of the list of business goals at many companies. Although security was always important, it has become even more important as networks become indispensable and as tools for breaking into networks become ubiquitous. Enterprises must protect their networks from both the	[redacted]

unsophisticated "script kiddies" and from more advanced attacks launched by criminals or political enemies. There is also a continued requirement to protect networks from Trojan horses and viruses.	[redacted]
Many enterprise managers now report that the network must be available 99.999 percent of the time. Although this goal might not be achievable without expensive redundancy in staff and equipment, it might be a reasonable goal for companies that would experience a severe loss of revenue or credibility if the network were down for even short periods of time. This goal is linked to goals for security, as the network can't be available if security breaches and viruses are disabling network devices and applications. When security and operational problems occur, networks must recover quickly. Networks must be resilient.	[redacted]

More than ever, IT and business managers require high-availability and resiliency features for their network equipment and protocols, as they realize the extent to which network downtime can jeopardize business success. In addition to security, another goal that has filtered to the top of the list of business goals is the need for business continuity during and after a disaster. Companies that have survived hurricanes, earthquakes, fires, and terrorist attacks have learned the importance of a disaster recovery plan that promotes business continuity despite the loss of critical network devices and services. Many companies have not had the misfortune of learning these lessons the hard way but are nonetheless embarking on network design projects with the goal of developing a network that will recover quickly if a natural or unnatural disaster occurs.	[redacted]

One aspect of analyzing a customer's business goals is the process of analyzing vulnerabilities related to disasters and the impact on business operations. Help your customer determine which network capabilities are critical and which facilities provide them. Consider how much of the network could be damaged without completely disrupting the company's mission. Determine whether other locations in the company are prepared to take on mission-critical functions.	[redacted]
In the past few years, networks have become more interconnected and complex, which can make meeting goals for business continuity and network resiliency more difficult. Many enterprise networks are linked to telecommuter home networks, branch-office networks, extranets that offer access to business partners and customers, and the Internet. The diversity and quantity of portals into the enterprise network pose	[redacted]

many security and stability risks. On the other hand, geographical diversity of mission-critical capabilities has turned out to be a lifesaver for some companies hit with disaster. One reason that The Wall Street Journal was able to publish its newspaper the day after the 9/11 attacks was because it had learned from 1990s power outages about the need to disperse critical functions across many different sites. In the current business environment, security and disaster recovery should be considered with every network design choice, and the network designer must propose solutions that provide resiliency and stability. A systematic and modular design process, as taught in this book, is even more important than it once was, as networks become increasingly more complex and vital to an organization's success.	

Typical Network Design Business Goals After considering the changes in business strategies and enterprise networking discussed in the previous sections, it is possible to list some typical network design business goals: ■ Increase revenue and profit ■ Increase market share ■ Expand into new markets ■ Increase competitive advantages over companies in the same market ■ Reduce costs ■ Increase employee productivity ■ Shorten product-development cycles ■ Use just-in-time manufacturing ■ Plan around component shortages ■ Offer new customer services ■ Offer better customer support	

■ Open the network to key constituents (prospects, investors, customers, business partners, suppliers, and employees)

■ Avoid business disruption caused by network security problems

■ Avoid business disruption caused by natural and unnatural disasters

■ Modernize outdated technologies

■ Reduce telecommunications and network costs, including overhead associated with separate networks for voice, data, and video

■ Make data centers more efficient in their usage of power, cabling, racks, storage, and WAN circuits

■ Comply with IT architecture design and governance goals

Identifying the Scope of a Network Design Project

One of the first steps in starting a network design project is to determine its scope. Some of the most

common network design projects these days are small in scope—for example, projects to allow a few people in a sales office to access the enterprise network via a VPN. On the other hand, some design projects are large in scope. Ask your customer to help you understand if the design is for a single network segment, a set of LANs, a set of WANs or remote-access networks, or the entire enterprise network. Also ask your customer if the design is for a new network or a modification to an existing one.

Explain to your customer any concerns you have about the scope of the project, including technical and business concerns. Subsequent sections in this chapter discuss politics and scheduling, which are tightly linked to the scope of a network design project. (Many network designers have learned the hard way what happens when you don't help your customers match the schedules of their projects to the scope.)

[redacted]

Make sure your customers tell you everything they can about the network and the design project. You might want to poke around outside the stated scope of the project, just to make sure nothing essential has been omitted. Double-check that you have gathered all the requirements and that you have accurate information about sites, links, and devices. If the project addresses network security, make sure you know about all external links, including any legacy dial-in access.	[redacted]
Note Designers rarely get a chance to design a network from scratch. Usually a network design project involves an upgrade to an existing network. However, this is not always the case. Some senior network designers have developed completely new next-generation networks to replace old networks. Other designers have designed networks for a new building or new campus. Even in these cases, however, the	[redacted]

new network usually has to fit into an existing infrastructure—for example, a new campus network that has to communicate with an existing WAN. Where there is an existing network, the design project must include plans for migrating to the new design with minimal disruption and risk.

When analyzing the scope of a network design, you can refer to the seven layers of the OSI reference model to specify the types of functionality the new network design must address. For example, you might decide that the design project is concerned only with network layer matters such as routing and IP addressing. Or you might decide that the design also concerns the application layer because the focus is on voice applications, such as Interactive Voice Response (IVR), which directs customers to the correct location in a call center, or unified messaging, where email can be retrieved via voice mail and text messages can be converted into speech. Figure 1-3 shows the

OSI reference model.

Figure 1-3 Open System Interconnection (OSI) Reference Model

In addition to using the OSI reference model, this book also uses the following terms to define the scope of a network and the scope of a network design project:

- Segment: A single network bounded by a switch or router and based on a particular Layer 1 and Layer 2 protocol such as Fast Ethernet.

- LAN: A set of switched segments based on a particular Layer 2 protocol such as Fast Ethernet and an interswitch trunking protocol such as the IEEE 802.1Q standard.

- Building network: Multiple LANs within a building, usually connected to a building- backbone network.

- Campus network: Multiple buildings within a local geographical area

(within a few miles), usually connected to a campus-backbone network.

■ Remote access: Networking solutions that support individual remote users or small remote branch offices accessing the network.

■ WAN: A geographically dispersed network including point-to-point, Frame Relay, ATM, and other long-distance connections.

■ Wireless network: A LAN or WAN that uses the air (rather than a cable) for its medium.

■ Enterprise network: A large and diverse network, consisting of campuses, remote- access services, and one or more WANs or long-range LANs. An enterprise network is also called an internetwork.

Identifying a Customer's Network Applications

At this point in the design process, you have identified

your customer's business goals and the scope of the project. It is now time to focus on the real reason networks exist: applications. The identification of your customer's applications should include both current applications and new applications. Ask your customer to help you fill out a chart, such as the one in Table 1-1.

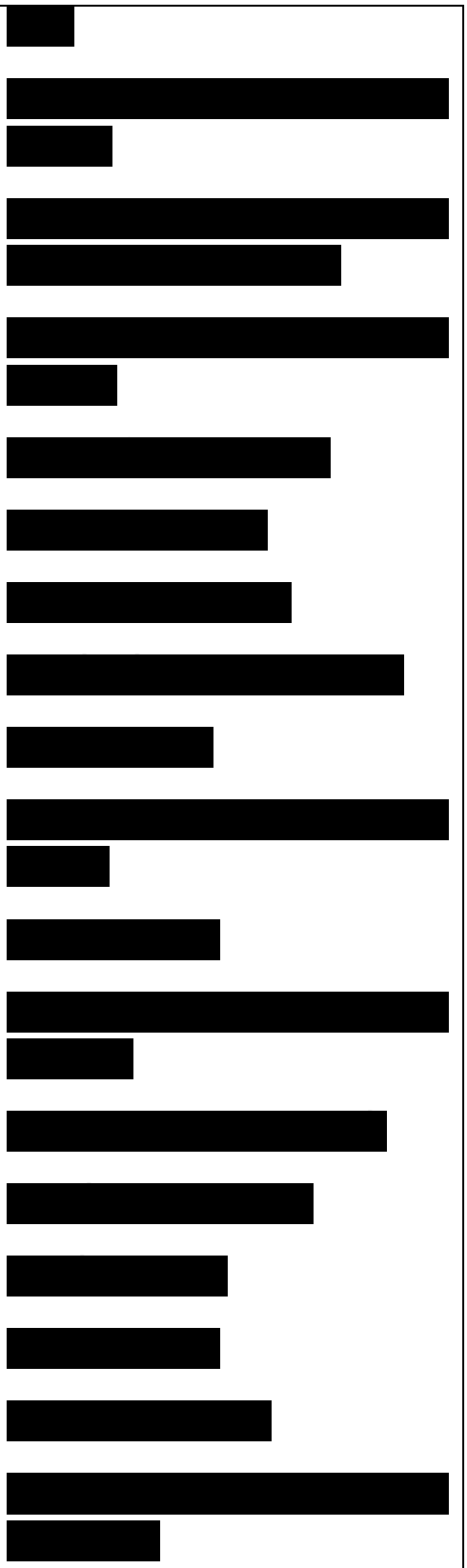
Note Table 1-1 identifies network applications. In Chapters 2 and 4, it will be enhanced to include technical requirements and network-traffic characteristics. At this point, your goal is simply to identify network applications.

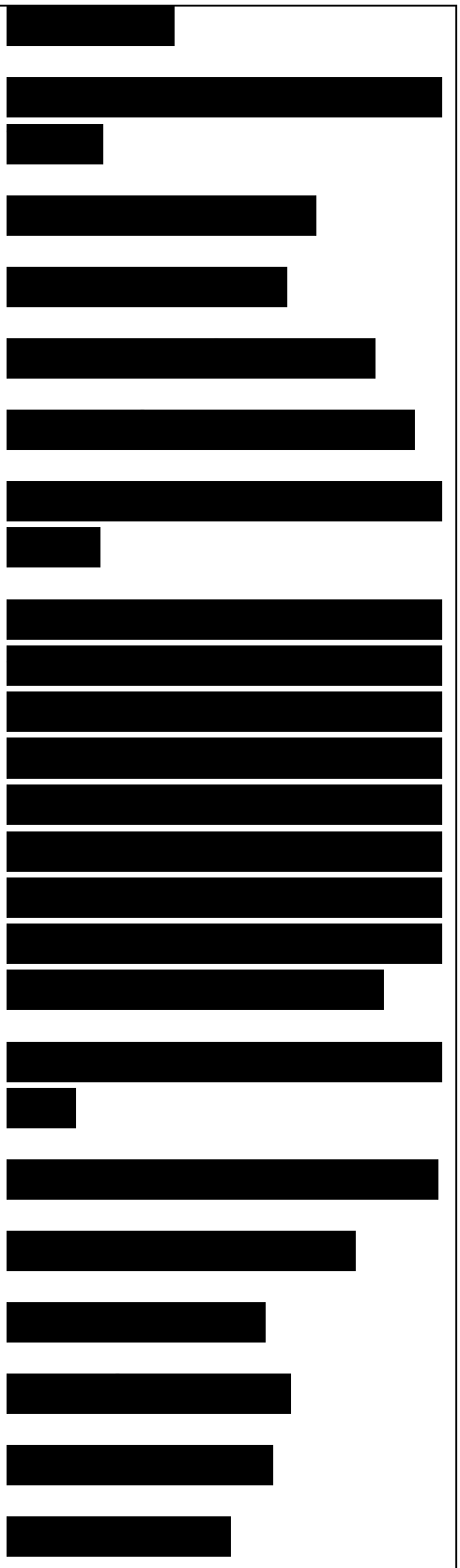
Table 1-1 Network Applications

Name of Application	Type of Application?	New Application? (Yes or No)	Criticality	Comments

For Name of Application, simply use a name that your customer gives you. This could be an industry-standard name, such as Lotus

Notes, or it could be an application name that means something only to the customer (especially for a home-grown application). For new applications, the name might be a code name for a software-development project.	[redacted]
For Type of Application, you can use any appropriate text that describes the type of application, or you can classify the application as one of the following standard network applications: ■ Email ■ File transfer, sharing, and access ■ Database access and updating ■ Web browsing ■ Network game ■ Remote terminal ■ Calendar ■ Medical imaging ■ Videoconferencing ■ Video on demand (VoD) ■ Scheduled multicast video	[redacted]

■ Surveillance and security camera video ■ Internet or intranet voice (IP telephony) ■ Internet or intranet fax ■ Sales order entry ■ Management reporting ■ Sales tracking ■ Computer-aided design ■ Document imaging ■ Inventory control and shipping ■ Telemetry ■ Interactive Voice Response (IVR) ■ Unified messaging ■ Desktop publishing ■ Web publishing ■ Electronic whiteboard ■ Terminal emulation ■ Online directory (phone book)	

■ Distance learning ■ Point of sales (retail store) ■ Electronic commerce ■ Financial modeling ■ Human resources management ■ Computer-aided manufacturing ■ Process control and factory floor The preceding list includes user applications. The chart in Table 1-1 should also include system applications. (Or if you prefer, you can do a separate chart for system applications.) System applications include the following types of network services: ■ User authentication and authorization ■ Host naming and name resolution ■ Dynamic host addressing ■ Remote booting ■ Remote configuration download ■ Directory services ■ Network backup ■ Network management ■ Software distribution	

In the Criticality column of the Network Applications chart, you can give each application a ranking from 1 to 3 with the following meanings: 1. Extremely critical 2. Somewhat critical 3. Not critical	[redacted]
Later, you can gather more specific information on mission criticality, including precisely how much downtime is acceptable (if the customer can quantify availability requirements).	[redacted]
In the Comments column, add any observations relevant to the network design. For example, include any information you have about corporate directions, such as plans to stop using an application in the future or specific rollout schedules and regional-use plans.	[redacted]
Analyzing Business Constraints In addition to analyzing	[redacted]

business goals and determining your customer's need to support new and existing applications, it is important to analyze any business constraints that will affect your network design.

Politics and Policies

It has been said that there are two things not to talk about with friends: politics and religion. It would be nice if you could escape discussing office politics and technological religion (technology preferences) with a network design customer, but avoiding these topics puts your project at risk.

In the case of office politics, your best bet is to listen rather than talk. Your goal is to learn about any hidden agendas, turf wars, biases, group relations, or history behind the project that could cause it to fail. In some cases, a similar project was already tried and didn't work. You should determine if this has happened in your case and, if it has, the reasons why the project failed or never had a chance

to come to fruition.	[redacted]
Pay attention to personnel issues that could affect the project. Which manager or managers started the project and how much do they have at stake? Are there any managers, network engineers, or users who want the project to fail for any reason? Find out who your advocates and opponents are. In some cases, no matter how technically sound your network design is, there will be people who have a negative reaction to it.	[redacted]
Be sure to find out if your project will cause any jobs to be eliminated. Some network design projects involve automating tasks that were once done by highly paid workers. These workers obviously will have reasons to want the project to fail.	[redacted]
Find out if there is a strategic business or IT plan. Does	[redacted]

your network design need to fit into an overall architecture that is based on strategic planning? Are there external regulatory or governmental pressures on the planning process or on the architecture? These sorts of pressures can often lead to messy political battles that can affect your network design.	[redacted]
Be prepared for the possibility of formidable office politics if your network design project involves the merging of voice and data networks. Voice experts and data experts have traditionally lived in their own worlds. They might face each other with some mistrust and fear for the future. You can often reduce the uncertainty by running short IP telephony seminars for voice technicians and traditional telephony seminars for the data network administrators.	[redacted]
While working with a client,	[redacted]

you will gain a feeling for the client's business style. One aspect of style that is important to understand is tolerance to risk. Is risk taking rewarded in the company, or are most people afraid of change? Knowing the employment history of the decision makers will help you select appropriate technologies. The employment history of the decision makers affects their tolerance to risk and their biases toward certain technologies. Understanding these issues will help you determine whether your network design should be conservative or if it can include new, state-of-the art technologies and processes.	[redacted]
Another aspect of the client's business style has to do with testing the design. At some companies, the testers might claim they have carefully tested a new Voice over IP (VoIP) implementation, for example, when what they actually did was get a VoIP	[redacted]

call to complete. Your idea of testing, on the other hand, might be to make numerous calls under various load conditions. See Chapter 12, "Testing Your Network Design," for more information on testing.	[redacted]
You need to discuss with your customer any policies about protocols, standards, and vendors. Try to learn of any "forbidden technologies" where the users or network engineers have decided, possibly for the wrong reasons, that a particular protocol is slow or unstable.	[redacted]
Find out whether the company has standardized on any transport, routing, desktop, or other protocols. Determine whether there is any doctrine regarding open versus proprietary solutions. Find out if there are any policies on approved vendors or platforms. In many cases, a company has already chosen technologies and products for the new network, and your design must fit into the plans. Ask your customer if there are any policies regarding distributed authority for	[redacted]

network design and implementation. For example, are there departments that control their own internetworking purchases? Find out if departments and end users are involved in choosing their own applications. Make sure you know who the decision makers are for your network design project.	[redacted]
A lot of organizations need to implement policies in response to legal, regulatory, or contractual requirements. In the United States, Generally Accepted Accounting Principles (GAAP) drive many accounting policies. In the medical profession, network designs might be affected by security and privacy policies that are regulated by HIPAA. In other parts of the world, network equipment choices may be regulated by governmental Postal, Telegraph, and Telephone (PTT) organizations.	[redacted]

In the rush to get to technical requirements, network designers sometimes ignore nontechnical issues, which is a mistake. Many brilliant network designs have been rejected by a customer because the designer focused on the lower layers of the OSI reference model and forgot about company politics and technical biases.	[redacted]
Budgetary and Staffing Constraints	[redacted]
Your network design must fit the customer's budget. The budget should include allocations for equipment purchases, software licenses, maintenance and support agreements, testing, training, and staffing. The budget might also include consulting fees (including your fees) and outsourcing expenses.	[redacted]
Throughout the project, work with your customer to identify requirements for new personnel, such as additional network managers. Point out the need for personnel training, which	[redacted]

will affect the budget for the project.	[redacted]
In general, it is a good idea to analyze the abilities of the networking staff. How much inhouse expertise is there? Should you recommend any training or outsourcing for network operations and management? The technologies and protocols that you recommend will depend on the abilities of internal staff. It is not a good idea to recommend a complex routing protocol, such as Open Shortest Path First (OSPF), for example, if the engineering staff is just starting to learn internetworking concepts (unless you also recommend a comprehensive training plan).	[redacted]
Analyzing in-house expertise is especially important and challenging for companies that merge their voice and data networks. Consider the need to train the traditional voice experts on data technologies and the data experts on voice technologies. Also, implementing voice and	[redacted]

video often requires advanced QoS knowledge that may necessitate training.	[redacted]
To ensure the success of your project, determine who controls the network budget—the Information Systems (IS) department, network managers, or users' departments? How much control do users and groups have over network expenditures? Are there any departmental charge-back schemes?	[redacted]
Regardless of who controls the budget, one common network design goal is to contain costs. Reduced budgets or limited resources often force network designers to select the most affordable solution instead of the best solution. It is useful to know the areas in which the network design can be changed with the least effect on performance to meet budget requirements. Chapter 2 discusses typical tradeoffs that must be made to meet the goal of affordability while achieving good performance and reliability.	[redacted]

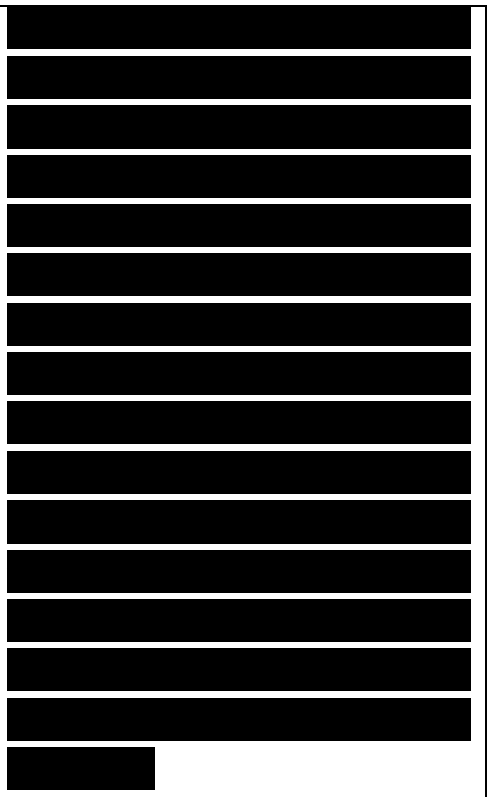
If possible, work with your customer to develop a return on investment (ROI) analysis for the network design. Make a business case to the customer that explains how quickly the new network will pay for itself, due to reduced operational costs, improved employee productivity, or the enabling of higher revenue potential and market expansion. Project Scheduling An additional business-oriented topic that you should review with your customer is the timeframe for the network design project. When is the final due date and what are the intermediate and major milestones? In most cases, management of the project schedule is the customer's obligation, not yours, but you should ask the customer to give you a copy of the schedule and to keep you informed about any slips in the schedule.	[redacted]

Note It's important to include intermediate milestones in the project schedule. They give you and your client a way to detect slips in the schedule. Consider the state of building wiring, which might be poor quality and not support new applications. If the wiring needs to be replaced, this will have a major impact on the schedule. Also, be sure to include circuit disconnect or circuit capacity changes in the project schedule. There is often a long lead time for these changes. Plan to document when the circuit changes and other major changes take place so that if problems occur, you can analyze what has changed to help you troubleshoot. Many tools exist for developing a schedule that includes milestones, resource assignments, critical-path analysis, and so on. Take a look at these aspects of the	[redacted]

schedule and voice your view on whether the schedule is practical, considering what you have learned about the scope of the project. An aggressive implementation schedule might require a reduction in the scope of the project or a reduction in the quality of the planning and testing that will be conducted. During the technical-analysis stage and the logical- and physical-design phases of the project, be sure to keep the schedule in mind. As you iteratively develop a concrete understanding of the technical scope of the network design project, point out any concerns you have about the schedule.

Business Goals Checklist

You can use the following checklist to determine if you have addressed your client's business-oriented objectives and concerns. If you can't gather every piece of data mentioned in the checklist, make sure you document what is missing in case it becomes critical, but don't stall the project to gather every last detail. This book teaches an ideal network design methodology that you should try to follow, but if real-world constraints, such

as uncooperative network design customers, budget cuts, and time constraints, hamper your ability to follow the methodology precisely, just follow it as much as you can. In general, the methodology still works even if some data is missing after you do your analysis.	
□ I have researched the customer's industry and competition.	□ Tôi đã nghiên cứu ngành nghề và những yếu tố cạnh tranh của khách hàng.
□ I understand the customer's corporate structure.	□ Tôi hiểu cấu trúc doanh nghiệp của khách hàng.
□ I have compiled a list of the customer's business goals, starting with one overall business goal that explains the primary purpose of the network design project.	□ Tôi đã biên soạn một danh sách các mục tiêu kinh doanh của khách hàng, bắt đầu với một mục tiêu kinh doanh tổng thể giải thích mục đích chính của dự án thiết kế mạng.
□ The customer has identified any mission-critical operations.	□ Khách hàng đã chỉ rõ mọi hoạt động then chốt.
□ I understand the customer's criteria for success and the ramifications of failure.	□ Tôi hiểu tiêu chí của khách hàng về thành công và những

	hệ quả của thất bại.
□ I understand the scope of the network design project.	□ Tôi hiểu phạm vi của dự án thiết kế mạng.
□ I have identified the customer's network applications (using the Network Applications chart).	□ Tôi đã xác định các ứng dụng mạng của khách hàng (bằng cách sử dụng biểu đồ ứng dụng mạng).
□ The customer has explained policies regarding approved vendors, protocols, or platforms.	□ Các khách hàng đã giải thích các chính sách liên quan đến các nhà cung cấp, các giao thức, hoặc các nền được chấp nhận.
□ The customer has explained any policies regarding open versus proprietary solutions.	□ Các khách hàng đã giải thích các chính sách liên quan đến các giải pháp mở và độc quyền.
□ The customer has explained any policies regarding distributed authority for network design and implementation.	□ Các khách hàng đã giải thích các chính sách liên quan đến quyền phân phối thiết kế và thực thi mạng.
□ I know the budget for this project.	□ Tôi biết ngân sách cho dự án này.
□ I know the schedule for this project, including the final due date and major milestones, and I believe it is practical.	□ Tôi biết lịch trình của dự án, bao gồm ngày đáo hạn cuối cùng và những cột mốc quan trọng, và tôi tin lịch trình đó khả thi.
□ I have a good understanding of the technical expertise of my clients and any relevant internal or external staff.	□ Tôi đã hiểu rõ về chuyên môn kỹ thuật của khách hàng và bất kỳ nhân viên nào có

	liên quan trong công ty cũng như ngoài công ty.
□ I have discussed a staff-education plan with the customer.	□ Tôi đã thảo luận kế hoạch đào tạo nhân viên với khách hàng.
□I am aware of any office politics that might affect the network design.	□ Tôi nhận thức được những vấn đề chính trị ở nơi làm việc có thể ảnh hưởng đến thiết kế mạng.
Summary This chapter covered typical network design business goals and constraints. It also talked about the top-down process for gathering information on goals, and the importance of using systematic methods for network design. Using systematic methods will help you keep pace with changing technologies and customer requirements. The next chapter covers analyzing technical goals and constraints.	Tóm tắt Chương này trình bày các mục tiêu và ràng buộc kinh doanh tiêu biểu trong thiết kế mạng. Nó cũng đề cập đến quy trình từ trên xuống để thu thập các thông tin về mục tiêu, và tầm quan trọng của việc sử dụng phương pháp hệ thống trong thiết kế mạng. Sử dụng phương pháp có hệ thống sẽ giúp bạn bắt kịp với sự thay đổi công nghệ và các yêu cầu của khách hàng. Chương tiếp theo sẽ đề cập đến việc phân tích các mục tiêu và ràng buộc kỹ thuật.
This chapter also talked about the importance of analyzing your customer's business style, tolerance to risk, biases, and technical expertise. You should also work with your customer to	Chương này cũng nói về tầm quan trọng của việc phân tích phong cách kinh doanh, khả năng chịu đựng rủi ro, những thành kiến, và chuyên môn kỹ thuật của khách hàng. Bạn

understand the budget and schedule for the network design project to make sure the deadlines and milestones are practical. Finally, you need to start gaining an understanding of your client's corporate structure. Understanding the corporate structure will help you analyze data flow and develop a network topology, which usually parallels the corporate structure. It will also help you identify the managers who will have the authority to accept or reject your network design, which will help you prepare and present your network design appropriately.	cũng nên làm việc với khách hàng của bạn để hiểu về vấn đề ngân sách và lịch trình của dự án thiết kế mạng để đảm bảo hạn chót và các cột mốc có tính khả thi. Cuối cùng, bạn cần phải bắt đầu tìm hiểu về cấu trúc doanh nghiệp của khách hàng. Hiểu rõ cấu trúc của công ty sẽ giúp bạn phân tích lưu lượng dữ liệu và xây dựng một tô-pô mạng, thường song song với cơ cấu doanh nghiệp. Nó cũng sẽ giúp bạn xác định các nhà quản lý có quyền chấp nhận hoặc loại dự án thiết kế mạng của bạn, điều này giúp bạn chuẩn bị và trình bày thiết kế mạng của bạn một cách thích hợp. checked

Chapter 2

Analyzing Technical Goals and Tradeoffs

This chapter provides techniques for analyzing a customer's technical goals for a new network design or network upgrade. Analyzing your customer's technical goals can help you confidently recommend technologies that will perform to your customer's expectations.

Typical technical goals include scalability, availability, network performance, security, manageability, usability, adaptability, and affordability. Of course, there are tradeoffs associated with these goals. For example, meeting strict requirements for performance can make it hard to meet a goal of affordability. The section "Making Network Design Tradeoffs" later in this chapter discusses tradeoffs in more detail.

One of the objectives of this

chapter is to give you terminology that will help you discuss technical goals with your customer. Network designers and users have many terms for technical goals, and, unfortunately, many different meanings for the terms. This chapter can help you choose terminology that has technical merit and is understandable by business and IT customers.

This chapter concludes with a checklist to help you determine whether you have addressed all your customer's technical goals and constraints.

Scalability

Scalability refers to how much growth a network design must support. For many enterprise network design customers, scalability is a primary goal. Many large companies add users, applications, additional sites, and external network connections at a rapid rate. The network design you propose to a customer should

be able to adapt to increases in network usage and scope.	[redacted]
Planning for Expansion	[redacted]
Your customer should help you understand how much the network will expand in the next year and in the next 2 years. (Ask your customer to analyze goals for growth in the next 5 years also, but be aware that not many companies have a clear 5-year vision.)	[redacted]
You can use the following list of questions to analyze your customer's short-term goals for expansion:	[redacted]
■ How many more sites will be added in the next year? The next 2 years?	[redacted]
■ How extensive will the networks be at each new site?	[redacted]
■ How many more users	[redacted]

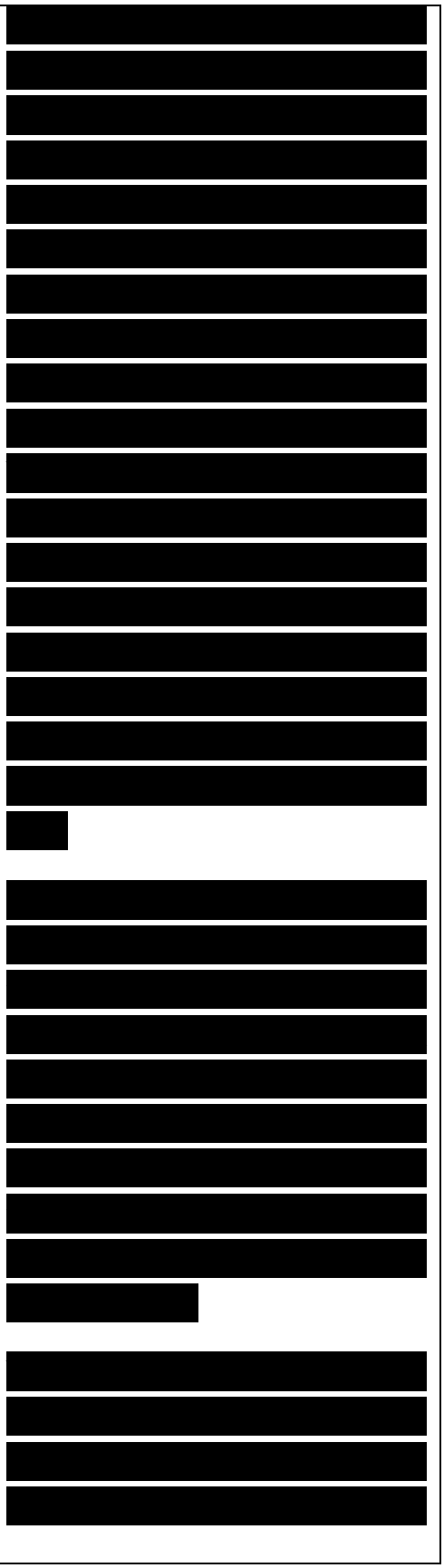
will access the corporate internetwork in the next year? The next 2 years?

- How many more servers will be added to the internetwork in the next year? The next 2 years?

Expanding Access to Data

Chapter 1, "Analyzing Business Goals and Constraints," talked about a common business goal of expanding access to data for employees who use enterprise networks. Managers empower employees to make strategic decisions that require access to sales, marketing, engineering, and financial data. In the 1970s and early 1980s, this data was stored on mainframes. In the late 1980s and the 1990s, this data was stored on servers in depart-mental LANs. Today, this data is again stored on centralized mainframes and servers.

In the 1990s, networking books and training classes taught the 80/20 rule for

capacity planning: 80 percent of traffic stays local in departmental LANs, and 20 percent of traffic is destined for other departments or external networks. This rule is no longer universal and is rapidly moving to the other side of the scale. Many companies have centralized servers residing in data centers. In addition, corporations increasingly implement intranets that enable employees to access centralized web servers using Internet Protocol (IP) technologies. At some companies, employees can access intranet web servers to arrange business travel, search online phone directories, order equipment, and attend distance-learning training classes. The web servers are centrally located, which breaks the classic 80/20 rule. As Chapter 1 also mentioned, there has been a trend of companies connecting internetworks with other companies to collaborate with partners,	

resellers, suppliers, and strategic customers. The term extranet is sometimes used to describe an internal internetwork that is accessible by outside parties. If your customer has plans to implement an extranet, you should document this in your list of technical goals so that you can design a topology and provision bandwidth appropriately. In the 1980s and 1990s, mainframes running Systems Network Architecture (SNA) protocols stored most of a company's financial and sales data. In recent years, the value of making this data available to more than just financial analysts has been recognized. The business goal of making data available to more departments often results in a technical goal of using the mainframe as an incredibly powerful database server. The business goal of making more data available to users results in the following technical goals for scaling and upgrading corporate	[redacted]

enterprise networks:

- Connect separated departmental LANs into the corporate internetwork.
- Solve LAN/WAN bottleneck problems caused by large increases in internetwork traffic.
- Provide centralized servers that reside in a data center.
- Make mainframe data accessible to the enterprise IP network.
- Add new sites to support field offices and telecommuters.
- Add new sites and services to support secure communication with customers, suppliers, resellers, and other business partners.

Constraints on Scalability

When analyzing a customer's scalability goals, it is important to keep in mind that there are

impediments to scalability inherent in networking technologies. Selecting technologies that can meet a customer's scalability goals is a complex process with significant ramifications if not done correctly. For example, selecting a flat network topology with Layer 2 switches can cause problems as the number of users scales, especially if the users' applications or network protocols send numerous broadcast frames. (Switches forward broadcast frames to all connected segments.)

Subsequent chapters in this book consider scalability again. Chapter 4, "Characterizing Network Traffic," discusses the fact that network traffic (for example, broadcast traffic) affects the scalability of a network. Part II, "Logical Network Design," provides details on the scalability of routing and switching protocols. Part III, "Physical Network Design," provides information on the scalability of LAN and WAN technologies and

internetworking devices. Remember that top-down network design is an iterative process. Scalability goals and solutions are revisited during many phases of the network design process. Availability Availability refers to the amount of time a network is available to users and is often a critical goal for network design customers. Availability can be expressed as a percent uptime per year, month, week, day, or hour, compared to the total time in that period. For example, in a network that offers 24-hour, 7-days-a-week service, if the network is up 165 hours in the 168-hour week, availability is 98.21 percent. Network design customers don't use the word availability in everyday English and have a tendency to think it means more than it does. In general, availability means how much	[redacted]

time the network is operational. Availability is linked to reliability but has a more specific meaning (percent uptime) than reliability. Reliability refers to a variety of issues, including accuracy, error rates, stability, and the amount of time between failures.

Note Sometimes network engineers classify capacity as part of availability. The thinking is that even if a network is available at Layer 1 (the physical layer), it is not available from a user's point of view if there is not enough capacity to send the user's traffic.

For example, Asynchronous Transfer Mode (ATM) has a connection admission control function that regulates the number of cells allowed into an ATM network. If the capacity and quality of service (QoS) requested for a connection are not available, cells for the connection are not allowed to enter the network. This problem could be considered an availability issue. However, this book

classifies capacity with performance goals. Availability is considered simply a goal for percent uptime.	[redacted]
Availability is also linked to redundancy, but redundancy is not a network goal. Redundancy is a solution to a goal of high availability. Redundancy means adding duplicate links or devices to a network to avoid downtime. Redundant network topologies are becoming increasingly important for many network design customers who want to ensure business continuity after a major fault or disaster. Chapter 5, "Designing a Network Topology," covers designing redundant network topologies in more detail.	[redacted]
Availability is also associated with resiliency, which is a word that is becoming more popular in the networking field. Resiliency means how much stress a network can handle and how quickly the network	[redacted]

can rebound from problems including security breaches, natural and unnatural disasters, human error, and catastrophic software or hardware failures. A network that has good resiliency usually has good availability. Disaster Recovery Most large institutions have recognized the need for a plan to sustain business and technical operations after natural disasters, such as floods, fires, hurricanes, and earthquakes. Also, some large enterprises (especially service providers) must plan how to recover from satellite outages. Satellite outages can be caused by meteorite storms, collisions with space debris, solar flares, or system failures. Unfortunately, institutions have also found the need to specify a recovery plan for unnatural disasters, such as bombs, terrorist attacks, riots, or hostage situations. A disaster recovery plan includes a process for keeping data backed up in one or more	[redacted]

places that are unlikely to be hit by disaster, and a process for switching to backup technologies if the main technologies are affected by a disaster.

Although this book doesn't cover the details of disaster recovery planning, the concepts in this book can be applied to the process of planning for a disaster. Not surprisingly, a top- down approach is recommended, with an emphasis on planning before implementing. One goal of the planning process should be to recognize which parts of the network are critical and must be backed up. A good understanding of the organization's business purpose is needed to understand which devices, network links, applications, and people are critical. As is the case with top-down network design, business goals must be analyzed before selecting technologies and devices that will be one

component of the implementation.	[redacted]
Note Don't underestimate the importance of having enough staff to activate a disaster recovery plan. Have you figured out what to do if the disaster involves a serious disease where the server and network administrators need to be quarantined? This could be a justification for providing high-speed VPN access from workers' homes and testing that capability before a disaster strikes.	[redacted]
One of the most important steps in disaster recovery planning is testing. Not only must the technology be tested, but employees must be drilled on the actions they should take in a disaster. If the people don't survive, the technology won't help much. Also, people should practice working with the network in the configuration it will likely have after a disaster when redundant servers or sites are in use. Although	[redacted]

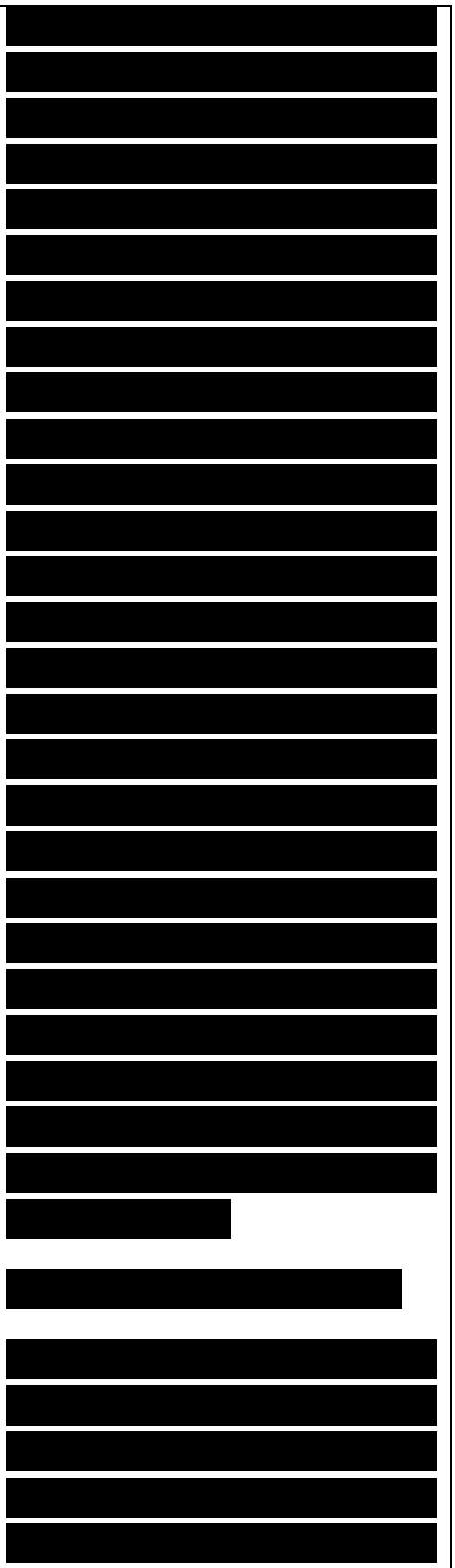
employees might object to emergency drills, especially if they are too frequent, periodic practice is a necessary part of achieving business continuity when a real disaster hits. The drills should be taken seriously and should be designed to include time and stress pressures to simulate the real thing.	[redacted]
Specifying Availability Requirements You should encourage your customers to specify availability requirements with precision. Consider the difference between an uptime of 99.70 percent and an uptime of 99.95 percent. An uptime of 99.70 percent means the network is down 30 minutes per week, which is not acceptable to many customers. An uptime of 99.95 percent means the network is down 5 minutes per week, which might be acceptable, depending on the type of business. Availability requirements should be	[redacted]

specified with at least two digits following the decimal point.	[redacted]
It is also important to specify a timeframe with percent uptime requirements. Go back to the example of 99.70 percent uptime, which equated to 30 minutes of downtime per week. A downtime of 30 minutes in the middle of a working day is probably not acceptable. But a downtime of 30 minutes every Saturday evening for regularly scheduled maintenance might be fine.	[redacted]
Not only should your customers specify a timeframe with percent uptime requirements, they should also specify a time unit. Availability requirements should be specified as uptime per year, month, week, day, or hour. Consider an uptime of 99.70 percent again. This uptime means 30 minutes of downtime during a week.	[redacted]

The downtime could be all at once, which could be a problem if it's not during a regularly scheduled maintenance window, or it could be spread out over the week. An uptime of 99.70 percent could mean that approximately every hour the network is down for 10.70 seconds. Will users notice a downtime of 10.70 seconds? Certainly some users will, but for some applications, a downtime of 10.70 seconds every hour is tolerable. Availability goals must be based on output from the first network design step of analyzing business goals, where you gained an understanding of the customer's applications.

Five Nines Availability

Although the examples cited so far use numbers in the 99.70 to 99.95 percent range, many companies require higher availability, especially during critical time periods. Some



customers might insist on a network uptime of 99.999 percent, which is sometimes referred to as five nines availability. For some customers, this requirement might be linked to a particular business process or timeframe. For example, the requirement might refer to the monthly closing of financial records or to the holiday season for a company that sells holiday gifts via catalog and web orders. On the other hand, some design customers might need, or think they need, five nines availability all the time. Five nines availability is extremely hard to achieve. You should explain to a network design customer that to achieve such a level, redundant equipment and links will be necessary, as will extra staffing possibly, and extremely reliable hardware and software.	[redacted]

Some managers will back down from such a requirement when they hear the cost, but, for others, the goal might be appropriate. If a company would experience a severe loss of revenue or reputation if the network were not operational for even short periods of time, five nines availability is a reasonable goal.	[redacted]
Many hardware manufacturers specify 99.999 percent uptime for their devices and operating systems and have real customer examples where this level of uptime was achieved. This might lead a naive network design customer to assume that a complex internetwork can also have 99.999 percent uptime without too much extra effort or cost. Achieving such a high level on a complex internetwork, however, is much more difficult than achieving it for particular components of the internetwork. Potential failures include carrier outages, faulty software in routers and switches, an	[redacted]

unexpected and sudden increase in bandwidth or server usage, configuration problems, human errors, power failures, security breaches, and software glitches in network applications. Note Some networking experts say that 80 to 90 percent of failures are due to human errors, either errors made by local administrators or errors made by service provider employees (or the infamous backhoe operator). Avoiding and recovering from human errors requires skill and good processes. You need smart people thinking about availability all the time and processes that are precise without stifling thought. Good network management and troubleshooting play a role. Network management tools should provide immediate alerts upon failures and enough information for a network administrator to make a quick fix.	[redacted]

Consider a network that is used 24 hours a day for 365 days per year. This equates to 8760 hours. If the network can be down only 0.001 percent of the time, it can be down for only 0.0876th of an hour or about 5 minutes per year. If the customer says the network must be available 99.999 percent of the time, you better make it clear that this doesn't include regularly scheduled maintenance time, or you better make sure that the network will have the capability to support in-service upgrades. In-service upgrades refer to mechanisms for upgrading network equipment and services without disrupting operations. Most internetworking vendors sell high-end internetworking devices that include hot-swappable components for in-service upgrading.

For situations where hot-swapping is not practical, it might be necessary to have extra equipment so there's never a need to disable

services for maintenance. In some networks, each critical component has triple redundancy, with one being active, one in hot standby ready to be used immediately, and one in standby or maintenance. With triple redundancy, you can bring a standby router down to upgrade or reconfigure it. After it is upgraded, you can then designate it as the hot standby, and take the previous hot standby down and upgrade it. You can then switch from the active to the hot standby and upgrade the active. Depending on the network design, you might load share among the redundant components during normal operations. The key design decision is whether your users can accept degraded performance when some of the components are unusable. If all this sounds too complicated or expensive, another possibility is not to do it all yourself but put resources at collocation centers that can amortize the highly redundant equipment over	[redacted]

many customers.	[redacted]
The Cost of Downtime	[redacted]
In general, a customer's goal for availability is to keep mission-critical applications running smoothly, with little or no downtime. A method to help you, the network designer, and your customer understand availability requirements is to specify a cost of downtime. For each critical application, document how much money the company loses per hour of downtime. (For some applications, such as order processing, specifying money lost per minute might have more impact.) If network operations will be outsourced to a third- party network management firm, explaining the cost of downtime can help the firm understand the criticality of applications to a business's mission. Specifying the cost of downtime can also help clarify whether in-service upgrades or triple redundancy must be	[redacted]

supported.	[redacted]
Mean Time Between Failure and Mean Time to Repair	[redacted]
In addition to expressing availability as the percent of uptime, you can define availability as a mean time between failure (MTBF) and mean time to repair (MTTR). You can use MTBF and MTTR to calculate availability goals when the customer wants to specify explicit periods of uptime and downtime, rather than a simple percent uptime value.	[redacted]
MTBF is a term that comes from the computer industry and is best suited to specifying how long a computer or computer component will last before it	[redacted]

fails. When specifying availability requirements in the networking field, MTBF is sometimes designated with the more cumbersome phrase mean time between service outage (MTBSO), to account for the fact that a network is a service, not a component. Similarly, MTTR can be replaced with the phrase mean time to service repair (MTTSR). This book uses the simpler and better-known terms MTBF and MTTR.

A typical MTBF goal for a network that is highly relied upon is 4000 hours. In other words, the network should not fail more often than once every 4000 hours or 166.67 days. A typical MTTR goal is 1 hour. In other words, the network failure should be fixed within 1 hour. In this case, the mean availability goal is as follows:
4000 / 4001 = 99.98 percent

A goal of 99.98 percent is typical for many companies.

When specifying availability using MTBF and MTTR, the equation to use is as follows: Availability = MTBF / (MTBF + MTTR)

Using this availability equation allows a customer to clearly state the acceptable frequency and length of network outages.

Remember that what is calculated is the mean. The variation in failure and repair times can be high and must be considered as well. It is not enough to just consider mean rates, especially if you depend on external service agents (vendors or contractors) who are not under your tight control. Also, be aware that customers might need to specify different MTBF and MTTR goals for different parts of a network. For example, the goals for the core of the enterprise network are probably much more stringent than the goals for a switch port that affects only one user.

Although not all customers can specify detailed application requirements, it is a good idea to identify availability goals for specific applications, in addition to the network as a whole. Application availability goals can vary widely depending on the cost of downtime. For each application that has a high cost of downtime, you should document the acceptable MTBF and MTTR.	[redacted]
For MTBF values for specific networking components, you can generally use data supplied by the vendor of the component. Most router, switch, and hub manufacturers can provide MTBF and MTTR figures for their products. You should also investigate other sources of information, such as trade publications, to avoid any credibility	[redacted]

problems with figures published by manufacturers. Search for variability figures and mean figures. Also, try to get written commitments for MTBF, MTTR, and variability values from the providers of equipment and services.

Network Performance

When analyzing technical requirements for a network design, you should isolate your customer's criteria for accepting the performance of a network, including throughput, accuracy, efficiency, delay, and response time.

Many mathematical treatises have been written on network performance. This book approaches network performance in a practical and mostly nonmathematical way, avoiding the daunting equations that appear in mathematical treatments of performance. Although the equations are much simpler than they seem, they are usually not necessary for understanding a customer's goals. The objective of this

section is to offer an uncomplicated view of network performance, including real-world conclusions you can draw when there is no time to do a mathematical analysis.

Analyzing a customer's network performance goals is tightly tied to analyzing the existing network, which is covered in Chapter 3, "Characterizing the Existing Internetwork." Analyzing the existing network can help you determine what changes need to be made to meet performance goals. Network performance goals are also tightly linked to scalability goals. You should gain an understanding of plans for network growth before analyzing performance goals.

Network Performance Definitions

Many network design customers cannot quantify their performance goals beyond, "It has to work with no complaints from users." If this is the case, you can make assumptions about throughput, response time, and so on. On the other

hand, some customers have specific performance requirements, based on a service level that has been agreed upon with network users.

The following list provides definitions for network performance goals that you can use when analyzing precise requirements:

- Capacity (bandwidth): The data-carrying capability of a circuit or network, usually measured in bits per second (bps)
- Utilization: The percent of total available capacity in use
- Optimum utilization: Maximum average utilization before the network is considered saturated
- Throughput: Quantity of error-free data successfully transferred between nodes per unit of time, usually seconds
- Offered load: Sum of all the data all network nodes have ready to send at a particular time
- Accuracy: The amount of useful traffic that is

correctly transmitted, relative to total traffic

■ Efficiency: An analysis of how much effort is required to produce a certain amount of data throughput

■ Delay (latency): Time between a frame being ready for transmission from a node and delivery of the frame elsewhere in the network

■ Delay variation: The amount of time average delay varies

■ Response time: The amount of time between a request for some network service and a response to the request

Optimum Network Utilization

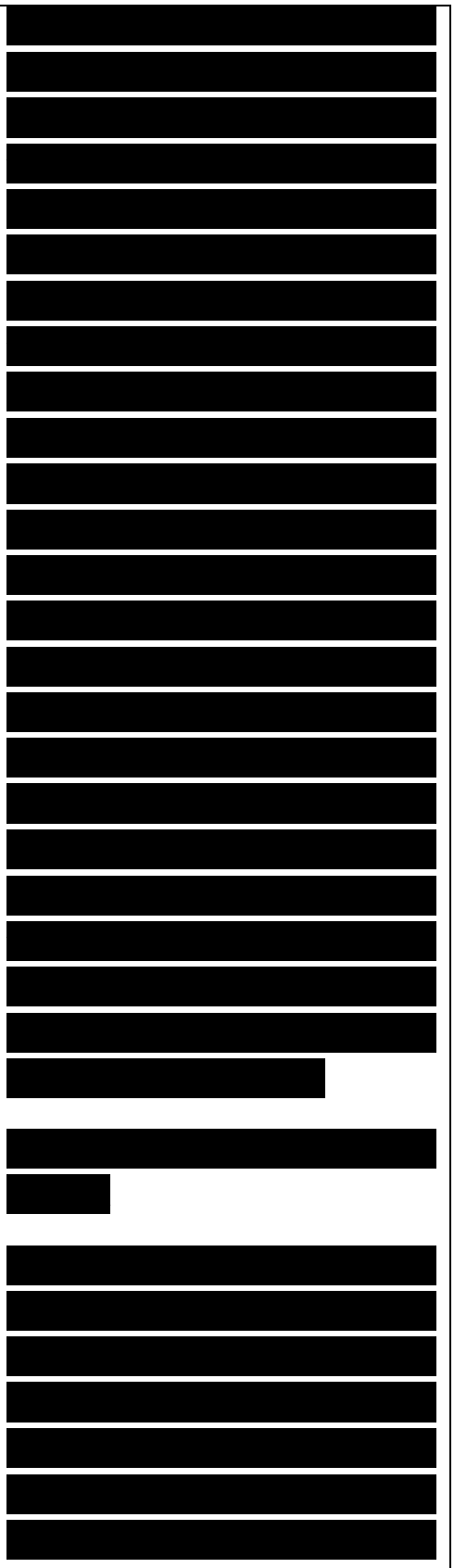
Network utilization is a measurement of how much bandwidth is used during a specific time period. Utilization is commonly specified as a percentage of capacity. For example, a network-monitoring tool might state that network utilization on an Ethernet segment is 30 percent, meaning that 30 percent of the capacity is in use.

Network-analysis tools use varying methods for measuring bandwidth usage and averaging the usage over elapsed time. Usage can be averaged every millisecond, every second, every minute, every hour, and so on. Some tools use a weighted average whereby more recent values are weighted more prominently than older values. Chapter 3 discusses measuring network utilization in more depth.	[redacted]
Your customer might have a network design goal for the maximum average network utilization allowed on a segment. Actually, this is a design constraint more than a design goal. The design constraint states that if utilization on a segment is more than a predefined threshold, the segment should be divided into multiple segments or bandwidth must be added.	[redacted]
Optimum average network	[redacted]

utilization is about 70 percent. A 70 percent threshold for average utilization means that peaks in network traffic can probably be handled without obvious performance degradation. Most WANs have less capacity than LANs, so more care is needed in selecting WAN bandwidth that can cover actual and reasonable variations. Customers have many options for technologies that can reduce bandwidth utilization on WANs, including advanced routing-protocol features and compression. Chapter 13, "Optimizing Your Network Design," covers optimizing bandwidth utilization in more detail.	[redacted]
With LANs, less attention is paid to monitoring network utilization because many LANs are already overbuilt with full-duplex Gigabit Ethernet links to servers and 100-Mbps or Gigabit Ethernet links to clients. If configured for full-duplex operations, which is typical these days, a Fast or Gigabit Ethernet link supports simultaneous transmitting	[redacted]

and receiving. So, in theory, a 100-Mbps Fast Ethernet segment could support 100 percent utilization of the transmit channel and 100 percent utilization of the receive channel, using 200 Mbps. However, total bandwidth in both directions isn't used all the time in most cases. Consider the case of a client system communicating with a server. The client sends requests and the server responds, in lock step. The client doesn't try to send at the same time as the server, so the bandwidth usage does not double on the client's link to the Ethernet switch.

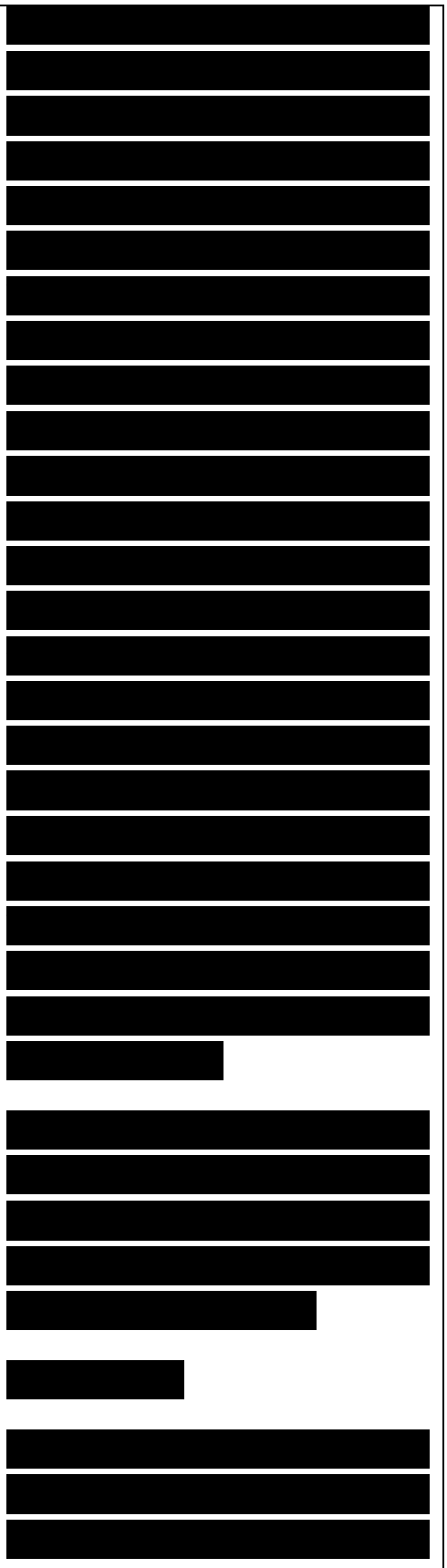
A point-to-point full-duplex link that connects a switch to a server or to another switch, on the other hand, could use all the bandwidth, depending on traffic patterns. Full-duplex Ethernet has become the standard method for connecting servers, switches,



and end users' machines. It's an essential performance boost for servers, in particular. With full-duplex Ethernet, a switch can transmit the next client's request at the same time the server is sending a response to a previous request. If the utilization exceeds about 70 percent of the full-duplex bandwidth, however, it's probably time to upgrade to more bandwidth. Network traffic is bursty. You should provision both LAN and WAN capacity with the assumption that the average utilization will be exceeded during bursts.

Throughput

Throughput is defined as the quantity of error-free data that is transmitted per unit of time. Throughput is often



defined for a specific connection or session, but in some cases the total throughput of a network is specified. Network novices consistently misuse the words throughput and bandwidth. Remember, bandwidth means capacity and is generally fixed. Throughput is an assessment of the amount of data that can be transmitted per unit of time. You measure throughput, which can vary depending on network performance characteristics and how you make the measurement. Bandwidth is a given. Note To understand bandwidth and throughput, think of a steel pipe that has a capacity of 100 gallons per minute. The pipe has fixed capacity (bandwidth). If just a trickle is coming through, throughput is low. If throughput is at 70 percent, you may have a flood. Ideally, throughput should be the same as capacity. However, this is not the case on real networks. Capacity	[redacted]

depends on the physical layer technologies in use. The capacity of a network should be adequate to handle the offered load, even when there are peaks in network traffic. (Offered load is the data that all nodes have to send at a particular moment in time.) Theoretically, throughput should increase as offered load increases, up to a max-imum of the full capacity of the network. However, network throughput depends on the access method (for example, token passing or carrier sensing), the load on the network, and the error rate.

Figure 2-1 shows the ideal situation, where throughput increases linearly with the offered load, and the real world, where actual throughput tapers off as the offered load reaches a certain maximum.

Figure 2-1 Offered Load and Throughput

Throughput of Internetworking Devices

Some customers specify throughput goals in terms of

the number of packets per second (pps) an internetworking device must process. (In the case of an ATM device, the goal is cells per second, or [cps].) The throughput for an internetworking device is the maximum rate at which the device can forward packets without dropping any packets. Most internetworking vendors publish pps ratings for their products, based on their own tests and independent tests. To test an internetworking device, engineers place the device between traffic generators and a traffic checker. The traffic generators send packets ranging in size from 64 bytes to 1518 bytes for Ethernet. By running multiple generators, the investigation can test devices with multiple ports. The generators send bursts of traffic through the device at an initial rate that is half of what is theoretically possible for test conditions. If all packets are received, the rate is increased. If all packets are not received, the	[redacted]

rate is decreased. This process is repeated until the highest rate at which packets can be forwarded without loss is determined. Pps values for small frames are much higher than pps values for large frames, so be sure you understand which value you are looking at when reading vendor test results for an internetworking device.

Many internetworking devices can forward packets at the theoretical maximum, which is also called wire speed . The theoretical maximum is calculated by dividing bandwidth by packet size, including any headers, preambles, and interframe gaps. Table 2-1 shows the theoretical maximum pps for one 100-Mbps Ethernet stream, based on frame size.

To rate the pps value for a multiport device, testers send multiple streams of data through the device to multiple output ports. The extreme numbers that you sometimes see in vendor marketing material (for

example, 400 million pps for the Cisco Catalyst 6500 switch) come from measurements made with multiple Gigabit Ethernet data flows, each using 64-byte packets.	
Table 2-1 Theoretical Maximum Packets per Second (pps)	
Frame Size (in Bytes) 100-Mbps Ethernet Maximum PPS 1-Gbps Ethernet Maximum pps	
Application Layer Throughput	
Most end users are concerned about the throughput for applications. Marketing materials from some networking vendors refer to application layer throughput as goodput. Calling it goodput sheds light on the fact that it is a measurement of good and relevant application layer data transmitted per unit of time.	
It is possible to improve throughput, such that more data per second is transmitted, but not increase	

goodput, because the extra data transmitted is overhead or retransmissions. Keep in mind what throughput means (bytes per second). Are these good (useful) application layer bytes or simply bytes used by the protocol to get its job done? It is also possible to increase throughput by not using compression. More data is transmitted per unit of time, but the user sees worse performance.	[redacted]
A simple goal for throughput based on data-per-second rates between stations does not identify the requirements for specific applications. When specifying throughput goals for applications, make it clear that the goal specifies good (error-free) application layer data per unit of time. Application layer throughput is usually measured in kilobytes per second (KBps) or megabytes per second (MBps).	[redacted]

Work with your customer to identify throughput requirements for all applications that can benefit from maximized application layer throughput, such as file transfer and database applications. (Throughput is not important for all applications; for example, some interactive character-based applications don't need large screen updates.) Explain to your customer the factors that constrain application layer throughput, which include the following:	[redacted]
■ End-to-end error rates	[redacted]
■ Protocol functions, such as handshaking, windows, and acknowledgments	[redacted]
■ Protocol parameters, such as frame size and retransmission timers	[redacted]
■ The pps or cps rate of internetworking devices	[redacted]
■ Lost packets or cells at internetworking devices	[redacted]
■ Workstation and server performance factors:	[redacted]

■ Disk-access speed ■ Disk-caching size ■ Device driver performance ■ Computer bus performance (capacity and arbitration methods)	
■ Processor (CPU) performance ■ Memory performance (access time for real and virtual memory) ■ Operating system inefficiencies ■ Application inefficiencies or bugs	
If necessary, work with your customer to identify application throughput problems caused by errors or inefficiencies in protocols, operating systems, and applications. Protocol analyzers are important tools for this. Chapter 3 discusses isolating performance problems in more detail.	
Accuracy The overall goal for accuracy is that the data	

received at the destination must be the same as the data sent by the source. Typical causes of data errors include power surges or spikes, impedance mismatch problems, poor physical connections, failing devices, and noise caused by electrical machinery. Sometimes software bugs can cause data errors also, although software problems are a less common cause of errors than physical layer problems. Frames that have an error must be retransmitted, which has a negative effect on throughput. In the case of IP networks, Transmission Control Protocol (TCP) provides retransmission of data. For WAN links, accuracy goals can be specified as a bit error rate (BER) threshold. If the error rate goes above the specified BER, the accuracy is considered unacceptable. Analog links have a typical BER threshold of about 1 in 105. Digital circuits have a much lower error rate than analog circuits, especially if fiber-optic cable is used. Fiber-optic links have an	[redacted]

error rate of about 1 in 1011. Copper links have an error rate of about 1 in 106.	[redacted]
For LANs, a BER is not usually specified, mainly because measuring tools such as protocol analyzers focus on frames, not bits; however, you can approximate a BER by comparing the number of frames with errors in them to the total number of bytes seen by the measuring tool. A good threshold to use is that there should not be more than one bad frame per 106 bytes of data.	[redacted]
On shared Ethernet, errors are often the result of collisions. Two stations try to send a frame at the same time and the resulting collision damages the frames, causing cyclic redundancy check (CRC) errors. Depending on the size of the Ethernet network, many of these collisions happen in the 8-byte preamble of the frames and are not registered by troubleshooting tools. If the collision happens past the	[redacted]

preamble and somewhere in the first 64 bytes of the data frame, this is registered as a legal collision, and the frame is called a runt frame. A general goal for Ethernet collisions is that less than 0.1 percent of the frames should be affected by a legal collision (not counting the collisions that happen in the preamble). A collision that happens beyond the first 64 bytes of a frame is a late collision. Late collisions are illegal and should never happen. Ethernet networks that are too large experience late collisions because stations sending minimum-sized frames cannot hear other stations within the allowed timeframe. The extra propagation delay caused by the excessive size of the network causes late collisions between the most widely separated nodes. Faulty repeaters and network interface cards (NIC) can also cause late collisions. Collisions should never	[redacted]

occur on full-duplex Ethernet links. If they do, there's probably a duplex mismatch. Collisions on a properly configured full-duplex link have no meaning. Both stations sending at the same time is normal. Receiving while sending is normal. So, there is no need for collision detection and collisions shouldn't occur. Chapter 3 has more to say about duplex mismatch problems and how to recognize if they cause errors on your networks.	[redacted]
Collisions also never occur on WAN links. Unfortunately, the output of the show interface serial command on Cisco routers includes a collision count. It should be ignored. Cisco programmers used a template for this part of the output. The template is based on the output from the show interface ethernet command. There are no collisions on a serial interface, regardless of the encapsulation or technology. Collisions occur only on carrier sense multiple access (CSMA)	[redacted]

networks including Ethernet, 802.3, LocalTalk, Aloha, and 802.11 networks. Collisions are a normal part of the "management-by-contention" approach that defines CSMA. (And although LocalTalk and 802.11 use CSMA with collision avoidance, collisions can still occur.)	[redacted]
Accuracy usually refers to the number of error-free frames transmitted relative to the total number of frames transmitted. Accuracy can also characterize how often the network reorders sequences of packets. Packet reordering occurs in many situations, including the use of parallel switching fabrics within a single network device and the use of parallel links between routers. Although upper-layer protocols, such as TCP and Real-Time Transport Protocol (RTP), correct for the reordering of packets, the problem can cause minor performance degradation. Some applications don't use	[redacted]

a protocol that corrects the problem and thus might be more severely affected. Because the problem is often corrected, it can be hard to detect. IP routers are not designed to detect, let alone correct, packet reordering, and because they do not detect this condition, they cannot report the problem to network management software. Measurements must be made at end hosts. For example, you could use a protocol analyzer on an end-station host to detect the reordering of packets. Efficiency Efficiency is a term borrowed from engineering and scientific fields. It is a measurement of how effective an operation is in comparison to the cost in effort, energy, time, or money. Efficiency specifies how much overhead is required to produce a required outcome. For example, you could measure	[redacted]

the efficiency of a method for boiling water. Does most of the energy go to actually boiling the water or does a lot of the energy get wasted heating the electrical wiring, the pot the water is in, and the air around it? How much overhead is required to produce the desired outcome? Efficiency also provides a useful way to talk about network performance. For example, shared Ethernet is inefficient when the collision rate is high. (The amount of effort to successfully send a frame becomes considerable because so many frames experience collisions.) Network efficiency specifies how much overhead is required to send traffic, whether that overhead is caused by collisions, token passing, error reporting, rerouting, acknowledgments, large frame headers, a bad network design, and so on. Large frame headers are one cause for inefficiency. We worry a lot less about frame headers than we used to when bandwidth was	[redacted]

scarcer. Nonetheless, for networks where bandwidth is still (or may become) scarce, a good network performance goal is for applications that send bulk data to minimize the amount of bandwidth used by headers by using the largest possible frame the MAC layer allows. Using a large frame maximizes the amount of useful application data compared to header data and improves application layer throughput.

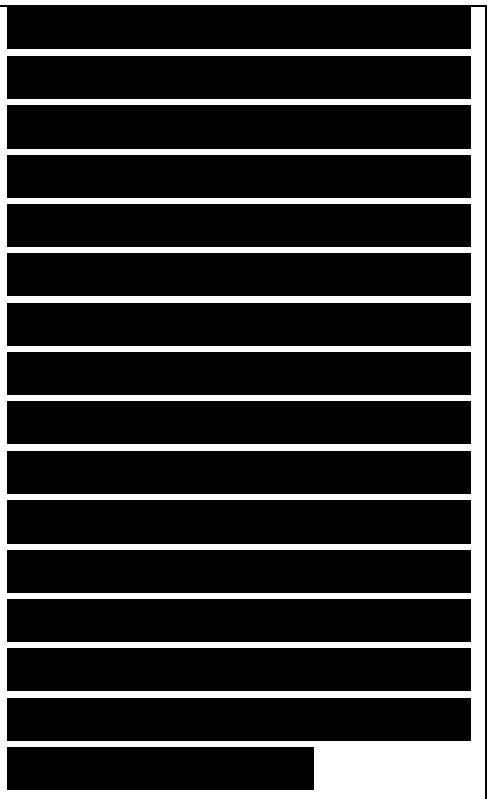
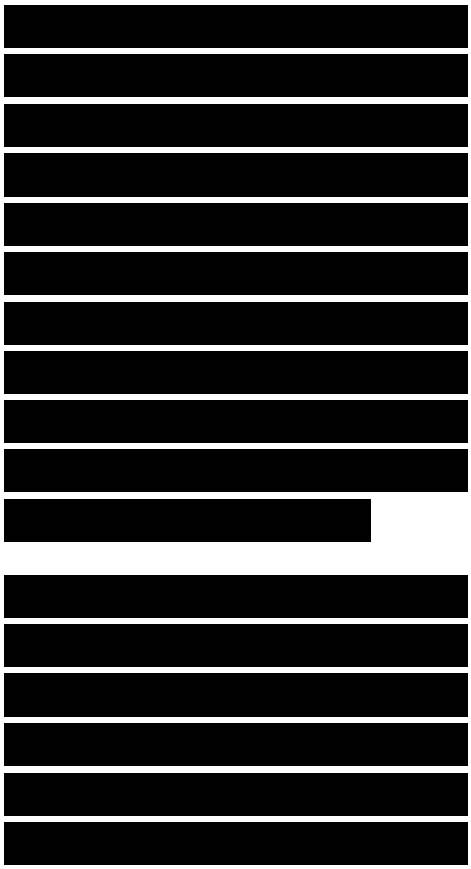


Figure 2-2 shows a bandwidth pipe used by small frames and the same pipe used by large frames. The header of each frame is shaded. Note that there is an interframe gap between each frame in addition to the headers. From the graphic, you can see that large frames use bandwidth more efficiently than small frames.



The maximum frame size is a tradeoff with the BER discussed in the previous section. Bigger frames have more bits and hence are more likely to be hit by an error. If there were no errors, an infinitely big frame would

be the most efficient (although not the most fair to other senders). If a frame is hit by an error, it must be retransmitted, which wastes time and effort and reduces efficiency. The bigger the frame, the more bandwidth is wasted retransmitting. So, because networks experience errors, frame sizes are limited to maximize efficiency and fairness. The maximum frame size for Ethernet, for example, is 1522 bytes, including the header, CRC, and an 802.1Q VLAN tag.

As is the case with many network design goals, there are tradeoffs associated with a goal of improving efficiency by using large frame sizes. On slow WAN links, the time to output a large frame is significant. The time to output a frame is called serialization delay. Serialization delay becomes an issue when applications that send large frames, such as file transfer, share a WAN link with applications that are delay-sensitive, such as voice and video. One

solution is to use ATM, which divides frames into cells. Other solutions include the use of link-layer fragmentation and interleaving options, such as Frame Relay FRF.12, Multilink Frame Relay (FRF.16), and Multilink PPP.	
Delay and Delay Variation	
Users of interactive applications expect minimal delay in receiving feedback from the network. Voice and video applications also require minimal delay. In addition, voice and video applications require a minimal variation in the amount of delay that packets experience. Variations in delay, called jitter, cause disruptions in voice quality and jumpiness in video streams.	
Applications that use the Telnet protocol are also sensitive to delay because the user expects quick feedback when typing	

characters. Telnet is becoming obsolete, but it hasn't disappeared yet. With the Telnet remote echo option, the character typed by a user doesn't appear on the screen until it has been acknowledged and echoed by the far end, and the near end has sent an acknowledgment for the echo. To help you recognize the need to design a network with low delay, you should determine if your customer plans to run any delay-sensitive applications, such as voice or video, or applications based on delay-sensitive protocols such as Telnet.

Causes of Delay

Any goals regarding delay must take into account fundamental physics. Despite science fiction stories that say differently, any signal experiences a propagation delay resulting from the finite speed of light, which is about 300,000 kilometers per second (186,000 miles per second). Network designers can also remember 1 nanosecond per foot. These values are for light traveling in a vacuum. A signal in a cable or optical

fiber travels approximately two-thirds the speed of light in a vacuum. Delay is relevant for all data transmission technologies but especially for satellite links and long terrestrial cables. Geostationary satellites are in orbit above the earth at a height of about 36,000 kilometers, or 24,000 miles. This long distance leads to a propagation delay of about 270 milliseconds (ms) for an intercontinental satellite hop. In the case of terrestrial cable connections, propagation delay is about 1 ms for every 200 kilometers (120 miles). Another fundamental cause for delay is serialization delay, the time to put digital data onto a transmission line, which depends on the data volume and the speed of the line. For example, to transmit a 1024-byte packet on a 1.544-Mbps T1 line takes about 5 ms. An additional fundamental	[redacted]

delay is packet-switching delay. Packet-switching delay refers to the latency accrued when switches and routers forward data. The latency depends on the speed of the internal circuitry and CPU, and the switching architecture of the internetworking device. Latency also depends on the type of RAM that the device uses. Dynamic RAM (DRAM) needs to be refreshed thousands of times per second. Static RAM (SRAM) doesn't need to be refreshed, which makes it faster, but it is also more expensive than DRAM. Low-end internetworking devices often use DRAM to keep the cost low.

Packet-switching delay can be quite small on high-end switches, in the 5- to 20-microsecond range for 64-byte Ethernet frames. Routers tend to introduce more latency than switches. The amount of latency that a router causes for packet switching depends on many variables, including the router architecture, configuration, and software features that optimize the forwarding of packets. Despite marketing claims by

switch salespeople, you should not assume that a router has higher latency than a switch. A high-end router with a fast CPU, SRAM, optimized software, and a highly evolved switching fabric can outperform many low-end or medium-end switches.

Of course, a router has a more complicated job than a Layer 2 switch. In general terms, when a packet comes into a router, the router checks its routing table, decides which interface should send the packet, and encapsulates the packet with the correct data link layer header and trailer. Routing vendors, such as Cisco, have advanced caching mechanisms so that a frame destined for a known destination can receive its new encapsulation quickly without requiring the CPU to do any table lookup or other processing. These mechanisms minimize packet-switching delay.

Packet-switching speed depends on the type and number of advanced features that are enabled on a packet-switching device. When designing an internetwork fabric, consider the power that you will need to incorporate into the design to implement quality of service (QoS), Network Address Translation (NAT), IPsec, filtering, and so on. Consider the policies that your design customer wants to enforce and the effect they will have on packet-switching delay. Packet-switching delay can also include queuing delay . The average number of packets in a queue on a packet-switching device increases exponentially as utilization increases, as you can see from Figure 2-3. If utilization is 50 percent, the average queue depth is one packet. If utilization is 90 percent, the average queue depth is nine packets. Without going into mathematical queuing theory, the general rule of thumb for queue depth is as follows:	[redacted]

Queue depth = Utilization / (1 - Utilization)

Average Utilization

Figure 2-3 Queue Depth and Bandwidth Utilization

Consider the following example. A packet switch has five users, each offering packets at a rate of 10 pps. The average length of the packets is 1024 bits. The packet switch needs to transmit this data over a 56-kbps WAN circuit. Putting all this together, you have the following equations:

Load = 5 x 10 x 1024 = 51,200 bps Utilization = 51,200 / 56,000 = 91.4 percent

Average number of packets in the queue = (0.914) / (1 - 0.914) = 10.63 packets

By increasing bandwidth on a WAN circuit, you can decrease queue depth and hence decrease delay. Alternatively, to improve performance, you can use an advanced queuing algorithm that outputs certain types of

packets first—for example, voice or video packets. Chapter 13 covers advanced router queuing techniques in more detail.

Delay Variation

As customers implement new digital voice and video applications, they are becoming concerned about delay and delay variation. Additionally, customers are becoming more aware of the issues associated with supporting bursty traffic on the same network that carries delay-sensitive traffic. If bursts in traffic cause jitter, audio and video streams experience problems that disrupt communications.

Desktop audio/video applications can minimize jitter by providing a jitter buffer. Display software or hardware pulls data from the

buffer. The insulating buffer reduces the effect of jitter because variations on the input side are smaller than the total buffer size and therefore not obvious on the output side. The data is smoothed in the output, and the user experiences no ill effects from the input jitter.	[redacted]
If possible, you should gather exact requirements for delay variation from a customer. For customers who cannot provide exact goals, a good rule of thumb is that the variation should be less than 1 or 2 percent of the delay. For example, for a goal of an average delay of 40 ms, the variation should not be more than 400 or 800 microseconds.	[redacted]
Short fixed-length cells, such as ATM 53-byte cells, are inherently better than frames for meeting delay and delay-variance goals. To help	[redacted]

understand this concept, consider the analogy of people trying to get onto an escalator. The escalator is like a bandwidth pipe. At first, each person gets onto the escalator in an orderly fashion and the delay is predictable. Then a school class arrives and the children are all holding hands, expecting to get onto the escalator all at once! What happens to your delay if you happen to be behind the children?

A gaggle of school children holding hands is analogous to a large frame causing extra delay for small frames. Consider the case of a user starting a file transfer using 1518-byte frames. This user's data affects bandwidth usage and queuing mechanisms at internetworking devices, causing unexpected delay for other traffic. Good throughput for one application causes delay problems for another application.

Cell-relay technologies (ATM, for example) were designed to support traffic that is sensitive to delay and jitter. Depending on the class of service, ATM lets a session specify a maximum cell transfer delay (MCTD) and maximum cell delay variation (MCDV).

Chapter 4 describes ATM service classes in more detail.

Response Time

Response time is the network performance goal that users care about most. Users don't know about propagation delay and jitter. They don't understand throughput in pps or in MBps. They aren't concerned about BERs, although perhaps they should be! Users recognize the amount of time to receive a response from the network system. They also recognize small changes in the expected response time and become frustrated when the response time is long.

Users begin to get frustrated

when response time is more than about 100 ms or 1/10th of a second. Beyond 100 ms, users notice they are waiting for the network to display a web page, echo a typed character, start downloading email, and so on. If the response happens within 100 ms, most users do not notice any delay. The 100-ms threshold is often used as a timer value for protocols that offer reliable transport of data. For example, many TCP implementations retransmit unacknowledged data after 100 ms by default. Note Good TCP implementations also adjust the retransmit timer based on network conditions. TCP should keep track of the average amount of time to receive a response and dynamically adjust the retransmit timer based on the expected delay. The 100-ms response time threshold applies to interactive applications. For bulk applications, such as transferring large files or graphical web pages, users are willing to wait at least 10	[redacted]

to 20 seconds. Technically savvy users expect to wait even longer if they know the file is large and the transmission medium is slow. If your network users are not technically savvy, you should provide some guidelines on how long to wait, depending on the size of files and the technologies in use (modems, high-speed digital networks, geostationary satellites, and so on).

Security

Security is a key technical goal, and security design is one of the most important aspects of enterprise network design. Increased threats from both inside and outside the enterprise network require the most up-to-date security rules and technologies. An overall goal that most companies have is that security problems should not disrupt the company's ability to conduct business. Network design customers need assurances that a design offers protection against business data and other

assets getting damaged or accessed inappropriately. Every company has trade secrets, business operations, and equipment to protect.	[redacted]
The first task in security design is planning. Planning involves identifying network assets that must be protected, analyzing risks, and developing requirements. This chapter briefly discusses security planning. Chapter 8, "Developing Network Security Strategies," covers planning for secure networks in more detail.	[redacted]
As is the case with most technical design requirements, achieving security goals means making tradeoffs. Security implementations can add to the cost of deploying and operating a network. Strict security policies can also affect the productivity of users, especially if some ease of use must be sacrificed to protect resources and data. Poor security implementations can annoy	[redacted]

users, causing them to think of ways to get around security policies. Security can also affect the redundancy of a network design if all traffic must pass through encryption devices, for example.	[redacted]
It is common practice to build systems with just enough security to bring potential losses from a security breach down to a desired level. A practical goal is to ensure that the cost to implement security does not exceed the cost to recover from security incidents. Alternatively, some organizations might want to implement stronger measures to mitigate unforeseen risks. As you work with your customer, you should analyze the cost associated with security incidents disrupting business and determine whether the customer wants to try to address unexpected problems.	[redacted]

Identifying Network Assets

The first step in security design is identifying the assets that must be protected, the value of the assets, and the expected cost associated with losing these assets if a security breach occurs. Network assets include hardware, software, applications, and data. Assets also include intellectual property, trade secrets, and a company's reputation.

Consider the possibility of a hacker damaging an enterprise's reputation by changing the enterprise's public web pages. You may have read about some of the cases of hackers changing U.S. government web pages. These security breaches affected the government's reputation in two ways: The changed web pages had silly graphics and text, and the government lost credibility because it appeared that it was easy to hack into government networks.

The data that a company uses to achieve its mission is an often-overlooked asset. Data can include engineering

blueprints, financial planning documents, customer relations information, competitive analysis documents, configuration information for hardware and software, employee Social Security numbers, employee badge information, and so on. The integrity and confidentiality of this data must be protected from intentional or unin-tentional damage. Some of the most important network assets are the networking devices themselves, including servers, switches and routers, and especially the firewalls and intrusion detection systems (IDS) that provide security services to network users. These devices are attractive targets to hackers and must be hardened (strengthened) against intrusions. As Chapter 8 discusses in more detail, hardening network devices involves running only the minimal necessary services, establishing trust only with authenticated partners, using secure device-management channels, and patching the device software to install fixes for known security problems.	[redacted]

You should consider more than just data and devices when identifying assets. The network user's time can be considered an asset. Whenever a virus attacks a system, it takes time to get rid of it, even if it's innocuous. The fact that this time is wasted is similar to a denial-of- service (D0S) attack. An asset might also be the capability to offer services to customers. This is especially true for Internet service providers (ISP), but also true for many other companies that offer medical, educational, financial, and other types of services.	[redacted]
Every design customer has different business assets and varying needs regarding the importance of assets. As a network designer, you should work with technical and business managers to identify which assets are critical to a business's mission. A financial services business, for example, has different assets than a health organization or a biomedical research company. As part of	[redacted]

the first step of network design, analyzing business requirements, you should have developed a good understanding of your network design customer's overall business mission (which might be different from the corporate mission statement, by the way, which is often written in a lofty manner to motivate employees and impress shareholders).

Analyzing Security Risks

In addition to identifying assets, an important step in security planning is analyzing potential threats and gaining an understanding of their likelihood and business impact. Risk analysis and the consequent building of a security policy and secure network design is a continuous process, as risks change in their severity and probability on a regular basis. For example, a company's encryption algorithm and the length of the encryption key might need to be reconsidered if a relatively inexpensive and exceptionally fast

codecracking computer becomes available, which allows easier decryption of valuable secrets. Risk assessment includes an analysis of the danger of not taking any action. Ask your customer to help you understand the risks associated with not implementing a secure network. How sensitive is the customer's data? What would be the financial cost of someone accessing the data and stealing trade secrets? What would be the financial cost of someone changing the data? What would be the financial cost associated with the network being down because of a security breach, causing employees to be unable to do their jobs? As mentioned previously, one of the biggest risks that must be managed is the risk that a hacker can undermine the security of a network device, such as a switch, router, server, firewall, or IDS. When a network device is compromised, the following threats arise:	[redacted]

■ Data flowing through the network can be intercepted, analyzed, altered, or deleted, compromising integrity and confidentiality. ■ Additional, related network services, which rely on trust among network devices, can be compromised. For example, bad routing data or incorrect authentication information could be injected into the network. ■ User passwords can be compromised and used for further intrusions and perhaps to reach out and attack other networks. ■ The configuration of the device can be altered to allow connections that shouldn't be allowed or to disallow connections that should be allowed. Some customers worry about hackers using protocol analyzers to sniff packets to see passwords, credit cards numbers, or other private data. This is not as big a risk as it appears. Credit card numbers are almost always sent encrypted, using technologies such as the Secure Sockets Layer (SSL)	[redacted]

protocol. Passwords are also sent encrypted and are often good for only one use anyway, if one-time passwords (OTP) are used. Even when passwords or credit cards are not encrypted, it is extremely difficult to find these minute pieces of data in the midst of millions of sniffed packets. Also, to sniff relevant packets, a hacker needs physical access to a link that carries relevant traffic or needs to have compromised a switch that supports port monitoring.

Hackers are getting more creative, though. Hackers disguised as customers, repair technicians, and contractors have been known to walk into organizations and gain network access via a network connection in an empty cubicle or a conference room. Sometimes companies have demo rooms, where they showcase their products. For ease of use of the people who configure the products, these rooms sometimes have access to the company's intranet and to the Internet. Hackers love that sort of

setup. Hackers who are less brazen, and don't want to walk into an organization's building, often sit outside in the parking lot with a wireless 802.11-enabled notebook computer or wireless handheld device and access corporate networks where the security was not well planned. In addition to considering outside hackers as a security risk, companies should heed problems caused by inept or malicious internal network users. Attacks might come from inadvertent user errors, including the downloading of software from untrusted sites that introduce malware. Attacks might also come from malicious acts by internal users, including employees disgruntled by cost cuts, employees who become greedy during tough economic times, and employees with a political agenda. Organizations should have information security training and awareness programs to mitigate the risk of internal user attacks.	[redacted]

Reconnaissance Attacks

A set of security risks falls into the category of a reconnaissance attack . A reconnaissance attack provides information about potential targets and their weaknesses and is usually carried out in preparation for a more focused attack against a particular target. Reconnaissance attackers use tools to discover the reachability of hosts, subnets, services, and applications. In some cases the tools are relatively sophisticated and can break through firewalls. A less-sophisticated hacker could convince users to download a file from an alleged music, video, pornographic, or game website. The file could actually be a Trojan horse that gathers reconnaissance data.

During a reconnaissance

attack, the attacker might make the following attempts to learn more about the network: ■ Gather information about the network's configuration and management from Domain Name System (DNS) registries. ■ Discover access possibilities using "war dialing" (attempts to discover and connect to dialup access points) and "war driving" (attempts to discover and connect to miscon- figured wireless access points). ■ Gather information about a network's topology and addressing using network mapping tools. Some tools, such as traceroute and Simple Network Management	[redacted] war driving: cũng là một loại tấn công gọi điện, tức "chạy xe" lang thang để tìm kiếm mạng không dây Wi-Fi nào mắc lỗi bảo mật, sử dụng máy tính xách tay hay thiết bị hỗ trợ cá nhân số – để xâm nhập vào mạng không dây của công ty [redacted]

Protocol (SNMP) queries, are primitive. Others are sophisticated and can send seemingly legitimate packets to map a network. ■ Discover the reachability of hosts, services, and applications using ping scans and port scans. ■ Discover operating system and application versions and probe for well-known security holes in the software. ■ Discover temporary holes created while systems, configurations, and software releases are being upgraded. Denial-of-Service Attacks Denial-of-service (DoS) attacks target the availability	[redacted]

of a network, host, or application, making it impossible for legitimate users to gain access. DoS attacks are a major risk because they can easily interrupt business processes and are relatively simple to conduct, even by an unskilled attacker. DoS attacks include the flooding of public servers with enormous numbers of connection requests, rendering the server unresponsive to legiti-mate users, and the flooding of network connections with random traffic, in an attempt to consume as much bandwidth as possible. Distributed denial-of-service (DDoS) attacks are even worse than DoS attacks because the attacker marshals multiple hosts, from various networks, to attack the target. DoS attacks are usually the consequence of a network's, host's, or application's inability to handle an enormous quantity of data, which crashes the system or halts services on the system. DoS attacks also take advantage of a host's or	[redacted]

application's failure to handle unexpected conditions, such as maliciously formatted input data or a buffer overflow.

DoS attacks are one of the most significant risks that a company must recognize and manage, because they have the capability to cause significant downtime.

Developing Security Requirements

Security problems should not disrupt an organization's capability to conduct business. That's the most basic security requirement that every organization has. A secondary security requirement is to protect assets from being incapacitated, stolen, altered, or harmed. Although every design customer has different detailed security requirements, basic requirements boil down to the need to develop and select procedures and technologies that ensure the following:

■ Confidentiality of data so that only authorized users can view sensitive information

■ Integrity of data so that only authorized users can change sensitive information and so that authorized users of data can depend on its authenticity	[redacted]
■ System and data availability, which should provide uninterrupted access to important computing resources	[redacted]
Other, more specific requirements could include one or more of the following goals:	[redacted]
■ Let outsiders (customers, vendors, suppliers) access data on public web or File Transfer Protocol (FTP) servers but not access internal data.	[redacted]
■ Authorize and authenticate branch-office users, mobile users, and telecommuters.	[redacted]
■ Detect intruders and isolate the amount of damage they do.	[redacted]

■ Authenticate routing-table updates received from internal or external routers. ■ Protect data transmitted to remote sites across a VPN. ■ Physically secure hosts and internetworking devices (for example, keep devices in a locked room). ■ Logically secure hosts and internetworking devices with user accounts and access rights for directories and files. ■ Protect applications and data from software viruses. ■ Train network users and network managers on security risks and how to avoid security problems. ■ Implement copyright or other legal methods of protecting products and intellectual property. ■ Meet compliance and regulatory requirements.	

Manageability

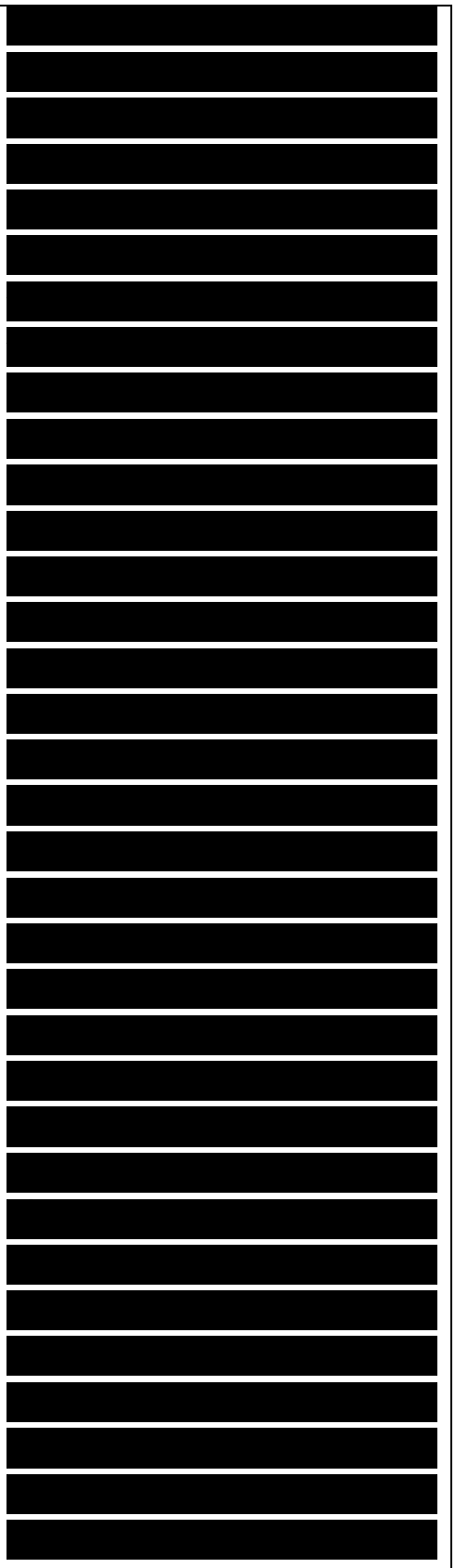
Every customer has different objectives regarding the manageability of a network. Some customers have precise goals, such as a plan to use SNMP to record the number of bytes each router receives and sends. Other clients have less-specific goals. If your client has definite plans, be sure to document them, because you will need to refer to the plans when selecting equipment. In some cases, equipment has to be ruled out because it does not support the management functions a customer requires.

Network management is discussed in more detail in Chapter 9, "Developing Network Management Strategies," but it's also important to consider management at the onset of a design project. During the initial gathering of technical requirements for a new network design or upgrade, you can use International Organization for Standardization (ISO) terminology to simplify the

discussion of network management goals with your design customer. ISO uses the FCAPS acronym to help you remember the following network management functions: ■ Fault management: Detecting, isolating, and correcting problems; reporting problems to end users and managers; tracking trends related to problems ■ Configuration management: Controlling, operating, identifying, and collecting data from managed devices ■ Accounting management: Accounting of network usage to allocate costs to network users and/or plan for changes in capacity requirements ■ Performance management: Analyzing traffic and application behavior to optimize a network, meet service-level agreements, and plan for expansion ■ Security management: Monitoring and testing security and protection policies, maintaining and distributing passwords and	[redacted]

other authentication and authorization information, managing encryption keys, and auditing adherence to security policies Usability A goal that is related to manageability, but is not exactly the same as manageability, is usability. Usability refers to the ease of use with which network users can access the network and services. Whereas manageability focuses on making network managers' jobs easier, usability focuses on making network users' jobs easier. It is important to gain an understanding of how important usability is to your network design customer, because some network design components can have a negative effect on usability. For example, strict security policies can have a negative effect on usability (which is a tradeoff that most customers are willing to make, but not all customers). You can plan to maximize usability by deploying user-friendly, host-naming	[redacted]

schemes and easy-to-use configuration methods that make use of dynamic protocols, such as the Dynamic Host Configuration Protocol (DHCP).	[redacted]
Usability might also include a need for mobility. As mentioned in Chapter 1, users expect to get their jobs done regardless of their physical location. They expect to have network access in conference rooms, at home, at a customer's site, and so on. Documenting this requirement as part of the technical requirements will help you recognize the need to select wireless and VPN solutions during the logical and physical design phases of the network design project. It will also help you recognize the need to conduct a site survey to prepare for a wireless infrastructure, as discussed in greater detail in the "Checking Architectural and Environmental Constraints" section of Chapter 3.	[redacted]
Adaptability	[redacted]

When designing a network, you should try to avoid incorporating any elements that would make it hard to implement new technologies in the future. A good network design can adapt to new technologies and changes. Changes can come in the form of new protocols, new business practices, new fiscal goals, new legislation, and a myriad of other possibilities. For example, some states have enacted environmental laws that require a reduction in the number of employees driving to work. To meet the legal requirement to reduce automobile emissions, companies need their remote-access designs to be flexible enough to adapt to increasing numbers of employees working at home. The adaptability of a network affects its availability. For example, some networks must operate in environments that change drastically from day to night or from winter to summer. Extreme changes in temperature can affect the behavior of electronic components of a network. A network that cannot adapt cannot offer good availability.	

A flexible network design can also adapt to changing traffic patterns and QoS requirements. For some customers, the selected WAN or LAN technology must adapt to new users randomly joining the network to use applications that require a constant-bit-rate service. Chapter 4 discusses QoS requirements in more detail.

One other aspect of adaptability is how quickly internetworking devices must adapt to problems and to upgrades. For example, how quickly do switches and bridges adapt to another switch failing, causing a change in the spanning-tree topology? How quickly do routers adapt to new networks joining the topology? How quickly do routing protocols adapt to link failures? Chapter 7, "Selecting Switching and Routing Protocols," discusses these issues in more detail.

Affordability

The final technical goal this chapter covers is affordability which is sometimes called cost-effectiveness. Most customers have a goal for affordability, although sometimes other goals such as performance and availability are more important. Affordability is partly a business goal and was discussed in Chapter 1. It is covered again in this chapter because of the technical issues involved.

For a network design to be affordable, it should carry the maximum amount of traffic for a given financial cost. Financial costs include nonrecurring equipment costs and recurring network operation costs. As mentioned in Chapter 1, you should learn about your customer's budget so that you can recommend solutions that are affordable.

In campus networks, low cost is often a primary goal. Customers expect to be able to purchase affordable

switches that have numerous ports and a low cost per port. They expect cabling costs to be minimal and service provider charges to be minimal or nonexistent. They also expect NICs for end systems and servers to be inexpensive. Depending on the applications running on end systems, low cost is often more important than availability and performance in campus network designs. For enterprise networks, availability is usually more important than low cost. Nonetheless, customers are looking for ways to contain costs for enterprise networks. Recurring monthly charges for WAN circuits are the most expensive aspect of running a large network.

To reduce the cost of operating a WAN, customers often have one or more of the following technical goals to achieve affordability:

- Use a routing protocol

that minimizes WAN traffic. ■ Consolidate parallel leased lines carrying voice and data into fewer WAN trunks. ■ Select technologies that dynamically allocate WAN bandwidth—for example, ATM rather than time-division multiplexing (TDM). ■ Improve efficiency on WAN circuits by using such features as compression. ■ Eliminate underutilized trunks from the internetwork and save money by eliminating both circuit costs and trunk hardware. ■ Use technologies that support oversubscription. With old-style TDM networks, the core backbone capacity had to be at least the sum of the speeds of the incoming access networks. With cell and frame switching, oversubscription is common. Because of the bursty nature of frame-based traffic, access port speeds	[redacted]

can add up to more than the speed of a backbone network, within reason. Enterprise network managers who have a goal of reducing operational costs are especially interested in solutions that will let them oversubscribe their trunks, while still maintaining service guarantees they have offered their users.

The second most expensive aspect of running a network, following the cost of WAN circuits, is the cost of hiring, training, and maintaining personnel to operate and manage the network. To reduce this aspect of operational costs, customers may require you to do the following as you develop the network design:

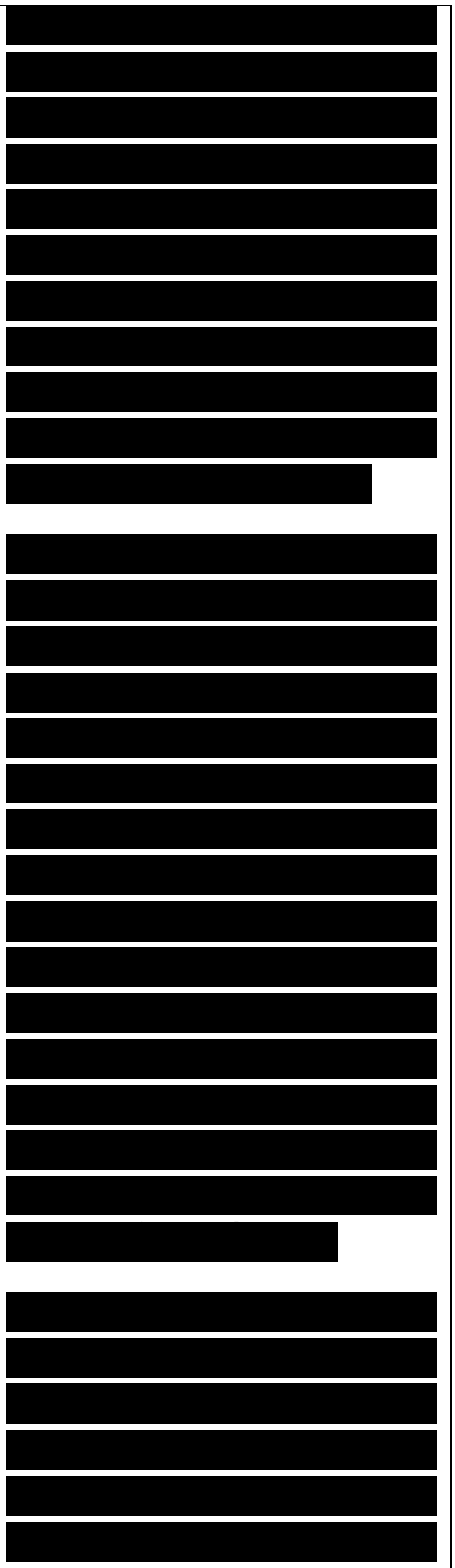
- Select internetworking equipment that is easy to configure, operate, maintain, and manage.
- Select a network design that is easy to understand and troubleshoot.
- Develop good network documentation that can help reduce troubleshooting time.
- Select network applications and protocols

that are easy to use so that users can support themselves to some extent.

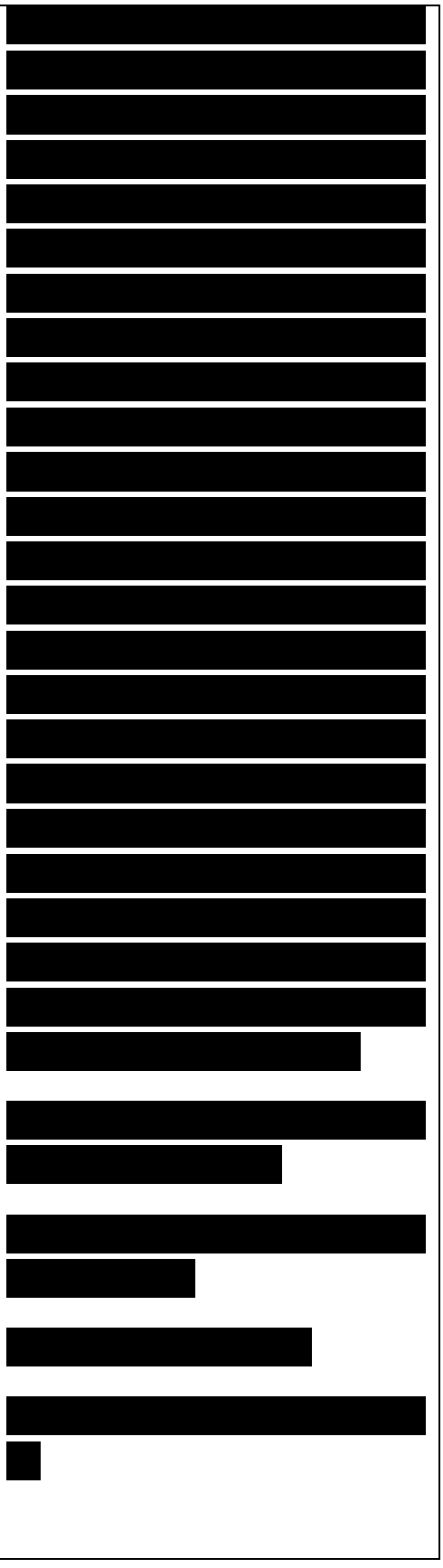
Making Network Design Tradeoffs

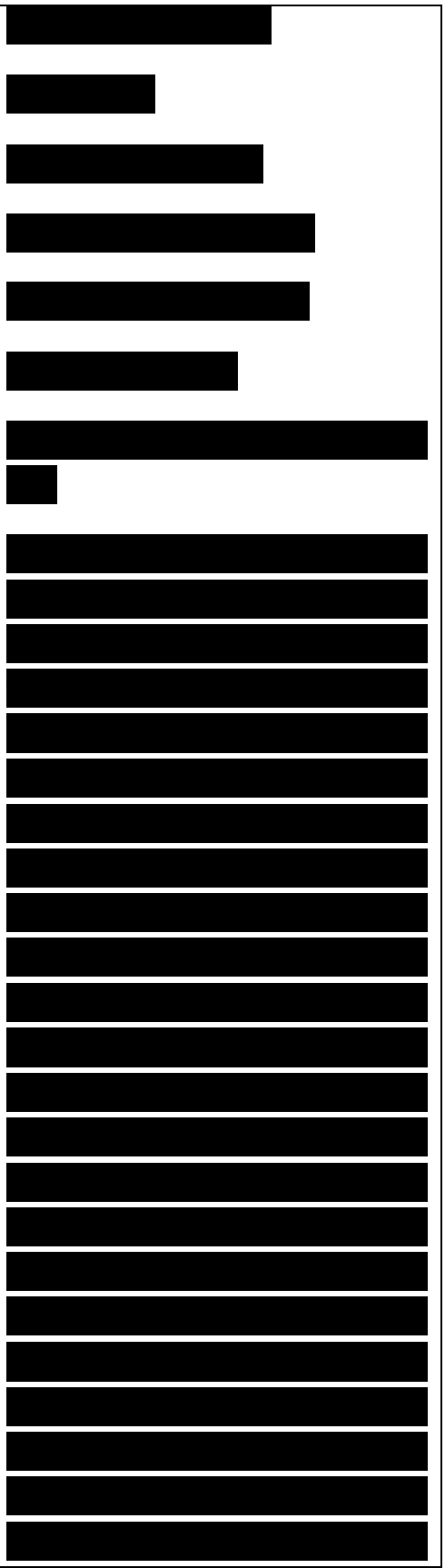
Despite what politicians tell us about state and federal budgets during an election year, in the real world meeting goals requires making tradeoffs. This section describes some typical network design tradeoffs.

To meet high expectations for availability, redundant components are often necessary, which raises the cost of a network implementation. To meet rigorous performance requirements, high-cost circuits and equipment are required. To enforce strict security policies, expensive monitoring might be required and users must forgo some ease of use. To implement a scalable network, availability might suffer, because a scalable network is always in flux as new users and sites are added. Implementing good throughput for one application might cause delay problems for another application. Lack of

qualified personnel might suggest the need for expensive training or the need to drop certain features. The network design that you develop must take these tradeoffs into consideration. One cause of network problems can be inadequate staffing and reduced training due to overzealous cost cutting. The tradeoff with cutting costs might be a network that isn't robust or has substandard performance until the problem is recognized, which often takes a year or two. If the in-house network staff was cut, outsourcing might become a necessity, which could end up being more costly than it would have been to keep the inhouse staff. The network design process is usually progressive. This means that legacy equipment must coexist with new equipment. Your design might not be as elegant as you would like because you might need for it to support	

old devices and old applications. If the new network is not being introduced at the same time as new applications, the design must provide compatibility with old applications. Also, be aware that insufficient bandwidth in parts of the network, where the bandwidth cannot be increased due to technical or business constraints, must be resolved by other means.	[redacted]
To help you analyze tradeoffs, ask your customer to identify a single driving network design goal. This goal can be the same overall business goal for the network design project that was identified in Chapter 1, or it can be a rephrasing of that goal to include technical issues. In addition, ask your customer to prioritize the rest of the goals. Prioritizing will help the customer get through the process of making tradeoffs.	[redacted]

One analogy that helps with prioritizing goals is the "kid in the candy store with a dollar bill" analogy. Using the dollar bill analogy, explain to the customers that they are like children in a candy store who have exactly one dollar to spend. The dollar can be spent on different types of candy: chocolates, licorice, jelly beans, and so on. But each time more money is spent on one type of candy, less money is available to spend on other types. Ask customers to add up how much they want to spend on scalability, availability, network performance, security, manageability, usability, adaptability, and affordability. For example, a customer could make the following selections: Scalability: 20 Availability: 30 Network performance: 15 Security: 5 Manageability: 5 Usability: 5	

Adaptability: 5 Affordability: 15 Total (must add up to 100): 100 Keep in mind that, sometimes, making tradeoffs is more complex than what has been described because goals can differ for various parts of an internetwork. One group of users might value availability more than affordability. Another group might deploy state- of-the-art applications and value performance more than availability. In addition, sometimes a particular group's goals are different from the overall goals for the internetwork as a whole. If this is the case, document individual group goals and goals for the network as a whole. Later, when selecting network technologies, you might see some opportunities to meet both types of goals—for example, choosing LAN technologies	

that meet individual group goals and WAN technologies that meet overall goals.	[redacted]
Technical Goals Checklist	[redacted]
You can use the following checklist to determine if you have addressed all your client's technical objectives and concerns:	[redacted]
□ I have documented the customer's plans for expanding the number of sites, users, and servers for the next 1 year and the next 2 years.	[redacted]
□ The customer has told me about any plans to migrate departmental servers to a centralized data center.	[redacted]
□ The customer has told me about any plans to integrate data stored on a legacy mainframe with the enterprise network.	[redacted]

□ The customer has told me about any plans to implement an extranet to communicate with partners or other companies.	
□ I have documented a goal for network availability in percent uptime and/or MTBF and MTTR.	
□ I have documented any goals for maximum average network utilization.	
□ I have documented goals for network throughput.	
□ I have documented goals for pps throughput of internetworking devices.	
□ I have documented goals for accuracy and acceptable BERs.	
□ I have discussed with the customer the importance of using large frame sizes to maximize efficiency.	
□ I have discussed with the customer the tradeoffs associated with large frame sizes and serialization delay.	

□ I have identified any applications that have a more restrictive response-time requirement than the industry standard of less than 100 ms. □ I have discussed network security risks and requirements with the customer. □ I have gathered manageability requirements, including goals for performance, fault, configuration, security, and accounting management. □ I have updated the Network Applications chart to include the technical application goals shown in Table 2-2.Table 2-2 Network Applications Technical Requirements □ Working with my customer, I have developed a list of network design goals, including both business and technical goals. The list starts with one overall goal and includes the rest of the goals in priority order. Critical goals are marked as such. Chapter 1 provided a	[redacted]

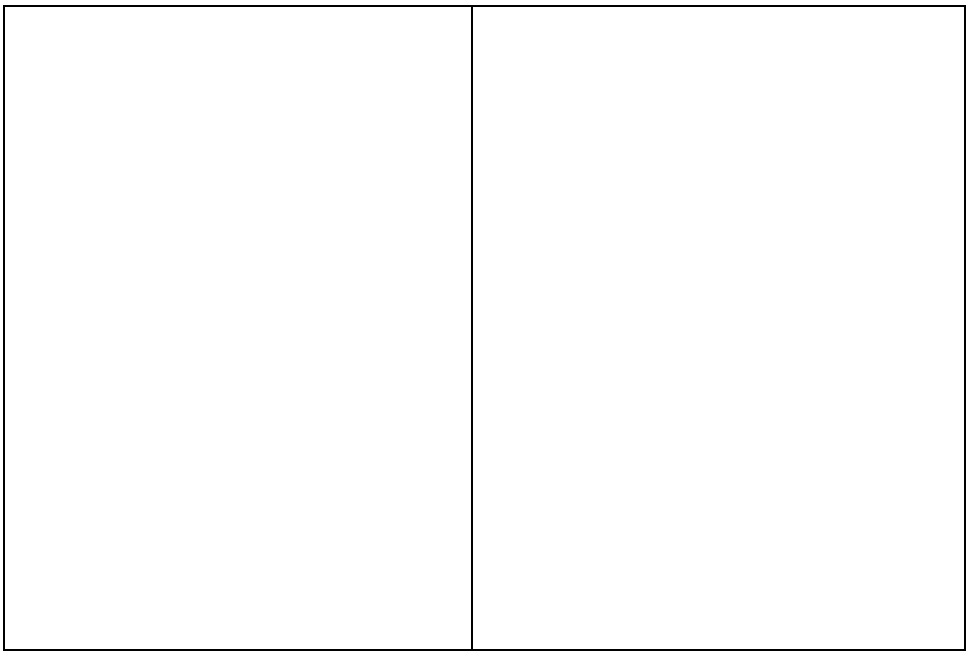
Network Applications chart. At this point in the design process, you can expand the chart to include technical application requirements, such as MTBF, MTTR, and throughput and delay goals, as shown in Table 2-2.

Summary

This chapter covered technical requirements for a network design, including scalability, availability, network performance, security, manageability, usability, adaptability, and affordability. It also covered typical tradeoffs that must be made to meet these goals.

Analyzing your customer's technical and business goals prepares you to carry out the next steps in the top-down network design process, including making decisions about network technologies to recommend to a customer. Researchers who study decision models say that one of the most important aspects of making a sound decision is having a good list of goals.

At this point in the network design process, you have gathered both business and technical goals. You should make a list of your customer's most important technical goals and merge this list with the list of business goals you made in Chapter 1. You should put the goals in the list in priority order, starting with the overall most important business and technical goal, and following with critical goals and then less critical goals. Later, you can make a list of options and correlate options with goals. Any options that do not meet critical goals can be eliminated. Other options can be ranked by how well they meet a goal. This process can help you select network components that meet a customer's requirements.	[redacted]



Chapter 3 check 2/1 (hôm nay)

Characterizing the Existing Internetwork

According to Abraham Lincoln:
"If we could first know where we are and whither we are tending, we could better judge what to do and how to do it."

An important step in top-down network design is to examine a customer's existing network to better judge how to meet expectations for network scalability, performance, and availability. Examining the existing network includes learning about the topology and physical structure and assessing the network's performance.

By developing an understanding of the existing network's structure, uses, and behavior, you can determine whether a customer's design goals are realistic. You can document any bottlenecks or network performance problems, and identify internetworking devices and links that will need to be replaced because the number of ports or capacity is insufficient for the new design. Identifying performance problems can help you select solutions to solve problems and develop a baseline for future measurements of performance.

Most network designers do not design networks from scratch. Instead, they design enhancements to existing networks. Developing a successful network design requires that you develop skills in characterizing an incumbent network to ensure interoperability between the existing and anticipated networks. This chapter describes techniques and tools to help you develop those skills. This chapter concludes with a Network Health checklist that documents typical thresholds for diagnosing a network as "healthy."

Characterizing the Network Infrastructure

Characterizing the infrastructure of a network means developing a set of network maps and learning the location of major internetworking devices and network segments. It also includes documenting the names and addresses of major devices and segments, and identifying any standard methods for addressing and naming. Documenting the types and lengths of physical cabling and investigating architectural and environmental constraints are also important aspects of characterizing the network infrastructure. Architectural and environmental constraints are becoming increasingly important in modern network designs that must accommodate wireless networking, which may not work if the signal is

blocked by cement walls, for example.

[redacted]

Developing a Network Map

Learning the location of major hosts, interconnection devices, and network segments is a good way to start developing an understanding of traffic flow. Coupled with data on the performance characteristics of network segments, location information gives you insight into where users are concentrated and the level of traffic that a network design must support.

[redacted]

At this point in the network design process, your goal is to obtain a map (or set of maps) of the existing network. Some design customers might have maps for the new network design as well. If that is the case, you might be one step ahead, but be careful of any assumptions that are not based on your detailed analysis of business and technical requirements.

[redacted]

To develop a network drawing, you should invest in a good network-diagramming tool. Tools include IBM's Tivoli products, WhatsUp Gold from Ipswitch, and LANsurveyor from SolarWinds. The Microsoft Visio Professional product is also highly recommended for network diagramming. For large enterprises and service providers, Visionael Corporation offers client/server network documentation products.

[redacted]

Note Tools that automatically diagram a network can be helpful, but the generated maps might require a lot of cleanup to make them useful.

Characterizing Large Internetworks

Developing a single network map might not be possible for large internetworks. There are many approaches to solving this problem, including simply developing many maps, one for each location. Another approach is to apply a top-down method. Start with a map or set of maps that shows the following high-level information:

- Geographical information, such as countries, states or provinces, cities, and campuses
- WAN connections between countries, states, and cities
- WAN and LAN connections between buildings and between campuses

For each campus network, you can develop more precise maps that show the following more detailed information:

- Buildings and floors, and possibly rooms or cubicles
- The location of major servers or server farms
- The location of routers and

switches ■ The location of firewalls, Network Address Translation (NAT) devices, intrusion detection systems (IDS), and intrusion prevention systems (IPS) ■ The location of mainframes ■ The location of major network-management stations ■ The location and reach of virtual LANs (VLAN) ■ Some indication of where workstations reside, although not necessarily the explicit location of each workstation Another method for characterizing large, complex networks is to use a top-down approach that is influenced by the OSI reference model. First, develop a logical map that shows applications and services used by network users. This map can call out internal web, email, FTP, and print and file-sharing servers. It can also include external web, email, and FTP servers. Note Be sure to show web caching servers on your network maps because they can affect traffic flow. Documenting the location of web caching servers will make it easier to troubleshoot any problems reaching web servers during the implementation and operation phases of the network design	[redacted]

cycle.	[redacted]
Next develop a map that shows network services. This map might depict the location of security servers; for example, Terminal Access Controller Access Control System (TACACS) and Remote Authentication Dial-In User Service (RADIUS) servers. Other network services include Dynamic Host Configuration Protocol (DHCP), Domain Name System (DNS), and Simple Network Management Protocol (SNMP) and other manage-ment services. The location and reach of any virtual private networks (VPN) that connect corporate sites via a service provider's WAN or the Internet can be depicted, including major VPN devices, such as VPN concentrators. Dial-in and dial-out servers can be shown on this map as well.	[redacted]
You may also want to develop a map that depicts the Layer 3 topology of the internetwork. This map can leave out switches and hubs, but should depict routers, logical links between the routers, and high-level routing protocol configuration information (for example, the location of the desired designated router [DR] if Open Shortest Path First [OSPF] is being used).	[redacted]

Layer 3 drawings should also include router interface names in Cisco shorthand nomenclature (such as s0/0) if Cisco routers are used. Other useful information includes Hot Standby Router Protocol (HSRP) router groupings, redistribution points between routing protocols, and demarcation points where route filters occur. The Layer 3 drawing should also include the location and high-level configuration of firewalls and NAT, IDS, and IPS devices.	[redacted]
A map or set of maps that shows detailed information about data link layer links and devices is often extremely helpful. This map reveals LAN devices and interfaces connected to public or private WANs. This map may hide the logical Layer 3 routing topology, which is shown in the previous map(s), but it should provide a good characterization of the physical topology. A data link layer map includes the following information: ■ An indication of the data link layer technology for WANs and LANs (Frame Relay, Point-to-Point Protocol [PPP], VPN, 100-Mbps or 1000-Mbps Ethernet, and so on) ■ The name of the service provider for WANs	[redacted]

■ WAN circuit IDs ■ The location and high-level configuration information for LAN switches (for example, the location of the desired root bridge if the Spanning Tree Protocol [STP] is used)	[redacted]
■ The location and reach of any VLANs and VLAN Trunking Protocol (VTP) configurations	[redacted]
■ The location and high-level configuration of trunks between LAN switches ■ The location and high-level configuration of any Layer 2 firewalls	[redacted]
Characterizing the Logical Architecture	[redacted]
While documenting the network infrastructure, take a step back from the diagrams you develop and try to characterize the logical topology of the network and the physical components. The logical topology illustrates the architecture of the network, which can be hierarchical or flat, structured or unstructured, layered or not, and other possibilities. The logical topology also describes methods for connecting devices in a geometric shape (for example, a star, ring, bus, hub and spoke, or mesh).	[redacted]
When characterizing the logical	[redacted]

topology, look for “ticking time bombs” or implementations that might hinder scalability. Ticking time bombs include large Layer 2 STP domains that will take a long time to converge and overly complex or oversized networks that might lead to Enhanced Interior Gateway Routing Protocol (EIGRP) stuck-in-active (SIA) problems and other routing problems. If the customer has fully redundant network equip-ment and cabling but the servers are all single-homed (attached to a single switch), keep this in mind as you plan your redesign of the network. This could be another ticking time bomb that can be fixed with a redesign. The logical topology can affect your ability to upgrade a network. For example, a flat topology does not scale as well as a hierarchical topology. A typical hierarchical topology that does scale is a core layer of high-end routers and switches that are optimized for availability and performance, a distribution layer of routers and switches that implement policies, and an access layer that connects users via hubs, switches, and other devices. Logical topologies are discussed in more detail in Chapter 5, “Designing a Network Topology.”	[redacted]

Figure 3-1 shows a high-level network diagram for an electronics manufacturing company. The drawing shows a physical topology, but it is not hard to step back and visualize that the logical topology is a hub-and-spoke shape with three layers. The core layer of the network is a Gigabit Ethernet network. The distribution layer includes routers and switches, and Frame Relay and T1 links. The access layer is composed of 10-Mbps and 100-Mbps Ethernet networks. An Ethernet network hosts the company's web server. As you can see from the figure, the network included some rather old design components. The company required design consultation to select new technologies and to meet new goals for high availability and security.

Frame Relay CIR = 56 Kbps
Frame Relay CIR = 56 Kbps DLCI = 4

Developing a Modular Block Diagram

In addition to developing a set of detailed maps, it is often helpful to draw a simplified block diagram of the network or parts of the network. The diagram can depict the major functions of the network in a modular fashion. Figure 3-2 shows a block, modularized network topology map that is based on the Cisco Enterprise Composite Network Model.

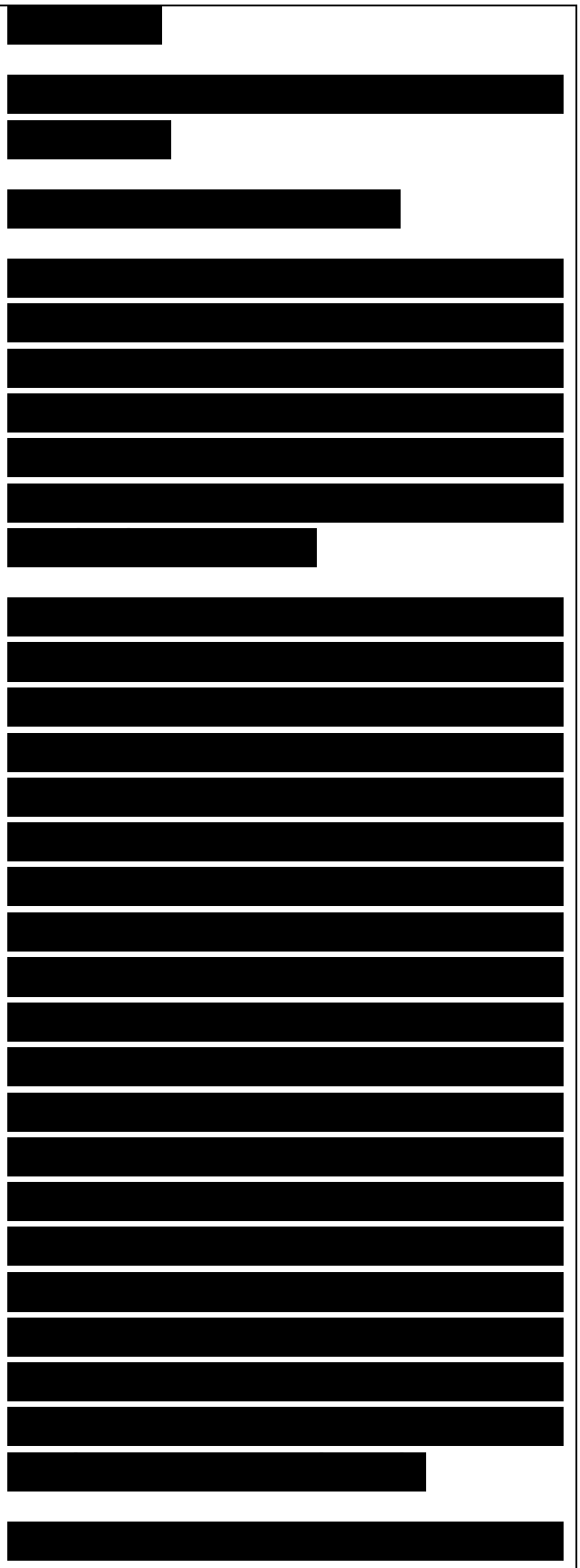


Figure 3-2 Modularized Network Topology Example

Characterizing Network Addressing and Naming

Characterizing the logical infrastructure of a network involves documenting any strategies your customer has for network addressing and naming. Addressing and naming are discussed in greater detail in Part II of this book, "Logical Network Design."

When drawing detailed network maps, include the names of major sites, routers, network segments, and servers. Also document any standard strategies your customer uses for naming network elements. For example, some customers name sites using airport codes (San Francisco = SFO, Oakland = OAK, and so on). You might find that a customer suffixes names with an alias that describes the type of device (for example, RTR for router). Some customers use a standard naming system, such as DNS, for IP networks, or NetBIOS Windows Internet Naming Service (WINS) on Windows networks. In such cases, you should document the location of the DNS and WINS servers and relevant high-level configuration information.

You should also investigate the network

layer addresses your customer uses. Your customer's addressing scheme (or lack of any scheme) can influence your ability to adapt the network to new design goals. For example, your customer might use unregistered IP addresses that will need to be changed or translated before connecting to the Internet. As another example, current IP subnet masking might limit the number of nodes in a LAN or VLAN. Your customer might have a goal of using route summarization, which is also called route aggregation or supernetting. Route summarization reduces routes in a routing table, routing-table update traffic, and overall router overhead. Route summarization also improves network stability and availability, because problems in one part of a network are less likely to affect the whole internetwork. Summarization is most effective when address prefixes have been assigned in a consistent and contiguous manner, which is often not the case. Your customer's existing addressing scheme might affect the routing protocols you can select. Some routing protocols do not support classless addressing, variable-length subnet masking (VLSM), or discontiguous subnets. A discontiguous subnet is a	[redacted]

subnet that is divided, as shown in Figure 3-3. Subnet 108 of network 10 is divided into two areas that are

Figure 3-3 Example of a Discontiguous Subnet

Characterizing Wiring and Media

To help you meet scalability and availability goals for your new network design, it is important to understand the cabling design and wiring of the existing network. Documenting the existing cabling design can help you plan for enhancements and identify any potential problems. If possible, you should document the types of cabling in use as well as cable distances. Distance information is useful when selecting data link layer technologies based on distance restrictions.

While exploring the cabling design, assess how well equipment and cables are labeled in the current network. The extent and accuracy of labeling will affect your ability to implement and test enhancements to the network.

Your network diagram should document the connections between buildings. The diagram should include information on

the number of pairs of wires and the type of wiring (or wireless technology) in use. The diagram should also indicate how far buildings are from one another. Distance information can help you select new cabling. For example, if you plan to upgrade from copper to fiber cabling, the distance between buildings can be much longer.

Probably the wiring (or wireless technology) between buildings is one of the following:

- Single-mode fiber
- Multimode fiber
- Shielded twisted-pair (STP) copper
- Unshielded twisted-pair (UTP) copper
- Coaxial cable
- Microwave
- Laser
- Radio
- Infrared

Within buildings, try to locate telecommunications wiring closets, cross-connect rooms, and any laboratories or computer rooms. If possible, determine the type of cabling that is installed between telecommunications closets and in work areas. (Some technologies, such as 100BASE-TX Ethernet, require

Category 5 or later cabling, so be sure to document the existence of any Category 3 cabling that needs to be replaced.) Gather information about both vertical and horizontal wiring. As shown in Figure 3-4, vertical wiring runs between floors. Horizontal wiring runs from telecommunications closets to wallplates in cubicles or offices. Work-area wiring runs from the wallplate to a worksta-tion in a cubicle or office.

Campus Backbone

Figure 3-4 Example of Campus Network Wiring

In most buildings, the cabling from a telecommunications closet to a workstation is approximately 100 meters (about 300 feet), including the work-area wiring, which is usually just a few meters. If you have any indication that the cabling might be longer than 100 meters, you should use a time-domain reflectometer (TDR) to verify your suspicions. (TDR functionality is included in most cable testers.) Many network designs are based on the assumption that workstations are no more than 100 meters from the telecommunications closet.

Ask the client for a copy of the copper or fiber certification tests that were completed when the cabling was first installed. Test results will help you learn the type of cabling that was installed, its

certification, and the warranty period for the installation work. Many modern network designers do just that one step of verifying that the cable was tested and certified rather than going through a detailed analysis of the cabling infrastructure. On the other hand, many network designers still focus on cabling because they have learned the hard way that meeting availability goals can be difficult when the cabling was not installed properly.

For each building, you can fill out the chart shown in Table 3-1. The data that you fill in depends on how much time you have to gather information and how important you think cabling details will be to your network design. If you do not have a lot of information, just put an X for each type of cabling present and document any assumptions (for example, an assumption that workstations are no more than 100 meters from the telecommunications closet). If you have time to gather more details, include information on the length and number of pairs of cables. If you prefer, you can document building wiring information in a network diagram instead of in a table.

Table 3-1 Building Wiring

Building name

Location of telecommunications closets Location of cross-connect rooms and demarcations to external networks	[redacted]
Logical wiring topology (structured, star, bus, ring, centralized, distributed, mesh, tree, or whatever fits)	[redacted]
Vertical Wiring	[redacted]
Coaxial Fiber STP Category 3 UTP Category 5 or 6 UTP Other	[redacted]
Vertical shaft 1 Vertical shaft 2 Vertical shaft n Horizontal Wiring	[redacted]
Coaxial Fiber STP Category 3 UTP Category 5 or 6 UTP Other	[redacted]
Floor 1 Floor 2 Floor 3 Floor n Work-Area Wiring Coaxial Fiber STP Category 3 UTP Category 5 or 6 UTP Other	[redacted]
Floor 1 Floor 2 Floor 3 Floor n Checking Architectural and Environmental Constraints	[redacted]
When investigating cabling, pay attention to such environmental issues as the possibility that cabling will run near creeks that could flood, railroad tracks or highways where traffic could jostle cables, or construction or	[redacted]

manufacturing areas where heavy equipment or digging could break cables.

Be sure to determine if there are any legal right-of-way issues that must be dealt with before cabling can be put into place. For example, will cabling need to cross a public street? Will it be necessary to run cables through property owned by other companies? For line-of-sight technologies, such as laser or infrared, make sure there aren't any obstacles blocking the line of sight.

Within buildings, pay attention to architectural issues that could affect the feasibility of implementing your network design. Make sure the following architectural elements are sufficient to support your design:

- Air conditioning
- Heating
- Ventilation
- Power
- Protection from electromagnetic interference
- Doors that can lock
- Space for:
- Cabling conduits
- Patch panels
- Equipment racks
- Work areas for technicians installing and troubleshooting equipment

Note Keep in mind that cabling and power are highly influenced by human factors. Installing new cabling might require working with labor unions, for example. Maintaining the reliability of cabling might require monitoring the infamous backhoe operator or the janitor who knocks cables around. It's also not unheard of for security guards to lean against a wall late at night and accidentally activate emergency power off (EPO) or discharge fire suppressant. To avoid problems, make sure EPO and fire suppressant buttons have safety covers and are out of the way.

Checking a Site for a Wireless Installation

A common goal for modern campus network designs is to install a wireless LAN (WLAN) based on IEEE 802.11 standards. An important aspect of inspecting the architectural and environmental constraints of a site is determining the feasibility of using wireless transmission. The term wireless site survey is often used to describe the process of analyzing a site to see if it will be appropriate for wireless

transmission.

In some ways, doing a wireless site survey is no different from checking an architecture for wired capabilities, where you might need to document obstructions or areas that have water leaks, for example. But in many ways, a wireless site survey is quite different from a wired site survey because the transmission isn't going through guided wires; it's being sent in radio frequency (RF) waves through air. Learning RF transmission theory in depth requires a lot of time and a good background in physics. For complex RF designs and concerns, it often makes sense to hire an RF expert. To do a basic site survey, you might not need help, though.

A site survey starts with a draft WLAN design. Using a floor plan or blueprint for the site, the designer decides on the initial placement of the wireless access points. An access point is a station that transmits and receives data for users of the WLAN. It usually serves also as the point of interconnection between the WLAN and the wired Ethernet network. A network designer can decide where to place access points for initial testing based on some knowledge of where the users will be located, characteristics of

the access points' antennas, and the location of major obstructions.	[redacted]
The initial placement of an access point is based on an estimate of the signal loss that will occur between the access point and the users of the access point. The starting point for an estimate depends on how much loss in power a signal would experience in the vacuum of space, without any obstructions or other interference. This is called the free space path loss and is specified in decibels (dB). The estimate is tuned with an understanding that the actual expected signal loss depends on the medium through which the signal will travel, which is undoubtedly not a vacuum. An RF signal traveling through objects of various sorts can be affected by many different problems, including the following:	[redacted]
■ Reflection: Reflection causes the signal to bounce back on itself. The signal can interfere with itself in the air and affect the receiver's capability to discriminate between the signal and noise in the environment. Reflection is caused by metal surfaces such as steel girders, scaffolding, shelving units, steel pillars, and metal doors. As an example, implementing a WLAN across a parking lot can be tricky because of metal cars (sources of reflection) that come and go.	[redacted]

■ Absorption: Some of the electromagnetic energy of the signal can be absorbed by the material in objects through which it passes, resulting in a reduced signal level. Water has significant absorption properties, and objects such as trees or thick wooden structures can have a high water content. Implementing a WLAN in a coffee shop can be tricky if there are large canisters of liquid coffee. Coffee shop WLAN users have also noticed that people coming and going can affect the signal level. (On Star Trek, a nonhuman character once called a human "an ugly giant bag of mostly water"!)

■ Refraction: When an RF signal passes from a medium with one density into a medium with another density, the signal can be bent, much like light passing through a prism. The signal changes direction and might interfere with the nonrefracted signal. It can take a different path and encounter other, unexpected obstructions and arrive at recipients damaged or later than expected. As an example, a water tank not only introduces absorption, but also the difference in density between the atmosphere and the water can bend the RF signal.

■ Diffraction: Diffraction, which is

similar to refraction, results when a region through which the RF signal can pass easily is adjacent to a region in which reflective obstructions exist. Like refraction, the RF signal is bent around the edge of the diffractive region and can then interfere with that part of the RF signal that is not bent.

The designers of 802.11 transmitting devices attempt to compensate for variable environmental factors that might cause reflection, absorption, refraction, or diffraction by boosting the power level above what would be required if free space path were the only consideration. The additional power added to a transmission is called the fade margin.

Performing a Wireless Site Survey

A site survey confirms signal propagation, strength, and accuracy in different locations. Many wireless network interface cards (NIC) ship with utilities that enable you to measure signal strength. Cisco 802.11 NICs ship with the Cisco Aironet Client Utility (ACU), which is a graphical tool for configuring, monitoring, and managing the NIC and its wireless environment. A site survey can be as simple as walking around with a wireless notebook computer and using the utility to measure signal strength.

Signal strength can also be determined with a protocol analyzer. The WildPackets AiroPeek analyzer, for example, presents the signal strength for each frame received.	[redacted]
An access point typically sends a beacon frame every 100 milliseconds (ms). You can divide the area being surveyed into a grid, and then move your protocol analyzer from gridpoint to gridpoint and plot on a diagram the signal strength of the beacon frames.	[redacted]
When evaluating the various metrics that are provided by wireless utilities, be sure to measure frame corruption and not just signal strength. With a protocol analyzer, capture frames and check for cyclic redundancy check (CRC) errors. CRC errors are the result of corruption from environmental noise or collisions between frames.	[redacted]
You can also indirectly measure signal quality by determining if frames are being lost in transmission. If your protocol analyzer is capturing relatively close to an access point and a mobile client is pinging a server, through the access point, onto the wired Ethernet, you can determine whether ping packets are getting lost.	[redacted]

As part of your site survey, you can also look at acknowledgments (ACK) and frame retries after a missing ACK. With 802.11 WLANs, both the client and the access point send ACKs to each other. An ACK frame is one of six special frames called control frames. All directed traffic (frames addressed to any nonbroadcast, nonmulticast destination) are positively acknowledged with an ACK. Clients and access points use ACKs to implement a retransmission mechanism not unlike the Ethernet retry process that occurs after a collision.	[redacted]
In a wired Ethernet, the transmitting station detects collisions through the rules of carrier sense multiple access with collision detection (CSMA/CD). 802.11 uses carrier sense multiple access with collision avoidance (CSMA/CA) as the access method and does not depend on collision detection to operate. Instead, an ACK control frame is returned to a sender for each directed packet received. If a directed frame does not receive an ACK, the frame is retransmitted.	[redacted]
Wireless networking is covered again in later chapters, but remember to consider it early in your design planning. Using a wireless utility, such as the Cisco ACU,	[redacted]

WildPackets OmniPeek, or NetStumbler, check signal strength and accuracy with potential access point placements to determine if the architecture of the physical site will be a problem. Performing a basic wireless site survey is an important part of the top-down network design process of checking for architectural and environmental constraints.	[redacted]
Checking the Health of the Existing Internetwork	[redacted]
Studying the performance of the existing internetwork gives you a baseline measurement from which to measure new network performance. Armed with measurements of the present internetwork, you can demonstrate to your customer how much better the new internetwork performs once your design is implemented.	[redacted]
Many of the network-performance goals discussed in Chapter 2, "Analyzing Technical Goals and Tradeoffs," are overall goals for an internetwork. Because the performance of existing network segments will affect overall performance, you need to study the performance of existing segments to determine how to meet overall network performance goals.	[redacted]

If an internetwork is too large to study all segments, you should analyze the segments that will interoperate the most with the new network design. Pay particular attention to backbone networks and networks that connect old and new areas.	[redacted]
In some cases, a customer's goals might be at odds with improving network performance. The customer might want to reduce costs, for example, and not worry about performance. In this case, you will be glad that you documented the original performance so that you can prove that the network was not optimized to start with and your new design has not made performance worse.	[redacted]
By analyzing existing networks, you can also recognize legacy systems that must be incorporated into the new design. Sometimes customers are not aware that older protocols are still running on their internetworks. By capturing network traffic with a protocol analyzer as part of your baseline analysis, you can identify which protocols are actually running on the network and not rely on customers' beliefs.	[redacted]

Developing a Baseline of Network Performance Developing an accurate baseline of a network's performance is not an easy task. One challenging aspect is selecting a time to do the analysis. It is important that you allocate a lot of time (multiple days) if you want the baseline to be accurate. If measurements are made over too short a timeframe, temporary errors appear more significant than they are. In addition to allocating sufficient time for a baseline analysis, it is also important to find a typical time period to do the analysis. A baseline of normal performance should not include atypical problems caused by exceptionally large traffic loads. For example, at some companies, end-of-the-quarter sales processing puts an abnormal load on the network. In a retail environment, network traffic can increase fivefold around Christmas time. Network traffic to a web server can unexpectedly increase tenfold if the website gets linked to other popular sites or listed in search engines. In general, errors, packet/cell loss, and	[redacted]

latency increase with load. To get a meaningful measurement of typical accuracy and delay, try to do your baseline analysis during periods of normal traffic load. On the other hand, if your customer's main goal is to improve performance during peak load, be sure to study performance during peak load. The decision whether to measure normal performance, performance during peak load, or both, depends on the goals of the network design.	[redacted]
Some customers do not recognize the value of studying the existing network before designing and implementing enhancements. Your customer's expectations for a speedy design proposal might make it difficult for you to take a step back and insist on time to develop a baseline of performance on the existing network. Also, your other job tasks and goals, especially if you are a sales engineer, might make it impractical to spend days developing a precise baseline.	[redacted]
The work you do before the baseline step in the top-down network design methodology can increase your efficiency in developing a baseline. A good understanding of your customer's technical and business goals can help you decide how thorough to make your study. Your discussions with your customer on business goals can help you	[redacted]

identify segments that are important to study because they carry critical and/or backbone traffic.

[redacted]

You can also ask your customer to help you identify typical segments from which you can extrapolate other segments.

[redacted]

Analyzing Network Availability

[redacted]

To document availability characteristics of the existing network, gather any statistics that the customer has on the mean time between failure (MTBF) and mean time to repair (MTTR) for the internetwork as a whole and major network segments. Compare these statistics with information you have gathered on MTBF and MTTR goals, as discussed in Chapter 2. Does the customer expect your new design to increase MTBF and decrease MTTR? Are the customer's goals realistic considering the current state of the network?

[redacted]

Talk to the network engineers and technicians about the root causes of the most recent and most disruptive periods of downtime. Assuming the role of a forensic investigator, try to get many

[redacted]

sides to the story. Sometimes myths develop about what caused a network outage. (You can usually get a more accurate view of problem causes from engineers and technicians than from users and managers.)

[redacted]

You can use Table 3-2 to document availability characteristics of the current network.

[redacted]

Table 3-2 Availability Characteristics of the Current Network

[redacted]

	MTBF	MTTR	Date and Duration of Last Major Downtime	Cause of Last Major Downtime	Fix for Last Downtime
Enterprise (as a whole)					
Segment 1					
Segment 2					
Segment 3					
Segment n					

[redacted]

Analyzing Network Utilization

Network utilization is a measurement of the amount of bandwidth that is in use during a specific time interval. Utilization is commonly specified as a percentage of capacity. If a network-monitoring tool says that network utilization on a Fast Ethernet segment is 70 percent, for example, this means that 70 percent of the 100-Mbps capacity is in use, averaged over a specified timeframe or window.

[redacted]

Different tools use different averaging windows for computing network utilization. Some tools let the user change the window. Using a long interval can be useful for reducing the amount of statistical data that must be analyzed, but granularity is sacrificed. As Figure 3-5 shows, it can be informative (though tedious) to look at a chart that shows network utilization averaged every minute.
16:40:00

Figure 3-6 shows the same data averaged over 1-hour intervals. Note that the network was not very busy, so neither chart goes above 7 percent utilization. Note also that changing to a long interval can be misleading because peaks in traffic get averaged out (the detail is lost). In Figure 3-5, you can see that the network was relatively busy around 4:50 p.m.

You cannot see this in Figure 3-6, when the data was averaged on an hourly basis.

In general, you should record network utilization with sufficient granularity in time to see short-term peaks in network traffic so that you can accurately assess the capacity requirements of devices and segments. Changing the interval to a

small amount of time, say a fraction of a second, can be misleading also, however. To understand the concern, consider a small time interval. In a packet-sized window, at a time when a station is sending traffic, the utilization is 100 percent, which is what is wanted.

The size of the averaging window for network utilization measurements depends on your goals. When troubleshooting network problems, keep the interval small, either minutes or seconds. A small interval helps you recognize peaks caused by problems such as broadcast storms or stations retransmitting quickly due to a misconfigured timer. For performance analysis and baselining purposes, use an interval of 1 to 5 minutes. For long-term load analysis, to determine peak hours, days, or months, set the interval to 10 minutes.

Utilization

Figure 3-6 Network Utilization in Hour Intervals

When developing a baseline, it is

usually a good idea to err on the side of gathering too much data. You can always summarize the data later. When characterizing network utilization, use protocol analyzers or other monitoring tools to measure utilization in 1- to 5-minute intervals on each major network segment. If practical, leave the monitoring tools running for at least 1 or 2 typical days. If the customer's goals include improving performance during peak times, measure utilization during peak times and typical times. To determine if the measured utilization is healthy, use the Network Health checklist that appears at the end of this chapter.

[redacted]

Measuring Bandwidth Utilization by Protocol

[redacted]

Developing a baseline of network performance should also include measuring utilization from broadcast traffic versus unicast traffic, and by each major protocol. As discussed in Chapter 4, "Characterizing Network Traffic," some protocols send excessive broadcast traffic, which can seriously degrade performance, especially on switched networks.

[redacted]

To measure bandwidth utilization by protocol, place a protocol analyzer or remote monitoring (RMON) probe on each major network segment and fill out

[redacted]

a chart such as the one shown in Table 3-3. If the analyzer supports relative and absolute percentages, specify the bandwidth used by protocols as relative and absolute. Relative usage specifies how much bandwidth is used by the protocol in comparison to the total bandwidth currently in use on the segment. Absolute usage specifies how much bandwidth is used by the protocol in comparison to the total capacity of the segment (for example, in comparison to 100 Mbps on Fast Ethernet).

Table 3-3 Bandwidth Utilization by Protocol

Protocol	Relative Network Utilization	Absolute Network Utilization	Broadcast/Multicast Rate
Protocol 1			
Protocol 2			
Protocol 3			
Protocol n			

Analyzing Network Accuracy

Chapter 2 talked about specifying network accuracy as a bit error rate (BER). You can use a BER tester (also called a BERT) on serial lines to test the number of damaged bits compared to total bits. As discussed in the "Checking the Status of Major Routers, Switches, and Firewalls" section later in this chapter, you can also use Cisco show commands to gain an understanding of errors on a serial interface, which is a more common practice on modern networks than using a BERT.

With packet-switched networks, it makes more sense to measure frame (packet) errors because a whole frame is considered bad if a single bit is changed or dropped. In packet- switched networks, a sending station calculates a CRC based on the bits in a frame. The sending station places the value of the CRC in the frame. A receiving station determines if a bit has been changed or dropped by calculating the CRC again and comparing the result to the CRC in the frame. A frame with a bad CRC is dropped and must be retransmitted by the sender. Usually an upper-layer protocol has the job of retransmitting frames that do not get acknowledged.

A protocol analyzer can check the CRC on received frames. As part of your baseline analysis, you should track the number of frames received with a bad CRC every hour for 1 or 2 days. Because it is normal for errors to increase with utilization, document errors as a function of the number of bytes seen by the monitoring tool. A good rule-of-thumb threshold for considering errors unhealthy is that a network should not have more than one bad frame per megabyte of data. (Calculating errors this way lets you simulate a serial BERT. Simply calculating a percentage of bad frames compared to good frames does not account for the size of frames and hence does not give a good indication of how many bits are actually getting damaged.)

In addition to tracking data link layer errors, such as CRC errors, a baseline analysis should include information on upper-layer problems. A protocol analyzer that includes an expert system, such as CACE Technologies' Wireshark analyzer or WildPackets' OmniPeek analyzer, speeds the identification of upper-layer problems by automatically generating diagnoses and symptoms for network conversations and applications.	[redacted]
Accuracy should also include a measurement of lost packets. You can measure lost packets while measuring response time, which is covered later in this chapter in the "Analyzing Delay and Response Time" section. When sending packets to measure how long it takes to receive a response, document any packets that do not receive a response, presumably because either the request or the response got lost. Correlate the information about lost packets with other performance measurements to determine if the lost packets indicate a need to increase bandwidth, decrease CRC errors, or upgrade internetworking devices. You can also measure lost packets by looking at statistics kept by routers on the number of packets dropped from input or output queues.	[redacted]

Analyzing Errors on Switched Ethernet Networks

Switches have replaced hubs in most campus networks. A switch port that is in halfduplex mode follows the normal rules of CSMA/CD. The port checks the medium for any traffic by watching the carrier sense signal, defers to traffic if necessary, detects collisions, backs off, and retransmits. Whether a collision can occur depends on what is connected to the switched port. If a shared medium is connected to the switch, collisions can occur. A good rule of thumb is that fewer than 0.1 percent of frames should encounter collisions. There should be no late collisions. Late collisions are collisions that happen after a port or interface has sent the first 64 bytes of a frame. Late collisions indicate bad cabling, cabling that is longer than the 100-meter standard, a bad NIC, or a duplex mismatch.

If the switch port connects a single device, such as another switch, a server, or a single workstation, both ends of this point-to-point link should be configured for full duplex. In this case, collisions should never occur. Full-duplex Ethernet isn’t CSMA/CD. There are

only two stations that can send because full duplex requires a point-to-point link, and each station has its own private transmit channel. So full duplex isn't multiple access (MA). There's no need for a station to check the medium to see if someone else is sending on its transmit channel. There isn't anyone else. So full duplex doesn't use carrier sense (CS). There are no collisions. Both stations sending at the same time is normal. Receiving while sending is normal. So, there is no collision detection (CD) either.

Unfortunately, the autonegotiation of half versus full duplex has been fraught with problems over the years, resulting in one end of a point-to-point link being set to half duplex and the other being set to full duplex. This is a misconfiguration and must be fixed. Autonegotiation problems can result from hardware incompatibilities and old or defective Ethernet software drivers. Some vendors' NICs or switches do not conform exactly to the IEEE 802.3u specification, which results in incompatibilities. Hardware incompatibility can also occur when vendors add advanced features, such as autopolarity, that are not in the IEEE 802.3u specification. (Autopolarity corrects reversed polarity on the

transmit and receive twisted pairs.)	[redacted]
The autonegotiation of speed isn't usually a problem. If the speed doesn't negotiate correctly, the interface doesn't work, and the administrator hopefully notices and corrects the problem immediately. Manually configuring the speed for 10 Mbps, 100 Mbps, or 1000 Mbps usually isn't necessary (except for cases where the user interface requires this before it will allow manual configuration of duplex mode). If a LAN still has Category 3 cabling, manually configuring the speed to 10 Mbps is recommended, however. Errors can increase on a LAN that has autonegotiated for 100 Mbps or 1000 Mbps if there is Category 3 cabling that does not support the high-frequency signal used on 100- or 1000-Mbps Ethernet.	[redacted]
Duplex negotiation happens after the speed is negotiated. Problems with duplex negotiation are harder to detect because any performance impact is dependent on the link partners transmitting at the same time. A workstation user who doesn't send much traffic might not notice a problem,	[redacted]

whereas a server could be severely impacted by a duplex mismatch. As part of analyzing the performance of the existing network, be sure to check for duplex mismatch problems. A surprisingly high number of networks have been hob-bling along for years with performance problems related to a duplex mismatch. To detect a duplex mismatch, look at the number and type of errors on either end of the link. You can view errors with the show interface or show port command on Cisco routers and switches. Look for CRC and runt errors on one side and collisions on the other side of the link. The side that is set for full duplex can send whenever it wants. It doesn't need to check for traffic. The side that is set for half duplex does check for traffic and will stop transmitting if it detects a simultaneous transmission from the other side. It will back off, retransmit, and report a collision. The result of the half-duplex station's stopping transmission is usually a runt frame (shorter than 64 bytes) and is always a CRC- errored frame. The full-duplex side receives runts and	[redacted]

CRC-errored frames and reports these errors. The half-duplex station reports collisions. Most of these will be legal collisions; some might be illegal late collisions. When checking the health of Ethernet LANs, check for these errors. Notice the asymmetry of the errors when there is a duplex mismatch. If you see collisions and CRC errors on both sides of the link, the problem is probably something other than a duplex mismatch, perhaps a wiring problem or bad NIC.	[redacted]
Until recently, most engineers recommended avoiding autonegotiation, but that is changing. Improvements in the interoperability of autonegotiation and the maturity of the technology mean that it is generally safer to rely on autonegotiation than to not rely on it.	[redacted]
There are numerous problems with not using autonegotiation. The most obvious one is human error. The network engineer sets one end of the link and forgets to set the other end. Another problem is that some NICs and switch ports don't participate in autonegotiation if manually set. This means they don't send the link pulses to report their setting.	[redacted]

How should the partner react to such a situation? The answer is undefined. Some NICs and switch ports assume the other side is too old to understand full duplex and must be using half. This causes the NIC or switch port to set itself to half. This is a serious problem if the other side is manually configured to full. On the other hand, there are cases where autonegotiation simply does not work, and you might need to carefully configure the mode manually.

Analyzing Network Efficiency

Chapter 2 talked about the importance of using maximum frame sizes to increase network efficiency. Bandwidth utilization is optimized for efficiency when applications and protocols are configured to send large amounts of data per frame, thus minimizing the number of frames and round-trip delays required for a transaction. The number of frames per transaction can also be minimized if the receiver is configured with a large receive window allowing it to accept multiple frames before it must send an acknowledgment. The goal is to maximize the number of data bytes compared to the number of bytes in headers and in acknowledgment packets sent by the other end of a conversation.

Changing frame and receive window sizes on clients and servers can result in improved efficiency. Increasing the maximum transmission unit (MTU) on router interfaces can also improve efficiency, although doing this is not appropriate on low-bandwidth links that are used for voice or other real-time traffic. (As Chapter 2 mentioned, you don't want to increase serialization delay.)	[redacted]
On the other hand, increasing the MTU is sometimes necessary on router interfaces that use tunnels. Problems can occur when the extra header added by the tunnel causes frames to be larger than the default MTU, especially in cases where an application sets the IP Don't Fragment (DF) bit and a firewall is blocking the Internet Control Message Protocol (ICMP) packets that notify the sender of the need to fragment. A typical symptom of this problem is that users can ping and telnet but not use HTTP, FTP, and other protocols that use large frames. A solution is to increase the MTU on the router interface.	[redacted]
To determine if your customer's goals for network efficiency are realistic, you should use a protocol analyzer to examine the current frame sizes on the	[redacted]

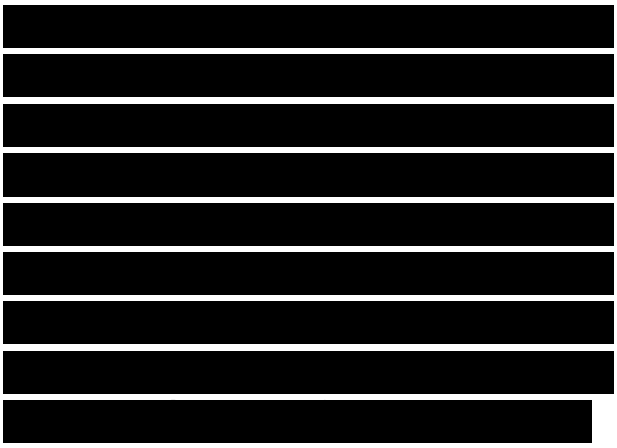
network. Many protocol analyzers let you output a chart, such as the one in Figure 3-7, that documents how many frames fall into standard categories for frame sizes. Figure 3-7 shows packet sizes at an Internet service provider (ISP). Many of the frames were 64-byte acknowledgments. A lot of the traffic was HTTP, which used 1500-byte packets in most cases, but also sent 500- and 600-byte packets. If many web-hosting customers had been transferring pages to a web server using a file-transfer or file-sharing protocol, there would have been many more 1500-byte frames. The other traffic consisted of DNS lookups and replies and Simple Mail Transfer Protocol (SMTP), Post Office Protocol (POP), and Address Resolution Protocol (ARP) packets.

A simple way to determine an average frame size is to divide the total number of megabytes seen on a segment by the total number of frames in a specified timeframe. Unfortunately, this is a case in which a simple statistical technique does not result in useful data. The average frame size is not a meaningful piece of information. On most networks, there are many small frames, many large frames, but few average-sized frames. Small frames consist of acknowledgments and control information. Data frames fall into the

large frame-size categories. Frame sizes typically fall into what is called a bimodal distribution, also known as a camel-back distribution. A "hump" is on either side of the average but not many values are near the average.

Figure 3-7 Graph of Packet Sizes on an Internet Service Provider's Ethernet Backbone

Note Network performance data is often bimodal, multimodal, or skewed from the mean. (Mean is another word for average.) Frame size is often bimodal. Response times from a server can also be bimodal, if sometimes the data is quickly available from RAM cache and sometimes the data is retrieved from a slow mechanical disk drive.



When network-performance data is bimodal, multimodal, or skewed from the mean, you should document a standard deviation with any measurements of the mean. Standard deviation is a measurement of how widely data disperses from the mean.

Analyzing frame sizes can help you understand the health of a network, not just the efficiency. For example, an excessive number of Ethernet runt frames (less than 64 bytes) can indicate too many collisions on a shared Ethernet segment. It is normal for collisions to

increase with utilization that results from access contention. If collisions increase even when utilization does not increase or even when only a few nodes are transmitting, there could be a more serious problem, such as a bad NIC or a duplex mismatch problem.

Analyzing Delay and Response Time

To verify that performance of a new network design meets a customer's requirements, you need to measure response time between significant network devices before and after a new network design is implemented. Response time can be measured many ways. Using a protocol analyzer, you can look at the amount of time between frames and get a rough estimate of response time at the data link layer, transport layer, and application layer.

(This is a rough estimate because packet arrival times on an analyzer can only approximate packet arrival times on end stations.)

A more common way to measure response time is to send ping packets and measure the round-trip time (RTT) to send a request and receive a response. While measuring RTT, you can also

measure an RTT variance. Variance measurements are important for applications that cannot tolerate much jitter (for example, voice and video applications). You can also document any loss of packets.

You can use Table 3-4 to document response time measurements. The table uses the term node to mean router, server, client, or mainframe.

Table 3-4 Response-Time Measurements

Depending on the amount of time you have for your analysis and depending on your customer's network design goals, you should also measure response time from a user's point of view. On a typical workstation, run some representative applications and measure how long it takes to get a response for typical operations, such as checking email, sending a file to a server, downloading a web page, updating a sales order, printing a report, and so on.

Sometimes applications or protocol implementations are notoriously slow or poorly written. Some peripherals are known to cause extra delay because of incompatibilities with operating systems or hardware. By joining mailing lists and newsgroups and reading information in journals and on the World Wide Web, you can learn about causes of responsetime problems. Be sure to do some testing on your own

also, though, because every environment is different.	[redacted]
In addition to testing user applications, test the response time for network-services protocols (for example, DNS queries, DHCP requests for an IP address, RADIUS authentication requests, and so on). Chapter 4 covers protocol issues in more detail.	[redacted]
You should also measure how much time a workstation takes to boot. Some workstation operating systems take a long time to boot due to the amount of network traffic that they send and receive while booting. You can include boot time measurements in your analysis of the existing network so that you have a baseline. When the new network design is implemented, you can compare the amount of time a workstation takes to boot with the baseline time. Hopefully you can use this data to prove that your design is an improvement.	[redacted]
Although your customer might not give you permission to simulate network problems, it makes sense to do some testing of response times when the network is experiencing problems or change. For example, if possible, measure response times while routing	[redacted]

protocols are converging after a link has gone down. Measure response times during convergence again, after your new design is implemented, to see if the results have improved. As covered in Chapter 12, "Testing Your Network Design," you can test network problems on a pilot implementation.	[redacted]
Checking the Status of Major Routers, Switches, and Firewalls	[redacted]
The final step in characterizing the existing internetwork is to check the behavior of the internetworking devices in the internetwork. This includes routers and switches that connect layers of a hierarchical topology, and devices that will have the most significant roles in your new network design. It's not necessary to check every LAN switch, just the major switches, routers, and firewalls.	[redacted]
Checking the behavior and health of an internetworking device includes determining how busy the device is (CPU utilization), how many packets it has processed, how many packets it has dropped, and the status of buffers and queues. Your method for assessing the health of an internetworking device depends on the vendor and architecture of the device. In the case of Cisco routers, switches, and firewalls, you can use the following Cisco IOS commands:	[redacted]

■ show buffers: Displays information on buffer sizes, buffer creation and deletion, buffer usage, and a count of successful and unsuccessful attempts to get buffers when needed.	[redacted]
■ show cdp neighbors detail: Displays information about neighbor devices, including which protocols are enabled, network addresses for enabled protocols, the number and type of interfaces, the type of platform and its capabilities, and the version of Cisco IOS Software.	[redacted]
■ show environment: Displays temperature, voltage, and blower information on the Cisco 7000 series, Cisco 7200 series, and Cisco 7500 series routers, and the Cisco 12000 series Gigabit Switch Router.	[redacted]
■ show interfaces: Displays statistics for interfaces, including the input and output rate of packets, a count of packets dropped from input and output queues, the size and usage of queues, a count of packets ignored due to lack of I/O buffer space on a card, CRC errors, collision counts, and how often interfaces have restarted.	[redacted]
■ show ip cache flow: Displays information about NetFlow, a Cisco technology that collects and measures data as it enters router and switch interfaces, including source and	[redacted]

destination IP addresses, source and destination TCP or UDP port numbers, differentiated services codepoint (DSCP) values, packet and byte counts, start and end time stamps, input and output interface numbers, and routing information (next-hop address, source and destination autonomous system numbers, and source and destination prefix masks). ■ show memory: Displays statistics about system memory, including total bytes, used bytes, and free bytes. Also shows detailed information about memory blocks. ■ show processes: Displays CPU utilization for the last 5 seconds, 1 minute, and 5 minutes, and the percentage of CPU used by various processes, including routing protocol processes, buffer management, and user-interface processes. (The show processes cpu and show processes cpu history commands are both useful variations of the show processes command.) ■ show running-config: Displays the router's configuration stored in memory and currently in use. ■ show startup-config: Displays the configuration the router will use upon	[redacted]

the next reboot.

■ show version: Displays software version and features, the names and sources of con-figuration files, the boot images, the configuration register, router uptime, and the reason for the last reboot.

Network Health Checklist

You can use the following Network Health checklist to assist you in verifying the health of an existing internetwork. The Network Health checklist is generic in nature and documents a best-case scenario. The thresholds might not apply to all networks.

□ The network topology and physical infrastructure are well documented.

□ Network addresses and names are assigned in a structured manner and are well documented.

□ Network wiring is installed in a structured manner and is well labeled.

□ Network wiring has been tested and certified.

□ Network wiring between telecommunications closets and end stations is no more than 100 meters.

□ Network availability meets current customer goals.

□ Network security meets current customer goals.

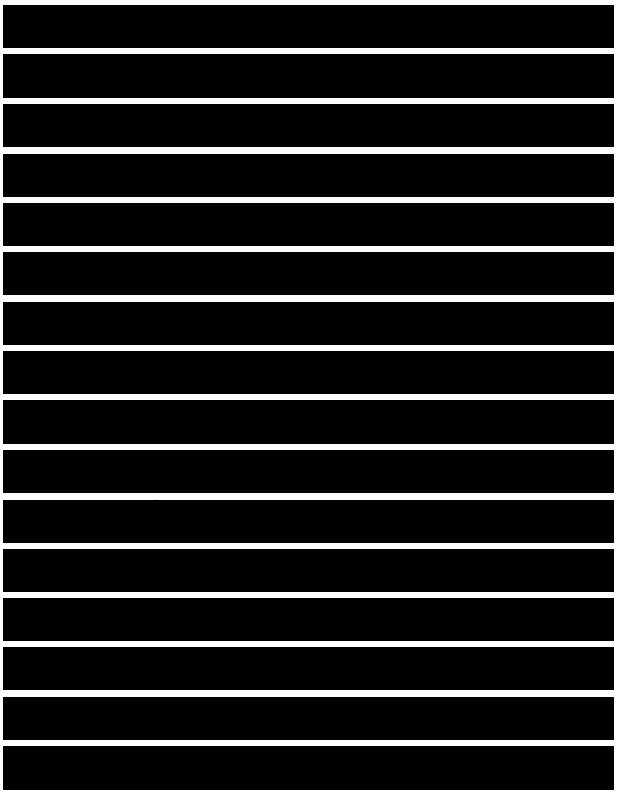
□ No LAN or WAN segments are becoming saturated (70 percent average network utilization in a 10-minute

window). □ There are no collisions on Ethernet full-duplex links. □ Broadcast traffic is less than 20 percent of all traffic on each network segment. (Some networks are more sensitive to broadcast traffic and should use a 10 percent threshold.) □ Wherever possible and appropriate, frame sizes have been optimized to be as large as possible for the data link layer in use. □ No routers are overused (5-minute CPU utilization is under 75 percent). □ On average, routers are not dropping more than 1 percent of packets. (For networks that are intentionally oversubscribed to keep costs low, a higher threshold can be used.) □ Up-to-date router, switch, and other device configurations have been collected, archived, and analyzed as part of the design study. □ The response time between clients and hosts is generally less than 100 ms (1/10th of a second).	[redacted]

Summary

This chapter covered techniques and tools for characterizing a network before designing enhancements to the network. Characterizing an existing network is an important step in top-down network design because it helps you verify that a customer's technical design goals are realistic. It also helps you understand the current topology and locate existing network segments and equipment, which will be useful information when the time comes to install new equipment.

As part of the task of characterizing the existing network, you should develop a baseline of current performance. Baseline performance measurements can be compared to new measurements once your design is implemented to demonstrate to your customer that your new design (hopefully) improves performance.

Chapter 4 12/1 chủ nhật 3 h 48 Characterizing Network Traffic This chapter describes techniques for characterizing traffic flow, traffic volume, and protocol behavior. The techniques include recognizing traffic sources and data stores, documenting application and protocol usage, and evaluating network traffic caused by common protocols. Upon completion of this chapter, you will be able to analyze network traffic patterns to help you select appropriate logical and physical network design solutions to meet a customer's goals. The previous chapter talked about characterizing the existing network in terms of its structure and performance. Because analyzing the existing situation is an important step in a systems analysis approach to design, this chapter discusses characterizing the existing network in terms of traffic flow. The chapter also covers new network design requirements, building on the first two chapters that covered business and technical design goals. This chapter refocuses on design requirements and describes requirements in terms of traffic flow, traffic load, protocol behavior, and quality of service (QoS) requirements.	Chương 4 Nghiên cứu lưu lượng mạng Chương này mô tả các kỹ thuật để nghiên cứu dòng lưu lượng, traffic volume (dung lượng mạng, tổng lưu lượng) và đặc điểm của giao thức. Các kỹ thuật bao gồm nhận diện các nguồn lưu lượng và kho dữ liệu, ghi nhận mức độ sử dụng ứng dụng và giao thức, cũng như đánh giá lưu lượng mạng của các giao thức thông thường. Sau khi hoàn thành chương này, bạn có thể phân tích các kiểu lưu lượng mạng để chọn những giải pháp thiết kế mạng logic và vật lý thích hợp đáp ứng nhu cầu của khách hàng. 

[redacted]

Characterizing Traffic Flow

[redacted]

Characterizing traffic flow involves identifying sources and destinations of network traffic and analyzing the direction and symmetry of data traveling between sources and destinations. In some applications, the flow is bidirectional and symmetric. (Both ends of the flow send traffic at about the same rate.) In other applications, the flow is bidirectional and asymmetric. (Clients send small queries and servers send large streams of data.) In a broadcast application, the flow is unidirectional and asymmetric. This section talks about characterizing the direction and symmetry of traffic flow on an existing network and analyzing flow for new network applications.

[redacted]

Identifying Major Traffic Sources and Stores

[redacted]

To understand network traffic flow, you should first identify user communities and data stores for existing and new applications.

[redacted]

Note Chapter 3, "Characterizing the Existing Internetwork," talked about locating major hosts, interconnect devices, and network segments on a customer's network. The tasks discussed in Chapter 3 facilitate the tasks discussed in this chapter of identifying major user communities and data stores.

[redacted]

A user community is a set of workers who use a particular application or set of applications. A user community can be a corporate department or set of departments. In many environments, however, application usage crosses departmental boundaries. As more corporations use matrix management and form virtual teams to complete ad hoc projects, it becomes increasingly necessary to characterize user communities by application and protocol usage rather than by departmental boundary.

To document user communities, ask your customer to help you fill out the User Communities chart shown in Table 4-1. For the Locations of Community column in Table 4-1, use location names that you already documented on a network map. For the Applications Used by Community column, use application names that you already documented in the Network Applications charts in Chapter 1, "Analyzing Business Goals and Constraints," and Chapter 2, "Analyzing Technical Goals and Tradeoffs." The case study in Chapter 10, "Selecting Technologies and Devices for Campus Networks," provides an example of a filled-in chart.

Table 4-1 User Communities

User Size of Community

Locations of Applications Used by Community (Number of Users) Community Community Name

In addition to documenting user communities, characterizing traffic flow also requires that you document major data stores. A data store (sometimes called a data sink) is an area in a network where application layer data resides. A data store can be a server, a server farm, a storage-area network (SAN), a mainframe, a tape backup unit, a digital video library, or any device or component of an internetwork where large quantities of data are stored. To help you document major data stores, ask your customer to help you fill out Table 4-2. For the Location, Applications, and Used by User Community columns, use names that you already documented on a network map and other charts.

Table 4-2 Data Stores

Data Store Location

Applications Used by User Community (or Communities)

Documenting Traffic Flow on the Existing Network Documenting traffic flow involves identifying and characterizing individual traffic flows between traffic sources and stores. Traffic flows have recently become a hot topic for discussion in the Internet community. A lot of progress is being made on defining flows, measuring flow behavior, and allowing an end station to specify performance requirements for flows. To understand traffic flow behavior better, you can read Request For Comments (RFC) 2722, "Traffic Flow	[redacted]

Measurement: Architecture." RFC 2722 describes an architecture for the measurement and reporting of network traffic flows and discusses how the architecture relates to an overall traffic flow architecture for intranets and the Internet.

Note You can find all RFCs online at http://www.ietf.org/rfc/rfcxxxx.txt, where xxxx is the number of the RFC.

Measuring traffic flow behavior can help a network designer determine which routers should be peers in routing protocols that use a peering system, such as the Border Gateway Protocol (BGP). Measuring traffic flow behavior can also help network designers do the following:

- Characterize the behavior of existing networks.
- Plan for network development and expansion.
- Quantify network performance.
- Verify the quality of network service.
- Ascribe network usage to users and applications.

An individual network traffic flow can be defined as protocol and application information transmitted between communicating entities during a single

session. A flow has attributes such as direction, symmetry, routing path and routing options, number of packets, number of bytes, and addresses for each end of the flow. A communicating entity can be an end system (host), a network, or an autonomous system (AS).	[redacted]
The simplest method for characterizing the size of a flow is to measure the number of megabytes per second (MBps) between communicating entities. To characterize the size of a flow, use a protocol analyzer or network management system to record load between important sources and destinations. You can also use Cisco NetFlow, which collects and measures data as it enters router and switch interfaces, including source and destination IP addresses, source and destination TCP or UDP port numbers, packet and byte counts, and so on.	[redacted]
You can use Table 4-3 to document information about the direction and volume of traffic flows. The objective is to document the megabytes per second between pairs of autonomous systems, networks, hosts, and applications. To get the information to fill out the charts, place a monitoring device in the core of the network and let it collect data for one or two days. To get the information to fill out the Path column, you can turn on the record-route option in an IP	[redacted]

network. The record-route option has some disadvantages, however. It doesn't support large internetworks and is often disabled for security reasons. You can also estimate the path by looking at routing tables and analyzing network traffic on multiple segments.

Table 4-3 Network Traffic Flow on the Existing Network

	Destination 1	Destination 2	Destination 3	Destination n
	MBps Path	MBps Path	MBps Path	MBps Path
Source 1				
Source 2				
Source 3				
Source n				

Characterizing Types of Traffic Flow for New Network Applications

As mentioned, a network flow can be characterized by its direction and symmetry. Direction specifies whether data travels in both directions or in just one direction. Direction also specifies the path that a flow takes as it travels from source to destination through an internetwork. Symmetry describes whether the flow tends to have higher performance or QoS requirements in one direction than the other direction. Many network applications have different requirements in each direction. Some data link layer transmission

technologies, such as Asymmetric Digital Subscriber Line (ADSL), are fundamentally asymmetric. A good technique for characterizing network traffic flow is to classify applications as supporting one of a few well-known flow types:

- Terminal/host traffic flow
- Client/server traffic flow
- Peer-to-peer traffic flow
- Server/server traffic flow
- Distributed computing traffic flow

In his book Network Analysis, Architecture, and Design, Third Edition, James D. McCabe does an excellent job of characterizing and distinguishing flow models. The descriptions of flow types in the sections that follow are partially based on McCabe's work.

Terminal/Host Traffic Flow

Terminal/host traffic is usually asymmetric. The terminal sends a few characters and the host sends many characters. Telnet is an example of an application that generates terminal/host traffic. The default behavior for Telnet is that the terminal sends each character a user types in a single packet. The host returns multiple characters, depending on what the user typed. As an illustration, consider the beginning of a Telnet session that starts with the user typing a username. When the host

receives each packet for the characters in the name, the host sends back a message (such as Password Required) in one packet.

Note Default Telnet behavior can be changed so that instead of sending one character at a time, the terminal sends characters after a timeout or after the user types a carriage return. This behavior uses network bandwidth more efficiently but can cause problems for some applications. For example, the vi editor on UNIX systems must see each character immediately to recognize whether the user has pressed a special character for moving up a line, down a line, to the end of a line, and so on.

Terminal/host traffic flows are less prevalent on networks than they once were, but they have not disappeared. In fact, so-called thin clients, which have become quite popular, can behave like terminal/host applications. Thin clients are covered following the next section on client/server traffic flow.

Client/Server Traffic Flow

Client/server traffic is the best-known and most widely used flow type. Servers are generally powerful computers

dedicated to managing disk storage, printers, or other network resources. Clients are PCs or workstations on which users run applications. Clients rely on servers for access to resources, such as storage, peripherals, application software, and processing power.	[redacted]
Clients send queries and requests to a server. The server responds with data or permission for the client to send data. The flow is usually bidirectional and asymmetric. Requests from the client are typically small frames, except when writing data to the server, in which case they are larger. Responses from the server range from 64 bytes to 1500 bytes or more, depending on the maximum frame size allowed for the data link layer in use.	[redacted]
Client/server protocols include Server Message Block (SMB), Network File System (NFS), Apple Filing Protocol (AFP), and NetWare Core Protocol (NCP). In a TCP/IP environment, many applications are implemented in a client/server fashion, although the applications were invented before the client/server model was invented. For example, FTP has a client and server side. FTP clients use FTP applications to talk to FTP servers. X Window	[redacted]

System is an example of a server (the screen manager) that actually runs on the user's machine. This can lead to a great deal of traffic in both directions, such as when the user enables a blinking cursor or ticking clock that needs continual updating across the network, even when the user isn't present.	[redacted]
These days, HTTP is the most widely used client/server protocol. Clients use a web browser application, such as Firefox, to talk to web servers. The flow is bidirectional and asymmetric. Each session often lasts just a few seconds because users tend to jump from one website to another.	[redacted]
The flow for HTTP traffic is not always between the web browser and the web server because of caching. When users access data that has been cached to their own systems, there is no network traffic. Another possibility is that a network administrator has set up a cache engine. A cache engine is software or hardware that makes recently accessed web pages available locally, which can speed the delivery of the pages and reduce WAN bandwidth utilization. A cache engine can also be used to control the type of content that users are allowed to view.	[redacted]

A content delivery network (CDN) can also affect the flow of HTTP traffic. A CDN is a network of servers that delivers web pages to users based on their geographic location. A CDN copies the pages of a website to a network of servers that can be dispersed at different locations. When a user requests a web page that is part of a CDN, the CDN redirects the request to a server that is closest to the user and delivers the cached content.

The CDN can also communicate with the originating server to deliver any content that has not been previously cached. CDNs speed the delivery of content and can protect a website from large surges in traffic.

Thin Client Traffic Flow

A special case of the client/server architecture is a thin client, which is software or hardware that is designed to be particularly simple and to work in an environment where the bulk of data processing occurs on a server. With thin client technology (also known as server-based computing), user applications

originate on a central server. In some cases, the application runs on the central server, and in other cases, the software is installed on the server and is downloaded into the client machine for execution.

An information appliance or computing appliance is a thin client designed to perform a particular set of dedicated tasks. A computing appliance could be a cash register, a dedicated email machine, or a database-retrieval device. Computing appliances often run the Linux operating system and a Java-enhanced Internet browser.

The main advantage of thin client technology is lower support costs. Network managers can have a centralized base of applications that are managed, configured, and upgraded only once. There is no need to individually configure each user's machine. In addition, because applications are controlled from the central server, security can be better managed. Thin clients provide a lower total cost of ownership (TCO) and a scalable TCO for large enterprises. Thin client technology is not applicable to every computing application, however, because users might need computers capable of operating without constant connection to a central server.

A downside of thin client technology is that the amount of data flowing from the server to the client can be substantial, especially when many computers start up at the same time every day. Networks that include thin clients should be carefully designed with sufficient capacity and an appropriate topology. Switched networking (rather than shared media) is recommended, and to avoid problems caused by too much broadcast traffic, each switched network should be limited to a few hundred thin clients and their servers. The switched networks can be connected via routers for communications between departments and accessing outside networks such as the Internet.

Peer-to-Peer Traffic Flow

With peer-to-peer traffic, the flow is usually bidirectional and symmetric. Communicating entities transmit approximately equal amounts of information. There is no hierarchy. Each device is considered as important as each other device, and no device stores substantially more data than any other device. In small LAN environments, network administrators often set up PCs in a peer-to-peer configuration so that everyone can access each other's data and printers. There is no central file or print server. Another example of a peer-to-peer environment is a set of multiuser UNIX hosts where users set up FTP,

Telnet, HTTP, and NFS sessions between hosts. Each host acts as both a client and server. There are many flows in both directions.

Recently, peer-to-peer applications for downloading music, videos, and software have gained popularity. Each user publishes music or other material and allows other users on the Internet to download the data. This is considered peer-to-peer traffic because every user acts as both a distributor and consumer of data. Traffic flow is bidirectional and symmetric. Most enterprises and many university networks disallow this type of peer-to- peer traffic for two reasons:

- It can cause an inordinate amount of traffic.
- The published material is often copyrighted by someone other than the person publishing it.

One other example of a peer-to-peer application is a meeting between business people at remote sites using videoconferencing equipment. In a meeting, every attendee can communicate as much data as needed at any time. All sites have the same QoS requirements. A meeting is different from a situation where videoconferencing is used to disseminate information. With

information dissemination, such as a training class or a speech by a corporate president to employees, most of the data flows from the central site. A few questions are permitted from the remote sites. Information dissemination is usually implemented using a client/server model. Server/Server Traffic Flow Server/server traffic includes transmissions between servers and transmissions between servers and management applications. Servers talk to other servers to implement directory services, to cache heavily used data, to mirror data for load balancing and redundancy, to back up data, and to broadcast service availability. Servers talk to management applications for some of the same reasons but also to enforce security policies and to update network management data. With server/server network traffic, the flow is generally bidirectional. The symmetry of the flow depends on the application. With most server/server applications, the flow is symmetrical, but in some cases there is a hierarchy of servers, with some servers sending and storing more data than others.	[redacted]

Distributed Computing Traffic Flow

Distributed computing refers to applications that require multiple computing nodes working together to complete a job. Some complex modeling and rendering tasks cannot be accomplished in a reasonable timeframe unless multiple computers process data and run algorithms simultaneously. The visual effects for movies are often developed in a distributed computing environment. Distributed computing is also used in the semiconduc-tor industry to serve the extreme computing needs of microchip design and verification, and in the defense industry to provide engineering and military simulations.

With distributed computing, data travels between a task manager and computing nodes and between computing nodes. In his book Network Analysis, Architecture, and Design, Third Edition, McCabe distinguishes between tightly coupled and loosely coupled computing nodes. Nodes that are tightly coupled transfer information to each other frequently. Nodes that are loosely coupled transfer little or no information.

With some distributed computing applications, the task manager tells the computing nodes what to do on an infrequent basis, resulting in little traffic flow. With other applications, there is

frequent communication between the task manager and the computing nodes. In some cases, the task manager allocates tasks based on resource availability, which makes predicting flow somewhat difficult.	[redacted]
Characterizing traffic flow for distributed computing applications might require you to study the traffic with a protocol analyzer or model potential traffic with a network simulator.	[redacted]
Traffic Flow in Voice over IP Networks	[redacted]
The most important concept to understand when considering traffic flow in VoIP networks is that there are two flows. The flow associated with transmitting the audio voice is separate from the flow associated with call setup and teardown. The flow for transmitting the digital voice is peer-to-peer, between two phones or between two PCs running software such as Skype or Cisco IP Communicator (CIPC). Call setup and teardown, on the other hand, can be characterized as a client/server flow because a phone needs to talk to a more complicated device, such as a server or traditional phone switch, that understands phone numbers, addresses, capabilities negotiation, and so on.	[redacted]

	[redacted]
The audio voice flow between two IP endpoints is carried by the Real-Time Transport Protocol (RTP), which is a connectionless protocol that runs on top of UDP. The main call setup, teardown, and control protocols in an IP network are H.323, the Cisco Skinny Client Control Protocol (SCCP), Simple Gateway Control Protocol (SGCP), Media Gateway Control Protocol (MGCP), and Session Initiation Protocol (SIP). These signaling protocols run between an IP endpoint and a voice-enabled server and follow the client/server paradigm.	[redacted]
Both traditional voice networks, which are based on private branch exchanges (PBX) and circuit switching, and modern VoIP networks, which use packet switching, must handle two fundamental functions: call control and call switching. Call control handles call setup and teardown, addressing and routing, and informational and supplementary services. A fundamental job of call control is to compare the digits dialed by the user making a call to configured number patterns to determine how to route a call. In a VoIP network,	[redacted]

call control maps a telephone number or username to an IP destination address, which is understood by the packet infrastructure layer. In a Cisco environment, an IP phone talks to the Cisco Unified Communications Manager software to get an IP destination address to use.	[redacted]
Call switching handles the actual switching of calls. In traditional voice networks, when a call is placed, a PBX connects the calling phone via a so-called line-side interface to another phone's line-side interface. If the call is destined for the public switched telephone network (PSTN), the call switching function connects the line-side interface with the trunk-side interface. In a VoIP network, switching of voice packets is handled by the packet infrastructure layer, which provides Ethernet switches and IP routers. The line- side devices are IP phones and PCs running VoIP software such as Cisco IP Communicator (CIPC). The trunk-side interface is handled by a PSTN gateway such as a voice-enabled router.	[redacted]
The audio voice packets, which are encapsulated in RTP, UDP, and IP and a data link layer header, might be switched through the internetwork using	[redacted]

a different path from that used by the call control packets. The audio packets represent a distinct traffic flow that should be analyzed differently from the call control flow when considering bandwidth and QoS requirements.

[redacted]

Documenting Traffic Flow for New and Existing Network Applications

[redacted]

To document traffic flow for new (and existing) network applications, characterize the flow type for each application and list the user communities and data stores that are associated with applications. You can use Table 4-4 to enhance the Network Applications charts already discussed in Chapters 1 and 2. When filling out Table 4-4, use the same application names you used in the other charts. (The Protocols Used by Application, Approximate Bandwidth Requirement for the Application, and QoS Requirements columns are described later in this chapter.)

[redacted]

Name of Application	Type of Traffic Flow	Protocols Used by Application	User Communities (Servers, Hosts, and So On) That Use the Application	Data Stores Used by Application	Approximate Bandwidth Requirement for the Application	QoS Requirement

When identifying the type of traffic flow for an application, select one of the

[redacted]

well- known types:

- Terminal/host
- Client/server
- Peer-to-peer
- Server/server
- Distributed computing

If necessary, add a comment to qualify the type of flow. For example, if the type is terminal/host and full screen, make sure to say this, because in a full-screen application, the host sends more data than in a so-called dumb-terminal application. If the flow type is distributed computing, add some text to specify whether the computing nodes are tightly or loosely coupled. If the flow is peer-to-peer and used for voice, include a comment to that effect because voice traffic needs different QoS characteristics than most peer-to- peer applications need.

Characterizing Traffic Load

To select appropriate topologies and technologies to meet a customer's goals, it is important to characterize traffic load with traffic flow. Characterizing traffic load can help you design networks with sufficient capacity for local usage and internetwork flows.

Because of the many factors involved in characterizing network traffic, traffic load estimates are unlikely to be precise. The goal is simply to avoid a design that

has any critical bottlenecks. To avoid bottlenecks, you can research application-usage patterns, idle times between packets and sessions, frame sizes, and other traffic behavioral patterns for application and system protocols. For customers with numerous applications, this level of analysis might not be practical, however. For these customers, you could limit the analysis to the top five or ten applications.	[redacted]
Another approach to avoiding bottlenecks is simply to throw large amounts of bandwidth at the problem (also known as overprovisioning). A strict interpretation of systems analysis principles wouldn't approve of such an approach, but bandwidth is cheap these days. LAN bandwidth is extremely cheap. There's no excuse for not using Fast Ethernet (or better) on all new workstations and switches, and most organizations can also afford to use Gigabit Ethernet on switch-to-switch and switch-to-server links. WAN bandwidth is still expensive in some parts of the world, including rural areas of the United States.	[redacted]
But in many parts of the United States	[redacted]

and the rest of the world, bandwidth has been overprovisioned and isn't overutilized. If you know that bandwidth will not be a constraint in your network designs, you can skip the next few sections and jump to "Characterizing Traffic Behavior."

[redacted]

Calculating Theoretical Traffic Load

As described in Chapter 2, traffic load (sometimes called offered load) is the sum of all the data all network nodes have ready to send at a particular time. A general goal for most network designs is that the network capacity should be more than adequate to handle the traffic load. The challenge is to determine if the capacity proposed for a new network design is sufficient to handle the potential load.

[redacted]

In his book Local and Metropolitan Area Networks, Sixth Edition, William Stallings provides some back-of-the-envelope computations for calculating traffic load. Stallings points out that you can make an elementary calculation based simply on the number of stations transmitting, how quickly each station generates messages, and the size of messages. For example, for a network with a proposed capacity of 1 Mbps, if 1000 stations send 1000-bit frames every second, the offered load equals the capacity. Although Stallings was referring to the capacity of a LAN, capacity could refer to the capacity of a WAN link, an entire internetwork or

[redacted]

parts of an internetwork, or the backplane of a switch or router.

In general, to calculate whether capacity is sufficient, only a few parameters are necessary:

- The number of stations
- The average time that a station is idle between sending frames
- The time required to transmit a message once medium access is gained

By studying idle times and frame sizes with a protocol analyzer, and estimating the number of stations, you can determine if the proposed capacity is sufficient.

If you research traffic flow types, as discussed earlier in this chapter, you can develop more precise estimates of load. Instead of assuming that all stations have similar load- generating qualities, you can assume that stations using a particular application have similar load-generating qualities. Assumptions can be made about frame size and idle time for an application after you have classified the type of flow and identified the protocols (discussed later in this chapter) used by the application.

For a client/server application, idle time for the server depends on the number of clients using the server, and the architecture and performance characteristics of the server (disk access speed, RAM access speed, caching mechanisms, and so on). By studying network traffic from servers with a protocol analyzer, you can estimate an average idle time. In general, servers should have little idle time. If a server is underutilized, you should consider moving it to a shared server platform using server virtualization technology.

Idle time on the client side depends partly on user action, which means it is impossible to precisely predict idle time. However, you can make estimates of idle time by studying traffic with a protocol analyzer and using scripts to simulate worst-case user actions, or by using a network-modeling tool. A good network-modeling tool knows what assumptions to make about idle time, MAC-layer delays, the distribution of packet arrival at servers and internetworking devices, and queuing and buffering behavior at internetworking devices.

After you have identified the approximate traffic load for an

application flow, you can estimate total load for an application by multiplying the load for the flow by the number of devices that use the application. The research you do on the size of user communities and the number of data stores (servers) can help you calculate an approximate aggregated bandwidth requirement for each application. Documenting the location of user communities and data stores, which you did in Tables 4-1 and 4-2, can help you understand the amount of traffic that will flow from one segment to another. This can aid in the selection of backbone technologies and internetworking devices. When estimating traffic load, in addition to investigating idle times between packets, frames sizes, and flow behavior, you should also investigate application-usage patterns and QoS requirements. Some applications are used infrequently, but require a large amount of bandwidth when they are used. For example, perhaps workers normally access an Ethernet LAN to read email, store small files on servers, and print small documents on network printers. But once every three months, they use real-time multimedia software on their PCs to watch the corporate president's speech on quarterly sales figures. This means that once a quarter traffic characteristics and QoS requirements are different than normal.	[redacted]

In general, to accurately characterize traffic load, you need to understand application- usage patterns and QoS requirements in addition to idle times and frame sizes. Some applications expect the network to simply make a best effort to meet load (bandwidth) requirements. Other applications, such as video applications, have inflexible requirements for a constant amount of bandwidth.

The next section covers characterizing usage patterns in more detail. The section "Characterizing Quality of Service Requirements" later in this chapter discusses characterizing QoS requirements.

Documenting Application-Usage Patterns

The first step in documenting application-usage patterns is to identify user communities, the number of users in the communities, and the applications the users employ. This step, which was already covered in the "Identifying Major Traffic Sources and Stores" section in this chapter, can help you identify the total number of users for each application.

In addition to identifying the total number of users for each application, you should also document the following information:

■ The frequency of application sessions (number of sessions per day, week, month, or whatever time period is appropriate)

■ The length of an average application session

■ The number of simultaneous users of an application

Armed with information on the frequency and length of sessions and the number of simultaneous sessions, you can more accurately predict the aggregate bandwidth requirement for all users of an application. If it is not practical to research these details, you can make some assumptions:

■ The number of users of an application equals the number of simultaneous users.

■ All applications are used all the time, so that your bandwidth calculation is a worst- case (peak) estimate.

■ Each user opens just one session, and that session lasts all day until the user shuts down the application at the end of the day.

Refining Estimates of Traffic Load Caused by Applications

To refine your estimate of application bandwidth requirements, you need to research the size of data objects sent by applications, the overhead caused by protocol layers, and any additional load caused by application initialization. (Some applications send much more traffic during initialization than during

xấu nhất tăng đột biến

Tinh chỉnh tính toán tải lưu lượng của các ứng dụng

xác định chính xác hơn yêu cầu băng thông của ứng dụng

kích thước của các đối tượng dữ liệu

bổ sung quá trình

steady-state operation.)	lưu lượng
Because applications and users vary widely in behavior, it is hard to accurately estimate the average size of data objects that users transfer to each other and to servers. (The true engineering answer to most questions related to network traffic is "it depends.") In the past, some networking books specified the average size of an email message, web page, multimedia object, database record, and so on. These days, with email messages that include video attachments, web pages that offer video on demand, and databases that might be used for voicemail, it's difficult to make any generalizations about the average size of objects sent on a network. A thorough analysis of actual application behavior is required if you want to get a precise answer as you provision your networks to handle the offered load.	những đặc trưng của kích thước các đối tượng lưu lượng mạng kích thước đặc điểm của
To completely characterize application behavior, you should investigate which protocols an application uses. When you know the protocols, you can calculate traffic load more precisely by adding the size of protocol headers to the size of data objects. Table 4-5 shows some typical protocol header sizes.	đặc điểm của tải lưu lượng kích thước tiêu đề kích thước cho thấy kích thước điển hình của tiêu đề giao thức
Table 4-5 Traffic Overhead for Various Protocols Protocol Overhead Details	Tổng lưu lượng

Total Bytes Ethernet Version II Preamble = 8 bytes, header = 14 bytes, CRC interframe gap (IFG) = 12 bytes= 4 bytes, 38 IEEE 802.3 with 802.2 Preamble = 8 bytes, header = 14 bytes, LLC SNAP (if present) = 5 bytes, CRC = 4 bytes, = 3 or 4 bytes, IFG = 12 bytes 46 HDLC Flags = 2 bytes, addresses = 2 bytes, control CRC = 4 bytes = 1 or 2 bytes, 10 IP Header size with no options 20 TCP Header size with no options 20 UDP Header size 8 Depending on the applications and protocols that a workstation uses, workstation initialization can cause a load on networks due to the number of packets and, in some cases, the number of broadcast packets. Although this is becoming less of a problem as network bandwidth has become so inexpensive, and extremely fast CPUs in workstations are readily available so that broadcast processing isn't a concern, it is still worthy of note for net-works that are bound by bandwidth and processing concerns, especially networks at schools and nonprofit organizations. In addition to applications that are set to start upon bootup, the following system-level protocols send packets as a workstation initializes: ■ Address Resolution Protocol (ARP)	[redacted] trên mạng gói tin gói tin broadcast nhưng chúng ta vẫn cần chú ý rằng Cùng với sẽ chạy gói tin [redacted]

■ Dynamic Host Configuration Protocol (DHCP)
■ Internet Control Message Protocol (ICMP), version 4 and 6
■ Internet Group Management Protocol (IGMP), version 4 and 6
■ Domain Name System (DNS)
■ Multicast DNS (mDNS)
■ NetBIOS name queries
■ Network Time Protocol (NTP)

■ Simple Service Discovery Protocol (SSDP)
■ Service Location Protocol (SLP)
Simple Network Management Protocol (SNMP)

Estimating Traffic Load Caused by Routing Protocols

At this point in the network design process, you might not have selected routing protocols for the new network design, but you should have identified routing protocols running on the existing network.

Estimating traffic load caused by legacy routing protocols is important in a topology that includes many networks on one side of a slow WAN link. A router sending a large dis- tance-vector routing table every half minute can use a significant percentage of WAN bandwidth. Because routing protocols limit the number of routes per packet, on large networks, a router sends multiple packets to send the entire table. Routing Information Protocol (RIP), for example, sends a routing packet every 30 seconds. Each route in the packet uses 20 bytes, and there are 25 routes per packet. When headers are added,

this means that a router running RIP sends one or more 532-byte packets every 30 seconds, depending on the size of the routing table.

Newer routing protocols, such as Open Shortest Path First (OSPF) and Enhanced Interior Gateway Routing Protocol (EIGRP), use little bandwidth. In the case of OSPF, your main concern should be the amount of bandwidth consumed by the database-synchronization packets that routers send every 30 minutes. By subdividing an OSPF network into areas and using route summarization, this traffic can be minimized. Other than the database-synchronization traffic, the only traffic OSPF sends after initialization is small Hello packets every 10 seconds.

EIGRP also sends Hello packets but more frequently than OSPF (every 5 seconds). On the other hand, EIGRP doesn't send any periodic route updates or database-synchroniza- tion packets. It sends route updates only when there are changes.

Characterizing Traffic Behavior

To select appropriate network design solutions, you need to understand protocol and application behavior in addition to traffic flows and load. For example, to select appropriate LAN topologies, you need to investigate the level of broadcast traffic on the LANs.

To provision adequate capacity for LANs and WANs, you need to check for extra bandwidth utilization caused

gửi sau mỗi 30 phút

chúng ta có thể giảm lưu lượng này đến mức tối thiểu đồng bộ hóa cơ sở dữ liệu lưu

tin đồng bộ hóa

lưu lượng

dòng và tải lưu lượng
đặc điểm của

dung lượng

by protocol inefficiencies and suboptimal frame sizes or retransmission timers.

Broadcast/Multicast Behavior
A broadcast frame is a frame that goes to all network stations on a LAN. At the data link layer, the destination address of a broadcast frame is FF:FF:FF:FF:FF:FF (all 1s in binary). A multicast frame is a frame that goes to a subset of stations. For example, a frame destined to 01:00:0C:CC:CC:CC goes to Cisco routers and switches that are running the Cisco Discovery Protocol (CDP) on a LAN.

Layer 2 internetworking devices, such as switches and bridges, forward broadcast and multicast frames out all ports. The forwarding of broadcast and multicast frames can be a scalability problem for large flat (switched or bridged) networks. A router does not forward broadcasts or multicasts. All devices on one side of a router are considered part of a single broadcast domain.

In addition to including routers in a network design to decrease broadcast forwarding, you can also limit the size of a broadcast domain by implementing virtual LANs (VLAN). VLAN technology, which Chapter 5, "Designing a Network Topology," discusses in more detail, allows a network administrator to subdivide users into subnets by associating switch ports with one or more VLANs.

[redacted]
định thời
[redacted]
chuyển tiếp
gây khó
khăn cho việc mở rộng
[redacted]
Cùng với việc gộp
vào
[redacted]

Although a VLAN can span many switches, broadcast traffic within a VLAN is not transmitted outside the VLAN.	lưu lượng broadcast
Too many broadcast frames can overwhelm end stations, switches, and routers. It is important that you research the level of broadcast traffic in your proposed design and limit the number of stations in a single broadcast domain. The term broadcast radiation is often used to describe the effect of broadcasts spreading from the sender to all other devices in a broadcast domain. Broadcast radiation can degrade performance at network endpoints.	ảnh hưởng của các broadcast
The network interface card (NIC) in a network station passes broadcasts and relevant multicasts to the CPU of the station. Some NICs pass all multicasts to the CPU, even when the multicasts are not relevant, because the NICs do not have driver software that is more selective. Intelligent driver software can tell a NIC which multicasts to pass to the CPU. Unfortunately, not all drivers have this intelligence. The CPUs on network stations become overwhelmed when processing high levels of broadcasts and multicasts. If more than 20 percent of the network traffic is broadcasts or multicasts, the network needs to be segmented using routers or VLANs.	một trạm trong mạng cho phép các broadcast và multicast có liên quan đến CPU của trạm Tuy nhiên Các đoạn
Another possible cause of heavy broadcast traffic is intermittent broadcast storms caused by misconfigured or misbehaving network	thất thường

stations. For example, a misconfigured subnet mask can cause a station to send ARP frames unnecessarily because the station does not correctly distinguish between station and broadcast addresses, causing it to send ARPs for broadcast addresses. In general, however, broadcast traffic is necessary and unavoidable. Routing and switching protocols use broadcasts and multicasts to share information about the internetwork topology. Servers send broadcasts and multicasts to advertise their services. Clients send broadcasts and multicasts to find services and check for uniqueness of addresses and names. Network Efficiency Characterizing network traffic behavior requires gaining an understanding of the efficiency of new network applications. Efficiency refers to whether applications and protocols use bandwidth effectively. Efficiency is affected by frame size, the interaction of protocols used by an application, windowing and flow control, and error-recovery mechanisms. Frame Size Using a frame size that is the maximum supported for the medium in use has a positive impact on network performance for bulk applications. For file-transfer applications, in particular, you should use the largest possible maximum transmission unit (MTU). Depending on the protocol stacks that your customer will use in the new network design, the MTU can be configured for some	được các địa chỉ chính xác đối với các Tuy nhiên, nhìn chung dịch vụ của họ chúng ta mức độ sử dụng băng thông hiệu quả của các ứng dụng và giao thức phục hỗ trợ cho môi trường đang dùng

applications.

In an IP environment, you should avoid increasing the MTU to larger than the maximum supported for the media traversed by the frames, to avoid fragmentation and reassembly of frames. When devices such as end nodes or routers need to fragment and reassemble frames, performance degrades.

Modern operating systems support MTU discovery. With MTU discovery, the software can dynamically discover and use the largest frame size that will traverse the network without requiring fragmentation. MTU discovery generally works, but there are some problems with it as mentioned in the "Analyzing Network Efficiency" section in Chapter 3.

Windowing and Flow Control

To really understand network traffic, you need to understand windowing and flow control. A TCP/IP device, for example, sends segments (packets) of data in quick sequence, without waiting for an acknowledgment, until its send window has been exhausted. A station's send window is based on the recipient's receive window. The recipient states in every TCP packet how much data it is ready to receive. This total can vary from a few bytes up to 65,535 bytes. The recipient's receive window is based on how much memory the receiver has and how quickly it can process received data. You can optimize network efficiency by increasing memory and CPU power on end stations, which can

đối với môi trường truyền khung

nút

Các hệ điều hành

sẽ truyền qua mạng

gắn liền với nó

suất

dòng

các đoạn

trong hàng đợi nhanh

chờ xác nhận

(cửa sổ truyền, cửa sổ gửi)

rỗng

Tổng

lượng dữ liệu

result in a larger receive window.

Note Theoretically, the optimal window size is the bandwidth of a link multiplied by delay on the link. To maximize throughput and use bandwidth efficiently, the send window should be large enough for the sender to completely fill the bandwidth pipe with data before stopping transmission and waiting for an acknowledgment.
RFC 1323 illustrates the need for a larger window than the standard maximum TCP window size of 65,535 bytes. The product of bandwidth times delay is larger than 65,535 bytes on links of high speeds but long delay, such as high-capacity satellite channels and terrestrial fiber-optic links that go a long distance (for example, across the United States). RFC 1323 refers to a path operating in this region as a long, fat pipe and a network containing this path as a long, fat network or LFN (pronounced “elephant”). If your customer’s network includes any LFNs, you should recommend implementations of TCP that are based on RFC 1323. The Window field in the TCP header is 16 bits. RFC 1323 defines a window-scale extension that expands the definition of the TCP window to 32 bits and uses a scale factor to carry this 32-bit value in the 16-bit Window field. During the TCP threeway handshake, hosts can include a TCP option that indicates their support for the win- dow-scale extension.

[redacted]

xác

[redacted]

Tích băng thông nhân thời gian trễ

đi qua

dưới dạng

to tròn

dưới dạng

to tròn

có

tiêu đề

xác định độ mở rộng tỷ lệ cửa sổ khái quát hóa định nghĩa cửa sổ TCP hệ số tỷ lệ mang

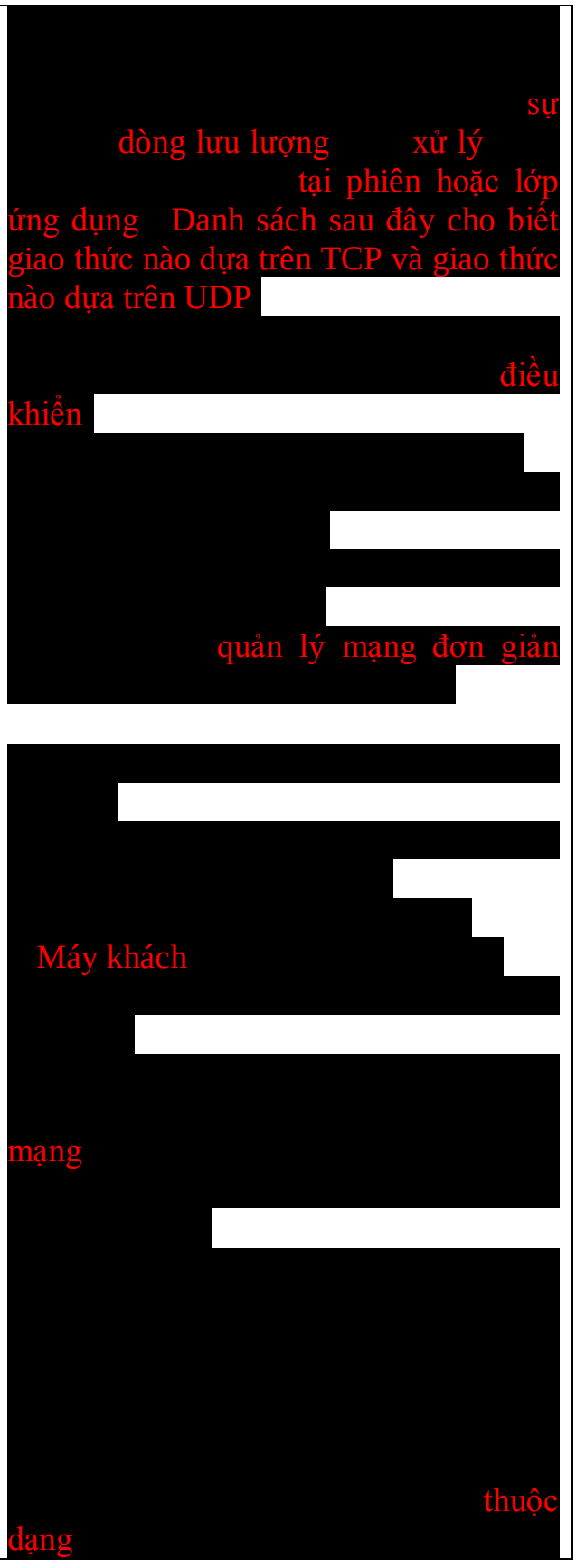
để chỉ khả năng hỗ trợ của chúng cho sự mở rộng tỷ lệ cửa sổ

Some IP-based applications run on top of UDP, not TCP. In this case, there is either no flow control or the flow control is handled at the session or application layer. The following list shows which protocols are based on TCP and which protocols are based on UDP:

- File Transfer Protocol (FTP): TCP port 20 (data) and TCP port 21 (control)
- Telnet: TCP port 23
- Simple Mail Transfer Protocol (SMTP): TCP port 25
- Hypertext Transfer Protocol (HTTP): TCP port 80
- Simple Network Management Protocol (SNMP): UDP ports 161 and 162
- Domain Name System (DNS): UDP port 53
- Trivial File Transfer Protocol (TFTP): UDP port 69
- DHCP server: UDP port 67
- DHCP client: UDP port 68
- Remote Procedure Call (RPC): UDP port 111

Protocols that run on UNIX systems, such as NFS and Network Information Services (NIS), often use RPC, although newer versions support TCP.

Protocols that send a reply to each request are often called Ping-Pong protocols. Ping- Pong protocols do not use bandwidth efficiently. As an example, consider SMB, the file-sharing protocol used on Windows platforms. Although SMB uses non-Ping-Pong protocols at the lower layers,



it behaves like a Ping-Pong protocol at the application layer. The client sends a Read request to receive data. A typical SMB server can send 32 KB of data at a time, divided into TCP segments. The client waits until this data is received before requesting more data. This greatly limits throughput.	Máy khách
Do the calculation for an IPsec VPN, for example, that connects a user in New York City with a server in Washington, DC. The one-way delay is about 50 ms. Ignoring client and server delays for disk I/O and serialization delay, the client receives at most 32 KB every 100 ms or 320 KBps. This means that the maximum throughput is 2.56 Mbps. With packet losses and consequent retransmissions, actual throughput might be even lower. This problem suggests that you should advise your network design customers not to plan on launching programs or copying large directories off a remote SMB file server. Error-Recovery Mechanisms	Chúng ta hãy xét ví dụ tính toán cho trường hợp một Thời gian trễ thời gian trễ của máy khách và máy chủ máy khách tổn hao gói tin
Poorly designed error-recovery mechanisms can waste bandwidth. For example, if a protocol retransmits data quickly without waiting a long enough time to receive an acknowledgment, this can cause performance degradation for the rest of the network due to the bandwidth used. Acknowledgments at multiple layers can also waste bandwidth.	chờ đủ thời gian để nhận xác nhận
Connectionless protocols usually do not implement error recovery. Most data	kết nối

link layer and network layer protocols are connectionless. Some transport layer protocols, such as UDP, are connectionless.	không tầng giao vận cũng thuộc dạng không kết nối
Error-recovery mechanisms for connection-oriented protocols vary. TCP implements an adaptive retransmission algorithm, which means that the rate of retransmissions slows when the network is congested, which optimizes the use of bandwidth.	hướng lại việc sử dụng
Newer TCP implementations also implement Selective ACK (SACK), as described in RFC 2018. Without SACK, error-prone, high-delay paths can experience low throughput due to the way that TCP acknowledges data. TCP acknowledgments (ACK) are cumulative up to the point where a problem occurs. If segments get lost, the ACK number is one more than the number of the last byte that was received before the loss, even if more segments arrived after the loss. There's no way for the receiver to report a hole in the received data. This causes the sender either to wait a round-trip time to find out about each lost segment or to unnecessarily retransmit segments that the recipient may have correctly received.	các đường truyền dễ bị lỗi, thời gian trễ cao có thể bị tình trạng thông lượng thấp TCP xác nhận dữ liệu Các xác nhận đến lúc nào sự cố xuất hiện bị mất số ACK sẽ lớn hơn số byte cuối cùng nhận được trước khi mất thu người nhận
With the SACK mechanism, the TCP recipient fills in the SACK option field in the TCP header to inform the sender of the noncontiguous blocks of data that have been received. The sender can then retransmit only the missing segments. RFC 2018 defines a TCP option for	trường tùy chọn tiêu đề đã nhận được

filling in the sequence numbers for received blocks and another TCP option for informing the recipient during the three-way handshake that the host supports SACK.

thủ tục bắt tay ba bước rằng

Using a protocol analyzer, you can determine whether your customer's protocols implement effective error recovery. In some cases you can configure retransmission and timeout timers or upgrade to a better protocol implementation.

bộ phân tích giao thức các giao thức của khách hàng khắc phục lỗi có hiệu quả hay không
chúng ta có thể cấu hình bộ định thời truyền lại và tạm ngưng

Characterizing Quality of Service Requirements

Mô tả các yêu cầu chất lượng dịch vụ

Analyzing network traffic requirements isn't quite as simple as identifying flows, measuring the load for flows, and characterizing traffic behavior such as broadcast and error- recovery behavior. You need to also characterize the QoS requirements for applications.

Phân tích các yêu cầu lưu lượng mạng không đơn giản như xác định dòng lưu lượng, đo tải của dòng lưu lượng, và xác định đặc trưng của lưu lượng, chẳng hạn như broadcast và khắc phục lỗi

Just knowing the load (bandwidth) requirement for an application is not sufficient. You also need to know if the requirement is flexible or inflexible. Some applications continue to work (although slowly) when bandwidth is not sufficient. Other applications, such as voice and video applications, are rendered useless if a certain level of bandwidth is not available. In addition, if you have a mix of flexible and inflexible applications on a network, you need to determine if it is practical to borrow bandwidth from the flexible application to keep the inflexible application working.

Chỉ biết yêu cầu về tải (băng thông) của một ứng dụng vẫn chưa đủ Bạn cũng cần phải biết yêu cầu này mềm dẻo hay khắt khe

ứng dụng
các ứng dụng thoại và video sẽ không hoạt động
ứng
dụng mềm dẻo và không mềm dẻo trên một mạng
ứng dụng mềm dẻo
mềm dẻo

As discussed in Chapter 2, voice is also inflexible with regard to delay. Voice is also sensitive to packet loss, which

ứng dụng thoại không mềm dẻo đối với thời gian trễ Ứng dụng thoại cũng nhạy với sự

results in voice clipping and skips. Without proper networkwide QoS configuration, loss can occur because of congested links and poor packet- buffer and queue management on routers.

The sections that follow cover analyzing QoS requirements using ATM and Internet Engineering Task Force (IETF) techniques. The goal of these sections is to introduce you to terminology that ATM and IETF engineers use for classifying traffic and specifying QoS requirements for classes of traffic. Although the material is highly technical and detailed, it should give you some fundamental ideas about classifying the types of applications that will play a part in your network design, and it should prepare you for future chapters that cover strategies for designing and optimizing networks that can meet the needs of various applications.

ATM QoS Specifications

In their document "Traffic Management Specification Version 4.1," the ATM Forum does an excellent job of categorizing the types of service that a network can offer to support different sorts of applications. In 2004, the ATM Forum joined forces with the MPLS and Frame Relay Alliance to form the MFA Forum. The MFA Forum later merged with the IP/MPLS Forum and in 2008 with the Broadband Forum. Nonetheless, network engineers still refer to the work of the ATM Forum.

mất gói tin cắt xén

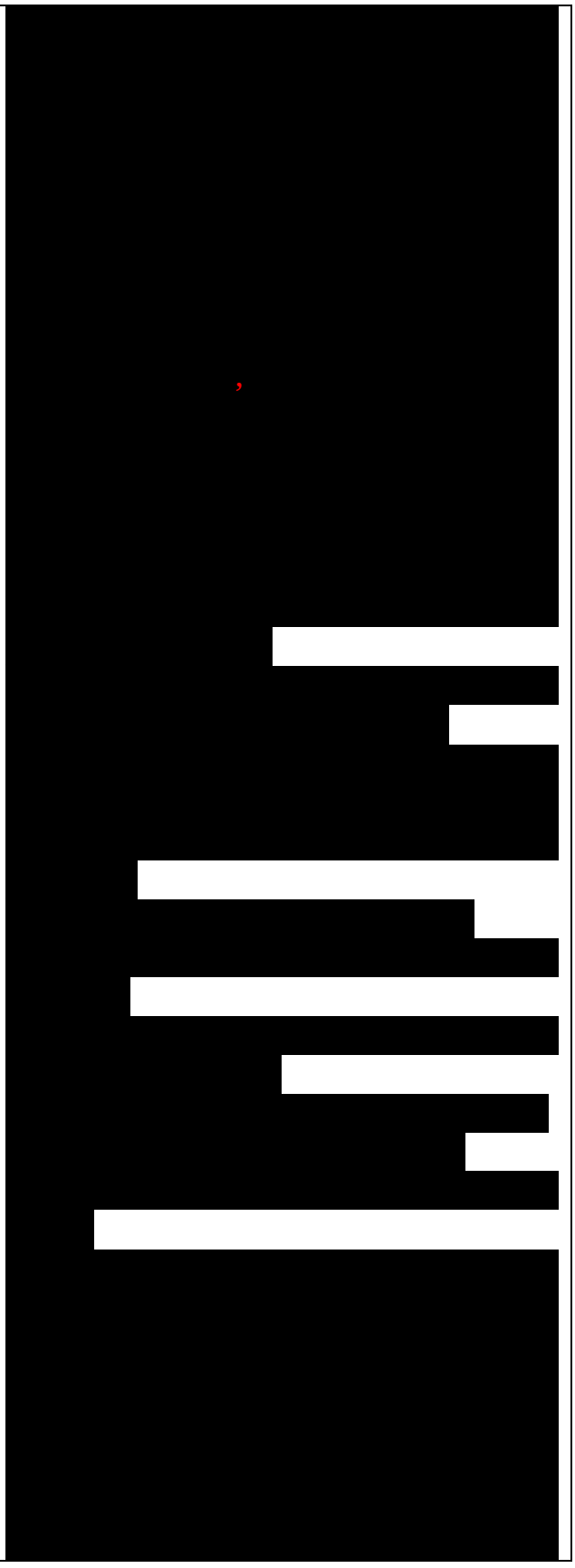
bộ đệm gói tin quản lý hàng đợi

Even if your customer has no plans to use Asynchronous Transfer Mode (ATM) technology, the ATM Forum terminology is still helpful because it identifies the parameters that different sorts of applications must specify to request a certain type of network service. These parameters include delay and delay variation, data-burst sizes, data loss, and peak, sustainable, and minimum traffic rates. Although you might have to replace the word cell with packet in some cases, the ATM Forum definitions can help you classify applications on any network, even non-ATM networks.

The ATM Forum defines six service categories, each of which is described in more detail later in this section:

- Constant bit rate (CBR)
- Real-time variable bit rate (rt-VBR)
- Non-real-time variable bit rate (nrt-VBR)
- Unspecified bit rate (UBR)
- Available bit rate (ABR)
- Guaranteed frame rate (GFR)

For each service category, the ATM Forum specifies a set of parameters to describe both the traffic presented to the network and the QoS required of the network. The ATM Forum also defines traffic control mechanisms that the network can use to meet QoS objectives. The network can implement such



mechanisms as connection admission control and resource allocation differently for each service category.

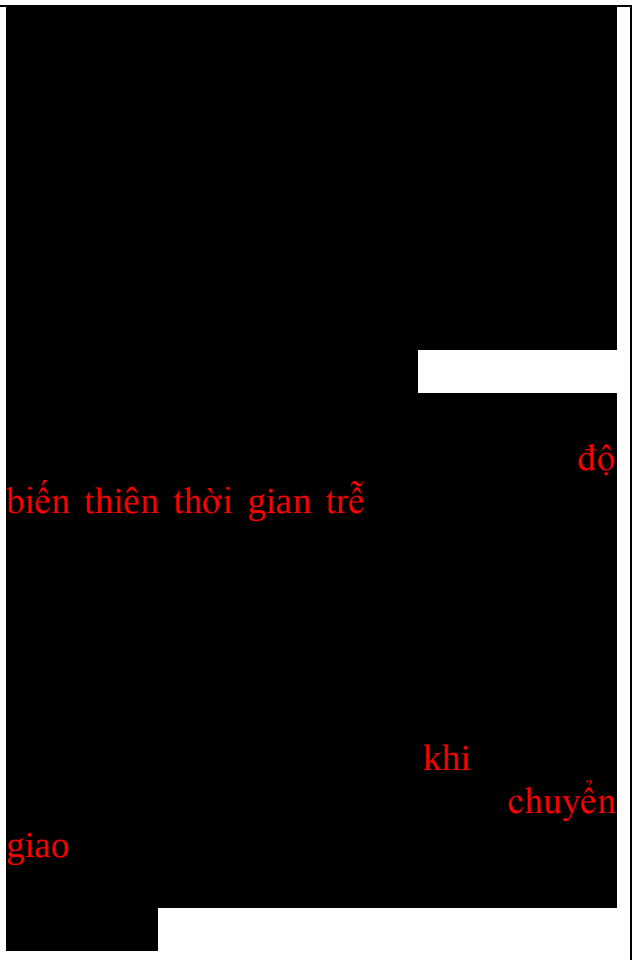
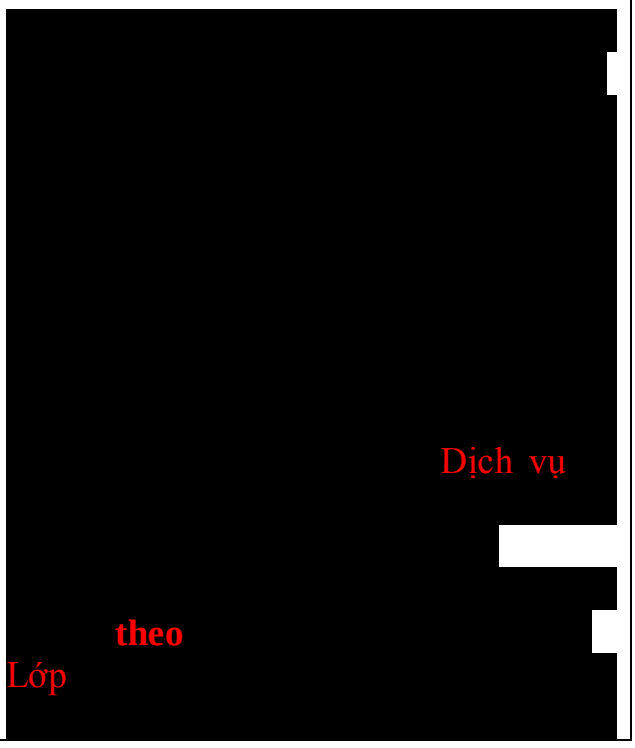
Service categories are distinguished as being either real-time or non-real-time. CBR and rtVBR are real-time service categories. Real-time applications, such as voice and video applications, require tightly constrained delay and delay variation. Non-real-time applications, such as client/server and terminal/host data applications, do not require tightly constrained delay and delay variation. Nrt-VBR, UBR, ABR, and GFR are non-real-time service categories.

It is important to work with your customer to correctly map applications and protocols to the correct service category to meet network performance objectives. A brief overview of ATM service categories is provided here. You can learn more about ATM service categories and traffic management by reading the document "Traffic Management Specification Version 4.1." This document is available from the ATM Forum at http://broadband-forum.org/ftp/pub/approved-specs/af-tm-0121.000.pdf.

Constant Bit Rate Service Category

When CBR is used, a source end system reserves network resources in advance and asks for a guarantee that the negotiated QoS be assured to all cells as long as the cells conform to the relevant conformance specifications. The source can emit cells at the peak cell rate (PCR)

của mạng.

at any time and for any duration and the QoS commitments should pertain. CBR is used by applications that need the capability to request a static amount of bandwidth to be continuously available during a connection lifetime. The amount of bandwidth that a connection requires is specified by the PCR value. CBR service is intended to support real-time applications requiring tightly constrained delay variation (for example, voice, video, and circuit emulation) but is not restricted to these applications. The source may emit cells at or below the negotiated PCR or be silent for periods of time. Cells that are delayed beyond the value specified by the maximum cell transfer delay (maxCTD) parameter are assumed to be of significantly reduced value to the application.	 độ biến thiên thời gian trễ khi chuyển giao
Real-time Variable Bit Rate Service Category rt-VBR connections are characterized in terms of a PCR, sustainable cell rate (SCR), and maximum burst size (MBS). Sources are expected to transmit in a bursty fashion at a rate that varies with time. Cells that are delayed beyond the value specified by maxCTD are assumed to be of significantly reduced value to the application. rt-VBR service may support statistical multiplexing of real-time data sources. Non-real-time Variable Bit Rate Service Category The nrt-VBR service category is intended for non-real-time applications	 Dịch vụ **theo** Lớp

that have bursty traffic characteristics. No delay bounds are associated with this service category. The service is characterized in terms of a PCR, SCR, and MBS. For cells that are transferred within the traffic contract, the application expects a low cell loss ratio (CLR). nrt-VBR service may support statistical multiplexing of connections. Unspecified Bit Rate Service Category UBR service does not specify any traffic-related service guarantees. No numeric commitments are made regarding the cell loss ratio or cell transfer delay (CTD) experienced by a UBR connection. A network might or might not apply PCR to the connection admission control and usage parameter control (UPC) functions. (UPC is defined as the set of actions taken by the network to monitor and control traffic at the end-system access point.)	có lưu lượng dạng cụm của chức năng đầu
Where the network does not enforce PCR, the value of PCR is informational only. (It is still useful to negotiate PCR to allow the source to discover the smallest bandwidth limitation along the path of the connection.) The UBR service category is intended for nonreal-time applications, including traditional computer communications applications such as file transfer and email. With UBR, congestion control can be performed at a higher layer on an end-to-end basis.	giới hạn Lớp dịch vụ điểm đầu và điểm cuối
Available Bit Rate Service Category With ABR, the transfer characteristics	**Lớp**

provided by the network can change subsequent to connection establishment. A flow-control mechanism offers several types of feedback to control the source rate in response to changing ATM-layer conditions. This feedback is conveyed to the source through control cells called resource management (RM) cells. An end system that adapts its traffic in accordance with the feedback should experience a low CLR and obtain a fair share of the available bandwidth according to a network-specific allocation policy. The ABR service does not require bounding the delay or the delay variation experienced by a given connection. ABR service is not intended to support realtime applications. On the establishment of an ABR connection, an end system specifies to the network both a maximum required bandwidth and a minimum usable bandwidth. These are designated as the peak cell rate (PCR) and the minimum cell rate (MCR), respectively. The MCR can be specified as zero. The bandwidth available from the network can vary, but not become less than the MCR. Guaranteed Frame Rate Service Category The GFR service category was added to the Traffic Management Specification in 1999, after the other categories, which were defined in 1996. GFR is designed for applications that require a minimum rate guarantee and can benefit from dynamically accessing additional bandwidth available in the network. Devices connecting LANs to an ATM	mạng dòng sự thay đổi đầu trạng thái chia sẻ ngang bằng băng thông sẵn có theo chính sách phân bổ tùy thuộc mạng Dịch vụ đầu mạng dưới dạng cho bằng tận dụng băng thông truy cập động có sẵn trong mạng

network can use GFR to transport multiple TCP/IP connections over a single GFR virtual circuit (VC). GFR does not require adherence to a flow-control protocol. Under congestion conditions, the network attempts to discard complete frames instead of discarding cells without reference to frame boundaries.

On establishment of a GFR connection, an end system specifies a PCR, MCR, MBS, and maximum frame size (MFS). The MCR can be zero. The end system can send cells up to the PCR, but the network only commits to sending cells (in complete unmarked frames) at the MCR. Traffic beyond the MCR and MBS is delivered within the limits of available resources.

IETF Integrated Services Working Group QoS Specifications

In an IP environment, you can use the work that the IETF Integrated Services working group is doing on QoS requirements. In RFC 2205, the working group describes the Resource Reservation Protocol (RSVP). In RFC 2208, the working group provides information on the applicability of RSVP and some guidelines for deploying it. RFCs 2209 through 2216 are also related to supporting QoS on the Internet and intranets.

RSVP is a setup protocol used by a host to request specific qualities of service from the network for particular application flows. RSVP is also used by

tham chiếu đến biên của khung

đầu

đầu

mạng

về

mạng

routers to deliver QoS requests to other routers (or other types of nodes) along the paths of a flow. RSVP requests generally result in resources being reserved in each node along the path.

RSVP implements QoS for a particular data flow using mechanisms collectively called traffic control. These mechanisms include the following:

- A packet classifier that determines the QoS class (and perhaps the route) for each packet
- An admission control function that determines whether the node has sufficient available resources to supply the requested QoS
- A packet scheduler that determines when particular packets are forwarded to meet QoS requirements of a flow

RSVP works with mechanisms at end systems to request services. To ensure that QoS conditions are met, RSVP clients provide the intermediate network nodes with an estimate of the data traffic they will generate. This is done with a traffic specification (TSpec) and a service-request specification (RSpec), as described in RFC 2216.

Note A TSpec is a description of the traffic pattern for which service is being requested. The TSpec forms one side of a “contract” between the data flow and

the service "provider." After a service request is accepted, the service provider agrees to provide a specific QoS as long as the flow's traffic continues to conform to the TSpec.

An RSpec is a specification of the QoS that a flow wants to request from a network element. The contents of an RSpec are specific to a particular service. The RSpec might contain information about bandwidth required for the flow, maximum delay, or packet-loss rates.

RSVP provides a general facility for reserving resources. RSVP does not define the different types of services that applications can request. The Integrated Services working group describes services in RFCs 2210 through 2216. For a complete understanding of the working group's view of how integrated services should be handled on the Internet or an intranet, you should read the RFCs. The sections that follow provide an overview of the two major types of service: controlled-load service and guaranteed service.

Controlled-Load Service

Controlled-load service is defined in RFC 2211 and provides a client data flow with a QoS closely approximating the QoS that same flow would receive on an unloaded network. Admission control is applied to requests to ensure that the requested service is received even when the network is overloaded.

The controlled-load service is intended for applications that are highly sensitive

[redacted]

to overloaded conditions, such as real-time applications. These applications work well on unloaded networks but degrade quickly on overloaded networks. A service, such as the controlled-load service, that mimics unloaded networks serves these types of applications well. Assuming the network is functioning correctly, an application requesting controlled-load service can assume the following: ■ A high percentage of transmitted packets will be successfully delivered by the network to the receiving end nodes. (The percentage of packets not successfully delivered must closely approximate the basic packet-error rate of the transmission medium.) ■ The transit delay experienced by a high percentage of the delivered packets will not greatly exceed the minimum transmit delay experienced by any successfully delivered packet. (This minimum transit delay includes speed-of-light delay plus the fixed processing time in routers and other communications devices along the path.) The controlled-load service does not accept or make use of specific target values for parameters such as delay or loss. Instead, acceptance of a request for controlled-load service implies a commitment by the network node to provide the requester with service closely equivalent to that provided to uncontrolled (best-effort) traffic under lightly loaded conditions.	giảm kiểm soát tải kiểm soát tải Số gói tin truyền đi được mạng phân phối thành công đến các nút nhận đầu cuối cao Tỷ lệ gói tin không được truyền thành công tỷ số gói tin-lỗi môi trường của đa số gói tin gửi đi của bất kỳ gói tin được gửi thành công nào dọc theo kiểm soát tải của các tham số như thời gian trễ hoặc tổn hao của kiểm soát tải của nút mạng

A network node that accepts a request for controlled-load service must use admission control functions to ensure that adequate resources are available to handle the requested level of traffic, as defined by the requester's TSpec. Resources include link bandwidth, router or switch port-buffer space, and computational capacity of the packet-forwarding engine. Guaranteed Service RFC 2212 describes the network node behavior required to deliver a service called guaranteed service that guarantees both bandwidth and delay characteristics. Guaranteed service provides a firm limit on end-to-end packet-queuing delays. (By firm, the RFC means that the limit can be proven mathematically.) It does not attempt to minimize jitter and is not concerned about fixed delay, such as transmission delay. (Fixed delay is a property of the chosen path, which is determined by the setup mechanism, such as RSVP.)	kiểm soát tải được xác định qua dung lượng tính toán cơ cấu chuyển tiếp gói tin tính chất cần thiết của nút mạng về thời gian trễ hàng đợi gói tin end-to-end Thông qua từ chắc chắn nhấn mạnh rằng giới hạn đó có thể chứng minh bằng toán học được chọn đại lượng này qua cơ chế thiết lập hệ thống
Guaranteed service guarantees that packets will arrive within the guaranteed delivery time and will not be discarded due to queue overflows, provided the flow's traffic conforms to its TSpec. A series of network nodes that implement RFC 2212 ensures a level of bandwidth that, when used by a regulated flow, produces a delay-bounded service with no queuing loss (assuming no failure of	hàng đợi mức băng thông sử

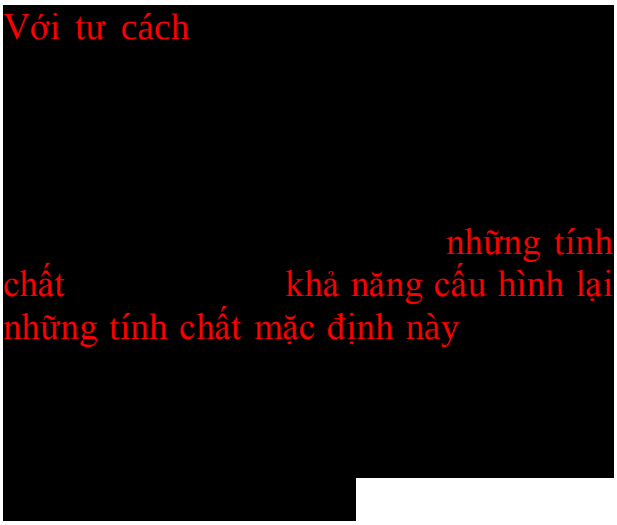
network components or changes in routing during the life of the flow).	trong các thành phần trong vòng đời của lưu lượng
Guaranteed service is intended for applications that need a guarantee that a packet will arrive no later than a certain time after it was transmitted by its source. For example, some audio and video playback applications are intolerant of a packet arriving after its expected playback time. Applications that have real-time requirements can also use guaranteed service.	
In RFC 2212, a flow is described using a token bucket. A token bucket has a bucket rate and a bucket size. The rate specifies the continually sustainable data rate, and the size specifies the extent to which the data rate can exceed the sustainable level for short periods of time.	
The rate is measured in bytes of IP datagrams per second and can range from 1 byte per second to as large as 40 TB per second (the maximum theoretical bandwidth of a single strand of fiber). The bucket size can range from 1 byte to 250 GB. The range of values is intentionally large to allow for future bandwidths. The range is not intended to imply that a network node has to support the entire range.	nằm trong khoảng cực đại lý thuyết nằm trong khoảng Khoảng giá trị này đủ lớn để thích ứng với băng thông trong tương lai
The expectation of the Integrated Services working group is that a software developer can use the relevant RFCs to develop intelligent applications	

that can accurately set the bucket rate and size. An application usually can accurately estimate the expected queuing delay the guaranteed service will provide. If the delay is larger than expected, the application can modify its token bucket to achieve a lower delay.

As a network designer, you won't generally be called upon to estimate token-bucket rates and sizes. On the other hand, you should recognize which applications need guaranteed service, and you should have some idea of their default behavior and whether a reconfiguration of the default behavior is possible. If an application can request terabytes-per-sec- ond bandwidth, you need to know this because of the negative effect it could have on other applications.

IETF Differentiated Services Working Group QoS Specifications

The IETF also has a Differentiated Services working group that works on QoS-related specifications. RFC 2475, "An Architecture for Differentiated Services," defines an architecture for implementing scalable service differentiation in an internetwork or the Internet. As Chapter 13, "Optimizing Your Network Design," covers in more detail, IP packets can be marked with a differentiated services codepoint (DSCP) to influence queuing and packet-dropping decisions for IP datagrams on an output interface of a router. RFC 2475 refers to these decisions as per-hop behaviors (PHB). The DSCP can have 1 of 64 possible



Đặc tả kỹ thuật QoS Bộ phận dịch vụ khác biệt IETF

IETF cũng có thêm một bộ phận dịch vụ khác biệt nghiên cứu các tiêu chuẩn QoS. RFC2475, " Kiến trúc cho những dịch vụ khác biệt", định nghĩa một cấu trúc để thực thi biệt hóa dịch vụ khả mở rộng . Như chương 13 - "Tối ưu hóa thiết kế mạng của bạn" sẽ nói chi tiết hơn, các gói IP có thể được đánh dấu bởi một điểm mã dịch vụ khác biệt (DSCP) để tác động đến việc xếp hàng và quyết định bỏ gói tin đối với các bó dữ liệu IP trên giao diện đầu ra của bộ định tuyến. RFC 2475 đề cập đến những quyết định này dưới dạng các đặc trưng per-hop PHB. DSCP có thể có một trong số 64 giá trị khả dĩ, mỗi giá trị có một PHB dù cho trong mạng thực

values, each of which outlines a PHB, although on a real network you would only use at most 6 to 8 DSCP values.	tế chỉ có tối đa 6 đến 8 giá trị DSCP.
Although the integrated services (RSVP) model, described in the previous section, offers finer granularity, it is less scalable than the differentiated service model. The integrated services model allows sources and receivers to exchange signaling messages that establish packet classification and forwarding state on each router along the path between them. State information at each router can be potentially large. The amount of information grows in proportion to the number of concurrent reservations, which can be a high number on high-capacity backbone links. Differentiated services doesn't require RSVP and can be utilized to aggregate integrated services/RSVP state in the core of a network.	Dù cho mô hình dịch vụ tích hợp (RSVP) (đã được miêu tả ở phần trước) có mức độ chi tiết cao, nó vẫn ít có khả năng thay đổi hơn là mô hình dịch vụ khác biệt. Các mô hình dịch vụ tích hợp cho phép bên nguồn phát và bên nhận trao đổi những tin nhắn báo hiệu nhằm mục đích phân loại gói tin và chuyển tiếp trạng thái trên mỗi bộ định tuyến dọc theo những đường dẫn giữa chúng. Thông tin trạng thái của mỗi bộ định tuyến có thể rất lớn. Lượng thông tin tăng tỷ lệ theo số lượng dự trữ hiện thời và nó có thể rất lớn trên những kết nối backbone (xương sống, đường trục) dung lượng cao. Những dịch vụ khác biệt không yêu cầu RSVP và có thể được sử dụng để tổng hợp những dịch vụ tích hợp/trạng thái RSVP trong lõi của mạng .
RFC 2475 compares its approach to the relative priority-marking model used by such QoS solutions as the IPv4 precedence marking defined in RFC 791, IEEE 802.5 Token Ring priority, and IEEE 802.1p traffic classes. Compared to those solutions, the differentiated services architecture more clearly specifies the role and importance of boundary nodes and traffic conditioners, and uses a per-hop behavior model that permits more general forwarding behaviors than a relative priority. An example of relative priority is IPv4 precedence, which can range from routine (the bits are set to 000) to high (the bits are set to 111).	RFC 2475 so sánh cách tiếp cận của nó với mô hình đánh dấu ưu tiên, một mô hình thường được sử dụng bởi các giải pháp QoS như đánh dấu ưu tiên IPv4 **đã** định nghĩa trong RFC 791 , ưu tiên Token Ring IEEE 802.5, và các lớp lưu lượng 802.1p IEEE. So với những giải pháp đó, kiến trúc dịch vụ phân biệt quy định **cụ thể hơn** vai trò và mức độ quan trọng của các nút biên và bộ điều hòa lưu lượng **,** và **nó** sử dụng mô hình ứng xử per-hop cho phép chuyển tiếp hành vi tổng quát hơn so với ưu tiên tương đối. Một ví dụ về ưu tiên tương đối là ưu tiên IPv4, có thể dao động từ thông thường (các bit được đặt là 000) đến cao (các bit được đặt là 111) .

RFC 2475 also compares its approach to the service-marking model in the IPv4 Type of Service (ToS) bits. As defined in RFC 1349, applications can use the ToS bits to mark each packet with a request for a type of service (for example, a request to minimize delay, maximize throughput, maximize reliability, or minimize cost). The intent of those bits, which were never used, was to allow a router to select routing paths or forwarding behaviors that were suitably engineered to satisfy the service request. The differentiated services model, on the other hand, does not describe the use of the DSCP field as an input to route selection. The ToS markings defined in RFC 1349 are generic and do not match the actual services that routers and service providers offer. Furthermore, the service request is associated with each individual packet, whereas some service semantics may depend on the aggregate forwarding behavior of a sequence of packets. The ToS marking model does not easily accommodate growth in the number and range of future services (because the codepoint space is small) and involves configuration of the forwarding behavior for each ToS in each core network node. The differentiated services model does not have these problems. Grade of Service Requirements for Voice Applications In a voice network, in addition to the need for QoS to ensure low and nonvariable delay and low packet loss,	RFC 2475 cũng so sánh cách tiếp cận của nó với mô hình đánh dấu dịch vụ trong các bit thuộc loại dịch vụ IPv4 (ToS). Theo như định nghĩa trong RFC 1349, các ứng dụng có thể sử dụng các bit ToS để đánh dấu mỗi gói tin cùng với yêu cầu về loại dịch vụ (ví dụ , yêu cầu cực tiểu hóa thời gian trễ, tối đa hóa thương lượng, tối đa hóa độ tin cậy, hoặc giảm thiểu chi phí). Mục đích của những bit này, những bit chưa tường được sử dụng, là cho phép bộ định tuyến chọn các đường định tuyến hoặc chuyển tiếp hành vi được thiết kế phù hợp để đáp ứng các yêu cầu dịch vụ. Mặt khác, mô hình dịch vụ phân biệt không mô tả việc sử dụng trường DSCP như một đầu vào để lựa chọn tuyến. Những dấu ToS định nghĩa trong RFC 1349 có tính chất chung chung và không phù hợp với các dịch vụ thực tế mà các bộ định tuyến và nhà phân phối dịch vụ cung cấp. Hơn nữa, yêu cầu dịch vụ gắn với từng gói tin riêng biệt, trong khi một số ngữ nghĩa dịch vụ có thể phụ thuộc vào đặc tính chuyển tiếp tổng hợp của một loạt gói tin. Mô hình đánh dấu ToS không dễ dàng thích nghi với sự tăng trưởng số lượng và phạm vi của các dịch vụ trong tương lai (vì không gian điểm mã nhỏ) và cần **phải thiết lập** hành vi chuyển tiếp cho mỗi ToS trong mỗi nút mạng lõi . Các mô hình dịch vụ **phân** biệt không có những vấn đề này. Phân cấp các yêu cầu dịch vụ cho các ứng dụng thoại Trong một mạng thoại, cùng với nhu cầu **về** QoS để đảm bảo độ trễ thấp và không đổi cũng như sự tổn hao gói tin

there is also a need for what voice experts call a high grade of service (GoS). GoS refers to the fraction of calls that are successfully completed in a timely fashion. Call completion rate (CCR) is another name for the requirement.	thấp, còn có một yêu cầu khác mà các chuyên gia thoại gọi là phẩm chất cao của dịch vụ (GOS) . GOS đề cập đến các phần của cuộc gọi được thực hiện thành công một cách kịp thời . Tỷ lệ hoàn thành cuộc gọi (CCR) là một tên khác của yêu cầu này.
A network must have high availability to support a high GoS. In an unreliable network, GoS is adversely affected when call setup and teardown messages are lost. A lost signal for call setup can result in an unsuccessful call attempt. A lost signal for call teardown can cause voice resources to be unavailable for other calls. Call setup and teardown messages aren't as delay-sensitive as the audio sent during the actual call, so retransmission of these messages is permitted but should generally be avoided to avoid impacting users. (The voice packets themselves are not retransmitted because there's no point. The voice wouldn't sound right if the packets arrived later than they were supposed to arrive.)	Một mạng phải có khả năng sẵn sàng hoạt động cao để cho ra GOS cao . Trong một mạng không đáng tin cậy , GOS **bị sụt giảm nếu** các tin nhắn thiết lập cuộc gọi và teardown bị mất. Một tín hiệu thiết lập cuộc gọi mất có thể làm cho cuộc gọi không thành công. Một tín hiệu để teardown cuộc gọi bị mất có thể làm cho các cuộc gọi khác không thể sử dụng được tài nguyên thoại. Các tin nhắn thiết lập và teardown cuộc gọi không nhạy với thời gian trì hoãn như âm thanh được gửi trong một cuộc gọi thực sự , vì vậy việc truyền lại những tin nhắn này là ~~**được phép**~~ **khả thi** nhưng nói chung nên tránh để không ảnh hưởng đến người dùng. (Chính gói tin thoại không được truyền lại vì chẳng có ích lợi gì. Giọng nói sẽ không chính xác nếu các gói tin đến trễ hơn thời gian dự kiến.)
To achieve high GoS, you should follow the recommendations that will be presented in subsequent chapters to use reliable components (cables, patch panels, switches, routers, and so on) and to build redundancy and failover into the network using such techniques as dynamic routing, the Spanning Tree Protocol (STP) for switched networks, Hot Standby Router Protocol (HSRP), and so on. As discussed in Chapter 9, "Developing Network Management	Để đạt được GOS cao, bạn nên làm theo những khuyến cáo được trình bày trong chương tiếp theo để sử dụng thành phần đáng tin cậy (cáp, bảng cắm, chuyển mạch, bộ định tuyến, v.v…) và xây dựng khả năng dự phòng và chuyển đổi dự phòng vào mạng bằng cách sử dụng các kỹ thuật như định tuyến động , Giao Thức Bắc Cầu Dạng Cây (STP) cho các mạng chuyển mạch, Giao Thức Bộ Định Tuyến Dự Phòng Nóng (HSRP) , và v.v….. Như sẽ trình bày trong Chương 9

Strategies," achieving a high GoS also requires implementing a network management strategy that will quickly alert you to network outages and degraded service. Documenting QoS Requirements You should work with your customer to classify each network application in a service category. When you have classified the application, you should fill in the QoS Requirements column in Table 4-4. If your customer has applications that can be characterized as needing controlled-load or guaranteed service, you can use those terms when filling in the QoS Requirements column. If your customer plans to use ATM, you can use the ATM Forum's terminology for service categories. Even if your customer does not plan to use ATM or IETF QoS, you can still use the ATM Forum or Integrated Services working group terminology. Another alternative is to simply use the following terms: ■ Inflexible: A generic term to describe any application that has specific requirements for constant bandwidth, delay, delay variation, accuracy, and throughput. ■ Flexible: A generic term to describe any application that simply expects the network to make a best effort to meet requirements. Many nonmultimedia applications have flexible QoS requirements. For voice applications, you should make more than one entry in Table 4-4 due to	, "Xây dựng chiến lược quản lý mạng , " việc đạt được GOS cao cũng đòi hỏi phải triển khai **một** chiến lược quản lý mạng **có thể** cảnh báo nhanh tình trạng rớt mạng hoặc dịch vụ **xuống cấp**. Ghi nhận các yêu cầu QoS Bạn nên làm việc với khách hàng để phân loại ứng dụng mạng theo loại dịch vụ. Khi bạn đã phân loại ứng dụng, bạn cần điền vào cột các yêu cầu QoS trong bảng 4-4. Nếu khách hàng của bạn có các ứng dụng thuộc loại dịch vụ đảm bảo hoặc kiểm soát tải, bạn có thể sử dụng những thuật ngữ đó khi điền vào cột Các yêu cầu QoS. Nếu khách hàng của bạn có kế hoạch dùng ATM , bạn có thể sử dụng thuật ngữ của Diễn đàn ATM để **phân loại** các dịch vụ. Ngay cả khi khách hàng của bạn không có ý định sử dụng ATM hoặc IETF QoS, bạn vẫn có thể sử dụng các thuật ngữ của diễn đàn ATM hoặc **của** nhóm làm việc dịch vụ tích hợp. Hoặc bạn có thể sử dụng các thuật ngữ sau đây : - Sự không linh hoạt: một thuật ngữ chung để mô tả bất kì ứng nào dụng có những yêu cầu cụ thể về băng thông, thời gian trễ, sự biến độngthời gian trễ, độ chính xác vàthông lượng xác định. - Sự linh hoạt: một thuật ngữ chung để mô tả bất kì ứng dụng nào chỉ đơn giản mong chờ mạng cố hết sức để đáp ứng các yêu cầu. Rất nhiều ứng dụng không thuộc về mảng đa phương tiện đều có những yêu cầu linh hoạt về QoS. - Về những ứng dụng thoại, bạn nên tạo ra nhiều hơn một cổng vào trong

the different requirements for the call control flow and the audio stream. The call control flow, used for setting up and tearing down calls, doesn't have strict delay constraints, but it does require high network availability and there may be a GoS requirement that should be specified. For the voice stream, the QoS classification should be listed using the ATM term CBR or the IETF term guaranteed service.	bảng 4-4 do những yêu cầu khác nhau về luồng điều khiển cuộc gọi và dòng âm thanh. Các luồng điều khiển cuộc gọi thường được dùng để thiết lập và tách các cuộc gọi không có những hạn chế nghiêm ngặt về thời gian trễ, nhưng nó yêu cầu một mạng lưới mang tính khả dụng cao và có thể chúng ta nên chỉ rõ yêu cầu QoS. Đối với các dòng âm thanh, sự phân chia QoS nên được liệt kê ra bằng cách sử dụng thuật ngữ của ATM là CBR hay thuật ngữ IETF về dịch vụđảm bảo.
When documenting QoS requirements for applications, it's also a good idea to let your clients know that QoS is an end-to-end proposition. Such issues as mapping LAN-based QoS (for example, the 802.1p bits) to IP DSCP and Multiprotocol Label Switching (MPLS) experimental (EXP) bits should be part of the QoS requirements analysis discussion. The detailed discussion of QoS solutions will happen later in the design cycle as covered in Chapter 13.	- Khi ghi nhận những yêu cầu QoS cho các ứng dụng thì việc cho các khách hàng của bạn biết rằng QoS là một mệnh đề end-to-endcũng là một ý kiến hay. Những vấn đề như ánh xạ QoS dựa trên mạng LAN (ví dụ như 802.1p bits) đến IP DSCP và các bit experimental (EXP) Công Nghệ Chuyển Mạch Nhãn Đa Giao Thức (MPLS) nên là một phần trong việc thảo luận về phân tích yêu cầu của QoS. Những thảo luận chi tiết về những giải pháp QoS sẽ được đưa ra sau trong chu trình thiết kế như đã đề cập trong chương 13.
Network Traffic Checklist You can use the following Network Traffic checklist to determine if you have completed all the steps for characterizing network traffic:	-Danh sách kiểm tra lưu lượng mạng Bạn có thể dùng danh sách kiểm tra lưu lượng mạng sau để xác định rằng liệu bạn đã hoàn thành tất cả các bước để định rõ đặc điểm của lưu lượng mạng hay chưa.
□ I have identified major traffic sources and stores and documented traffic flow between them. □ I have categorized the traffic flow for each application as being terminal/host, client/server, peer-to-	+ Tôi đã xác địnhđược các nguồn lưu lượng mạng, những nơi lưu trữ và các dòng lưu lượng được ghi nhận giữa chúng. Tôi đã phân loại các dòng lưu lượng cho mỗi ứng dụng dưới dạng đầu cuối/máy

peer, server/server, or distributed computing. □ I have estimated the bandwidth requirements for each application. □ I have estimated the bandwidth requirements for routing protocols. □ I have characterized network traffic in terms of broadcast/multicast rates, efficiency, frame sizes, windowing and flow control, and error-recovery mechanisms. □ I have categorized the QoS requirements of each application. □ I have discussed the challenges associated with implementing end-to-end QoS and the need for devices across the network to do their part in implementing QoS strategies. Summary This chapter provided techniques for analyzing network traffic caused by applications and protocols. The chapter discussed methods for identifying traffic sources and data stores, measuring traffic flow and load, documenting application and protocol usage, and evaluating QoS requirements.	chủ, máy khách/máy chủ, máy chủ/máy chủ hay tính toán phân tán. Tôi đã tính toán những yêu cầu về băng thông cho mỗi ứng dụng Tôi đã tính toán những yêu cầu về băng thông cho giao thức định tuyến Tôi đã mô tả đặc điểm của lưu lượng mạng theo tỷ lệ broadcast/multicast, hiệu suất, kích thước khung, sự điều khiển lưu lượng hay phân chia cửa sổ, và cơ chế khắc phục lỗi. Tôi đã phân loại những yêu cầu về QoS cho mỗi ứng dụng Tôi đã thảo luận về những thách thức liên quan đến việc thực hiện QoS end-to-end và nhu cầu về những thiết bị đi xuyên qua mạngđể thực hiện nhiệm vụ của chúng trong việc thực hiện các chiến lược QoS. Tóm tắt Chương này cung cấp những công cụ cho việc phân tích lưu lượng mạng của các ứng dụng và các giao thức. Chương này bàn luận về những phương pháp để xác định nguồn lưu lượng và lưu trữ dữ liệu, tính toán các dòng lưu lượng và tải, cung cấp dữ liệu về việc sử dụng ứng dụng và giao thức, và đánh giá những yêu cầu QoS.

Summary for Part I checked At this point in the network design process, you have identified a customer's network applications and the technical requirements for a network design that can support the applications. You should take another look at Table 4-4, "Network Applications Traffic Characteristics," and Table 2-2, "Network Applications Technical Requirements," to make sure you understand your customer's application requirements. If you want, you can merge these two tables so that there is one row for each application. A top-down methodology for network design focuses on applications. Chapter 1 covered identifying applications and business goals. Chapter 2 analyzed technical goals for applications and the network as a whole, such as availability, performance, and manageability. Chapter 3 concentrated on techniques for characterizing the existing network, and Chapter 4 refocused on technical requirements in terms of the network traffic characteristics of applications and protocols.	[redacted]

This summary wraps up Part I, "Identifying Your Customer's Needs and Goals," which presented the requirements-analysis phase of network design. The requirements-analysis phase is the most important phase in top-down network design. Gaining a solid understanding of your customer's requirements helps you select technologies that meet a cus-tomer's criteria for success.

You should now be able to analyze a customer's business and technical goals and be ready to start developing a logical and physical network design. Part II, "Logical Network Design," covers designing a logical network topology, developing a network layer addressing and naming model, selecting switching and routing protocols, and planning network security and management strategies.

Logical Network Design

Chapter 5 Designing a Network Topology

Chapter 6 Designing Models for Addressing and Numbering

Chapter 7 Selecting Switching and Routing

Protocols

Chapter 8 Developing Network Security Strategies

Chapter 9 Developing Network Management Strategies

This page intentionally left blank

Chapter 5
Designing a Network Topology

In this chapter, you will learn techniques for developing a network topology. A topology is a map of an internetwork that indicates network segments, interconnection points, and user communities. Although geographical sites can appear on the map, the purpose of the map is to show the geometry of the network, not the physical geography or technical implementation. The map is a high-level blueprint of the network, analogous to an architectural drawing that shows the location and size of rooms for a building, but not the construction materials for fabricating the rooms.

Designing a network topology is the first step in the logical design phase of the top- down network design methodology. To meet a customer's goals for scalability and adaptability, it

is important to architect a logical topology before selecting physical products or technologies. During the topology design phase, you identify networks and interconnection points, the size and scope of networks, and the types of internetworking devices that will be required, but not the actual devices.

This chapter provides tips for both campus and enterprise WAN network design and focuses on hierarchical network design, which is a technique for designing scalable campus and WAN networks using a layered, modular model. In addition to covering hierarchical network design, the chapter also covers redundant network design topologies and topologies that meet security goals. (Security is covered in more detail in Chapter 8, "Developing Network Security Strategies.") This chapter also discusses the Cisco SAFE security reference architecture.

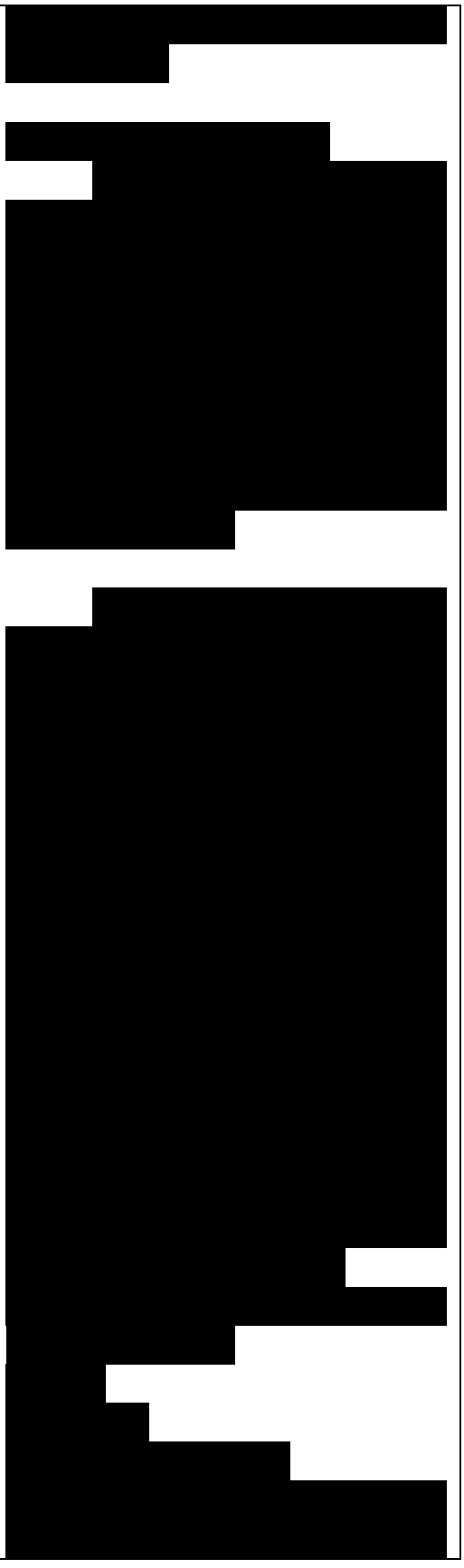
Upon completion of this chapter, you will know more about designing secure, redundant, hierarchical, and modularized topologies. A topology diagram is a useful tool to help you and your customer begin the process of moving from a logical design

to a physical implementation of the customer's network environment.

Hierarchical Network Design

To meet a customer's business and technical goals for a corporate network design, you might need to recommend a network topology consisting of many interrelated components. This task is made easier if you can "divide and conquer" the job and develop the design in layers.

Network design experts have developed the hierarchical network design model to help you develop a topology in discrete layers. Each layer can be focused on specific functions, allowing you to choose the right systems and features for the layer. For example, in Figure 5-1, high-speed WAN routers can carry traffic across the enterprise WAN backbone, medium-speed routers can connect buildings at each campus, and switches can connect user devices and servers within buildings.



Core Layer

Campus A

Campus B Campus C

■ A core layer of high-end routers and switches that are

optimized for availability and performance.

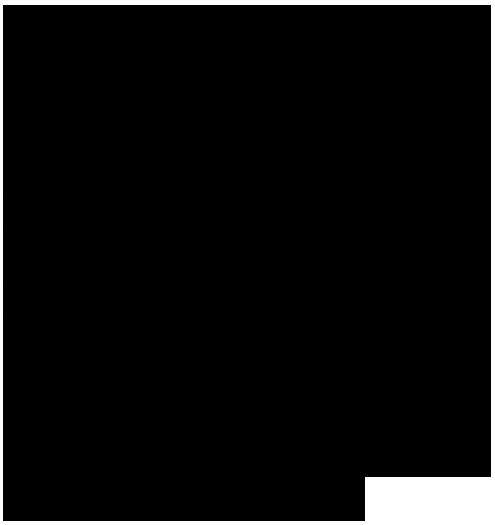
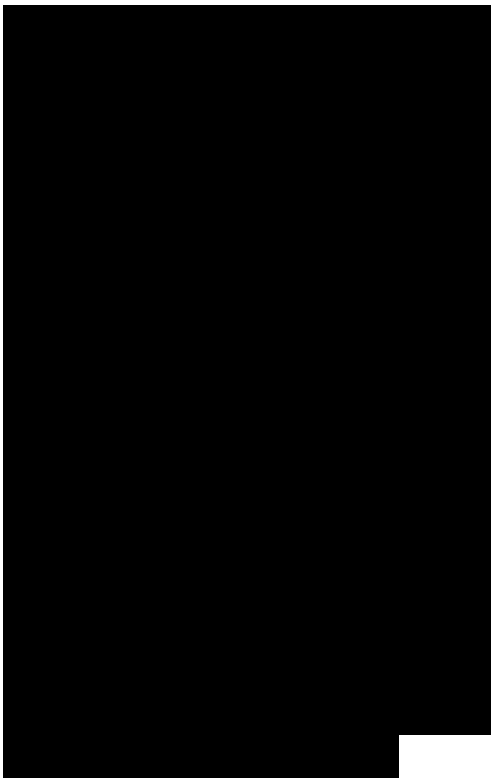
■ A distribution layer of routers and switches that implement policies. In small and medium-sized organizations, the core and distribution layers can be combined.

■ An access layer that connects users via lower-end switches and wireless access points.

Why Use a Hierarchical Network Design Model?

Networks that grow unheeded without any plan in place tend to develop in an unstructured format. Dr. Peter Welcher, the author of network design and technology articles for Cisco World and other publications, refers to unplanned networks as fur-ball networks.

Welcher explains the disadvantages of a fur-ball topology by pointing out the problems that too many CPU adjacencies cause. When network devices communicate with many other devices, the workload required of the CPUs on the devices can be burdensome. For example, in a large flat (switched) network, broadcast packets are burdensome. A broadcast packet interrupts the CPU on

each device within the broadcast domain and demands processing time on every device (including routers, workstations, and servers) for which a protocol understanding for that broadcast is installed. Another potential problem with nonhierarchical networks, besides broadcast packets, is the CPU workload required for routers to communicate with many other routers and process numerous route advertisements. A hierarchical network design methodology enables you to design a modular topology that limits the number of communicating routers. Using a hierarchical model can help you minimize costs. You can purchase the appropriate internetworking devices for each layer of the hierarchy, thus avoiding spending money on unnecessary features for a layer. Also, the modular nature of the hierarchical design model enables accurate capacity planning within each layer of the hierarchy, thus reducing wasted bandwidth. Network management responsibility and network management systems can be distributed to the different layers of a modular network architecture to control management costs.	  

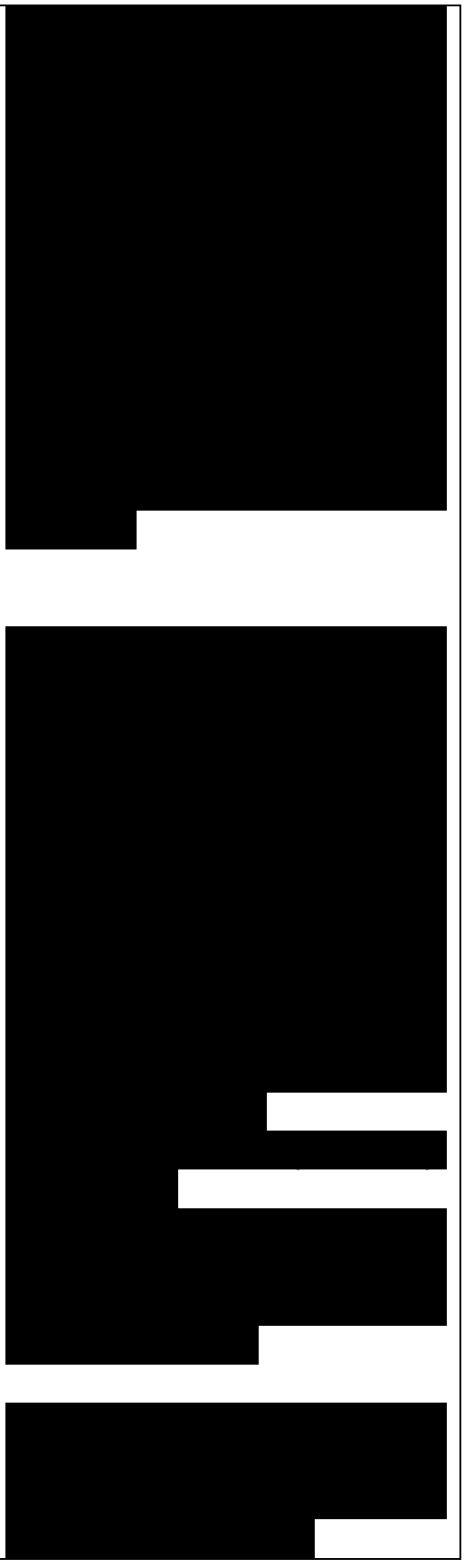
Modularity enables you to keep each design element simple and easy to understand. Simplicity minimizes the need for extensive training for network operations personnel and expedites the implementation of a design. Testing a network design is made easy because there is clear functionality at each layer. Fault isolation is improved because network technicians can easily recognize the transition points in the network to help them isolate possible failure points.

Hierarchical design facilitates changes. As elements in a network require change, the cost of making an upgrade is contained to a small subset of the overall network. In large flat or meshed network architectures, changes tend to impact a large number of systems. Replacing one device can affect numerous networks because of the complex interconnections.

How Can You Tell When You Have a Good Design?

Here are some wise answers from Peter Welcher that are based on the tenets of hierarchical, modular network design:

- When you already know how to add a new building, floor, WAN link, remote site, e-commerce service, and so on



■ When new additions cause only local change to the directly connected devices ■ When your network can double or triple in size without major design changes \ ■ When troubleshooting is easy because there are no complex protocol interactions to wrap your brain around When scalability is a major goal, a hierarchical topology is recommended because modularity in a design enables creating design elements that can be replicated as the network grows. Because each instance of a module is consistent, expansion is easy to plan and implement. For example, planning a campus network for a new site might simply be a matter of replicating an existing campus network design. Today's fast-converging routing protocols were designed for hierarchical topologies. Route summarization, which Chapter 6, "Designing Models for Addressing and Naming," covers in more detail, is facilitated by hierarchical network design. To control routing CPU overhead and	[redacted]

bandwidth consumption, modular hierarchical topologies should be used with such protocols as Open Shortest Path First (OSPF), Intermediate System-to- Intermediate System (IS-IS), Border Gateway Protocol (BGP), and Enhanced Interior Gateway Routing Protocol (Enhanced IGRP).

Flat Versus Hierarchical Topologies

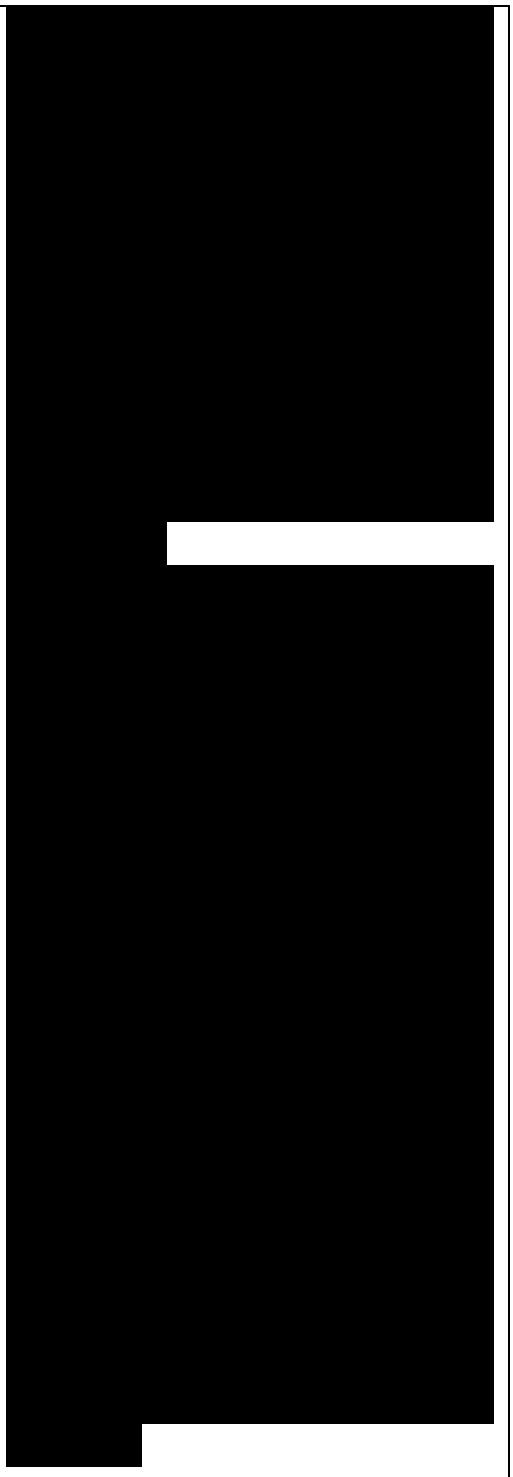
A flat network topology is adequate for small networks. With a flat network design, there is no hierarchy. Each network device has essentially the same job, and the network is not divided into layers or modules. A flat network topology is easy to design and implement, and it is easy to maintain, as long as the network stays small. When the network grows, however, a flat network is undesirable. The lack of hierarchy makes troubleshooting difficult. Rather than being able to concentrate troubleshooting efforts in just one area of the network, you might need to inspect the entire network.

Flat WAN Topologies

A WAN for a small company can consist of a few sites connected in a loop. Each site has a WAN router that connects to two other adjacent

sites via point-to-point links, as shown at the top of Figure 5-2. As long as the WAN is small (a few sites), routing protocols can converge quickly, and communication with any other site can recover when a link fails. As long as only one link fails, communication recovers. When more than one link fails, some sites are isolated from others.

A flat loop topology is generally not recommended for networks with many sites, however. A loop topology can mean that there are many hops between routers on opposite sides of the loop, resulting in significant delay and a higher probability of failure. If your analysis of traffic flow indicates that routers on opposite sides of a loop topology exchange a lot of traffic, you should recommend a hierarchical topology instead of a loop. To avoid any single point of failure, you can place redundant routers or switches at upper layers of the hierarchy, as shown at the bottom of Figure 5-2.



Headquarters in Medford
Klamath Falls Branch Office
Flat Loop Topology
Headquarters in Medford
rants Pass Klamath
Ashland White City

Branch Falls Branch Branch Office Branch Office Office Office

Hierarchical Redundant Topology

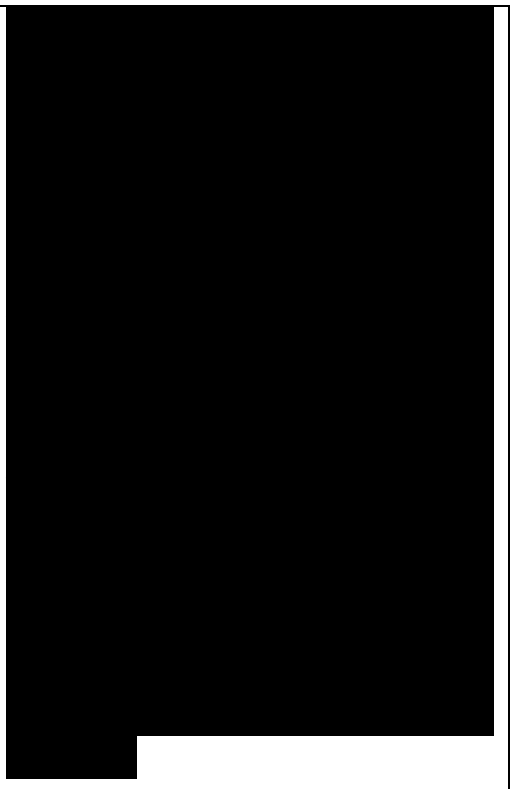
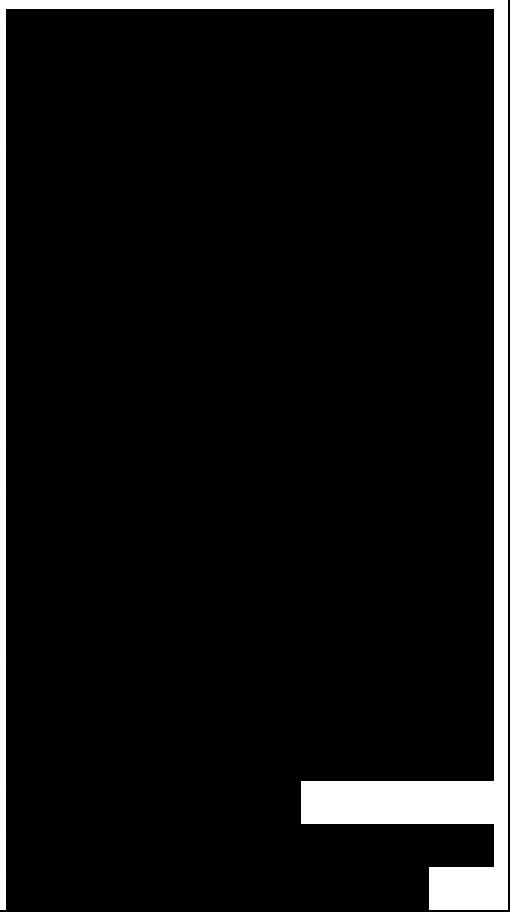
Figure 5-2 Flat Loop Topology (Top) and Hierarchical Redundant Topology (Bottom)

The flat loop topology shown at the top of Figure 5-2 meets goals for low cost and reasonably good availability. The hierarchical redundant topology shown at the bottom of Figure 5-2 meets goals for scalability, high availability, and low delay

Flat LAN Topologies

In the early and mid-1990s, a typical design for a LAN was PCs and servers attached to one or more hubs in a flat topology. The PCs and servers implemented a media-access control process, such as token passing or carrier sense multiple access with collision detection (CSMA/CD) to control access to the shared bandwidth. The devices were all part of the same bandwidth domain and had the capability to negatively affect delay and throughput for other devices.

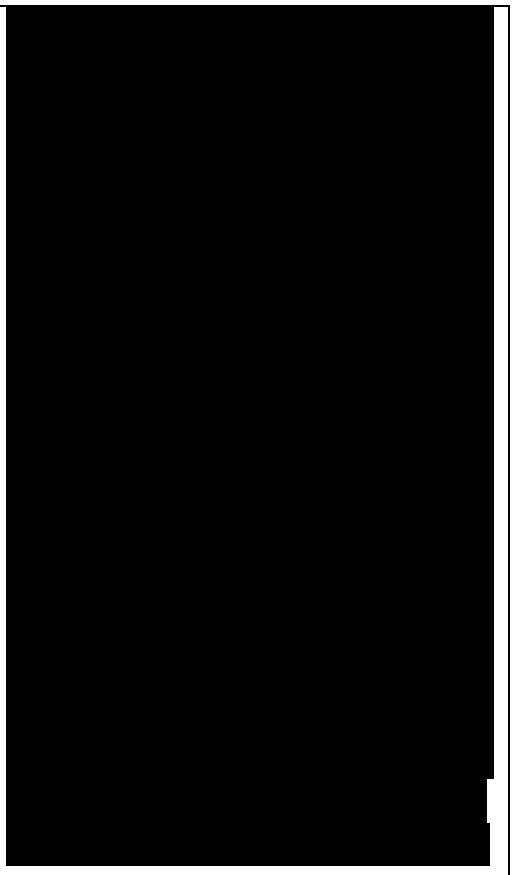
These days, network designers recommend attaching the PCs

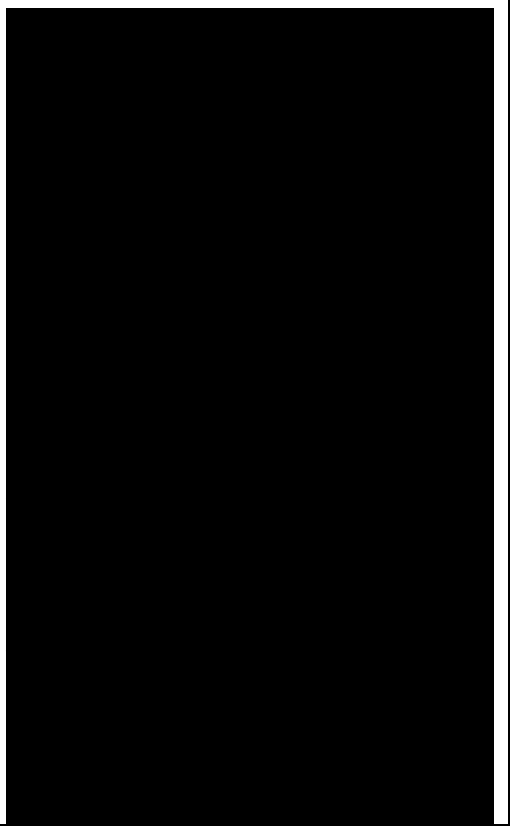
and servers to data link layer (Layer 2) switches instead of hubs. In this case, the network is segmented into small bandwidth domains so that a limited number of devices compete for bandwidth at any one time. The devices do compete for service by the switching hardware and software, however, so it is important to understand the performance characteristics of candidate switches, as discussed in Chapter 10, "Selecting Technologies and Devices for Campus Networks."	
As discussed in Chapter 4, "Characterizing Network Traffic," devices connected in a switched or bridged network are part of the same broadcast domain. Switches forward broadcast frames out all ports. Routers, on the other hand, segment networks into separate broadcast domains. A single broadcast domain should be limited to a few hundred devices so that devices are not overwhelmed by the task of processing broadcast traffic. Introducing hierarchy into a network design by adding routers curtails broadcast radiation.	

With a hierarchical design, you can deploy internetworking devices to do the job they do best. You can add routers to a campus network design to isolate broadcast traffic. You can deploy high-end switches to maximize bandwidth for high-traffic applications, and use low-end switches when simple, inexpensive access is required. Maximizing overall performance by modularizing the tasks required of internetworking devices is one of the many benefits of using a hierarchical design model.

Mesh Versus Hierarchical-Mesh Topologies

Network designers often recommend a mesh topology to meet availability requirements. In a full-mesh topology, every router or switch connects to every other router or switch. A full-mesh network provides complete redundancy and offers good performance because there is just a single-link delay between any two sites. A partial-mesh network has fewer connections. To reach another router or switch in a partial-mesh network might require traversing intermediate links, as shown in Figure 5-3.





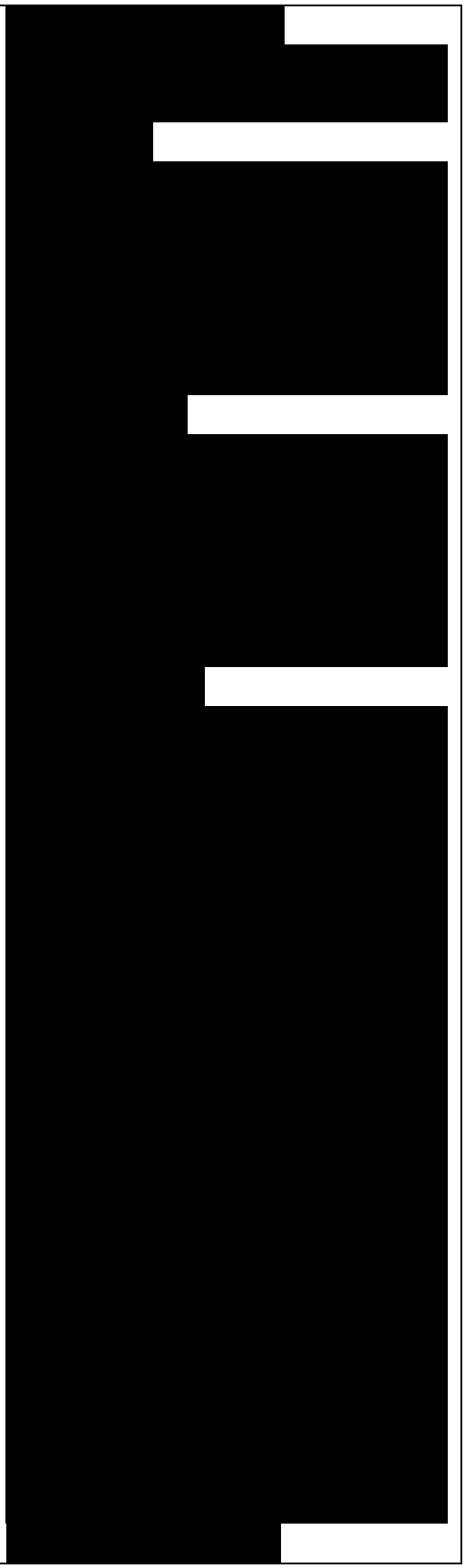


Figure 5-3 Partial-Mesh Topology (Left) and Full-Mesh Topology (Right)

Note In a full-mesh topology, every router or switch is connected to every other router or switch. The number of links in a full-mesh topology is as follows:

N is the number of routers or switches. (Divide the result by 2 to avoid counting Router X to Router Y and Router Y to Router X as two different links.)

Although mesh networks feature good reliability, they have many disadvantages if they are not designed carefully. Mesh networks can be expensive to deploy and maintain. (A full- mesh network is expensive.) Mesh networks can also be hard to optimize, troubleshoot, and upgrade, unless they are designed using a simple, hierarchical model. In a nonhierar- chical mesh topology, internetworking devices are not optimized for specific functions. Containing network problems is difficult because of the lack of modularity. Network upgrades are problematic because it is difficult to upgrade just one

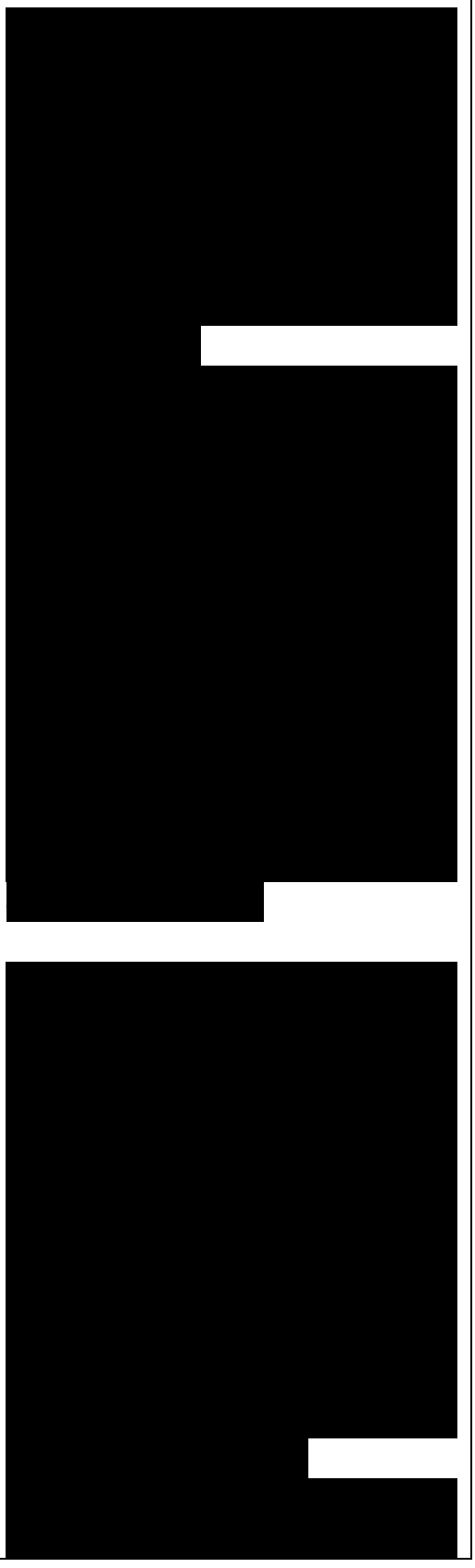
part of a network.

Mesh networks have scalability limits for groups of routers that broadcast routing updates or service advertisements. As the number of router CPU adjacencies increases, the amount of bandwidth and CPU resources devoted to processing updates increases.

A good rule of thumb is that you should keep broadcast traffic at less than 20 percent of the traffic on each link. This rule limits the number of adjacent routers that can exchange routing tables and service advertisements. This limitation is not a problem, however, if you follow guidelines for simple, hierarchical design. A hierarchical design, by its very nature, limits the number of router adjacencies.

With routing protocols, such as OSPF and EIGRP, the problem is not with the broadcast/multicast traffic and CPU resources used for day-to-day routing. The problem is the amount of work and bandwidth required to reestablish routing after an outage. Be careful not to let your network grow into a complicated mesh just because it's still working.

There will probably be an outage someday, and then you



might learn the hard way the downfalls associated with a complex mesh of routers.

Figure 5-4 shows a classic hierarchical and redundant enterprise design. The design uses a partial-mesh hierarchy rather than a full mesh. The figure shows an enterprise routed network, but the topology could also be used for a switched campus network.

For small and medium-sized companies, the hierarchical model is often implemented as a hub-and-spoke topology with little or no meshing. Corporate headquarters or a data center form the hub. Links to remote offices and telecommuters' homes form the spokes, as shown in Figure 5-5.

Classic Three-Layer Hierarchical Model

Literature published by Cisco and other networking vendors talks about a classic three-layer hierarchical model for network design topologies. The three-layer model permits traffic aggregation and filtering at three successive routing or switching levels. This makes the three-layer hierarchical model scalable to large international internetworks.

Branch Offices (Access Layer)

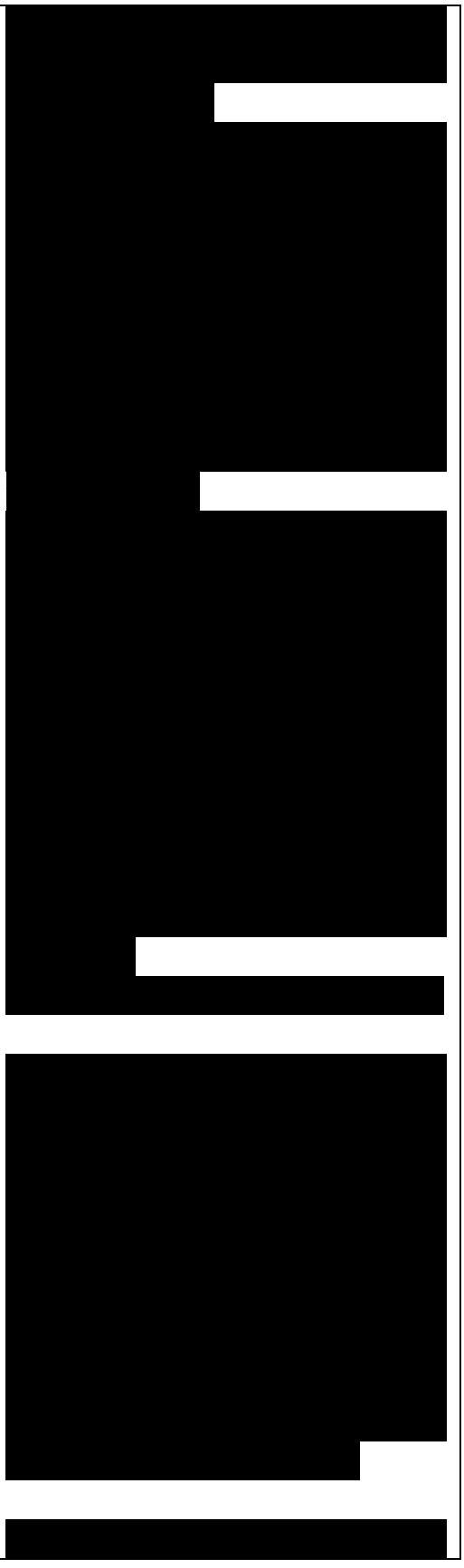


Figure 5-4 Partial-Mesh Hierarchical Design

Branch Office | Home Office | Branch Office

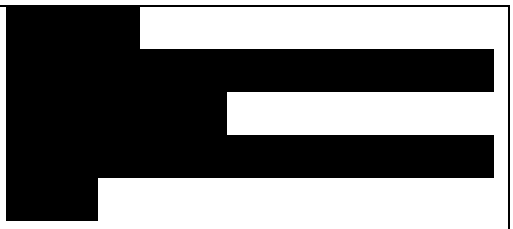


Figure 5-5 Hub-and-Spoke Hierarchical Topology for a Medium-Sized Business

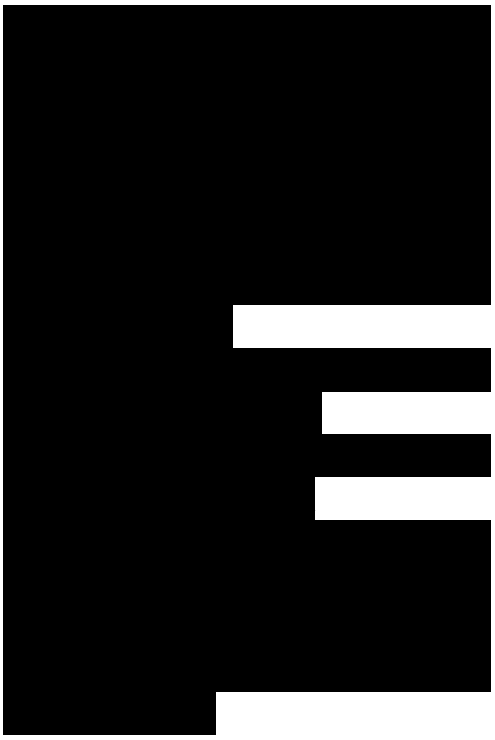
Although the model was developed at a time when routers delineated layers, the model can be used for switched networks and routed networks. Figure 5-1 and Figure 5-4 show three-layer hierarchical topologies.

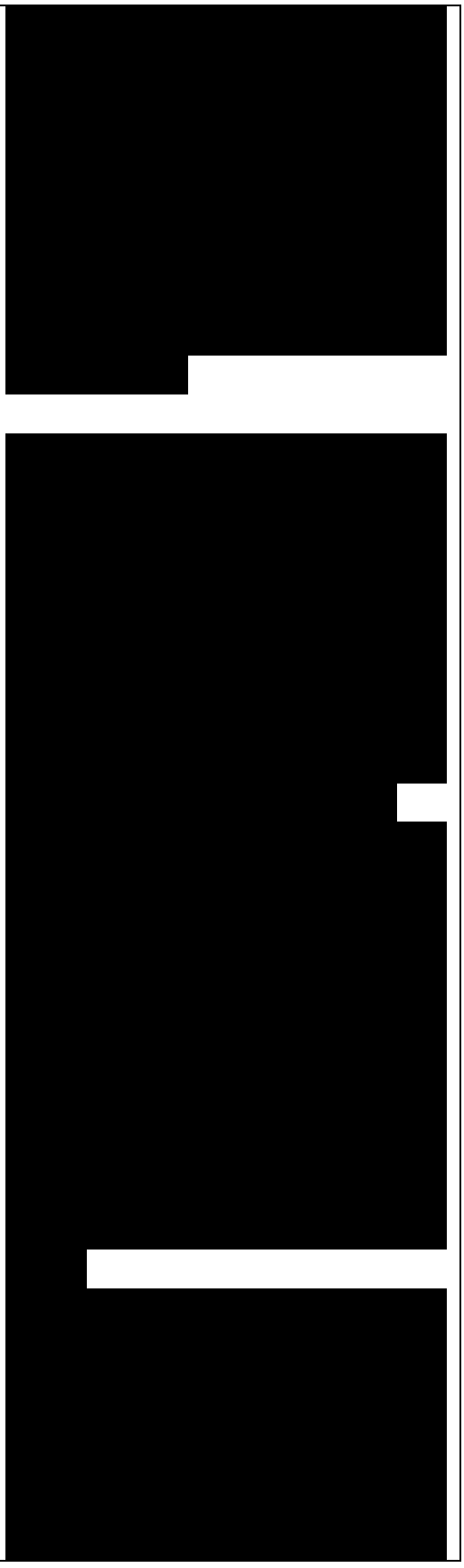
Each layer of the hierarchical model has a specific role:

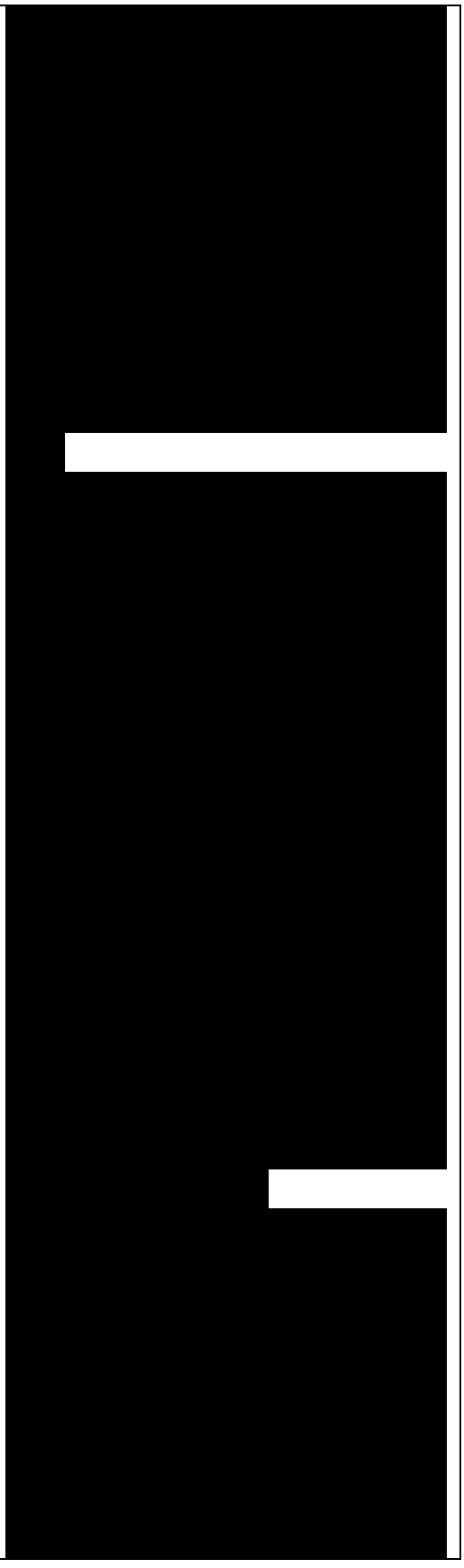
- The core layer provides optimal transport between sites.
- The distribution layer connects network services to the access layer and implements policies regarding security, traffic loading, and routing.
- In a WAN design, the access layer consists of the routers at the edge of the campus networks. In a campus network, the access layer provides switches or hubs for end- user access.

The sections that follow discuss the core, distribution, and access layers in greater detail.

Core Layer



The core layer of a three-layer hierarchical topology is the high-speed backbone of the internetwork. Because the core layer is critical for interconnectivity, you should design the core layer with redundant components. The core layer should be highly reliable and should adapt to changes quickly. When configuring routers in the core layer, you should use routing features that optimize packet throughput. You should avoid using packet filters or other features that slow down the manipulation of packets. You should optimize the core for low latency and good manageability. The core should have a limited and consistent diameter. Distribution layer routers (or switches) and client LANs can be added to the model without increasing the diameter of the core. Limiting the diameter of the core provides predictable performance and ease of troubleshooting. For customers who need to connect to other enterprises via an extranet or the Internet, the core topology should include one or more links to external networks. Corporate network administrators should	

discourage regional and branch-office administrators from planning their own extranets or connections to the Internet. Centralizing these functions in the core layer reduces complexity and the potential for routing problems, and is essential to minimizing security concerns. Bringing business-partner links into the branch office where collaboration is taking place might seem logical, but it means you have to allow the partner's traffic into the branch office but not beyond. Over time, you'll end up with a hodgepodge of distributed access control lists (ACL) and firewalls, which complicates policy enforcement. It also greatly raises costs if you want to use intrusion detection systems (IDS) and other security technologies. Similarly, some remote offices with IPsec VPN connectivity are shifting away from split access at the remote sites where users have local access to the Internet in addition to remote IPsec access to corporate headquarters. Despite bandwidth costs, forcing all	

external access to go through the core of the network means having only one security structure to administer, which is a good way to avoid security problems. Distribution Layer The distribution layer of the network is the demarcation point between the access and core layers of the network. The distribution layer has many roles, including controlling access to resources for security reasons and controlling network traffic that traverses the core for performance reasons. The distribution layer is often the layer that delineates broadcast domains. (Although this can be done at the access layer as well.) In network designs that include virtual LANs (VLAN), the distribution layer can be configured to route between VLANs. The distribution layer allows the core layer to connect sites that run different protocols while maintaining high performance. To maintain good performance in the core, the distribution layer can redistribute between bandwidth-intensive access layer routing protocols and optimized core routing protocols. For example, perhaps one site in the access layer is still running an older	kể cả người ta có thể cấu hình lớp phân phối để định tuyến các giao thức định tuyến lớp truy cập chuyên sâu băng thông và các giao thức định tuyến lõi tối ưu

protocol, such as IGRP. The distribution layer can redistribute between IGRP at the access layer and EIGRP in the core layer.	
To improve routing-protocol performance, the distribution layer can summarize routes from the access layer. For some networks, the distribution layer offers a default route to access layer routers and runs only dynamic routing protocols when communicating with core routers.	khái quát hóa các tuyến tuyến đến truy cập
To maximize hierarchy, modularity, and performance, the distribution layer should hide detailed topology information about the access layer from core routers. The distribution layer should summarize numerous access layer destinations into a few advertisements into the core. Likewise, the distribution layer should hide detailed topology information about the core layer from the access layer by summarizing to a small set of advertisements or just one default route, if possible. The distribution layer can provide the access layer with a route to the closest distribution layer router that has access to the core.	sự phân cấp khái quát hóa vô số điểm đến lớp truy cập vào một vài quảng bá trong lõi cho tuyến bộ định tuyến lớp phân phối
Access Layer The access layer provides users on local segments with access	ở từng phân đoạn mạng

to the internetwork. The access layer can include routers, switches, bridges, shared-media hubs, and wireless access points. As mentioned, switches are often implemented at the access layer in campus networks to divide up bandwidth domains to meet the demands of applications that need a lot of bandwidth or cannot withstand the variable delay characterized by shared bandwidth.	các chuyển mạch các hub dùng chung môi trường các chuyển mạch chịu được thời gian trễ biến thiên, một tính chất đặc trưng của cấu hình dùng chung băng thông
For internetworks that include small branch offices and telecommuter home offices, the access layer can provide access into the corporate internetwork using wide-area technologies such as ISDN, Frame Relay, leased digital lines, and analog modem lines. You can implement routing features, such as dial-on-demand routing (DDR) and static routing, to control bandwidth utilization and minimize cost on access layer remote links. (DDR keeps a link inactive except when specified traffic needs to be sent.)	định tuyến quay số theo yêu cầu các liên kết từ xa của lớp truy cập cần gửi
Guidelines for Hierarchical Network Design This section briefly describes some guidelines for hierarchical network design.	

Following these simple guidelines will help you design networks that take advantage of the benefits of hierarchical design.

The first guideline is that you should control the diameter of a hierarchical enterprise network topology. In most cases, three major layers are sufficient (as shown in Figure 5-4):

- The core layer
- The distribution layer
- The access layer

Controlling the network diameter provides low and predictable latency. It also helps you predict routing paths, traffic flows, and capacity requirements. A controlled network diameter also makes troubleshooting and network documentation easier.

Strict control of the network topology at the access layer should be maintained. The access layer is most susceptible to violations of hierarchical network design guidelines. Users at the access layer have a tendency to add networks to the internetwork inappropriately. For example, a network administrator at a branch office might connect the branch network to another branch, adding a fourth layer. This is a common network

design mistake known as adding a chain. Figure 5-6 shows a chain.

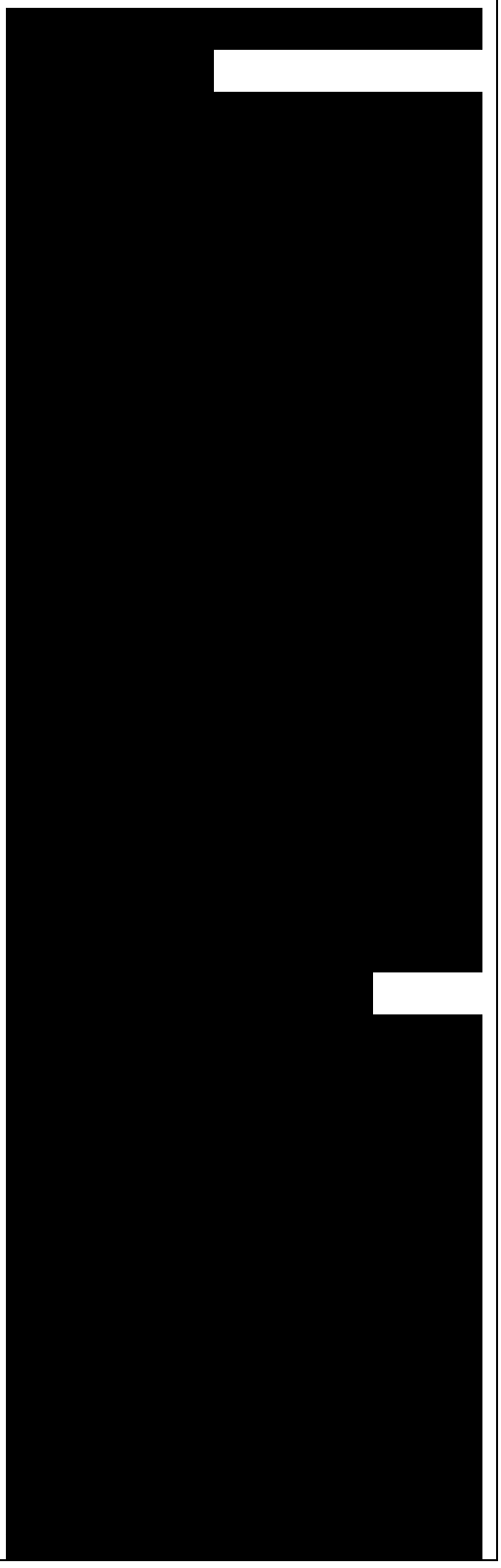


Figure 5-6 Chain and Backdoor at the Access Layer

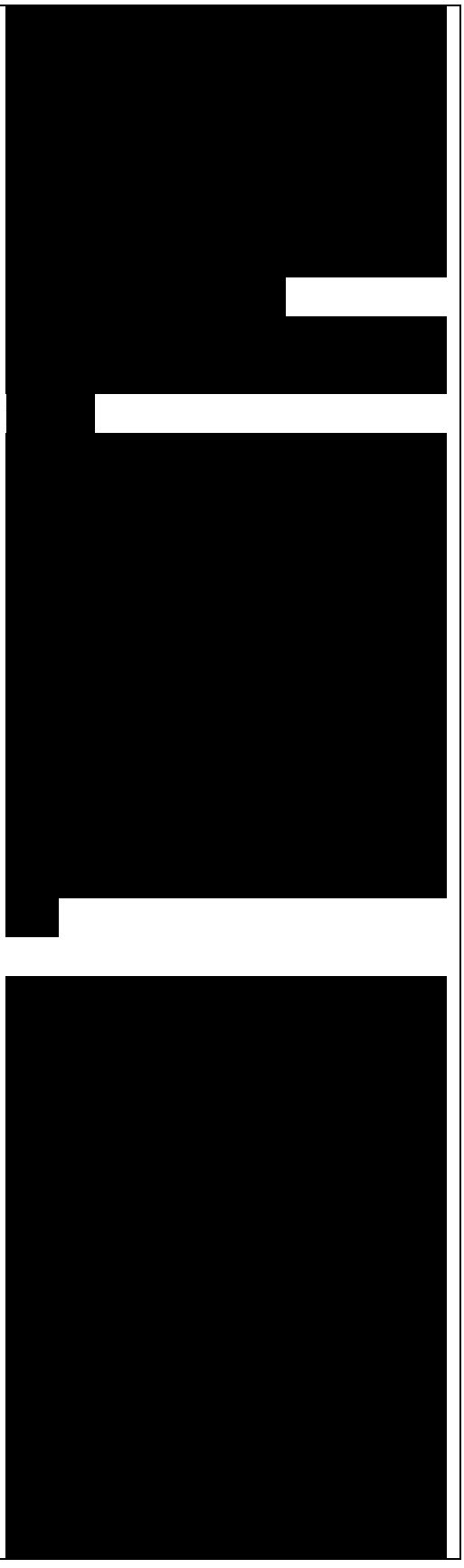
In addition to avoiding chains, you should avoid backdoors. A backdoor is a connection between devices in the same layer, as shown in Figure 5-6. A backdoor can be an extra router, bridge, or switch added to connect two networks. A backdoor can also be a hub; for example, someone might install a minihub in a conference room and accidentally connect the hub to two jacks instead of just one. Backdoors should be avoided because they cause unexpected routing and switching problems and make network documentation and troubleshooting more difficult.

Note Sometimes there are valid reasons for adding a chain or a backdoor. For example, international network topologies sometimes get skewed by the availability of fiber-optic links, the ease and cost of provisioning new networks, and the availability of competent carriers. An international network might require a chain to add another country. A backdoor is

sometimes added to increase performance and redundancy between two parallel devices in a layer. In general, however, other design options can usually be found that let the design retain its hierarchical structure. To maximize the benefits of a hierarchical model, you should usually avoid chains and backdoors.

Finally, one other guideline for hierarchical network design is that you should design the access layer first, followed by the distribution layer, and then finally the core layer. By starting with the access layer, you can more accurately perform capacity planning for the distribution and core layers. You can also recognize the optimization techniques you will need for the distribution and core layers.

You should design each layer using modular and hierarchical techniques and then plan the interconnections between layers based on your analysis of traffic load, flow, and behavior. To better understand network traffic characteristics, you can review the concepts covered in Chapter 4. As you select technologies for each layer, as discussed in Part III, "Physical Network Design," you might need to go back and tweak the design for other

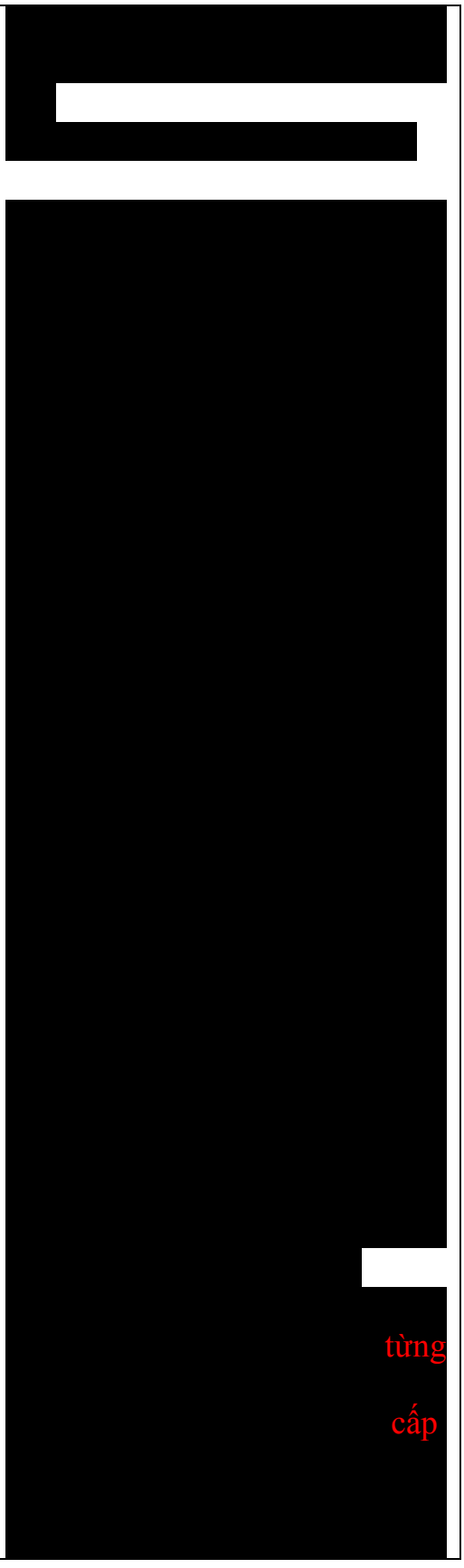


layers. Remember that network design is an iterative process.

Redundant Network Design Topologies

Redundant network designs enable you to meet requirements for network availability by duplicating elements in a network. Redundancy attempts to eliminate any single point of failure on the network. The goal is to duplicate any required component whose failure could disable critical applications. The component could be a core router, a switch, a link between two switches, a channel service unit (CSU), a power supply, a WAN trunk, Internet connectivity, and so on. To enable business survivability after a disaster and offer performance benefits from load sharing, some organizations have completely redundant data centers. Other organizations try to constrain network operational expenses by using a less-comprehensive level of redundancy.

You can implement redundancy inside individual campus networks and between layers of the hierarchical model. Implementing redundancy on campus networks can help you meet



availability goals for users accessing local services. You can also implement redundancy on the edge of the enterprise network to ensure high availability for Internet, extranet, and virtual private network (VPN) access. Note Because redundancy is expensive to deploy and maintain, you should implement redundant topologies with care. Be sure to select a level of redundancy that matches your customer's requirements for availability and affordability. Before you select redundant design solutions, you should first analyze the business and technical goals of your customer, as discussed in Part I, "Identifying Your Customer's Needs and Goals." Make sure you can identify critical applications, systems, internetworking devices, and links. Analyze your customer's tolerance for risk and the consequences of not implementing redundancy. Make sure to discuss with your customer the tradeoffs of redundancy versus low cost, and simplicity versus complexity. Redundancy adds complexity to the network topology and to network addressing and routing.	phục vụ các yêu cầu truy cập dịch vụ cục bộ của người dùng đảm bảo Phải đảm bảo rằng bạn đã xác định được quan trọng Phải chắc chắn rằng bạn đã bàn bạc tác động qua lại Việc triển khai dự phòng sẽ làm cho tô-pô mạng cũng như việc định địa chỉ và định tuyến mạng phức tạp thêm

Backup Paths

To maintain interconnectivity even when one or more links are down, redundant network designs include a backup path for packets to travel when there are problems on the primary path. A backup path consists of routers and switches and individual backup links between routers and switches, which duplicate devices and links on the primary path.

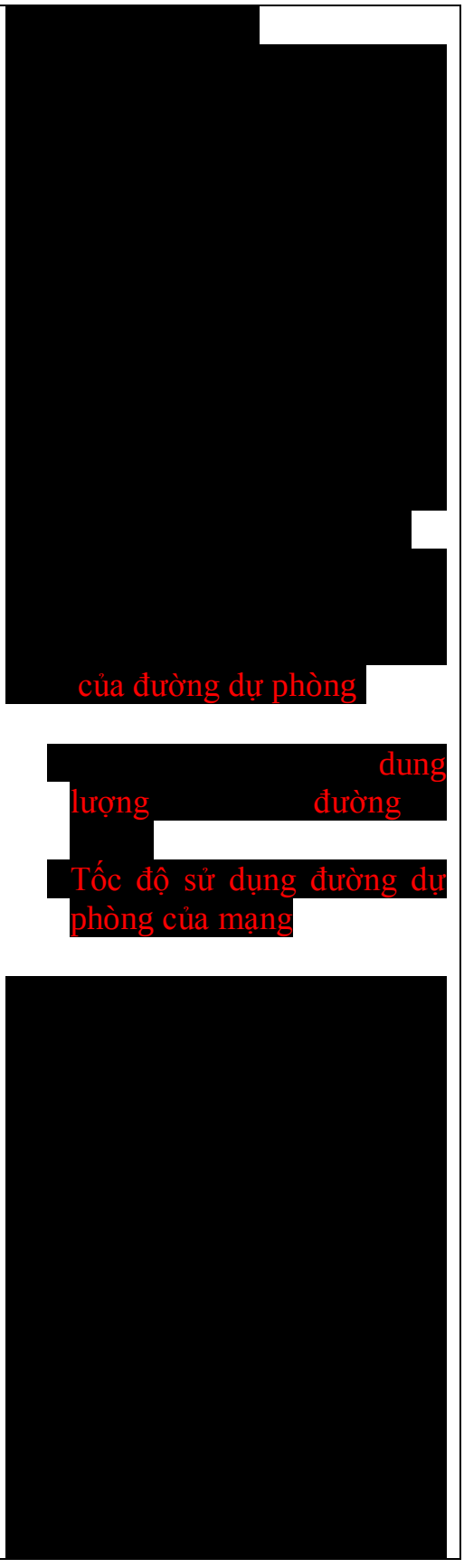
When estimating network performance for a redundant network design, you should take into consideration two aspects of the backup path:

- How much capacity the backup path supports
- How quickly the network will begin to use the backup path

You can use a network-modeling tool to predict network performance when the backup path is in use. Sometimes the performance is worse than the primary path, but still acceptable.

It is quite common for a backup path to have less capacity than a primary path.

Individual backup links within the backup path often use different technologies. For example, a leased line can be in parallel with a backup dialup



line or ISDN circuit. Designing a backup path that has the same capacity as the primary path can be expensive and is appropriate only if the customer's business requirements dictate a backup path with the same performance characteristics as the primary path.	mạch cần
If switching to the backup path requires manual reconfiguration of any components, users will notice disruption. For mission-critical applications, disruption is probably not acceptable. An automatic failover is necessary for mission-critical applications. By using redundant, partial-mesh network designs, you can speed automatic recovery time when a link fails.	cần phải cấu hình lại bất kỳ thành phần nào bằng phương pháp thủ công dự phòng Thông qua việc sử dụng các thiết kế mạng mắc lưới một phần, bạn có thể tăng tốc thời gian phục hồi tự động khi liên kết hỏng.
One other important consideration with backup paths is that they must be tested. Sometimes network designers develop backup solutions that are never tested until a catastrophe happens. When the catastrophe occurs, the backup links do not work. In some network designs, the backup links are used for load sharing and redundancy. This has the advantage that the backup path is a tested solution that is regularly used and	và được giám sát hàng ngày hoặc

monitored as a part of day to day operations. Load sharing is discussed in more detail in the next section.	từng khoản thời gian trong ngày
Load Sharing The primary purpose of redundancy is to meet availability requirements. A secondary goal is to improve performance by supporting load sharing across parallel links. Load sharing, sometimes called load balancing, allows two or more interfaces or paths to share traffic load.	hổ trợ
Note Purists use the term load sharing instead of load balancing because the load is usually not precisely balanced across multiple links. Because routers can cache the interface that they use for a destination host or even an entire destination network, all traffic to that destination tends to take the same path. This results in the load not being balanced across multiple links, although the load should be shared across the links if there are many different destinations.	Những người theo chủ nghĩa thuần túy sử dụng thuật ngữ chia sẻ tải thay cho cân bằng tải thậm chí toàn bộ mạng đích qua các liên kết
In WAN environments, you can facilitate load sharing by configuring channel aggregation. Channel aggregation means that a router can automatically bring up multiple channels as bandwidth	tạo điều kiện thuận lợi cho kích hoạt

requirements increase. The Multilink Point-to-Point Protocol (MPPP) is an Internet Engineering Task Force (IETF) standard for channel aggregation. MPPP ensures that packets arrive in sequence at the receiving router. To accomplish this, data is encapsulated within the Point-to-Point Protocol (PPP), and datagrams are given a sequence number. At the receiving router, PPP uses the sequence number to re-create the original data stream. Multiple channels appear as one logical link to upper-layer protocols.	bộ định tuyến nhận đóng số thứ tự bộ định tuyến nhận số thự tự luồng dữ liệu
Most vendors' implementations of IP routing protocols support load sharing across parallel links that have equal cost. (Cost values are used by routing protocols to determine the most favorable path to a destination. Depending on the routing protocol, cost can be based on hop count, bandwidth, delay, or other factors.) With EIGRP, Cisco supports load sharing even when the paths do not have the same cost. Using a feature called variance, EIGRP can load share across paths that do not have the same cost. Cisco supports load sharing across six parallel paths.	định tuyến thuận tiện nhất t

	song song
Some routing protocols base cost on the number of hops to a particular destination. These routing protocols load balance over unequal bandwidth paths as long as the hop count is equal. When a slow link becomes saturated, however, higher-capacity links cannot be filled. This is called pinhole congestion, which can be avoided by designing equal bandwidth links within one layer of the hierarchy, or by using a routing protocol that bases cost on bandwidth and has the variance feature.	hop Tuy nhiên bão hòa các liên kết dung lượng cao không thể được làm đầy giá thành
Modular Network Design Top-down network design lets you drill down to the components of the network design and apply fundamental design principles to the components and the overall design. Hierarchy and redundancy, as mentioned in the previous sections, are fundamental network design concepts. Another fundamental concept related to hierarchy is	toàn bộ thiết kế khác

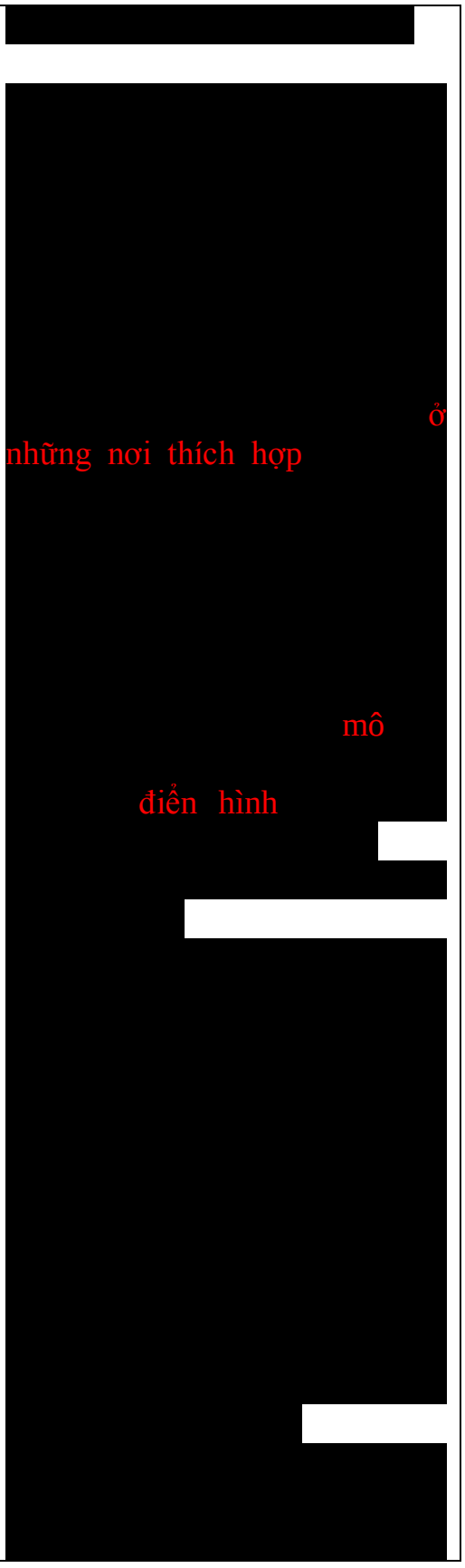
modularity.

Large network design projects and large networks in general consist of different areas or modules. Each area should be designed using a systematic, top-down approach, applying hierarchy and redundancy where appropriate. Network solutions and services can be selected on a per-module basis but validated as part of the overall network design. Cisco developed the SAFE security reference architecture to depict the components or modules of a typical enterprise network. The next section describes the architecture.

Cisco SAFE Security Reference Architecture

SAFE is a reference architecture that network designers can use to simplify the complexity of a large internetwork. The architecture lets you apply a modular approach to network design. With SAFE, you can analyze the functional, logical, and physical components of a network and thus simplify the process of designing an overall enterprise network.

Cisco SAFE architecture is especially concerned with security. SAFE takes a



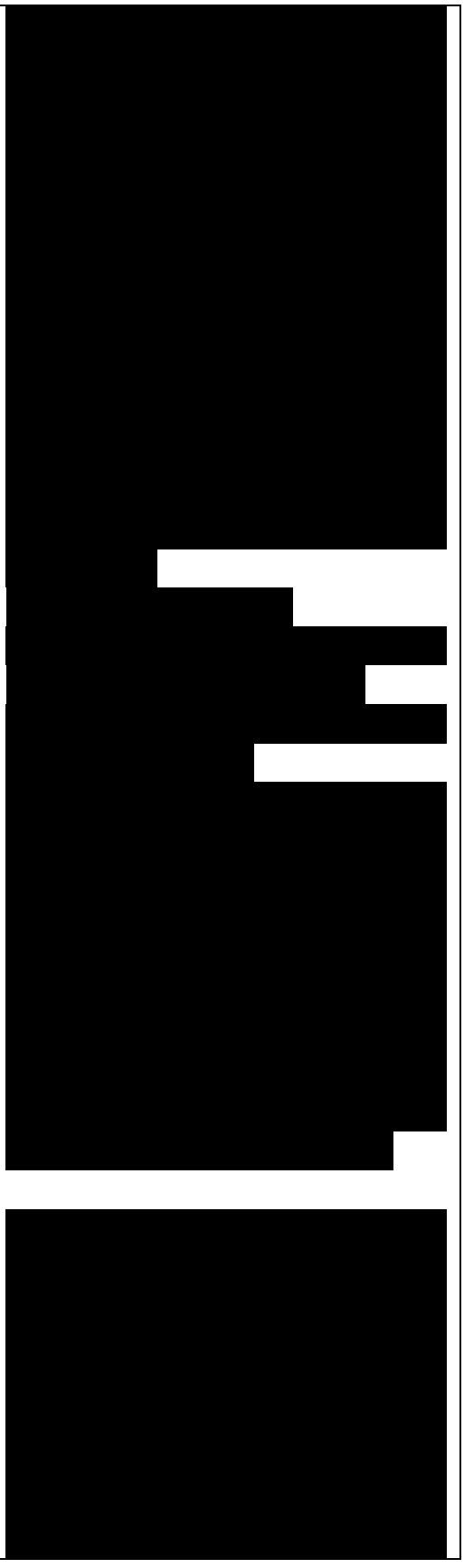
defense-in- depth approach, in which multiple layers of protection are strategically located throughout the network. The layers are under a unified strategy for protecting the entire network and the various components of the network, including individual network segments, infrastructure devices, network services, endpoints, and applications. Figure 5-7 shows the main modules in the SAFE architecture.

High-Level View

Figure 5-7 High-Level View of Cisco SAFE Architecture

SAFE architecture comprises the following major modules:

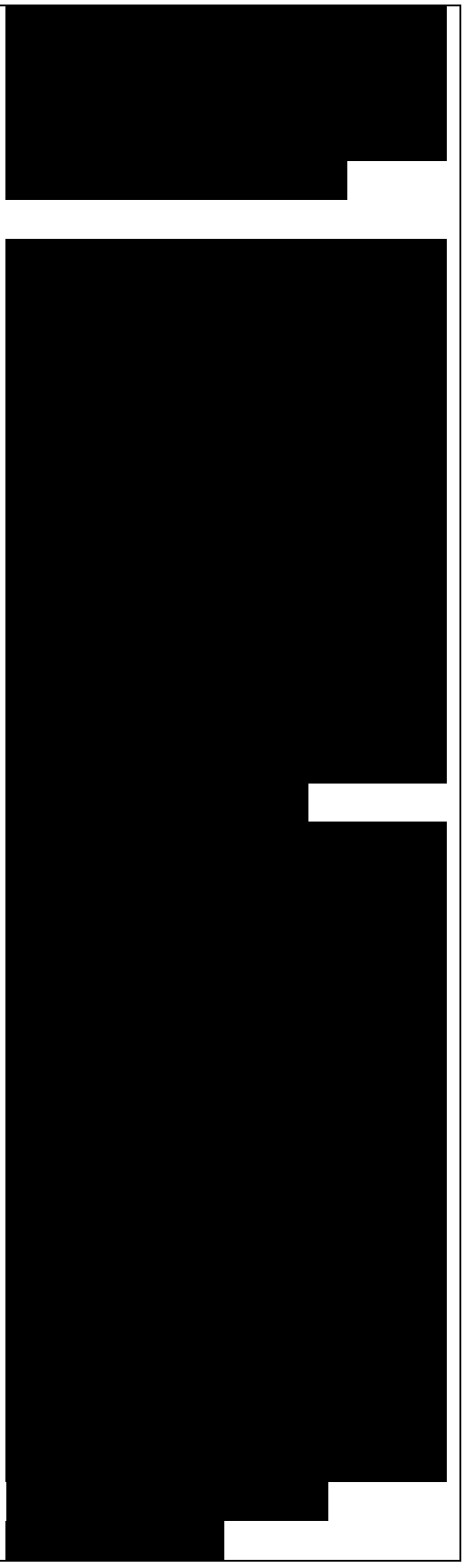
- Core: The core stitches together all the other modules. The core is a high-speed infrastructure that provides reliable and scalable Layer 2 and Layer 3 transport. The core is typically implemented with redundant switches that aggregate the connections to the campus, data center, WAN edge, and Internet edge.
- Data center: The data center hosts servers, applications, and storage devices for use by internal users. The data center also connects the network infrastructure that these devices require, including routers, switches, load balancers,

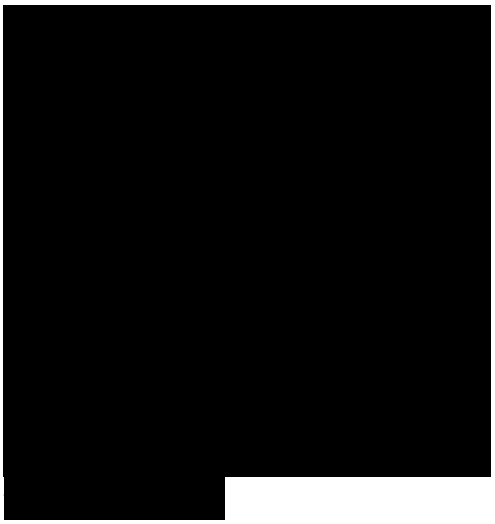


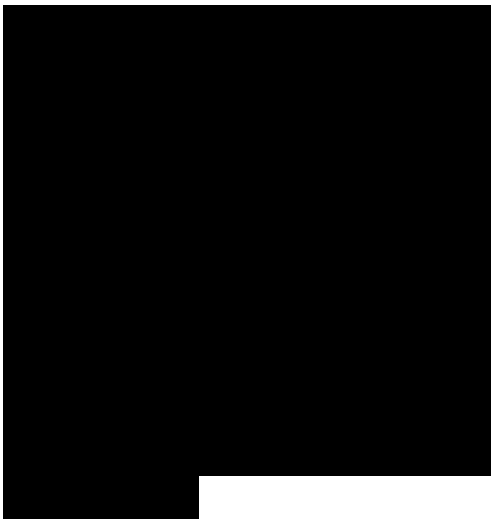
content delivery devices, and application acceleration devices. The data center is not directly accessible from the Internet to the general public.

■ Campus: The campus network provides network access to end users and devices located in a single geographical location. The campus may span several floors in a single building or multiple buildings for larger enterprises. The campus hosts local data, voice, and video services. The campus design should allow campus users to securely access data center and Internet resources from the campus infrastructure.

■ Management: The management network provides monitoring, analysis, authentication, and logging services. Management servers support RADIUS, Kerberos, Network Time Protocol (NTP), Simple Network Management Protocol (SNMP), and syslog traffic. The management network combines out-of-band (OOB) management and in-band (IB) management, spanning all the building blocks of the SAFE architecture. The OOB management network can be implemented as a collection of dedicated switches or through



the use of isolated VLANs. ■ WAN edge: The WAN edge is the portion of the network that aggregates WAN links that connect geographically distant branch offices to a central site or regional hub. The WAN can be owned by the enterprise or by a service provider, the latter being the more common option.	
■ Internet edge: The Internet edge is the infrastructure that provides connectivity to the Internet and that acts as a gateway for the enterprise to the rest of the world. Internet edge services include a public DMZ, corporate Internet access, and remote- access VPN.	
■ Branches: Branches provide connectivity to users and devices at remote locations. A branch office typically includes one or more LANs and connects to the central site via a private WAN or an Internet connection using VPN technology. Branches host local data, voice, and video services.	
■ Extranet: An extranet allows selected business	

partners, customers, and suppliers to access a portion of the network via secure protocols. Extranet services include remote-access VPN, threat detection and mitigation, stateful failover for servers and network devices, and topological redundancy.	
■ Partner site: Partner sites are networks owned by business partners, customers, and suppliers. They access services in the extranet via secure WAN or Internet connectivity.	
■ E-Commerce: The e-commerce module hosts applications, servers, and data used in the selling and buying of products. Services include Layer 2 through 7 security, server farms with traffic filtering, and server load balancing. Virtual contexts provide segmentation and policy enforcement for server-to-server communication.	
■ Teleworker: The teleworker module is the home office of a full-time or part-time employee. Services in the teleworker module include remote-access VPN, desktop security, secure wireless networking, IP telephony, and	

IP video.

■ Cisco SensorBase: Cisco SensorBase consists of threat collection servers that receive daily updates from globally deployed sensors regarding threats such as botnets, dark nets, malware, and serial attackers. Sensors include intrusion prevention systems, email servers, and web security appliances.

Designing a Campus Network Design Topology

Campus network design topologies should meet a customer's goals for availability and performance by featuring small bandwidth domains, small broadcast domains, redundancy, mirrored servers, and multiple ways for a workstation to reach a router for off-net communications. Campus networks should be designed using a hierarchical, modular approach so that the network offers good performance, maintainability, and scalability.

A campus network can consist of access, distribution, and core

layers: ■ Campus access layer: This module contains end-user workstations and IP phones connected to switches or wireless access points. Higher-end switches provide uplinks to the distribution layer. Services offered by this module include network access, broadcast control, protocol filtering, and the marking of packets for quality of service (QoS) features. ■ Campus distribution layer: The job of this module is to aggregate wiring closets within a building and provide connectivity to the campus core via routers (or switches with routing modules). This module provides routing, QoS, and access control methods for meeting security and performance requirements. Redundancy and load sharing are recommended for this module. For example, each building should have two equal-cost paths to the campus core. ■ Campus core layer: The campus core interconnects the	[redacted]

access and distribution modules with the data center, network management, and edge modules. The campus core provides redundant and fast-converging connectivity. It routes and switches traffic as quickly as possible from one module to another. This module usually uses high-speed routers (or switches with routing capability) and provides QoS and security features.

Spanning Tree Protocol

The topology of a campus network design is often determined by the Spanning Tree Protocol (STP), which is a protocol and algorithm, documented in IEEE 802.1D, for dynamically "pruning" an arbitrary topology of connected Layer 2 switches into a spanning tree. The topology that results spans the entire switched domain and is shaped like a mathematical tree, with branches that spread out from a stem without forming loops or polygons. The network designer physically connects switches in a meshed, redundant topology, but STP creates a logical tree with no redundancy.

Note Because STP was first developed when bridges were used instead of switches, STP discussions use the word bridge for the Layer 2 device that today we call a switch.

The spanning tree has one root bridge and a set of ports on other bridges that forward traffic toward the root bridge. Bridges send bridge protocol data unit (BPDU) frames to each other to build and maintain the spanning tree. BPDUs identify the root bridge and help the other bridges compute their lowest-cost path to the root. Bridges send topology change notification BPDUs when bridge ports change state. Bridges send configuration BPDUs every 2 seconds to maintain the spanning tree. BPDUs are sent to the Bridge Group Address 01:80:C2:00:00:00.

Spanning Tree Cost Values
As already mentioned, bridges compute their lowest-cost path to the root bridge. The lowest-

cost path is usually the highest-bandwidth path, although the cost is configurable. The original IEEE 802.1D specification used a 16-bit field for the cost of links of different speeds. The 2004 version uses a 32-bit field. Cisco switches by default use a 16-bit field with the default values shown in Table 5-1. Cisco switches can also be configured for 32-bit values. Table 5-1 shows the default Cisco 16-bit values and the default 32-bit values for the 2004 version of IEEE 802.1D.

Table 5-1 IEEE 802.1D and Cisco 16-Bit Cost Values

Rapid Spanning Tree Protocol

In 2004, the IEEE incorporated its 802.1w standard, "Rapid Reconfiguration of Spanning Tree," into the IEEE 802.1D standard. The goal of the 802.1w committee was to standardize an improved mode of switch operation that reduces the time STP takes to converge to a least-cost spanning tree and to restore

service after link failures. With the original 802.1D standard, networks took almost a minute to converge. In a properly configured network, the new 802.1D Rapid Spanning Tree Protocol (RSTP) can achieve convergence or reconvergence in a few hundred milliseconds.

With RSTP, bridge ports can be in one of three states:

- Discarding: A port that is neither learning MAC addresses nor forwarding users' frames
- Learning: A port that is learning MAC addresses to populate the MAC-address table but is not yet forwarding user frames
- Forwarding: A port that is learning MAC addresses and forwarding user frames

The original STP passively waited for the network to converge before it transitioned a port into the forwarding state. To achieve quick convergence, a network administrator had to carefully tune the conservative default values for the Maximum Age and Forward

Delay timers, which put the stability of the network at stake. RSTP, on the other hand, can actively confirm that a port can safely transition to the forwarding state without having to rely on any timer configuration. There is now a synchronization mechanism that takes place between RSTP-compliant bridges so that they actively build the topology as quickly as possible. As was the case with STP, RSTP elects the bridge with the lowest bridge ID as the root bridge. Every bridge has a root path cost associated with it. For the root bridge this is zero. For all other bridges, it is the sum of the port path costs on the least-cost path to the root bridge. When the switched network has converged, each bridge port has one of the following roles: ■ Root: Assigned to the one port on a nonroot bridge that provides the lowest-cost path to the root bridge. If a bridge has two or more ports with the same cost, the port with the lowest port ID is	[redacted]

selected as the root port. A root port is a forwarding port. ■ Designated: Assigned to the one port attached to a LAN that provides the lowest- cost path from that LAN to the root bridge. All ports on the root bridge are designated ports. If there are two or more bridges with the same cost, the bridge with the lowest bridge ID is the designated bridge. The bridge port on the designated bridge that is connected to the LAN is assigned the role of designated port for that LAN. If the designated bridge has two or more ports connected to the LAN, the bridge port with the lowest port ID is selected as the designated port. A designated port is a forwarding port. ■ Alternate: Assigned to a port that offers an alternative path in the direction of the root bridge to that provided by the bridge's root port. An alternative port is a discarding port. ■ Backup: Assigned to a port on a designated bridge that acts as a backup for the path provided by a designated port in the direction of the leaves of the spanning tree. Backup ports exist only where two ports on a	[redacted]

bridge are connected together in loopback by a point-to-point link, or where the bridge has two or more connections to a shared-media LAN. A backup port is a discarding port. ■ Disabled: Assigned to a port that is not operational or is excluded from the active topology by network management. A disabled port is a discarding port. The 2004 version of 802.1D also supports the concept of an edge port. A network manager can configure a port as an edge port if it is attached to a LAN that has no other bridges attached. (RSTP can also automatically detect edge ports.) Edge ports transition directly to the forwarding state, which is a major benefit for access layer ports that connect end-user systems. RSTP still continues to monitor the port for BPDUs in case a bridge is connected. As soon as the bridge detects a BPDU arriving at an edge port, the port becomes a nonedge port. An edge port corresponds to the Cisco PortFast feature (and is configured with the Cisco spanning-tree portfast	[redacted]

command).	[redacted]
In a stable network where RSTP has communicated consistent information throughout the network, every LAN has one and only one designated port, and every bridge with the exception of the root bridge has a single root port connected to a LAN. Because each bridge provides connectivity between its root port and its designated ports, the resulting active topology connects all LANs and has no loops. In other words, it is a spanning tree.	[redacted]
RSTP Convergence and Reconvergence RSTP can converge quickly to a tree topology where the lowest-cost paths are forwarding frames. RSTP achieves rapid transition to the forwarding state on edge ports, root ports, and point-to-point links. Edge and root ports can transition to forwarding without transmitting or receiving messages from other bridges.	[redacted]
A designated port attached to a point-to-point link can transition to the forwarding	[redacted]

state when it receives an explicit role agreement transmitted by the other bridge attached to the link. In the case of a shared LAN, the forwarding transition delay used by a designated port is long enough for other bridges attached to the LAN to receive and act on transmitted messages but is independent of the overall network size. If all the LANs in a network are point-to-point links, RSTP timers define worst-case delays that occur only if protocol messages are lost or rate transmission limits are exceeded.	[redacted]
The link type for a LAN is automatically derived from the duplex mode of a port. A port that operates in full-duplex mode is assumed to be point-to-point, whereas a half-duplex port is considered to be a shared port by default. (The automatic link type setting can also be overridden by explicit configuration.) In modern switched networks, most links operate in full-duplex mode and RSTP treats them as point-to-point links. This makes them	[redacted]

candidates for rapid transition to the forwarding state. If the physical connectivity of the network changes or management parameters change, new spanning-tree information propagates rapidly. Each bridge accepts better information from any bridge on a LAN or revised information from the prior designated bridge for the LAN. Updated configuration BPDUs are transmitted through designated ports until the leaves of the spanning tree are reached. The transmission of information ceases as new configuration BPDUs reach designated ports that have already received the new information through redundant paths in the network, or reach LANs that are not redundantly connected. In the original 802.1D specification, a bridge that detected a topology change generated a topology change notification up to the root. The root then flooded the change until the Maximum Age and Forward Delay timers expired. In the newer 802.1D	[redacted]

specification, which includes the 802.1w enhancements, the change propagation is a one-step process. The initiator of the topology change floods the information throughout the network. There is no need to wait for the root bridge to be notified or for bridges to maintain a topology change state until the timers expire. RSTP includes another form of immediate transition to the forwarding state that is similar to the Cisco proprietary UplinkFast extension to STP. When a bridge loses its root port, it can immediately transition an alternative port into the forwarding state. The selection of an alternative port as the new root port generates a topology change. The RSTP topology change mechanism clears the appropriate entries in the MAC address tables of the affected bridges, a necessary step to ensure that the table accurately reflects the new topology and the path to destination MAC addresses.	[redacted]

Selecting the Root Bridge It is good practice to control which switch becomes the root bridge. The root bridge should be a reliable, high-speed switch in the center of the switched topology. If you let switches elect the root on their own, you have little control over the direction that traffic flows and the amount of frame-forwarding delay in your network. If you aren't careful, a slow bridge can become the root bridge. Also, high-speed ports can accidentally be removed from the spanning tree in deference to low-speed ports that are closer to the root bridge. The root bridge is the switch with the lowest bridge ID. The bridge ID has two parts, a priority field and the MAC address of the switch. If all priorities are left at their default value, the switch or bridge with the lowest MAC address becomes the root. This could easily be one of the earliest products from Cisco, because Cisco had such a low vendor ID. (The vendor ID makes up the first 3 bytes of a MAC address, and the original vendor ID for Cisco was 00:00:0C.)	[redacted]

Manual control of the root bridge selection process is critical to maintaining high throughput on switched networks. This can be accomplished by ensuring that a particular switch has the lowest bridge ID. It is not recommended (or even possible on some switches) to change the MAC address portion of a bridge ID. Instead, to control the bridge ID, set the bridge priority. On Cisco switches, you can use the spanning-tree vlan vlan-id priority command. You should give a single, high-speed, centrally located switch the lowest priority so that it becomes the root bridge. You should also lower the priority on another high-speed, centrally located switch so that it becomes the root if the primary root fails. Generally these two switches are distribution layer switches.	[redacted]
Note Cisco also supports a Root Guard feature that protects your network from a low- speed switch hijacking the job of root bridge. A switch port configured for Root Guard cannot become a root port. Instead the port becomes a	[redacted]

designated port for its LAN segment. If there is a better BPDU received on the port, Root Guard disables the port, rather than taking the BPDU into account and restarting election of the root bridge. Root Guard needs to be enabled on all ports on all switches that should not become the root bridge. Scaling the Spanning Tree Protocol STP works best when the switched network is kept relatively small and the switches have sufficient CPU power and RAM to do their jobs effectively. On an oversized network with switches that have too little CPU power, BPDUs might not be sent and received properly, resulting in loops. A switch with insufficient RAM or a software problem could also drop BPDUs. In addition, a congested network could cause problems for the transmission of BPDUs. To ensure that old information	[redacted]

does not endlessly circulate through redundant paths in a network, thus preventing the propagation of new information, each BPDU includes a Message Age and a Maximum Age. When the Message Age exceeds the Maximum Age, the BPDU is discarded. On large, slow switched networks, BPDUs can get discarded before they reach all the switches. This problem causes STP to reconverge much more frequently than it should. STP relies on the timely reception of BPDUs. Cisco has a BPDU skew detection feature that enables a switch to keep track of late-arriving BPDUs and notify the administrator by means of syslog messages. This feature is more of a workaround than a solution. A better plan is to design a switched network with care. Keep the switched network small with a root bridge selected for its high performance and central placement in the network. It is also common to develop campus networks that rely on STP only for inadvertent loops.	[redacted]

In other words, the Layer 2 topology is intentionally designed to be a tree topology by default, so there's no need for STP to remove loops to create a tree. STP is still enabled in case a novice network engineer (or a user) places a switch into the network in a looped fashion, but STP is relegated to a protection role instead of its legacy operational role. Routers and routing protocols are added to the design at the distribution and core layers. Routers have advantages compared to switches in their capability to implement fast-converging routing protocols and security policies, load sharing, and QoS features.	[redacted]
Note Cisco also supports a BPDU Guard feature that protects your network from a low- speed switch joining a network and sending BPDUs. A switch port configured for BPDU Guard goes into errdisable mode.	[redacted]
Virtual LANs A campus network should be designed using small	[redacted]

bandwidth and small broadcast domains. A bandwidth domain is a set of devices that share bandwidth and compete for access to the bandwidth. A traditional bus topology or hub-based Ethernet, for example, is a single bandwidth domain. A switch divides up bandwidth domains and is often used to connect each device so that the network consists of many, extremely small bandwidth domains. With switches, as opposed to hubs, the bandwidth domain consists of the switch port and the device that connects it. If full-duplex transmission mode is used, a bandwidth domain becomes even smaller and consists of just the port or the device. Note On networks that experience collisions, including traditional Ethernet, a bandwidth domain is also called a collision domain. A broadcast domain is a set of devices that can all hear each other's broadcast frames. A broadcast frame is a frame that is sent to the MAC address FF:FF:FF:FF:FF:FF. By default, switches do not divide broadcast domains. The campus access layer should use switches and provide broadcast	[redacted]

control, however. To accomplish this, virtual LANs are necessary.	[redacted]
A virtual LAN (VLAN) is an emulation of a standard LAN that allows data transfer to take place without the traditional physical restraints placed on a network. A VLAN is a set of LAN devices that belong to an administrative group. Group membership is based on configuration parameters and administrative policies rather than physical location. Members of a VLAN communicate with each other as if they were on the same wire or hub, when they might be located on different physical LAN segments. Members of a VLAN communicate with members in a different VLAN as if they were on different LAN segments, even when they are located in the same switch. Because VLANs are based on logical instead of physical connections, they are extremely flexible.	[redacted]
In the early days of VLANs in the mid-1990s, there was a lot	[redacted]

of talk about using VLANs to group users working on a project together, even though they weren't physically located together. With VLANs, the physical location of a user doesn't matter. A network administrator can assign a user to a VLAN regardless of the user's location. In theory, VLAN assignment can be based on applications, protocols, performance requirements, security requirements, traffic-loading characteristics, or other factors.	[redacted]
There was also a lot of talk about VLANs simplifying moves, adds, and changes in a campus network. The theory was that with VLANs, network administrators can stay seated in their offices or in the wiring closet when an end user moves into a new office or cubicle. If a user in the marketing department, for example, moves to a new office that is physically located among engineers, the marketing person might not have the skills to configure IP addressing for compatibility with the new location. Asking the engineers for help might not work because engineers don't like marketers, and	[redacted]

asking the network administrator to come to the office and make the change might take a long time because administrators are so busy. Instead, the network administrator can configure the switch port for the moved device to be part of the marketing VLAN. Additional changes might be necessary to make sure the other switches learn that the marketing VLAN has expanded into a new area; however, no change is required on the marketer's computer.	[redacted]
In modern networks, VLANs aren't often used this way. Manually configuring IP addresses isn't common because DHCP has become so popular. Also, when a VLAN is dispersed across many physical networks, traffic must flow to each of those networks, which affects the performance of the networks and adds to the capacity requirements of links that connect VLANs. Networks with VLANs that migrate all over the campus topology are hard to manage and optimize.	[redacted]

In modern networks, instead of allowing for the spread of a logical LAN or administrative group across many physical LANs, a VLAN has become a method to subdivide physical switch-based LANs into many logical LANs. VLANs allow a large, flat, switch-based network to be divided into separate broadcast domains. Instead of flooding all broadcasts out every port, a VLAN-enabled switch floods a broadcast out only the ports that are part of the same VLAN as the sending station.	[redacted]
When switches first became popular in the mid-1990s, many companies implemented large, switched campus networks with few routers. The goals were to keep costs down by using switches instead of routers, and to provide good performance because presumably switches were faster than routers. Without the router capability of containing broadcast traffic, however, the companies needed VLANs, which enable the large flat network to be divided into broadcast domains. A router (or a routing module within a	[redacted]

switch) is still needed for inter-VLAN communication. In IP-based campus networks, a VLAN is usually its own IP subnet, due to the way the Address Resolution Protocol (ARP) works. When an IP host in a subnet needs to reach another host in the same subnet, it sends an ARP message to determine the Media Access Control (MAC) address of the host it is trying to reach. The ARP message is sent as a broadcast. All devices that find each other this way need to be in the same VLAN. Thus, in an IP network, VLANs are implemented as separate IP subnets. A router (or a routing module within a switch) provides intersubnet communication just as it would for a set of interconnected real (not virtual) LANs.	[redacted]

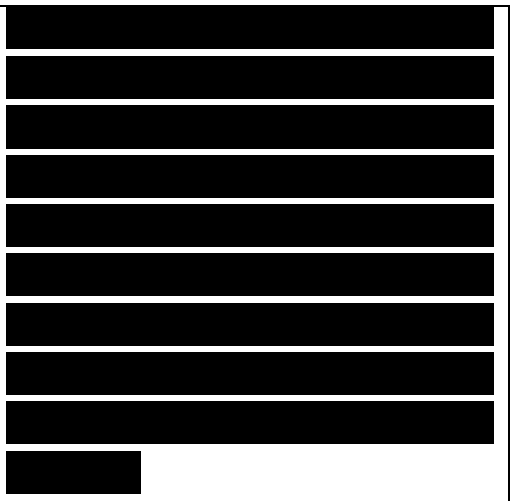
Fundamental VLAN Designs

To understand VLANs, it helps to think about real (nonvirtual) LANs first. Imagine two switches that are not connected to each other in any way. Switch A connects stations in Network A, and Switch B connects stations in Network B, as shown in Figure 5-8.

When Station A1 in Figure 5-8 sends a broadcast, Station A2 and Station A3 receive the broadcast, but none of the stations in Network B receive the broadcast, because the two switches are not connected. This same configuration can be implemented through configuration options in a single switch, with the result looking like Figure 5-9.

Through the configuration of the switch there are now two virtual LANs implemented in a single switch, instead of two separate physical LANs. This is the beauty of VLANs. The broadcast, multicast, and unknown-destination traffic originating with any member of VLAN A is forwarded to all other members of VLAN A,

and not to a member of VLAN B. VLAN A has the same properties as a physically separate LAN bounded by routers. The protocol behavior in Figure 5-8 is exactly the same as the protocol behavior in Figure 5-9.



Switch A

Station A1 Station A2 Station A3

Switch B

Station B1 Station B2 Station B3

Network A

Network B

Figure 5-8 Two Switches with Stations Attached

VLAN A

Station A1 Station A2 Station A3

Station B1 Station B2 Station B3 VLAN B

Figure 5-9 Single Switch with Stations from Network A and Network B Attached

VLANs can span multiple switches. In Figure 5-10, both switches contain stations that are members of VLAN A and VLAN B. This design introduces a new problem, the

solution to which is specified in the IEEE 802.1Q standard and the Cisco proprietary InterSwitch Link (ISL) protocol. The problem has to do with the forwarding of broadcast, multicast, or unknown-destination frames from a member of a VLAN on one switch to the members of the same VLAN on the other switch.	[redacted]
In Figure 5-10, all frames going from Switch A to Switch B take the same interconnection path. The 802.1Q standard and the Cisco ISL protocol define a method for Switch B to recognize whether an incoming frame belongs to VLAN A or to VLAN B. As a frame leaves Switch A, a special header is added to the frame, called the VLAN tag. The VLAN tag contains a VLAN identifier (ID) that specifies to which VLAN the frame belongs.	[redacted]
Because both switches have been configured to recognize VLAN A and VLAN B, they can exchange frames across the	[redacted]

interconnection link, and the recipient switch can determine the VLAN into which those frames should be sent by examining the VLAN tag.

The link between the two switches is sometimes called a trunk link or simply a trunk.

Trunk links allow the network designer to stitch together VLANs that span multiple switches. A major design consideration is determining the scope of each VLAN and how many switches it should span. As mentioned earlier, most designers try to keep the scope small. Each VLAN is a broadcast domain, and per recommendations specified in the previous chapter, a single broadcast domain should be limited to a few hundred workstations (or other devices, such as IP phones).

Another major design consideration is the capacity of trunk links. Using methods discussed in Chapter 4, you should study network traffic to determine if Fast Ethernet, Gigabit Ethernet, or multiples of Fast or Gigabit Ethernet will be required for trunk links.

Although Cisco supports 10-Mbps Ethernet trunks on some equipment, 10 Mbps is usually sufficient only for trunks that support small networks or for lab networks used for learning and testing purposes.	[redacted]
Wireless LANs As discussed in Part I of this book, user mobility has become an important goal for many enterprises. In a campus network design, one or more wireless LANs (WLAN) can meet this goal by offering intranet and Internet access in open areas on the campus and in high-density areas such as auditoriums, conference rooms, and cafeterias. WLAN technology also enables deployment of LANs in offices or other parts of buildings where it may not be cost effective or practical to install cabling. A WLAN consists of access	[redacted]

points that communicate using radio frequency (RF) with wireless clients. The area that a single access point can cover is often called a wireless cell. Designing a WLAN topology requires a designer to determine the coverage area of each wireless cell and to decide how many cells will be required to meet total coverage needs. Factors that affect the coverage of a single access point include data rate, power level, antenna choice, and antenna positioning. Architectural characteristics of the wireless site also affect coverage, as described in the "Checking a Site for a Wireless Installation" section in Chapter 3, "Characterizing the Existing Internetwork."	[redacted]
Positioning an Access Point for Maximum Coverage	[redacted]
Most access points use an isotropic antenna, which means that the signal strength is theoretically the same when measured along axes in all directions. If you suspend an access point in space, the coverage should resemble that of a three-dimensional sphere with the access point at its	[redacted]

center. In reality, the limitations of antenna design usually result in less-uniform coverage, however. The most common type of access point antenna is omnidirectional, which isn't actually "omni" or "iso." Instead of a sphere, the coverage looks more like a donut or tire inner tube. An omnidirectional antenna is usually a 4- to 6-inch transmitting element, often attached to a rotating or positionable pivot. The signal propagating from an omnidirectional antenna is strongest in a direction perpendicular to the antenna shaft and weakest in the same direction as the antenna shaft. Remembering this can help you position your antennae for maximum coverage (and help you decide if you might need a directional antenna instead of an omnidirectional antenna). Caution Think about the meaning of omni in omnidirectional. Placing an access point near an exterior wall means that some of the signal will probably radiate strongly outside the building, where an unauthorized user	[redacted]

sitting in the parking lot can easily receive it.	[redacted]
Also, keep in mind that an access point in one room can propagate through a wall to potentially interfere with an access point in the next room. Walls attenuate (decrease) the strength of the signal, but they don't block it completely.	[redacted]
Access points can be mounted in a horizontal or a vertical position. It's important to make sure that an omnidirectional antenna points straight up. In addition to the access point antenna, also consider the antenna in receiving stations, usually notebook computers. Every wireless NIC and computer is different. Some laptops have long antennas that extend from the card up through the back of the laptop, behind the screen. Other computers might not have a built-in antenna and must rely on a smaller antenna in the NIC. You should test your WLAN design with a variety of computers and other devices that the actual users will be using.	[redacted]

For a given data rate, you can alter the power level or choose a different antenna to change the coverage area and coverage shape. A large cell size might result in too many clients sharing the available bandwidth. (IEEE 802.11 WLANs are shared networks, with all devices in the same bandwidth domain.) By reducing the access point power or antenna gain, you can reduce the cell size and share the bandwidth with fewer clients. This will result in more access points for a given coverage area but will provide better performance for clients. WLANs and VLANs You can place multiple access points throughout a facility to give users the ability to roam freely throughout an extended area while maintaining uninterrupted access to network resources. The easiest method for making sure users can roam is to put all of the users in the same IP subnet and the same VLAN. Otherwise, devices that move from subnet	[redacted]

to subnet must acquire a new IP address and can lose packets that might have been transmitted while they were acquiring an address.	[redacted]
Whenever possible, a WLAN should be a separate subnet to simplify addressing while roaming and also to improve management and security. Keeping all wireless clients in their own subnet makes it easier to set up traffic filters to protect wired clients from an attack launched from the WLAN.	[redacted]
Redundant Wireless Access Points In both wired and wireless campus LAN architectures, redundancy is usually desirable to ensure high availability. For campus networks with WLANs that are mission critical, Cisco has a feature called Hot Standby mode that supports two access points being configured to use the same channel in a single coverage area. Only one of the access points is active. The standby access point passively monitors	[redacted]

the network and the primary access point. If the primary access point fails, the secondary access point takes over to provide coverage.	[redacted]
Note Don't confuse access-point Hot Standby mode with the Cisco Hot Standby Router Protocol (HSRP), which is covered in the upcoming "Workstation-to-Router Redundancy" section. Access-point Hot Standby mode addresses Layer 2 redundancy, whereas HSRP addresses Layer 3 redundancy.	[redacted]
You should place the standby access point near the access point it will monitor and give it the same configuration (except for a different IP address). The standby access point associates with the monitored access point as a client and queries the monitored access point regularly through both the Ethernet interface and the RF interface. If the monitored access point fails to respond, the standby access point becomes active, signals the primary access point radio to become inactive, and takes the	[redacted]

monitored access point's place in the network.	[redacted]
As soon as the primary access point failure is detected, user intervention is required. The user should return the backup access point to standby mode. Failure to reset the standby access point results in both the primary and standby access points operating concurrently on the same channel when the primary access point comes back online.	[redacted]
Redundancy and Load Sharing in Wired LANs	[redacted]
In wired campus networks, it is common practice to design redundant links between LAN switches. Most LAN switches implement the IEEE 802.1D spanning-tree algorithm to avoid network loops. The 802.1D standard is a good solution for redundancy, but not for load sharing, because only one path is active. Some switch vendors, including Cisco, let you have one spanning tree per VLAN, which can be used to implement redundancy. A switch can act as the root	[redacted]

bridge for one VLAN and as a backup for the root bridge for another VLAN.	[redacted]
Cisco Per VLAN Spanning Tree+ (PVST+) builds a separate logical tree topology for each VLAN. PVST+ allows load sharing by having different forwarding paths per VLAN. PVST+ is less scalable than the classic 802.1D method, where there is just one root and tree, because CPU time is required to process BPDUs for each VLAN. Cisco overcame this limitation with the Multi-Instance Spanning Tree Protocol (MISTP), which allows a set of VLANs to be grouped into a single spanning tree.	[redacted]
IEEE has also enhanced the original spanning-tree algorithm with its Multiple Spanning Trees (MST) standard, which is documented in IEEE 802.1s. The Multiple Spanning Tree Protocol (MSTP) uses RSTP for rapid convergence but improves	[redacted]

RSTP scalability by aggregating a set of VLAN-based spanning trees into distinct instances, and by running only one (rapid) spanning-tree algorithm per instance. This architecture provides multiple forwarding paths for data traffic, enables load sharing, and reduces the number of spanning trees required to support a large number of VLANs.	[redacted]
If you use VLANs in a campus network design with switches that support 802.1s, PVST+, or MISTP, redundant links can offer load sharing in addition to fault tolerance. Figure 5-11 shows a redundant campus LAN design that uses the spanning-tree algorithm and VLANs.	[redacted]
The design in Figure 5-11 takes advantage of the concept of one spanning tree per VLAN. Switch A acts as the root bridge for VLANs 2, 4, and 6. (Switch B can become the root bridge for those VLANs if Switch A fails.) Switch B acts as the root bridge for VLANs 3, 5, and 7. (Switch A can become the root bridge for those VLANs if Switch B	[redacted]

fails.) The result is that both links from an access layer switch carry traffic, and failover to a new root bridge happens automatically if one of the distribution layer switches fails. Both load sharing and fault tolerance are achieved.

The design in Figure 5-11 can scale to a large campus network. The design has been tested on a network that has 8000 users, 80 access layer switches, 14 distribution layer switches, and 4 core campus routers (not counting the routers going to the WAN).

Access Layer

Figure 5-11 Campus Hierarchical Redundant Topology

Server Redundancy

This section covers guidelines for server redundancy in a campus network design. File, web, Dynamic Host Configuration Protocol (DHCP), name, and database servers are all candidates for redundancy in a campus

design, depending on a customer's requirements. In a network that supports Voice over IP (VoIP), the servers that provide the mapping between a phone number and an IP address and handle call processing should be provisioned in a redundant fashion. Cisco Unified Communications Manager, for example, supports a redundancy group where servers are assigned the role of primary, secondary, or tertiary server. DHCP servers can be placed at the access, distribution, or core layer. For large, globally distributed networks, redundant DHCP servers are usually placed in the access layer. This avoids excessive traffic between the access and distribution or core layers and allows each DHCP server to serve a smaller percentage of the user population. If the core of the network is in New York, for example, and the access and distribution layers are spread out across the world, it makes sense to have DHCP servers in the access layer. For small networks, however, DHCP servers are often centrally located in the core of the network to facilitate management by a centralized IT department.	[redacted]

In large campus networks, the DHCP server is often placed on a different network segment than the end systems that use it. If the server is on the other side of a router, the router can be configured to forward DHCP broadcasts from end systems. The router forwards the broadcasts to a server address configured via the ip helper address command on a Cisco router. The router inserts the address of the interface that received the request into the giaddr field of the DHCP request. The server uses the giaddr field to determine from which pool of addresses to choose an address.	[redacted]
Name servers are in theory less critical than DHCP servers because users can reach services by address instead of name if the name server fails; because many users do not realize this, however, it is a good idea to plan for redundant name servers. Name servers implement the Internet Domain Name System (DNS), the	[redacted]

Windows Internet Naming Service (WINS), and the NetBIOS Name Service (NBNS). Name servers can be placed at the access, distribution, or core layer. For large, globally distributed networks, it makes sense to have name servers at the access layer. For small networks, however, it's more common to place name servers in the core of the network.	[redacted]
In any application where the cost of downtime for file servers is a major concern, mirrored file servers should be recommended. For example, in a brokerage firm where traders access data to buy and sell stocks, the data can be replicated on two or more mirrored file servers. Mirrored file servers hold identical data. Updates to the data are synchronized across the servers. The servers should be on different networks and power supplies to maximize availability.	[redacted]

If complete server redundancy is not feasible due to cost considerations, mirroring or duplexing of the file server hard drives is a good idea. (Duplexing is the same as mirroring with the additional feature that the two hard drives are controlled by different disk controllers.) Implementing a storage-area network (SAN) is another option. SANs are a popular solution for organizations seeking highly reliable, uninterrupted access to large amounts of stored information.	[redacted]
Redundancy has both availability and performance advantages. With mirrored file servers, it is possible to share the workload between servers. Using a content delivery network (CDN) and content services devices, users can be directed to one of many mirrored servers that all hold the same data.	[redacted]
Redundancy can also be	[redacted]

achieved by adding some sophistication to DNS. When a client requests access to a resource by its DNS name, a DNS server can return multiple host addresses in its response. Whether this will provide good redundancy depends on the host software. Some implementations try additional addresses if the first one doesn't respond.

Another possibility is a feature called DNS round robin, where the server has a list of addresses through which it cycles. The server gives out a different address with each request, going through its list of addresses. When it gets to the end of the list, it cycles back to the beginning of the list. Due to DNS caching, where clients and other DNS servers remember a previous name-to-address mapping, DNS round robin isn't perfect, but it can be quite simple to implement and configure on a typical DNS server.

Redundancy and load balancing with DNS can also work with multiple DNS servers. Assuming that clients

access different DNS servers, one server can respond with one address, while other servers respond with different addresses. Again, DNS caching can limit the effectiveness of this method.	[redacted]
Note There is one caveat to keep in mind with mirrored file, DHCP, web, and other types of servers. Mirrored servers offer redundancy for the hardware, cabling, LAN connection, and power supply, but they do not offer software or data redundancy. Because mirrored servers hold replicated data, if the problem is in the data or the software's capability to access the data, all the mirrored servers are affected.	[redacted]
Workstation-to-Router Redundancy	[redacted]
Workstations in a campus network must have access to a router to reach remote services.	[redacted]

Because workstation-to-router communication is critical in most designs, you should consider implementing redundancy for this function.

A workstation has many possible ways to discover a router on its network, depending on the protocol it is running and also the implementation of the protocol. The next few sections describe methods for workstations to learn about routers and redundancy features that guarantee a workstation can reach a router.

IP implementations vary in how they implement workstation-to-router communication. Some IP workstations send an ARP frame to find a remote station. A router running proxy ARP can respond to the ARP request with the router's data link layer address. Cisco routers run proxy ARP by default.

The advantage of depending on proxy ARP to reach remote stations is that a workstation doesn't have to be configured with the address of a router. However, because proxy ARP has never been standardized,

most network administrators don't depend on it. Also, many security experts recommend turning it off because it makes it easier for an attacker to reach another network. Instead, IP workstations are given the address of a default router. This can be manually configured or supplied by DHCP. A default router is the address of a router on the local segment that a workstation uses to reach remote services. The default router is usually called the default gateway for historical reasons.

An IP workstation usually knows the address of only one router: the default gateway. The result is that a workstation does not always use the most expedient method to reach a remote station. The workstation can select a path that includes an extra hop. Figure 5-12 shows the extra-hop problem. To get around the extra-hop problem and to add redundancy, some workstation IP implementations allow a network administrator to add static routes to a configuration file or to configure the

workstation to run a routing protocol. Note In UNIX environments, workstations sometimes run the RIP daemon to learn about routes. It is best if they run the RIP daemon in passive rather than active mode. In active mode, a workstation sends a RIP broadcast frame every 30 seconds. When many UNIX workstations run RIP in active mode, the amount of broadcast traffic can degrade network performance. In addition, there are security risks in allowing uncontrolled stations to run a routing protocol in active mode. Another alternative for IP workstation-to-router communication is the Router Discovery Protocol (RDP). RFC 1256 specifies the RDP extension to the Internet Control Message Protocol (ICMP). With RDP, each router periodically multicasts an ICMP router advertisement packet from each of its interfaces, announcing the IP address of that interface.	[redacted]

Workstations discover the addresses of their local routers simply by listening for advertisements, in a similar fashion to the method AppleTalk workstations use to discover the address of a router. The default advertising rate for RDP is once every 7 to 10 minutes, though, which is quite different from AppleTalk's default, which is once every 10 seconds.	[redacted]
When a workstation starts up, it can multicast an ICMP router solicitation packet to ask for immediate advertisements, rather than wait for the next periodic advertisement to arrive. RDP does not attempt to solve the extra-hop problem. Although most routers support RDP, few workstation IP implementations support it, so RDP is not widely used.	[redacted]
One reason that RDP has not become popular is that DHCP includes an option for a DHCP server to return the address of a default gateway to a client. As specified in RFC 2131, a server's response to a DHCP client's request for an IP address can include an options	[redacted]

field in which the server can place one or more default gateway addresses. A preference level can be used to specify which default gateway is the best option. The server can also include a list of static routes in the options field.	[redacted]
These days, most IP workstations are configured with the address of a default gateway The configuration can be done at each workstation or at a DHCP server that supports many workstations, which is the more common method. Running routing protocols or router discovery protocols at workstations has proven to be a poor alternative because of traffic and processing overhead, security issues, and the lack of implementations for many platforms.	[redacted]
The problem with a default gateway configuration is that it creates a single point of failure, particularly because many implementations keep track of only one default gateway. Loss of the default gateway results	[redacted]

in workstations losing connections to remote sites and being unable to establish new connections.

Hot Standby Router Protocol

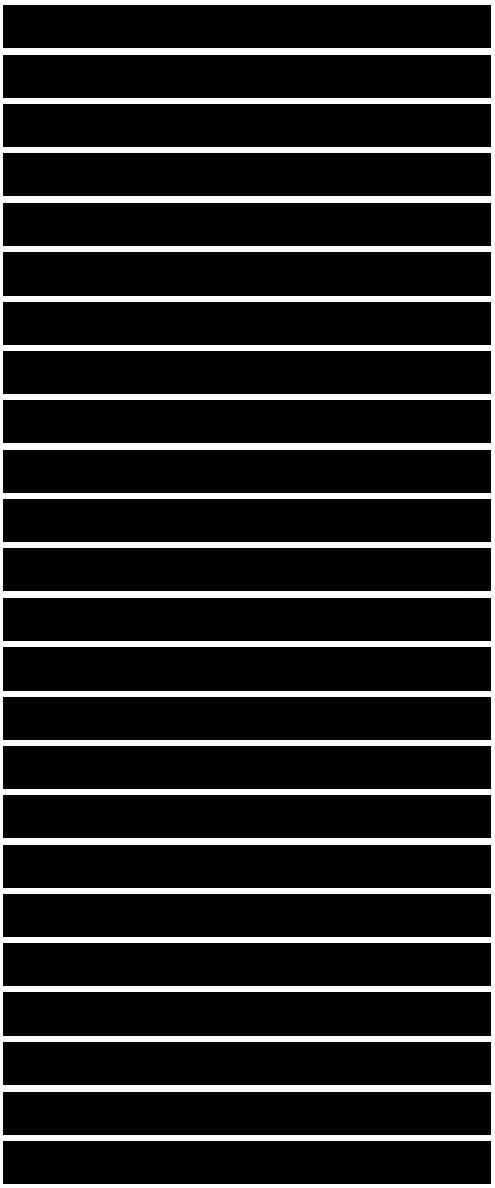
Cisco Hot Standby Router Protocol (HSRP) provides a way for an IP workstation to keep communicating on an internetwork even if its default gateway becomes unavailable. In RFC 2338, the IETF standardized a similar protocol, the Virtual Router Redundancy Protocol (VRRP). Routers in the core, distribution, or access layer can run HSRP or VRRP. The campus design shown in Figure 5-11 features HSRP at the core layer.

HSRP works by creating a virtual router, also called a phantom router, as shown in Figure 5-13. The virtual router has its own IP and MAC addresses. Each workstation is configured to use the virtual router as its default gateway. When a workstation broadcasts an ARP frame to find its default gateway, the active HSRP router responds with the virtual router's MAC address. If the active router goes offline, a standby router takes over as active router, continuing the

delivery of the workstation's packets. The change is transparent to the workstation.

Figure 5-13 Hot Standby Router Protocol (HSRP)

HSRP routers on a LAN communicate among themselves to designate an active and standby router. The active router sends periodic Hello messages. The other HSRP routers listen for the Hello messages. If the active router fails, causing the other HSRP routers to stop receiving Hello messages, the standby router takes over and becomes the active router. Because the new active router assumes both the IP and MAC addresses of the phantom, workstations see no change. They continue to send packets to the virtual router's MAC address, and the new active router delivers those packets. The Hello timer should be configured to be short enough so that workstation applications and protocols do not drop connections before the standby router becomes active.



HSRP also works for proxy ARP. When an active HSRP router receives an ARP request for a station that is not on the local network, the router replies with the virtual router's MAC address. If the router becomes unavailable, the new active router can still deliver the traffic.	[redacted]
Cisco also has a useful enhancement to HSRP, standby tracking, which monitors one or more WAN interfaces on a router that has HSRP enabled on the LAN interfaces. If the software senses a problem with the WAN circuit connected to one of the WAN interfaces that it is tracking, it fails over to an active WAN interface on a standby router. The default gateway, for which HSRP provides redundancy, is the user's method of getting outside the LAN and is often connected to a WAN interface that provides access to the rest of the intranet or the Internet, so the standby tracking feature	[redacted]

is extremely useful. Cisco also supports an HSRP-enabled router preserving Network Address Translation (NAT) and IPsec state information. WAN edge devices can maintain NAT translations and IPsec tunnels, used in VPNs usually, when HSRP switches to a different router. Gateway Load Balancing Protocol To achieve load sharing along with redundancy, Cisco also has a newer protocol, the Gateway Load Balancing Protocol (GLBP), which is similar, but not identical, to HSRP and VRRP. With HSRP and VRRP, the standby routers in a group are superfluous until the active router fails. These standby routers may have access to bandwidth that is wasted until a problem arises. Although multiple virtual router groups can be configured for the same set of routers, which is less wasteful, the hosts must be configured for different default gateways,	[redacted]

which results in an extra administrative burden. GLBP provides load balancing over multiple routers using a single virtual IP address and multiple virtual MAC addresses. Each host is configured with the same virtual IP address, and all routers in the virtual router group participate in forwarding packets. Members of a GLBP group elect one router to be the active virtual gateway (AVG) for that group. Other group members provide backup for the AVG in the event that the AVG becomes unavailable. The AVG assigns a virtual MAC address to each member of the GLBP group. Each gateway assumes responsibility for forwarding packets sent to the virtual MAC address assigned to it by the AVG. These gateways are known as active virtual forwarders (AVF) for their virtual MAC address. The AVG is responsible for answering ARP requests for the virtual IP address. Load sharing is achieved by the AVG replying to the ARP requests with different virtual MAC addresses.	[redacted]

Designing the Enterprise Edge Topology The enterprise edge consists of a WAN edge for connecting branch offices and an Internet edge for connecting to the public Internet via a service provider's edge infrastructure. The enterprise edge might also include an extranet edge for connecting partners and an e-commerce module for selling products. This section covers enterprise edge topologies that include redundant WAN segments, multihomed connections to the Internet, and VPNs. The section also includes a few comments about the service provider edge. Redundant WAN Segments Because WAN links can be critical pieces of an enterprise internetwork, redundant	[redacted]

(backup) WAN links are often included in an enterprise edge network topology. A WAN network can be designed as a full mesh or a partial mesh. A full-mesh topology provides complete redundancy. It also provides good performance because there is just a singlelink delay between any two sites. However, as already discussed in this chapter, a full mesh is costly to implement, maintain, upgrade, and troubleshoot. A hierarchical partial- mesh topology, as shown previously in Figure 5-4, is usually sufficient.

Circuit Diversity

When provisioning backup WAN links, you should learn as much as possible about the actual physical circuit routing. Different carriers sometimes use the same facilities, meaning that your backup path is susceptible to the same failures as your primary path. You should do some investigative work to ensure that your backup actually is a

backup. Network engineers use the term circuit diversity to refer to the optimum situation of circuits using different paths. Because carriers lease capacity to each other and use third-party companies that provide capacity to multiple carriers, it is getting harder to guarantee circuit diversity. Also, carriers often merge with each other and mingle their circuits after the merge. As carriers increasingly use automated techniques for physical circuit rerouting, it becomes even more difficult to plan diversity because the rerouting is dynamic. \ Nonetheless, you should work with the providers of your WAN links to gain an understanding of the level of circuit diversity in your network design. Carriers are usually willing to work with customers to provide information about physical circuit routing. Be aware, however, that carriers sometimes provide inaccurate information, based on databases that are not kept current. Try to write circuit-	[redacted]

diversity commitments into contracts with your providers.	[redacted]
When analyzing circuit diversity, be sure to analyze your local cabling in addition to your carrier's services. Perhaps you have designed an ISDN link to back up a Frame Relay link. Do both of these links use the same cabling to get to the demarcation point in your building network? What cabling do the links use to get to your carrier? The cabling that goes from your building to the carrier is often the weakest link in a network. It can be affected by construction, flooding, ice storms, trucks hitting telephone poles, and other factors.	[redacted]
Multihoming the Internet Connection The generic meaning of multihoming is to "provide more than one connection for a system to access and offer network services." The term multihoming is also used in many specific ways. A server, for example, is said to be multihomed if it has more than	[redacted]

one network layer address. Content delivery networks can multihome application layer data and services. The term multihoming is increasingly used to refer to the practice of providing an enterprise network with more than one entry into the Internet. Redundant entries into the Internet provide fault tolerance for applications that require Internet access. An enterprise network can be multihomed to the Internet in many different ways, depending on a customer's goals. Figure 5-14 and Table 5-2 describe some methods for multihoming the Internet connection. In the case of Options C and D, the goal might be to improve network performance by allowing European enterprise sites to access the Internet using the Paris router and North American sites to use the New York router. This can be accomplished by correctly configuring a default gateway on end stations and a default route on enterprise routers in Europe and North America. A default route specifies where a packet should go if there is no explicit entry for the destination network in a	[redacted]

router's routing table. The default route is also sometimes called the gateway of last resort.

Your customer might have more complex goals than the simple goal in the previous paragraph. Perhaps your customer wants to guarantee that European enterprise sites access North American Internet sites via the New York router. A parallel goal is that North American enterprise sites access European Internet sites via the Paris router. This could be a reasonable goal when a constant, low latency is required for an application. The latency is more predictable if the first part of the path is across the enterprise intranet instead of the Internet. This goal is harder to meet than the first goal, however. It requires that the enterprise routers understand routes from the ISP and set preferences on those routes.

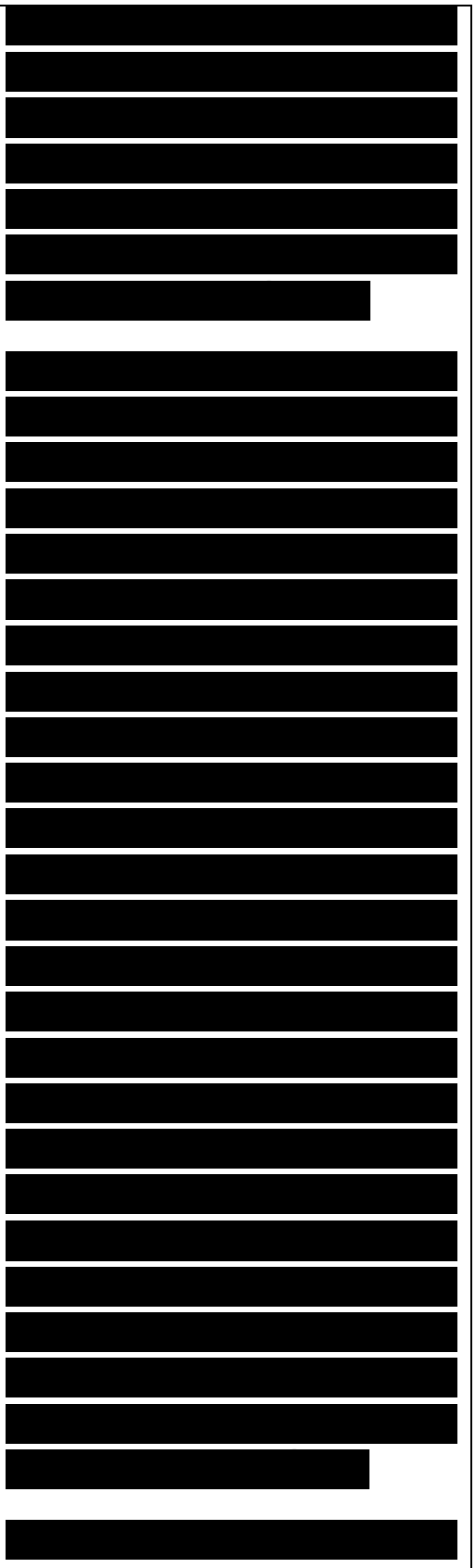


Figure 5-14 Options for Multihoming the Internet

Connection A related goal is to use the "best route" across the Internet to the sites that the enterprise users rely on the most. Unless an enterprise contracts (and pays) for end-to-end managed QoS, this goal cannot be met. The routing protocol used on the Internet, BGP, doesn't offer route optimality. Its only purpose is to provide reachability and stability in the global routing system. Intermediate providers with whom an enterprise has no business relationship don't care if the enterprise's traffic follows optimal routes, nor do they have any incentive to do so. Another, more complex goal is to guarantee that incoming traffic from the Internet destined for European enterprise sites uses the Paris router and incoming traffic for North American enterprise sites uses the New York router. This goal requires the enterprise routers to advertise to the Internet routes to	[redacted]

enterprise sites. The routes must include metrics so that routers on the Internet know the preferred path to sites on the enterprise intranet.

One other caveat when an enterprise network is multihomed is the potential to become a transit network that provides interconnections for other networks. Referring to Figure 5-14, consider that the enterprise router learns routes from the ISP. If the enterprise router advertises these learned routes, it risks allowing the enterprise network to become a transit network and being loaded by unintended external traffic. When an enterprise network becomes a transit network, routers on the Internet learn that they can reach other routers on the Internet via the enterprise network. To avoid this situation, enterprise routers should advertise only their own routes. Alternatively they can run without a routing protocol and rely on default and static routing.

In general, multihoming the Internet connection can be challenging if a customer's goals are complex. Encourage your customers to simplify their goals to ensure ease of implementation, scalability, availability, and affordability. If the main goal is high availability, don't assume that this means more redundancy is required. According to Howard Berkowitz in his book WAN Survival Guide, "Uncontrolled increases in redundancy lead to uncontrolled increases in complexity, and may actually decrease availability."

Virtual Private Networking

Virtual private networks (VPN) use advanced encryption and tunneling to permit organizations to establish secure, end-to-end, private network connections over a third- party network. The third-party network can be a private service provider network or the public Internet. An organization can connect to the

third-party network using a variety of WAN and remote-access technologies, including leased lines, Frame Relay, cable modems, digital subscriber line (DSL), analog modems, ISDN, and so on. Organizations can also use VPNs to connect outside users, such as business partners, customers, resellers, and suppliers. VPNs also support mobile users and telecommuters. Point-to-point connectivity across the third-party network is typically provided by a tunneling protocol. Tunneling is a technique for encapsulating packets of one protocol inside another protocol. For example, a tunnel can carry IPv4 packets across an internetwork that supports only IPv6. In the context of a VPN, tunneling is used to encapsulate private messages and apply encryption algorithms to the payload. Tunnels provide a logical, point-to-point connection	[redacted]

across a connectionless IP network, enabling application of advanced security features. Encryption is applied to the tunneled connection to scramble data, thus making data legible only to authorized systems. In applications where security and privacy are less of a concern, tunnels can be used without encryption to provide multiprotocol support.	[redacted]
Layer 2 tunneling methods encapsulate at the data link layer of the OSI model. Examples include Point-to-Point Tunneling Protocol (PPTP), Layer 2 Forwarding (L2F), MPLS VPNs, and Layer 2 Tunneling Protocol (L2TP). L2TP is an IETF standard (RFC 2661) that many vendors support for their VPN solutions, including Cisco and Microsoft. The IETF is also developing a new version of L2TP, called L2, which is emerging as a lightweight yet robust solution for Layer 2 tunneling.	[redacted]
Layer 3 tunneling encapsulates at the network layer. Two examples are IPsec and Cisco generic routing encapsulation	[redacted]

(GRE). If only IP-unicast packets are being tunneled, IPsec is the best choice. GRE is used when multicast, broadcast, and non-IP packets need to be tunneled.

VPN applications for enterprise networks can be divided into two main categories:

- Site-to-site VPNs: Site-to-site VPNs focus on connecting geographically dispersed offices and extending the classic enterprise WAN. A site-to-site VPN can also add interconnections between multiple organizations, in which case it is sometimes called an extranet VPN.

- Remote-access VPNs: Remote-access VPNs focus on remote users and business partners who access the network on an as-needed basis.

The sections that follow describe these two types of VPNs in greater detail.

Site-to-Site VPNs

Site-to-site VPNs have emerged as a relatively inexpensive way for a company to connect geographically

dispersed branch offices and home offices via a service provider or the Internet, as opposed to maintaining an expensive private WAN. The company's private data can be encrypted for routing through the service provider's network or the Internet. Traditionally, businesses relied on private 1.544-Mbps T1 leased lines to link remote offices together. Leased lines are expensive to install and maintain. For many companies, a leased line provides more bandwidth than is needed at too high a price. Companies also used Frame Relay and point-to-point networks for their private WANs, but they were also somewhat expensive and hard to manage. A site-to-site VPN is a more cost-effective and manageable solution.	[redacted]
When designing the topology of a site-to-site network, you should consider the same needs that you would for a private WAN, including the need for high availability with automatic failover, performance, security, and	[redacted]

scalability. The most common topologies for a site-to-site VPN are as follows:

- Hub-and-spoke
- Mesh
- Hierarchical network

The hub-and-spoke topology is used when there is a single regional or headquarters location with many remote offices, and most traffic is between the remote sites and the regional or headquarters location. This design minimizes configuration complexity by having a single IPsec connection or a single GRE tunnel from each remote location back to the regional or headquarters location. This design isn't appropriate when there is a high level of traffic between remote sites or when there is a need for redundancy and automatic failover. An enhancement to the design is to include multiple VPN routers at headquarters to provide better redundancy.

Mesh VPN designs can either be fully meshed, providing any-to-any connectivity, or partially meshed, providing

some-to-some connectivity, depending upon customer requirements. The meshed topology is a good design to use when there are a small number of total locations (regional, headquarters, or remote locations), with a large amount of traffic flowing between some (partial mesh) or all (full mesh) of the sites. In a fully meshed design, the loss of a single location affects only traffic to or from that location. All other locations remain unaffected. This design does not scale well when there are numerous sites, due to the large number of IPsec connections or GRE tunnels with IPsec that have to be configured on each device. A hierarchical VPN topology is a hybrid topology for a large company that has many headquarters and regional offices with a lot of traffic flowing between them, and many remote offices, with little interaction between them. The topology consists of a full- or partial-mesh core, with peripheral sites connecting into the core using a hub-and-spoke design. A hierarchical design is the most complex of the designs in terms of	[redacted]

configuration, and might have a combination of IPsec and GRE tunnels.

Remote-Access VPNs

Remote-access VPNs permit on-demand access to an organization's internetwork, via secure, encrypted connections. Mobile or remote users, and branch offices that don't need to always be connected, can access their corporate networks via a third-party network, such as a service provider's network or the Internet. Enterprises use remote-access VPNs to reduce communications expenses by leveraging the local infrastructures of service providers who support dialup, ISDN, cable modem, DSL, or wireless access to the Internet or the provider's private network.

When implementing a remote-access VPN architecture, an important consideration is where to initiate tunneling and encryption. Should the tunnel initiate on the client PC or on a network access server (NAS) operated by the VPN service provider? In a client-initiated model, the encrypted tunnel is established by client software

using IPsec, L2TP, or PPTP, thereby making the service provider network solely a means of transport to the corporate network. An advantage of a client-initiated model is that the "last mile" service provider access network used for accessing the provider point of presence (POP) is secured. A disadvantage of the client-initiated model is the need to manage software on client machines. In a NAS-initiated model, a remote user accesses a service provider's POP, is authenticated by the service provider, and, in turn, initiates a secure tunnel to the corporate network from the POP. With a NAS-initiated architecture, VPN intelligence resides in the service provider network. There is no end-user client software for the organization to maintain, thus eliminating client management issues associated with remote access. The drawbacks, however, are lack of security on the local access network connecting the client to the service provider network and the need to interoperate with servers operated by the	[redacted]

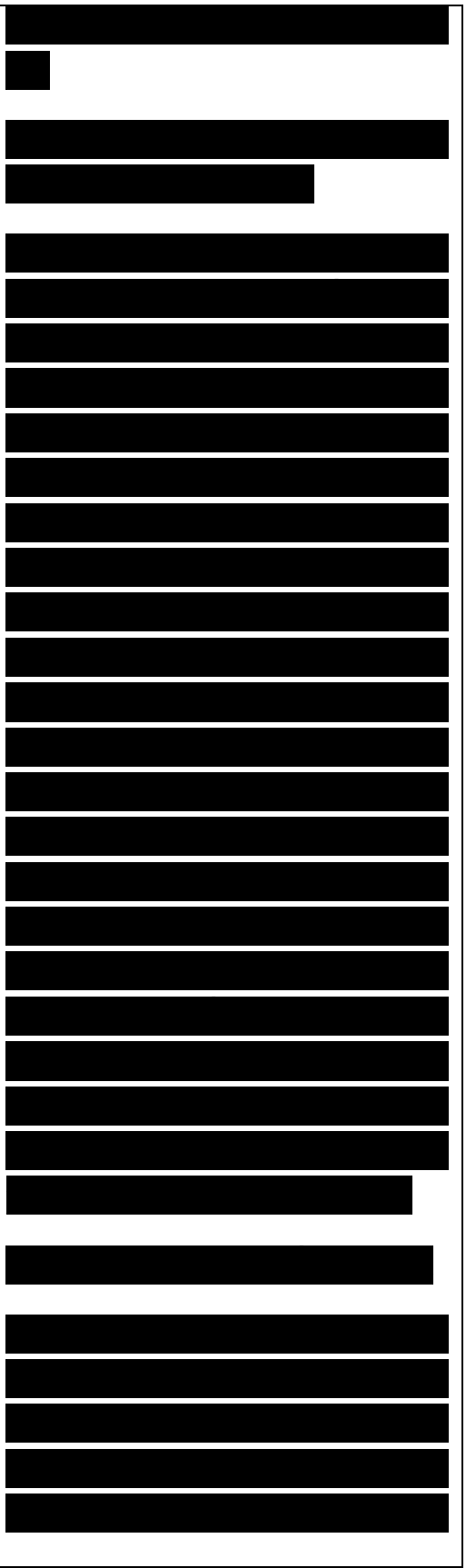


Figure 5-15 Remote-Access VPN for a Retail Company

At the headquarters, user VPN connections terminate in the Remote-Access VPN section of the Internet edge module, according to the Cisco SAFE architecture. Cisco recommends that one or more VPN concentrators reside within this area. A VPN concentrator is a dedicated hardware platform that aggregates a large volume of simultaneous VPN connections. Generally, enterprises place the concentrator between a router that has access to the VPN and a router that forwards traffic into the campus network. The Remote-Access VPN section also includes an authentication, authorization, and accounting (AAA) server and an intrusion detection system (IDS) or intrusion prevention system (IPS).

Service Provider Edge

Although the focus of this chapter is designing a logical topology for an enterprise network, a quick discussion of service providers is warranted at this point. An enterprise network connects to a module

that can be called the service provider edge, and, although you aren't expected to design this module as an enterprise network designer, you need to have some understanding of it and be able to select the appropriate provider (or providers) for your design customers. The selection of a service provider is something you should consider during the logical design phase, which is the focus of Part II. Part III addresses the topic again because during that phase you should make some definite selections of WAN technologies, devices, and providers. In the early days of data communications, there were the regional or national telephone companies and their customers, and nothing else. A customer's choice of provider was dictated by location. The level of service and pricing was dictated by the provider. Today, there is a broad range of service providers and service levels, and pricing is more negotiable. Finding a provider that matches an enterprise's needs requires a	[redacted]

good understanding of those needs and the culture of the enterprise and the potential provider. Many ISPs are small startups that offer wireless and cable modem service to end users. These ISPs might not have the expertise to support a large enterprise, although they might be appropriate for home users accessing the corporate network using VPN software. Some ISPs focus mostly on hosting servers and don't support end users. Some ISPs are actually network service providers (NSP), which means that their main business is connecting other ISPs rather than enterprises or end users. Selecting providers for your network design requires you to understand which of these types of services you actually need. ISPs and NSPs are sometimes classified as being Tier 1 through Tier 5. Although these categories don't have universal meaning, if a provider calls itself a Tier 1 provider and you are looking for an inexpensive provider to connect a small	[redacted]

office or home, you know to look elsewhere. Tier 1 ISPs are large, international providers, whereas Tier 5 ISPs are small, specialized providers, sometimes located in a town or rural area. A Tier 5 provider could be as small as an Internet cafe.	[redacted]
One important difference between the tiers has to do with the relationship a provider has with other ISPs. Using an economic definition of peer (rather than the BGP definition), a peer relationship means that two ISPs do not charge each other to carry each other's traffic. They are both about the same size and it is to their mutual advantage to let their customers have access to each other, without worrying about billing. This differs from the other common ISP relationship, which is a provider-customer one, where a smaller ISP pays a larger ISP for the privilege of sending traffic through the larger ISP's network. This is often called buying transit.	[redacted]
A Tier 1 provider doesn't buy transit. A Tier 1 provider has a	[redacted]

24/7 network operations center and a national or international backbone with at least DS-3 connectivity, and more likely OC-3 to OC-48. The provider gets all its routes from bilateral peering arrangements. Its customers are primarily other providers, but it might support a large enterprise also. Examples of Tier 1 providers include Verizon, Cable & Wireless, British Telecom, Verio, Level 3, and AT&T. Tier 2 providers also have high-bandwidth backbones and 24/7 operations, but they are limited to a regional or national presence, and they buy transit (often at a bulk discount) from a Tier 1 provider for traffic that goes outside the region. A Tier 2 provider gets all its regional routes through peering arrangements.	[redacted]
A Tier 3 provider is typically a regional provider for a small or medium-sized region. A Tier 3 provider buys transit from multiple upstream providers and runs a default-free routing	[redacted]

table. There's no general definition of Tier 4 or Tier 5, but Tier 4 could be a metropolitan provider that is multihomed to two regional providers, and Tier 5 might be a small, single-homed provider that connects end users via a wireless or cable modem service. At this point in the design process, you should have analyzed requirements and topologies to the extent that you have a good idea of the tier you will need. During the logical design phase, you should start making a list of criteria for selecting providers and develop a plan and set of standards for evaluating candidates. Investigate the availability of service providers in the relevant regions and start making inquiries. Be specific in your requests for information from candidates. Prioritize the requested information and indicate how quickly you need a response. You may also want to ask for references and to start asking questions about the provider of other users in the region. See the "Selecting a WAN Service	[redacted]

Provider" section in Chapter 11, "Selecting Technologies and Devices for Enterprise Networks," for more information on this topic.	[redacted]
Secure Network Design Topologies This section discusses network security in relation to network topologies. Chapter 8 covers network security in more detail. The focus of this section is logical topologies, but physical security is also briefly mentioned.	[redacted]
Planning for Physical Security When developing the logical topology of a network, you should begin to get an idea of where equipment will be installed. You should start working with your design customer right away to make sure that critical equipment will be installed in computer rooms that have protection from unauthorized access, theft, vandalism, and natural disasters such as floods, fires, storms, and earthquakes. Physical security is not really an aspect of logical network design, but it is mentioned here	[redacted]

because your logical topology might have an impact on it, and because the planning for physical security should start right away, in case there are lead times to build or install security mechanisms.

Meeting Security Goals with Firewall Topologies

A firewall is a system or combination of systems that enforces a boundary between two or more networks. A firewall can be a router with ACLs, a dedicated hardware box, or software running on a PC or UNIX system. A firewall should be placed in the network topology so that all traffic from outside the protected network must pass through the firewall. A security policy specifies which traffic is authorized to pass through the firewall.

Firewalls are especially important at the boundary between the enterprise network and the Internet. A basic firewall topology is simply a router with a WAN connection to the Internet, a LAN connection to the enterprise network, and software that has

security features. This elementary topology is appropriate if your customer has a simple security policy. Simple security policies can be implemented on the router with ACLs. The router can also use NAT to hide internal addresses from Internet hackers. For customers with the need to publish public data and protect private data, the firewall topology can include a public LAN that hosts web, FTP, DNS, and SMTP servers. Older security literature often referred to the public LAN as the free-trade zone, which is a good name for it. Unfortunately, the less apropos term demilitarized zone (DMZ) has become more popular. Security literature refers to a host in the DMZ as a bastion host, a secure system that supports a limited number of applications for use by outsiders. The bastion host holds data that outsiders can access, such as web pages, but is strongly protected from outsiders using it for anything other than its limited purposes. For larger customers, it is recommended that you use a dedicated firewall in addition to a router between the Internet and the enterprise network. To	[redacted]

maximize security, you can run security features on the router and on the dedicated firewall. (To maximize performance, on the other hand, you would not run security features on the router.) Figure 5-16 shows a DMZ secure topology.

Web, File, DNS, Mail Servers

Figure 5-16 DMZ Topology

An alternative topology is to use two routers as the firewalls and place the DMZ between them, as shown in Figure 5-17. This topology is called a three-part firewall topology. A disadvantage with this approach is that the configuration on the routers might be complex, consisting of many ACLs to control traffic in and out of the private network and the DMZ. Another disadvantage is that traffic for the enterprise network flows through the DMZ. The DMZ connects public servers that can

be compromised and act as launching pads for attacks into the enterprise network. You can strengthen this topology by using routers with simple ACLs at either end of the DMZ and also including firewalls at either end that are configured with more complex ACLs. Also, the bastion hosts inside the DMZ should run firewall software and be configured for a limited set of services.

Summary

This chapter focused on techniques for developing a topology for a network design. Designing a network topology is the first step in the logical design phase of the top- down network design methodology. By designing a logical topology before a physical implementation, you can increase the likelihood of meeting a customer's goals for scalability, adaptability, and performance.

Web, File, DNS, Mail Servers

Figure 5-17 Three-Part Firewall Topology

This chapter discussed four characteristics of network topologies:

- Hierarchy
- Modularity
- Redundancy
- Security

All of these characteristics can be applied to both campus and enterprise WAN design. The characteristics are not mutually exclusive. Your goal should be to design hierarchical, modular, redundant, and secure network architectures based on your customer's goals.

Hierarchy and modularity let you develop a network consisting of many interrelated components in a layered and structured fashion. Using a hierarchical model can help you maximize network performance, reduce the time to implement and troubleshoot a design, and minimize costs.

Redundant network designs let you meet requirements for network availability by duplicating network components. Redundancy eliminates single points of

failure on the network. Redundancy also facilitates load sharing, which increases network performance. Redundancy adds complexity and cost to the network, however, and should be designed with care.	[redacted]
Depending on your particular network design, you should plan a secure topology that protects core routers, demarcation points, cabling, switches, servers, and so on. Adding one or more firewalls to your topology can help you protect enterprise networks from outside attackers.	[redacted]
After completing a logical topology for a customer, you should continue in the logical design phase by designing network addressing and naming models, selecting switching and routing protocols, and developing network security and management strategies.	[redacted]
These topics are covered in the next few chapters. Doing a thorough job in the logical design phase can ease your transition into the design of the physical implementation of the network. It can also prepare	[redacted]

you for the job of selecting the right products and technologies for your customer.	[redacted]

Chapter 6

Designing Models for Addressing and Numbering

This chapter provides guidelines for assigning addresses and names to internetwork components, including networks, subnets, routers, servers, and end systems. The chapter focuses on Internet Protocol (IP) addressing and naming. To benefit most from this chapter, you should already have a basic understanding of IP addressing.

This chapter illustrates the importance of using a structured model for network layer addressing and naming. Without structure, it is easy to run out of addresses, waste addresses, introduce duplicate addresses and names, and use addresses and names that are hard to manage. To meet a customer's goals for scalability, performance, and manageability, you should assign addresses and names systematically.

This chapter also demonstrates the importance of developing policies and procedures for addressing and naming. Policies often involve a plan for distributing authority for addressing and naming to avoid one department having to manage all addresses and names. A central authority can assign blocks of addresses and names in a hierarchical fashion to departments and branch offices.

Part I of this book, “Identifying Your Customer’s Needs and Goals,” recommends gaining an understanding of your customer’s organizational structure (for example, departments, branch offices, business units, and so on). This information is helpful when planning addresses and names. A topology map of the network is also useful, because it helps you see the hierarchy in the network and recognize where address boundaries exist. The addressing and naming step of the top-down network design process falls here in the methodology because it makes use of information you gathered in the previous phases of the network design process.

The addressing and naming step precedes selecting routing and switching protocols, which is covered in the next chapter, because the addressing model you develop might dictate which routing protocols you can select. (For example, some routing protocols do not support variable-length subnet masking [VLSM].)

đã trình bày những phương pháp để tìm hiểu

định địa chỉ và đặt tên

các giao thức định tuyến và chuyển mạch

Guidelines for Assigning Network Layer Addresses

Network layer addresses should be planned, managed, and documented. Although an end system can learn its address dynamically, no mechanisms exist for assigning network or subnet numbers dynamically. These numbers must be planned and administered. Many vintage networks still exist where addressing was not planned or

documented. These networks are hard to troubleshoot and do not scale.	[redacted]
The following list provides some simple rules for network layer addressing that can help you architect scalability and manageability into a network design. These rules are described in more detail in later sections of this chapter.	[redacted]
■ Design a structured model for addressing before assigning any addresses. ■ Leave room for growth in the addressing model. If you do not plan for growth, you might later have to renumber many devices, which is labor-intensive.	[redacted] quá trình [redacted] tốn
■ Assign blocks of addresses in a hierarchical fashion to foster good scalability and availability.	[redacted] khả năng sẵn sàng hoạt động tốt
■ Assign blocks of addresses based on the physical network, not on group membership, to avoid problems when groups or individuals move.	[redacted]
■ If the level of network management expertise in regional and branch offices is high, you can delegate authority for addressing regional and branch-office networks, subnets, servers, and end systems.	[redacted]
■ To maximize flexibility and minimize configuration, use dynamic	[redacted]

addressing for end systems.

■ To maximize security and adaptability, use private addresses with Network Address Translation (NAT) in IP environments.

Using a Structured Model for Network Layer Addressing

A structured model for addressing means that addresses are meaningful, hierarchical, and planned. IP addresses that include a prefix and host part are structured. Assigning an IP network number to an enterprise network, then subnetting the network number and subnetting the subnets, is a structured (hierarchical) model for IP addressing.

A clearly documented structured model for addressing facilitates management and troubleshooting. Structure makes it easier to understand network maps, operate network management software, and recognize devices in protocol analyzer traces and reports. Structured addresses also facilitate network optimization and security because they make it easier to implement network filters on firewalls, routers, and switches.

A lot of companies have no model for addressing. When there is no model, and addresses are assigned in a haphazard way, the following problems can occur:

■ Duplicate network and host addresses	bị
■ Illegal addresses that cannot be routed on the Internet	Những
■ Not enough addresses in total, or by group	
■ Addressees that cannot be used, and so are wasted	
Administering Addresses by a Central Authority	
A corporate Information Systems (IS) or enterprise networking department should develop a global model for network layer addressing. As the network designer, you should help the IS department develop the model. The model should identify network numbers for the core of the enterprise network and blocks of subnets for the distribution and access layers. Depending on the organizational structure of the enterprise, network managers within each region or branch office can further divide the subnets.	để hiệu
IP addresses are either public or private. Public addresses are globally unique and are registered with a numbering authority. Private addresses are never routed on the global Internet and are assigned from a special range, documented in RFC 1918. Private addresses are covered in more detail in the section “Using Private Addresses in an IP Environment” later in this chapter.	

Early in the addressing design process, you need to answer the following questions about public versus private addresses:

- Are public, private, or both address types required?
- How many end systems need access to the private network only?
- How many end systems need to be visible to the public network?
- How will translation between private and public addresses occur?
- Where in the network topology will the boundary between private and public addresses exist?

The Internet Assigned Numbers Authority (IANA) is responsible for the global coordination of public IP addresses. The IANA allocates IP addresses to Regional Internet Registries (RIR). If you need a large number of public addresses, you will work with one of five RIRs:

- American Registry for Internet Numbers (ARIN) serves North America and parts of the Caribbean. Go to www.arin.net for more information.
- RIPE Network Coordination Centre (RIPE NCC) serves Europe, the Middle East, and Central Asia. Go

to www.ripe.net for more information.	[redacted]
■ Asia-Pacific Network Information Centre (APNIC) serves Asia and the Pacific region. Go to www.apnic.net for more information.	[redacted]
■ Latin American and Caribbean Internet Addresses Registry (LACNIC) serves Latin America and parts of the Caribbean. Go to www.lacnic.net for more information.	[redacted]
■ African Network Information Centre (AfriNIC) serves Africa. Go to www.afrinic.net for more information.	[redacted]
The term provider-independent address space refers to addresses that are assigned directly by one of the RIRs. In practice, most enterprises do not use addresses from the provider-independent address space. To become eligible for provider-independent address space, an organization must demonstrate to the RIRs that it will have thousands of Internet-connected hosts. Therefore, most enterprises work with an Internet service provider (ISP) to obtain public addresses, in which case their addresses are part of the provider-assigned address space. The enterprise uses these addresses for as long as it remains a subscriber of the provider. Changing to a new provider requires renumbering, which is one problem with provider-assigned addresses. Nevertheless, unless you have numerous hosts that need public	[redacted]

addressing, you will probably use provider-assigned addresses.

việc
khó khăn đi kèm với

Distributing Authority for Addressing

One of the first steps in developing an addressing and naming model is to determine who will implement the model. Which network administrators will actually assign addresses and configure devices? If addressing and configuration will be carried out by inexperienced network administrators, you should keep the model simple.

thiết bị

If there is a shortage of network administrators (which occurs in many organizations), simplicity and minimizing the amount of configuration required is important. In these situations, dynamic addressing is a good recommendation. Dynamic addressing mechanisms, such as the Dynamic Host Configuration Protocol (DHCP) for IP environments, allow each end system to learn its address automatically. Little, if any, configuration is necessary.

chúng ta nên dùng cơ chế định địa chỉ động

If network administrators in regional and branch offices are inexperienced, you might consider not delegating authority for addressing and naming. A lot of small and mediumsized companies maintain strict control of addressing and naming at a corporate (centralized) level. Maintaining strict

control avoids mistakes that can cause user frustration and network failures. Maintaining strict control can be challenging, especially as regional network administrators and users become more experienced with networking and start installing devices that can assign addresses.

Using Dynamic Addressing for End Systems

Dynamic addressing reduces the configuration tasks required to connect end systems to an internetwork. Dynamic addressing also supports users who change offices frequently, travel, or work at home occasionally. With dynamic addressing, a station can automatically learn the network segment to which it is currently attached, and adjust its network layer address accordingly.

Dynamic addressing was built into legacy desktop protocols such as AppleTalk and Novell NetWare. The designers of these protocols recognized the need to minimize configuration tasks so that inexperienced users could set up small internetworks. The IP protocols, on the other hand, were designed to run on computers managed by experienced system administrators and did not originally support dynamic addressing. In recent years, however, the importance of dynamic addressing has been recognized, and most companies use DHCP to minimize configuration tasks for IP end systems.

[redacted]

những sự cố

[redacted]

phân

[redacted]

sẵn vào

trong các cũ

[redacted]

Mặt khác,

[redacted]

Many networks use a combination of static and dynamic addressing. Static addresses are typically used for servers, routers, and network management systems. Static addresses are also used at the enterprise edge in the e-commerce, Internet edge, VPN/remote-access, and WAN edge modules of a modular network design. Although switches work below the network layer at the data link and physical layers, it is also a good idea to assign a static network layer address to switches for management purposes. Dynamic addresses are typically used for end systems, including workstations and IP phones.

Other criteria for using static versus dynamic addressing include the following:

■ The number of end systems: When there are more than 30 systems, dynamic addressing is usually preferable.

■ Renumbering: If it is likely you will need to renumber systems in the future and there are many end systems, dynamic address assignment is the better choice. Renumbering for public addresses will become necessary if a new ISP is selected. In addition, you might plan to renumber because the current plan is not well structured or will run out of numbers soon.

■ High availability: Statically assigned IP addresses are available

(phía)
mạng
các tầng
cho các
Một số tiêu chuẩn khác để lựa chọn giữa định địa chỉ tĩnh và định địa chỉ động là
ấn định
các
tại chưa có
những
địa chỉ IP được ấn định là tĩnh có thể

anytime. Dynamically assigned IP addresses have to be acquired from a server first. If the server fails, an address cannot be acquired. To avoid this problem, you can deploy redundant DHCP servers or use static addresses.

■ Security: With dynamic address assignment, in most cases, any device that connects to the network can acquire a valid address. This imposes some security risk and might mean that dynamic addressing is not appropriate for a company with a strict security policy.

■ Address tracking: If a management or security policy requires that addresses be tracked, static addressing might be easier to implement than dynamic addressing.

■ Additional parameters: If end systems need information beyond an address, dynamic addressing is useful because a server can provide additional parameters to clients along with the address. For example, a DHCP server provides a subnet mask, a default gateway, and optional server information such as the address of a TFTP server for VoIP and the address of one or more name servers, including Domain Name System (DNS) and Windows Internet Naming Service (WINS) servers.

IP Dynamic Addressing

When the IP protocols were first developed, a network administrator

was required to configure each host with its unique IP address. In the mid-1980s, protocols were developed to support diskless stations dynamically learning an address, which was necessary because a diskless station has no storage for saving a configuration. These protocols included the Reverse Address Resolution Protocol (RARP) and BOOTP. BOOTP has evolved into DHCP, which has gained considerable popularity since the late 1990s. RARP is limited in scope; the only information returned to a station using RARP is its IP address. BOOTP is more sophisticated than RARP and optionally returns additional information, including the address of the default gateway, the name of a boot file to download, and 64 bytes of vendor-specific information. Dynamic Host Configuration Protocol DHCP is based on BOOTP, which hosts can interoperate with DHCP hosts, although DHCP adds many enhancements to BOOTP, including a larger vendor-specific information field (called the options field in DHCP) and the automatic allocation of reusable network layer addresses. DHCP has bypassed BOOTP in popularity because it is easier to configure. Unlike BOOTP, DHCP does not require a network administrator to maintain a MAC-to-IP address table.	Giao thức phân giải địa chỉ ngược trả về thông tin bổ sung theo kiểu tùy chọn cổng đặc trưng cho từng nhà cung cấp tương tác hổ biến hơn so với BOOTP đòi hỏi

DHCP uses a client/server model. Servers allocate network layer addresses and save information about which addresses have been allocated. Clients dynamically request configuration parameters from servers. The goal of DHCP is that clients should require no manual configuration. In addition, the network manager should not have to enter any per-client configuration parameters into servers.

DHCP supports three methods for IP address allocation:

- Automatic allocation: A DHCP server assigns a permanent IP address to a client.
- Dynamic allocation: A DHCP server assigns an IP address to a client for a limited period of time.
- Manual allocation: A network administrator assigns a permanent IP address to a client, and DHCP is used simply to convey the assigned address to the client. (Manual allocation is rarely used because it requires per-client configuration, which automatic and dynamic allocations do not require.)

Dynamic allocation is the most popular method, partly because its reallocation feature supports environments where hosts are not online all the time, and there are more hosts than addresses. With dynamic allocation, a client requests the use of

an address for a limited period of time. The period of time is called a lease. The allocation mechanism guarantees not to reallocate that address within the requested time, and attempts to return the same network layer address each time the client requests an address. The client can extend its lease with subsequent requests. The client can choose to relinquish its lease by sending a DHCP release message to the server. The allocation mechanism can reuse an address if the lease for the address has expired. As a consistency check, the allocating server should probe the address before allocating the address by sending an Internet Control Message Protocol (ICMP) echo request (also known as a ping packet) to the address. The client should also probe the newly received address to make sure it is not currently in use by sending a ping packet to the address or by sending an Address Resolution Protocol (ARP) request for the address. If there is a reply, the address is already in use and should not be used by the client. When a client boots, it broadcasts a DHCP discover message on its local subnet. A station that has previously received a network layer address and lease can include them in the DHCP discover message to suggest that they be used again. A router can pass the	[redacted]

DHCP discover message on to DHCP servers not on the same physical subnet to avoid a require-ment that a DHCP server resides on each subnet. (The router acts as a DHCP relay agent.)	[redacted]
Each server responds to the DHCP request with a DHCP offer message that includes an available network layer address in the your address (yiaddr) field. The DHCP offer message can include additional configuration parameters in the options field such as the subnet mask and the addresses of a default gateway, a DNS server, or a TFTP server.	[redacted]
After the client receives DHCP offer messages from one or more servers, the client chooses one server from which to request configuration parameters. The client broadcasts a DHCP request message that includes the server identifier option to indicate which server it has selected. This DHCP request message is broadcast and relayed through routers if necessary.	[redacted]
The server selected in the DHCP request message commits the configuration parameters for the client to persistent storage and responds with a DHCP ACK message, containing the configuration parameters for the requesting client.	[redacted]
If a client receives no DHCP offer or DHCP ACK messages, the client	[redacted]

times out and retransmits the DHCP discover and request messages. To avoid synchronicity and excessive network traffic, the client uses a randomized exponential backoff algorithm to determine the delay between retransmissions. The delay between retransmissions should be chosen to allow sufficient time for replies from the server, based on the characteristics of the network between the client and server. For example, on a 10-Mbps Ethernet network, the delay before the first retransmission should be 4 seconds, randomized by the value of a uniform random number chosen from the range -1 to +1. The delay before the next retransmission should be 8 seconds, randomized by the value of a uniform number chosen from the range -1 to +1. The retransmission delay should be doubled with subsequent retransmissions up to a maximum of 64 seconds.

DHCP Relay Agents

DHCP clients send their messages as broadcasts. Broadcasts don't cross routers and remain local to a subnet. Without a relay agent, each subnet needs a DHCP server. As mentioned, a router can act as a DHCP relay agent. This means that the router passes DHCP broadcast messages from clients to DHCP servers that are not on the same subnet as the clients. This avoids a requirement for a DHCP server to reside on each subnet with clients.

With Cisco routers, you can use the ip helper-address command on each router interface where clients reside to cause the router to become a DHCP relay agent. An address parameter for the command should point to the IP address of the DHCP server. Alternatively, the address can be a broadcast address so that the router broadcasts the DHCP discover message on to the specified network. A broadcast address should be used only if the server is on a directly connected network because modern routers do not forward directed broadcasts to other networks.

When a router relays a discover message to another network or subnet, the router places the IP address for the interface on which the message arrived in the gateway address (giaddr) field of the DHCP header. The DHCP server can use the giaddr information to determine from which scope the assigned address should come.

Caution When you enable an IP helper address, Cisco routers forward numerous UDP broadcasts by default, including TFTP, DNS, NetBIOS, and TACACS broadcasts. To configure the router to be more discerning in its forwarding, use the ip forward-protocol command for the protocols that should be forwarded and the no ip forward-protocol command for the protocols that should not be forwarded. To make sure DHCP

packets are forwarded, use the ip forward-protocol udp 67 command. DHCP discover and request messages use UDP destination port number 67, the port number reserved many years ago for BOOTP.

IP Version 6 Dynamic Addressing

Like IPv4, IP version 6 (IPv6) supports both static and dynamic addressing. IPv6 calls dynamic addressing autoconfiguration. IPv6 autoconfiguration can be stateful or stateless.

With a stateful autoconfiguration method, hosts obtain addresses and other parameters from a server. The server stores a database containing the necessary information and maintains control over address assignments. This should sound familiar. The stateful autoconfiguration model is defined in DHCPv6.

Stateless autoconfiguration requires no manual configuration of hosts, minimal (or no) configuration of routers, and no servers. For a network engineer who is not concerned about which addresses are used, as long as they are unique and routable, stateless autoconfiguration offers many benefits. Stateless autoconfiguration is discussed in RFC 4862.

With stateless autoconfiguration, a host generates its own address using locally available information plus information advertised by routers. The

process begins with the generation of a link-local address for an interface. The link-local address is generated by combining the well-known link-local prefix (FE80::/10) with a 64-bit interface identifier. For more information about IPv6 prefixes, see the "Hierarchy in IP Version 6 Addresses" sec-tion later in this chapter.

The next step determines the uniqueness of the tentative address that has been derived by combining the link-local address prefix with the interface identifier. The host transmits a Neighbor Solicitation message with the tentative address as the target address. If another host is using this address, a Neighbor Advertisement is returned. In this event, autoconfiguration stops and some manual intervention is required. Because the address is usually based on a NIC address, duplicates are very unlikely. If no responses are returned, the tentative address is considered unique, and IP connectivity with local hosts is now possible.

The final step of the autoconfiguration process involves listening for IPv6 Router Advertisement messages that routers periodically transmit. A host can also force an immediate Router Advertisement by transmitting a Router Solicitation message to the all routers multicast address. Router Advertisements contain zero or more Prefix Information options that contain information used by the host to

generate a global address.

Additional option fields contain a subnet prefix and lifetime values, indicating how long addresses created from the prefix remain preferred and valid.

Zero Configuration Networking

The Zero Configuration Networking (Zeroconf) working group of the Internet Engineering Task Force (IETF) carried the concept of dynamic addressing one step further than DHCP. Like AppleTalk and IPv6, Zeroconf can allocate IP addresses without a server. It can also translate between names and IP addresses without a DNS server. To handle naming, Zeroconf supports multicast DNS (mDNS), which uses multicast addressing for name resolution. Zeroconf is implemented in Mac OS, Windows operating systems, Linux, most printers from major printer vendors, and other network devices.

Zeroconf is a link-local technology. Link-local addresses and names are meaningful only on a particular network. They are not globally unique. Zeroconf is appropriate for home and small office networks, ad hoc networks at meetings and conferences (especially wireless networks), embedded systems as in an automobile, and whenever two devices need to spontaneously share or exchange information.

Although the main goal of Zeroconf is to make personal computer networking easier to use, according to the www.zeroconf.org web page, a long-term goal is "to enable the creation of entirely new kinds of networked products, products that today would simply not be commercially viable because of the inconvenience and support costs involved in setting up, configuring, and maintaining a network."

Zeroconf has many advantages, but one potential risk is that it will interfere with more structured systems for address and name assignments, although the www.zeroconf.org page does say that Zeroconf "must coexist gracefully with larger configured networks" and that Zeroconf protocols "must not cause harm to the network when a Zeroconf machine is plugged into a large network." Zeroconf has some intriguing features and, as a network designer, you should have some familiarity with it. Zeroconf can solve various problems for small, ad hoc networks.

Using Private Addresses in an IP Environment

Private IP addresses are addresses that an enterprise network administrator assigns to internal networks and hosts without any coordination from an ISP or one of the RIRs. An ISP or the RIR provides public addresses for web servers or other servers that external

users access, but public addresses are not necessary for internal hosts and networks. Addressing for internal hosts that need access to outside services, such as email, FTP, or web servers, can be handled by a NAT gateway. NAT is covered later in this chapter.

In RFC 1918, the IETF reserves the following numbers for addressing nodes on internal private networks:

- 10.0.0.0 through 10.255.255.255
- 172.16.0.0 through 172.31.255.255
- 192.168.0.0 through 192.168.255.255

One advantage of private network numbers is security. Private network numbers are not advertised to the Internet. Private network numbers must not be advertised to the Internet because they are not globally unique. By not advertising private internal network numbers, a modicum of security is achieved. Additional security, including firewalls and intrusion detection systems, should also be deployed, as discussed in Chapter 5, "Designing a Network Topology," and Chapter 8, "Developing Network Security Strategies."

Private addressing also helps meet goals for adaptability and flexibility. Using private addressing makes it

easier to change ISPs in the future. If private addressing has been used, when moving to a new ISP, the only address changes required are in the router or firewall providing NAT services and in any public servers. You should recommend private addressing to customers who want the flexibility of easily switching to a different ISP in the future. Another advantage of private network numbers is that an enterprise network can advertise just one network number, or a small block of network numbers, to the Internet. It is good practice to avoid advertising many network numbers to the Internet. One of the goals of modern Internet practices is that Internet routers should not need to manage huge routing tables. As an enterprise network grows, the network manager can assign private addresses to new networks, rather than requesting additional public network numbers from an ISP or RIR. This avoids increasing the size of Internet routing tables. Private network numbers let a network designer reserve scarce Internet addresses for public servers. During the mid-1990s, as the Internet became commercialized and popularized, a scare rippled through the Internet community about the shortage of addresses. Dire predictions were made that no more addresses would be available by the turn of the century.	[redacted]

Because of this scare, many companies (and many ISPs) were given a small set of addresses that needed to be carefully managed to avoid depletion. These companies recognize the value of private addresses for internal networks.

Note The shortage of Internet addresses was mainly due to the way IPv4 addresses were divided into classes. Although the IP address space is fixed in size and will become depleted at some point, the 4-byte size of an IPv4 address is theoretically large enough to address more than 4 billion nodes. The method used to divide the address space into classes meant that many addresses were wasted, however. Approximately 50 percent of the address space was used for Class A host addresses. Another 12 percent was used for Class C host addresses.

With the invention of classless addressing, the threat of running out of addresses is less imminent. Classless addressing is covered in the "Using a Hierarchical Model for Assigning Addresses" section of this chapter.

Caveats with Private Addressing

Although the benefits of private addressing outweigh the disadvantages, it is important to be aware of the drawbacks. One drawback is that outsourcing network management is difficult. When a

company delegates network management responsibility to an outside company, the outside company typically sets up network consoles at its own site that communicate with internetworking devices inside the client's network. With private addressing, however, the consoles cannot reach the client's devices, because no routes to internal networks are advertised to the outside. NAT can be used to translate the private addresses to public addresses, but then there might be problems with interoperability between NAT and network management protocols such as the Simple Network Management Protocol (SNMP).

Another drawback for private addressing is the difficulty of communicating with partners, vendors, suppliers, and so on. Because the partner companies are also probably using private addresses, building extranets becomes more difficult. Also, companies that merge with each other face a difficult chore of renumbering any duplicate addresses caused by both companies using the same private addresses.

One other caveat to keep in mind when using private addresses is that it is easy to forget to use a structured model with the private addresses. Enterprise network managers, who were once starved for addresses that were carefully doled out by ISPs and the RIRs, get excited when they move to private addressing and have all of

điều

những địa chỉ này các

bây giờ

họ

network 10.0.0.0 at their disposal.

The excitement should not overshadow the need to assign the new address space in a structured, hierarchical fashion. Hierarchical addressing facilitates route summarization within the enterprise network, which decreases bandwidth consumption by routing protocols, reduces processing on routers, and enhances network resiliency.

Network Address Translation

Network Address Translation (NAT) is an IP mechanism that is described in RFC 3022 for converting addresses from an inside network to addresses that are appropriate for an outside network, and vice versa. NAT is useful when hosts that need access to Internet services have private addresses. NAT functionality can be implemented in a separate appliance, router, or firewall.

The NAT administrator configures a pool of outside addresses that can be used for translation. When an inside host sends a packet, the source address is translated dynamically to an address from the pool of outside addresses. NAT also has a provision for static addresses for servers that need a fixed address (for example, a web or mail server that must always map to the same well-known address).

Some NAT products also offer port translation for mapping several

addresses to the same address. With port translation, all traffic from an enterprise has the same address. Port numbers are used to distinguish separate conversations. Port translation reduces the number of required outside addresses. Port translation is sometimes called NAT overload or Port Address Translation (PAT). When using NAT, all traffic between an enterprise network and the Internet must go through the NAT gateway. For this reason, you should make sure the NAT gateway has superior throughput and low delay, particularly if enterprise users depend on Internet video or voice applications. The NAT gateway should have a fast processor that can examine and change packets quickly. Keep in mind that, in addition to modifying IP addresses, a NAT gateway must modify the IP, TCP, and UDP checksums. (The checksums for TCP and UDP cover a pseudo header that contains source and destination IP addresses.) In many cases, NAT must also modify IP addresses that occur inside the data part of a packet. IP addresses can appear in ICMP, FTP, DNS, SNMP, and other types of packets. Because NAT has the job of translating something so basic as network layer addresses, it can be tricky to guarantee correct behavior with all applications. A NAT gateway should be thoroughly tested in a pilot environment before it is generally deployed.	vào cùng một địa chỉ tính năng biên dịch cổng Tính năng biên dịch cổng giúp địa chỉ cần thiết bên ngoài tính năng biên dịch cổng các ứng dụng video và thoại trên Internet các xuất hiện đến mức độ cơ bản như các địa chỉ mạng nó (NAT) hoạt động tốt với mọi ứng dụng là điều khó khăn

Using a Hierarchical Model for Assigning Addresses Hierarchical addressing is a model for applying structure to addresses so that numbers in the left part of an address refer to large blocks of networks or nodes, and numbers in the right part of an address refer to individual networks or nodes. Hierarchical addressing facilitates hierarchical routing, which is a model for distributing knowledge of a network topology among internetwork routers. With hierarchical routing, no single router needs to understand the complete topology. This chapter focuses on hierarchical addressing and routing for IP environments, but the concepts apply to other environments also.	[redacted] sử dụng [redacted] cho [redacted] sao cho [redacted] thể hiện một lượng lớn các khối mạng hoặc nút [redacted] biểu diễn từng mạng hoặc nút riêng biệt [redacted] nhưng những khái niệm này cũng có thể áp dụng [redacted]
Why Use a Hierarchical Model for Addressing and Routing? Chapter 5 examined the importance of hierarchy in topology design. The benefits of hierarchy in an addressing and routing model are the same as those for a topology model: ■ Support for easy troubleshooting, upgrades, and manageability ■ Optimized performance ■ Faster routing-protocol convergence	[redacted] việc [redacted] cũng như [redacted]

■ Scalability

■ Stability

■ Fewer network resources needed (CPU, memory, buffers, bandwidth, and so on)

Hierarchical addressing permits the summarization (aggregation) of network numbers. Summarization allows a router to group many network numbers when advertising its routing table. Summarization enhances network performance and stability. Hierarchical addressing also facilitates variable-length subnet masking (VLSM). With VLSM, a network can be divided into subnets of different sizes, which helps optimize available address space.

giúp triển khai dễ dàng quá trình

Hierarchical Routing

Hierarchical routing means that knowledge of the network topology and configuration is localized. No single router needs to understand how to get to each other network segment. Hierarchical routing requires that a network administrator assign addresses in a hierarchical fashion. IP addressing and routing have been somewhat hierarchical for a long time, but in recent years, as the Internet and enterprise intranets have grown, it has become necessary to add more hierarchy.

To understand hierarchical routing in simple terms, consider the telephone

Không bộ định tuyến nào cần biết cách đi vào mỗi phân đoạn mạng.

việc tăng thêm mức độ phân cấp đang ngày càng cấp thiết

system. The telephone system has used hierarchical routing for years. When you dial 541-555-1212 from a phone in Michigan, the phone switch in Michigan does not know how to reach this specific number in Oregon. The switch in Michigan simply knows that the number is not in Michigan and forwards the call to a national phone company. A switch at the national phone company knows that 541 is for southern Oregon but does not know specifically where calls to 555 should go. A switch in Oregon determines which central office switch handles the 555 prefix, and that switch routes the call to 1212.	[redacted] thuộc [redacted] gia. Bộ phận chuyển đổi ở công ty điện thoại quốc gia [redacted] đến của các cuộc gọi vào số 555 [redacted] sẽ [redacted] bộ phận chuyển đổi văn phòng trung tâm nào sẽ bộ phận chuyển đổi nào [redacted]
With data networking, similar decisions are made as a packet travels across a routed network. However, in traditional IP networks, the decisions could not be quite as local as the decisions in the telephone example. Until recently, IP addresses for the Internet were assigned in a nonhierarchical fashion. For example, two companies in Oregon might have widely different Class C network numbers, despite the fact that they both use the same upstream provider to reach the Internet. This meant that the provider had to tell all other Internet sites about both Oregon companies. So, unlike the phone	[redacted] khi [redacted] về đến thời gian [redacted] dịch vụ cấp trên [redacted]

system example, a router in Michigan would know how to reach specific networks in Oregon.

Classless Interdomain Routing

In the mid-1990s, the IETF and IANA realized that the lack of a hierarchical model for assigning network numbers in the Internet was a severe scalability problem. Internet routing tables were growing exponentially, and the amount of overhead to process and transmit the tables was significant. To constrain routing overhead, it became clear that the Internet must adopt a hierarchical addressing and routing scheme. To solve the routing overhead problem, the Internet adopted the classless interdomain routing (CIDR) method for summarizing routes. CIDR specifies that IP network addresses should be assigned in blocks, and that routers in the Internet should group routes to cut down on the quantity of routing information shared by Internet routers.

RFC 2050 provides guidelines for IP address allocation by RIRs and ISPs. RFC 2050 states that An Internet Provider obtains a block of address space from an address registry, and then assigns to its customers addresses from within that block based on each customer requirement. The result of this process is that routes to many customers will be aggregated together, and will appear to other providers as a single route. For route aggregation to

be effective, Internet providers encourage customers joining their network to use the provider's block, and thus renumber their computers. Such encouragement may become a requirement in the future.

At the same time that the IETF and IANA addressed the problem of nonhierarchical routing, they also addressed the problem of IP address depletion. As mentioned previously, the system of assigning addresses in classes meant that many addresses were going to waste. The IETF developed classless addressing, which provides more flexibility in specifying the length of the prefix part of an IP network number.

Classless Routing Versus Classful Routing

As shown in Figure 6-1, an IP address contains a prefix part and a host part. Routers use the prefix to determine the path for a destination address that is not local. Routers use the host part to reach local hosts.

Figure 6-1 Two Parts of an IP Address

A prefix identifies a block of host numbers and is used for routing to that block. Traditional routing, also known as classful routing, does not transmit any information about the prefix length. With classful routing, hosts and routers calculate the prefix length by looking at the first few bits of an address to determine its class. The first

khác việc tập trung tuyến có hiệu quả

Nhìn vào chúng ta thấy

đến

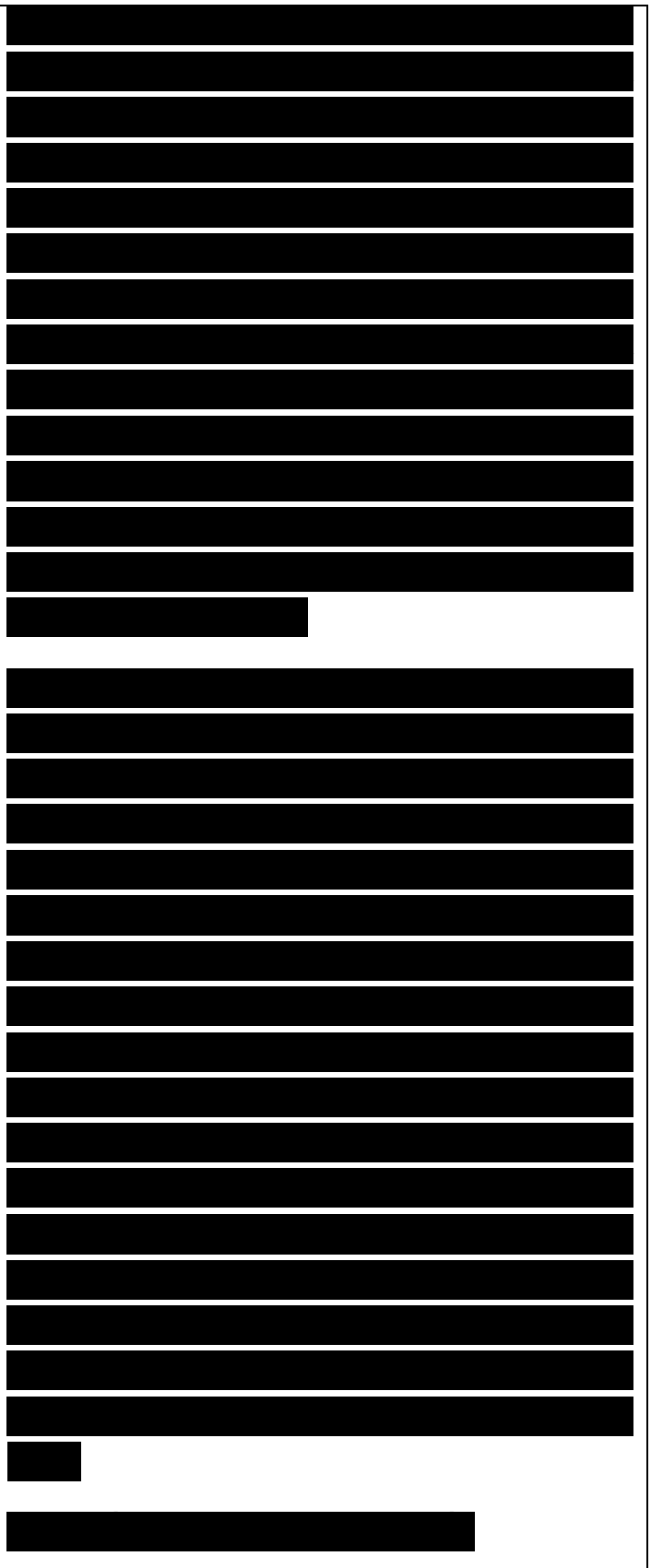
few bits for Class A through C addresses are as shown in the following chart: Class A First bit = 0 Prefix is 8 bits Class B First 2 bits = 10 Prefix is 16 bits Class C First 3 bits = 110 Prefix is 24 bits In early IP implementations, IP hosts and routers understood only three prefix lengths: 8, 16, and 24. This became a limitation as networks grew, so subnetting was introduced. With subnets, a host (or router) can be configured to understand that the local prefix length is extended. This configuration is accomplished with a subnet mask. For example, routers and hosts can be configured to understand that network 10.0.0.0 is subnetted into 254 subnets by using a subnet mask of 255.255.0.0. CIDR notation indicates the prefix length with a length field, following a slash. For example, in the address 10.1.0.1/16, the 16 indicates that the prefix is 16 bits long, which means the same as the subnet mask 255.255.0.0. Traditional IP hosts and routers had a limited capability to understand prefix lengths and subnets. They understood the length for local configurations but not for remote configurations. Classful routing did not transmit any information about the prefix length.	đối với các địa chỉ từ Lớp A đến lớp C Điều này dần trở thành một nhược điểm khi mạng phát triển vì vậy người ta đưa ra phương pháp chia mạng con các cục bộ người ta có thể nó bằng cách dùng

The prefix length was calculated from the information about the address class provided in the first few bits of an address, as mentioned earlier. Classless routing protocols, on the other hand, transmit a prefix length with an IP address. This allows classless routing protocols to group networks into one entry and use the prefix length to specify which networks are grouped. Classless routing protocols also accept any arbitrary prefix length, rather than only accepting lengths of 8, 16, or 24, which the classful system dictated. Classless routing protocols include Routing Information Protocol (RIP) version 2, Enhanced Interior Gateway Routing Protocol (EIGRP), Open Shortest Path First (OSPF), Border Gateway Routing Protocol (BGP), and Intermediate System-to-Intermediate System (IS-IS). Classful routing protocols include RIP version 1 and the Interior Gateway Routing Protocol (IGRP). Classful routing protocols are almost obsolete. RIP version 2 has replaced RIP version 1 and EIGRP has replaced IGRP. Route Summarization (Aggregation) When advertising routes into another major network, classful routing protocols automatically summarize	[redacted]

subnets. They only advertise a route to a Class A, B, or C network, instead of routes to subnets. Because classful routers and hosts do not understand nonlocal prefix lengths and subnets, there is no reason to advertise information about prefix lengths. The automatic summarization into a major class network has some disadvantages; for example, discontiguous subnets are not supported. Chapter 3, "Characterizing the Existing Internetwork," mentioned discontiguous subnets, and they are covered in more detail later in this chapter in the section "Discontiguous Subnets."

Classless routing protocols advertise a route and a prefix length. If addresses are assigned in a hierarchical fashion, a classless routing protocol can be configured to aggregate subnets into one route, thus reducing routing overhead. The importance of route summarization on the Internet was already discussed. Summarizing (aggregating) routes on an enterprise network is also important because route summarization reduces the size of routing tables, which minimizes bandwidth consumption and processing on routers. Route summarization also means that problems within one area of the network do not tend to spread to other areas.

Route Summarization Example



This section covers a route summarization example that is based on the network shown in Figure 6-2. Looking at Figure 6-2, you can see that a network administrator assigned network numbers 172.16.0.0 through 172.19.0.0 to networks in a branch office.

Figure 6-2 Route Summarization Example

The branch-office router in Figure 6-2 can summarize its local network numbers and report that it can reach 172.16.0.0/14. By advertising this single route, the router is saying, "Route packets to me if the destination has the first 14 bits set to 172.16." The router is reporting a route to all networks where the first 14 bits are equal to 10101100 000100 in binary.

To understand the summarization in this example, you should convert the number 172 to binary, which results in the binary number 10101100. You should also convert the numbers 16 through 19 to binary, as shown in the following chart:

Second Octet in Decimal Second Octet in Binary

Notice that the leftmost 6 bits for the numbers 16 through 19 are identical. This is what makes route summarization with a prefix length of 14 possible in this example. The first 8 bits for the networks are identical (all the networks have 172 for the first

octet), and the next 6 bits are also identical.

Route Summarization Tips

For route summarization to work correctly, the following requirements must be met:

- Multiple IP addresses must share the same leftmost bits.

- Routers must base their routing decisions on a 32-bit IP address and prefix length that can be up to 32 bits. (A host-specific route has a 32-bit prefix.)
- Routing protocols must carry the prefix length with 32-bit addresses.

By spending some time analyzing network numbers (and converting the addresses to binary), you can see the simplicity and elegance of classless addressing and route summarization. When you look at a block of subnets, you can determine if the addresses can be summarized by using the following rules:

- The number of subnets to be summarized must be a power of 2 (for example, 2, 4, 8, 16, 32, and so on).
- The relevant octet in the first address in the block to be summarized must be a multiple of the number of subnets.

Consider an example. The following network numbers are defined at a branch office. Can they be summarized?

The number of subnets is 5, which is not a power of 2, so the first condition is not met. The relevant octet (third in this case) is 32, which is a not a multiple of the number of subnets. So, the second condition is not met. However, the first four subnets can be summarized. A router can summarize the first four networks as 192.168.32.0/22. The leftmost 22 bits for the first four networks are identical. The router can advertise the 192.168.36.0 network in a separate route from the 192.168.32.0/22 summary route.

Discontiguous Subnets

As mentioned earlier, classful routing protocols automatically summarize subnets. One side effect of this is that discontiguous subnets are not supported. Subnets must be next to each other (contiguous). Figure 6-3 shows an enterprise network with discontiguous subnets.

Figure 6-3 Network with Discontiguous Subnets

With a classful routing protocol such as RIP version 1 or IGRP, Router A in Figure 6-3 advertises that it can get to network 10.0.0.0. Router B ignores this advertisement because it can already get to network 10.0.0.0. It is directly attached to network 10.0.0.0. The opposite is also true: Router B advertises that it can get to network 10.0.0.0, but Router A ignores this

information. This means that the routers cannot reach remote subnets of net-work 10.0.0.0.

To solve this problem, a classless routing protocol can be used. With a classless routing protocol, Router A advertises that it can get to networks 10.108.16.0/20. Router B advertises that it can get to networks 10.108.32.0/20. (To understand why the prefix length is 20, convert the network numbers to binary.) Because classless routing protocols understand prefixes of any length (not just 8, 16, or 24), the routers in Figure 6-3 can route to discontiguous subnets, assuming they are running a classless routing protocol, such as OSPF or EIGRP.

Note To configure the devices in the previous example with an old-style subnet mask rather than a prefix length, use a mask of 255.255.240.0. The first 4 bits of the third octet are set to 1s. A trick for determining the value of the relevant octet in a subnet mask is to subtract the number of summarized subnets from 256. In this example, there are 16 summarized subnets, so the relevant octet is 256 minus 16, or 240.

Mobile Hosts

Classless routing and discontiguous subnets support mobile hosts. A mobile host, in this context, is a host

that moves from one network to another and has a statically defined IP address. A network administrator can move a mobile host and configure a router with a host-specific route to specify that traffic for the host should be routed through that router.

In Figure 6-4, for example, host 10.108.16.1 has moved to a different network. Even though Router A advertises that network 10.108.16.0/20 is behind it, Router B can advertise that 10.108.16.1/32 is behind it.

When making a routing decision, classless routing protocols match the longest prefix.

The routers in the example have in their tables both 10.108.16.0/20 and 10.108.16.1/32.

Host 10.108.16.1

Figure 6-4 Mobile Host

In Figure 6-4, a better design would be to use DHCP so that hosts can be moved without requiring any reconfiguration on the hosts or routers. The example is simply used to explain the longest-prefix-match concept. It is not meant to be a design recommendation.

Variable-Length Subnet Masking

Using a classless routing protocol means that you can have different sizes of subnets within a single

network. Varying the size of subnets is also known as variable-length subnet masking (VLSM). VLSM relies on providing prefix length information explicitly with each use of an address. The length of the prefix is evaluated independently at each place it is used. The capability to have a different prefix length at different points supports efficiency and flexibility in the use of the IP address space. Instead of each subnet being the same size, you can have both big and small subnets.

One use for small subnets is point-to-point WAN links that only have two devices (one router on each end of the link). Such a link can use a subnet mask of 255.255.255.252, because only two devices need addresses. The two devices can be numbered 01 and 10.

Note A disadvantage of using a separate subnet for each WAN link is that each subnet adds an entry to the routing table. With some vendors' routers, you do not need to number the serial ports on a point-to-point WAN link, which obviates the need for small WAN point-to-point subnets. One drawback with unnumbered ports, however, is that you cannot ping them, which makes troubleshooting more difficult. But if SNMP or other network management tools can identify port problems, the capability to ping a WAN port is not essential. Unnumbered WAN ports are a better

solution than small WAN point-to-point subnets in this case.

Hierarchy in IP Version 6 Addresses

IPv6 increases the IP address size from 32 bits to 128 bits. The long address means that multiple levels of hierarchy can be built in to the address. Despite popular myths, the developers of IPv6 considered support for multiple levels of hierarchy a more important reason for a huge address space than the ability to address every device on the planet and possibly other planets.

An IPv6 address is written in hexadecimal rather than the dotted decimal format used by IPv4. The hexadecimal values of the eight 16-bit pieces of the address are represented as a series of fields separated by colons, in x:x:x:x:x:x:x:x format. For example, here are two IPv6 addresses:

FEDC:BA98:7654:3210:FEDC:BA98:7654:3210

1080:0:0:0:8:800:200C:417A

Note that it is not necessary to write the leading 0s in an individual field, but there must be at least one numeral in every field (except when suppressing multiple fields of 0s). Because IPv6 addresses tend to contain long strings of 0s, you can

substitute double colons (::) at the start, middle, or end of an address to indicate consecutive 16-bit fields of 0s.

To avoid confusion, only one set of double colons may be used. For example, the IPv6 address 2031:0000:130F:0000:0000:09C0:876A:130B can be written as 2031:0:130F::9C0:876A:130B.

However, it cannot be written as 2031::130F::9C0:876A:130B.

As is the case with IPv4, in IPv6 a source can address datagrams to either one or many destinations. IPv6 supports unicast (one to one) and multicast (one to many). IPv6 has no concept of broadcast addresses; multicast addresses are used instead. IPv6 also supports anycast (one to nearest), which is used for sending a packet to any one of a group of interfaces. An IPv6 anycast address is an address that is assigned to more than one interface (typically belonging to different nodes), with the property that a packet sent to an anycast address is routed to the "nearest" interface having that address, according to the routing protocol's measure of distance.

There are three types of unicast addresses in IPv6:

- Link-local addresses
- Global unicast addresses
- IPv6 addresses with embedded IPv4 addresses

In the past, site-local addresses were also supported but they were deprecated by RFC 3879. The sections that follow describe the three types of IPv6 unicast addresses in more detail.

Link-Local Addresses

A link-local address is useful only in the context of a single link or network. An IPv6 link- local unicast address can be automatically configured on an interface by using the link- local prefix (FE80::/10) and an interface identifier. The interface ID is 64 bits long and is often derived from the hardware address in ROM on an interface. For example, it can be based on an IEEE 802 48-bit MAC address.

A link-local address serves as a method for connecting devices on the same local network without the need for globally unique addresses. An IPv6 router must not forward packets that have either a link-local source or destination address. Link-local addresses are used in neighbor discovery and in the stateless autoconfiguration process, which was explained earlier in the chapter in the "IP Version 6 Dynamic Addressing" section.

The following Wireshark output shows a computer with a link-local

unicast address sending a packet to a link-local multicast address. The computer is trying to find its router. Ethernet II Destination: 33:33:00:00:00:02 Source: 00:22:41:36:97:17 Type: IPv6 (0x86dd) Internet Protocol Version 6 Version: 6 Traffic class: 0x00000000 Flowlabel: 0x00000000 Payload length: 16 Next header: ICMPv6 (0x3a) Hop limit: 255 Source: fe80::222:41ff:fe36:9717 Destination: ff02::2 Internet Control Message Protocol v6 Type: 133 (Router solicitation) Code: 0 Checksum: 0xca4e [correct] ICMPv6 Option (Source link-layer address) Type: Source link-layer address (1) Length: 8 Link-layer address: 00:22:41:36:97:17 Notice that the 64-bit interface ID in the source IPv6 address is based on the computer's 48-bit MAC address. You can see the MAC address in the Ethernet II header and in the ICMPv6 option field at the end of the packet. The 0222:41ff:fe36:9717 IPv6	[redacted]

interface ID is the MAC address with FF:FE in the middle. Also, the value of bit 6 has been changed to a binary 1. In the MAC address, bits 0-7, counting from the left, were 00 in hexadecimal. In the IPv6 address, bit 6 is a binary 1, so bits 0-7 are 02 in hexadecimal. Bit 6 is the universal/local bit and is usually set to 1 in IPv6 to indicate a universal address rather than an address that has only local significance to an enterprise.

Global Unicast Addresses

Global unicast addresses are equivalent to public registered addresses in IPv4. These addresses are designed to support the type of provider-based aggregation currently used on the Internet. The structure of global unicast addresses enables aggregation of routing prefixes so that the number of routing-table entries in the global Internet routing table and in provider and enterprise routing tables can be minimized. Global unicast addresses are aggregated upward through organizations, then to intermediate-level ISPs, and eventually to top-level ISPs.

The general format for IPv6 global unicast addresses is as follows:
Global routing prefix: n bits Subnet ID: m bits Interface ID: 128-n-m bits
The global routing prefix is typically a hierarchically structured value assigned to a site. It is a cluster of

subnets or links. It is designed to be structured hierarchically by RIRs and ISPs. The subnet ID is an identifier of a subnet within the site and is designed to be structured hierarchically by site administrators.

RFC 3513 requires that all unicast addresses, except those that start with binary value 0, have interface IDs that are 64 bits long and are constructed in modified EUI-64 format. So, according to this RFC, the global routing prefix is n bits; the subnet ID is 64-n bits; and the Interface ID is 64 bits.

An example of a global unicast address is 2001:4860:800d::63, which belongs to Google. Note from the following DiG output that Google has many IPv6 global unicast addresses:

……..

IPv6 Addresses with Embedded IPv4 Addresses

Some IPv4-to-IPv6 transition strategies use an IPv4 address. For example, when tunneling IPv6 packets over an IPv4 routing infrastructure, IPv6 nodes can be assigned special IPv6 unicast addresses that carry an IPv4 address in the low-order 32 bits. The IPv6 address consists of 96 zeros, followed by a 32-bit globally unique IPv4 unicast address. This type of address is called an IPv4-compatible IPv6 address. An example of such an address is 0:0:0:0:0:0:66.241.68.22, or more succinctly, ::66.241.68.22.

A second type of IPv6 address that holds an embedded IPv4 address is used to represent the address of an IPv4 node as an IPv6 address. This type of address is called an IPv4-mapped IPv6 address and consists of 80 zero bits, 16 one bits, and a 32-bit IPv4 unicast address. An example of such an address is 0:0:0:0:0:FF:66.241.68.22, or more succinctly, ::FF:66.241.68.22.

Note For the specific cases of IPv4 in IPv6 addresses, you write the IPv4 address in the low-order 32 bits in dotted decimal instead of hexadecimal.

Designing a Model for Naming

Names play an essential role in meeting a customer's goals for usability. Short, meaningful names enhance user productivity and simplify network management. A good naming model also strengthens the performance and availability of a network. The goal of this section is to help you design naming models for internetworks that will meet your customer's goals for usability, manageability, performance, and availability.

Names are assigned to many types of resources in a typical internetwork: routers, servers, hosts, printers, and other resources. This section covers

the naming of devices and networks. Providing names for users, groups, accounts, and passwords is not covered, although some of the guidelines for naming devices apply to these items also.

A good naming model should let a user transparently access a service by name rather than address. Because networking protocols require an address, the user's system should map the name to an address. The method for mapping a name to an address can be either dynamic, using some sort of naming protocol, or static (for example, a file on the user's system that lists all names and their associated addresses). Usually, a dynamic method is preferable, despite the additional network traffic caused by dynamic naming protocols.

When developing a naming model, consider the following questions:

- What types of entities need names? Servers, routers, printers, hosts, others?
- Do end systems need names? Will the end systems offer any services, such as personal web serving?
- What is the structure of a name? Does a portion of the name identify the type of device?
- How are names stored, managed, and accessed?

- Who assigns names?
- How do hosts map a name to an address? Will a dynamic or static system be provided?
- How does a host learn its own name?
- If dynamic addressing is used, will the names also be dynamic and change when an address changes?
- Should the naming system use a peer-to-peer or client/server model?
- If name servers will be used, how much redundancy (mirroring) will be required?
- Will the name database be distributed among many servers?
- How will the selected naming system affect network traffic?
- How will the selected naming system affect security?

Distributing Authority for Naming

During the early stages of designing a naming model, consider who will actually assign names by asking the following questions:

- Will the name space be completely controlled by a centralized authority, or will the naming of some devices be carried out by decentralized agents?
- Will a corporate IS department name devices at regional and branch

offices, or can departmental administrators implement naming at those sites?

■ Will users be allowed to name their own systems, or are all names assigned by network administrators?

The disadvantage of distributing authority for naming is that names become harder to control and manage. If all groups and users agree on, and practice, the same policies, however, there are many advantages to distributing authority for naming.
The obvious advantage of distributing authority for naming is that no department is burdened with the job of assigning and maintaining all names. Other advantages include performance and scalability. If each name server manages a portion of the name space instead of the whole name space, the requirements for memory and processing power on the servers are lessened. Also, if clients have access to a local name server instead of depending on a centralized server, many names can be resolved to addresses locally, without causing traffic on the internetwork. Local servers can cache information about remote devices, to further reduce network traffic.

Guidelines for Assigning Names
To maximize usability, names should be short, meaningful, unambiguous,

and distinct. A user should easily recognize which names go with which devices. A good practice is to include in a name some sort of indication of the device's type. For example, you can prefix or suffix router names with the characters rtr, switches with sw, servers with svr, and so on. Using meaningful prefixes or suffixes decreases ambiguity for end users and helps managers more easily extract device names from network management tools.

Names can also include a location code. Some network designers use airport codes in their naming models. For example, all names in San Francisco start with SFO, all names in Oakland start with OAK, and so on. The location code could be a number instead, but most people remember letters better than numbers.

Try to avoid names that have unusual characters, such as underscores, ampersands, asterisks, and so on, even if the naming protocol allows these characters (which many do). These characters are hard to type and can cause applications and protocols to behave in unexpected ways. Unusual characters might mean something special to a protocol. For example, the dollar sign when used as the last character in a NetBIOS name means that the name does not appear in network browser lists or in response to NetBIOS network- survey commands. NetBIOS names with a dollar sign at the end are for administrative use

only.	[redacted]
It is also best if names are not case-sensitive because people usually cannot remember which case to use. Names that require a user to remember mixed cases (for instance, DBServer) are not a good idea. They are hard to type, and some protocols might not be case-sensitive anyway and might transmit the name as all lowercase or all uppercase, losing the significance of the mixed case. You should also avoid spaces in names. Spaces confuse users and might not work correctly with some applications or protocols. Names should generally be eight characters or fewer, if possible. This is especially true for operating systems, applications, or protocols that map names to filenames and restrict the size of a filename to eight characters.	[redacted]
If a device has more than one interface and more than one address, you should map all the addresses to one common name. For example, on a multiport router with multiple IP addresses, assign the same name to all the router's IP addresses. This way network management software does not assume that a multiport device is actually more than one device.	[redacted]
Note Security policy may dictate recommendations for naming. Names that are easily recognized by users are easily recognized by attackers also.	[redacted]

For key devices and data sources (for example, routers and servers), it is often a good idea to use long, cryptic names. In some cases, names are used only by system software and not by users, so usability is not affected. In other cases, tradeoffs must be made between usability and security goals.

Assigning Names in a NetBIOS Environment

NetBIOS is an application programming interface (API) that includes functions for naming devices, ensuring the uniqueness of names, and finding named services. NetBIOS was developed by IBM and Sytek in the 1980s for use on PC networks. It gained popularity in the late 1980s as a way of connecting PCs using software from IBM, Microsoft, and 3Com. It is still widely used in Microsoft Windows environments. When NetBIOS is used in a TCP/IP network, which is typical these days, the implementation is often called NetBT.

NetBT makes extensive use of broadcast packets by default. Broadcast packets are used to announce named services, find named services, and elect a master browser in a Windows environment. Using broadcasts is not the preferred method for implementing naming functions in a TCP/IP environment, however,

because of the performance implications and because routers do not forward broadcast packets by default. A router can be configured to forward NetBT broadcasts, which go to UDP port 137, but this is not an optimal solution because it requires extra configuration and spreads broadcasts, unless the configuration specifies a unicast address.

To avoid clients having to send broadcast frames to look for named services, a network administrator can place an lmhosts file on each station. The lmhosts file is an ASCII text file that contains a list of names and their respective IP addresses. The lmhosts file is similar to the hosts file on UNIX TCP/IP devices, although it includes some Windows-specific functionality.

The use of lmhosts files requires a lot of maintenance because the files don't dynamically change as names change. As a network grows, lmhosts files should be removed in favor of using WINS or DNS servers for dynamic resolution of NetBIOS names to IP addresses. When a PC is configured with a WINS server, the PC sends a message directly to the WINS server to resolve a name, instead of using the lmhosts file or sending broadcast packets. The PC also sends a message to the WINS server when it boots to make sure its own name is unique. To avoid configuring each PC with the address of a WINS server, a PC can receive the address of a WINS server

in the options field of a DHCP response.	[redacted]
To ensure that a PC can reach a WINS server, you can establish redundant WINS servers. To use redundant servers, you must plan to synchronize the WINS databases on the servers. This is accomplished by establishing WINS partners that use WINS replication. If the redundant WINS servers are on opposite sides of a slow WAN link, replication should occur infrequently or after business hours. For international networks, WINS replication is often set to every 12 hours.	[redacted]
In a NetBT environment, hosts have both a NetBIOS and an IP hostname. Typically these names are the same, but they do not have to be. IP hostnames are mapped to addresses using the Domain Name System (DNS). DNS is a standard Internet service and is covered in the next section, which covers naming in a generic IP environment as opposed to a NetBT environment. It is expected that over time, naming in a Windows environment will be accomplished solely with DNS, and WINS will become obsolete. If you are designing a network from scratch, there's no need for NetBT or WINS.	[redacted]
Microsoft also supports dynamic hostnames, which are necessary when	[redacted]

DHCP is used for dynamic addressing. With the DNS/WINS integration, a DNS server can query a WINS server to determine if the WINS server has learned a dynamic name. This avoids having to configure names in a DNS server, which is difficult with dynamic names.

Assigning Names in an IP Environment

Naming in an IP environment is accomplished by configuring hosts files, DNS servers, or Network Information Service (NIS) servers. DNS is used on the Internet and has also gained widespread popularity for managing names in enterprise networks. It is the recommended naming system for modern networks.

A hosts file tells a UNIX workstation how to convert a hostname into an IP address. A network administrator maintains a hosts file on each workstation in the internetwork. Both DNS and NIS were developed to allow a network manager to centralize the naming of devices, using a distributed database approach, instead of a flat file that resides on each system.

Sun Microsystems developed NIS to allow a UNIX network administrator to centralize the management of names and other configuration parameters. An administrator can use NIS to maintain names, user and password information, Ethernet addresses, mail aliases, group definitions, and protocol names and numbers. NIS was once quite common

but has become less common as the Internet standard for naming (DNS) gained momentum.

The Domain Name System

DNS was developed in the early 1980s when it became clear that managing a hosts file containing the names and addresses of all the systems on the Internet would no longer work. As the Internet hosts file grew, it became difficult to maintain, store, and transmit to other hosts.

DNS is a distributed database that provides a hierarchical naming system. A DNS name has two parts: a hostname and a domain name. For example, in information.priscilla.com, information is the host, and priscilla.com is the domain. Table 6-1 shows some of the most common top-level domains.

Newer top-level domains, such as .biz, .info, .museum, and .name, may help prevent the many disputes that occur over the right to use popular and marketable names. There are also many geographical top-level domains (for example, .uk for the United Kingdom and .de for Germany).

The DNS architecture distributes the knowledge of names so that no single system has to know all names. The Internet Corporation for Assigned Names and Numbers (ICANN) is a nonprofit corporation responsible for overall DNS management and top-level domains. ICANN has accredited a set of competitive registrars that

have authority over names under the top level.

Table 6-1 Top-Level Domains

Each layer of the hierarchy can also delegate authority. For example, a registrar might delegate authority to a corporate IS department for a name such as cisco.com. The IS department can delegate authority to the engineering department for names in the engi- neering.cisco.com subdomain. Within the engineering department, there might be multiple hosts with names such as development.engineering.cisco.com and testing.engineer- ing.cisco.com. Delegation allows DNS to be autonomously managed at each layer, which increases scalability and helps keep names meaningful.

DNS uses a client/server model. When a client needs to send a packet to a named station, resolver software on the client sends a name query to a local DNS server. If the local server cannot resolve the name, it queries other servers on behalf of the resolver. When the local name server receives a response, it replies to the resolver and caches information for future requests. The length of time that a server should cache information received from other servers is entered into the DNS database by a network administrator. Long time intervals decrease network traffic but can also make it difficult to change a name. The old name might be cached on thousands of servers in

the Internet.

The management of DNS names and servers is a complex task. For more information on managing DNS names in a UNIX environment, see the classic book by Paul Albitz and Cricket Liu, DNS and BIND, now in its fifth edition at press time.

Dynamic DNS Names

With many DHCP implementations, when a host requests an IP address from a DHCP server, the host also receives a dynamic hostname, something like pc23.dynamic.priscil-la.com. A dynamic name is not appropriate for some applications. For example, web servers, FTP servers, and some Internet telephony applications rely on static hostnames.

To reach a web server, a user types in a Uniform Resource Locator (URL) that is based on the server's domain name. If the name changes dynamically, it becomes impossible to reach the server. With some Internet telephony applications, a user needs to tell people the hostname to use when placing a call to the user's system. Another example is home users who want to access a computer on their home network while traveling. The home computer might get a different IP address every time it makes a connection to its ISP. This means that there is no stable address to connect

to.	[redacted]
For these types of applications, it is important to have a DNS implementation that can associate a static name with a dynamic address. Dynamic DNS is a service that provides the capability for a networked device, such as a home computer or router, to notify a DNS server to change, in real time, the active DNS configuration of its configured hostnames, addresses, or other information stored in the server. Service providers and vendors offer a variety of dynamic DNS solutions. Providers supply client software (or firmware) that automates the discovery and registration of a client's public IP address. The client program runs on a computer or router and connects to the service provider's DNS server and causes the server to link the discovered IP address with a hostname. Depending on the provider, the hostname is registered within a domain owned by the provider or the customer's own domain name. These services use a variety of methods and protocols. Often they use an HTTP request because restrictive environments sometimes allow only the HTTP protocol in outbound traffic from a client.	[redacted]
In Microsoft Windows networks, dynamic DNS is an integral part of Active Directory. Domain controllers	[redacted]

register their network service types in DNS so that other computers in the domain (or forest) can access them. Microsoft uses Kerberos authentication to secure this transaction. Other dynamic DNS services use the Generic Security Service Algorithm for Secret Key Transaction Authentication for DNS (GSS-TSIG), defined in RFC 3645. GSS-TSIG uses shared secret keys and one-way hashing to provide a cryptographically secure method of identifying each endpoint of a connection as being allowed to make or respond to a DNS update.

IPv6 Name Resolution

Name resolution with IPv6 can be handled statically with manual entries in a client host's local configuration files, or dynamically. Dynamic name resolution is accomplished with a DNS server that has built-in support for IPv6, usually along with IPv4. An IPv6-aware application maps a name to an IPv6 address with a request for an A6 record (an Address record for the IPv6 host). The network administrator must set up the appropriate DNS servers with IPv6 support and connect the named hosts to the IPv6 network with valid IPv6 addresses. On the client side, the administrator must either manually enter the address of a DNS server that can handle IPv6 addresses or use

DHCPv6 to inform the client of the DNS server's address.

Summary

This chapter provided guidelines for assigning addresses and names in an internetwork. The chapter illustrated the importance of using a structured model for addressing and naming to make it easier to understand network maps, operate network management software, recognize devices in protocol analyzer traces, and meet a customer's goals for usability.

Structured addresses and names facilitate network optimization and security because they make it easier to code network filters on firewalls, routers, and switches. Structured addresses also help you implement route summarization, which decreases bandwidth utilization, processing on routers, and network instability.

This chapter also discussed distributing authority for addressing and naming to avoid one department having to manage all addresses and names. Another way to simplify addressing and naming tasks is to use dynamic addressing and naming. Dynamic address- ing—for example, DHCP for IP environments—allows each end system to learn its address automatically. DHCP is recommended for the addressing of end systems in a campus network design.

Addressing and naming are essential elements of the logical design phase of the top- down network design process. If designed correctly, addressing and naming models can strengthen your ability to satisfy a customer's needs. They can also help you decide which routing and switching protocols to select, which is covered in the next chapter.	[redacted]

Chapter 7 checked 29/1

Selecting Switching and Routing Protocols

The goal of this chapter is to help you select the right switching and routing protocols for your network design customer. The selections you make will depend on your customer's business and technical goals. To help you select the right protocols for your customer, this chapter covers the following attributes of switching and routing protocols:

- Network traffic characteristics
- Bandwidth, memory, and CPU usage
- The approximate number of peer routers or switches supported
- The capability to quickly adapt to changes in an internetwork
- The capability to authenticate route updates for security reasons

At this point in the network design process,

you have created a network design topology and have developed some idea of where switches and routers will reside, but you haven't selected any actual switch or router products. An understanding of the switching and routing protocols that a switch or router must support will help you select the best product for the job.

This chapter begins with a generic discussion about decision making to help you develop a systematic process for selecting solutions for both the logical and physical components of a network design. Making sound decisions about protocols and technologies is a crucial network design skill that this chapter can help you develop.

A discussion of switching protocols follows the section on decision making. The switching section covers transparent bridging, multilayer switching, spanning-tree algorithm enhancements, and switching protocols for transporting virtual

LAN (VLAN) information. A section on routing protocols follows the bridging section. The routing section provides techniques for comparing and contrasting routing protocols. The chapter concludes with a table that summarizes the comparison of routing protocols. Making Decisions as Part of the Top-Down Network Design Process The next few chapters provide guidelines for selecting network design solutions for a customer. The decisions you make about protocols and technologies should be based on the information you have gathered on your customer's business and technical goals. Researchers studying decision models say that one of the most important aspects of making a sound decision is having a good list of goals. In her book The Can-Do Manager, published by the American Management Association, Tess Kirby	[redacted]

says that there are four factors involved in making sound decisions:

- Goals must be established.
- Many options should be explored.
- The consequences of the decision should be investigated.
- Contingency plans should be made.

To match options with goals, you can make a decision table, such as the one in Table 7-1. Table 7-1 shows a decision table that matches routing protocols to a fictional customer's business and technical goals. You can develop a similar table for switching protocols, campus-design technologies, enterprise-design technologies, WAN protocols, and so on. To develop the table, place options in the leftmost column and your customer's major goals at the top. Place the goals in priority order, starting with critical goals.

You can fill in Table 7-1

first by simply putting an X in each option that meets a critical goal. Any options that do not meet critical goals can immediately be eliminated. Other options can be evaluated on how well they meet other goals, on a scale from 1 to 10.

After a decision has been made, you should troubleshoot the decision. Ask yourself the following:

- If this option is chosen, what could go wrong?
- Has this option been tried before (possibly with other customers)? If so, what problems occurred?
- How will the customer react to this decision?
- What are the contingency plans if the customer does not approve of the decision?

This decision-making process can be used during both the logical and physical network design phases. You can use this

process to help you select protocols, technologies, and devices that will meet a customer's requirements.

……

*X = Meets critical criteria. 1 = Lowest. 10 = Highest.

Selecting Switching Protocols

Switches became popular in the mid-1990s as an inexpensive way of partitioning LANs without incurring the latency associated with bridges. Switches take advantage of fast integrated circuits to offer low latency. Bridges were much slower than switches and had fewer ports and a higher cost per port. For these reasons, switches have replaced bridges these days. The fundamental concepts, however, have not changed that much, and many discussions about switching concepts, including discussions in this chapter, use the term bridge.

Switches have the capability to do store-and-

forward processing or cut-through processing. With cut-through processing, a switch quickly looks at the destination address (the first field in the Ethernet frame header), determines the outgoing port, and immediately starts sending bits to the outgoing port.

A disadvantage with cut-through processing is that it forwards illegal frames (for example, Ethernet runts) and frames with CRC errors. On a network that is prone to runts and errors, cut-through processing should not be used. Some switches have the capability to automatically move from cut-through mode to store-and-forward mode when an error threshold is reached. This feature is called adaptive cut-through switching by some vendors.

A switch also supports parallel forwarding, whereas bridges usually do not. When a typical bridge is forwarding a frame from one port to another, no other frame can be forwarded. There is only one forwarding path. A switch, on the other

hand, allows multiple, parallel forwarding paths, which means a switch can handle a high volume of traffic more quickly than a bridge. High-end switches can support numerous simultaneous forwarding paths, depending on the structure of the switching fabric. (Manufacturers use the term switching fabric to describe the architecture of their switches.)

Switching and the OSI Layers

In this book, a switch refers to a device that operates at Layers 1 and 2 of the OSI reference model, unless otherwise specified. The term switch does have a more generic meaning, though. The verb to switch simply means to move something to a different position. An internetworking device moves data that comes in one interface to a different interface.

An Ethernet hub or repeater switches bits that come in one interface to all other interfaces. A hub works at Layer 1 of the OSI model and does not

understand anything beyond bits. An Ethernet switch is a high-speed, multiport bridge that switches frames based on the Layer 2 destination address. After a switch has learned the correct interface to use for a particular unicast destination, it switches frames for that destination exclusively to that interface, unlike a hub, which switches bits to all interfaces. A router switches packets based on the Layer 3 destination address. For unicast packets, the router switches exclusively to one interface. The noun switch is a good engineering term that should not be clouded by marketing hype. In electronics, a switch is a device that permits or interrupts the flow of current through an electrical circuit. In the transportation industry, a switch is a device made of two movable rails designed to turn a locomotive from one track to another. In the	[redacted]

networking field, a switch permits or interrupts the flow of data and, although it doesn't turn a locomotive, it does turn bits, frames, and packets.

Modern routers can switch packets extremely quickly. Some manufacturers add the word switch to their router product names to emphasize that the routers are as fast (or almost as fast) as Layer 2 switches. Modern routers use high-speed internal data paths, parallel processors, and advanced caching methods, all essential to the high-speed switching of data. Manufacturers call their products Layer 3 switches, routing switches, switching routers, multilayer switches, and many other creative names. In general, a Layer 3 switch, routing switch, or switching router is a device that can handle both Layer 2 and Layer 3 switching of data. A Layer 3 switch is a high-speed router that can include interfaces that make a forwarding decision based solely on Layer 2 information.

Transparent Bridging

Ethernet switches and bridges use a classic technology called transparent bridging. A transparent bridge connects one or more LAN segments so that end systems on different segments can communicate with each other transparently. An end system sends a frame to a destination without knowing whether the destination is local or on the other side of a transparent bridge. Transparent bridges are so named because their presence is transparent to end systems.

To learn how to forward frames, a transparent bridge listens to all frames and determines which stations reside on which segments. The bridge learns the location of devices by looking at the source address in each frame. The bridge develops a switching table such as the one shown in Table 7-2. The switching

table also sometimes goes by the names bridging table, MAC address table, or Content Addressable Memory (CAM) table.

MAC Address	Port
08-00-07-06-41-B9	1
00-00-0C-60-7C-01	2
00-80-24-07-8C-02	3

When a frame arrives at a bridge, the bridge looks at the destination address in the frame and compares it to entries in the switching table. If the bridge has learned where the destination station resides (by looking at source addresses in previous frames), it can forward the frame to the correct port. A transparent bridge sends (floods) frames with an unknown destination address and all multicast/broadcast frames out every port (except the port on which the frame was received).

Bridges operate at Layers 1 and 2 of the OSI reference model. They determine how to forward a frame based on information in the Layer 2 header of the frame. Unlike a router, a bridge does not look at Layer 3

information or any upper layers. A bridge segments bandwidth domains so that devices on opposite sides of a bridge do not compete with each other for media access control. A bridge does not forward Ethernet collisions or MAC frames in a Token Ring network.	[redacted]
Although a bridge segments bandwidth domains, it does not segment broadcast domains (unless programmed by filters to do so). A bridge sends broadcast frames out every port. This is a scalability issue that was already discussed in Part I of this book. To avoid excessive broadcast traffic, bridged and switched networks should be segmented with routers or divided into VLANs.	[redacted]
A bridge is a store-and-forward device. Store and forward means that the bridge receives a complete frame, determines which outgoing port to use, prepares the frame for the outgoing port, calculates a cyclic redundancy check (CRC), and transmits the	[redacted]

frame when the medium is free on the outgoing port.

Selecting Spanning Tree Protocol Enhancements

As discussed in Chapter 5, "Designing a Network Topology," transparent bridges and switches implement the Spanning Tree Protocol (STP) to avoid loops in a topology. An important design consideration is which enhancements to STP you should select to ensure high availability in your campus network designs. Chapter 5 discussed two important enhancements:

- IEEE 802.1w, which provides rapid reconvergence of the spanning tree and is now built into the IEEE 802.1D standard
- IEEE 802.1s, which aggregates a set of VLAN-based spanning trees into distinct instances and runs only one (rapid) spanning-tree algorithm per instance

As described in the next few sections, there are many other enhancements to STP that can increase

the availability and resiliency of a campus network design that relies on STP. PortFast The 2004 version of 802.1D supports the concept of a switch edge port. An edge port corresponds to the Cisco PortFast feature (and is configured with the Cisco spanning- tree portfast command). A network engineer can configure a port as an edge port if it is attached to a LAN that has no other switches attached. The Rapid Spanning Tree Protocol (RSTP) can also automatically detect edge ports. Edge ports transition directly to the forwarding state, which is a major benefit for access layer ports that connect end- user systems and IP phones. Without PortFast, a switch port lingers in the discarding and learning states before starting to forward frames, which can cause important frames to get dropped. With IP networks, the most serious	[redacted]

problem with switch port startup delay is that a client might timeout while waiting to receive an IP address from a DHCP server. With some implementations, if this happens, the client uses an address from the Automatic Private IP Addresses range (169.254.0.1 through 169.254.255.254) rather than an address assigned by a DHCP server. This address does not allow communication across a router, which means users cannot reach the Internet and possibly corporate servers as well. PortFast is meant to be used only on ports that do not connect another switch; however, sometimes this is unpredictable, especially as users and junior network administrators become more networking savvy and decide to install their own equipment. To protect a network that uses PortFast, Cisco supports a feature called BPDU Guard that shuts down a PortFast-enabled port if a	[redacted]

bridge protocol data unit (BPDU) from another switch is received. The 2004 version of RSTP also supports a similar feature and monitors an edge port for BPDUs in case a switch is connected. As soon as the switch detects a BPDU arriving at an edge port, the port becomes a non-edge port.

UplinkFast and BackboneFast

UplinkFast is a Cisco feature that can be configured on access layer switches. UplinkFast improves the convergence time of STP if a failure of a redundant uplink from an access layer switch occurs. An uplink is a connection from an access layer switch to a higher-end switch in the distribution layer of a hierarchical network design. Figure 7-1 illustrates a typical redundant, hierarchical network design. Users are connected to Switch A in the access layer. The access layer switch is attached to two

distribution layer switches. One of the uplinks is blocked by STP. (STP has also blocked one of the links between the distribution and core layers.) If the uplink to Switch B in Figure 7-1 fails, STP eventually unblocks the uplink to Switch C, hence restoring connectivity. With the default STP parameters, the recovery takes between 30 and 50 seconds. With UplinkFast, the recovery takes about 1 second. The UplinkFast feature is based on the definition of an uplink group. On a given switch, the uplink group consists of the root port and all the ports that provide an alternate connection to the root bridge. If the root port fails or the primary uplink fails, a port from the uplink group is selected to immediately replace the root port. UplinkFast should be configured only on access layer switches at the edge of your network and not on distribution or core layer switches. Figure 7-1 Access Layer Switch with Two Uplinks	[redacted]

to Distribution Layer Switches	[redacted]
Note RSTP includes a form of immediate transition to the forwarding state that is similar to the Cisco proprietary UplinkFast extension to STP. When a bridge loses its root port, it can immediately transition an alternate port into the forwarding state.	[redacted]
Cisco also supports a feature called BackboneFast, which can save a switch up to 20 seconds (Maximum Age) when recovering from an indirect link failure that occurs on a nonlocal port. When enabled on all switches in a switched network, BackboneFast speeds converge after a failure by taking advantage of the fact that a switch involved in a nonlocal failure might be able to move into the listening state immediately. In some topologies, it is not necessary for a switch to wait for the Maximum Age timer to lapse. The switch first checks with	[redacted]

other switches to determine if its status is valid. The checking is accomplished with two Cisco-proprietary protocol data units (PDU): Root Link Query (RLQ) and RLQ Response. Unidirectional Link Detection Sometimes hardware fails in such a way that connectivity between two switches works in only one direction. This is called a unidirectional link. Switch A can hear Switch B, but Switch B can't hear Switch A. This situation can be caused by Switch B's receiver being dead or weak, Switch A's transmitter being dead or weak, or some other component, such as a repeater or cable, being unable to transmit or receive. For example, a cable might be working at the physical layer (so the link is up) but be constructed incorrectly so that a switch port can transmit but not receive, even though its partner is unaware of the problem and can transmit and receive.	[redacted]

A unidirectional link can cause a loop in a switched network. If a switch port can't receive data, it can't hear BPDUs, and it might go into the forwarding state when its partner is already forwarding. If a switch port can't send data, it can't send BPDUs, and its partner might be unaware of its existence. IEEE doesn't say how to handle this situation, but vendors recognized the potential for a problem and offered fixes. Cisco provides the UniDirectional Link Detection (UDLD) protocol on high-end switches. UDLD allows devices connected through fiber-optic or copper Ethernet cables to monitor the physical configuration of the cables and detect when a unidirectional link exists. When a unidirectional link is detected, UDLD shuts down the affected port and alerts the user. A switch port that is configured to	[redacted]

use UDLD transmits UDLD messages periodically to neighbor devices. Devices on both ends of a link must support UDLD for the protocol to successfully identify and disable unidirectional links. LoopGuard Cisco also supports a feature called LoopGuard that is intended to provide additional protection against loops caused by a blocking port erroneously moving to the forwarding state. This usually happens because one of the ports of a physically redundant topology (not necessarily the blocking port) stops receiving BPDUs, perhaps because of a unidirectional link problem. If LoopGuard is enabled and BPDUs are not received on a nondesignated port, the port moves into the loop-inconsistent state instead of moving to the listening, learning, and forwarding states. Without STP LoopGuard, the port would assume the designated port role and move to the forwarding state, and thus create a	[redacted]

loop. When a BPDU is received on a port in a loop-inconsistent state, the port transitions into another STP state. This means that recovery is automatic, and no intervention is necessary. You can choose either UDLD or STP LoopGuard, or both, which is recommended. UDLD works better than LoopGuard on EtherChannel (which is a grouping of Ethernet interfaces into one logical channel). UDLD disables only the failed interface. The channel remains functional with the remaining interfaces. STP LoopGuard blocks the entire channel in such a failure (by putting it into the loop-inconsistent state). LoopGuard does not work where the link has been unidirectional since it was first brought up. The port may never receive BPDUs but not recognize that there is a problem and become a designated port. UDLD provides protection against such a problem. On the other hand, UDLD does not protect against	[redacted]

STP failures caused by software problems that result in a designated port not sending BPDUs. Software problems are less likely than hardware problems, but they could happen. Enabling both UDLD and LoopGuard provides the highest level of protection.

Protocols for Transporting VLAN Information

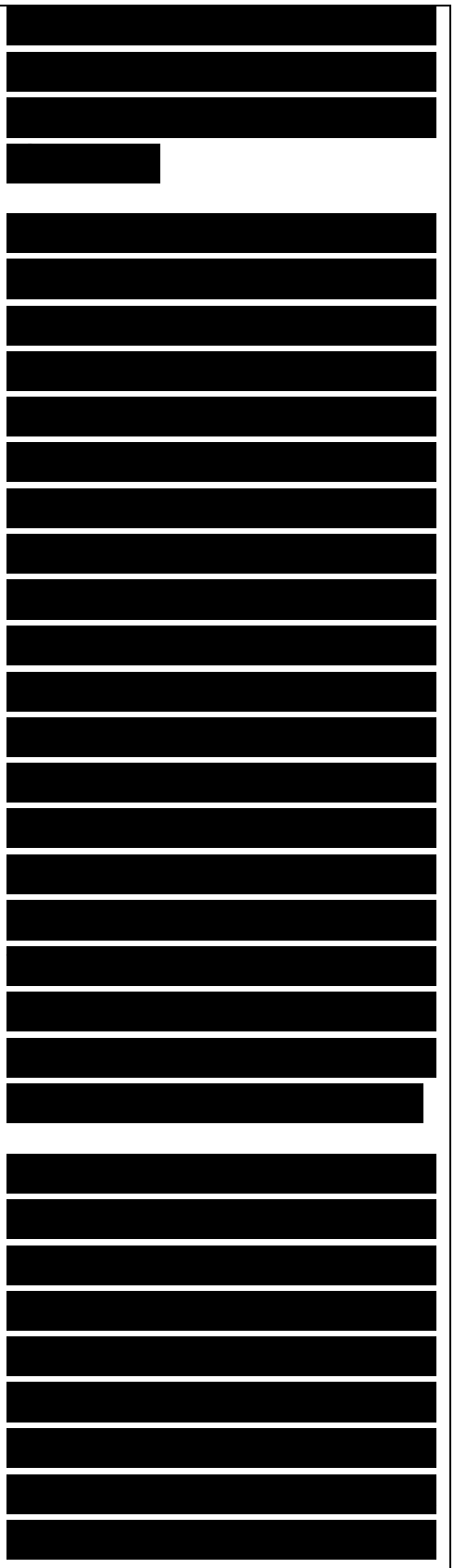
Before moving to a discussion of Layer 3 routing protocols, it is important to cover some additional Layer 2 protocols that can be deployed in switched networks that use VLANs.

When VLANs are implemented in a switched network, the switches need a method to make sure intra-VLAN traffic goes to the correct interfaces. To benefit from the advantages of VLANs, the switches need to ensure that traffic destined for a particular VLAN goes to that VLAN and not to any other VLAN. This can be accomplished by tagging frames with VLAN information using the IEEE 802.1Q standard,

which is discussed in the next section. Another important aspect of VLANs is configuration and management. This section also covers the Cisco VLAN management protocol: VLAN Trunking Protocol (VTP).

IEEE 802.1Q

In 1998, IEEE defined a standard method for tagging frames with a VLAN ID. The method is published in the revised 2006 version of the IEEE 802.1Q document "Virtual Bridged Local Area Networks." With 802.1Q, a VLAN tag is added inside an Ethernet frame. The frame is not encapsulated, as it is with the older Cisco Inter-Switch Link (ISL) protocol. Instead, 802.1Q adds a header that is inserted immediately following the desti-nation and source MAC addresses of the frame to be transmitted. This is normally where an EtherType would reside for Ethernet II frames or where the length field would reside for 802.3

frames. The EtherType or length field from the original frame is pushed forward and follows the 802.1Q header. The first 2 bytes of the 802.1Q header are the Tag Protocol Identifier (TPID) field. The TPID is set to 0x8100. Because this number is bigger than the maximum size of an Ethernet frame, a recipient knows that the field is not an 802.3 length field and that the frame is not a typical (untagged) 802.3 frame. If the recipient supports 802.1Q, it recognizes the 0x8100 as the TPID field and continues to process the rest of the fields in the 802.1Q header. If the recipient does not support 802.1Q, it sees the 2 TPID bytes as an unsupported EtherType and drops the frame. Because 802.1Q changes an Ethernet/802.3 frame, rather than encapsulating the frame as ISL does, a switch must recompute the frame check sequence (FCS) at the end of the frame, which is a minor disadvantage of 802.1Q compared to ISL. However, the CPUs on	

switches these days are so fast that recomputing the FCS does not take a significant amount of time. Some Cisco switches support only 802.1Q, and some older, obsolete Cisco switches support only ISL. Some support both. Check the Cisco product catalog for information on which trunking method is supported on a switch. Also, on some switches, you can use the show port capabilities command to display which trunking technologies are supported. Aside from differences in how they tag frames, the most important difference between ISL and 802.1Q is in their interaction with STP. Depending on how recent the software is on a switch, 802.1Q might require all VLANs to be in one spanning tree, whereas ISL allows one spanning tree per VLAN. With the addition of IEEE's 802.1s Multiple Spanning Tree (MST) standard, this might no longer be a problem, but support for 802.1s with	[redacted]

802.1Q depends on the software on the switch. Dynamic Trunk Protocol The Cisco proprietary Dynamic Trunk Protocol (DTP) supports a switch negotiating with the remote side to enable or disable 802.1Q. DTP should be recommended to your design customers, but be careful with its configuration. 802.1Q on a trunk interface can be set to on, off, desirable, auto, and nonegotiate. Nonegotiate enables 802.1Q but does not send any configuration requests to the other side. Use nonegotiate when connecting to a switch that does not support DTP. You should use the off mode whenever you do not want the local interface to be an 802.1Q trunk, but you do want it to participate in DTP to inform the remote side of its off status. Use the on mode when the remote side supports DTP, and	[redacted]

when you want the local side to remain in trunk mode regardless of the remote side's mode. The auto mode means that the switch can receive a request to enable 802.1Q and automatically enter into trunk mode. A switch configured in auto mode never initiates a request. The other side must be set to on or desirable. The desirable mode causes a switch interface to inform the remote end of its intent to enable 802.1Q but does not actually enable 802.1Q unless the remote side agrees to enable it. The remote side must be set to on, auto, or desirable for the link to use 802.1Q. Do not use the desirable mode if the remote side doesn't support DTP, because receiving DTP frames may confuse the remote switch. In general, however, when both switches sup-port DTP, Cisco recommends setting both sides to desirable. In this mode, network engineers can trust syslog and command-line status messages that a port is up	[redacted]

and trunking, whereas with on mode, a port can appear up even though the neighbor is misconfigured.

DTP was supposed to make configuring 802.1Q simpler. With all those options, however, it is easy to make a mistake, especially because some combinations are invalid and result in an 802.1Q mode mismatch. If one side is trunking and the other side is not, the switch ports won't understand each other's traffic. One side will be tagging frames and the other side will not be tagging frames. You should avoid the following combinations:

- Nonegotiate (enable 802.1Q but don't negotiate) and off (don't enable 802.1Q)
- Nonegotiate (enable 802.1Q but don't negotiate) and auto (enable 802.1Q only if the other side says to)

VLAN Trunking Protocol

The Cisco VLAN Trunking Protocol (VTP) is a switch-to-switch and switch-to-router VLAN

management protocol that exchanges VLAN configuration changes as they are made to the network. VTP manages the addition, deletion, and renaming of VLANs on a campus network without requiring manual intervention at each switch. VTP also reduces manual configuration by automatically configuring a new switch or router with existing VLAN information when the new switch or router is added to the network. In large switched networks, you should divide the network into multiple VTP domains, which reduces the amount of VLAN information each switch must maintain. A switch accepts VLAN information only from switches in its domain. VTP domains are loosely analogous to autonomous systems in a routed network where a group of routers shares common administrative policies. Multiple VTP domains are recommended on large networks. On medium-sized and small networks, a single VLAN	[redacted]

domain is sufficient and minimizes potential problems. Cisco switches can be configured in VTP server, client, or transparent mode. Server mode is the default. In VTP server mode, you can create, modify, and delete VLANs. VTP servers save their VLAN configurations when they are powered down. VTP clients exchange information with other VTP clients and servers, but you cannot create, change, or delete VLANs on a VTP client. You must do that on a VTP server. VTP clients do not save their VLAN configurations when powered down. Nevertheless, most switches should be clients to avoid VLAN information becoming desynchronized if updates are made on many switches. A VTP transparent switch does not advertise its VLAN configuration and does not synchronize its VLAN configuration based on received advertisements; however, VTP transparent switches	[redacted]

do forward received VTP advertisements to other switches. Use transparent mode when a switch in the middle of the topology does not need to match other switch configurations but would cause problems for other switches if it did not forward VLAN information. You can also use transparent mode on all switches if you prefer to manually control the configuration of VLAN information.

Selecting Routing Protocols

A routing protocol lets a router dynamically learn how to reach other networks and exchange this information with other routers or hosts. Selecting routing protocols for your network design customer is harder than selecting switching protocols, because there are so many options. The decision is made easier if you can use a decision table, such as the one shown in Table 7-1. Armed with a solid understanding of your customer's goals and information on the

characteristics of different routing protocols, you can make a sound decision about which routing protocols to recommend.

Characterizing Routing Protocols

All routing protocols have the same general goal: to share network reachability information among routers. Routing protocols achieve this goal in a variety of ways. Some routing protocols send a complete routing table to other routers. Other routing protocols send specific information on the status of directly connected links. Some routing protocols send periodic Hello packets to maintain their status with peer routers. Some routing protocols include advanced information such as a subnet mask or prefix length with route information. Most routing protocols share dynamic (learned) information, but in some cases, static configuration information is more appropriate.

Routing protocols differ in their scalability and performance characteristics. Many routing protocols were designed for small internetworks. Some routing protocols work best in a static environment and have a hard time converging to a new topology when changes occur. Some routing protocols are meant for connecting interior campus networks, and others are meant for connecting different enterprises. The next few sections provide more information on the different characteristics of routing protocols. Table 7-5 at the end of this chapter summarizes the comparison of various routing protocols.

Distance-Vector Routing Protocols

Routing protocols fall into two major classes: distance-vector protocols and link-state protocols. This chapter covers

distance-vector protocols first.

The following protocols are distance-vector protocols (or derivatives of distance-vector protocols):

- Routing Information Protocol (RIP) version 1 and 2
- Interior Gateway Routing Protocol (IGRP)
- Enhanced IGRP (EIGRP) (an advanced distance-vector protocol)
- Border Gateway Protocol (BGP) (a path-vector routing protocol)

The term vector means direction or course. A distance vector is a course that also includes information on the length of the course. Many distance-vector routing protocols specify the length of the course with a hop count. A hop count specifies the number of routers that must be traversed to reach a destination network. (For some protocols, hop count means the number of links rather than the number of

routers.)

A distance-vector routing protocol maintains (and transmits) a routing table that lists known networks and the distance to each network. Table 7-3 shows a typical distance- vector routing table.

A distance-vector routing protocol sends its routing table to all neighbors. It sends a broadcast packet that reaches all other routers on the local segment (and any hosts that use routing information). Distance-vector protocols can send the entire table each time, or they can simply send updates after the first transmission and send the complete routing table only occasionally.

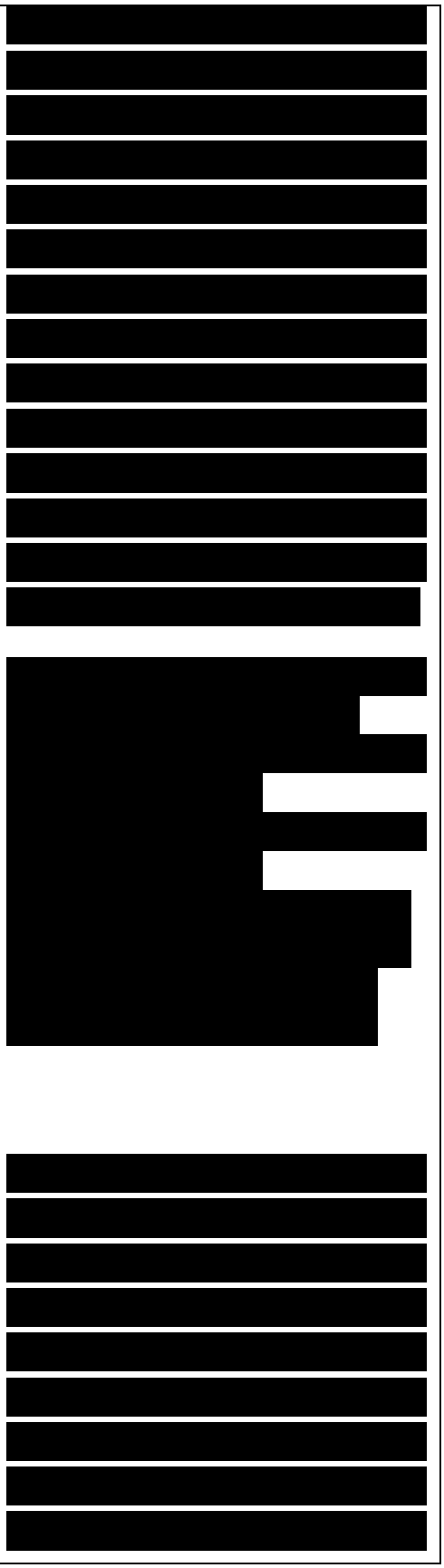
Split-Horizon, Hold-Down, and Poison-Reverse Features of Distance-Vector Protocols

A router running a distance-vector protocol sends its routing table out each of its ports on a periodic basis. If the protocol supports the split-

horizon technique, the router sends only routes that are reachable via other ports. This reduces the size of the update and, more important, improves the accuracy of routing information. With split horizon, a router does not tell another router information that is better learned locally. Table 7-3 Distance-Vector Routing Table

Network	Distance (in Hops)	Send to (Next Hop)
10.0.0.0	0 (directly connected)	Port 1
172.16.0.0	0 (directly connected)	Port 2
172.17.0.0	1	172.16.0.2
172.18.0.0	2	172.16.0.2
192.168.1.0	1	10.0.0.2
192.168.2.0	2	10.0.0.2

Most distance-vector protocols also implement a hold-down timer so that new information about a route to a suspect network is not believed right away, in case the information is based on stale data. Hold-down timers are a standard way to avoid loops that



can happen during convergence. To understand the loop problem, consider the network shown in Figure 7-2.

……..

Figure 7-2 Partial Distance-Vector Routing Tables on Router A and Router B

When routers broadcast their routing tables, they simply send the Network and Distance columns of the table. They do not send the Send To (Next Hop) column, which is one of the causes of the loop problem.

The sequence of events that can lead to a routing loop is as follows:

1. Router A's connection to Network 172.16.0.0 fails.
2. Router A removes Network 172.16.0.0 from its routing table.

3.Based on previous announcements from Router A, Router B broadcasts its routing table saying that Router B can reach network 172.16.0.0.

4. Router A adds Network 172.16.0.0 to its routing table with a Send To (Next Hop) value of Router B and a distance of 2. 5. Router A receives a frame for a host on network 172.16.0.0. 6. Router A sends the frame to Router B. 7. Router B sends the frame to Router A. The frame loops back and forth from Router A to Router B until the IP Time-To-Live (TTL) value expires (TTL is a field in the IP header of an IP packet that is decremented each time a router processes the frame). To make matters worse, without split horizon, at some point Router A sends a route update saying it can get to Network 172.16.0.0, causing Router B to update the route in its table with a distance of 3. Both Router A and Router B continue to send route updates until finally the	[redacted]

distance field reaches infinity. (Routing protocols arbitrarily define a distance that means infinity. For example, 16 means infinity for RIP.) When the distance field reaches infinity, the routers remove the route.	[redacted]
The route-update problem is called the count-to-infinity problem. A hold-down function tells a router not to add or update information for a route that has recently been removed, until a hold-down timer expires. In the example, if Router A uses hold-down, it does not add the route for network 172.16.0.0 that Router B sends. Split horizon also solves the problem in the example, because if Router B uses split horizon, it does not tell Router A about a route to 172.16.0.0.	[redacted]
Poison-reverse messages are another way of speeding convergence and avoiding loops. With	[redacted]

poison reverse, when a router learns a route from another router, it responds by sending an update back to that router that lists the distance to the network as infinity. By doing so, the router explicitly states that the route is not directly reachable via itself.	[redacted]
Triggered updates are another advanced feature of distance-vector protocols that can speed convergence. With triggered updates, a routing protocol announces route failures immediately. Rather than simply waiting for the next regularly scheduled routing update and not including in the update any routes that have failed, a router can immediately send an update. The immediate (triggered) update lists the failed route with the distance set to infinity.	[redacted]
Link-State Routing Protocols	[redacted]
Link-state routing protocols do not exchange routing tables. Instead,	[redacted]

routers running a link-state routing protocol exchange information about the links to which a router is connected. Each router learns enough information about links in the internetwork from peer routers to build its own routing table.

The following protocols are link-state routing protocols:

- Open Shortest Path First (OSPF)
- Intermediate System-to-Intermediate System (IS-IS)
- NetWare Internetwork Packet Exchange (IPX) Link Services Protocol (NLSP)

A link-state routing protocol uses a shortest-path first algorithm, such as the Dijkstra algorithm, to determine how to reach destination networks. The Dijkstra algorithm, which is named after the computer scientist who invented the algorithm, Edsger Dijkstra, solves the problem of finding the

shortest path from a source point in a mathematical graph to a destination point. One of the beauties of the algorithm is that while finding the shortest path to one destination, a source can also find the shortest path to all points in the graph at the same time. This makes the algorithm a perfect fit for a routing protocol, although it does have other uses. With link-state routing, routers use a Hello protocol to establish a relationship (called an adjacency) with neighbor routers. Each router sends link-state advertisements (LSA) to each adjacent neighbor. The advertisements identify links and metrics. Each neighbor that receives an advertisement propagates the advertisement to its neighbors. The result is that every router ends up with an identical link-state database that describes the nodes and links in the internetwork graph. Using the Dijkstra algorithm, each router independently calculates its shortest path	[redacted]

to each network and enters this information in its routing table.	[redacted]
Link-state routing requires more router CPU power and memory than distance-vector routing and can be harder to troubleshoot. Link-state routing does have some advantages over distance-vector routing, however. In general, link-state routing was designed to use less bandwidth, be less prone to loops, and converge more quickly that distance-vector routing. (Although there are distance-vector protocols, such as EIGRP, that have those qualities also.)	[redacted]
Choosing Between Distance-Vector and Link-State Protocols	[redacted]
According to Cisco design documents, you can use the following guidelines to help you decide which type of routing protocol to deploy. Choose distance-vector	[redacted]

protocols when

- The network uses a simple, flat topology and does not require a hierarchical design.
- The network uses a simple hub-and-spoke topology.
- The administrators do not have enough knowledge to operate and troubleshoot link- state protocols.
- Worst-case convergence times in the network are not a concern.

Choose link-state protocols when

- The network design is hierarchical, which is usually the case for large networks.
- The administrators are knowledgeable about the selected link-state protocol.
- Fast convergence of the network is crucial.

Routing Protocol Metrics

Routing protocols use metrics to determine which path is preferable when more than one path is available. Routing

protocols vary on which metrics are supported. Traditional dis- tance-vector routing protocols used hop count only. Newer protocols can also take into account delay, bandwidth, reliability, and other factors. Metrics can affect scalability. For example, RIP supports only 15 hops. Metrics can also affect network performance. A router that uses hop count as its sole metric misses the opportunity to select a route that has more hops but also more bandwidth than another route.

Hierarchical Versus Nonhierarchical Routing Protocols

Some routing protocols do not support hierarchy. All routers have the same tasks, and every router is a peer of every other router. Routing protocols that support hierarchy, on the other hand, assign different tasks to routers, and group routers in areas, autonomous systems, or domains. In a hierarchical arrangement, some routers communicate with local routers in the same area, and other routers have the job of connecting areas,

domains, or autonomous systems. A router that connects an area to other areas can summarize routes for its local area. Summarization enhances stability because routers are shielded from problems not in their own area. Interior Versus Exterior Routing Protocols Routing protocols can also be characterized by where they are used. Interior routing protocols, such as RIP, OSPF, and EIGRP, are used by routers within the same enterprise or autonomous system (AS). Exterior routing protocols, such as BGP, perform routing between multiple autonomous systems. BGP is used on the Internet by peer routers in different autonomous systems to maintain a consistent view of the Internet's topology.	[redacted]

Classful Versus Classless Routing Protocols

The previous chapter discussed the differences between IP classful and classless routing protocols. To summarize the concepts in Chapter 6, "Designing Models for Addressing and Naming," a classful routing protocol, such as RIP or IGRP, always considers the IP address class (Class A, B, or C). Address summarization is automatic by major network number. This means that discontiguous subnets are not visible to each other, and variable-length subnet masking (VLSM) is not supported.

Classless protocols, on the other hand, transmit prefix length or subnet mask information with IP network addresses. With classless routing protocols, the IP address space can be mapped so that discontiguous subnets and VLSM are supported. The IP address space should be mapped carefully so that subnets

are arranged in contiguous blocks, allowing route updates to be summarized at area boundaries. Dynamic Versus Static and Default Routing A static route is a route that is manually configured and does not rely on updates from a routing protocol. In some cases, it is not necessary to use a routing protocol. Static routes are often used to connect to a stub network. A stub network resides on the edge of an internetwork and isn't used as a transit path for traffic trying to get anywhere else. An example of a stub network is a company that connects to the Internet via a single link to an Internet service provider (ISP). The ISP can have a static route to the company. It is not necessary to run a routing protocol between the company and the ISP. A disadvantage with static	[redacted]

routing is the amount of administration that might be required, especially on large networks. On small (and even some large) networks, static routes have many advantages, however, and should not be overlooked when designing or upgrading a network. Static routes reduce bandwidth usage and are easy to troubleshoot. Static routes allow you to use a route other than the one that dynamic routing would choose, which can be beneficial when you want traffic to follow a specific path. Static routes can also let you use a route that is more specific than the dynamic routing protocol permits. Static routes also facilitate security because they give you more control over which networks are reachable. Most ISPs have many static routes in their routing tables to reach their customers' networks. At an ISP, when traffic arrives from other sites on the Internet with a destination address that matches the network	[redacted]

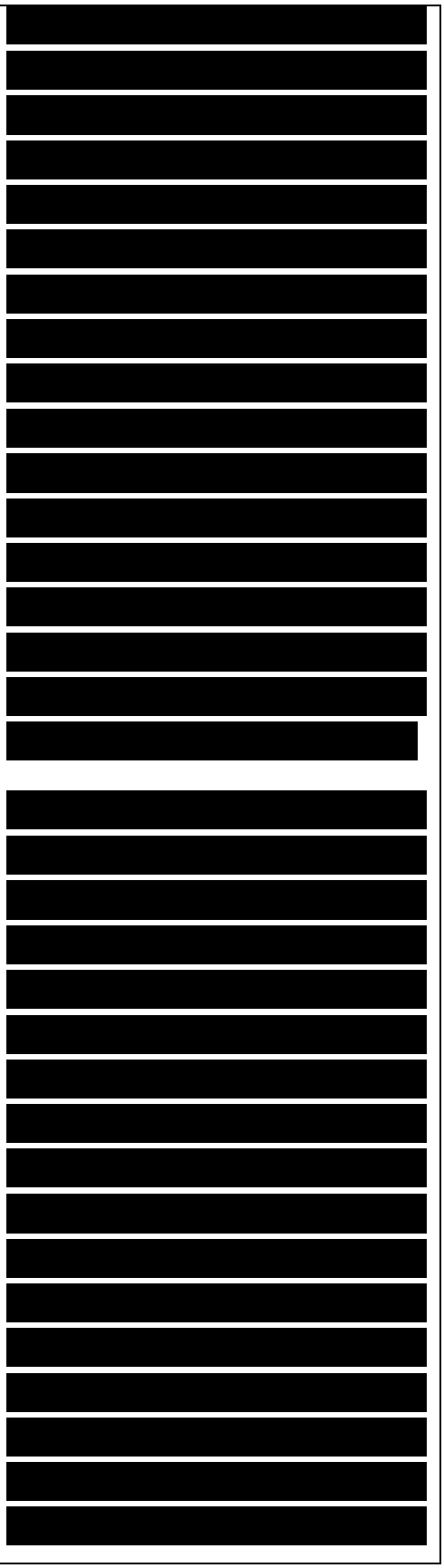
address assigned to a customer, the routing decision is simple. The traffic goes in just one direction—to the router at the customer's site. There's no need for a routing protocol.

On Cisco routers, static routes take precedence over routes to the same destination that are learned via a routing protocol. Cisco IOS Software also supports a floating static route, which is a static route that has a higher administrative distance than dynamically learned routes and can thus be overridden by dynamically learned routes. One important application of floating static routes is to provide backup routes when no dynamic infor-mation is available.

A default route is a special type of static route that is used when there is no entry in the routing table for a destination network. A default route is also

called “the route of last resort.” In some cases, a default route is all that is necessary. Take the example of the customer network connected to an ISP again. At the customer side, it might not be necessary or practical to learn routes to all the networks on the Internet. If there’s just one connection to the Internet (the link to the ISP), all Internet traffic has to go in that direction anyway. So, the enterprise network designer can simply define a default route that points to the ISP’s router.

Although static and default routes reduce resource usage, including bandwidth and router CPU and memory resources, the tradeoff is a loss of detailed information about routing. Routers with a default route always send traffic that is not local to a peer router. They have no way of knowing that the other router might have lost some of its routes. They also have no way of knowing if a destination is always unreachable (for example, when someone is doing a ping scan and sending multiple pings to



numerous IP destination addresses, some of which are not reachable). A router with a default route forwards these packets. It has no way of distinguishing destinations that it cannot reach from destinations that no routers can reach. Default routing can also cause a router to use suboptimal paths. To avoid these types of problems, use dynamic routing.

On-Demand Routing

On-Demand Routing (ODR) is a Cisco proprietary feature that provides IP routing for stub networks. ODR uses the Cisco Discovery Protocol (CDP) to carry minimal routing information between a main site and stub routers. ODR avoids the overhead of dynamic routing without incurring the configuration and management overhead of static routing.

The IP routing information required to represent the network topology for a router in a stub network is fairly simple. For example, in a hub-and-spoke topology (see Figure 5-5 in Chapter 5), stub routers at the spoke sites have a WAN connection to the hub router and a small number of LAN segments directly connected to the stub router. These stub networks might not require the stub router to learn any dynamic IP routing information. With ODR, the hub router provides default route information to the stub routers, thereby eliminating the need to configure a default route on each stub router. Stub routers use CDP to send IP prefix information for directly connected interfaces to the hub. The hub router installs stub network routes in its routing table. The hub router can also be configured to redistribute these routes into any configured dynamic IP	[redacted]

routing protocols. On the stub router, no IP routing protocol is configured. This simplifies configuration and is often a good solution for the access layer of a hierarchically designed network.

Scalability Constraints for Routing Protocols

When selecting a routing protocol for a customer, consider your customer's goals for scaling the network to a large size and investigate the following questions for each routing protocol. Each of the following questions addresses a scalability constraint for routing protocols:

- Are there any limits placed on metrics?

- How quickly can the routing protocol converge when upgrades or changes occur? Link-state protocols tend to converge more quickly than distance-vector protocols. Convergence is discussed in more detail in the next section.

- How often are

routing updates or LSAs transmitted? Is the frequency of updates a function of a timer, or are updates triggered by an event, such as a link failure?	[redacted]
■ How much data is transmitted in a routing update? The whole table? Just changes? Is split horizon used?	[redacted]
■ How much bandwidth is used to send routing updates? Bandwidth utilization is particularly relevant for low-bandwidth serial links.	[redacted]
■ How widely are routing updates distributed? To neighbors? To a bounded area? To all routers in the AS?	[redacted]
■ How much CPU utilization is required to process routing updates or LSAs?	[redacted]
■ Are static and default routes supported?	[redacted]
■ Is route	

summarization supported?	[redacted]
These questions can be answered by watching routing protocol behavior with a protocol analyzer and by studying the relevant specifications or RFCs. The next few sections in this chapter can also help you understand routing protocol behavior better.	[redacted]
Routing Protocol Convergence Convergence is the time it takes for routers to arrive at a consistent understanding of the internetwork topology after a change takes place. A change can be a network segment or router failing, or a new segment or router joining the internetwork. To understand the importance of quick convergence for your particular customer, you should develop an understanding of the likelihood of frequent changes on the customer's network. Are there links that tend to fail often? Is	[redacted]

the customer's network always "under construction" either for enhancements or because of reliability problems? Because packets might not be reliably routed to all destinations while convergence is taking place, convergence time is a critical design constraint. The convergence process should complete within a few seconds for time-sensitive applications, such as voice applications and Systems Network Architecture (SNA)-based applications. When SNA is transported across an IP internetwork, a fast-converging protocol such as OSPF is recommended. Link-state protocols were designed to converge quickly. Some newer distance-vector protocols, such as EIGRP, were also designed for quick convergence. A router starts the convergence process when it notices that a link to one of its peer routers has failed. A Cisco router sends keepalive frames every 10 seconds (by default) to help it	[redacted]

determine the state of a link. On a point-to-point WAN link, a Cisco router sends keepalive frames to the router at the other end of the link. On LANs, a Cisco router sends keepalive frames to itself.	[redacted]
If a serial link fails, a router can start the convergence process immediately if it notices the Carrier Detect (CD) signal drop. Otherwise, a router starts the convergence after sending two or three keepalive frames and not receiving a response. On an Ethernet network, if the router's own transceiver fails, it can start the convergence process immediately. Otherwise, the router starts the convergence process after it has been unable to send two or three keepalive frames. If the routing protocol uses Hello packets and the Hello timer is shorter than the keepalive timer, the routing protocol can start convergence sooner. Another factor that influences convergence	[redacted]

time is load balancing. If a routing table includes multiple paths to a destination, traffic can immediately take other paths when a path fails. Load balancing was discussed in more detail in Chapter 5.

IP Routing

The most common IP routing protocols are RIP, EIGRP, OSPF, IS-IS, and BGP. The following sections describe some of the performance and scalability characteristics of these protocols to help you select the correct protocols for your network design customer.

Routing Information Protocol

The IP Routing Information Protocol (RIP) was the first standard routing protocol developed for TCP/IP environments. RIP was developed originally for the Xerox Network System (XNS) protocols and was adopted by the IP community in the early

1980s. RIP was the most common interior routing protocol for many years, probably because it is easy to configure and runs on numerous operating systems. It is still in use on older networks and networks where simplicity and ease of troubleshooting are important. RIP version 1 (RIPv1) is documented in RFC 1058. RIP version 2 (RIPv2) is documented in RFC 2453. RIP broadcasts its routing table every 30 seconds. RIP allows 25 routes per packet, so on large networks, multiple packets are required to send the whole routing table. Bandwidth utilization is an issue on large RIP networks that include low-bandwidth links. To avoid routing loops during convergence, most implementations of RIP include split horizon and a hold-down timer. RIP uses a single routing metric (hop count) to measure the distance to a destination network. This limitation should be	[redacted]

considered when designing networks that use RIP. The limitation means that if multiple paths to a destination exist, RIP maintains only the path with the fewest hops, even if other paths have a higher aggregate bandwidth, lower aggregate delay, less congestion, and so on.

Another limitation of RIP is that the hop count cannot go above 15. If a router receives a routing update that specifies that a destination is 16 hops away, the router purges that destination from its routing table. A hop count of 16 means the distance to the destination is infinity—in other words, the destination is unreachable.

RIPv1 is a classful routing protocol, which means that it always considers the IP network class. Address summarization is automatic by major network number. This means that discontiguous subnets are not visible to each other, and VLSM is not supported. RIPv2, on the other hand, is classless.

The Internet Engineering

Task Force (IETF) developed RIPv2 to address some of the scalability and performance problems with RIPv1. RIPv2 adds the following fields to route entries within a routing table:

- Route tag: Distinguishes internal routes that are within the RIP routing domain from external routes that have been imported from another routing protocol or a different autonomous system

- Subnet mask: Contains the subnet mask that is applied to the IP address to yield the nonhost (prefix) portion of the address
- Next hop: Specifies the immediate next-hop IP address to which packets to the destination in the route entry should be forwarded

Route tags facilitate merging RIP and non-RIP networks. Including the subnet mask in a route entry provides support for classless routing. The

purpose of the next-hop field is to eliminate packets being routed through extra hops. Specifying a value of 0.0.0.0 in the next-hop field indicates that routing should be via the originator of the RIP update. Specifying a different value than 0.0.0.0 is useful when RIP is not in use on all routers in a network. RIPv2 also supports simple authentication to foil hackers sending routing updates. The authentication scheme uses the space of a route entry. This means that there can be only 24 route entries in a message when authentication is used. Currently, the only authentication supported is a simple plain-text password. Enhanced Interior Gateway Routing Protocol Cisco developed the proprietary distance-vector Interior Gateway Routing Protocol (IGRP) in the mid-1980s to meet	[redacted]

the needs of customers requiring a robust and scalable interior routing protocol. Many customers migrated their RIP networks to IGRP to overcome RIP's 15-hop limitation and reliance on just one metric (hop count). IGRP's 90-second update timer for sending route updates was also more attractive than RIP's 30-second update timer for customers concerned about bandwidth utilization. Cisco developed the proprietary Enhanced IGRP (EIGRP) in the early 1990s to meet the needs of enterprise customers with large, complex, multiprotocol internetworks. EIGRP is compatible with IGRP and provides an automatic redistribution mechanism to allow IGRP routes to be imported into EIGRP, and vice versa. EIGRP can also redistribute routes for RIP, IS-IS, BGP, and OSPF. EIGRP uses a composite metric based on the following factors: ■ Bandwidth: The	[redacted]

bandwidth of the lowest-bandwidth segment on the path. A network administrator can configure bandwidth or use the default value, which is based on the type of link. (Configuration is recommended for high-speed WAN links if the default bandwidth value is less than the actual speed.) ■ Delay: A sum of all the delays for outgoing interfaces in the path. Each delay is inversely proportional to the bandwidth of each outgoing interface. Delay is not dynamically calculated. ■ Reliability: The reliability of the path, based on the interface reliability reported by routers in the path. In an EIGRP update, reliability is an 8-bit number, where 255 is 100 percent reliable and 1 is minimally reliable. By default, reliability is not used unless the metric weights command is configured, in which case it is dynamically calculated.	[redacted]

■ Load: The load on the path, based on the interface load reported by routers in the path. In an EIGRP update, reliability is an 8-bit number, where 255 is 100 percent loaded and 1 is minimally loaded. By default, load is not used unless the metric weights command is configured, in which case load is dynamically calculated. Note Use the metric weights command to change the default behavior of EIGRP with caution and only under the guidance of an experienced engineer. EIGRP allows load balancing over equal-metric paths and nonequal-metric paths. The EIGRP variance feature means that if one path is three times better than another, the better path can be used three times more than the other path. Only routes with metrics that are within a certain range of the best route can be used as	[redacted]

multiple paths. Refer to the Cisco configuration documentation for more information. EIGRP has a better algorithm for advertising and selecting a default route than RIP does. RIP allows a network administrator to configure one default route, which is identified as network 0.0.0.0. EIGRP, on the other hand, allows real networks to be flagged as candidates for being a default. Periodically, EIGRP scans all candidate default routes and chooses the one with the lowest metric to be the actual default route. This feature allows more flexibility and better performance than RIP's static default route. To reduce convergence time, EIGRP supports triggered updates. A triggered router sends a update in response to a change (for example, the failure of a link). Upon receipt of a triggered update, other routers can also send triggered updates. A failure causes a wave of update messages to propagate throughout	[redacted]

the network, thus speeding convergence time and reducing the risk of loops.	[redacted]
EIGRP has many advanced features and behaviors not found in IGRP or other distance-vector protocols. Although EIGRP still sends vectors with distance information, the updates are as follows:	[redacted]
■ Nonperiodic means that updates are sent only when a metric changes rather than at regular intervals.	[redacted]
■ Partial means that updates include only routes that have changed, not every entry in the routing table.	[redacted]
■ Bounded means that updates are sent only to affected routers.	[redacted]
These behaviors mean that EIGRP uses little bandwidth.	[redacted]
Unlike IGRP, EIGRP updates carry a prefix length with each network number, which makes EIGRP a classless	[redacted]

protocol. By default, EIGRP summarizes routes on the classful network boundaries, however. Automatic summarization can be turned off and manual summarization used instead, which can be helpful when a network includes discontiguous subnets.	[redacted]
One of the main goals of EIGRP is to offer quick convergence on large networks. To meet this goal, the designers of EIGRP adopted the diffusing-update algorithm (DUAL) that Dr. J. J. Garcia-Luna-Aceves developed at SRI International. DUAL specifies a method for routers to store neighbors' routing information so that the routers can quickly switch to alternative routes. Routers can also query other routers to learn alternative routes and send Hello packets to determine the reachability of neighbors. DUAL guarantees a loop- free topology, so there is no need for a hold-down mechanism, which is another feature that	[redacted]

minimizes convergence time.	[redacted]
DUAL is one reason that EIGRP uses significantly less bandwidth than IGRP or other dis- tance-vector protocols. A router using DUAL develops its routing table using the concept of a feasible successor. A feasible successor is a neighboring router that has the least-cost path to a destination. When a router detects that a link has failed, if a feasible successor has an alternate route, the router switches to the alternate route immediately, without causing any network traffic. If there is no successor, the router sends a query to neighbors. The query propagates across the network until a new route is found.	[redacted]
An EIGRP router develops a topology table that contains all destinations advertised by neighboring routers. Each entry in the table contains a destination and a list of neighbors that have	[redacted]

advertised the destination. For each neighbor, the entry includes the metric that the neighbor advertised for that destination. A router computes its own metric for the destination by using each neighbor's metric in combination with the local metric the router uses to reach the neighbor. The router compares metrics and determines the lowest-cost path to a destination and a feasible successor to use in case the lowest-cost path fails.

EIGRP can scale to thousands of routing nodes. To ensure good performance in large internetworks, EIGRP should be used on networks with simple hierarchical topologies.

Open Shortest Path First

In the late 1980s, the IETF recognized the need to develop an interior link-state routing protocol to meet the needs of large enterprise networks that were constrained by the limitations of RIP. The

Open Shortest Path First (OSPF) routing protocol is a result of the IETF's work. OSPF is defined in RFC 2328.

The advantages of OSPF are as follows:

- OSPF is an open standard supported by many vendors.
- OSPF converges quickly.
- OSPF authenticates protocol exchanges to meet security goals.
- OSPF supports discontiguous subnets and VLSM.
- OSPF sends multicast frames, rather than broadcast frames, which reduces CPU utilization on LAN hosts (if the hosts have NICs capable of filtering multicasts).
- OSPF networks can be designed in hierarchical areas, which reduces memory and CPU requirements on routers.
- OSPF does not use a lot of bandwidth.

To minimize bandwidth utilization, OSPF propagates only changes. Other network traffic is limited to database-synchronization packets that occur infrequently (every 30 minutes) and Hello packets that establish and maintain neighbor adjacencies and are used to elect a designated router on LANs. Hellos are sent every 10 seconds. On dialup and ISDN links configured as demand circuits, OSPF can be even quieter. In this case, OSPF routers sup-press Hellos and the database-synchronization packets. Upon startup and when there are changes, an OSPF router multicasts LSAs to all other routers within the same hierarchical area. OSPF routers accumulate link-state information to calculate the shortest path to a destination network. The calculation uses the Dijkstra algorithm. The result of the calculation is a database of the topology,	[redacted]

called the link- state database. Each router in an area has an identical database.	[redacted]
All routers run the same algorithm, in parallel. From the link-state database, each router constructs a tree of shortest paths, with itself as the root of the tree. The shortest-path tree provides the route to each destination. Externally derived routing information appears on the tree as leaves. When several equal-cost routes to a destination exist, traffic is distributed equally among them.	[redacted]
According to RFC 2328, the cost of a route is described by "a single dimensionless metric" that is "configurable by a system administrator." A cost is associated with the output side of each router interface. The lower the cost, the more likely the interface is to be used to forward data traffic. A cost is also associated with externally derived routes	[redacted]

(for example, routes learned from a different routing protocol).	[redacted]
On a Cisco router, the cost of an interface defaults to 100,000,000 divided by the bandwidth for the interface. For example, both FDDI and 100-Mbps Ethernet have a cost of 1. The cost can be manually configured. Usually it is best if both ends of a link use the same cost. If a Cisco router is at one end of a link and a non-Cisco router is at the other end, you might need to manually configure the cost. Because OSPF defines the cost metric so broadly, vendors are not required to agree on how the cost is defined.	[redacted]
Note Cisco OSPF implementation uses a reference bandwidth of 100 Mbps for cost calculation. The formula to calculate the cost for an interface is the reference bandwidth divided by the interface bandwidth. For example, in the case of a 10-Mbps Ethernet	[redacted]

interface, the interface cost is 100 Mbps divided by 10 Mbps, or 10. In an internetwork with high-speed links of 100 Mbps or higher, you should change the reference bandwidth to a number higher than 100 Mbps. You can use the ospf auto-cost reference-bandwidth command to make that change.

OSPF Architectures

OSPF allows sets of networks to be grouped into areas. The topology of an area is hidden from the rest of the autonomous system. By hiding the topology of an area, routing traffic is reduced. Also, routing within the area is determined only by the area's own topology, providing the area protection from bad routing data. By dividing routers into areas, the memory and CPU requirements for each router are limited.

A contiguous backbone area, called Area 0, is required when an OSPF network is divided into areas. Every other area connects to Area 0 via an Area Border Router

(ABR), as shown in Figure 7-3. All traffic between areas must travel through Area 0, which should have high availability, throughput, and bandwidth. Area 0 should be easy to manage and troubleshoot. A set of routers in a rack connected via a high-speed LAN makes a good Area 0 for many customers.

Area 0 (Backbone)

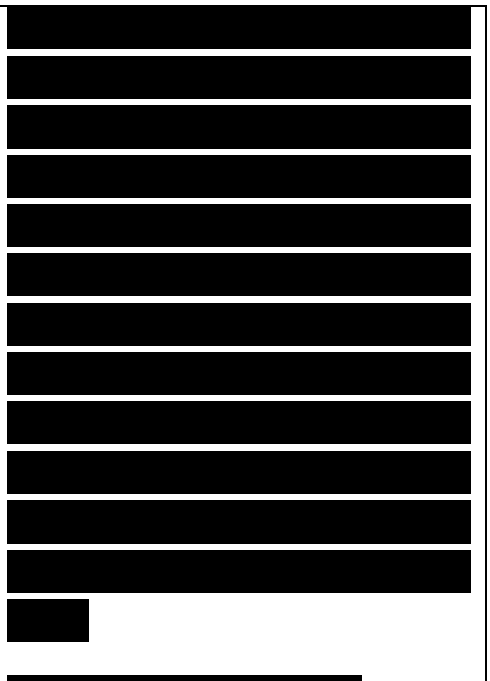


Figure 7-3 OSPF Areas Connected via ABRs

In addition to ABRs, an OSPF network might include one or more Autonomous System Boundary Routers (ASBR), which connects an OSPF network to a different AS or to a network that uses a routing protocol other than OSPF. For example, an ASBR could connect an internal OSPF campus network to the Internet.

When designing an OSPF network, make sure to assign network numbers in blocks that can be summarized. An ABR should summarize routes

behind it to avoid routers in the backbone and other areas having to know details about a particular area. The summarization must be configured on Cisco routers with the area-range command.	[redacted]
An ABR that connects a stub network can be configured to inject a default route into the stub area for all external networks that are outside the AS or are learned from other routing protocols. The router can also be configured to inject a default route for internal summarized or nonsummarized routes to other areas. If a router injects a default route for all routes, Cisco calls the area a totally stubby area, which are a Cisco technique that works as long as all the stubby areas' ABRs are Cisco routers.	[redacted]
Cisco also supports not-so-stubby areas, which allows the redistribution of external routes into OSPF in an otherwise	[redacted]

stubby area. Not-so-stubby areas are specified in RFC 1587. Not-so-stubby areas are not common, but they can be used on a stub network that includes a legacy link to another routing protocol or AS that is different from the link used by the rest of the internetwork to reach the outside world. Because of the requirement that OSPF be structured in areas and the recommendation that routes be summarized, it can be difficult to migrate an existing network to OSPF. Also, enlarging an existing OSPF network can be challenging. If a network is subject to rapid change or growth, OSPF might not be the best choice. For most networks, however, OSPF is a good choice because of its low-bandwidth utilization, scalability, and compatibility with multiple vendors. Intermediate System-to-Intermediate System Intermediate System-to-	[redacted]

Intermediate System (IS-IS) is a dynamic link-state protocol developed for use with the Open System Interconnection (OSI) protocol suite. Integrated IS-IS is an implementation of IS-IS for mixed OSI and IP networks that has gained limited popularity for use within large, hierarchical IP networks, especially within the core of large ISPs. IS-IS is a classless, interior routing protocol that is similar in operation to OSPF although somewhat more flexible, efficient, and scalable.

As with OSPF, IS-IS can be implemented in a hierarchical fashion. A router can play different roles:

- Level 1 routers route within an area.
- Level 2 routers route between areas.
- Level 1-2 routers participate in Level 1 intra-area routing and Level 2 interarea routing.

In IS-IS, the boundary

between areas is on a link between routers. A router belongs to only one area. In comparison, in OSPF, area boundaries reside inside an ABR. Some router interfaces in the ABR belong to one area and other interfaces belong to another area. With IS-IS, all router interfaces are in the same area. This makes IS-IS somewhat more modular and means that in some cases, an area can be upgraded without affecting as many routers.	[redacted]
Level 1 routers within an area (including Level 1-2 routers) maintain an identical link-state database that defines their area's topology. Level 2 (including Level 1-2 routers) also maintain a separate link-state database for the Level 2 topology.	[redacted]
Unlike OSPF ABRs, Level 1-2 routers do not advertise Level 2 routes to Level 1 routers.	[redacted]

A Level 1 router has no knowledge of destinations outside of its own area. This makes ISIS more efficient than OSPF with regard to CPU use and the processing of routing updates, although similar functionality can be implemented on Cisco OSPF routers by configuring a totally stubby area. Note From a Level 1 router's point of view, sending traffic outside the area involves finding the nearest Level 2 router. The Level 1 router relies on the Level 2 router to reach the destination. The path taken can be suboptimal if some of an area's exit-point Level 2 routers are poorly located or have poor connectivity to the destination network. The choice of exit router in IS-IS is not based on detailed information. On the positive side, this behavior means Level 1 routers require less processing. The set of Level 2 routers (including Level 1-2 routers) and their interconnecting links	[redacted]

comprise the IS-IS backbone. As with OSPF, interarea traffic must traverse this backbone. OSPF has a central backbone (Area 0) that physically attaches all the other areas. A consistent IP addressing structure is necessary to summarize addresses into the backbone and reduce the amount of information that is carried in the backbone and advertised across the network. In comparison, an IS-IS backbone can be a set of distinct areas interconnected by a chain of Level 2 and Level 1-2 routers, weaving their way through and between Level 1 areas. Compared to OSPF, IS-IS allows a more flexible approach to extending the backbone by adding additional Level 2 routers.

Border Gateway Protocol

The IETF developed the Border Gateway Protocol (BGP) to replace the now-obsolete Exterior Gateway Protocol (EGP) as the standard exterior routing protocol for the Internet.

BGP solves problems that EGP had with reliability and scalability. BGP4, the current version of BGP, is specified in RFC 1771. Internal BGP (iBGP) can be used at a large company to route between domains. External BGP (eBGP) is used to route between companies and to participate in global Internet routing. eBGP is often used to multihome an enterprise's connection to the Internet. It is a common misconception that multihoming requires BGP, but this is not true. Depending on a customer's goals and the flexibility of their ISP's policies, you can multihome with default routes, as discussed in Chapter 5. Running eBGP can be challenging, requiring an understanding of the complexities of Internet routing. eBGP should be recommended only to companies that have senior network engineers, and a good relationship with their ISPs. An inexperienced network engineer can configure eBGP in such a	[redacted]

way as to cause problems for the entire Internet. Also, eBGP should be used only on routers with a lot of memory, a fast CPU, and a high-bandwidth connection to the Internet. An Internet routing table contains more than 100,000 routes and is continually growing as the Internet expands and more companies use BGP to multihome.	[redacted]
The main goal of BGP is to allow routers to exchange information on paths to destination networks. Each BGP router maintains a routing table that lists all feasible paths to a particular network. BGP routers exchange routing information upon initial startup, and then send incremental updates, using TCP for reliable delivery of BGP packets. An update specifies path attributes, which include the origin of the path information, a sequence of autonomous-system path segments, and next-hop information.	[redacted]

When a BGP router receives updates from multiple autonomous systems that describe different paths to the same destination, the router must choose the single best path for reaching that destination. After the best path is chosen, BGP propagates the best path to its neighbors. The decision is based on the value of attributes in the update (such as next hop, administrative weight, local preference, the origin of the route, and path length) and other BGP-configurable factors.

Using Multiple Routing Protocols in an Internetwork

When selecting routing protocols for a customer, it is important to realize that you do not have to use the same routing protocols throughout the internetwork. The criteria for selecting protocols are different for different parts of an internetwork. Also, when merging a new network with an old network, it is often necessary to run more than one routing protocol. In

some cases, your network design might focus on a new design for the core and distribution layers and need to interoperate with existing access layer routing protocols. As another example, when two companies merge, sometimes each company wants to run a different routing protocol.

This section summarizes some recommendations for selecting a routing protocol for different layers of the hierarchical design model and then discusses redistribution between routing protocols. The section ends with a quick discussion of Integrated Routing and Bridging (IRB), a Cisco IOS method for connecting bridged and routed networks in a single router.

Routing Protocols and the Hierarchical Design Model

As discussed in Chapter 5, large internetworks are designed using a modular, hierarchical approach. One such approach is the three-layer hierarchical design model, which has a core

layer that is optimized for availability and performance, a distribution layer that implements policies, and an access layer that connects users. The next three sections discuss routing protocols for the three layers of the model.

Routing Protocols for the Core Layer

The core layer should incorporate redundant links and load sharing between equal-cost paths. It should provide immediate response if a link failure occurs and adapts quickly to change. Routing protocols that meet these needs include EIGRP, OSPF, and IS-IS. The decision to use EIGRP, OSPF, or IS-IS should be based on the underlying topology, IP addressing design, vendor preferences, and other business and technical goals.

OSPF imposes a strict hierarchical design. OSPF areas must map to the IP addressing plan, which

can be difficult to achieve. EIGRP and IS-IS are more flexible with regard to the hierarchical structure and IP addressing design. EIGRP is a Cisco proprietary protocol, however. Although Cisco has licensed it to a few vendors, if other vendors' products are to be used in the implementation of the network design, EIGRP might not be a good solution. Alternatively, EIGRP can be used in the core with redistribution to another routing protocol in the distribution layer. RIP is not recommended as a core routing protocol. Its response to change is slow, which can result in disrupted connectivity. Routing Protocols for the Distribution Layer The distribution layer represents the connection point between the core and access layers. Routing protocols used in the distribution layer include RIPv2, EIGRP, OSPF, and ISIS. The distribution layer also sometimes uses ODR. The distribution	[redacted]

layer often has the job of redistributing between routing protocols used in the core layer and those used in the access layer.

Routing Protocols for the Access Layer

The access layer provides access to network resources for local and remote users. As with the distribution and core layers, the underlying topology, IP addressing design, and vendor preferences drive the choice of routing protocol. Access layer equipment may be less powerful than distribution and core layer equipment, with regard to processing power and memory, which influences the routing protocol choice.

Routing protocols that should be used in the access layer include RIPv2, OSPF, and EIGRP. The distribution layer also sometimes uses ODR. Use of static routing is also a possibility. IS-IS is not often appropriate for the access layer because it demands more knowledge to configure and is also not well suited to dialup

networks. The limitations of using OSPF as an access layer routing protocol are connected to its high memory and processing power requirements and strict hierarchical design requirements. The high memory and processing power requirements of OSPF can be avoided with the use of summarization and careful area planning, however. Redistribution Between Routing Protocols Redistribution allows routers to run more than one routing protocol and share routes among routing protocols. Implementing redistribution can be challenging because every routing protocol behaves differently and routing protocols cannot directly exchange information about routes, prefixes, metrics, link states, and so on. Redistribution can lead to routing loops if not configured carefully, and can complicate planning and troubleshooting.	[redacted]

Despite the challenges, redistribution may be desirable when connecting different layers of the hierarchical model, when migrating to a new routing protocol, when different departments use different protocols, or when there is a mixed-vendor environment. A network administrator configures redistribution by specifying which protocols should insert routing information into other protocols' routing tables. Redistribution design involves determining the routing protocols to be used in a network and the extent of each routing domain. A routing domain in this context is a set of routers that share information through the use of a single routing protocol. The designer must determine where the boundaries between routing domains reside and where redistribution must occur. Redistribution is most often needed in the	[redacted]

distribution layer where routing domains intersect.	[redacted]
Another decision is whether to use one- or two-way redistribution. With one-way redistribution, routing information is redistributed from one routing protocol to another, but not vice versa. Static or default routes can be used in the opposite direction to provide connectivity. With two-way redistribution, routing information is redistributed from one routing protocol to another and vice versa. Complete routing information can be exchanged or filtering can be used to limit the information that is exchanged.	[redacted]
In most hierarchical designs, you will probably use one-way rather than two-way redistribution. When you do use two-way redistribution, you will probably not redistribute all routes from one domain to the other. Most	[redacted]

designs require you to filter so that only a subset of the routes in one domain is redistributed into the other. Filtering routes may be necessary to avoid routing loops and to maintain security, availability, and performance. You should make sure that a routing protocol does not inject routes into another routing protocol that already has a better way to reach the advertised networks. This can be accomplished with filtering. If a router is misconfigured (or maliciously reconfigured by a hacker) so that it begins advertising routes that do not belong to its domain, traffic flow can be affected. The problem might result in suboptimal routing or, worse, prevent traffic from reaching its destination. A good design practice is to filter routes so that a router receives only expected routes from other domains.	[redacted]

Filtering is also used to enhance performance. A large network could have hundreds or even thousands of routes. If all routes are redistributed into a network with smaller routers, the size of the routing table can overwhelm the smaller routers, degrading network performance. The router can get bogged down searching the large routing table for the next-hop address. The large routing table could also exceed the router's memory, causing the router to fail altogether.	[redacted]
Redistribution configuration should also be done with care to avoid feedback. Feedback happens when a routing protocol learns about routes from another protocol and then advertises these routes back to the other routing protocol. For example, if a router is configured to redistribute EIGRP routes into a RIPv2 domain, and also configured to redistribute routes back	[redacted]

into EIGRP, the router must filter any routes that it learned from EIGRP before redistributing routes into EIGRP. This avoids any problems caused by the differences in metrics used by different routing protocols.

Resolving Incompatible Metrics

When redistributing from one routing protocol to another, you will need to make some decisions about metrics. Routing protocols use different metrics that cannot easily be converted to each other. Instead of attempting a conversion, you should simply make a decision on what the metric should be for routes that originated with a different routing protocol. For example, you may decide that all EIGRP routes start out with a hop count of 1 when redistributed into a RIPv2 domain. Or you may decide that all OSPF routes start out with a bandwidth of 1000 and a delay of 100 when redistributed into EIGRP.

Administrative Distances

Another factor that makes redistribution challenging is the possibility that a router can learn about a destination via more than one routing protocol. Every routing protocol and every vendor handles this issue differently. Cisco assigns an administrative distance to routes learned from different sources. A lower administrative distance means that a route is preferred. For example, if a router learns about a route via both OSPF and RIPv2, the OSPF route is preferred because OSPF has a default administrative distance of 110 and RIPv2 has a default administrative distance of 120. If a router also has a static route to the destination, the static route is preferred because the default administrative distance for a static route is 1. Table 7-4 outlines some common administrative distances by type of route.

Note Cisco IOS Software also supports a floating

static route, which is a static route that has a higher administrative distance than a dynamically learned route. Floating static routes are available for IP, IPX, and AppleTalk. A floating static route is a statically configured route that has a high administrative distance so that it can be overridden by dynamically learned routing information. A floating static route can be used to create a "path of last resort" that is used only when no dynamic information is available. One important application of floating static routes is to provide backup routes in topologies where dial-on-demand routing (DDR) is used.

Table 7-4 Administrative Distance by Route Type

Integrated Routing and Bridging

For customers who need to merge bridged and routed networks, Cisco IOS Software offers support for IRB, which connects VLANs and

bridged networks to routed networks within the same router. An older Cisco IOS feature, Concurrent Routing and Bridging (CRB), supported routing and bridging within the same router, but it simply meant that you could connect bridged networks to other bridged networks and routed networks to other routed networks. IRB extends CRB by providing the capability to forward packets between bridged and routed interfaces via a software-based interface called the bridged virtual interface (BVI). One advantage of IRB is that a bridged IP subnet or VLAN can span a router. This can be useful when there is a shortage of IP subnet numbers and it is not practical to assign a different subnet number to each interface on a router. IRB can also be useful during migration from a bridged environment to a routed or VLAN environment.	[redacted]

A Summary of Routing Protocols

Table 7-5 provides a comparison of various routing protocols to help you select a routing protocol based on a customer's goals for adaptability, scalability, affordability, security, and network performance.

Table 7-5 Routing Protocol Comparisons

Summary

This chapter provided information to help you select the right switching and routing protocols for your network design customer. The chapter covered scalability and performance characteristics of the protocols and talked about how quickly protocols can adapt to changes.

Deciding on the right switching and routing protocols for your customer will help you select the best switch and router products for the customer. For example, if you have decided that the design must support a routing protocol that can converge within seconds

in a large internetwork, you will probably not recommend a router that runs only RIP. This chapter began with a generic discussion about decision making to help you develop a systematic process for selecting network design solutions. A discussion of bridging and switching protocols followed, which covered transparent bridging, multilayer switching, enhancements for STP, and VLAN protocols. A section on routing protocols followed the switching section. Table 7-5 summarized the comparisons that were made of various routing protocols in the routing section.	[redacted]