Internet Service Provider Configuration Guidelines

Version 3.0

October 2002

Bruce Baikie

Steve Gaede

Jim Baty

Sun microsystems

Contents

Introduction 1

If anything remains constant in the Internet Service Provider (ISP) market, it is rapid growth and constant change. In February 2001, the registry at *thelist.internet.com* listed more than 9,500 providers, indicating that the overall number of ISPs is increasing at a moderate pace — however in traffic and number of subscribers, the market continues to explode. Internet traffic exceeded voice traffic for the first time in 1998, *surpassing in less than a decade* the telephony market that has been *growing for nearly a century!* With skyrocketing growth in Europe, Asia, and Latin America, one estimate of 400 million worldwide Internet users (*www.nua.com*) indicates that users in the United States and Canada are now in the minority.

ISPs have managed to escape the long-predicted market consolidation by diversifying and providing a range of services limited only by the imagination:

- Many local and regional ISPs have solidified their niches by maintaining close relationships with their customers and providing friendly service and technical support.
- Huge national and international ISPs provide rich content and services for their subscribers, hoping to keep a significant portion of traffic within their own networks.
- Some ISPs have specialized, focusing on areas such as automated Web hosting, portals, and Web-based e-mail.
- What started as co-location services has now blossomed into such a huge segment — the Internet Data Center (IDC) market — that many ISPs are now hosted in IDC facilities.

In fact, the outsourcing movement that has benefitted so many businesses also works for ISPs, with functions such as world-wide access to modem pools, credit-card clearing, Digital Subscriber Line (DSL) connections, and even billing now outsourced to specialty providers.

Cost-Effective Growth

Regardless of their market niche, all types of service providers must position themselves to handle increasing numbers of subscribers, additional services, and more challenging workloads, all while keeping costs in check so that they can retain their competitive edge. ISPs must consider a wide range of costs, including total cost of ownership of their servers, the cost per square foot of data center space, and the number of customers per server. With servers and software designed to reduce total cost of ownership, there is no better partner for ISPs than Sun Microsystems.

With its single vision, "The Network is the Computer,™" Sun has been uniquely positioned at the forefront of server and software technologies to support service providers of all types. Sun has been committed to the Internet since 1982, when its first products included TCP/IP-based networking, and freely-distributed software was made available on FTP sites through the SunSITE program. Sun has developed such key technologies as the NFS network file system and the Java™ technologies that power so many Web applications around the world.

Sun commits enormous resources to support the needs of Internet service providers — and has designed several new lines of rack-mount servers and storage devices designed with the needs of service providers in mind. Sun's Netra™ servers and storage products are NEBS Level-3 certified to provide carrier-grade reliability for these rugged environments, with rack-mount servers supporting between one and four UltraSPARC™ CPUs each. For ISPs needing easy-to-deploy and configure Internet appliances, the Sun Cobalt™ line of 1U servers addresses areas including Web hosting, caching, and security with systems supporting one and two x86 architecture processors.

For environments not requiring NEBS compliance, the entry-level Sun Fire™ V100 server breaks the USD $1,000 price point, putting Sun's UltraSPARC technology within reach of more ISPs than ever before. Sun's the two- and four- processor Sun Fire 280R and V480 servers provide rack-mount capability with redundant power supplies and hot-swappable disk drives where more processing power per server is required. Scalability is built in to Sun's product

line, with Sun Fire™ server products extending up to the 106-processor Sun Fire 15K server capable of supporting the most high-end Web hosting and database applications.

Built from the ground up to be networked, the Solaris™ Operating Environment provides the cornerstone of Sun's line of ISP-ready software products, delivering the reliable, mature networking capabilities that built the Internet. To better support quality-of-service delivery in ISP environments, Solaris™ 9 Resource Manager software can be used to automatically allocate resources according to pre-defined priorities. For proven network security. Sun's SunScreen™ security products can run on any Sun server, and can be installed so that they are impervious to intrusion and undetectable on the network.

The Sun Netscape Alliance has produced the Sun™ ONE (formerly iPlanet™) product line that enables service providers to deliver world-class quality of service in a variety of environments. Sun ONE application and integration products support high-performance Web sites and integration with application server and e-commerce functions. Communication servers include Sun™ ONE Messaging Server, Sun™ ONE Wireless server, and Sun Internet Mail Server™ software, products that have for years been proven in demanding ISP environments. To help ISPs deploy unified LDAP-based user authentication and certificate management services, Sun ONE™ Directory Server is now integrated with the Solaris 9 Operating Environment.

ISP products and services involve more than just low-cost, high-performance servers, storage, and software — Sun understands that the way these solutions are deployed and integrated in ISP environments is just as important as the servers themselves. Sun and its integration affiliates know that each ISP installation is unique and requires the utmost care in design and deployment.

Well-Kept Secrets

Many of the configuration details that make a successful ISP installation are well-kept secrets — however there are some basic guidelines and principles that can be applied to many ISP architectures. This document presents some of the approaches that can be used to position ISP infrastructures for performance, scalability, reliability, manageability, security, and flexibility. The principles in this document extend well beyond the ISP environment, and

apply to environments that include corporate intranets, application service provisioning, and dot-com companies making large, Web-based applications available over the Internet.

This document cannot — indeed no document can — make the reader an expert in ISP configuration. Innovations in both servers and software happen so quickly that it is difficult to stay on top of current developments. ISPs differ greatly because they pursue different market niches — and because of the pace of technology — an architecture copied from a book or another ISP will be out of date before it is deployed. The stakes are high, and up-front investments in well-planned architectures will pay off in scalable, high-performance, secure, reliable, manageable, and flexible networks that can withstand the test of time.

The best way to deploy the most up-to-date architecture possible is by working with Sun's integration affiliates or with Sun Professional Services[SM] consultants from the beginning. Sun Professional Services has been a key facilitator for the highest-performance and largest-scale ISPs in the world.

Document Overview

This version of I*nternet Service Provider Configuration Guidelines* has been organized around a structure that includes sections on architectures, services, and deployment issues. The document begins with a discussion of the nature of ISPs and trends in the market. The architectures section contains chapters on architectural principles followed by detailed chapters on architectures for high availability and security for ISPs. The middle section focuses on the services that ISPs provide, including traditional mail, news, and Web services, as well as automated Web hosting. The final section covers deployment issues, discussing the role of Internet data centers in the ISP market, and some samples of local and regional ISP configurations.

The Nature of ISPs 2

The ISP Market

The ISP market continues to defy prediction, and the driving force behind the rapid change is phenomenal growth. This growth comes from three sources: growth in number of subscribers, an increase in the services they demand, and growth in the businesses that the ISPs support:

- *Subscriber growth*

  The number of ISP subscribers continues to grow at a rapid pace. In the U.S., a year 2000 Infonetics Research survey of the largest ISPs predicted that a dial-up port expansion of up to 400 percent, ISDN growth at 76 percent, and DSL growth at 42 percent would be needed to handle the increase in subscribers expected in the year 2001. Though growth in the U.S. is moderating, it is continuing at a rapid pace in Europe and Asia.

- *Demand for Services*

  Business organizations everywhere have recognized the benefits of outsourcing to alleviate the pressures of today's competitive marketplace. The need to do business on the Internet in spite of difficulty hiring, training, and retaining skilled IT staff leaves companies faced with conflicting priorities. By outsourcing jobs that require special skills (like security analysis) or common functions (like e-mail and Web hosting), and jobs which cannot be done better in-house (like establishing a global Internet presence), companies can focus to their core business and benefit from their ISP's economies of scale.

- *Customer Growth*

  As companies successfully negotiate outsourcing agreements, ISPs are faced with managing the growth of their customers. For example, an increase in numbers of e-mail accounts and the traffic that results, or an increase in visitors to the Web sites that the company outsources to the ISP.

The end result is that the biggest concern of ISPs today is managing their own growth, and hiring, training, and retaining qualified staff. The dramatic increase in the number of Internet subscribers and the traffic they generate has fueled the emergence of three distinct players: Internet Service Providers (ISPs), Network Service Providers (NSPs), and Application Service Providers (ASPs). Each of these providers develop their specific market while making inroads into complementary areas. Companies that provide services in all three areas are known as Full Service Providers (Figure 2-1):

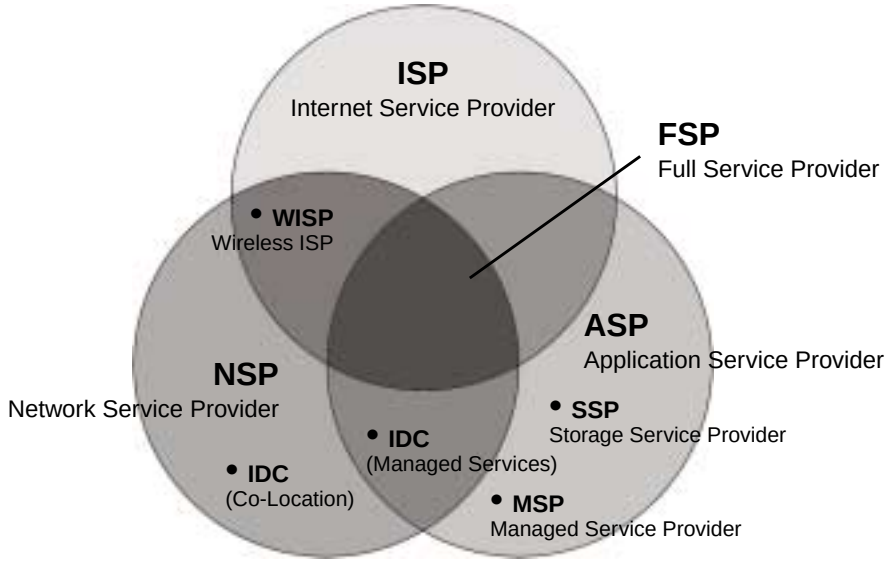

Figure 2-1 The ISP market can be divided into ISPs, ASPs, and NSPs, with some Full Service Providers addressing the needs of each segment. New niches are constantly appearing, including IDCs, MSPs, and SSPs

- *Internet Service Providers (ISPs)*

  Internet service providers supply millions of customers with on-ramp access through dial-up, high-speed copper, broadband, and wireless connections. ISPs deliver a wide range of services including e-mail, Web hosting, local content provision, security management, and virtual private networks. ISPs are under pressure to provide the most broad set of services possible, and many ISPs are alleviating their own pressures by outsourcing their modem banks and DSL access.

  Outside of the U.S., many ISPs have been regulated entities affiliated with government postal services and telephone companies, but the current climate of deregulation has fostered an international market of competing ISPs. Where once each foreign ISP had their own separate — and costly — high-speed connections to the United States, now these ISPs are free to aggregate service using international and global network service providers, resulting in lower costs and increased business.

- *Network Service Providers (NSPs)*

  The enormous increase in Internet traffic has stimulated significant growth in providers making up the Internet backbone. National providers like BBN and UUNET have reached out into international markets. There is a whole range of “new generation” providers like Qwest and Williams Communications, which use existing pipeline and railroad right-of-ways for the installation of extensive fiber optic networks. Meanwhile, regional NSPs like Merit continue to provide geographically-limited services. The growth in NSPs has resulted in a large increase in network Points Of Presence (POPs), where ISPs obtain their bandwidth, and network exchanges, where traffic crosses between NSPs.

  Internet data centers have begun to make inroads into both the ISP and ASP arenas by providing Web site hosting and co-location services where servers are located close to the backbone. This approach results in faster access, which leads to increased traffic, and leveraged infrastructure, which reduces costs to customers. IDCs offering pure co-location services might be classified as NSPs, while those offering managed hosting services would overlap with the ASP category. Some ISPs are located in IDCs, while other ISPs own them and provide sophisticated co-location and hosting services to their customers.

With the rapid deployment of low-cost wireless technologies in devices ranging from cell phones and PDAs to laptops, existing NSPs and some upstart companies are providing wireless Internet access to their customers. Established ISPs like AT&T are beginning to offer wireless access to a comprehensive set of ISP services through General Packet Radio Service (GPRS), while upstarts like Boingo are establishing a national network of 802.11b access points, offering network access only. Both of these technologies promise eventually to provide anytime, anywhere Internet access to customers.

- *Application Service Providers (ASPs)*

  Recognizing the capability of the Internet to deliver application services to businesses in a secure manner, a growing number of ISPs and application vendors are acting in the capacity of Application Service Providers. Examples are numerous:

  - Consumers are starting to use ASPs for their tax preparation.
  - Small businesses are accessing accounting software over the Internet.
  - Medium-sized companies compete with their larger counterparts by utilizing high-end applications like Oracle™ Financials and Enterprise Resource and Planning (ERP) on a subscription basis.

  Recognizing the need to manage hosted applications independent of the hosting facility, a new breed of ASP — the Managed Service Provider (MSP) — has arrived on the scene to manage hosted applications. These MSPs, like FreshWater Software, can monitor the performance and availability of Web-based applications all the way from the user perspective (can users purchase an item from the site?) to internal administration and management services (is there sufficient database table space?).

  As the Internet data center market matures, many customers look for more from their hosting facility than just bandwidth, power, and air conditioning. Storage Service Providers (SSPs) offer a range of storage-on-demand services including scalable disk storage, backups, mirroring, and remote replication.

Capitalizing on the tremendous opportunities for growth and innovation, there is an increasing amount of crossover and overlap between the major categories. Regardless of market niche, all types of providers must position themselves for growth and agility to handle increasing numbers of subscribers, additional services, and more challenging workloads. System architectures that meet these demands are critical to success.

ISP Services

The connectivity that customers expect of Internet service providers has changed dramatically in the last decade. When fixed, high-speed connections to the Internet were the rule, only businesses, educational institutions, and government agencies participated in the world-wide interconnection of networks. As Serial Line Internet Protocol (SLIP) and Point-to-Point Protocol (PPP) enabled Internet connections over dial-up lines, the role of the ISP as the on-ramp access provider was born.

Today, the demand for high-speed connectivity using technologies including Integrated Services Data Network (ISDN), Digital Subscriber Lines (DSL), broadband cable, and mobile as well as fixed wireless has increased both the Internet's reach and customer demand for more data-intensive services. Whereas once ISPs had only to provide high-speed connectivity to a small number of highly-technical customers, they now have to provide network-based applications, content, and value-added network services to attract and keep the new wave of high-speed business and residential users.

Basic ISP Functionality

Internet service providers must now deliver attractive packages that include the ability to store and retrieve mail, to access huge numbers of news articles, to view and host Web sites; to view streaming video clips, to interact in chat rooms, and to communicate using voice over IP. In order to provide these value-added services, companies entering the ISP marketplace must ensure that their infrastructure supports the full range of basic functionality:

- *ISP Applications*

  The base of applications that ISPs are expected to provide to their users includes electronic mail, netnews, World Wide Web access, and Web page hosting. More sophisticated users may also demand access to other Internet services such as chat, IP telephony, and bulletin-board systems. Business customers may demand security services and Virtual Private Network (VPN) support. Decisions must be made regarding which services will be provided to ISP customers — and in large part these choices are what differentiate one ISP from another.

- *System Integrity*

  ISPs must take the utmost care to ensure both the availability and integrity of customer data. Therefore the network through which an ISP provides services must be protected against breaches of security and against hardware and software failures. Firewalls are a necessity for protecting an ISP's core services from intrusion from both the Internet and its customers. High Availability (HA) for ISP applications can be achieved by deploying horizontally-scaled servers in configurations supported by Sun™ Cluster software, which provide integrated configurations for both NFS network file systems and many Internet server applications. System integrity must also be ensured through frequent backups.

- *Full Range of Connectivity*

  Depending on their market, ISPs carefully consider the connectivity that they provide to their subscribers, and must maintain agility to adopt new technologies as they become available:

  - At minimum, today's ISPs provide dial-up service using the Public Switched Telephone Network (PSTN).
  - Broadband ISPs provide access through cable modems and DSL lines, satisfying the needs of most residential and small office/home office customers.
  - High-bandwidth connections using leased lines, fixed wireless, ATM, and Frame Relay are needed to support commercial subscribers.

  Points of presence extend an ISP's home service area, reaching a broader customer base, and allowing customers to access ISP services while travelling. The largest ISPs average more than 150 POPs each. Some ISPs establish roaming arrangements with other service providers, allowing national and world-wide access to a uniform set of services. Others outsource their access networks entirely, contracting with other providers for modem pool access, and for high-speed access including DSL.

- *System and User Management*

  Service Providers must have well thought-out strategies for managing their own network and the potentially explosive growth that results from providing a high level of service to customers:

  - Basic facilities like Domain Name Service (DNS) must be established to provide addressibilty to ISP services and customer-owned domain names.

- Lightweight Directory Access Protocol (LDAP) is used to coordinate subscriber information across a variety of services and operating system platforms.
- Network management tools help to manage routers, firewalls, and the core servers themselves.

Effectively managing users and creating new accounts becomes increasingly critical as an ISP grows. Sun's affiliates observe that user administration costs tend to peak at around 20,000 subscribers, after which it is essential to establish automatic subscription mechanisms and credit-card billing.

Local, Regional, and National ISPs

Internet service providers tend to be clustered in three general size categories (Table 2-1). Local ISPs begin with configurations for 5,000 to 10,000 subscribers and scale upwards from this range. Regional ISPs begin in the range of 50,000 subscribers. National ISPs have often have more than one million subscribers. Each category of provider has its own characteristics and challenges, many of which are discussed in this document.

ISP Categories	Number of Subscribers
Local	5,000 to 10,000
Regional	50,000 and up
National	1,000,000 and up

Table 2-1 ISPs categorized by size

There are many ways to categorize ISPs. In the definition of local, regional, and national ISPs used in this paper, it is important not to tie any hard and fast number of users to the architectures that are illustrated. This is because all ISPs are different — the infrastructure needed to support a pure on-ramp dial-up ISP is a world apart from the infrastructure in a cable ISP, where customers tend to stay online longer and make much more intensive use of available bandwidth.

Architectural Principles 3

The architectures deployed by Internet service providers generally consist of a set of sub-networks having different functional responsibilities — for example user access, services, and administration. These subnets are connected via routers and/or firewalls that are used to control access from one subnet to another. There are many ways in which ISP services can be deployed onto a physical network of machines. Factors to be considered include the number of subscribers, expected workloads, services to be provided, desired performance level, availability requirements, expectations for growth, and security concerns.

The result is that no two architectures are the same. Further, designs that are appropriate at one point in time are often superseded by those that exploit new developments in hardware and software. There is, however, a set of common underlying principles used in the design of ISP architectures, which is the topic of this chapter.

Architectures to Support the -ilities

All ISPs wish to provide a set of services that meet the functional requirements of their customers, for example dial-up access, e-mail, Web hosting, and news services. ISP architectures are designed to deliver these services to customers, but a major factor in their design is meeting what are technically called the *non-functional requirements*, including scalability, reliability, availability, serviceability, security, flexibility, and manageability. How well an ISP delivers on these non-functional requirements — or the *-ilities* — determines everything from customer satisfaction to the cost of providing services.

If there is any magic to ISP architectural design, it lies in successfully balancing the conflicting demands of the non-functional requirements. For example, the most easy-to-manage ISP infrastructure would be to host all services on a single, very powerful server. But requirements for scalability, availability, and security dictate the deployment of services across several systems. Knowing where to strike the balance is a fine art, and one reason that so many ISPs choose Sun products and services is because the Sun platform provides so much freedom when designing ISP architectures.

The importance of each of the -ilities is discussed below, and the rest of this chapter discusses architectural approaches that support them.

Scalability

Scalability is without a doubt the most important issue facing service providers today. Dramatic growth is the norm, and ISPs must be prepared for their subscriber base to grow by as much as an order of magnitude in just one year. Beginning with a base of 10,000 subscribers, the requirement for an ISP to quickly adapt to handling 100,000 subscribers will put early design decisions to the test. Those ISPs with architectures supporting scalability can grow with the speed of their business. Those which don't can be paralyzed with costs that rise more quickly than their customer base, service outages, and the failure of their administration staff to support their infrastructure. As this chapter discusses in detail, the horizontal and vertical scalability of Sun's server product line provides a competitive edge for service providers.

Availability

The Internet has no 'off' switch, and providing Internet services is a 24 hours a day, seven days per week, 365 days per year proposition. Customers expect their ISP services to be available around the clock, without interruption. Web site visitors access ISP-hosted Web sites from every time zone around the world, and therefore around the clock. Companies outsourcing their e-commerce sites to service providers stand to lose business in the event of a service outage. Zona Research (1999) estimated that e-commerce sites stand to lose more than USD $4.35 billion per year due to long download times and Web site failures. This tangible cost has led companies to choose their ISPs based on the service-level agreements they provide.

Sun's SunTone[SM] certification and branding program provides a symbol of quality for outsourced services and applications. Both service providers and their customers can have confidence in the quality of the products and services they buy when they carry the SunTone brand. Many service providers including ISPs and IDCs have gone through the rigorous certification process in order to advertise to their customers that a well-respected third party like Sun has found that their service levels exemplify best practices for operations and infrastructure.

Reliability

A necessary — but not sufficient — requirement for providing highly-available services is having reliable servers, software, and network components. Reliability is usually specified as the Mean Time Between Failures (MTBF), and service providers understand that MTBF is affected by hardware, software, and human factors. An IEEE/Oracle research report shows that only 25 percent of down time is due to physical factors including disk failure, power surges, and CPU failure. The same study shows that 34 percent can be traced to human factors, and 36 percent to software bugs and design issues. Given that it is impossible to build servers and software that are 100 percent reliable, ISPs must rely on their network architectures to be resilient and provide continuous service ability in the event of a failure.

Sun provides high reliability through its Netra and Sun Fire server products with features like ECC memory, redundant power supplies, hot-swappable components, and dynamic reconfiguration. Sun's hardware reliability is complemented by the stable, mature Solaris Operating Environment, which has incorporated TCP/IP-based networks since 1982. Features like alternate pathing enables network and I/O operations to continue even in the event that one of two redundant connections fails. Dynamic reconfiguration, on supported servers, enables failed parts to be taken offline for replacement without power cycling or rebooting.

Serviceability

Given the inevitability of server, software, and human failures, ISPs must build network infrastructures and deploy servers that can be brought back on line as quickly as possible. Serviceability is expressed as the Mean Time To Repair (MTTR), and is facilitated both by network designs that are resilient to single

server failures, but also by servers that can be dynamically reconfigured for continuous operation in the event of a component failure, and by servers which, once down, can be repaired with easily-accessible replacement parts.

Serviceability is influenced by network architecture, server design, and ISP choice of vendor support services. Network architectures can be designed to support serviceability by being resilient to failure of a single server, and by providing fail-over mechanisms that enable servers to be taken offline for maintenance and upgrades. Server designs for serviceability include systems designed to accept field-replaceable units (FRUs), servers with easily-accessible components that can quickly be replaced and the server rebooted, and dynamically-reconfigurable systems that even allow replacement of CPUs and memory without interruption in service. With Sun servers, ISPs can choose between these 'replace,' 'reboot,' and 'reconfigure' serviceability models, and can select support from Sun Services that most appropriately meets their needs.

Security

Security is one of the most difficult jobs that an ISP must face, and is a task that requires constant work and vigilance. At the same time that an ISP must protect its services from attack over the Internet, it must also protect itself from attack by its customers. Services must be open to customers yet closed to the outside, leading to paradoxical security requirements that are difficult to meet. Yet the pressure to maintain a secure infrastructure is greater than ever, with security flaws making national news on a daily basis. Denial of service attacks, Web site defacements, and unauthorized access to customer data — including credit card numbers — are all threats to ISP operations, their reputation, and future business.

ISPs protect themselves with network architectures that resist attack and that can respond quickly to any unforeseen intrusions. They also utilize the most sophisticated firewall and network intrusion detection mechanisms available, including SunScreen security technologies from Sun.

Flexibility

Flexibility might seem like the most intangible of the non-functional requirements, however it has potentially the most tangible impact on an ISP's bottom line. ISP architectures that are designed to be flexible are the ones that

can quickly respond to spikes in demand, for example, when a hosted Web site is mentioned on the major news networks. Architectures that include networks, servers, and storage that are based on open standards result in components that are interchangeable and can be easily re-deployed to serve new functions — saving capital costs and protecting investments.

ISPs that keep flexibility in mind can move quickly to adapt to rapidly-changing market conditions, potentially edging out competition. For example, decisions as minute as the relationship between a Web server and application server can have a huge impact on the bottom line: where the relationship is fixed, scalability is limited; where the relationship is flexible, each layer in the infrastructure can be scaled independently, meeting customer demands more quickly and at lower cost.

Sun believes that open standards promote the most flexible ISP architectures possible, which is why it promotes open standards including TCP/IP, the NFS network file system, freely-available Java technologies, and the Jiro™ open storage management platform. A key component of ISP flexibility is the ability to scale services in multiple dimensions, discussed further in the section on scalability.

Manageability

ISPs, like their business customers, are limited in their ability to hire, train, and retain the skilled administration staff they need to maintain competitiveness. ISP architectures designed for manageability are poised for success because they can grow quickly without outstripping the capacity of their staff. With the right synergy between people, process, and technology, ISPs can gain a high degree of leverage with their existing staff. The technology to support manageability includes network management tools and separate networks to support management functions. When designing architectures with manageability in mind, there is a constant tension between the need for more servers to provide high availability, and fewer servers to reduce management overhead. These trade-offs are considered in the remainder of this chapter.

Scalability

Managing growth is the number one issue for service providers of all types. Scalability is the factor that affects whether service providers can grow in a managed, cost-effective manner, or whether they will encounter problems with

functionality and performance as they grow. Sun views scalability with a multi-dimensional model that includes vertical scalability, horizontal scalability, and Z-Axis scalability — all are necessary to support scalable ISP infrastructures (Figure 3-1).

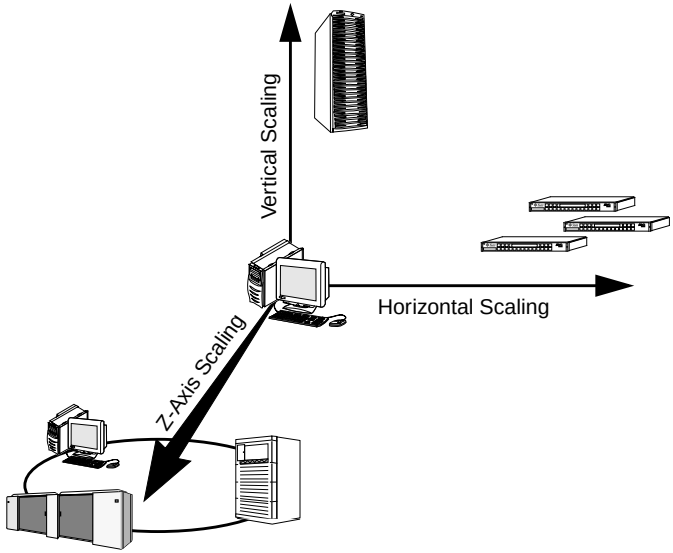

Figure 3-1 Scalability to meet the needs of service providers and their customers is best considered in three dimensions. Illustrated data points include horizontally-scaled Web servers and vertically-scaled database servers.

Definitions

Scalability is the property that adding more processing resources — including servers, memory, CPUs, disks, or even network segments — results in an increased capacity to provide services.

Linear scalability is the desirable property that the cost of capacity increments is uniform and predictable. Simply put, adding another dollar's worth of resources into a system results in the same increment of capacity out, whether

it's the first dollar in, or the hundredth. All systems have capacity limitations, and ISPs must pay attention to whether the limitations they face are due to the hardware or the software architecture of their services.

Vertical Scalability

Vertical scalability refers to adding capacity by increasing the processing power of a single server. Vertical scalability means using more and faster processors, high performance interconnects, and more efficient, highly-optimized system software. Vertical scalability can be accomplished by adding resources to a server or, at the point where the server has reached its maximum capacity, upgrading to one capable of accommodating more processors, CPUs, memory, disk, and network resources.

Sun has long been a leader in offering servers with vertical scalability, with ever faster processors, continued innovations in system design, and an ongoing commitment to symmetric multiprocessing (SMP) capabilities. The combination of vertical scalability in Sun servers and binary compatibility across the product line means that the same application can run without change whether it's in a single-processor Sun Fire V100 server or a 106-processor Sun Fire 15K server. This leaves ISPs with greater flexibility in how they deploy their services, and the freedom to grow without changing their applications or operating environment.

Vertical scalability requires both scalable servers and scalable software. Sun has produced steady improvements in the Solaris Operating Environment to unleash the power of its SMP servers. For example, its TCP/IP stack is now multi-threaded, meaning that multiple processors can work in parallel to speed the flow of information from the server to the network. The operating environment itself is implemented using threads as its low-level scheduling unit, enabling higher levels of concurrency in the kernel, and higher performance for applications.

Network applications must be capable of utilizing multiple threads in order for vertical scalability to be effective. Historically, ISP applications like netnews and standard sendmail were not sufficiently multi-threaded and did not scale well with the addition of multiple CPUs. Today, most commercial ISP applications are multi-threaded and can scale to varying degrees. Most Web servers, for example, scale well to around 24 processors, with performance improvements diminishing beyond that point.

Horizontal Scalability

Horizontal scalability is achieved by adding more servers to a particular service, such as mail, Web, or news services. Since scaling an ISP installation inevitably means adding more machines, an architecture that provides for horizontal growth is superior to one that depends on vertical scalability alone. In fact, horizontal scaling is typically the first approach to growing an ISP installation because the addition of multiple servers delivers a dramatic improvement in availability.

The availability benefits of horizontal scaling diminish quickly with the number of servers, and the larger the number of servers, the greater the cost of maintaining the separate hardware, operating system, and application software instances. Horizontal scaling, therefore, must be used in conjunction with vertical scaling. When deploying large Web server farms or Web-based applications, one rule of thumb is to scale the number of front-end Web servers to five, and scale each server vertically so that it can process 25 percent of the workload. With this formula, a single server can be taken offline for maintenance or upgrades with the remaining four servers capable of processing 100 percent of the workload.

When deploying horizontally-scaled servers, ISPs must pay careful attention to whether the application can be horizontally scaled. Web servers, for example, can scale horizontally only if each user's session state can be pushed down to the user's browser in the form of a cookie, and up to an application server that is accessible by all Web servers. Some applications, such as database management systems, store an immense amount of internal state, and do not scale well horizontally except in very narrowly-prescribed configurations, such as parallel database servers. In general, vertical scaling provides the greatest benefit for these complex services, and also provides the best leverage for applications that often have high per-server licensing costs.

When horizontal scaling is used to increase capacity and availability, a load-balancing mechanism must be used to spread the workload across the set of servers. There is a broad range of both hardware and software products available today, and a wide selection of load-balancing algorithms. Some algorithms, such as round-robin, are static, and others adapt to quickly-changing workloads. Performance-based load balancing directs requests to the server that responds the most quickly, with the goal of evening out the response times seen by users. Resource-based approaches have visibility into server workloads, and can level out the resource consumption levels (like CPU

utilization) across the array of servers. In cases where session state is not maintained elsewhere, load-balancing mechanisms can be configured so that they are 'sticky' and fix one user's session to a specific server. This approach can make software implementations more straightforward, but have less capacity to recover from errors — the failure of a 'sticky' server can result in the loss of all transactions in process on that server.

Z-Axis Scalability

The demands on service providers has resulted in a new set of requirements that Sun is pursuing as a third dimension of scalability. *Z-axis scalability* refers to the additional characteristics needed to implement service provider networks and leverage the benefits of horizontal and vertical scaling — including greater bandwidth, reduced latencies, and better data coherency, along with features needed to implement and manage a dynamic, flexible environment. Examples include load balancing, rapid resource reconfiguration, and better investment in system architectures and applications. Some of the examples in this chapter illustrate techniques of Z-axis scalability for rapid response to spikes in Web traffic and redeployment of existing systems for rapid response to quickly-changing markets.

Managing Growth

Horizontal, vertical, and Z-axis scalability are good tools for scaling service provider infrastructures, but managing growth is a much broader problem. In order to stay on top of the growth curve, an iterative process should be implemented to constantly model, design, deploy, monitor, and update architectures before bottlenecks are reached:

- Build a model of the expected user demand.
- Design architectures to accommodate the maximum user demand expected.
- Deploy the architecture to a reasonable time horizon, accounting for the fact that scaling hardware, software, and network bandwidth requires lead time from suppliers.
- Monitor usage of the deployed architecture, and build an intuitive model based on actual measurements, not just educated guesses.
- Update the architecture to accommodate the real workload, and repeat the process by modeling the next major increment of growth.

Functional Decomposition

The principles of scalability provide necessary, but insufficient tools for designing ISP architectures. *Functional decomposition* offers a set of techniques for deploying ISP services across multiple servers. Breaking up an ISP service into components — mail, for example — enables network architects to design for availability, reliability, serviceability, security, flexibility, and manageability. In addition, with services decomposed into separate components, different scalability techniques can be used as appropriate for different parts. For example, if the user access function of a mail service is separated from the mailbox storage function, horizontal scaling could be used for the first component and vertical scaling could be used for the latter.

The principles of functional decomposition provide some guidance as to the merits of deploying an entire ISP infrastructure using a single, large SMP server. One could argue that, given the power and scalability of Sun's Sun Fire server products, an entire medium-sized ISP could be based on a single SMP server. Indeed, some of Sun's largest ISP customers use several Sun Enterprise™ 10000 servers in their network infrastructure. But in these ISPs, Sun's largest servers are used as part of a plan that includes horizontal and vertical scaling, along with attention to the rest of the -ilities. Successful ISPs are always pushing the scalability limits of their software, networks, and their servers, and given that both horizontal and vertical scaling eventually will be needed, it is best to start with a network infrastructure that supports multi-dimensional scalability through the use of functional decomposition.

The benefits of functional decomposition include more scalable infrastructures where separate components can be scaled differently. More high performance services can be provided by tuning each server or functional layer for its designated tasks. Security is enhanced because a successful intrusion attempt can more easily be isolated to a specific server or service without affecting the ISP's overall availability. The use of multi-dimensional scalability increases availability because servers can be taken offline for maintenance or upgrades with the remaining servers taking up the workload.

Decomposition without limits can put a good network architecture out of balance. The more servers that are deployed, the more administration time that is required for maintaining and upgrading both hardware and software components. The improved manageability achieved by balancing the multiple

dimensions of scalability results in lower operating costs — and is one of the reasons why so many ISPs leverage the multi-dimensional scalability supported by Sun servers.

There are several ways to implement functional decomposition — by service, task layer, and by special function. All of these techniques can be used separately or they can be superposed to form more complex — and more reliable, available, scalable, and secure ISP infrastructures. The examples in this section illustrate the basic principles of functional decomposition — in real use they would be combined with other techniques such as isolating multiple subnets for security.

Partitioning by Service

An obvious way to functionally decompose an ISP's services is to partition each service onto a separate server (Figure 3-2). This example illustrates independent mail, news, and Web servers, with customer access from the Public Switched Telephone Network (PSTN) and Internet access through an access router.

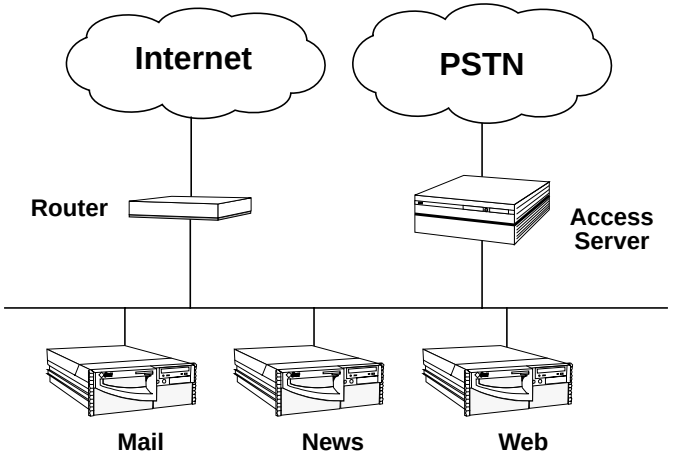

Figure 3-2 Partitioning by service

Each server can be configured and tuned for the service it is to host, yielding the best possible performance. Because each service is independent, management of each service is straightforward and the ISP can flexibly reconfigure each server — for example increasing main memory — while affecting only one service at a time.

Partitioning by Task Layer

Another way to decompose ISP services is by task layer (Figure 3-3). Each of the three basic services — mail, news and Web — have functions that can be layered into a three-tier architecture.

An access server provides the interfaces with which clients interact. A mail access server supports POP and IMAP access to user mailboxes. For a news service, this is where an NNTP server providing user access to news resides. For Web services, the Web server would reside on a front-end access server.

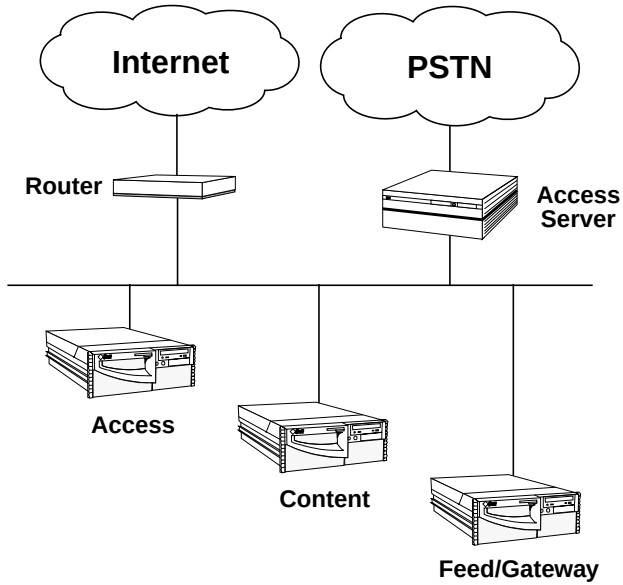

Figure 3-3 Partitioning by task layer

A content server holds the data accessed by each service. A mail server may require a database server to store user mailboxes. A Web service may use a server specifically tuned for CGI generation of dynamic Web content. An NFS server might be used to store news articles or static Web pages.

The feed/gateway server supports the interface between the ISP and the Internet. It accepts incoming mail via SMTP and delivers mail to user mailboxes on the storage server. A news gateway server handles a news feed from an upstream site and places articles into a spool area.

This example is fairly academic; a more realistic implementation (Figure 3-4) would partition by service and by task layer. This architecture has several advantages. Performance tuning can be done at the layer requiring optimization. Security is improved because penetration of one host does not necessarily yield access to the entire ISP network. This architecture is both vertically and horizontally scalable, and offers the ISP increased flexibility in configuring and managing its network.

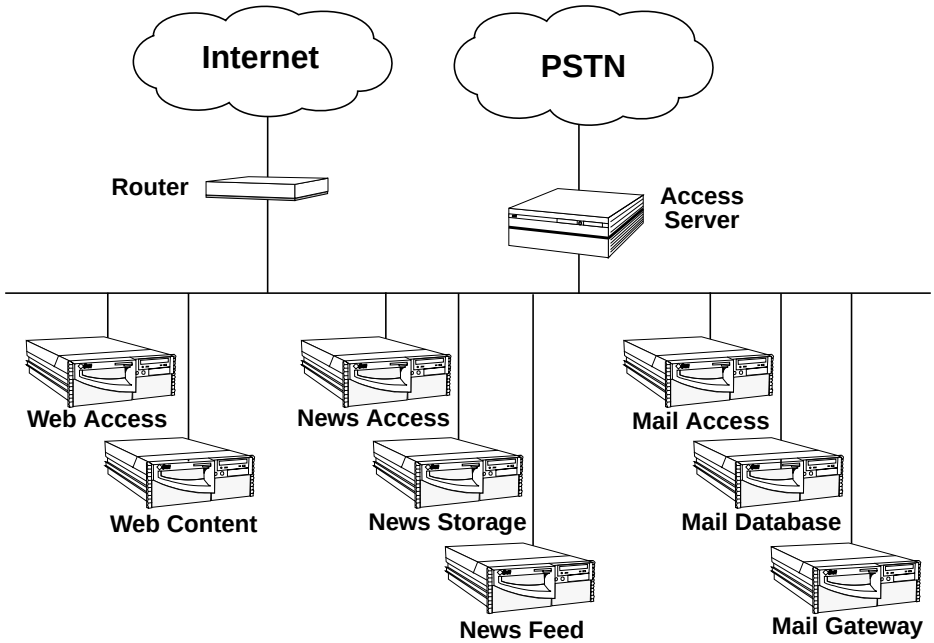

Figure 3-4 Partitioning by service and task layer

Partitioning by Special Function

A third way to partition ISP services is by special functions. One of the more common approaches is to configure an LDAP or RADIUS authentication server that is used by all ISP services — including dial-up modem banks and e-mail servers — to authenticate users. Having multiple DNS servers hosted in separate subnets is also an example of partitioning by special function.

ISPs supporting e-commerce Web sites may have a single server dedicated to handling the credit card billing function. And Web sites with the potential of very large spikes in their workload may divert requests above some threshold to an overflow server whose entire purpose is to present a page informing the user of the overload situation and asking them to come back later.

Reliability and Availability

Reliability and availability can be improved by incorporating redundancy into the ISP architecture so that increased reliability is achieved through resilient architectures rather than trying to make each ISP component impervious to hardware or software errors. Probability theory provides some direction in how to accomplish this (Figure 3-5):

- *Reliability of series components*

  Recall that the probability of failure for a system depending on series components is the sum of each components' probability. More components in series increases the probability of failure. For example, a single mail server with probably of failure m, requiring the functions of a single authentication server with probability of failure a, will have a probability of failure p calculated as $p = m + a$.

- *Reliability of parallel components*

  The probability of failure for a system depending on the operation of only one of multiple parallel components is the product of each component's failure probability. Thus the failure probability p of two horizontally-scaled Web servers each with probability of failure w, is $p = w^2$.

With the probabilities of failure being expressed as p, the reliability, or the probability that the service won't fail, is expressed as $1-p$. So if all of the components above have a 10 percent probability of failure, the system composed of parallel components has reliability of 80 percent, while the

parallel system has reliability of 99 percent. Given that 90 percent server availability is a low estimate, very highly reliable services can be deployed with only a few parallel components.

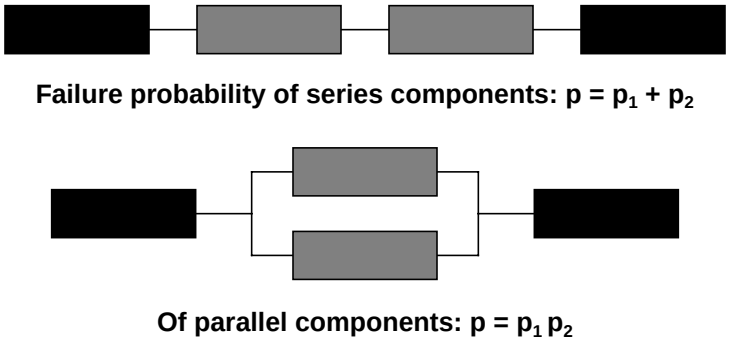

Figure 3-5 Reliability increases exponentially through the use of parallel components resulting in a cost-effective and highly manageable architectural model for ISPs

This brief foray into probability theory yields an important lesson for ISPs: adding redundancy into a network architecture through the use of parallel components has a dramatic effect on reliability. Given that reliability of parallel systems increases exponentially with the addition of more components, it is also clear that the point of diminishing returns is quickly reached. This is good news for ISPs, because the result is that large gains can easily be accomplished with modest investments, and manageability can be kept in check because the underlying probability theory does not require a large number of servers for a high degree of reliability.

Increasing Redundancy

Redundancy can be added into ISP architectures in several ways (Figure 3-6). Multiple Internet connections can be established to different upstream providers to ensure that, if one link fails, the other link can carry traffic (with reduced performance). Multiple connections also may be partitioned by function. For example, a news feed could be supported on a dedicated connection so that the fluctuations in news traffic do not affect performance of other services. Of course the ISP network requires multiple routers in order to

accrue the reliability benefits of multiple Internet connections; likewise redundant access servers enable at least some users to access the ISP network in the event of one access server failure.

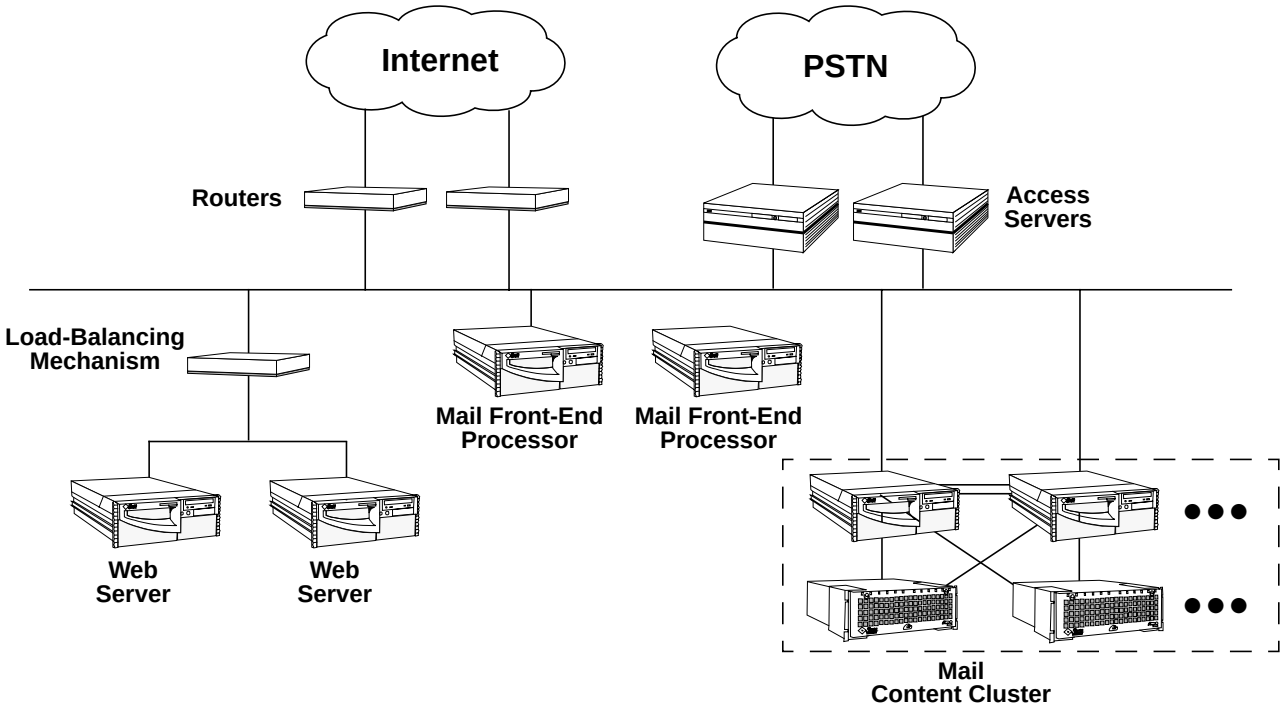

Figure 3-6 Redundancy techniques that improve reliability and availability

Increasing service availability through the use of multiple servers always requires some method for load balancing or fail-over so that the loss of one server does not cause the service itself to become unavailable. Figure 3-6 illustrates three techniques for clustering servers for high availability, illustrated from left to right:

- *Hardware load balancing*

  In this illustration, two Web servers having identical content are deployed on a subnet controlled by a hardware load-balancing mechanism. This device allocates incoming requests to the two servers depending on an

algorithm determined at configuration time. The load balancing mechanism monitors the workload that it allocates and directs all traffic to the remaining servers in the event of a single server failure. The configuration illustrated contains a single point of failure — the load-balancing mechanism itself — and in real life would need to be deployed in pairs for maximum benefit.

- *Software load balancing*

  The mail front-end servers use software load balancing, which through the use of third-party software, enables one server to act as a master which directs incoming requests to itself or the other servers in the configuration. In the event of a failure, the master server stops directing requests to the failed server. Software load balancing, although it requires a small amount of CPU time on the servers themselves, is architecturally more straightforward to deploy.

  In the past, some ISPs used a fixed allocation of users to mail servers in order to statically balance the workload. This approach has given way to an architecture using front-end processors and a back-end message store because of the complexity in partitioning users and their mailboxes as the workload shifts.

- *Virtualization*

  Clusters, such as the one illustrated as the mail content cluster, use IP address virtualization to support one service with multiple servers. Multiple servers in the cluster appear to clients as a single IP address, with load balancing and fail-over managed through the cluster interconnect shown connecting the two servers. One configuration providing high-availability NFS services using Sun Cluster software provides high availability with a primary and backup server. In the event of a failure, the backup server assumes the IP address of the failed server and continues operation without interruption.

Sun has found that the most effective solution for reliability and availability is to configure multiple servers to support each service. These configurations will be discussed in forthcoming chapters.

Architectures for Security

Architectural support for security is usually provided by creating multiple sub-networks that are isolated by firewalls. A typical ISP configuration (Figure 3-7) consists of an access network, demilitarized zone (DMZ), and services network. One of the advantages of partitioning services by task layer is that firewalls can be interposed between layers — in this example a firewall controls access between the front-end processors and the content storage network.

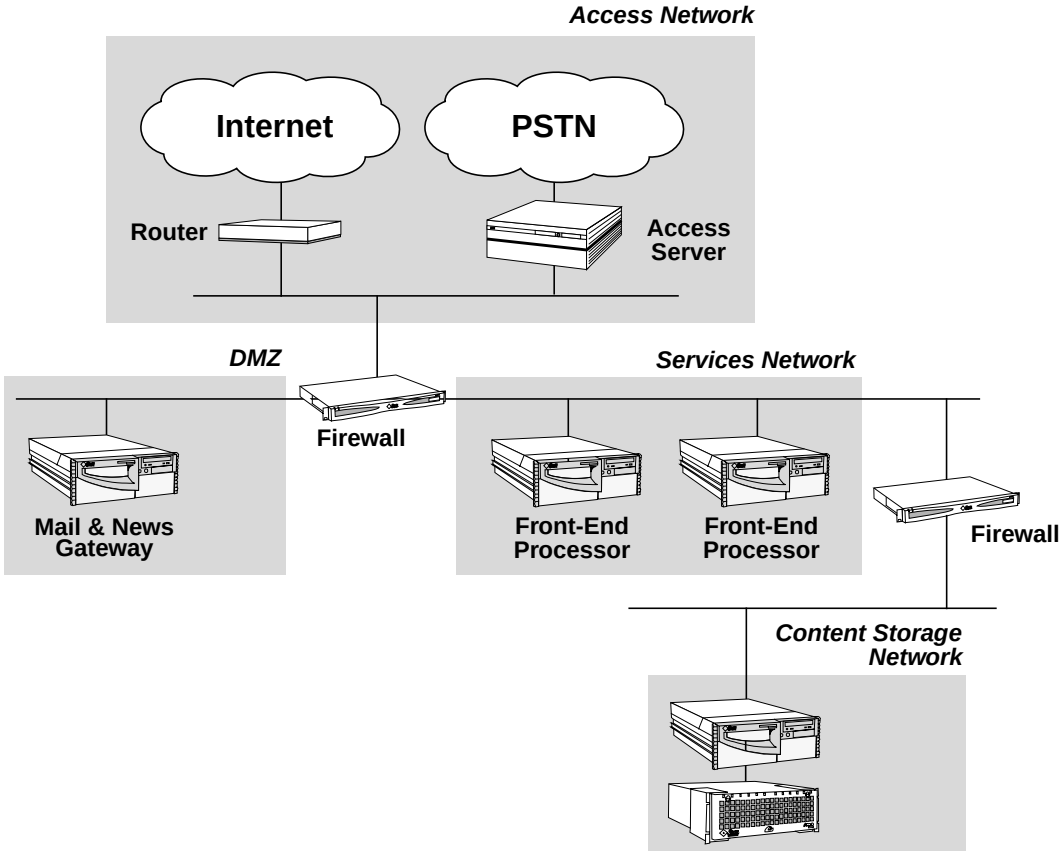

Figure 3-7 Architectures for security

The access network provides the connection by which services may be used by both Internet and dial-up users. The packet-filtering router allows Internet traffic to proceed only to the access servers and the DMZ.

The DMZ creates a moat between the Internet and the services network. This example shows a mail and news gateway in the DMZ. All SMTP and NNTP traffic from the Internet is examined by the firewall and routed only to the gateway server. Once the mail and/or news is deposited on the gateway server, it proceeds to transfer it to the storage server.

The DMZ is appropriately named because the indirect transfer of data from the Internet to the storage network requires any intruder to "drop their weapons" in the DMZ — making direct penetration into the services or content storage network more difficult. A fully-configured DMZ might contain separate news feed servers, mail gateways, proxy caching servers, DNS, and authentication servers.

The services network provides access to mail, news, and Web services. The combination of router and firewall could be configured to allow only subscribers connected via the access server to access the mail and news servers. Internet access is usually provided to Web servers, and in this example the firewall could allow only HTTP service to a specific front-end processor. This limitation of protocols to specific hosts makes intrusion via multiple protocols difficult to accomplish.

Because the content storage network must satisfy requests from both the DMZ and the services network, a separate firewall is configured so that performance is not limited. This firewall, interposed between the access and the storage aspects of each service, limits traffic to NFS requests from the authorized front-end processors.

Architectures for Manageability

Managing ISP networks requires an intricate balance of people, process, and technology — and an entire book could be written on the subject. ISPs need to adopt the same kinds of procedures that large IT organizations use, including run books that describe the appropriate system administrator response to specific error situations, and the use of standard, well-tested software configurations and patch levels to minimize inconsistency — and hence the opportunity for error — across the servers that are maintained.

This discussion focuses on architectural techniques for increasing manageability, including the use of separate physical networks for different functions, and the use of a separate console network for access to each server console.

Separate Ethernet Networks

A key architectural approach for increasing manageability is to separate different types of traffic onto different networks. For example, network management systems can utilize a dedicated network to access the various devices that they monitor. Separate networks are frequently used to support backup traffic, and access to back-end storage and databases is often provided over physically isolated network segments. The ability to configure a server onto multiple network segments is so useful that some Internet data centers configure each of their Sun servers with a Sun Quad FastEthernet™ card so that they can easily be patched into the required networks.

The use of multiple networks for administration has a significant benefit to ISP security. When administration traffic is allowed to co-exist on the same physical network as user traffic, intrusions can be accomplished by snooping administration traffic — including passwords — and then masquerading as an administrator using the compromised information. When root logins and administrative access are allowed only over a separate network, the ability to obtain and use administrative privileges is significantly curtailed.

The use of separate networks must be carefully controlled so that they cannot be used by intruders to gain access from one server to another. This control can be achieved through the use of Virtual Local-Area Networks (VLANs), limited network segments, and routers and firewalls that restrict the flow of traffic to only the specific protocols that may legitimately be used on the segment.

Separate Console Network

For increased security, ISPs configure their servers to allow root logins only on each server system console. An effective way to enable console access from a central location is to utilize the serial port of each server in the ISP network. Each serial port is wired to a terminal concentrator (Figure 3-8) which is configured onto a private network. From any host on the private network, a telnet session can be established to the appropriate port on the terminal concentrator to obtain access to one of the server's serial ports.

Sun has found this to be a useful approach in many installations where access to the consoles of a large number of servers must be provided — including Sun's own engineering design automation laboratories. Using this technique in an ISP network involves allowing root access — and hence administrative access — only to logins on the serial port, eliminating one set of intrusion paths from the Internet.

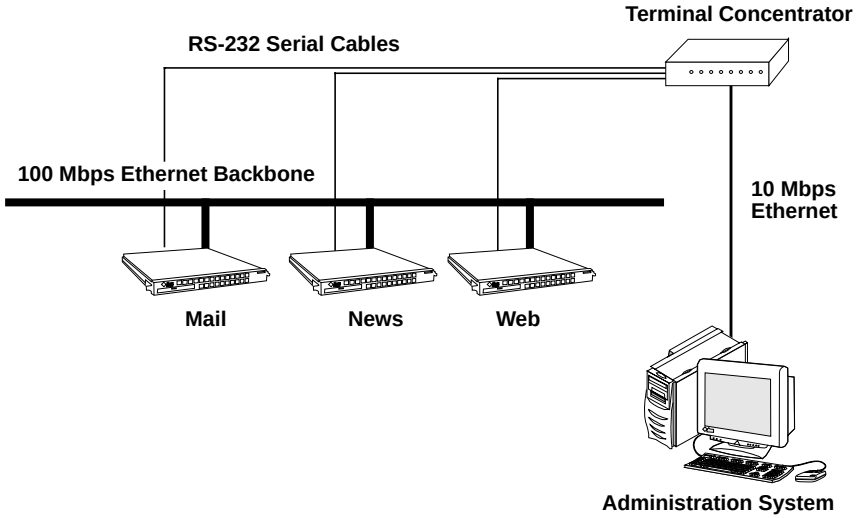

Figure 3-8 Console network using terminal concentrator

Summary

Every ISP architecture is unique, and is determined by multiple variables, including the number of subscribers, the services to be offered, the workload expected, and the differing goals of supporting scalability, availability, reliability, serviceability, security, flexibility, and manageability. Scalability is the most important aspect of an ISP configuration, and it can be achieved by architectures that support both vertical and horizontal scaling. Even with architectures designed to utilize both vertical and horizontal scaling, network configurations will ultimately require change in order to accommodate ISP growth.

The benefits of functional decomposition — by service, task layer, and special function — accrue by deploying each ISP service using multiple servers. Once a service is decomposed into parts, other techniques including those for reliability and availability, security, and manageability can be applied. In actual use, many of these approaches are superposed, resulting in ISP architectures that exploit the best of all of these techniques.

High-Availability Solutions For ISPs

4

Today's Internet users have high expectations about the availability of the services they purchase. Service interruptions and lagging response times make national news, and so there is intense competitive pressure among ISPs to provide high-availability services and to stay out of the limelight. Indeed, as ISPs offer increasingly sophisticated services, the availability of customer data — from Web sites to electronic mail — becomes even more important. Providing highly-available services is important to all ISPs, but it is especially so to telephone company ISPs, where customers expect Internet services to be as reliable as a dial tone.

Fault Tolerance vs. High Availability

There is an important distinction between fault tolerance and high availability, and where each approach should be used:

- *Fault-tolerant* computers are specifically designed to provide uninterrupted service even after catastrophic system or environmental failures that would completely shut down other configurations. Fault-tolerant systems typically employ specialized hardware with processors running in lock-step execution. Applications requiring fault tolerance, such as telephone switching and air traffic control, cannot sustain *any* interruption in service. Sun offers the carrier-grade Netra ft 1800 server for exactly these kinds of applications.

- *High availability* (HA), is used when the uninterruptability of fault tolerance is not needed, but a much higher degree of service is required than is normally expected from a single system. The processors in HA systems run asynchronously, are loosely-coupled, and execute separate copies of the operating system. HA environments provide full hardware and software redundancy, and recovery from failures takes only seconds or minutes.

While fault-tolerant systems provide a very high level of hardware availability, they provide no coverage for software failures. Because the servers in HA configurations run asynchronously, and execute their own copies of the operating system, a software problem in one host — such as an application failure — is unlikely to affect the other host in exactly the same way. Indeed, since the majority of system failures are due to software faults, approaches which can quickly detect and recover from hardware *and* software failures are often preferable to pure fault-tolerant systems.

High-availability configurations can provide nearly the same levels of total availability as fault-tolerant systems, but at a much lower cost — and without compromising the flexibility and scalability that is so important to Internet service providers (Figure 4-1).

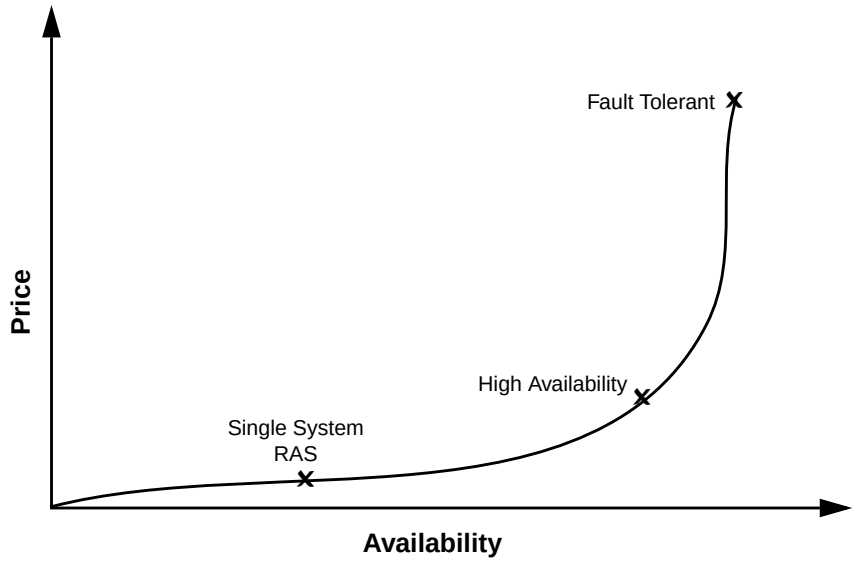

Figure 4-1 High-availability configurations provide the best price-performance trade-off for Internet service providers.

Using Parallel Constructions

Rather than use expensive fault-tolerant systems, ISPs can create resilient network architectures with parallel constructions using off-the-shelf network components and servers to deliver highly-available services. Chapter 3 provided the probability theory background for the use of parallel servers in delivering highly available services. The lessons of that discussion are dramatized by illustrating the percent availability of a set of servers each of which has only 90 percent availability (Figure 4-2).

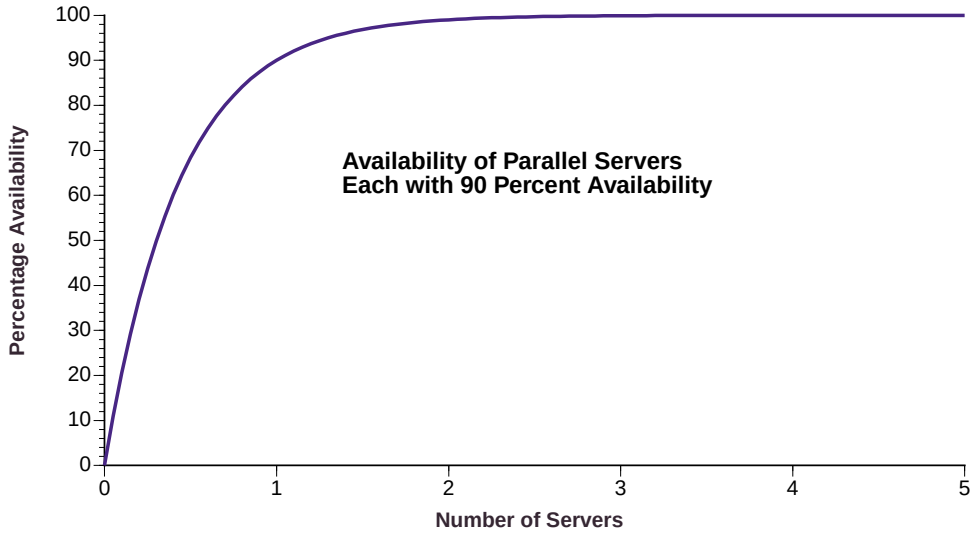

Figure 4-2 Availability increases exponentially through the use of parallel servers as illustrated by this graph of availability vs. number of servers for 90 percent available servers.

The use of parallel servers increases availability exponentially, yielding very high levels of availability with only a small number of systems. In this illustration, 99.999 percent availability is achieved with only five servers; this example allows for nearly 17 hours of down time per server per week, which enables ISPs enormous flexibility for maintenance and upgrades.

The exponential availability curve has another benefit to ISPs. Because it doesn't take a large number of servers to achieve high levels of availability, the number of servers used for any HA service is limited and more manageable. Likewise, software upgrades and hardware maintenance is limited to those

servers, keeping administration costs down. With binary compatibility across the product line, and the choice of both vertical and horizontal scalability, the choice of Sun servers by ISPs gives more flexibility in creating highly-available configurations.

Technologies for High Availability

The use of parallel servers to provide a single service is known as a *service group*. Given the use of parallel servers, the technical challenge is how to direct incoming requests to the servers in the service group so that the failure of a single system will not impact service delivery. There is a wide range of hardware and software approaches to this issue, but they basically boil down to one of two techniques that are discussed in detail in this chapter:

- *Load Balancing*

  Load balancing uses one of a variety of mechanisms for spreading the incoming workload across the servers in the service group. As long as the load-balancing mechanism recognizes the loss of a server, and can re-direct incoming requests to the remaining servers in the service group, the same mechanism that balances the workload for performance can also balance the workload for availability.

  Load-balancing techniques are best used when the data that the servers require is either shared or replicated, and where there is little or no coordination necessary between servers in the service group, for example in Web servers. For load balancing to be most effective, the user's session state is stored elsewhere so that it can be recovered in the event that their server fails.

  Load balancing principles apply to single servers as well. For example, with today's symmetric multiprocessing servers able to more than drive a single network interface, IP multi-pathing in the Solaris Operating Environment can be used to balance the workload across several interfaces on the same network. Because IP multi-pathing stops routing traffic over an un-responsive interface, and resumes once connectivity is restored, it provides a form of single-server fail-over.

- *Clustering*

  Clustering techniques make two or more servers act as a single server, responding to requests using a single IP address. The allocation of incoming requests to specific servers is handled internally by the cluster, as is the response to the failure of a server or application.

  - In *symmetric configurations* all servers participate in handling the workload, and in the event of a failure the remaining server(s) can access the customer data and continue providing services.
  - In *asymmetric configurations*, a stand-by server notices the failure of a server in the cluster and takes over its workload through a *fail-over mechanism*.

  Clusters are used for applications where there is a significant amount of coordination between the elements of the cluster; this is supported by a high-speed interconnect used for inter-node communication. Parallel and HA databases are common examples of clustered applications where inter-node coordination is necessary. Clusters are also used for services where customer data is shared and writable, and therefore only one server in the cluster can have write access at a time. ISPs frequently use this kind of cluster to provide HA NFS services. Applications supported on clusters are aware of the existence of the cluster and use cluster Application Programming Interfaces (APIs) to implement their HA services.

Architectural Implementations

There are three architectural approaches to providing high availability services which are often used as building blocks in the more complex architectures discussed in this paper (Figure 4-3):

- *Hardware load balancing*

  Hardware load balancing uses a hardware mechanism to distribute the load across a service group.

- *Software load balancing*

  Uses a software mechanism for distributing the load by redirecting requests to the multiple servers in the service group. This method is usually invisible in architecture diagrams.

- *Clustering*

  Cluster architectures are distinguished from load-balancing architectures by the existence of the high-speed interconnect linking the nodes in the cluster. Because storage devices are accessible from multiple nodes in a cluster, they are always shown with the cross connections between servers.

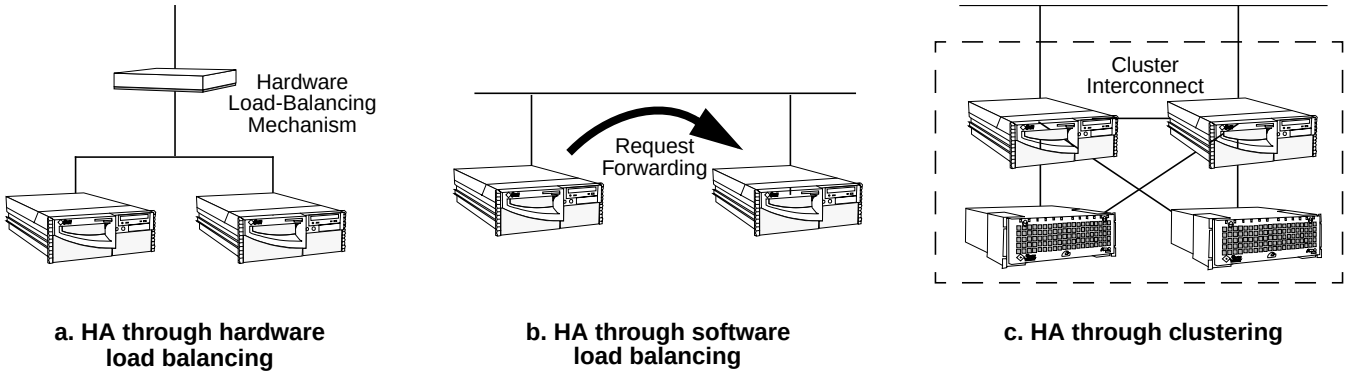

Figure 4-3 Three architectural approaches for delivering high availability services: hardware load balancing, software load balancing, and clustering.

All of these approaches can be implemented with no single point of failure. With hardware load balancing, multiple networks and multiple load-balancing switches can be configured. For software load balancing, back-up software mechanisms can be put in place. For clusters, multiple network connections can be made to each server so that the failure of a network interface or network component (like an IP switch) does not result in a service failure. For the sake of simplicity, the use of multiple networks is usually not illustrated in this paper.

High Availability through Load Balancing

With combinations of horizontal and vertical scaling, ISPs can choose the balance of high availability (through replication) and manageability (through limiting the number of servers). For applications that do not need to be aware that they are horizontally scaled, high availability through load balancing is the most commonly-deployed architecture.

A Web hosting service is a good example of an application that can be effectively horizontally scaled with load balancing. Consider the simplest situation, where a number of Web servers are configured with each one having a replicated copy of the Web content (Figure 4-4). The access network consists of a router with an Internet connection and a set of dial-up servers. The services network contains a DNS server and three Web servers, each with exact replicas of the static Web page data that they serve. Traffic between the access network and the services network is controlled with a firewall system. (True high-availability services require redundant Internet connections, routers, firewalls, DNS servers, and backbone network; however the purpose of this example is to focus on the Web service).

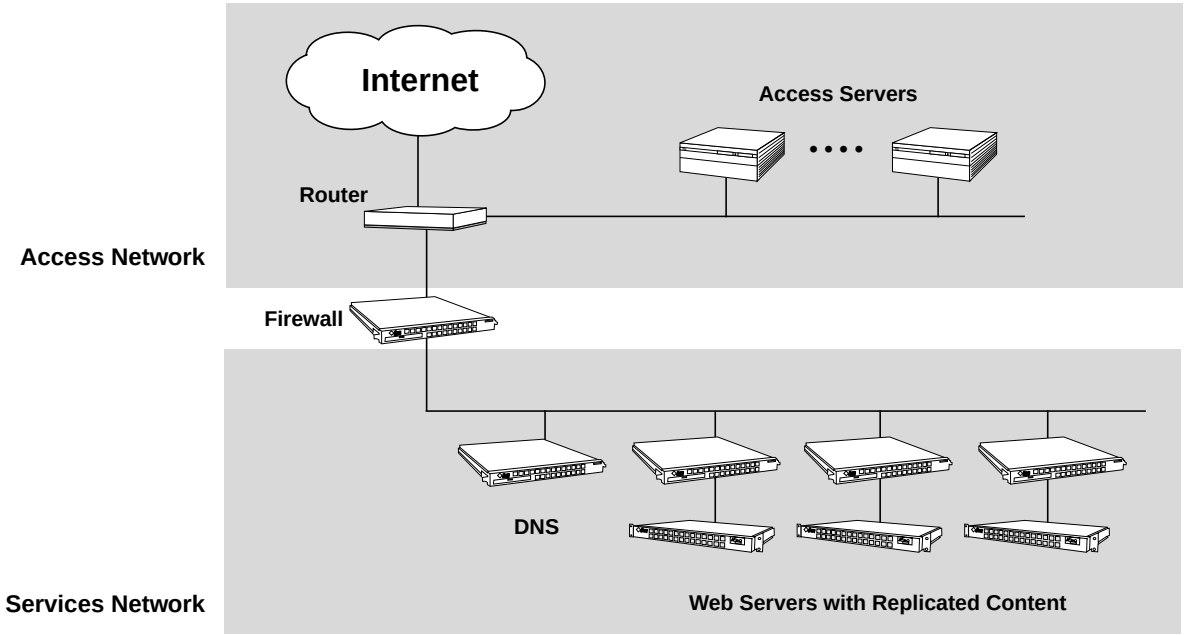

Figure 4-4 Simple configuration for a high-availability Web service

In order to provide the best performance possible, an ISP must balance incoming Web server requests across the service group. Techniques used to balance workloads also improve the availability of the service because a failed server is functionally the same as one that is overloaded — neither should allocated any more client requests.

There are several approaches to load balancing across a service group. Historically, DNS-based approaches were used to provide different addresses to clients. Today most load-balancing techniques use some form of address virtualization, where clients access the HA service using a single IP address and the load-balancing mechanism routes requests to the different servers in the service group.

Historical Approaches with Round-Robin DNS and Lbnamed

The oldest approach to load balancing, seldom used today, is to load balance through manipulation of DNS addresses. Standard DNS and Berkeley BIND can be configured to resolve a single name used for the service group — for example *www.isp.net* — to one of the three Web server addresses. This is known as *round-robin DNS*, and it involves simply configuring multiple IP addresses for the same host name.

The name server rotates through the three addresses each time a name resolution is performed, and the result is that each Web server is allocated to every third client. The address records provided by DNS have their time-to-live (TTL) set to zero, so DNS clients will not cache the address records. This results in static load balancing, which is independent of the actual workload on the Web servers — but is also independent of whether the servers are even available!

The usefulness of this scheme for high availability is limited, because it depends on an administrator manually modifying the DNS tables to remove a failed server from its list. During the time that the server is down and the DNS configuration is unchanged, one third of the clients will receive an address for the failed server, and will find the Web service to be unavailable.

A more sophisticated approach that was used for some time was to use open source Load-Balancing Name Daemon — *lbnamed* — software. This software was configured to dynamically update DNS configuration files based on the availability and performance of the various servers in the service group. This mechanism achieved a basic form of dynamic load balancing, and had as the advantage that it could quickly and automatically remove a failed server from its configuration list.

The usefulness of both of these schemes was somewhat limited, as most Web browsers cache resolved addresses independently of the TTL in the DNS address records. Users experiencing a failure on their HTTP connection do not

receive a new, valid, IP address unless they completely quit and re-start their Web browser. These drawbacks have lead to the development of more resilient load-balancing techniques, all of which utilize the principle first provided in lbnamed, to adjust the load according to server performance.

Address Virtualization Techniques

The most commonly-used approach in use today is to advertise only a single public address for the service group and use address translation mechanisms to enable multiple hosts to handle the traffic. This approach ensures that a given IP address, once obtained by a client, will always point to an operational server — overcoming the limitations of the DNS-based schemes. Load-balancing address virtualization mechanisms can be implemented using load-balancing switches or load-balancing software.

Hardware Load Balancing

Load-balancing switches are used to manage access to the subnet on which the service group is configured, and hence to which server any given request is directed. They measure each server's latency in handling requests and use response times to calculate a performance index which is used to decide how to route incoming requests.

These devices vary in the sophistication of their load-balancing mechanisms. Some make their decisions based purely on their observations of server response times, while others also factor in data obtained by agents running on the servers themselves or by measured network latencies. They usually provide flexible configuration options, for example to support 'sticky' sessions where the same client is served by the same server once the initial load-balancing decision is made. Given that they are used to support HA services, pairs of switches can usually be configured so that they do not create a single point of failure for the network.

Cisco's LocalDirector can be configured to provide simple round-robin load balancing that distributes request the next available server regardless of load; it can allocate requests to the server servicing the least number of connections, and it can allocate requests to the server with the lowest response times. Other load-balancing switches are available from vendors like Alteon Networks, Cisco, and Foundry. Load-balancing appliances based on general-purpose computing systems are available from F5 Labs and RADware. The devices vary

in load-balancing capabilities, administration capabilities, performance, and the range of protocols they support. In recent years these devices have become so fast and powerful that they have become the most popular way to handle load balancing in ISP networks.

Software Load Balancing

While hardware load balancing requires the use of routing hardware to manage the service group, software load balancing uses components hosted on the servers themselves to redirect incoming requests to the appropriate servers. Software load-balancing techniques are especially useful in situations where the organization deploying the service does not have control over the network infrastructure that directs traffic to its servers.

Software-based approaches can support routing based on arbitrary characteristics of the packet stream by enabling inspection of multiple protocol types and layers. For example, an ISP-configured service could be designed to guarantee specific service levels for premium customers based on some distinguishing characteristic of their network packets.

Resonate's Central Dispatch is the most popular software load-balancing mechanism in use today. It supports traffic management based on the servers' CPU load, the number of open connections, network latencies, round-robin, class-of-service thresholds, and ISP-designed metrics. A site controlled with Central Dispatch has a primary and a backup scheduler that responds to the IP address advertised by the ISP service. Based on its scheduling criteria, it redirects incoming requests through its TCP Connection Hop to the appropriate servers (Figure 4-5). Clients receive responses as if they came from the original IP address to which the request was sent. Central Dispatch supports multiple schedulers so that the service group continues to be managed if the primary scheduler fails. As with hardware load-balancing approaches, Central Dispatch supports sticky sessions to support services with session state.

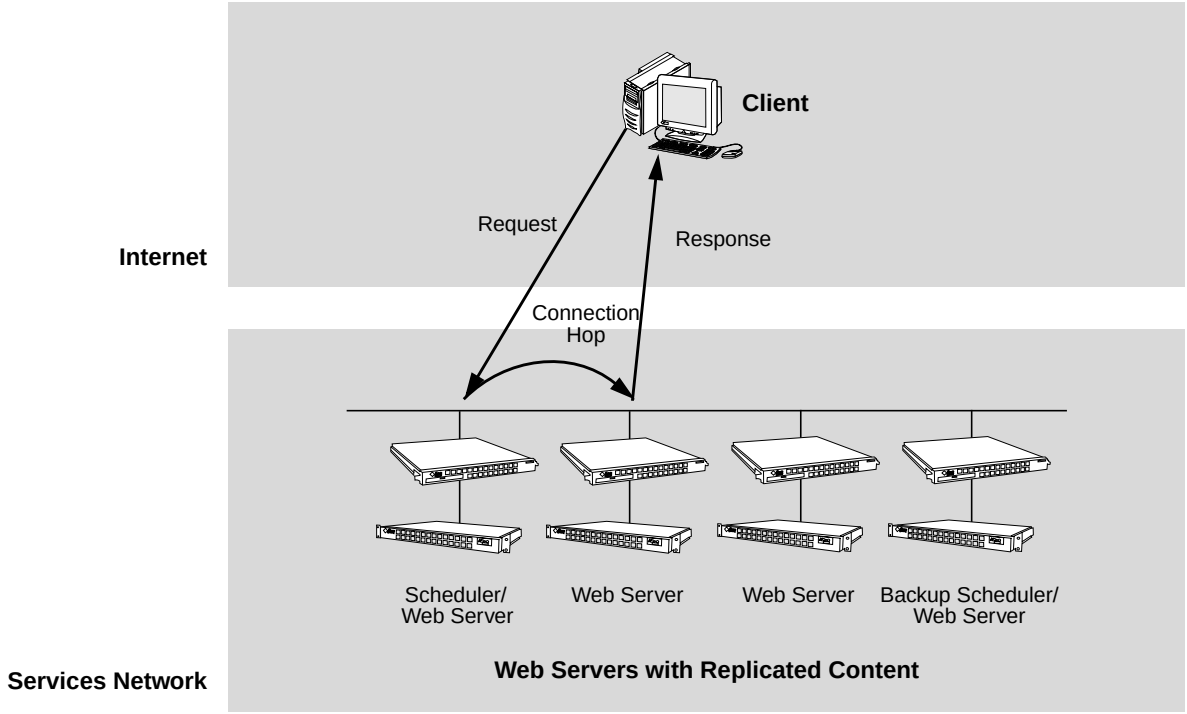

Figure 4-5 Software load balancing using Resonate Central Dispatch

High Availability through Clustering

Service groups can be hosted on cluster configurations that handle load balancing and routing decisions internally so that the cluster service is addressed with a single IP address. ISPs use clusters most frequently with software that is written to be cluster-aware and participate in the cluster's load balancing and fail-over mechanisms. The most common clustered applications in ISP environment include HA NFS; high-availability and parallel database software. These configurations are often found backing up horizontally-scaled front-end servers like Web and mail servers.

Sun Cluster software also provides a general-purpose computing environment where existing applications can run unchanged. Sun Cluster configurations appear as a single server to the outside world, while internal load-balancing mechanisms route requests to appropriate servers. They provide a unique

solution that combines load-balancing and HA services, supporting up to eight servers in both symmetric and asymmetric configurations. For example, a two-node cluster could be built using Sun Fire 280R servers and Netra st D1000 storage arrays (Figure 4-6).

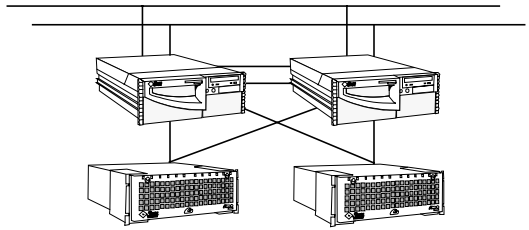

Figure 4-6 Two node Sun Cluster illustrated with Sun Fire 280R servers and Sun Netra st D1000 Arrays.

Sun Cluster configurations feature redundant paths between all systems, between all disk subsystems, and to all external networks No single point of failure — hardware, software, or network — can bring a cluster down. Fully integrated fault management software in the cluster detects failures and manages the recovery process without operator intervention, allowing failed components to be replaced on-line, without impacting availability. These configurations provide high levels of service without requiring costly, proprietary technology. The architectural approach taken in Sun Cluster software includes:

- *Redundant Disk Systems*

  Sun Cluster software utilizes multiple redundant, multi-homed disk systems, which are often mirrored to allow uninterrupted operation in the event that one of them fails. Clusters can be configured to use a range of Sun disk storage devices, including Sun StorEdge™ T3 Arrays, the Netra st D1000 array, and even low-cost Multi-Disk packs.

- *Redundant Servers*

  Sun Cluster software can be deployed with two to eight servers, enabling both horizontal and vertical scaling at the ISP's discretion. Most ISPs deploy clusters in symmetric configurations, which allows all servers to satisfy content requests from a partitioned data set. The symmetric configuration

used by most ISPs combines the performance of *N* servers for normal operation, and degrades to the performance of *N-1* servers in the event of a failure.

- *Private Redundant Network*

  The Sun Cluster utilizes redundant private, high-speed interconnects to monitor the status of each server.

- *Sun Cluster Software*

  Sun Cluster software is gradually being integrated as a standard feature of the Solaris Operating Environment. Sun Cluster software provides the intelligence by which the hardware, operating environment, and applications can be monitored, as well as the mechanisms by which failures are detected and resolved.

  Sun provides a range of services which are specifically designed to take advantage of the Sun Cluster HA environment. High availability NFS and database software, for example, is designed to be re-started in the event of a software failure, and to migrate to different servers in the event of hardware failures or scheduled maintenance activities. With Sun Cluster software, many services can continue uninterrupted despite failures without having to re-establish connections. Sun's HA software monitors the health of services by acting as clients, constantly evaluating availability and performance, initiating the re-routing of requests, re-starting of the service, or migrating or initiating migration to a different service as specified by the Cluster configuration.

Symmetric and Asymmetric Configurations

Clustered approaches to HA make the assumption that one server can potentially take on its workload plus the entire workload of a failed server. This results in a dramatic decrease in overall performance in the event of a failure. Fortunately, there are two ways to deploy Internet services in high-availability clusters:

- *Asymmetric Configurations* use one server as a hot standby which has access to, or contains a copy of, the data that is needed in order to take over the operation of a failed server. Asymmetric configurations provide consistent performance in the event of a server failure, as the workload of the failed

server is assumed by a hot standby server with no other users to support. This approach provides performance at the cost of having an idle server to handle fail-over.

- *Symmetric Configurations* use all servers to handle user requests which all continue to share the workload in the event of a server failure. In symmetric configurations, each server may, in fact, be master of a different set of data, and even support different services. Traditional clusters deploy only two servers, however today's clusters enable the workload to be allocated across *N* servers, with *N-1* servers available in the event of a single server failure. The larger the number of servers, the less the impact of a single server failure. This provides ISPs with the greatest leverage of their hardware investment, and enables them to bring servers down for maintenance without dramatic performance loss for their customers.

When using clusters, most Internet service providers use symmetric configurations, as they believe that degraded performance in the event of a server failure is preferable to the additional cost of a hot standby server.

Replicating, Partitioning, and Sharing Data

Regardless of whether service groups are deployed using load balancing or clusters, the issue of how the various servers access their data must be addressed. The Web server example used so far in this chapter uses identical, replicated, Web content. In real life, data required by multiple servers must be either replicated, partitioned, or shared; these decisions have an impact on how a service group configured for high availability is managed.

Replicated Data

In cases where the data used by the application does not change much over time, and does not require a large number of disk volumes, replication makes HA simpler to implement. DNS is a prime example of a service where a small amount of data can easily be replicated so that one server can take over in the event the other fails. News servers can use replicated databases by having news feed servers provide incoming articles to each server in a service group simultaneously. If a server goes down, the standard mechanisms provided by Inter-Network News (INN) will ensure that all articles will be propagated to it

when it comes back on-line. If Web server data does not change frequently, replicating content is a reasonable approach. If the volume is large, or changes frequently, a shared approach may be preferable.

Partitioned Data

When the volume of data is large, and the changes are frequent, partitioning is an option. User mailboxes are sometimes partitioned by ISPs because it is too costly to maintain replicated copies of user mailboxes on each server in a mail service group. Instead, user mailboxes can be allocated to individual servers, each of which is required to handle storing and retrieving mail for a partition of the ISP's subscriber base. The disadvantage of this approach is that the relationship of subscriber to mail server is fixed, requiring manual re-balancing of the workload as the subscriber base changes.

Shared Data

When there is a large amount of data to be provided by the service group, and file locking mechanisms are available to prevent data inconsistencies, shared data solutions can be deployed. The majority of Web traffic today is dynamic, requiring a back-end set of servers that provide dynamic content to all of front-end processors. Dynamic content is often provided by database management systems.

A growing number of mail servers also utilize a shared data approach, with front-end servers enabling uniform access to all user mailboxes, and the storage handled by a set of back-end servers. In some cases, the mailbox storage is provided by a high-availability database cluster; another approach is to store user mailboxes as files on a back-end set of NFS servers.

High Availability in Multi-Tier Architectures

With ISP services commonly decomposed into multiple layers for optimum scalability, security, manageability, and performance, different approaches for HA are appropriate for the different architectural tiers. Load-balanced, horizontally-scaled servers are most frequently used for front-end functions such as mail gateways; clusters are most frequently used for back-end functions such as content storage and database management systems.

Multi-tier architectures are effective ways to implement ISP services, and are often seen in the deployment of key ISP services like Web, mail, and news. An example Web hosting environment illustrates the use of different approaches to HA at different architectural levels (Figure 4-7):

- A Web site visitor over the Internet accesses one of several front-end Web servers through a redundant pair of load-balancing switches. The user is directed to one of two geographically-distant implementations of the Web site.
- The user's session state is pushed down to a pair of application servers which implement the site's business logic and provide secure insulation between the Web servers and the database management system.
- The database management system provides the application servers with information such as product catalogs used to generate dynamic Web pages, pricing information, and information on registered users.

High Availability through Multiple Tiers

With a multi-tier service implementation, different approaches for HA are implemented as appropriate for each tier, utilizing load balancing, fail-over, clustering, DNS, and simple request re-tries:

- From the point of view of a user, if a failure occurs, the most common response is to hit the browser's reload button and refresh the Web content by transparently attaching to a different front-end Web server.
- If multiple redundant sites are involved, a DNS-based approach like Resonate's Global Dispatch can be used to load balance between the sites by providing users with different IP addresses depending on whether load balancing or fail-over is desired. Many approaches can be taken to geographic fail-over, ranging from the most complex, which can maintain the state of user sessions across geographic distance, to a more simple approach which, in the event of a site failure, would lose user session state. In determining which approach to take, the impact of the loss must be considered. If multi-million dollar transactions are at stake, the former solution is cost-effective. If a site loss results in the failure of a set of in-process book sales, a more low-cost approach is often more appropriate.

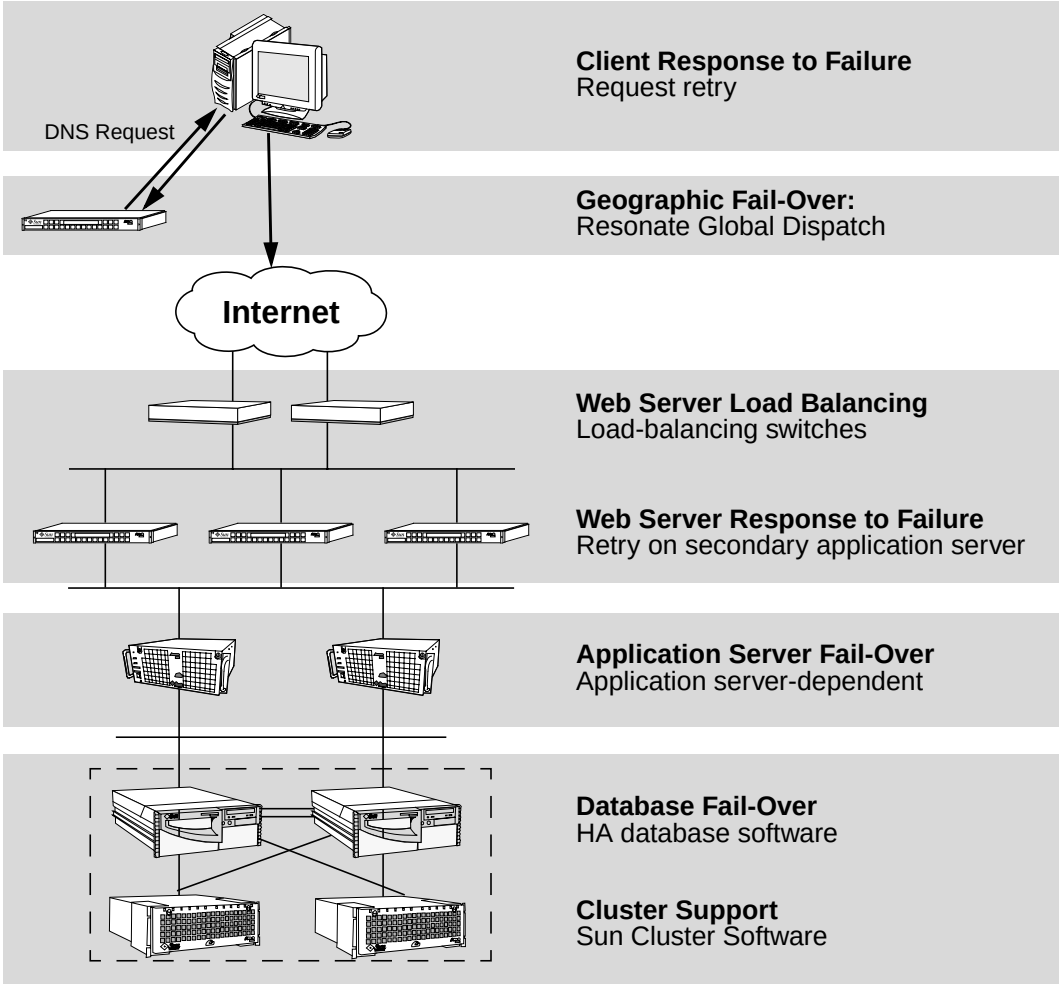

Figure 4-7 Different approaches to high availability apply at different architectural levels such as the Web hosting infrastructure illustrated here.

- The failure of a single Web server is transparent to the user because of the load balancing implemented across the service group.
- Within the Web server tier, the failure of an application server can be dealt with in many ways, and is often dictated by the application server product itself. Application server software usually includes a fail-over mechanism,

which might be as simple as statically allocating Web servers to application severs. In the event of a failure, each Web server would have a secondary address to try after a request time-out to its primary application server.

- At the application sever tier, uninterrupted service to the database management system is provided by either a parallel database server or a cluster-aware HA database implementation. Acting as a single IP address, the cluster manages its own fail-over which is transparent to the application server tier above.

Summary

Rather than use costly fault-tolerant systems which do not protect against software failures, ISPs use high-quality servers and network components in parallel constructions that provide highly-available services. The use of parallel constructions implies the use of some form of load balancing. Horizontal scaling with load balancing can be used when the application does not need to be aware of its other instances running in other servers. Clustering is used for applications where shared state is an issue, and where cluster-aware software is available. When implementing high-availability services, their data must be managed either by replication, partitioning, or sharing. Sometimes the issue can be managed by splitting up the service into multiple tiers, with front-end tiers — such as mail access servers — accessing shared data stored in a back-end database management system. Multi-tier service implementations are common in ISP architectures, and different approaches to HA can be used in each tier.

Ensuring ISP Security 5

Security is an increasingly important concern for Internet service providers. Breaches of security make the news on a daily basis, with reports ranging from Web site defacements to theft of credit card information. ISPs are right in the middle of the action, needing to fend off potential intrusions launched from customer connections, external sources on the Internet, and sometimes even from internal attacks. ISPs must guard against security breaches for more than just staying out of the nightly news — they must protect their customers' data from unauthorized use and corruption, respond quickly to Denial of Service (DoS) attacks, facilitate the use of the secure applications that they host, and even filter volumes of un-solicited e-mail. This chapter discusses the security requirements that ISPs face, and how they can meet them through the techniques of perimeter security, server hardening, and data encryption.

Paradoxical Security Requirements

The security requirements of ISPs are somewhat paradoxical. An ISP's internal network — containing billing, network management, security logging, and customer service functions — must be locked down as tightly as any major corporation's networks. Indeed, the "corporate jewels" that any ISP needs to protect the most are the customer access and billing data for services. In contrast, the front-end of the ISP network — providing network access and services to customers — must be relatively open. Many ISPs allow unrestricted traffic between dial-up customers and the Internet, which gives them freedom to utilize whatever protocols they wish in accessing services over the network. ISPs also must allow access to customer Web pages by both dial-up users and

those accessing services from the Internet. These paradoxical security concerns — open at the front, closed at the back — must be addressed in the ISP architecture through a set of carefully-constructed, access-controlled sub-networks.

Further complicating the open/closed nature of ISP networks is the need to guard against the growing variety of attacks made against them and their customers, including DoS attacks, viruses, and un-solicited e-mail. The need to protect against these problems requires ISPs to monitor their network traffic, watch for anomalies, and respond quickly to problems as they are detected. For example, many routers and load-balancing switches deployed by ISPs feature quick reconfiguration to thwart DoS attacks. ISPs also take precautionary measures to prevent their own customers from intentionally or unintentionally being involved in security issues. For example, many responsible ISPs deny outgoing SMTP requests from their subscribers to all servers other than those owned by the ISP itself, making it difficult for their customers to use other servers to relay un-solicited e-mail.

The open/closed nature of ISP networks is also influenced over time by the changing requirements placed on them. Initially, ISPs have acted as simple "on ramps," offering little or no value-added services to customers. Today, most ISPs are providing content in some capacity — offering various levels of service from hosting customer Web pages to providing multimedia interactive games and entertainment-based sites. Increasing numbers of ISPs are acting as application service providers, enabling customers to access applications ranging from business accounting services to business-to-business exchanges. This evolution demands increasing security measures (Figure 5-1):

- On-ramp ISPs must ensure the security of their own user authentication and billing information, and can do so with relatively simple security measures.
- Content provider ISPs must establish multiple sub-networks having carefully-controlled access from one security layer to the next. These ISPs make extensive use of firewall technology to implement their security policies.
- ISPs that support electronic commerce and act as ASPs must provide the highest level of security for financial transactions — requiring extensive use of encryption from Secure Socket Layer (SSL) connections between the ISP and the user, to Virtual Private Networks (VPNs) that establish encrypted IP tunnels between layers in the ISP network.

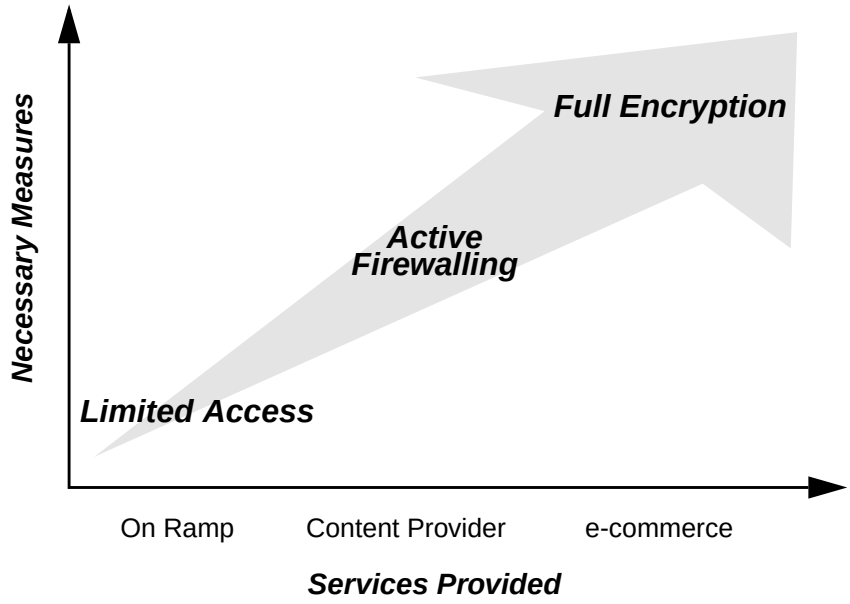

Figure 5-1 Over time, ISPs have provided increasingly sophisticated services, requiring more powerful security mechanisms to protect themselves and their customers.

Five Aspects of Security

There are five aspects of security that ISPs consider in designing a security architecture:

1. *Physical Security* protects the infrastructure from compromise. It protects systems from being removed or compromised by intruders or employees, or from being vandalized by terrorists.

2. *Perimeter Security* uses a layered architecture to force any attacker to successively penetrate several layers of security mechanisms to gain access to privileged information.

3. *Protocol Security* involves the use of proxies to protect against protocol-layer bugs being exploited to gain access to secure systems, and encryption mechanisms to prevent sensitive data from being accessible by other servers sharing the same network.

4. *Application Security* is intended to prevent ISP and ASP applications from unwittingly divulging information that should be kept private. For example, an on-line tax preparation application should protect against bugs in its software making it possible for one user to view another's private financial information.

5. *Data Security* uses encryption to protect data at rest — data that is stored on disk and contains information that should be kept private. For example, a database of ISP users and their credit card information can protect data at rest by encrypting the credit card number field so that theft of the database will not easily yield its secrets.

An ISP infrastructure is only as secure as its weakest link, so ignoring any one of the five areas could provide an attacker with the means to penetrate many other levels of ISP security. The area of physical security yields many colorful stories — including the security audit during which the data center personnel assisted the auditor with the physical removal of hard disk drives. The need for good physical security is one of the reasons why many ISPs are locating their servers within Internet data centers.

Although all areas of security are important, this document focuses on ISP configurations; therefore this chapter emphasizes network architectures that can be used to enhance security.

Categories of Attacks

In the range of security measures that an ISP can deploy, it's important to keep in mind the basic kinds of attacks that need to be guarded against:

- *Denial of Service Attacks* are aimed at making parts of the network unavailable by flooding it with superfluous requests, for example the ICMP requests that result from pinging hosts on the network. Other denial of service attacks include gaining root access to a server and making its services unavailable.

- *Confidentiality Attacks* compromise the privacy of user or ISP data by gaining unauthorized access to servers on the network. At the user level, these attacks could result in an intruder gaining valuable information from a subscriber's mail messages. At the core of the ISP internal network, confidentiality attacks could result in an intruder gaining access to subscribers' credit card numbers.

- *Integrity Attacks* occur when an intruder gains access to unauthorized data and then modifies the data. For example, a subscriber's or business customer's Web site might be defaced — a common ploy of hackers. More serious attacks on ISP data might invalidate credit card information or eliminate billing data.
- *Attacks on Authenticity* occur when an intruder replaces a standard part of the system with one that compromises security in some way. For example, a Trojan Horse login program might accept user passwords and pass them through a covert channel to an intruder who could use them to gain unauthorized access to the ISP network.

There are several basic areas that an ISP can exploit to reduce the likelihood that such attacks will be successful. The steps range from straightforward — but often overlooked — network design principles to the application of encryption technology that plays an increasing role in ISP networks.

Policy and Implementation

The first step in determining what security measures are necessary in an ISP network is development of a security policy — an unambiguous statement of what the ISP is attempting to accomplish. This statement should declare a general policy, as well as provide precise statements regarding the points between which network traffic is allowed to flow. This may seem elementary, but without taking this step, it is impossible to determine which security measures to deploy. Without a written statement, it is also difficult to assess whether the various components of the security policy (routers and firewalls, for example) are properly configured to implement the policy.

The two most common security policies are to *allow access to all services unless expressly denied*, or to *deny access to all services unless expressly permitted*. The first policy, allowing all services unless expressly denied, allows the widest range of services with the least amount of intervention and security mechanisms. This policy may be appropriate for ISPs that provide connectivity to the Internet with no value-added services. The problem with this policy is that, as new protocols and services become available, ISPs are vulnerable until explicit action is taken to prevent them from being exploited. This is why most ISPs — especially those providing more than just Internet connectivity — choose to implement a security policy that involves denying all services unless they are

specifically allowed. As new services and protocols become available, the ISP will need to assess whether to allow them and what measures need to be taken to prevent unauthorized use or intrusions.

A good security policy should take into account cost and convenience trade-offs. For example, an ISP needs to consider whether customers are to be automatically protected from intrusion originating from the Internet, or whether this is more appropriately a value-added service provided at extra cost — particularly to corporate customers, who require this level of security. Similarly, ISPs need to consider the cost/benefit of electronic mail security. Password-protection of customer mailboxes is clearly a benefit worth the cost, however is the cost of providing completely encrypted SMTP, POP, and IMAP mail connections and mail storage worth the benefit to the few customers who need this level of security?

Implementation

Once a good security policy is written, network designers can begin implementing it. There are two main considerations to make when building a secure infrastructure:

1. *Establish access control to each network element*. This is accomplished by controlling traffic to each of the ISP's sub-networks with packet-filtering routers and firewalls. This step ensures that, for example, only HTTP requests reach a Web server, limiting the reliance on the security of each individual host.

2. *Secure each network element itself*. This can involve simple measures such as removing support for unapproved protocols — for example removing telnet software from a Web server. The most sophisticated techniques restrict traffic to a small set of authenticated and encrypted connections between servers, and encrypting sensitive data at rest. This kind of "lock-down" requires state-of-the-art encryption technology discussed later in this chapter.

It is important to consider where to begin to implement a security policy. The logical sequence is for an ISP to first ensure the security of the internal network, then the services network, and finally the subscribers.

Security Audits

It is easy for designers to be so close to the problem of security that some aspect invariably gets overlooked. As part of an ISP's security policy, it is important to incorporate plans for ongoing, targeted audits of network security. This is an area where third-party perspectives can be extremely valuable, and specialized security consultants can be quite helpful. Sun Professional Services is equipped for both post-design consultation and ongoing security audits of ISP networks. Open source packages can be utilized by ISPs for their own security audits, as well.

Perimeter Security

Chapter 3 introduced the concept of separate firewall-protected sub-networks as an architecture for security. This is known as *perimeter security*. Perimeter security creates a series of concentric layers of firewall-protected layers — like moats around a castle — making the services network *most* accessible and the internal network the *least* vulnerable to intrusion. These practices are almost always used in ISP installations, and are complemented by the proper use and configuration of the network elements used to construct the layers.

Perimeter security involves not only creating concentric security layers; it requires the decomposition of applications onto servers at each level. The multi-tier model — its use so widespread that it is essentially a design idiom — benefits availability, scalability, and performance as well as security. In ISP environments, it is used for virtually all services, including Web, news, and mail network architectures.

Consider a simplified network diagram of a simple Web service deployed on an ISP network (Figure 5-2). This example is somewhat contrived to show the use of different techniques at various levels. There might be different horizontal or vertical scaling of the servers and network elements at each level; there also may be servers hosting other services (like mail) sharing each subnet. The key is to see the relationship between the different layers and how each layer contributes to overall security. The following straightforward and effective techniques are based on the principle of isolating network traffic so that it can pass — and hence can be snooped or intercepted — only between a limited number of hosts.

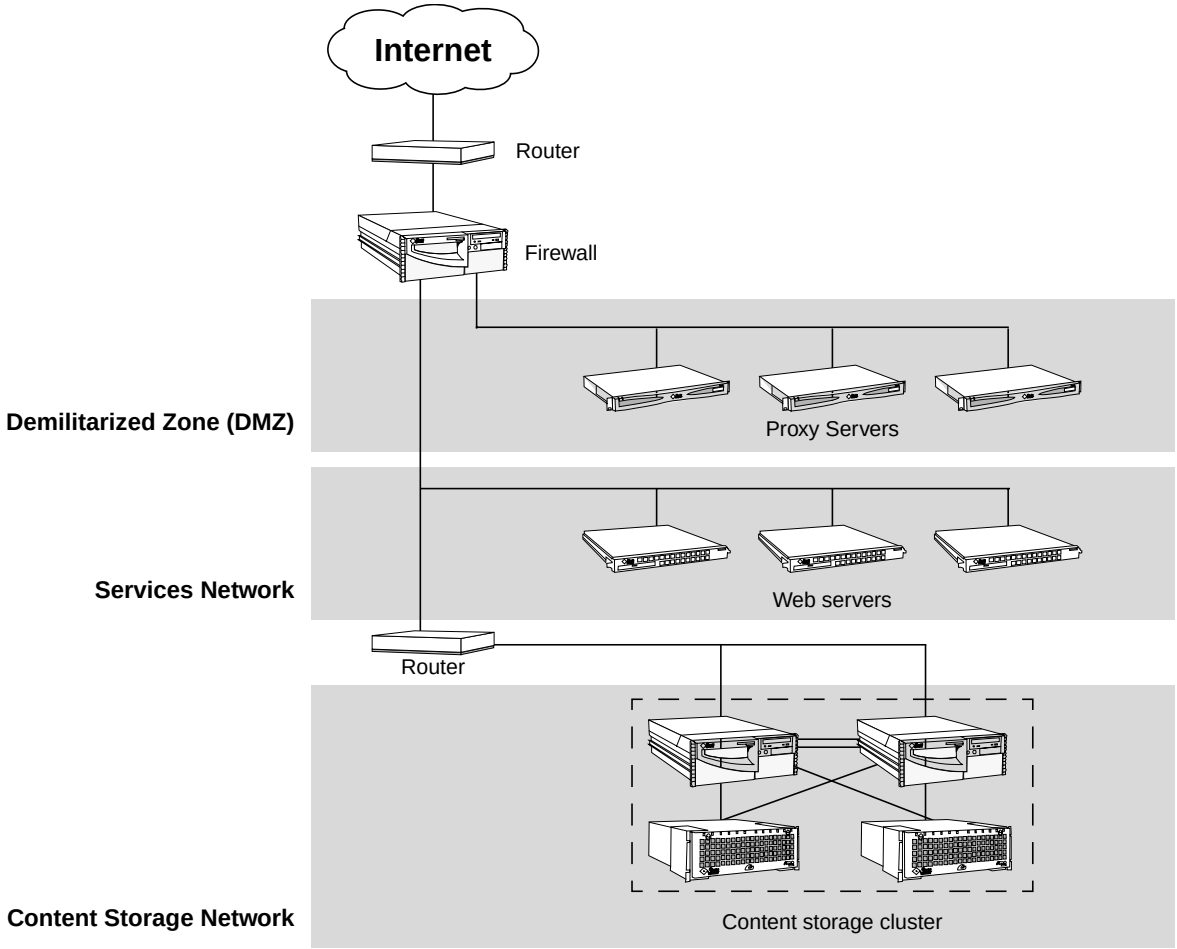

Figure 5-2 Example multi-tier Web environment illustrating the use of perimeter security through routers, firewalls, and proxy servers.

- A router directs HTTP traffic to the firewall system that controls access to the Web hosting hierarchy. This might be the core router for the entire ISP installation, or it might be a router deep inside of an Internet data center.
- The firewall, equipped with multiple network interfaces, directs incoming traffic to a set of Web proxy servers, and allows their requests to be passed to the origin Web servers in the services network.

- The Proxy servers handle incoming requests from the Internet, and respond to the requests immediately if a copy of the requested content happens to be cached, or with a page obtained from the Web servers in the services tier. The proxy servers add a layer of efficiency (through caching) and security (through protocol indirection).

  Because no packet passes directly from a client on the Internet to a Web server, an intruder would have to first penetrate the security of the proxy server in order to establish a connection to the Web servers. Even if this were accomplished, the firewall system would limit the traffic in and out of the services network to HTTP only, limiting the impact of such a penetration.

  In addition to the protocol indirection provided by the proxy server, the use of a very simple software component, such as a proxy server, at the point where exposure to the Internet is allowed enhances security. The simpler the software, the fewer the number of exploits that can be created; hiding more complex components behind relatively simple security components makes intrusion more difficult.

  There are similar mechanisms to the Web proxy server used for other ISP services; for example a mail gateway could run the Simple Mail Access Protocol Daemon (SMAPD) to protect the true mail server from protocol-based attacks.

- In this example, the Web servers themselves store no content, instead they may execute software to deliver dynamic content (through Java Servlets or CGI scripts), and obtain their static content from a shared content cluster. Separating storage from the servers enhances security as well as facilitates horizontal scaling of the Web servers.

- A router filters traffic between the Web servers and the storage cluster, allowing only network file system protocol traffic to pass. A router, rather than a firewall, is used here to highlight the trade-offs between the use of routers vs. firewalls. This trade-off is discussed later in this chapter.

The following sections discuss the use of these different network elements to establish good perimeter security.

Packet Filtering Routers

Packet filtering routers are the first line of defense for an ISP, as they allow packets to be routed based on their source and destination IP addresses and TCP or UDP port numbers. This is the basic mechanism by which, for example, an ISP can ensure that only HTTP requests are sent to a Web server. Packet filtering routers are necessary, but not sufficient, measures for establishing a secure ISP network.

There are several shortcomings of traditional routing technologies. A primary shortcoming is that many routers are stateless, which means that they do not inspect the contents of each packet in a protocol stream to ensure that they are valid within the context of the protocol itself. Another shortcoming is that routers generally do not provide logging facilities that can aid in the detection and tracking of intrusion attempts. Finally, complex rule sets dramatically decrease the performance of routers and can severely limit the available bandwidth. The most effective way to use packet-filtering routers is to use the simplest access control lists possible, and leave the more complex filtering activities to the firewall.

There are a variety of products from router manufacturers that alleviate some of these concerns. Intrusion detection modules can enable high-level analysis of traffic to detect anomalies such as DoS attacks and automatically take pre-defined actions to limit the scope of the attack. Specialized, high-speed processors speed execution of filtering rules and reduce the impact of using complex rules.

Most routers today are capable of performing Network Address Translation (NAT) functions so that internal networks can be configured with addresses that are not visible — or routable — from the outside. NAT, for example, could be established between the Internet and the proxy servers in the example above to further increase the difficulty of penetrating to successively deeper subnets.

Switching Technology

To an increasing degree, functions that were previously thought of as routing functions are appearing in switch hardware. For example, Virtual Local-Area Networks (VLANs) can be used to create virtual sub-networks of servers defined by a collection of ports on the switch. With load-balancing switches able to make intelligent decisions on the routing of requests to servers, the

capabilities exist to make decisions based on packet contents. And the line between routing and switching is becoming obscured as some switches assume level 3 — or IP level — switching capability.

Many IDCs build their networks with a hierarchy of routing and switching devices, leveraging powerful — but costly — devices by using them to manage multiple independent VLANs. For example, a distribution switch might be used to simultaneously manage many independent networks for many customers at once, amortizing the cost of complex functionality across many customers. The router illustrated in Figure 5-2 might in fact represent a combination of a set of switch ports on a VLAN combined with a small set of filtering rules on a large router that is shared between many customers.

The use of powerful routing and switching technology in IDCs and ISPs is compelling because network architectures can be created by attaching every server to the same switching fabric and configuring the ISP or ASP networks through switch and router configurations — not by error-prone plugging and unplugging Ethernet connections.

ISPs must take care when deploying their networks using shared infrastructure. Internet architects have for some time been advocating the use of switched Ethernet segments as a privacy technique. The logic is that, by using switches rather than hubs, a compromised server does not have access to snoop its entire subnet, only the traffic on the drop cable from the switch to its network interface. Thus switches, besides enhancing performance by enabling multiple independent 100 Mbps connections to be active simultaneously, could be considered security devices.

But as switches become more sophisticated, they host more complex firmware that is subject to intrusion just like any network component. There are a variety of intrusion techniques that have been successful at causing switches to momentarily bridge all traffic across them, giving hackers a chance to snoop all activity between all servers connected to the switch. At the same time, many IDCs feel that they have brought these issues under sufficient control to create safe, secure networks. Since security issues change much more quickly than this document is revised, ISPs are advised to pay attention to this issue and consult with Sun Professional Services for the most up-to-date solutions.

Firewalls

Firewalls can be used to establish highly-secure gateways in and out of the ISP network, and to control access from one level of the ISP installation to the next — from the DMZ to the services domain, for example. With routers taking the first line of defense by dropping packets that have invalid source addresses or whose destinations are not allowed by the ISP, firewalls can handle a more fine-grained, content-based filtering of traffic. In fact, since firewalls inspect application-specific data inside packets, they can be viewed as a form of access control lists for the application domain. Some of the services provided by firewalls that are key to ISPs include:

- *Stateful Inspection*

  Firewalls typically inspect the contents of packets and validate the connection type, address, protocol, and port numbers. Firewalls recommended by Sun perform stateful inspection of packets, which means that they can ensure that each packet is valid within the context of the protocol stream in which it is contained.

  Consider, for example, filtering of FTP requests. When an outgoing FTP request is made, a firewall notes that it is expecting a response from the remote FTP server. When the response is received, the firewall inspects it and then passes it on to the client. Incoming FTP packets that are not responses to outgoing requests are dropped, eliminating the possibility of FTP spoofing.

  Since firewalls typically contain knowledge of application protocols, they can also filter UDP requests based on the packet contents. For example, requests to the portmapper daemon can be filtered by protocol, helping ensure that disabled applications cannot be accessed from improper sources. This stateful handling of protocols is simply not possible with packet filtering routers alone.

- *Logging and Detection*

  Just as important as preventing breaches in security is detecting, logging, and sounding alarms as intrusion attempts occur. This allows the source of intrusion attempts to be localized, for the intruder's methods to be monitored, and for the firewall to be hardened against future attempts using similar techniques. A solid architecture for ISPs is to store firewall logs on a server in the most highly-protected internal network, and to deploy network intrusion detection software to scan for anomalies.

- *Address Translation*

  Some ISPs use firewalls to translate private, internal addresses to public, external addresses at the firewall. Private addressing helps to obscure the ISP's network configuration to any would-be intruders; it can be used to ease the shortage of IP addresses by mapping a large number of private addresses to a smaller set of external addresses; and it can be used in server load-balancing strategies that were discussed in chapter 4.

- *Encryption*

  All of Sun's firewall products support encrypted communication with remote firewalls, which is necessary for the creation of virtual private networks. The SunScreen™ SKIP package allows encrypted communication with remote clients as well. Encryption technology is quickly becoming a key differentiator between ISPs, and its implications are discussed more fully in the next section.

Firewall Scalability

Many popular firewalls — like SunScreen™ Secure Net — are software products that enable the use of standard Sun server platforms as firewalls. As software products, these firewalls can perform arbitrarily complex analysis of protocol streams at the same time they provide routing, proxy, and network address translation functions. These firewalls, along with bundled hardware/software products, have limits to scalability that are sometimes less than the entire bandwidth they need to serve. For example, stateful packet inspection of a 45 Mbps protocol stream cannot always be handled with a single-processor firewall because some protocols (like HTTP) require more packets — and thus more CPU cycles for inspection — than FTP file transfers. The key capacity limit for firewalls is packet rate, not bandwidth.

Sun Professional Services has designed a highly-scalable firewall solution that combines the use of load-balancing switches discussed in the previous chapter with horizontally-scaled firewalls (Figure 5-3). The 'sandwich' configuration uses load-balancing switches to distribute incoming requests across a horizontally scaled set of firewalls using IP address hashing so that each protocol stream is 'sticky' and continues to be inspected by the same firewall once a connection is established.

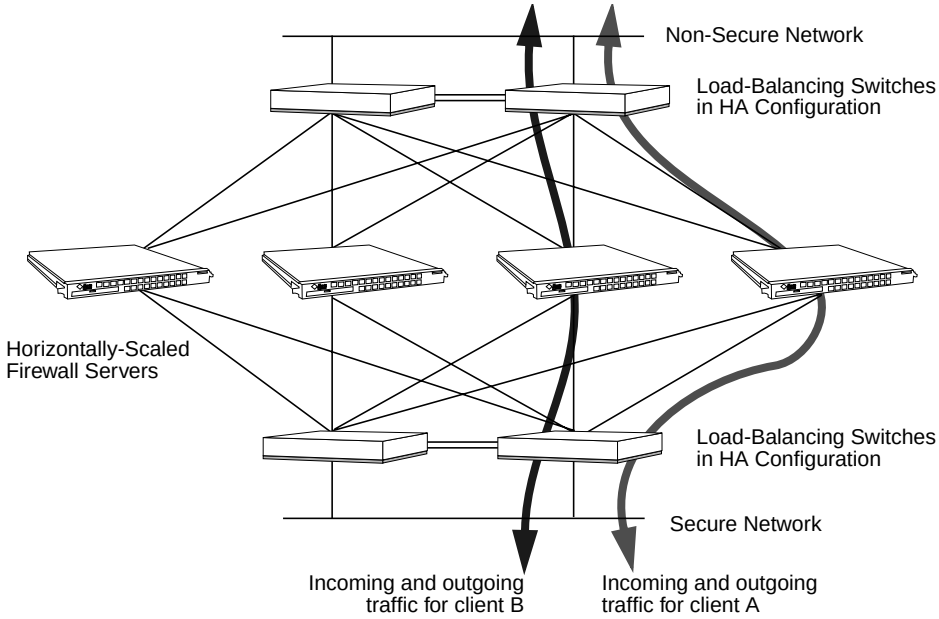

Figure 5-3 Firewall sandwich that enables horizontal scaling of firewalls for throughput and availability.

This connection management enables each firewall to maintain the state that pertains to each session and filter traffic accordingly. The back-end load-balancing switches perform the same request distribution based on address hashing for outgoing connections and response packets. Given that traffic is balanced by hashing algorithms, there is no fail-over between firewalls, so users with active sessions will have to re-start or re-authenticate their sessions when their traffic is routed through an alternate firewall.

This ingenious network architecture is extremely scalable, highly available, and is also very flexible. The capabilities of load-balancing switches today are so sophisticated that the sandwich can be configured to have specific protocols processed only be specific servers. This can be useful, for example, if an ISP prefers the HTTP filtering capabilities of one firewall product and the SMTP capabilities of a different firewall product. Both types of firewalls can be configured within the same sandwich, with protocol-level routing handled accordingly.

Server Hardening

Packet-filtering routers and firewalls are effective tools for controlling network access to each server on the ISP network. There are three ways to further harden each server to prevent or limit the extent of intrusion attempts should the network components be insufficient:

- *Access Control*

  Each server should only allow connections appropriate for the service that it is to provide. For example, telnet and FTP access should be turned off on a Web server so that any weaknesses in these protocols cannot be exploited to gain access to the server — or to load software to help further intrusion or DoS attacks. One of the first things that hackers attempt to do with a compromised server is to load their own software onto it. Denying unnecessary *outgoing* connections makes this process more difficult.

  Sun's SunScreen Secure Net firewall product can help with the process of server hardening by providing packet filtering on each ISP server. Individual services can be configured or denied with ease, and stateful packet filtering can further limit access to the server by adding an additional level at which packets are inspected before being accepted by the host. Bundled with the Solaris 9 Operating Environment, it can be installed on all ISP servers to enhance security at no additional cost.

- *Removing Non-Essential Applications*

  Further control can be achieved by actually removing the software used to implement non-essential services. It is one thing to rely on the portmapper to limit connections to unauthorized services; it is another to ensure that these services are not even loaded onto the server. This is a relatively simple security measure, but it is often overlooked. Sun has produced Blueprints documents describing how to configure servers for security, leveraging the JumpStart™ process that provides a foolproof method for multiple servers to be installed and configured with exactly the same software.

- *Non-Standard Ports*

  Just as NAT helps to obscure network topology and make it difficult for hackers to navigate, services used within the ISP network should use non-standard port numbers. Standard port numbers must be used for services — for example Web and mail — provided to subscribers and Internet clients. Non-standard port numbers can be used where traffic using standard

protocols is generated within the ISP network. For example, where SMTP is used to transfer mail from a mail gateway server to a mail storage server, using a non-standard port number makes it more difficult for intruders to gain access to lower layers of the ISP network from a higher-level machine that has been compromised.

The Role of Encryption

Once a secure set of sub-networks is established, Internet service providers can then begin to utilize encryption technology to integrate perimeter-based defense with security that is distributed throughout the ISP network. The level of security afforded by encryption technology is important for both the ISP and its corporate and individual subscribers.

Corporate Virtual Private Networks

ISPs can deploy secure virtual private networks (VPN) as value-added services for corporate subscribers. As shown in a simplified network diagram (Figure 5-4), an ISP can use a firewall to encrypt all traffic between a local company office and a remote office connected via the Internet. This is known as a LAN-to-LAN VPN, and it creates an encrypted tunnel for traffic to pass between two networks.

Traffic within each company office is un-encrypted, but once a packet addressed to a remote office passes through a SunScreen SKIP-enabled firewall or a VPN router, it is encrypted and can cross through the Internet in complete privacy. The corporate customer can enjoy the same security benefits as an expensive leased-line connection between remote offices at a fraction of the cost by using a VPN over public networks.

VPN client software is available to support connections to both firewall software and VPN router-based solutions, enabling individual users to join their company's VPN and enjoy the same level of privacy that remote offices have. These clients can access the Internet through the same ISP providing the virtual private network, or via dial-up connections through ISPs anywhere in the world. The end result is that companies have increased flexibility in supporting the computing resources of telecommuters and travelling employees.

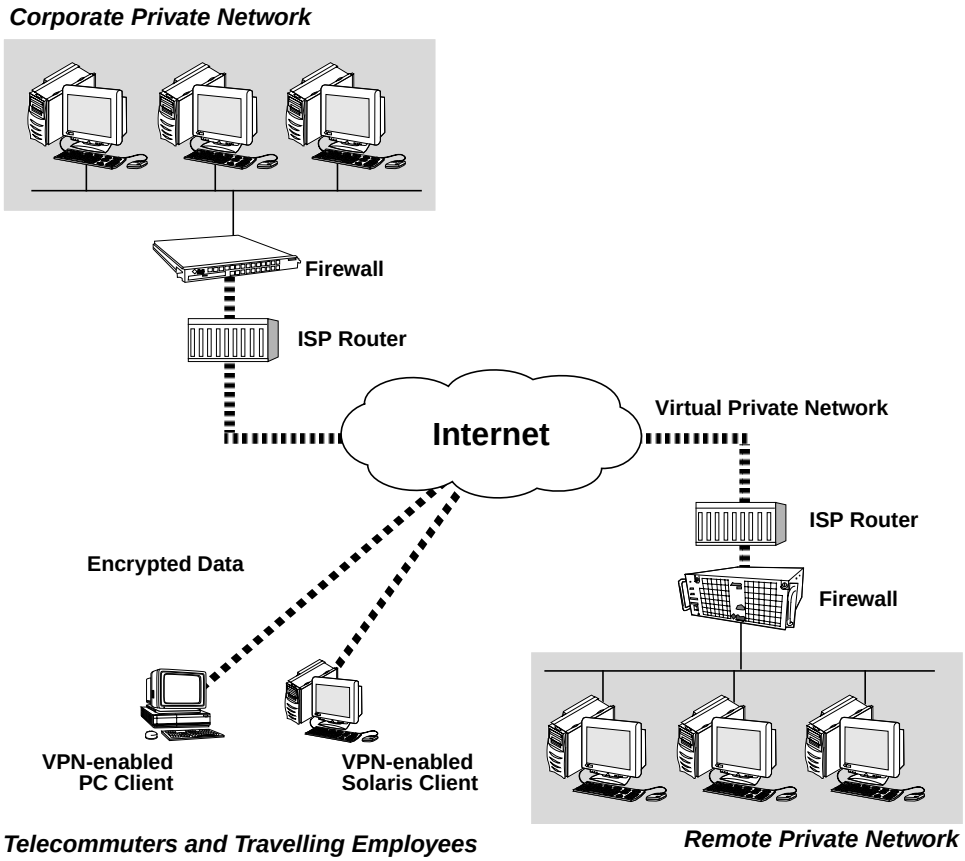

Figure 5-4 Virtual private networks can connect both remote LANs and remote clients with SunScreen SKIP-enabled firewalls.

Outsourced Portal Services

One weakness of VPN solutions is that they require fixed, pre-configured access control for clients, whether the client is a remote LAN or a travelling employee. Spontaneous access is impossible for employees at customer locations, airport kiosks, or anywhere that an Internet connection is available unless VPN software is loaded and configured on the client system.

Sun™ ONE Portal Server software provides employees, customers, partners, and suppliers secure remote access to company resources from anywhere that a Java technology-enabled Web browser is available. Using a secure Web site

with access to authorized protocols, applications, and services, companies can use Sun ONE Portal Server to extend a company's reach to far more users than can effectively be supported with traditional VPN solutions. This remote access technology is made secure with strong authentication and Secure Socket Layer (SSL) encryption.

Sun ONE Portal Server can be configured in an enterprise configuration, where remote access servers are hosted at a company site (Figure 5-5). They also can be configured following an ASP model where the servers are hosted and administered at the ISP facility. Both the initial configuration and maintenance of an enterprise configuration, or the subscription-based provisioning of an outsourced service are potential revenue sources for ISPs.

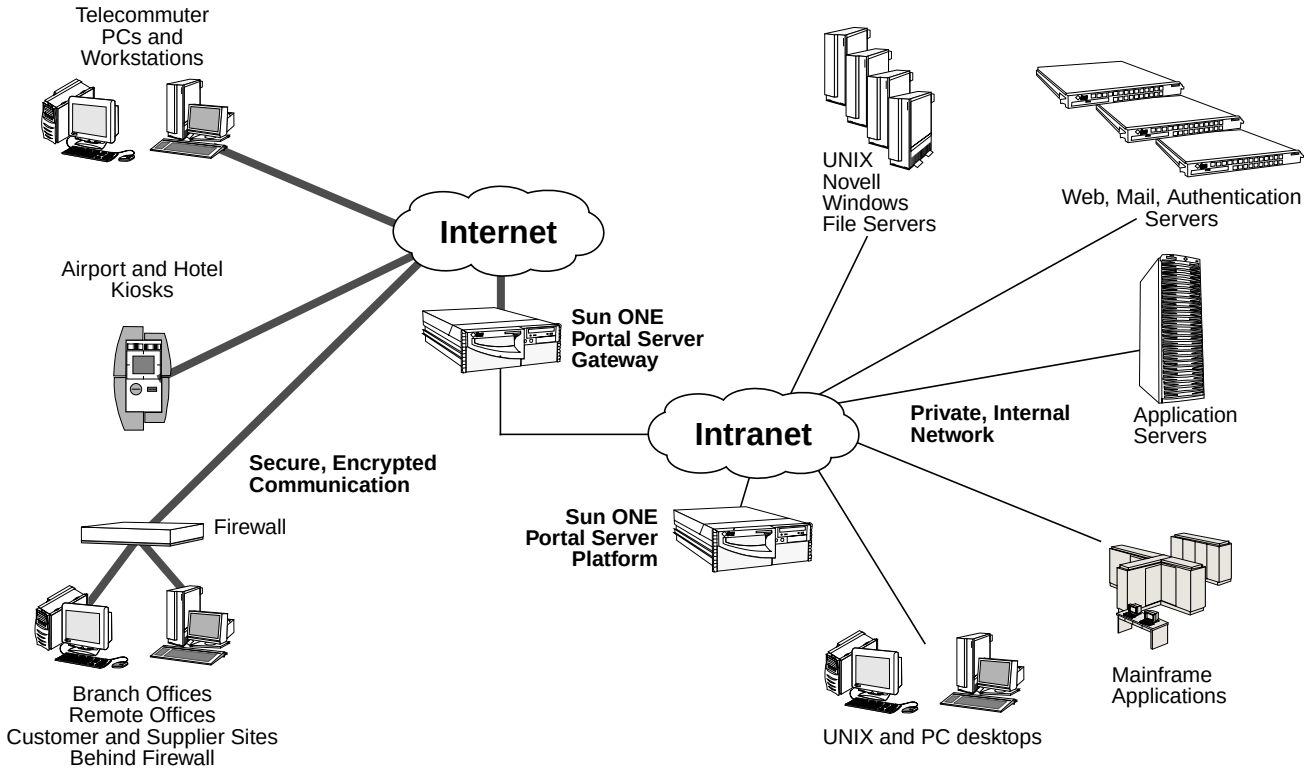

Figure 5-5 Secure, remote access is provided by Sun ONE Portal Server software hosted on a gateway and a platform servers. Remote users may be located anywhere on the Internet, including behind other corporate firewalls. Both Enterprise and ASP models can be deployed.

Secure Communication Within ISPs

Secure communication has an important role within ISP networks because it can be used to ensure authenticated, secure traffic between sub-networks and between hosts in the ISP network itself. Consider the benefits of secure communication for conducting e-commerce transactions across an ISP's internal network (Figure 5-6):

- Billing information such as credit card data can be passed into the secure internal network with little risk of an intruder using the channel to gain access to enterprise-critical data stored at this level. Indeed, many corporations use VPNs in their internal networks to ensure secure transmission of sensitive data between departments — for example corporate financial data or payroll records. The benefits for corporate intranets and for ISPs are the same — even internal users cannot view data that is meant to be private.
- Virtual private networks are an effective way to implement a security policy in which all access is denied unless specifically allowed. By "hard-wiring" the allowed connections between servers, a successful intrusion into one server is stopped short of affecting other servers in the same or more secure sub-networks, resulting in a highly-secure "lock-down" of the ISP's servers. This can be used to secure all servers at a particular layer, for example the DMZ or the services network. It can also be used between layers to provide secure communication between the central ISP location and a remote point of presence.

This kind of encrypted communication can be established between servers in an ISP facility in two ways. First, actual VPNs can be established between servers using technologies like SunScreen SKIP encryption, available with SunScreen Secure Net software. Second, Secure IP (IPSec) connections can be established between servers using the standard IPSec software available with every version of the Solaris Operating Environment beginning with Solaris 8. The ultimate lock-down of a server can be accomplished by allowing only authenticated, secure IPSec connections using specific protocols only from specific servers.

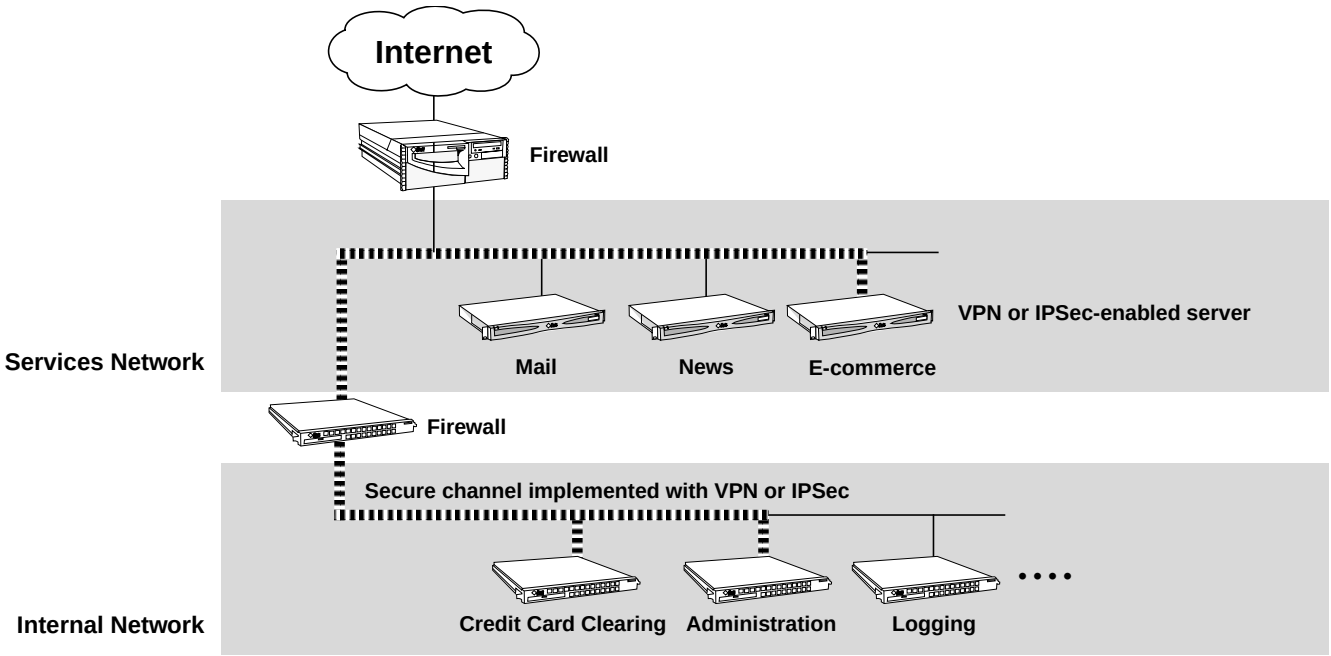

Figure 5-6 Encrypted channels can be used for increased security within the ISP network, for example transmitting credit card information from one security layer to another.

Support for Electronic Commerce

An ISP network that is properly wired for secure communication is one that is ready to support electronic commerce transactions. Because any loss in security in an electronic commerce environment has direct and immediate financial implications, it's important to make the deepest data more secure than in any other application. Encryption between network elements — for example between a Web server accepting credit card numbers and a credit card clearing server — becomes absolutely essential.

There are a variety of electronic commerce servers available from different vendors, and they generally provide services such as catalog management, search engines, automatic generation of product pages from catalog databases, sales analysis, and automated shipping and sales tax calculation. With respect to security, commerce servers typically provide secure ordering and payment methods, secure payment processing methods, and additional restrictions on

access to the commerce server itself. Although there have been some starts at creating Internet-based credit-card clearing mechanisms. the majority of clearing still happens today through hard-wired connections on the back end of a Web hosting architecture.

Summary

In the past, Internet service providers were simply on-ramps to the Internet with very straightforward security requirements. Today's ISPs provide value-added services like Web site hosting, application services, and even secure remote access for their customers. As the complexity of ISPs increases, so do their security needs. Once an ISP has established a written security policy, implementation should evolve along a path that begins with solid network design principles and techniques to deny unauthorized access, to active firewalling, and finally to the deployment of encryption technology for internal and external traffic and electronic commerce transactions. Sun believes in securing both the networks and the individual servers within the ISP installations, and is ready with security products and services to meet these needs.

ISP Services 6

This chapter discusses the major Internet service provider components and infrastructure from a software standpoint. Core components include Internet access, electronic mail, netnews, the World Wide Web, FTP, domain name service, and firewalls. Just as important as the components providing the ISP services is the ISP infrastructure, including operating environments, high-availability components, authentication, and management.

Internet Access

Perhaps the most basic service that an ISP provides is access to the ISP network and to the Internet itself. ISPs today provide access through dial-up lines, broadband connections including cable and DSL, fixed wireless, and even wireless access in retail locations such as coffee shops. Many ISPs today choose to focus on providing their core services and outsource the connectivity aspect of their business. For example, many ISPs outsource their national dial-up access to telephone company modem pools that are efficiently supported by the telephone company's existing infrastructure.

ISPs often overlook the capacity issues inherent in the access they provide, causing scalability — and customer service — problems later on. ISPs must consider these issues early on so that they can plan for current and future bandwidth requirements.

For the same number of customers, more bandwidth is required for broadband customers than dial-up customers, and their traffic patterns are different as well — the convenience of "always on" connections has broadband customers

using their service more frequently than dial-up customers. When considering bandwidth requirements, ISPs must also consider how much traffic customers will generate to access e-mail and local Web pages versus how much traffic will need to be forwarded on to the Internet.

Electronic Mail

Electronic mail has always been a key service for ISPs, and in today's increasingly-connected world, it has never been so important. Just a few years ago, supporting customers with Post Office Protocol (POP) and Internet Mail Access Protocol (IMAP) access to their mailboxes was sufficient. Today, customers demand secure access to their e-mail through Web browser interfaces that allow them to access mail at home, the office, and while travelling. With Personal Digital Assistants (PDAs) becoming more common, access to e-mail and notification via wireless messaging systems is also becoming commonplace. And with some ISPs able to integrate electronic mail with telephony services, unified messaging services that incorporate voice mail and fax transmissions into Web-based interfaces and other value-added services is quickly maturing.

E-mail was once a low-bandwidth proposition for ISPs, requiring little resource planning. Three factors have increased the bandwidth requirements of e-mail services causing ISPs to carefully consider the scalability of the solutions they implement:

- First, the ease at which e-mail client software supports attaching files to e-mail messages and the low cost of digital photography have facilitated the use of e-mail for circulating images among friends and family members.
- Second, the increasing availability of broadband connections to customers has reduced upload and download times, making it painless for users to send large e-mail messages.
- Finally, the amount of un-solicited-mail, or *spam*, places a significant burden on ISP bandwidth.

Another important trend with electronic mail is the recognition by many companies that it is less costly to outsource common services such as e-mail than to support them through their internal IT departments. A study by the Radicati Group (2000) found that 65 percent of all active mailboxes are corporate, and 36 percent of all mailboxes are paid, outsourced mailboxes. Because supporting e-mail services is routine, but requires specific expertise,

ISPs can leverage their administrative staff and their vast experience to provide corporate e-mail services at competitive prices. Some features that are merely popular for casual users are necessities for corporate users, for example integrating personal alias lists with the mail server so that the same aliases are universally-available.

Changing Requirements

Trends in the demands and sophistication of e-mail services that ISPs provide to customers have created a more stringent set of requirements on the e-mail software and server architectures that support them:

- *Traditional Client Support*

  ISPs can reduce support costs by providing a bundled Web browser and mail reader to customers with pre-configured addresses. Some ISPs configure products like Netscape™ Communicator for these purposes. Bundled packages simplify the process of bringing new customers on-line, and the reduced administration overhead results in a net savings to the ISP.

- *Web-Based Access*

  Most ISPs today provide Web-based access to customer e-mail, augmenting traditional POP and IMAP protocol support with access from anywhere a Web browser is available. For corporate customers, access to e-mail through a secure SSL-enabled site is often required.

- *Wireless Integration*

  ISP mail services increasingly enable integration with wireless networks for both short message delivery to devices such as cell phones and pagers, and full wireless e-mail support for more full-featured clients such as wireless PDAs. Wireless gateways for Web-based access detect the type of client and prepare content in either Web Markup Language (WML) or Compact HTML (CHTML) formats. With full integration of e-mail and voice messaging becoming a reality, one challenge is for ISPs to deliver messages in a secure manner.

- *Value-Added Services*

  With customers increasingly demanding e-mail access to any device, including Web browsers, traditional clients, PDAs, and even telephones, it is more important for ISPs to offer value-added services that facilitate

universal access. For example, enabling customers to store their e-mail aliases with the ISP service can enable them to use the same, up-to-date aliases independent of the client with which they access their e-mail. Likewise, calendaring services enable universal access to schedules, and integration with e-mail facilities allows companies and work groups to easily exchange schedules and appointments. Instant Messaging (IM), increasingly popular within corporate organizations, can be integrated with e-mail services to enable users to interact in real time.

- *Outsourcing Support*

  Support for outsourced e-mail is important for ISPs wishing to benefit from the trend of moving these functions from internal IT departments to ISPs. E-mail software used in outsourcing situations handles mail for multiple domains. It can handle the fact that the same account — *info*, for example — may exist in multiple domains, forwarding messages to the correct mailbox. Mail server software must support delegation of administrative duties so that, for example, customers can be empowered to create and delete their own e-mail addresses. Less of an issue with casual home users, protection of sensitive company data is critical for outsourced corporate accounts.

- *Spam Control*

  The amount of un-solicited e-mail that an ISP must manage continues to grow, consuming an ever greater percentage of bandwidth. If not filtered out before being stored in user mailboxes, it can consume significant amounts of disk space. ISPs respond to spam in three ways. First is use of mail server software with good spam control features, including the ability to subscribe to various real-time blackhole lists, as well as to define dynamic filters that can be used to distinguish spam from legitimate messages. Second, responsible ISPs take care to ensure that spam does not originate from their networks. This can be done by blocking Simple Mail Transfer Protocol (SMTP) traffic from customer connections to servers other than the ISP's own mail servers, preventing customers from relaying spam through foreign relay servers. Finally, e-mail bounce rates are an indication of spam activity. Users found to be sending spam can quickly be denied access if e-mail bounce rates are monitored and administrators paged if rates exceed a threshold. Some ISPs find that customers attempt to avoid detection by beginning their activities over the weekend when staff might not be as responsive.

- *High-Volume Message Stores*

  The use of high-volume message stores is dictated by the increasing growth in the number of ISP customers, the volume of mail sent by each customer, and the number of routes by which customers require access to their e-mail (e.g.,. Web, IMAP, and wireless). A rule of thumb is that users will consume, on average, up to 5 MB of disk space each for queued mail messages, depending somewhat on the mechanisms used to store the messages, and on the speed of user connections. It used to be common for users to receive approximately 10 e-mail messages per day; users with broadband connections typically receive 25-30 messages per day.

  A message store can be as simple as an NFS server storing user mailboxes in flat files, or it can be a high-performance database server; in either case high-availability approaches can be used to limit the exposure to the ISP should one of the servers or storage devices in the message store fail.

  Some mail servers allow flexible storage limits and others support fixed limits. The ability of a mail server to provide billing information based on actual disk space used is one way to manage users with large mail files. In estimating disk space for mail users, additional space must be allowed for incoming and outgoing system-level queues.

- *Scalability*

  The same growth that requires the use of high-volume message storage also dictates the need for highly-scalable server software that can support both vertical and horizontal scaling. Vertical scaling enables ISPs to increase the throughput of their mail service through the use of multi-threaded server software in conjunction with symmetric multiprocessors such as those available from Sun Microsystems. Horizontal scaling with load balancing is frequently used by ISPs to increase overall availability of the mail service.

- *Administration Support*

  An ISP's mail service should integrate with the rest of the ISP's provisioning, billing, network management, and Customer Relationship Management (CRM) software. When new subscribers are added, mail accounts and message store space should be allocated automatically. When subscribers are terminated, the accounts should be deleted and message store space reclaimed. Some ISPs use their CRM software in conjunction with scripts to integrate the provisioning and billing functions, and some write their own software to the mail server vendor's APIs.

Lightweight Directory Access Protocol (LDAP) has become an industry-standard means of maintaining user information across multiple services and operating environments, and many mail server products now integrate LDAP into their administration facilities. LDAP is a standard feature of Sun's Solaris Operating Environment.

Mail Server Architectures

Because ISPs support a rapidly-increasing subscriber base, a large number of messages per hour, and a high degree of security and availability, they typically deploy mail services across a number of servers that are layered for availability, flexibility and security (Figure 6-1). The ISP's DMZ hosts proxy servers to deny hackers the ability to make direct attacks on systems in the services network. Individual mail services — including the Mail Transfer Agent (MTA), IMAP, POP, Web, and wireless functions — are often deployed into individual servers, each of which can be tuned and configured for optimal performance and security. Finally, the back-end message store supports all of the functions in the services network through the file system or database access protocol integrated into the mail server software. Most mail servers can accommodate multiple message stores, using information provided by the directory server to route requests for specific mailboxes to the correct message store. The architecture illustrated in Figure 6-1 can be vertically and horizontally scaled, and can be augmented for greater security with techniques described in "Firewall Scalability" on page 65.

Mail Server Software

Given the importance of e-mail to ISP customers, the most critical decision in establishing a mail service is the choice of mail server software. An early decision is whether to buy or build the needed software. Commercial, off-the-shelf software enables a rapid time-to-market, has known scalability characteristics, vendor support, and the ability to cross-check quality with customer references. The 'build' choice often uses standard sendmail as a starting point, offering a low-cost way to begin building a business. Building mail service software offers greater flexibility, including the ability to incorporate proprietary value-added services. Though developing software comes at a cost, and risks limited scalability, some ISPs do support more than one million subscribers using software developed in-house.

The various mail server packages allow for varying numbers of users, different features, architectural flexibility, and scalability (Table 6-1). All of the products listed for regional and national ISPs support client access through IMAP, POP, Web and wireless gateways, and they all support delegated authority to support outsourced mailboxes. All mail server products listed provide some features to control un-solicited e-mail.

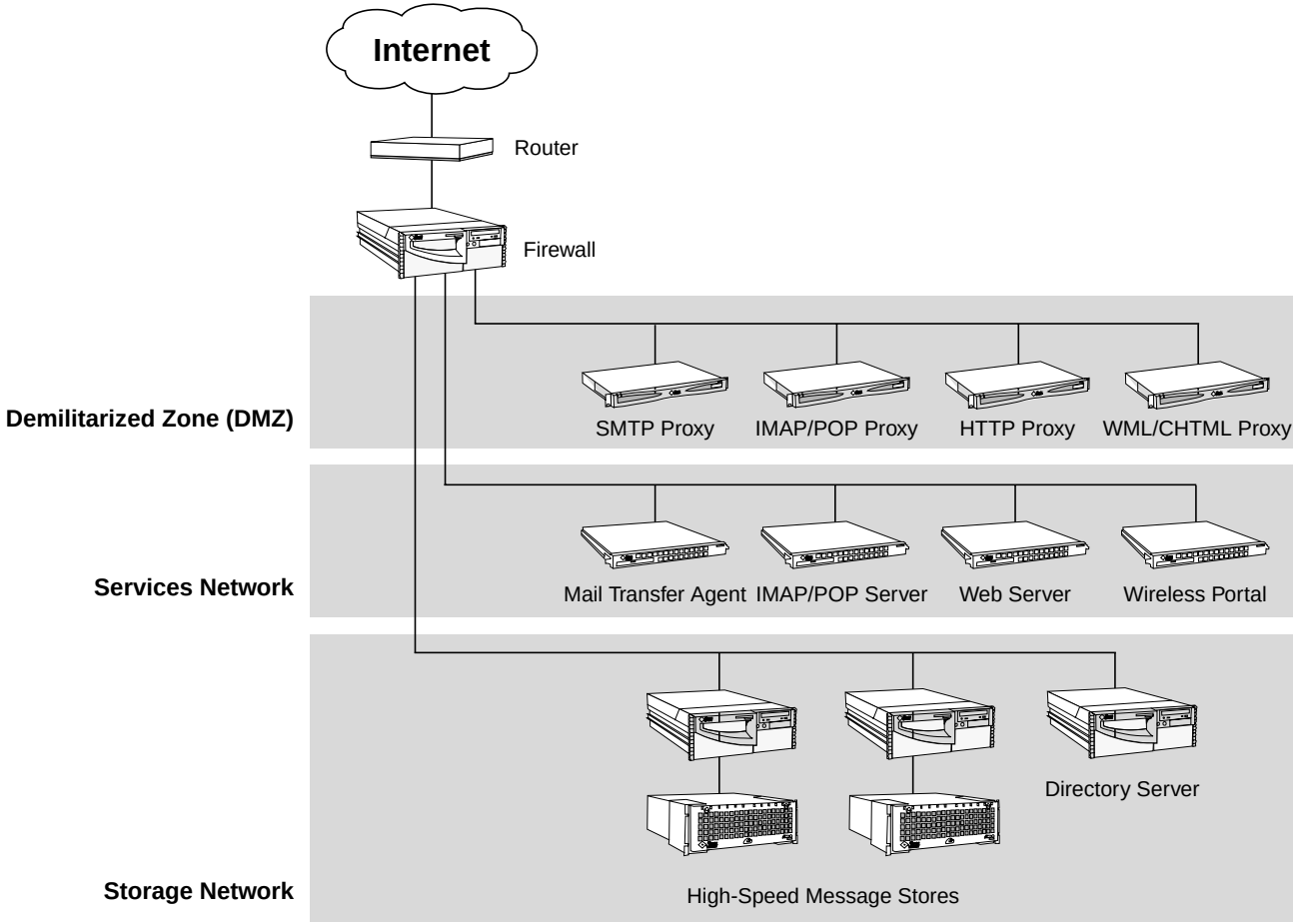

Figure 6-1 Deploying components of the mail service onto separate servers enables horizontal and vertical scaling while enhancing performance, availability, and security.

Standard Solaris Operating Environment utilities can support local ISPs up to approximately 10,000 users; Openwave's (formerly Software.com) legacy *Post.Office* extends support to approximately 50,000 users for local ISPs. Openwave's *Email Kx* is appropriate for ISPs with 50,000 to 200,000 users, with *Email Mx* supporting ISPs with up to 25 million users. *Sun ONE Messaging Server* supports both regional and national ISPs with millions of users.

Mail Software	Local	Regional	National
Standard Solaris Operating Environment Utilities	✔		
Openwave Post.Office	✔		
Openwave Email Kx		✔	
Sun ONE Messaging Server		✔	✔
Openwave Email Mx			✔

Table 6-1 Mail software for various sizes of ISPs

Standard Utilities

Small ISPs often use standard sendmail for outgoing mail, with open source Post Office Protocol (POP) and Internet Message Access Protocol (IMAP) servers for providing incoming mail to subscribers.

Sendmail accepts outgoing mail from the customer's mail client software using Simple Mail Transfer Protocol (SMTP). Sendmail also transfers and queues mail messages to SMTP servers at other sites.

POP and IMAP servers are used to provide access for mail clients to read, save, and delete messages from their mailboxes. IMAP enables clients to compose, delete, and send mail while disconnected from the server — synchronizing changes once the connection is re-established. POP and IMAP protocols are used for mail retrieval only.

Standard sendmail does not scale well beyond approximately 10,000 subscribers for two reasons:

1. Sendmail's local delivery agent stores all incoming mail messages in a single directory (*/var/mail*), making performance and access using normal UNIX® tools quite difficult.

2. Both sendmail and standard POP/IMAP servers require subscribers to be registered through the standard UNIX mechanisms, either */etc/passwd*, NIS, or NIS+. These mechanisms all use sequential searches — which do not scale well — to locate records in their respective databases. Newer versions of sendmail (8.10) support LDAP integration, which helps to alleviate this issue.

Various ISPs have customized sendmail and POP servers to overcome the problem with storing all mailboxes in a single UNIX directory. One ISP has modified sendmail's delivery agent to deposit mail in a directory structure where the full path name of a mailbox can be generated by the subscriber's account. Another approach is to store mailboxes in large indexed files. Using off-the-shelf databases is another solution for overcoming the limitations of */etc/passwd* and NIS.

Post.Office

Openwave's (formerly Software.com) Post.Office is a drop-in replacement for standard sendmail and POP, and scales up to 50,000 subscribers per mail host machine. It supports mail accounts for users without login IDs, eliminating a major restriction of the standard UNIX utilities. The most important advantage of Post.Office is that it provides increased functionality and simplified administration for local ISPs.

Account management is handled by a mail or Web-based forms interface. Post.Office does not run with root permissions, which makes it significantly more secure than standard sendmail. It supports size limits on individual mail messages, mailboxes, and for the entire mail system. Its ability to handle mail for multiple domains makes it easy to support commercial and small business customers with their own domain names.

Email Kx

Openwave's Email Kx has many of the same features of the company's premium Email Mx, but scaled for regional ISPs with 50,000 to 200,000 users. Email Kx is the lowest-level product discussed here that enables both vertical and horizontal scaling by enabling the hosting of different functions on

separate servers, including the Message Transfer Agent (MTA), POP server, Web server, and message store. IMAP and wireless servers can be added onto the basic Email Kx configuration.

Email Kx supports multi-level delegated authority for e-mail administration. This enables, for example, delegating authority for creating new e-mail domains to local points of presence, and in turn delegating authority for creating and deleting users to the ISP business customers owning the domain. Additional levels enable large companies to delegate within themselves, so different organizations within a company can have independent control their own mailboxes.

Email Kx uses a message store built for high availability; up to two message stores can be configured with the primary failing over to the secondary in the event of a failure. User information is stored in an LDAP database.

Sun ONE Messaging Server

Sun ONE Messaging Server takes the best features from Sun Internet Mail Server (SIMS) and Netscape Messaging Server (NMS) and combines them into a carrier-grade platform for ISPs with one of the industry's lowest total cost of ownership. It can be deployed to exploit both horizontal and vertical scaling, and it can be deployed in a Sun Cluster environment for high availability. It was built from the ground-up to support large ISPs and telecommunications companies, so it uses multi-threading for superior performance, reliability, and scalability. Understanding that security of intellectual property is paramount to outsourced business customers, Sun ONE Messaging Server supports encrypted POP and IMAP communication to customer mail clients, and SSL connectivity to Web-based mail servers. Sun also designed Sun ONE Messaging Server so that it can host individual domains on separate servers, giving these important customers even greater peace of mind over the security of their outsourcing decisions.

Sun ONE Messaging Server supports all standard messaging protocols, and features include delegated management, virtual domain hosting, and individual user address books. As part of a larger platform, Sun ONE Messaging Server functionality can be extended through wireless gateways and even unified messaging with the storage of voice mail messages that can be retrieved either through a telephone or e-mail interface. Sun ONE

Messaging Server can also be extended with products like Sun™ ONE Calendar Server and Sun ONE Portal Server to provide value-added services such as calendaring and secure remote access for ISP business customers.

Sun ONE Messaging Server supports the use of multiple message stores for rapid growth; indeed one ISP customer using Sun ONE Messaging Server to support a free e-mail service *has added 12,000 new users per day*, testimony to both the message store scalability and the ease of adding new users. To simplify user management of configurations that require multiple messaging servers, Sun ONE Messaging Server utilizes a Mail Multiplexor (MMP). With MMP, each user is assigned to a specific messaging server but requests mail from the MMP server. The MMP server transparently re-directs incoming requests based on user entries in the LDAP directory.

The Sun ONE Messaging Server Message Transfer Agent (MTA) that transfers incoming messages into the message store has withstood the test of time, with various versions of the software in use for more than fifteen years. The MTA's design helps ensure that no mail is lost because messages are committed to disk before the MTA acknowledges their receipt.

Email Mx

Email Mx (formerly Software.com's InterMail) is a high-end mail server product, able to support thousands of concurrent connections to tens of millions of mailboxes. Email Mx supports all standard e-mail protocols, including IMAP, POP, and SMTP, and can provide Web and wireless-based access to user e-mail. Email Mx has been installed for use with tens of millions of customers. It can scale vertically, horizontally, and services can be geographically distributed.

Using the flexible architecture of Email Mx, ISPs can configure feature sets in any combination to best meet customer needs, including the ability to define classes of service with business customers receiving priority. The delegated authority and customer 'self care' features enable delegation through multiple levels of business customer organization, with departments and workgroups able to administer their mail services to the degree to which they are authorized. Likewise, multi-layer delegation can enable various ISP points of presence to administer their own local customers, including adding new domain names and delegating administration authority to local business customers.

The Email Mx message store is based on an Oracle database, and unlike Email Kx, many message stores can be configured and geographically distributed as well. This feature is key for ISPs with worldwide operations because user accounts can be assigned to the geographically-closest message store, with a local LDAP database containing the knowledge of local user accounts, and a master LDAP database in a central location consulted when local users roam to other points of presence. Using the journaling feature of the Oracle database, a record of all transactions can be replayed in the event of a serious outage, providing a means for recovery.

Netnews

Hosting a netnews service is a high-bandwidth proposition, with news volumes steadily increasing. Each ISP has one or more news feeds from which it obtains new articles and newsgroups. New articles are distributed using a flood model, where the ISP receiving new articles is responsible for weeding out duplicates. Many ISPs also host local news groups containing information and discussions of local interest. Netnews consumes so much bandwidth that some ISPs are re-considering whether the traffic volume is worth the revenue that it generates. Some smaller ISPs outsource news services to their upstream provider to save both administrative and bandwidth expenses. Some larger ISPs utilize commercial netnews providers and establish dedicated satellite connections to support the deluge of data.

As of April 2001, a full news feed in the U.S. included more than 50,000 groups, averaged more than 650,000 messages per day, and transferred an average of almost 72 GB of news articles per day, up from the 18 GB per day reported by this document in 1999. The average bandwidth for such a feed exceeds 7 Mbps, and ISPs must not only allocate bandwidth for such a feed but also must consider how much bandwidth their commercial customers might require if they establish private news feeds. The amount of disk space required by netnews is enormous, however it can be adjusted to accommodate the disk space available by carefully tuning expiration times for messages.

Almost all ISPs use news server software based on the Internet Software Consortium's open source Inter-Network News (INN) or commercial news server products including the Cyclone News Router and the Typhoon News Server from Highwinds Software (formerly part of Software.com.)

News software is typically divided into at least two components. A news router handles news feeds, and a news server handles interactions with clients, transferring articles via Network News Transfer Protocol (NNTP). Client software is integrated with major Web browsers, and independent news clients are available as well. When evaluating which news server products to use, ISPs must carefully consider whether the software enables different functions to be partitioned onto different servers, enabling the use of horizontal and vertical scaling to accommodate large numbers of users and rapid growth.

World Wide Web Services

The single reason that has compelled most residential and commercial subscribers to sign up with an ISP is the World Wide Web. The Web is a proven information-gathering resource, and Web-based electronic commerce is firmly established as a way for companies to do business. Informational Web pages range from movies and personal health to company product information. E-commerce sites provide consumer items of all types, and business-to-business sites speed the flow of information and purchases between companies, lowering costs and helping to maintain competitiveness. Transactional activities are expected to increase, ranging from mom-and-pop grocery stores speeding their soft drink orders via thin-client network computers to micro-transactions where fractions of a cent are charged for database searches — with the Internet service provider serving as the clearing house.

ISPs preparing for these changes are establishing high-performance Web services using architectures that can scale as their customers' Web use expands and changes. ISPs generally provide three levels of service to their customers depending on their needs:

- *Vanity Sites*

  Most ISPs provide modest Web hosting services to their casual users as part of their service package. Typically used as personal — or 'vanity' — sites, these sites are hosted within the ISP's domain, have a limited amount of storage allocated to them, and have low bandwidth requirements. Customers often use these sites for displaying family photographs and personal information. A rough rule of thumb is that 10 percent of dial-up customers will utilize this service, 20-30 percent of broadband customers

establish such sites, and only about 20 percent of the sites are changed frequently. Although ISPs usually allow up to 5 MB of space, an average vanity site hosts approximately from 500 KB to 2 MB of content.

These numbers can vary dramatically — Sun's affiliates have observed that there is a significant difference between urban and rural points of presence. Urban subscribers are more likely to have high-speed connections, and rural subscribers tend to access the Internet less, and more frequently visit Web sites of local interest hosted by their ISP. These numbers also vary by country, with areas just beginning Internet use using less space than those where it is well-established.

- *Virtual Hosting*

  Small commercial enterprises often host informational sites or simple e-commerce sites where they wish to use their own domain name, but where their limited need for storage, network bandwidth, and security can be satisfied by hosting several customers on a single, shared server that supports virtual, or shared hosting of multiple domains.

- *Dedicated Hosting*

  Larger companies typically have storage, bandwidth, functionality, and security requirements that dictate the use of a dedicated infrastructure that ranges from a single, dedicated Web server appliance to a full multi-tier Web hosting infrastructure that includes front-end Web servers, application servers, and database servers.

Most ISPs support all three of these hosting models in some capacity, and choose Web server software that can enable them to meet their customer requirements. Some ISPs provide value-added services including Web site design, development, infrastructure for transaction support, and limited CGI scripts for customers wishing to host forms, guest books, etc.

The major ISPs predict that some future revenue growth will come from content services. The Internet is becoming a source for pay-per-use video games, personalized on-line news services, on-line shopping services, and other sources of entertainment. As content providers themselves, broadband ISPs have already recognized this potential, and they tend to build ISP configurations geared to providing interesting and entertaining Web sites. It is possible with sophisticated configuration techniques to open different avenues of Web page access to each subscriber depending on the level of service they purchase.

Web Service Architecture

A basic Web service architecture that might be deployed by an ISP includes several components to enhance security and performance: a firewall to control traffic between layers, a proxy/cache server in the DMZ, and the Web servers themselves in the services network (Figure 6-2).

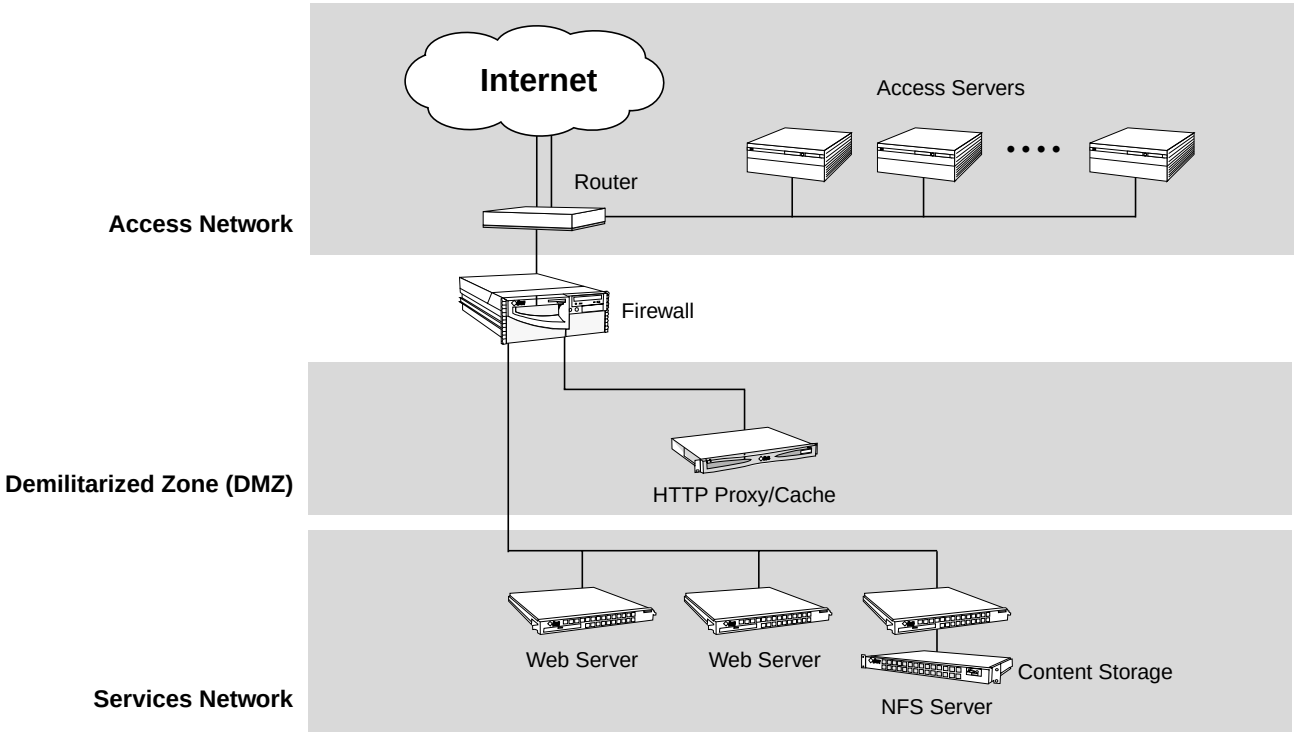

Figure 6-2 An ISP's Web server infrastructure usually includes a proxy/cache server in the DMZ and load-balanced Web servers in the services network.

Customer requests come through the access servers to the HTTP proxy/cache server which satisfies the request immediately if the request is in its cache. The proxy/cache server acts as a proxy for requests of Web servers in the services network, and as a cache server to satisfy requests with cached content. Most ISPs configure their proxy/cache servers to cache requests made of servers outside of the ISP itself, significantly reducing the Internet bandwidth needed when multiple customers access the same sites. An ISP's Web servers might be

horizontally scaled using a load-balancing mechanism to increase both performance and availability; proxy / cache servers could be horizontally scaled as well. With horizontally-scaled Web servers, ISPs must either replicate content on each server or share it from a single storage device. Figure 6-2 illustrates the use of an NFS server to provide shared content to both Web servers.

Web Server Software

The key piece of software required for supporting Web hosting services and enhanced services including e-commerce transactions is the Web server itself. Most Web servers provide support for virtual hosting, and extensions such as server-side include files and Java Servlets. Products vary in the degree to which they can scale and in the customer support available. Some of the popular Web server solutions include:

- *Apache Web Server.* The most popular Web server on the Internet — 60 percent of all Web servers by a Netcraft study — is the open source Apache Web server. Because of the number of developers working on Apache, and the number of sites that it supports, this open source software has become known for its reliability. Apache supports most functions required by ISPs ranging from fast CGI support to Java Servlets. Apache runs on most UNIX platforms as well as servers running Microsoft Windows.
- *Sun™ ONE Web Server.* Supporting 66 of the top 100 Web sites (Gartner Group) is Sun ONE Web Server. With the Netscape Web server as its foundation, Sun ONE Web Server has been at the forefront of Web server design from the beginning. Featuring unparalleled scalability, Sun ONE Web server supports sites like *cnn.com*, which serviced 150 million hits and 6.3 million individual users, setting a record level of traffic on Election Day 2000 in the United States. Unlike open source software, Sun ONE Web server is supported, commercial software designed for business-critical sites.
- *Sun Cobalt Server Appliances.* Offering a complete hardware and software solution for ISPs providing both shared virtual and dedicated hosting services to customers, the Sun Cobalt RaQ™ 4 and Sun Cobalt RaQ™ XTR servers provide low-cost and easy-to-maintain solutions for ISPs in a single rack unit-high configuration. Through its Web-based administration interface, ISPs can configure virtual Web sites as well as e-mail, FTP,

e-commerce services, and even delegate administration capabilities to customers so that they have authority to administer parameters (like adding and deleting e-mail users) within their own domain.

Other Considerations

Web server software is just one part of an effective Web service. Other software that an ISP must consider include:

- *Proxy Servers*

  Proxy servers act as intermediaries between requests from users and the Web servers. These servers (as illustrated in Figure 6-2) are used to protect ISP Web servers from both ISP customers and other users on the Internet. HTTP requests issued from client Web browsers are directed to the proxy server which, in turn, fetches the requested data from the Web server. By acting as an intermediary, proxy servers limit the ability of intruders to make direct attacks on Web servers. Users making outbound requests through proxy servers are likewise somewhat more protected against TCP hijacking attempts on their session. Because they are used to filter traffic, proxy servers are usually combined with caching servers so that repeated requests to the same content can be satisfied with cached data. Many Web servers can be used as proxy servers, including the open source Apache Web server; commercial products like Sun™ ONE Proxy Server and Inktomi Proxy Server provide specialized solutions for ISPs. Both of these products provide proxy and caching functions.

- *Caching Servers*

  The larger the ISP, the more likely it is that Internet bandwidth will be used to request the same Web pages on behalf of many different users. A caching server makes more efficient and cost-effective use of Internet resources while hiding network latency from subscribers. Caching has been shown to save ISPs up to 40 percent on network infrastructure costs while at the same time providing superior quality of service to subscribers. ISPs can use caching server appliances like the Sun Cobalt CacheRaQ™ 4 server, software including the open source SQUID server, or they can take advantage of an increasing number of network switches and load-balancing mechanisms that provide on-board content caching capabilities.

Clearly not all content can be cached, with dynamic content being a common example. Some sites, in order to obtain accurate hit counts, mark their content as non-cacheable, so that all accesses pass through the caching server.

An increasing number of ISPs and IDCs are using caching servers at the *edges* of their networks to push content closer to customers, resulting in increased performance for visitors to the sites they host. For example, a U.S. company with an IDC-hosted site supporting edge caching can mirror its content at points of presence around the world. This instantly delivers a global presence because requests to the customer's site from other continents are served as rapidly as those from within the U.S.

Two approaches are in common use to push content closer to customers. The ISP or IDC can host caching servers in their remote points of presence, selling edge caching services to customers willing to pay a premium for the increased performance. Some IDCs implement this strategy with racks of Sun Netra T1 servers running edge caching software from Inktomi. Another approach is taken by Akamai, which places servers in strategic locations around the Internet, caching their customers' content from anywhere on the Internet to Akamai servers located in or upstream from end users' ISPs.

- *Content Filtering*

  Although used primarily for security, proxy servers also provide a convenient point for ISPs to filter content, allowing them to offer subscriber-specific packages that limit access to sites having content inappropriate for children. Commercial tools are available to provide content filtering services, and packages like SQUID can be configured by the ISP to act as a filter.

- *Log Analysis Tools*

  There is no better upselling tool than providing customers with detailed statistics on visitors to their Web sites. Likewise, statistics on Web site usage can enable ISPs to bill based on traffic, and also to balance network activity based on complete knowledge of Web site access patterns. Log file analysis is one way to obtain this information, and packages such as those from WebTrends can analyze and graphically present information to ISPs and their customers.

- *Performance Monitoring*

  With quality of service such an important differentiator between ISPs today, continuous performance monitoring of Web services is important for ISPs and their customers. Stand-alone packages like Freshwater Software's SiteScope can monitor responsiveness and even track correct execution of e-commerce transactions. Add-on components to network management systems like BMC Patrol or Micromuse Netcool can also be used to track Web site performance.

File Transfer Protocol

Many ISPs provide their customers with FTP sites that enable them to store files downloaded by others. These sites must be carefully administered so that one user cannot inadvertently access another user's files. Directories with global write permission should be avoided to prevent the storing of files by unauthorized users.

Numerous commercial FTP server products are available, including Sun Cobalt servers, which provide FTP services in conjunction with virtual Web domain hosting. Perhaps the most commonly-used FTP server is the open source wu-ftp package. When selecting an FTP server for integration into an ISP environment, care should be taken to ensure that provisioning of sites can be automated, and that authentication information can be used from the same directory used for other user information, typically an LDAP database. Some ISPs provide customers with delegated administrative authority so that they can add new FTP users and change passwords on their own; this is a standard feature of the virtual hosting functions of Sun Cobalt server appliances.

Domain Name Service

Domain name service (DNS) maps domain names to host addresses for business and residential customers as well as for clients on the Internet. In order to access an Internet service, an address for its server must be obtained through DNS. One of the most important functions of DNS in an ISP network is to provide a limited view of the ISP network so that clients are only allowed access to a small set of servers. Many large ISPs have a public DNS that reveals addresses that can be accessed from outside the ISP, while an internal DNS server provides mappings that make all of the ISP's systems accessible for internal use.

Separate domain names are typically defined on a *per-service* basis. For example, a news service might be accessed through the domain name *news.isp.net*. A Web service might be accessed through *www.isp.net*. These domain names provide a layer of abstraction that an ISP can use for ease of configuration, load balancing, and fail-over. For example, a small ISP might have the names *news.isp.net* and *www.isp.net* map to the same physical server. As the ISP grows, these two services may be hosted on separate servers and with only a change in DNS, customers will correctly access the new configuration.

Domain Name Services	Local	Regional	National
Standard Solaris Operating Environment Utilities	✔		
ISC BIND		✔	✔

Table 6-2 DNS software for various sizes of ISPs

Local ISPs typically use the standard domain name service supplied with the Solaris Operating Environment (Table 6-2). Regional ISPs often begin with the open source Berkeley Internet Name Daemon (BIND) available from the Internet Software Consortium, and modify it to suit their particular needs. Another option is new commercially-available DNS and DHCP software from Nortel. Some of the variations and common modifications include:

- *Static Load Balancing* can be achieved with both versions of DNS. When multiple servers are used to host a single service, DNS can be configured to provide different addresses for the same domain name on a round-robin basis. The effect of round-robin DNS is to statically balance the workload across multiple servers. Though DNS-based approaches to load balancing generally have been superseded by load-balancing mechanisms integrated with switches and routers, this simple technique is still used for load balancing messaging systems. For example multiple MX (Mail eXchange) records can be configured in a DNS server to achieve static load balancing across several ISP mail servers.
- *Dynamic Load Balancing* can be achieved by modifying BIND to monitor the load on various servers and map addresses on the basis of measured server loads. In the past, ISPs used developed various schemes using the open source *lbnamed* package. Today, most ISPs use load-balancing routers and software like Resonate's Global Dispatch to achieve the same results by

virtualizing IP *addresses* rather than *names*. Today's load-balancing mechanisms are simple to configure, and avoid many of the drawbacks of DNS modifications. These approaches are discussed in detail in Chapter 4.

- *Fail-Over* can be supported by modifying BIND to interact with fail-over software on the hosts that provide a particular service. If, for example, mail is hosted on two servers, and one fails, DNS would stop providing address mappings for the failed server and route all traffic to the second server — which could be either a hot spare or part of a load-balanced set of servers.

  One problem with name-based approaches to fail-over is that many Web browsers cache DNS responses as long as they are running. In the event of a server failure, new addresses are not obtained and a failed-over Web site continues to appear inaccessible to the user. IP address virtualization solves many of these shortcomings and is the technique used in load-balancing mechanisms, as well as for clusters.

- *Client-Sensitive Mappings* can be used for several purposes. Different subscribers, for example, can be given different server addresses depending on the specific content package they have purchased from their ISP. Edge caching techniques rely on an intelligent DNS server that can point clients to the closest cache rather than the origin server that might be a continent away. These DNS servers use detailed knowledge of Internet routing tables to effect this client re-direction.

One of the primary shortcomings of DNS is the lack of intuitive user interfaces to simplify configuration — making DNS administration a tricky and error-prone activity. One solution to this problem is to integrate the name mapping tables with a DBMS so that the database, not the DNS configuration, contains the definitive name-mapping tables. A graphical user interface can be developed for the DBMS to allow entries to be added, deleted, and modified; and back-end software can periodically generate new configurations for the DNS service and re-start the daemons. This solution greatly simplifies administration overhead, and is especially useful for ISPs with customer domain names that are frequently added and modified.

Security with Routers and Firewalls

Routers and firewalls are used to provide protection between the various sub-networks that make up an ISP architecture. As discussed in chapter 5, there are numerous hardware and software products that can be used to help ensure ISP

security. Sun's SunScreen Secure Net software is a high-performance solution that can be configured in two different modes depending on the level of security that is required.

- *Routing mode* enables a Sun server to act as a router while performing stateful packet filtering on packets that pass through its interfaces.
- *Stealth mode* provides a significantly higher level of security by using a dedicated, hardened version of the Solaris Operating Environment, and operates as a transparent network device with no IP address. Since it is invisible to intruders, it is highly impervious to intrusion.

SunScreen Secure Net is multi-threaded, allowing ISPs to configure firewalls using symmetric multiprocessing (SMP) servers to meet desired performance levels. SunScreen Secure Net provides proxies for a range of protocols, and a free download is available for use with two or fewer network interfaces per server. SunScreen Secure Net software can be used to harden servers throughout the ISP infrastructure by limiting the ports on which they will respond and by filtering the traffic to and from each server.

ISP Infrastructure

There are many choices that an ISP must make when designing the infrastructure to support the services provided to customers. Choice in operating environments, high-availability strategies, billing, and network management have significant impact on an ISP's ability to deliver services and scale with a rapidly-growing subscriber base.

Operating Environment

Choice of operating environment is one of the most important decisions facing Internet service providers. Depending on the choice, ISPs can constrain their architectures to a limited number of possibilities, or maintain the greatest amount of flexibility for future growth. Sun's Solaris Operating Environment is designed from the ground up to provide reliable, scalable, multi-user, multi-platform network computing — one of the reasons why Infonetics Research (2000) reports that 89 percent of the UNIX servers at Tier-1 Internet service providers are Sun servers.

- Scalability is key for successful ISPs, whose subscriber bases can grow by orders of magnitude per year. The Solaris Operating Environment can support vertical scalability for performance and capacity, and horizontal scalability for reliability. For example, using multi-threaded services such as e-mail and Web servers, the use of symmetric multiprocessing with the Solaris Operating Environment enables the utmost in performance. Sun Cluster software supports horizontal scaling by enabling standard services to be configured to utilize as many as eight servers all sharing a single IP address and sharing the workload equally.
- Reliable systems are a requirement for ISPs, and the Solaris Operating Environment delivers maturity that can only be achieved by years of use by many different users in many different environments — backed up by the hard work of eliminating bugs over time. Hardware failures are inevitable facts of life, and the dynamic reconfiguration features of the Solaris Operating Environment support hot-pluggable components, management of individual processor sets, support for alternate IP paths through redundant network interfaces, and separate operating environment domains on supported platforms.
- Support for standard Internet protocols is the basis for an ISP's operations. Sun invested in Internet standards like TCP/IP beginning more than fifteen years ago, and the result is a protocol stack whose reliability, functionality, and interoperability has been refined for many years. In fact, the 1998 release of the Solaris 7 Operating Environment included a multi-threaded TCP/IP stack that enables servers to leverage symmetric multiprocessing right down to the network protocols.
- Multi-user systems result in more cost-effective computing because the resources of networked systems can be shared remotely. Authorized users can log in and access services as needed. Client systems can access a wide range of facilities on servers — including file, print, and name services, Web pages, and database services. Administrators can manage networked systems remotely, resulting in significantly lower cost of ownership. Sun servers are easily configured to provide multiple services — such as mail, file, and Web services — at the same time to many simultaneous users. Features such as these are not available on systems where multi-user support is an add-on.
- Ease of administration is required for ISPs whose architectures include distributed points of presence; with network-based administration provided by the Solaris Operating Environment, ISPs can retain central control of

resources and reduce the expense of administrators having to perform on-site maintenance tasks. Lights-out management features of Sun's Netra servers combined with Solaris software support helps ISPs to maintain distributed points of presence in unattended locations such as telephone switching centers.

- Security in ISP networks is a fundamental requirement, and the Solaris Operating Environment provides facilities that enable servers to be locked down and made impervious to attack. Between the inherent security provided by the Operating Environment and the additional security that can be provided by utilizing the included SunScreen Secure Net software, Solaris Operating Environment-based firewalls provide the best protection available.
- Resource management is a key issue for ISPs needing to support various quality-of-service levels for customers. Solaris 9 Resource Manager software enables management of such critical resources as CPUs, memory, and I/O throughput in a hierarchical manner, giving high-priority processes and/or customers and increased share of resources when system utilization is high.
- Solaris 9 Resource Manager's bandwidth management functions enable ISPs to allocate network bandwidth levels for specific services. They can be used, for example, to maintain a minimum T1 level of service for one customer's Web server while throttling another site to a maximum 56 Kbps of bandwidth.

High Availability

High Availability (HA) is becoming more of a necessity in ISP networks, especially in telco and cable ISPs where subscribers have the expectation that Internet services will be as available as normal telephone and cable services. ISPs use a variety of horizontal scaling techniques to achieve high availability. For some services where significant state is stored in the application — like a database server — high availability is best accomplished using a server cluster.

Sun Cluster software implements automatic fail-over mechanisms that allow backup servers to take on the load for a failed server, with more sophisticated solutions allowing multiple servers in a cluster configuration to share the workload equally. HA solutions typically have a small time lag from the time a failure is recognized to the point at which the service is restored on other servers.

- Sun Cluster software is designed for environments requiring high-availability data, file, and application services. When services are managed in the Sun Cluster environment, they are provided with rapid detection and recovery from any hardware, network, operating system, or application software failure.
- Sun Cluster software supports pre-packaged applications that utilize its high-availability features, most importantly NFS, database, and Internet services including mail, news, and Web servers. The Sun Cluster environment also supports an HA toolkit with which ISPs can create their own HA solutions.

Solutions for high availability in ISP environments are discussed in detail in chapter 4.

Authentication

Lightweight Directory Access Protocol (LDAP)-based repositories are quickly becoming the norm for maintaining user information — including passwords — for ISPs. Because of the interoperability that it offers in multi-platform and multi-service environments, many vendors are providing the capability to query LDAP repositories for user authentication data. For example, ISP-related services like e-mail servers from both Sun and Openwave use LDAP-based authentication. Beginning with the Solaris 8 Operating Environment, Sun servers have built-in LDAP support. Some open source tools can be configured to refer to LDAP repositories, and for legacy devices such as terminal servers, integration packages are available to bridge between LDAP repositories and the RADIUS protocol authentication they require.

There are many open source LDAP servers available, however such a key function is better hosted by commercial software like Sun ONE Directory Server. Built with multiple threads to leverage the power of Sun SMP servers, Sun ONE Directory Server offers linear scalability, and responsiveness to handle up to 5,000 queries per second. Sun ONE Directory Server supports integration with RADIUS authentication as well.

Management and Administration

Though small ISPs can operate through manual configuration of services, those larger than approximately 20,000 subscribers can become severely hampered without automatic procedures in place for customer care, billing, service level

monitoring, network management, and provisioning. At this point in an ISPs growth, administration overhead can skyrocket due to the configuration complexity that accompanies so many users.

Fortunately, the telco and Internet data center industries have experience with these issues, and the lessons they have learned apply to ISPs as well. Because many ISPs are located in IDCs, those with strategic relationships can even leverage the same tools for managing their infrastructure.

Management and administration functions are best viewed in a hierarchy, with functions that require customer interaction at the top, and management of individual network elements and services at the bottom. A diagram used to describe Internet data center management and administration functions is illustrated in Figure 6-3. At the top of the diagram is the range of customer care processes for which there is a general flow from left to right. Customer care activities feed requirements to the service management — or operations — layer, where service packages are defined, implemented, and managed, with information on quality of service reported up to the customer care level. Services provided to customers require interoperation of a range of network components that support and deliver the services; these are planned, configured, provisioned, and managed at the next layer. Finally, the servers and services are coordinated to provide continuous service availability to customers.

Regardless of whether all functions illustrated in Figure 6-3 apply to a specific ISP, the diagram serves as a reminder of those management functions that should be considered. For example, an ISP must decide what, if any, service level agreements are to be provided to customers. If none are to be provided, then the management structure to support SLAs can be omitted.

It is a significant undertaking to determine which software tools should be purchased to help manage ISP operations, and an even greater challenge to integrate various tools from multiple vendors so that an ISP can operate with as little administrator intervention as possible. Sun has prepared a whitepaper entitled *Architectures for Managing IDCs* which can help ISPs make these decisions and to design a tool integration strategy.

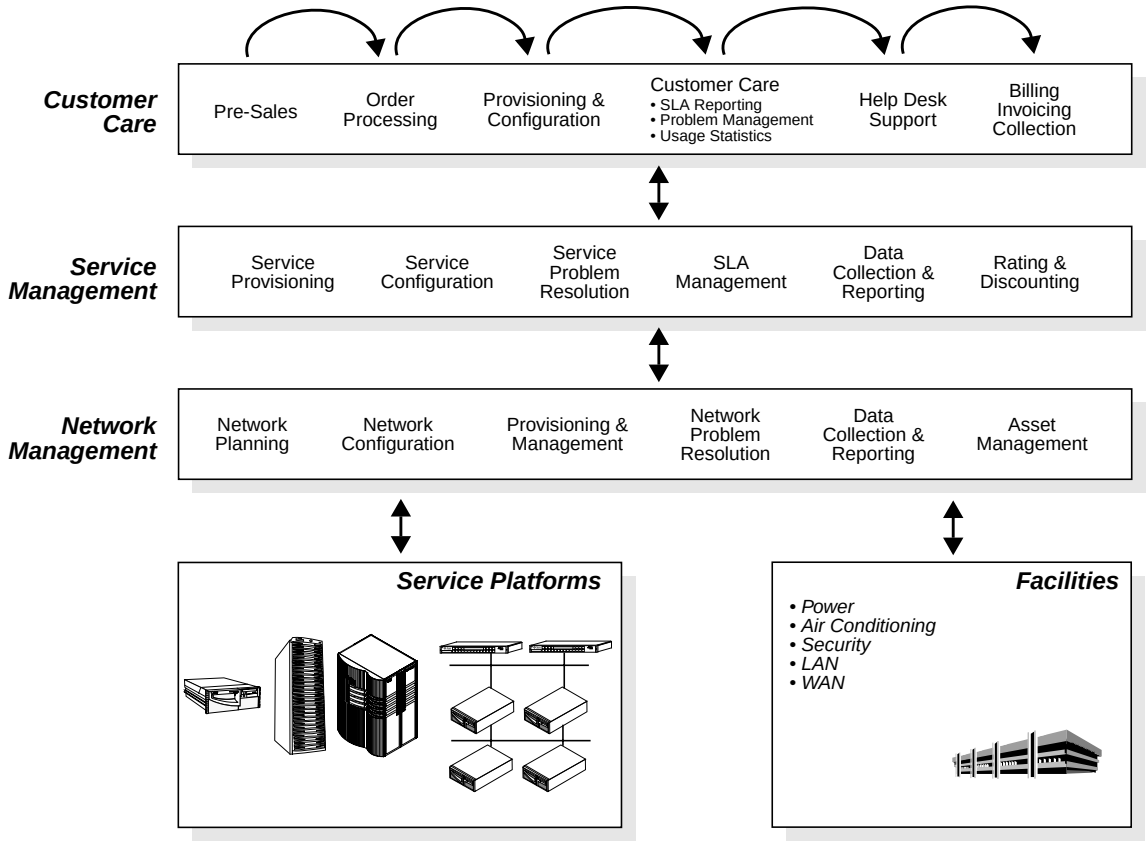

Figure 6-3 The hierarchical model of managing telco and IDC infrastructure applies to ISPs as well.

Summary

Many software components must be assembled to provide a full range of ISP services, and the deployment of each service can vary substantially depending on the subscriber base, workload, and expectations of growth. Solutions — such as standard UNIX utilities — that are appropriate for small Internet service providers may not scale acceptably into mid- and large-sized configurations. Given that scalability is key, it is often best to make only investments in tools and infrastructure that will carry an ISP well into the future.

Value-Added Web Hosting

7

Beyond the basic ISP services that include mail, news, Web, and Internet access, ISPs are increasingly providing other value-added services to enable competitive advantage and produce additional revenue. The most widely-adopted additional service is Web hosting, with virtually 100 percent of ISPs offering some form of hosting. Web hosting services can be categorized into three types: personal, or ‘vanity pages; light commercial; and commercial. Each of these three categories have different requirements for service and management.

As the popularity of hosting has increased, so has the complexity. The industry has evolved from hosting basic static content to delivering dynamic content, advanced applications, and electronic commerce functionality. Accompanying this increased complexity are issues of performance, availability, and security surrounding proprietary data and electronic transactions. Though this chapter is focused on Web hosting, the same principles apply to any other application that an ISP might host for its customers

Static vs. Dynamic Content

One of the major factors influencing the design of Web hosting services is whether the sites deliver static or dynamically-generated content. The most straightforward Web hosting service to implement is static content hosting. Static content does not change, or is only changed manually by the Web site creator. This includes HTML files, text files, pdf files, graphical images, and other media. Each customer is allocated disk space according to their service level, and FTP or direct access to their own document root directory.

Dynamically-generated Web pages, in contrast, use some programmatic method to create content. These pages can contain dynamic content such as stock quotes and customer-specific information such as order status. A good example of a dynamically-generated page is an application where the viewer can query the Web site and cause the material to be assembled using some template mechanism — for example generating catalog pages from images, descriptions, and prices in a back-end database. Dynamic pages are generated using a wide range of facilities including basic Common Gateway Interface (CGI), Java Servlets, JavaServer Pages™, and Web server APIs. The most common approach is to use an application or process, separate from the Web server, to process a request and perform some function. This is a fundamental difference from static pages, for which the Web server simply delivers the file content.

Resource Requirements

Resource requirements for Web hosting environments vary depending on the type of content provided by the hosted sites:

- *Static content* presents minimal infrastructure demands on the ISP. Static content is generally easily hosted in a shared environment, with relatively few security concerns.
- *Dynamic content* is typically associated with e-commerce sites, giving rise to special security considerations both on the part of the ISP and its customers. Given that various programming methods are used to provide dynamic content, there is the possibility that these programming methods have the potential to cause security or functionality problems within the site. The source of these problems may be either programming errors generated by customers, or attacks mounted by hackers.

For these reasons it is often desirable for commercial Web sites to use dedicated resources (servers, disk systems, etc.) for hosting each customer's dynamic content. Indeed, most ISPs and IDCs provide only dedicated hosting options for commercial sites. Dedicated resources also provide specific hardware methods for ensuring quality-of-service levels for each customer. They help ensure that each customer's programs and data do not interfere with, or physically intermingle with those of other customers.

ISPs often choose different Web server software for different functions. For example some sites might inter-mix the use of commercial servers like Sun ONE Web Server for hosting and electronic commerce functions with the use of

servers like Zeus or Apache for delivery of static content including graphics and product images. Many ISPs find that open source servers like Apache are perfectly satisfactory for delivering static content, preferring the company backing of commercial servers like Sun ONE Web server for e-commerce sites.

Application Architecture Complexity

There are several application architectures for dynamic Web service from which to choose. These are presented in order of increasing complexity, where the more complex architectures offer the greatest benefits in terms of scalability, security, performance, and support for dynamic content generation.

- *Static Content using CGI*

  This is a simple form of dynamic content generation that includes Web pages created by CGI processes that access static file content. While this architecture does provide dynamic content, it does not maximize performance, generality, and maintainability. There are several alternatives which provide additional functionality and performance, but which also increase the application architecture complexity.

- *Dynamic Content with Java Servlets and JavaServer Pages*

  More complex than CGI, but more secure and high-performance is the use of Java Servlets to generate Web site content. Because Java Servlet code essentially extends the functionality of the Web server itself, it can be used for generating content where session state is involved, or where the application requires integration with other functions such as those provided by application or database servers. JavaServer Pages (JSPs) enable portions of the Java code to be integrated in the Web page itself, making it easy for customers to modify the appearance of pages without re-writing the server portion of the software.

- *Database Access Using JDBC™ and SQL*

  The approach uses an interface from the page generation mechanism to access a generic database system. This can be done, for example, via standard SQL interfaces or a JDBC interface. A major advantage of this model is that the database content is readily linkable (with other applications) and easily maintainable. This allows for an increase in the sophistication of the underlying data resources, but does not result in significantly-increased application functionality.

- *Application Servers*

  Application servers — like Sun™ ONE Application Server — can provide dramatic improvements in application functionality by running code that supports business rules in a separate server tier. Application servers provide a container for executing software like Enterprise JavaBeans™ components, providing improvements in programmer productivity and improved availability, performance, and scalability.

- *Transaction Servers*

  In complex Web applications, the Web server may interact with several back-end resources and can involve complex, high-value transactions. In such applications, transaction servers like BEA's Tuxedo or M3 can improve the performance of complex transactions and can accommodate multi-phase commit. The additional transaction processing monitor improves performance and reliability.

- *Multi-Tier Architectures*

  While each of the above application architectures could be deployed on a single, monolithic server, there are significant advantages to multi-tier models. Each layer of the architecture — Web server, application server, transaction server, and database server — provides specific server resources that can be independently tuned and scaled.

  While the resulting multi-tier application architectures may seem quite complex, a multi-tier environment offers superior scalability, system performance, and availability. Each tier can use the scalability approach best suited to the function: Web servers may be made horizontally redundant through the use of multiple servers, while database servers can be made highly available through the use of clusters. In addition, the multi-tier model naturally supports geographic distribution and multiple-site operations.

Dynamic Hosting Performance

The mechanism used to provide dynamic content has a significant impact on application performance:

- *CGI Programming*

  CGI is the most basic dynamic content mechanism, offering rapid and easy-to-code functionality. The use of scripting languages, however, tends to result in poor performance. The CGI mechanism spawns additional processes which are constrained by basic operating system resources. A Web server engine that waits on CGI functionality may deliver only one-eighth to one-tenth of the performance (in hits per second) of a similar server offering static pages.

- *Compiled code*

  One alternative for improved performance is the use of a Web server that allows compiling of the application code into the Web server base. For example, Apache allows the compiling of extensions into the Web server itself. This results in a two- to four-fold increase in performance over CGI scripts.

- *Server Extensions through Server APIs and Java Servlets*

  Perhaps the best combination of performance and commercial support is the use of Web server APIs like *NSAPI*, now supported by Sun ONE Web Server. The use of proprietary APIs results in code for dynamic content or other application functionality that can deliver performance up to eight times that of CGI — nearly that of static Web page serving. Another way to flexibly manage services is through the Java Servlet API, which is supported by most Web servers.

Dedicated vs. Shared Hosting

The decision on whether to offer static or dynamic content is accompanied by a decision regarding how to allocate hardware resources to each hosting customer. Application hosting has evolved to include both dedicated and shared-hardware resource hosting models.

Vanity site hosting is generally provided via a shared, static content model. A user authoring tool may be supported and pages can be uploaded using simple FTP or even e-mail. Light commercial hosting may include either static or

dynamically-generated content. This level of service is typically provided in a shared environment, although some ISPs may offer dedicated light hosting on small server platforms. Heavy commercial hosting generally includes dynamic page generation and uses a dedicated platform which provides both isolation of load impact and higher quality of service.

In both the shared and dedicated models, the ISP determines system hardware and software configurations. Customers typically can select components but cannot interpose their own configurations or products if the ISP is to manage the service. To facilitate provisioning and maintenance, pre-defined, standard configurations should be prepared and tested.

Dedicated Hosting

With low-cost servers in common use for Web hosting, and the ability to separate one customer from another so effective, the trend for commercial Web hosting is towards the dedicated hosting model. Dedicated hosting provides dedicated hardware for each customer application (Figure 7-1). Dedicated hosting automatically provides a mechanism for physically outsourcing Web sites and for ensuring independent quality-of-service levels to each customer. Any service interruptions — whether due to system failure, maintenance, Web site data or programming — usually affect only the individual customer.

Management of dedicated Web sites is straightforward, but the administration of a very large number of systems does result in inefficiencies, for example, the time it takes for upgrading the OS on multiple systems. This can be overcome by crafting software agents that can automatically apply software upgrades from a trusted source, or through provisioning modules of some sophisticated network management software products.

ISPs can simplify the management functions somewhat by delegating authority to customers for configuring and provisioning specific services. For example, Sun Cobalt servers, used in many ISPs and IDCs, can be initially configured by an administrator, with the remainder of its functions — including selection of server extensions and enabling FTP access — activated by the customer.

There are two variations on the dedicated hosting model. First, multiple sites can be hosted in separate, electrically isolated domains in Sun Fire™ Midframe and Sun Fire 15K servers, enabling ISPs to leverage the reliability and availability advantages of larger servers when offering hosting services to smaller customers. With dynamic reconfiguration, the CPU, memory, and I/O

resources of a Sun server can be changed as customer requirements change. A second, less-frequently used approach is to have multiple dedicated Web server applications running on the same server, providing each dedicated customer with an independent software environment that is not shared with other customers. ISPs can provide various levels of service when deploying this model by using resource management tools like Solaris Resource Manager software (See “Resource Controls” on page 119.)

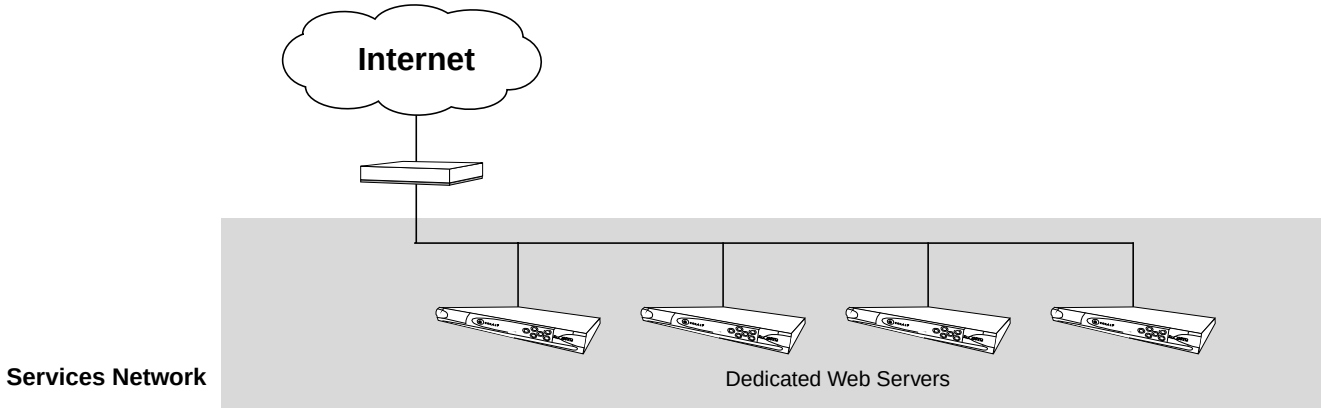

Figure 7-1 Dedicated hosting provides specific, independent hardware resources for each Web site that is hosted. Illustrated are a set of dedicated Sun Cobalt RaQ XTR servers.

Advantages of Dedicated Hosting

- *Ease of administering individual sites.* Each customer is an independent unit with its own configuration. No additional management framework is required beyond normal system administration and customer provisioning.
- *No cross-contamination between customers.* All event, configuration, and performance impacts are confined to individual customers. Compromises in security, the overloading of an application, or application failure only impacts a single site.
- *Isolated quality of service.* Dedicated hardware enables specific quality-of-service levels to be offered to individual customers according to a pre-set fee structure.

Disadvantages of Dedicated Hosting

- *Cost of dedicated hardware.* The most significant disadvantages to dedicated hosting are both the long-term cost of ownership in terms of management costs, and the initial cost of hosting each customer using separate server and disk resources.
- *Administration overhead of large server farms.* While the administration of dedicated hosting is straightforward and does not require a specialized framework, it also becomes more difficult to perform basic tasks (e.g., software upgrades) over a large number of servers. ISPs that take on the task of integrating provisioning capabilities into their management systems have a competitive edge over those which don't.
- *No flexible allocation of resources.* Dedicated hosting results in a lack of flexibility in re-allocating resources between customers. For example, in addition to the core servers for a customer, any redundant system components must also be allocated to each customer. This means sophisticated additional services may not be offered because they are not economical for individual customers.

Shared Hosting

Many early hosting demands were for static commercial sites, with early Web servers used for this purpose offering little in the way of virtual Web site hosting capabilities. Today, individual users require Web sites hosted with their own domain names, and most Web servers have added the capability to provide many virtual sites from one server. Shared hosting provides an improved outsourcing model with additional economies of scale, lower cost of entry, and potential interoperability. Larger, more cost-effective servers and disk arrays can also be utilized (Figure 7-2).

Advantages of Shared Hosting

- *Lower cost.* The greatest advantage of shared hosting is the resulting lower cost of supporting individual customers.
- *Improved use of resources.* Inherent in the shared hosting model is the ability to flexibly allocate hardware and other system resources across multiple customers. Adding more disk to a customer site is as easy as changing a configuration setting for a pooled disk resource, rather than the physical installation of additional hardware.

- *More advanced services.* In addition to lower cost for basic core hosting services, the economies of shared hosting may make it possible to offer advanced services. This includes both point technologies, such as electronic commerce transactions or disk caching, as well as the ability to offer superior, more functional application architectures.

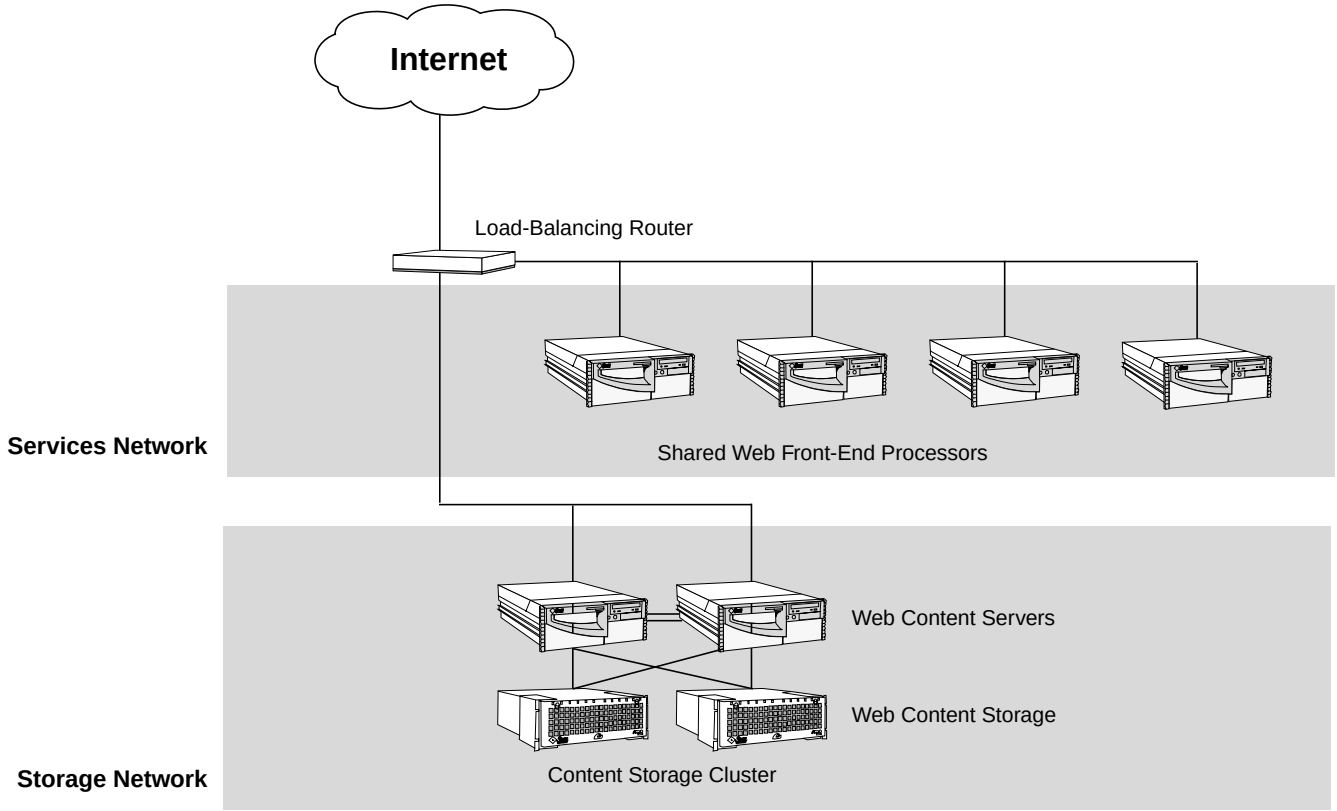

Figure 7-2 Shared hosting environments can exploit the economies of scale and reduced cost of ownership of larger servers and disk arrays.

Disadvantages of Shared Hosting

- *Requires special administrative framework.* In general, both server hardware and application software are managed for individual customer utilization — not for shared hosting. For example, shared Web hosting requires the ability to run multiple virtual Web servers answering to different names all on one system.
- *Possible cross-contamination.* While dedicated hosting means that no customer's problems will affect others, shared hosting means that a security compromise or performance problem affecting one customer has the potential to do so.

Management Issues

Management of Web hosting facilities is an important consideration for ISPs. Issues include provisioning, content development, and facilities for meeting quality-of-service guarantees.

Provisioning

Provisioning includes all of the activities required to create and configure a hosted customer and their environment. Dedicated hosting environments often involve significant intervention by systems support staff. Shared hosting environments often can be automatically configured based on information provided to sales personnel, or even by users accessing a Web-based interface to purchase and provision Web hosting services.

In addition to the configuration of hardware, provisioning includes the assignment of appropriate identity and naming conventions (for example DNS entries and user profiles), and the configuration of system services (billing, for example).

Manual Provisioning

Web hosting is often provisioned manually by the ISP after the customer has selected the desired level of service and functionality. This involves configuring any dedicated system resources — including the server — or simply configuring file space on a shared server for a shared hosting customer. This process may be augmented by an on-line service request form. Personal

home page hosting may be implemented automatically, while more significant commercial Web site hosting may require the answers to many questions asked by a sales agent, and the involvement of administrators and programmers.

Automated Provisioning

While basic provisioning may be accomplished using manual procedures and independent service applications, service providers focusing on Web hosting as a source of revenue typically implement automated provisioning systems to streamline the process:

- Automated provisioning improves the speed at which the ISP may implement services and reduces or eliminates configuration errors.
- Use of automated provisioning provides an audit trail of provisioning and administration activities.
- The delegation of certain customer care activities to the customer reduces the platform operator's workload and enables the customer to make and review service changes more easily.

In the past, automatic provisioning was accomplished with custom scripts developed by ISPs. Today, some sophisticated network management systems facilitate this process by providing an infrastructure that integrates network management, customer relationship management, and billing software. This allows ISPs to create modules to execute specific configuration instructions and lets the rest of the ISP management infrastructure support the Web hosting service.

For ISPs focusing their business on hosting a large number of sites, there are a growing number of commercial software packages that support automated provisioning of ISP servers and self-provisioning for options within the customer's feature set. The architecture for managing such an operation is usually based on an administration server that hosts a Web site for customers and administrators to choose and configure services, along with configuration and accounting databases (Figure 7-3). When sites are to be configured, or their parameters changed (disk space, for example), the administration server communicates with agents on the ISP servers to accomplish whatever actions are necessary to provision or modify the service. With central administration, actions can be coordinated, for example registering a domain name, setting up entries in the DNS server, establishing e-mail hosting, and setting up a virtual site on a shared Web server. The administration server communicates to the

agents over the ISP's management network using distributed communication mechanisms like CORBA (Common Object Request Broker Architecture). Communication with registration authorities may be through automatically-generated e-mail messages with administrator action required when a response is returned, or it may use APIs provided by the registration authority.

Automated provisioning may be used in conjunction with self-provisioning, where customers provision and configure sites with no contact with sales staff or intervention by administrators.

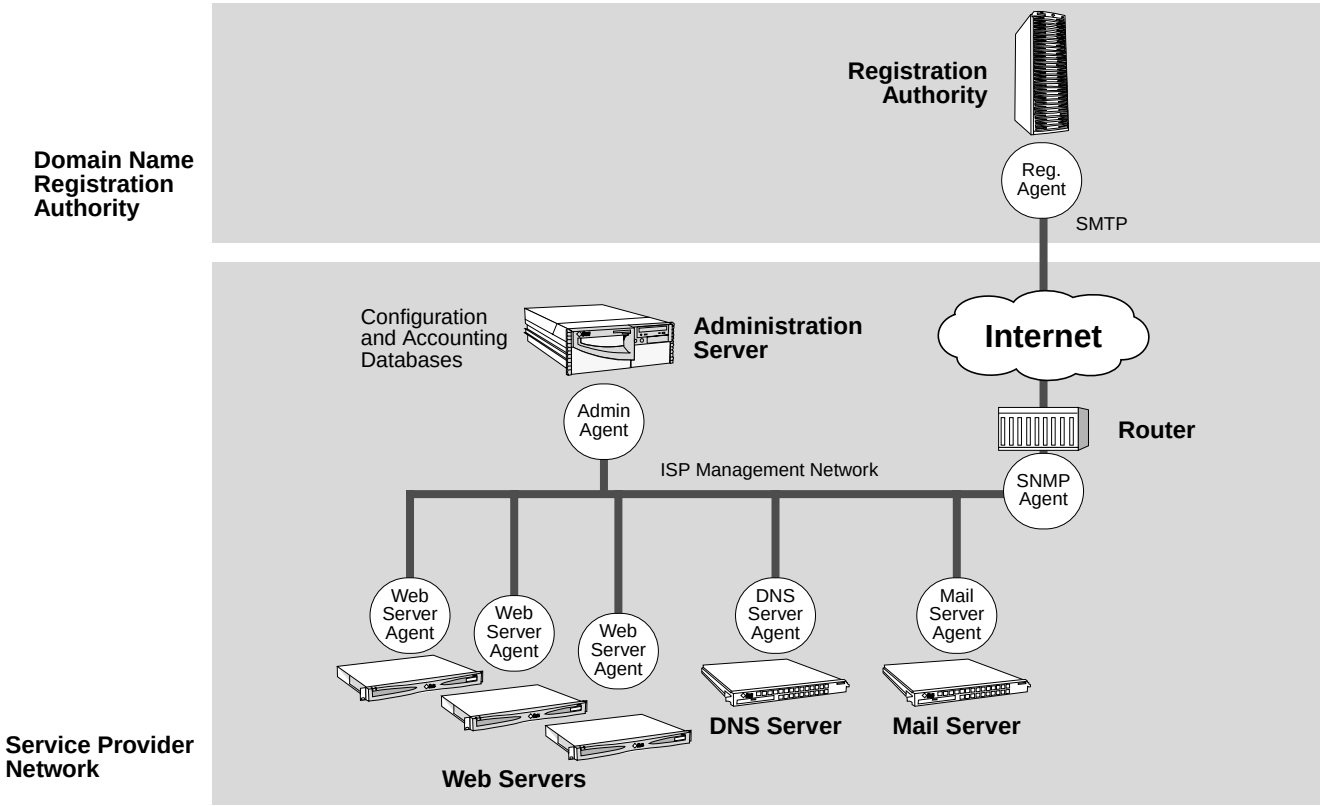

Figure 7-3 The architecture of products that enable automated and self-provisioning of sites is based on a central administration server and agents that make changes that are local to the ISP servers.

Self Provisioning

Self provisioning delegates specific provisioning capabilities not just to the ISP staff, but to customers as well. Generally, each delegated administrative authority has the ability to perform specific functions by means of a Web-based interface. For example, customers may change passwords without involving a sales agent or a system administrator. When used in conjunction with automatic provisioning even the use of system administrator time is avoided for the creation of new sites.

Some ISPs focused on the Web hosting market use automatic and self-provisioning systems that take customers all the way through the sales cycle beginning with a compelling Web site offering a variety of attractive service packages, and ending with the collection of billing information and handing off delegated administrative authority to the customer (Figure 7-4). The process from the users' perspective is roughly as follows:

- *Greeting Page*

  The ISP's Web sales site greets the customer with compelling content or design that encourages further exploration of the site.

- *Service Options*

  From the home page, customers can explore pages that describe the benefits of each of the service plans offered. Well-planned sites allow customers to view service options before and after they select a domain name.

- *Domain Name Selection*

  The crucial step in signing up a new customer is finding a suitable domain name, or taking information to transfer an existing domain name to the new hosting service. When a domain name requested by a customer is found already to be taken, linguistic matching tools can generate alternative names and suggest them to the customer.

- *Plan Confirmation*

  With an appropriate domain name and service plan selected, the customer often gets the chance to confirm choices, and also to choose configuration details such as how the site will be populated, for example through FTP or Microsoft FrontPage extensions.

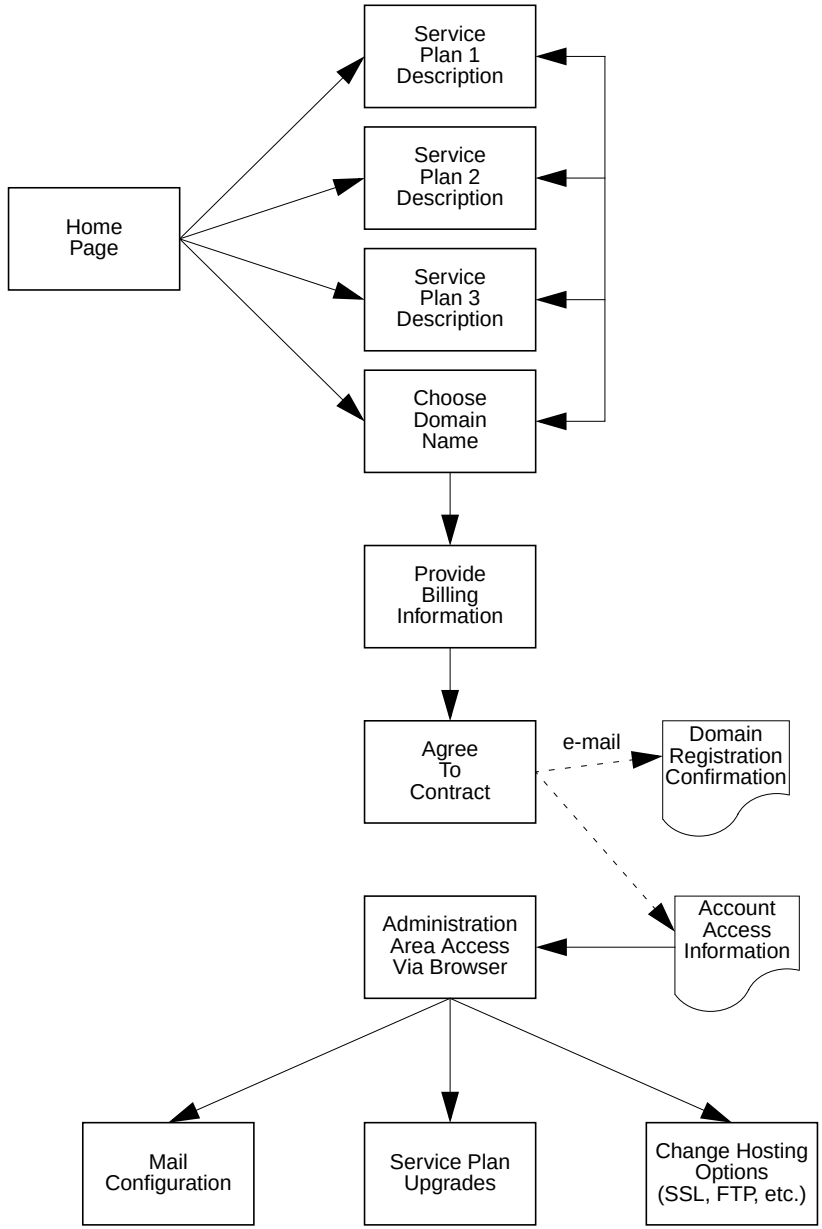

Figure 7-4 Customers proceed through a fairly standard set of steps in the Web site self-provisioning process.

- *Billing Information*

  Using an SSL-encrypted Web page, the customer now makes a commitment to the service plan and domain name by providing billing information. In the U.S., the authorization of a recurring credit card charge is usually required. Another alternative is to allow customers to provide a billing address. Regardless of the means of billing, the hosting facility must verify that the customer can and will pay for the service — either by making a credit card charge or receiving payment from the customer — before making the site available.

  Concurrent with the transfer of billing information is the acceptance of a service contract that legally binds the two parties.

- *Asynchronous Activities*

  At this point there is some delay while the credit card charge is processed, and a domain name registration form is sent to the appropriate name registry. Two independent e-mail confirmations might be sent at this time:

  - If the automated Web site hosting software registers the domain name with the registration authority via e-mail (rather than through Internet-based APIs), the customer is often given a token number for tracking the request, and the token is tracked by the ISP's customer relationship management system
  - Customers might receive an e-mail confirmation that gives an IP address of the site (pending domain name registration), along with user names and passwords for populating the site with content.

- *Access to Administration Area*

  Once the billing is complete, and the domain name request is initiated, the Web site is provisioned and the customer is allowed access to a delegated administration area. This page provides customers with control over their Web sites, including options such as:

  - Configuring mailboxes and aliases
  - Changing plan options and upgrading from one service level to another
  - Setting hosting options; choosing how the site will be populated, and establishing SSL certificates for merchant sites
  - Viewing log information and traffic summaries

Once up and running, customers can populate their sites with their own content and change it as often as they wish, giving them superior service and complete control over the services they receive.

Content Development

The type of content development support offered by ISPs usually depends on the nature of the hosting supported. Heavy commercial customers requiring dedicated hosting of dynamic content will most often develop their own content or work with an outside design firm. Light commercial users are frequently attracted to content authoring support if it is offered. Vanity site users do not have a sophisticated content creation capability and are often offered content creation support including packages like Allaire's HomeSite. For these users, ISPs may support server-side extensions like those for Microsoft FrontPage, however such extensions have support, security, and performance implications that ISPs must consider.

Staging

Staging involves the placement of Web content into a staging area prior to it actually being placed on-line. This may involve the testing or analysis of the content. Staging facilities can be provided to customers at additional cost.

Log Analysis and Viewing

In addition to content creation tools, the ISP may offer log analysis capability to its commercial customers, or to all customers using a self-provisioning system. Users may be presented with simple bandwidth usage data provided by tools like the open source MRTG (Multi-Router Traffic Grapher) or complex log analysis performed by commercial packages like those from WebTrends.

Quality of Service

Quality of service is a critical differentiator between Web hosting services. Consideration must be given to controlling the quality of service offered to Web hosting clients, monitoring service levels, and managing resources.

Quality-of-Service Controls

ISPs may offer quality-of-service controls as a value-added service for which there may be additional service charges. The ISP's management systems must include mechanisms to ensure that customers get the performance they purchased. Before the platform operator may offer quality-of-service guarantees to customers, two facilities must exist:

- Controls that limit the consumption of systems and network resources
- A measurement system that documents the utilization of these resources

The consumption of system resources can be controlled in a number of ways: First, by utilizing a multiple-tier architecture, and by partitioning services according to their relative impact on the underlying resources (i.e., resources are built to isolate load impact). Second, by applying system and network resource management within various systems (i.e., monitoring is used to control load impacts). These methods are discussed in the following sections.

Another critical component of quality of service is reporting to the customer and to the platform operator. Metrics and statistics should be delivered to the customer to demonstrate compliance with contracted quality-of-service levels. Management reporting enables the customer to determine which quality of service best suits the application while allowing the platform operator to tailor services to the customer's needs. For example, a customer who is consistently reaching their service limit should be offered an increased level of service and be charged accordingly.

Resource Controls

Delivering specific service levels is achieved by partitioning services, carefully managing server resources, and allocating specific levels of bandwidth to customers.

- *Service Partitioning*

  In service partitioning, services are separated according to their functional boundaries, like Web server, application server, and database server, and then arranged in a multi-tier configuration on the available hardware. For example, there are several different types of application execution platforms including CGI, JavaScript, Java Servlet, and application server. It may be necessary to separate these platforms for reasons of performance, security, and feature set. It is likely, for example, that the CGI execution platform will

be separated from other application servers since this type of Web application presents specific concerns with respect to security. Therefore, CGI programs are sometimes executed on their own server or a cluster of servers dedicated to this purpose.

- *Resource Management*

  Controlling resources within a shared server is necessary in order to guarantee quality of service. Solaris Resource Manager software provides the ability to control and allocate CPU time, processes, virtual memory, connect time, and logins. This control can be applied on a fine-grained, hierarchical basis — making it possible to define a resource pool for each customer Web site with each pool allocated a pre-defined portion of the available system resources. Solaris Resource Manager software helps ensure, for example, that applications running on behalf of one customer Web site do not consume more than the pre-determined amount of CPU and memory. This mechanism allows applications belonging to other Web sites to run without being starved of resources.

  Another capability of Solaris Resource Manager software is to define hierarchical relationships between resource pools. Resources may be assigned to a *parent* pool that contains a number of subordinate pools, or *children*. The parent is assigned resources as usual. The child pools are assigned resources *from* the parent pool. This technique is useful for creating classes of service. The parent represents the overall class of service and the children represent the individual customers that belong to the class.

  Solaris Resource Manager software helps ensure a fair sharing of the available system resources over hosted *applications* as well as Web services. With this system component, ISPs can control costs and server resource consumption using methods similar to those found on costly mainframes. Multiple applications, groups, and individuals users can be guaranteed a consistent level of service on dedicated server. By dynamically allocating unused CPU capacity and virtual memory, resource utilization is increased. Systems are easier to manage because of the ability to set and enforce policies that control how system resources are utilized, ensuring that customers always receive the assigned service level within a shared resource environment.

Another method used to provide quality service is to deploy a network caching infrastructure. Many ISPs are placing caches in their distributed Web hosting centers to provide both improved response time and as a way to level out peaks in user demand cycles.

- *Network Bandwidth Allocation*

  Quality of service starts with a reliable, scalable, high-performance network operating environment that can quickly and safely deliver required services, and enable customers to pro-actively manage their ISP-provided bandwidth and service-level agreements. System resource allocation provided by Solaris Resource Manager does not strictly control the use of network resources; however without some control over the consumption of bandwidth, an application may consume excessive network resources.

  Solaris 9 Resource Manager's bandwidth management functions are specifically designed to allocate network bandwidth, helping to ensure a fair sharing of the network bandwidth available to each Web server. This provides ISPs with the means to uphold service-level agreements set with customers.

  Bandwidth allocation is most critical in the Web server front-end, since the bulk of network resources are consumed by the delivery of static content. Users typically consume static content at a greater rate than they consume content resulting from the execution of applications. The Web server front-end is far more efficient at delivering content than applications in the middle tier and is therefore the chief consumer of Internet and intranet bandwidth. The amount of network bandwidth consumed by each customer's virtual Web server can be limited by Solaris 9 Resource Manager software to implement fair sharing

Monitoring

To gather sufficient information to check quality-of-service levels, basic system monitoring and monitoring of the Web services is required:

- *System and Network Monitoring*

  The ISP's network management system, for example HP Open View or CA Unicenter, can usually provide basic monitoring functions like reporting if the server is up and running and what bandwidth is being used. Many

ISPs also monitor Web sites using Multi-Router Traffic Grapher (MRTG), making bandwidth utilization graphs available to customers. These tools all use SNMP and basic monitoring capabilities like ICMP echo requests.

- *Service Monitoring*

  Server hardware may continue to operate, without reporting alert conditions even though an application has hung or failed, making it necessary to monitor Web-based services at the application level. Some commercial packages like Freshwater Software SiteScope and Micromuse Netcool perform end-to-end service monitoring, reporting on responsiveness and functionality of complete customer scenarios including the actual purchase of products and credit-card payment on e-commerce sites. Servers performing these service probes can be located at the ISP's facilities.

  External monitoring that includes Internet latencies evaluates the quality of the Internet connectivity to the ISP site and server responsiveness. Servers designed to gather this information can be located outside the ISP's facilities at an IDC. Tools like SiteScope and Agilent's Firehunter can provide SNMP-based information so that their findings can integrate with the ISP's network management systems. Performance monitoring services are on the rise, with companies like Gomez and Keynote evaluating site performance relative to other popular sites.

Security Issues

With attacks that sometimes leave Web sites completely disabled or defaced with content placed by hackers, security is a paramount concern of ISPs providing Web hosting services.

Security Controls

Because it involves customer access to content servers, the ISP's Web hosting environment should be part of a comprehensive security strategy such as that discussed in chapter 5. This strategy includes both the implementation of security features including firewalls, as well as a security policy defining procedural characteristics such as specific responses to security events.

The implementation of the security policy is facilitated over the ISP infrastructure in general, and the Web-hosting environment in particular, by dividing the systems architecture into *security zones* (Figure 7-5). A security zone is a collection of system, content, and processes that require the same, or similar, level of protection. Security zones can be used to compartmentalize access to systems by users authorized to update content.

Security zones can be separated by firewalls that filter traffic entering and exiting the platform. Only authorized protocols and source and destination addresses are accepted. All other attempted connections are dropped with the firewall software generating alerts to administrators.

Although the Internet is generally considered the least secure layer in the security model, and the legacy systems the most secure, the model becomes divided when legacy systems represent services such as stock ticker feeds, which are susceptible to outside intrusion. In this case, the Web hosting infrastructure needs to be protected from the legacy systems, with the back-end firewall and optional fire-ridge providing the needed protection.

In addition to the careful deployment of firewalls, a number of additional precautions should be taken, including:

- *Server Hardening*

  Each server should be pruned of any unnecessary programs, services or configuration options. In addition, all security-related operating system and application patches should be installed. Sun's Solaris™ Security Toolkit provides tools for automatic configuration of hardened servers.

- *Cryptographic File System Integrity Testing*

  Automated processes can be used to test the integrity of files and institute required permissions. Several times a day a cryptographic (MD5) signature is computed for each critical system or application file. These signatures are compared to a set of "known good" signatures stored in a read-only file system. Software like Tripwire can be used to accomplish this.

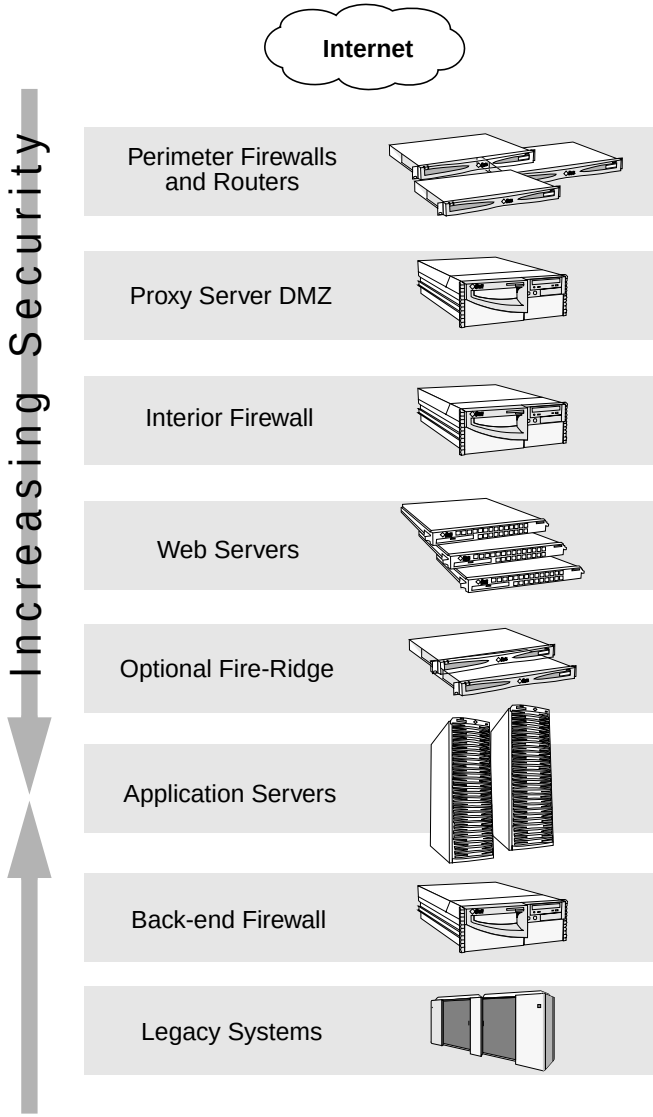

Figure 7-5 Security zones are separated by firewalls, with careful consideration given to the accesses allowed from one layer to the next.

- *Event Logging*

  System and application event messages should be forwarded to a centralized log facility.

- *Threat Detection and Risk Mitigation*

  Several open source programs can test systems periodically for known security weaknesses. Such programs should be run on a frequent basis.

- *Penetration Testing*

  Security policies and their implementation are most effective when augmented by periodic penetration testing by an outside service.

Secure User Authentication

When Web applications require authentication and encryption, centralized LDAP directory servers can be used to authorize users and applications. This directory service can allow for traditional password challenges as well as verification by way of digital certificates.

In order to support authentication via X.509 certificates, the client browser and/or the server must obtain a digital certificate from a certificate authority (CA). There are various commercial and non-commercial organizations that generate and provide custody for digital certificates. In other cases, the ISP may wish to establish their own certificate authority in order to maximize control over the issuance, custody, and revocation of certificates.

Data Encryption Using SSL

Encrypted Web site sessions using the Secure Socket Layer (SSL) are possible, with client-side and server-side X.509 certificate-based authentication as necessary. Secure communication consists of three basic components: authentication, data concealment, and integrity. SSL offers authentication based on a secure exchange of digital certificates and session encryption.

Users selecting a secure page or site are directed to a Web server that supports SSL. Since the encryption process associated with supporting an SSL connection can be a large drain on the Web server in terms of resources, it may be wise to dedicate separate server systems to handle SSL sites. These front-end servers may be outfitted with additional CPUs or even hardware-based encryption accelerators.

Summary

Many ISPs now provide Web-hosting capabilities to casual, light commercial, as well as heavy commercial users. Web hosting is a specific case of general application hosting, and it requires ISPs to consider whether to provide dedicated hosting, shared hosting, or both. Management facilities — ranging from site management to automatic provisioning — must be provided to ensure effective administration of services. Monitoring services are necessary to monitor service levels and alert staff to any anomalies. Finally, security against attack on the application services and the user data itself must be in place.

The Role of Internet Data Centers 8

Internet data centers are computing centers where many organizations can share a single, comprehensive infrastructure to economically host Internet servers and services. They tend to be deployed at multiple, geographically-dispersed sites connected with high-speed, wide-area networks, hosting a heterogeneous array of servers. IDCs specialize in around-the-clock operation, comprehensive management services and systems, and skilled system administration staff.

In many ways, IDCs provide exactly the infrastructure that today's ISPs need in order to capture increasing amounts of the market for outsourced business services. In a world where characteristics such as performance, reliability, availability, scalability, security, and manageability are key, ISPs are either using IDCs to help meet their increasingly stringent service-level agreements, or emulating IDCs so that they use the same techniques to build world-class facilities of their own. This chapter discusses the symbiotic relationship between ISPs and IDCs, the compelling reasons for ISPs to utilize IDCs, along with an overview of IDC facilities and network architectures.

A Symbiotic Relationship

The existence of ISPs pre-dates by years the existence of IDCs, however the two have a considerably intertwined relationship. Recognizing the need for customers to locate their own servers closer to the Internet backbone, some ISPs have been offering co-location services to their customers for some time. Some national and global ISPs essentially have become their own IDCs, providing for themselves the infrastructure that IDCs provide for their

customers. For the vast number of local and regional ISPs, IDCs offer a way to reach into the rapidly-increasing outsourced services market, allowing them to capture a larger market and more easily meet service-level agreements in select areas where IDCs offer the greatest leverage.

The many facets of this symbiotic relationship include:

- *ISPs Emulating IDCs*

  The larger the ISP, the more likely it is that they emulate IDCs in terms of how they handle their physical and network infrastructure. Taking lessons from IDCs, these ISPs have closely-integrated customer relationship management (CRM) software, network management facilities, problem-tracking tools, and billing systems. At minimum, the largest ISPs run their operations as IDCs do because even brief outages affecting millions of customers quickly makes the evening news. Far more than just avoiding bad publicity, however, the largest ISPs must be structured in the same way as IDCs simply to provide the world-class quality of service necessary to win and maintain customers.

- *ISPs Becoming IDCs*

  Historically, ISPs were the first organizations to offer Web hosting services to their customers. Beginning with the free vanity sites provided to their subscribers, they began offering a higher level of service to business customers, including virtual domain hosting on shared servers, as well as dedicated hosting. For some customers wishing to buy and maintain their own Web servers, some IDCs also offered co-location services. Some of the ISPs most successful at selling dedicated, managed Web services and co-location have become IDCs, winning customers through their superior service and proximity to the Internet backbone.

- *ISPs Locating Services in IDCs*

  For smaller local and regional ISPs wishing to capture more of the outsourced business market, locating premium Web customers in an IDC is a strategic move that upwardly expands the ISP's range of service offerings. Rather than lose their profitable business customers as they grow in size, or risking capital to implement the global presence that an IDC can provide, the ISP can keep their growing business customers by managing their Web services in an IDC and handling other services — like hosted e-mail — in their own data centers.

- *ISPs Located in IDCs*

  Some of the newer ISPs have simply located their servers in IDCs for the same reasons that other businesses do: rapid time-to-market, reduced risk, increased performance due to proximity to the backbone, economical infrastructure, and 24-by-7 support.

Compelling Reasons for Using IDCs

The overriding reason why ISPs strategically use IDCs to support some — or all — of their services is the flexibility to offer higher service levels to their business customers. In today's environment, where outsourcing is viewed as the way to contain costs while improving service, ISPs can support their business customers through the high performance, scalability, reliability, availability, and manageability offered by IDCs. Some of the many reasons why ISPs use IDCs to support their services include:

- *Quality of Service*

  With the physical infrastructure to support higher service levels than their customers or their own facilities provide, ISPs have an even greater competitive edge through the strategic use of IDCs. For example, most ISP customers' enterprise data centers maintain sufficient backup power to support a graceful shutdown of their servers. For those companies outsourcing their e-mail and Web services to ISPs that utilize IDCs, battery back-up and rapid-start generator power can keep servers running indefinitely. So companies win by outsourcing to ISPs, and ISPs win by locating services within IDCs.

- *Time to Market*

  For many ISP customers, reducing time to market is key to success. With an Internet data centers' existing physical plant, redundant and back-up power, and high-speed Internet connectivity, ISPs can bring their customers' products to market much more quickly than the traditional hosting model. And with traffic to sites often doubling every six months, the time needed to scale services is minimized with IDCs because they specialize in making needed space and computing facilities available quickly.

- *Redundant, High-Speed Global Connectivity*

  By combining the purchasing power of many tenants, IDCs can purchase high-bandwidth network services and deliver Internet connectivity from multiple carriers simultaneously. With high traffic volume, IDCs have the clout to make strategic peering arrangements with other facilities and providers, resulting in topologically-superior connectivity that results in fewer hops — and therefore less latency — between large numbers of end users and the IDC. These peering relationships are so strategic that they are often closely-guarded business secrets.

- *Economical Infrastructure*

  The physical infrastructure necessary to support high-availability services is much more economical with larger facilities. High-security data centers; redundant, uninterruptible power supplies; multiple air conditioning systems; fire suppression; and backup generators with days or weeks worth of fuel are more costly than a local or regional ISP can afford — yet these are requirements for providing the continuous availability that business customers now expect.

- *Comprehensive Support*

  Just as the connectivity and physical infrastructure of IDCs is shared, so is the support staff. In today's competitive labor market, it is difficult to hire, train, and retain skilled staff members for maintaining Internet services 24 hours a day — and when ISPs use IDCs to support their most critical services, their own staff is unburdened from the around-the-clock attention that providing those services can entail.

- *Lower Cost*

  Even with the benefits of less time-to-market, improved connectivity, world-class infrastructure and support, using IDCs can result in significantly lower cost. Because the large infrastructure expenses of high-end routing equipment and uninterruptible power are amortized over a large number of customers, ISPs utilizing IDCs can benefit from the quality of service that these facilities provide without having to make the huge capital investments themselves.

Range of Services Offered

There is a wide range of services available from Internet data centers, and the ways in which they differentiate themselves can be characterized in three dimensions:

- *Quality of Service*

  IDCs can provide quality of service that can be quantified on a range from *data grade* availability of 99 percent to *carrier grade* availability of 99.999 percent. Considering that a 99 percent available facility means up to 87 hours — more than three days — of down time per year, the use of highest-quality facilities can have an enormous attraction for customers. Zona Research estimated a USD $4.35 billion loss due to Web site outages and slow response times for the year 1999, raising the importance of this issue. In order support ISP customers who require guaranteed uptime, most IDCs today offer Service Level Agreements (SLAs) that specify measurable service goals and penalties if they are not met.

- *Hosting and Co-Location*

  IDCs offer pre-packaged and custom Web hosting using dedicated servers, shared servers, and even multi-tier e-commerce implementations. They also provide Web-based application hosting, customer server co-location, and standard IT services such as e-mail outsourcing. The use of dedicated servers is rising in popularity not only because of improving economics — there are fewer security issues with dedicated servers. Today, hosting and other managed services make up the majority of an IDC's business, with co-location services lagging significantly behind.

  Many IDCs provide a range of options for server ownership, which usually has an influence on the support services that can be provided. IDCs that do not own the servers in their facilities provide what is known as *co-location* services; those which own the servers and lease them directly or indirectly to their ISP customers provide *hosting* services. Internet data centers generally provide both co-location and hosting. The IDCs that provide sophisticated management services often do so only for the systems that they own and maintain because this combination gives them the most control and enhances their ability to support strict service-level agreements.

- *Storage Services*

  Just as power, cooling, and Internet connectivity are 'utilities' that IDCs provide, storage is a service that can help ISPs more easily leverage IDCs. The more forward-looking IDCs provide an imaginative array of storage services:

  - Storage-on-Demand services can be provided through a number of mechanisms, including dedicated, direct-attached storage, network-attached file system storage, and Storage Area Network (SAN)-attached storage using dedicated services securely allocated from a common pool.
  - Backup services are deployed using shared tape library facilities, backing up user data over a separate dedicated network, with off-site copies maintained by the IDC. Backup services are such a necessity that many IDCs offer this service even when they don't offer disk storage services.
  - Database services enable ISPs to outsource the implementation and maintenance of back-end database software associated with high-end Web customers When supplying database services, the IDC's monitoring facilities provide direct benefits to ISPs, but also provide upselling opportunities to the IDC as additional customer storage space is required.

- *Support Services*

  IDCs vary in the set of support choices they provide, ranging from almost no support to full-service application-level monitoring and management. IDCs that provide simple co-location services are often referred to as "rack, stack, and ping" facilities, where servers are hosted and only their ability to respond to an ICMP echo request is monitored. In the middle of the spectrum are facilities that closely monitor bandwidth usage and response times, and sometimes track the user experience by loading specific URLs and even performing transactions. These IDCs automatically create trouble tickets and respond quickly when server performance varies from pre-determined limits. At the high end are facilities that provide full-service monitoring capabilities that go beyond measuring parameters that are visible by end users; these facilities monitor and manage the entire customer infrastructure including database systems management systems and credit card clearing functions.

- *Content Distribution*

  ISP services outsourced to IDCs can often benefit from content distribution services such as those provided by Digital Island. With caching servers hosting Web content at strategic points in their global network, end users

access the content that is topologically closest to them — resulting in faster response times from anywhere in the globe. Key to Digital Island's strategy is its DNS server technology which intelligently resolve users' address requests to provide the address of the nearest caching server. Digital Island's services enable its customers to easily extend their reach around the globe, giving end users high performance as if they were accessing a nearby site — regardless of location.

Physical Infrastructure

There are several major components to an Internet data center's physical infrastructure, including the building, security, environmental protection, power, and the physical distribution of data centers. All IDCs consider each of these factors carefully; what weight they give to each area determines to some extent the customer base they will attract. Some IDCs have additional space available for a fast recovery from major disasters; all IDCs maintain off-site backups for recovery of customer data.

A typical Internet data center incorporates design considerations including building, security and access control, electric power, and protection from environmental problems including fire, flood, and wind storms.

Building

The buildings in which Internet data centers are located are usually designed or chosen for their ability to provide a high level of security from intruders, vandals, terrorists, and from environmental concerns such as floods, earthquakes, and tornadoes. The amount of protection built into these buildings varies widely. Many have strategically-placed barriers to prevent damage from car bombs. Some have Kevlar-reinforced walls. Some are located on upper floors to prevent access at ground level and to offer protection from flooding. One hosting facility in Canada uses a retired Royal Bank vault with a double underground perimeter that prevents underground tunneling into the facility. Some IDCs that focus on co-location are situated in prime areas for easy access by customers, while those which focus on hosting place less emphasis on location.

Inside, Internet data centers typically use raised floors and overhead cable ducts; in the well-organized Globix facility, all copper wire is routed beneath the raised floor, and all fiber feeds are delivered from above.

Providing customers security from each other is a huge concern. Individually-locked racks allow access by only the customer whose equipment is located in the rack. Larger spaces are controlled by wire cages, and some IDC customers even demand completely-enclosed cabinets with Faraday cage protection to eliminate electronic snooping. These customers sometimes require cabinets to extend through a suspended ceiling to the concrete floor above, with video surveillance detecting any entrance into the cavity beneath the raised floor. Indeed, facilities such as this never route cables such that any other customers' data would pass above or below another customer's rack, cage, or cabinet.

Security and Access Control

All aspects of physical security are considered in the design of Internet data centers. A staff of security guards is on duty around the clock, and surveillance cameras observe and record all activity inside and outside of the data center — including underneath the raised floor.

Customer access is carefully controlled, and every entrance and exit by any customer, employee, or visitor is permanently on record. Physical access is usually granted only after the visitor's ID is checked and a magnetic card swiped; further authentication is often performed, including matching biometric factors such as hand and retinal scans. To prevent 'tailgating,' some IDCs employ a "man trap" that allows only one person at a time to enter through an air-lock style passageway. Once inside, the customer has keys that only unlock their specific rack, cage, or cabinets — in fact, some IDCs even have no master key to prevent any unauthorized access even by security personnel.

Security in the back-end of the data center is equally important. Because of the premium for space inside the data center, staging areas are often provided for customers to assemble and test their equipment before installing it into racks. Access to these areas, and areas leading to the loading docks is also tightly controlled.

Electric Power

Internet data centers provide conditioned power to their customers' equipment, with both battery and generator backup. Battery backup enables the data center to continue operation during the brief period between a power failure and the automatic spin-up of diesel generators. In order to ensure the

rapid availability of diesel-generated power, hot water can be circulated through the generators so that they are always ready for a warm start. Some IDCs have sufficient fuel storage for a day or more; some are prepared to operate for weeks on backup power.

Redundant power is an important feature of Internet data centers. Some IDCs go so far as to obtain independent power feeds from different sub-stations. Most have multiple generators, and feed power from two different sources to each rack. Customers using equipment with dual power supplies — like many Sun Fire servers — can connect to both power sources and operate uninterrupted through the loss of one of the two feeds.

Power Supply Concerns

The IDC market is moving from co-location towards hosting and increasing levels of managed services, giving them more control over the servers and therefore the power consumption in their facilities. With electric power supply becoming less of a certainty in areas like California, and with huge numbers of data centers stressing power grids in places like Dublin and London, IDCs have both a financial and a political incentive to manage their power consumption carefully. Low power consumption — significantly less than its closest competition — is another reason why many ISPs choose Sun servers for their IDC facilities (Figure 8-1).

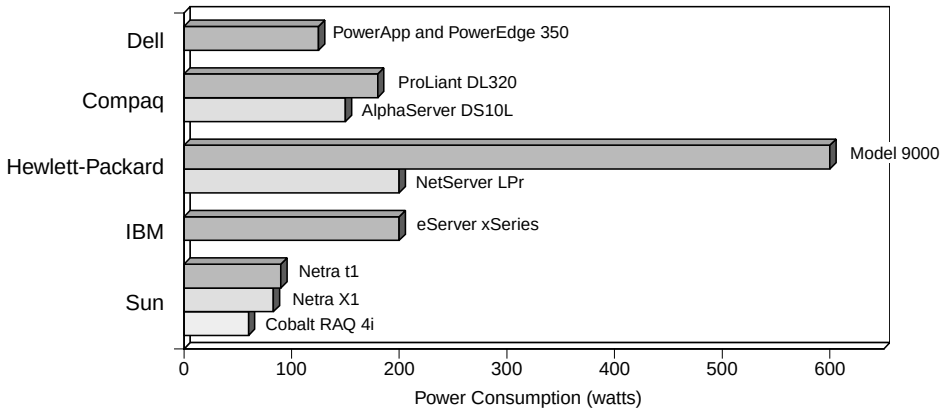

Figure 8-1 Power consumption for Sun's 1 U severs is significantly lower than its competitors 1- and 2-U models. Source: vendor Web sites

Environmental Protection

Every watt used to power IDC servers is ultimately dissipated as heat which must be removed by IDC air conditioning units. Redundant air conditioning capacity is installed so that mechanical failures can occur without an unacceptable rise in room temperatures. Systems using chilled water or freon often have separate external cooling or condenser systems, helping to ensure that there are at least two complete systems with no single point of failure.

Fire suppression systems in U.S. Internet data centers are usually based on FM-200 chemistry; halon is no longer used because of its adverse environmental impact. These systems are always redundant, and the use of racks with permeable doors and cages with wire walls enables rapid access to the source of fire without the delays of obtaining rack or cage keys.

Internet data centers are built with consideration to the potential for natural disasters such as earthquake, flood, and wind storms:

- In earthquake country, IDCs are often built on thick concrete slabs that are designed to move on a bed of gravel during an earthquake; these centers also provide seismically-braced racks and expect their customers to maintain their equipment so that it cannot be ejected from its rack or shelf.
- In areas where flood is even a remote possibility, data centers and all support equipment are located on upper floors.
- Reinforced buildings are used in locations where tornadoes or hurricanes are a possibility, with data centers located in the building core.

Even with all of the preparations that an Internet data center can make, there is still the possibility that a crippling failure could bring the entire center down, which is one of the reasons why most IDCs maintain multiple sites. Customers requiring the highest availability distribute systems across many locations, with geographic fail-over enabling operation to continue in the event of the loss of one data center.

Network Architecture

Hand-in-hand with a fully-redundant physical infrastructure is a fully-redundant network architecture. There are three levels to consider in IDC network architectures:

- Virtually all Internet data centers have more than one location, and high-speed interconnects facilitate geographically-distributed customer sites that keep running even through the failure of a single data center.
- Inside the data center is another redundant infrastructure that distributes traffic with minimal delay from the center's core routers to the customer servers.
- Also inside the data center are secure, private networks used for management functions, administration, backups, and for connecting the components of multi-tier environments, for example Web servers, database servers, and application servers.

Whether discussing the routing between distributed data centers or within a single one, several concepts are important:

- *Fully-Meshed Connectivity*

  An ideal IDC architecture has fully-meshed connectivity both between and within each data center. A fully-meshed architecture has at least one path directly connecting every component (Figure 8-2.) This enables routing or switching to take place using the shortest path first, falling back to secondary, multiple-hop paths in the event that a primary path becomes disconnected. In practice, only some IDCs have fully-meshed connectivity between data centers, while many have fully-meshed architectures within the data center.

- *Layered Architectures*

  Layered architectures in IDCs are used to optimize performance, security, and to provide a level of abstraction for the network. Performance can be optimized by seeking a balance between the layers where packets are routed versus where they are switched. Security is enhanced by layering in packet filters and firewalls where appropriate, and by isolating customer network segments through Virtual Local Area Networks (VLANs), which are used to limit the impact of denial-of-service attacks on individual customers. Finally, layering isolates routers and switches from their peers so that, for example, a change in a server's MAC address doesn't require all network components to adjust their IP routing and switching tables.

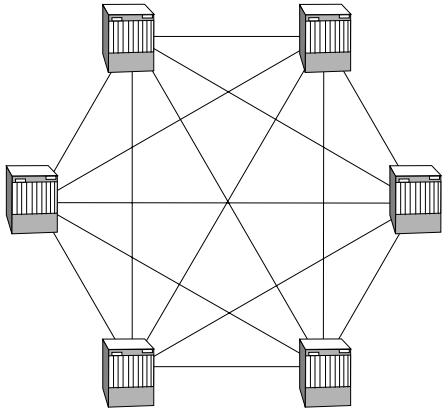

Figure 8-2 A fully-meshed architecture interconnecting six elements (for example data centers, routers, switches, etc.)

IDC network architectures, both local and global, layer their components into a *core layer*, where the number of routers is small and the latency is minimal; a *border layer* where policy-based connectivity is delivered; and an *access layer*, where packets are distributed — via routing and switching — to the appropriate physical servers.

Global Routing Strategies

Almost all Internet data centers have multiple locations that are interconnected via high-speed switched networks. Having multiple locations provides IDCs and their customers with multiple benefits:

- More locations means more capacity and performance because there is more data center availability across a geographic region.
- ISPs utilizing a geographically-distributed IDC may host a redundant server infrastructure in a second location, enabling fail-over from one to another in the event of a complete failure in one center.
- The IDC's backbone network can be used to carry traffic bound for its customers from its various Points of Presence (POPs) to the customer servers. This gives the end user a lower-latency connection than the public Internet, and gives the IDC customer more of a global presence. Some IDCs, like Digital Island, capitalize on their global network by providing caching servers in their POPs that can deliver content with even lower

latency than across the IDC backbone network. Other companies, like Akamai, locate caching servers in ISPs and IDCs worldwide, selling the service of bringing their customers' content closer to Web site visitors.

Geographically-distributed data centers are interconnected with redundant switched networks, and utilize Open Shortest-Path First routing (OSPF) between them. Based on both economic and network latency costs, each link between data centers can be assigned a cost and traffic routed based on those costs. Network traffic can be balanced between connections when equal costs are assigned, and backup paths are created when unequal costs are assigned. For establishing peering relationships and connectivity to other backbone providers, Border Gateway Protocol (BGP) is used.

Figure 8-3 illustrates a set of four data centers distributed across the United States, in San Francisco, Los Angeles, Chicago, and New York. Although the network is not fully meshed, it can survive the loss of two links. For example, if all connectivity between San Francisco and Los Angeles is lost, traffic can be routed via New York and Chicago.

At each POP, each ATM Permanent Virtual Circuit (PVC) terminates in each of two core routers. Each POP typically has its own peering relationships and Internet backbone connections so that it can operate independently and can accept transit traffic for the other locations.

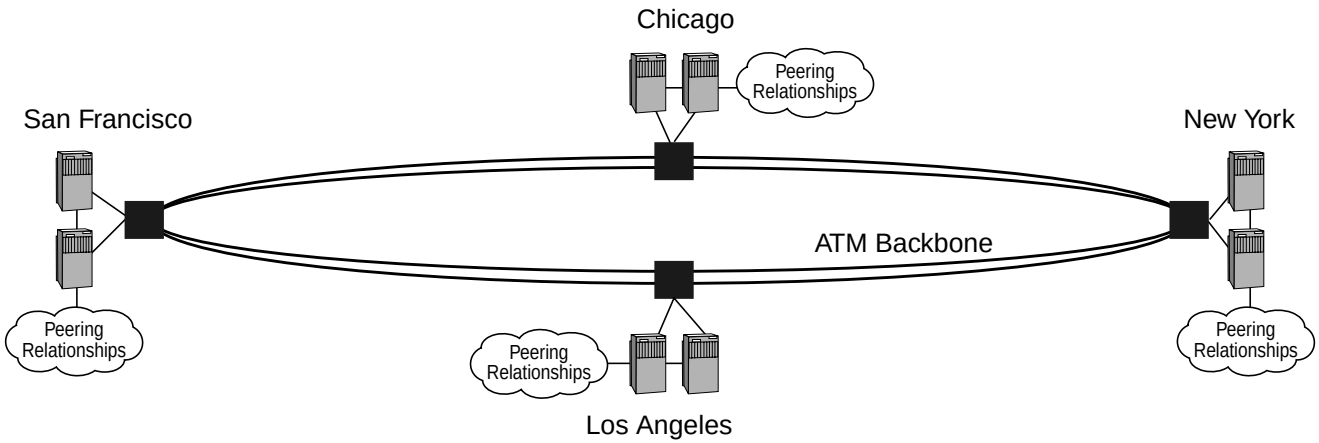

Figure 8-3 Topology of a geographically-distributed set of Internet data centers.

Switched Data Center Infrastructure

Within the data center is a fully-meshed, switched fabric that is layered for performance, security, and isolation of customer devices. A layer of core routers transfers traffic from the IDC backbone and, following policy-based routing rules, routes it down to a set of distribution switches which pass the traffic on to the access layer (Figure 8-4). Customer equipment connects into the access layer, often with two feeds, each from a separate access switch. As the sophistication of network switches increases with the addition of features typically found in routers, this distinction will increasingly blur.

The hierarchy of the data center's public network enables a fan-out to a large number of servers without introducing any significant latency. IDCs typically use routers and switches with VLAN capability. The use of Virtual LANs enables customers to be isolated from each other, and it enables a single subnet to extend across switch boundaries to serve separate racks or cages. The use of a fully-meshed, switched infrastructure ensures that any single failure — all the way to the server cable drop — will result in network traffic being routed over an alternate path.

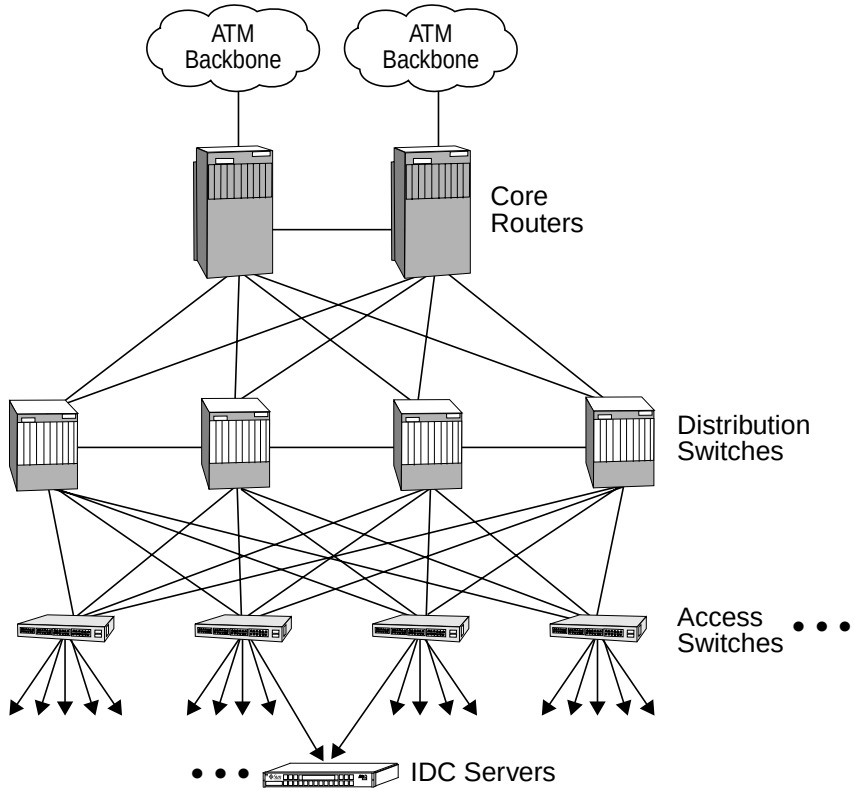

Figure 8-4 Hierarchical, fully-meshed network architecture within the data center.

With the hierarchical structure of an IDC's network, high-traffic services can be inserted into appropriate layers of the infrastructure. For example, a backup or database service can be served from the distribution layer, providing a high-capacity path accessible to each hosted server. A fact of life in Internet data centers is that larger customers will demand access to higher layers of the hierarchy — a request that is often accommodated for an extra charge.

Connectivity for private services such as storage-on-demand or backup may be separated logically from the existing public network using VLANs, or may separated physically through the use of separate physical networks.

Leveraging Infrastructure for ISPs

The increasing ability of routers and switches to support VLANs, stateful packet filtering, load balancing, and even content caching makes it possible for IDCs to easily establish multiple security zones for ISP customers using one consolidated infrastructure. With these infrastructure components deployed in pairs and cross-connected for coordinated fail-over and load balancing, IDCs can deploy a 'ladder' structure of routers and switches so that each layer creates a set of isolated virtual sub-networks using VLANs (Figure 8-5). This infrastructure can be used to isolate sub-networks at the same level in the infrastructure, and also to control traffic between virtual sub-networks.

It can be cost-effective for IDCs to manage the network infrastructure for ISPs using this architecture because the cost of the high-end networking components is amortized across many customers. It can also provide a superior solution for ISPs. With virtual, rather than real, sub-networks created through VLANs, smaller, more granular security domains can be created. For example, as Figure 8-5 illustrates, separate Web and Mail security domains can be created at both the services network level and the storage level. These techniques can be used to build the ISP infrastructures discussed in the following two chapters.

Summary

There is a rich symbiosis between ISPs and IDCs, with larger ISPs emulating IDCs, some ISPs becoming IDCs, ISPs using IDCs, and some ISPs entirely contained within IDCs. The driving force behind this symbiosis is the opportunity to capture increasing amounts of the growing outsourced business services market, where the key to success is quality of service. Using both the physical and the networking infrastructure provided by IDCs, ISPs can more easily deliver services that business customers demand, including those which are more high-performance, reliable, available, scalable, and secure. Using the high-end switched infrastructures of IDCs, ISPs can more effectively address these requirements and meet the service level agreements demanded by their customers.

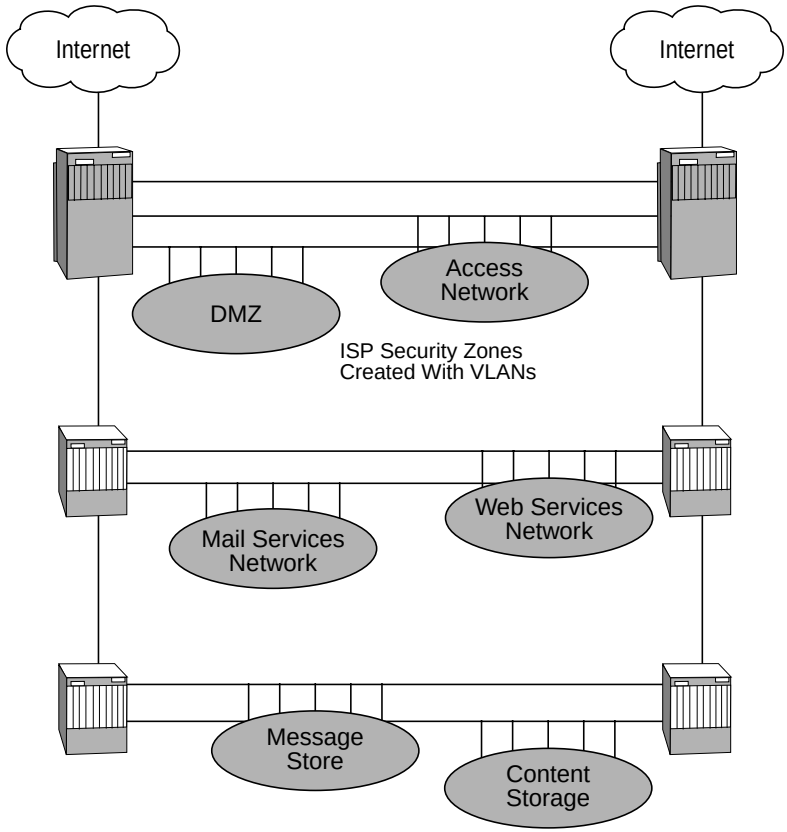

Figure 8-5 IDCs can create multiple security zones for ISP customers by creating sub-networks using VLANs with routers and switches capable of packet filtering and even content caching.

ISP Deployment Issues 9

So far, this paper has described ISPs from both a network architecture and a services-level perspective. It has described techniques that can be used to enhance the key properties that ISPs require, namely scalability, availability, manageability, and security. It has also described in detail how software components can be deployed onto network architectures to provide the right quality of service at the right price point.

This, and the remaining chapters address deployment issues, including illustrations of how local, regional, and national ISPs might configure their networks to host mail, news, and Web services with quality-of-service levels appropriate for their customers. These chapters are important not for the exact network architectures they propose, but more for the insight they provide into how ISP configurations can be designed to support the services provided by an ISP today. The following chapters also illustrate how they can grow or be modified to provide additional services and to grow as rapidly as their subscriber base.

Basic Building Blocks

ISP network architectures begin with three basic building blocks — three networks that serve different purposes and where traffic between the networks is regulated by packet-filtering routers and/or firewalls (Figure 9-1a).

- An *access network* connects the ISP to both the Internet and its subscribers, whether connected via dial-up, broadband, or wireless connections

- A *services network* hosts the services that the ISP provides to its customers — like Web, mail and news services — and also the services that it provides to other Internet users, namely access to hosted Web sites
- An *internal network* supports the private functions that have to do with maintaining and managing the ISP itself, including network management, billing, and customer-relationship management

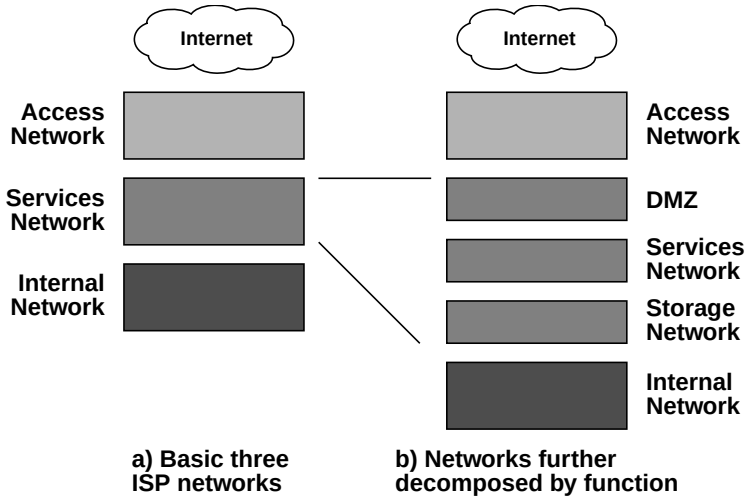

Figure 9-1 ISP network designs begin with three networks and further decompose them to meet goals of scalability, availability, manageability, and security.

Depending on an ISP's customer base, number of customers, service access patterns, type of services provided, and quality-of-service goals, ISP networks often further decompose the access, services, and internal networks by function. For example, the services network might be separated into three tiers (Figure 9-1b):

- A *DMZ* supports gateway functions that protect services from direct protocol attacks, and also caching and proxy functions that protect ISP users from attack while minimizing the external Internet bandwidth required.
- A *services network* includes the front-end processing systems that actually handle mail, news, and Web hosting.
- A *storage network* holds the resources used by the services network, for example mail files or databases, shared Web content, and news articles.

Decomposing the basic three ISP layers into a more fine-grained hierarchy of networks enables them to scale and tune the services at each layer for optimal performance, improve security through layered security perimeters, and increase flexibility of how services are deployed. Using a single mail storage cluster, for example, enables multiple mail front-end processors to access the same data store, eliminating the need to statically partition customers onto specific mail servers. Many of today's commercial software packages supporting ISP functions are flexible and can be deployed onto a single server, or layered onto multiple tiers of servers.

Design to Maximum Scale

When designing ISP networks, a key principle is to design the architecture to maximum scale. Beginning with an end point in mind — 10,000, 100,000, or millions of subscribers — ISP designs should either begin with the network configuration that can support their maximum scale, or have a plan that supports rapid growth and re-deployment of networks and servers to accommodate a rapidly-growing customer base.

Determining the scale to which an ISP network must be designed is a tricky business because it involves a complex trade-off between the ISP's business objectives, desired quality-of-service levels, customer usage patterns, and the number and bandwidth of access points. There is no factor more key in this calculation than the nature and size of ISP access networks.

ISP Access Networks

ISP access networks provide the means for ISP customers to connect to both the Internet and the ISP network itself, making them a key factor dictating the maximum network capacity to which an ISP must be designed. The maximum bandwidth that ISPs encounter today has been growing both larger and more variable over the years — making the task of designing to maximum scale even more challenging.

In the early days, most ISPs provided dial-up access through their own modem banks where the number of concurrent customers and bandwidth was both predictable and limited by the number of dial-up connections configured. When demand exceeded a fixed amount, customers simply encountered a busy signal on their dial-up modems. Today, with broadband DSL, cable, and fixed wireless connections becoming more prevalent, bandwidth requirements have

increased and so have customer usage patterns. As mobile wireless connectivity becomes a reality, even more challenges to ISP networks will be presented.

Dial-up, broadband, and mobile wireless connectivity represent not only a chronological progression of 'last mile' connection technology utilized by ISPs, but they also represent a set of choices service offerings that an ISP may choose to provide.

Dial-Up Access

ISPs have historically configured and supported their own modem pools, extending their reach by establishing points of presence with modem pools in markets they wish to penetrate. The traditional rule-of-thumb for sizing modem pools is to have one modem available for every 10 dial-up customers.

Today, support of modem pools has become such a commodity business that many ISPs outsource their modem pools to organizations such as telephone companies, where the access, infrastructure, and support is local to the area where customers are to be reached. These modem pools can be connected via high-speed lines to an ISP's core facility; they may offer direct Internet access, with ISP services accessed across the open Internet; or they may be connected via VPN tunnels from the modem pool across the Internet to the ISP facility where services are provided.

The fact that outsourced modem pools have become a commodity operation means not only that regional and national ISPs can extend their reach with ease. It also enables local ISPs to keep their customer base from migrating to larger ISPs by providing nationwide service while maintaining close, personal relationships with their customers.

Broadband Access

In contrast to dial-up access, broadband customers are always attached, making it more difficult to limit peak workloads. Whether connected via DSL, cable, fixed wireless, or leased line connections, these customers are always on, making it convenient to more frequently and more heavily use their connectivity. Experience shows that broadband ISP subscribers use services more heavily than dial-up customers, and Sun estimates that an ISP can support three times the number of dial-up than broadband subscribers with the same infrastructure.

There are some physical limits to the maximum bandwidth that customers can demand, for example the maximum carrying capacity of a broadband cable segment, or the connectivity between a DSL modem pool and the ISP network. Using these characteristics to limit the need for ISP capacity is unwise, however, because customers whose connectivity is limited by highly over-subscribed networks will likely take their business elsewhere.

Mobile Wireless Access

Forward-looking ISPs are rapidly developing strategies to capitalize on the growing number of subscribers able to access the Internet through mobile wireless networks. With the cost of PCMCIA cards for wireless connectivity plummeting and also being incorporated into many laptop computers, the number of customers interested in accessing the Internet from coffee shops, libraries , hotels, airports, or park benches is rapidly increasing. Two technologies are likely to co-exist to support mobile wireless access:

- *General Packet Radio Service (GPRS)* is part of the GSM family of standards used by cellular providers in Europe and Asia, with deployment in the United States underway. Having a theoretical maximum speed of 171 Kbps, GPRS is likely to be deployed by telco ISPs for Internet access.
- *IEEE 802.11b* is a wireless local-area network standard for which a variety of inexpensive devices and access points are available. 802.11b, or WiFi technology is commonly used in homes and small businesses, and with a maximum bandwidth of 11 Mbps it can easily bridge wired and wireless networks.

Although the market for both of these technologies is relatively young, there are several scenarios through which both of these technologies could be exploited by ISPs to provide a higher level of service and support a wider customer base. With GPRS offering a relatively wide range of wireless coverage, with low bandwidth, and 802.11b offering a short coverage range with high bandwidth, either or both of these technologies are likely to be exploited for complementary purposes. Regardless of which scenarios play out in the market, each one will have an impact on the overall workload parameters to which ISP networks are designed.

GPRS

Because of the infrastructure required to support cellular access through GSM and GPRS, these technologies are likely to evolve through adoption by telco and large ISPs. In the European and Asian markets, where local telephone service is costly and thus limits the dial-up ISP market, Internet access through cellular technology like GPRS could help ISPs leapfrog past local loop limitations and fuel an explosion in growth.

802.11b

The higher bandwidth and lower cost of 802.11b technology makes it likely to occupy a niche in the future of mobile wireless Internet access. With a more limited range than cellular GPRS technology, 802.11b Internet connectivity is likely to be adopted at more of a grass-roots level. Three interesting trends in the United States are unfolding in this area:

- *Wireless Community Networks* are being created by *ad hoc* groups of Internet users in close proximity. One or more community members host a wireless access point, or 'hot spot' to which other members can connect and enjoy high-speed wireless connectivity.
- *Commercial 802.11b Networks* are being established using a similar model, where Internet users are given an incentive to sharing their Internet connectivity with users authenticated by a service provider that manages and runs the operation. In the U.S., Boingo, and in Japan, DoCoMo are examples of such service providers.
- *Local ISPs* are establishing wireless hot spots in public areas like coffee shops and metropolitan areas. These ISPs typically bridge from the wireless network to the ISP network through a wired connection, though GPRS and line-of-sight 802.11b connections can be used as well (Figure 9-2),

Regardless of the business model, access to services via 802.11b networks is often managed using 'captive portal' software. Hosted on a generic Linux or a Sun Cobalt server, the captive portal allows users to spontaneously connect to the wireless network, granting them Internet access only after they have provided an authentication credential to the portal.

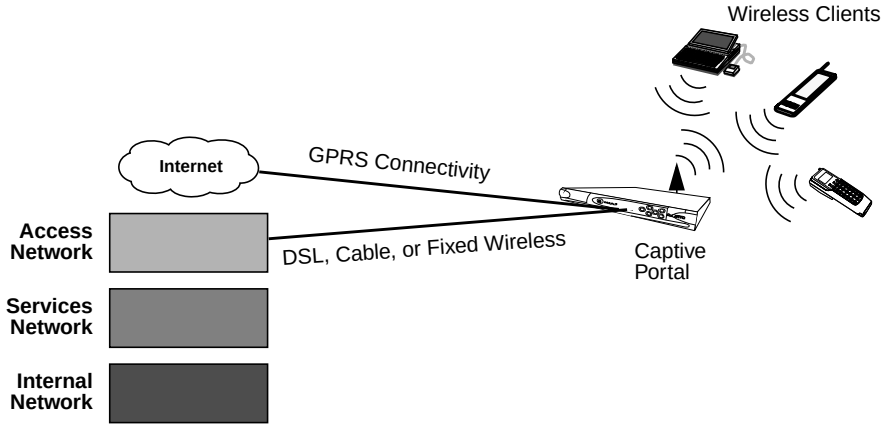

Figure 9-2 An ISP can provide wireless connectivity via wireless 'hot spots' deployed in locations such as coffee shops, metropolitan areas, hotels, conference centers, and businesses. Example return connections include third-party GPRS, broadband DSL, cable, or fixed wireless connections.

Summary

All ISPs begin with a network model that includes an access, services, and internal network, each of which can be further decomposed as business model, desired quality of service, number of customers, and required bandwidth dictate. A key factor in determining an appropriate ISP architecture is the services network. The services network determines the number of customers and the bandwidth they require through connectivity including dial-up, broadband, and mobile wireless networks. A potential new growth area for ISPs of all sizes is reaching more customers through mobile wireless networks using GPRS or 802.11b technologies. The remaining two chapters discuss example architectures for local, regional, and national ISPs, each of which can be configured with access networks using any or all of the technologies discussed in this chapter.

Local ISP Configurations 10

Local Internet service providers handle between 5,000 and 10,000 subscribers. With a concurrency rate of 10 percent, this requires accommodating a maximum 500 to 1,000 simultaneous users. This chapter examines a minimal local Internet service provider architecture with examples on how it might be scaled. The example provided in this chapter is a good prototype for those considering entering the ISP market with approximately 5,000 subscribers.

Characteristics of Local ISPs

Local Internet service providers can use standard tools, deploy small networks — often with entry-level servers. Although performance and reliability are important, some compromises can be tolerated:

- *Standard Tools.* Local ISPs can use many standard UNIX tools and off-the-shelf software. The choices range from standard utilities provided with the Solaris Operating Environment to sophisticated packages that can ensure scalability long into the future.
- *Small Configurations.* Local providers tend to use networks with a small number of servers to achieve maximum functionality at low cost.
- *Entry-Level Servers.* Local ISPs often deploy low-cost, entry-level servers while maintaining the ability to upgrade to more high-performance servers as needed.

- *Single Points of Failure.* Compromises are often made in local ISP architectures, allowing areas in which a single failure can bring down some or all of the ISP's operations.

The architecture described in this chapter reflects a set of choices that can provide a reasonable level of service with emphasis on low cost. Illustrated with fewer servers than Sun would typically recommend, some of the trade-offs of where to add components can then be discussed. The configuration's capabilities can be improved depending on performance, reliability, flexibility, and security needs, and it can be scaled down to further reduce costs or to handle fewer subscribers.

Architectural Overview

Like Internet service providers of all sizes, the basic configuration (Figure 10-1) consists of a set of sub-networks isolated by firewalls. The cornerstone of this local ISP is a Sun Fire™ V120 server that is configured with SunScreen Secure Net software (included with the Solaris 9 Operating Environment) and a low-cost, Sun Quad FastEthernet card that provides security and isolation between three basic networks:

- The *access network* provides connectivity to the ISP from the Internet and from dial-up users, as well as routing between dial-up users and the Internet. In order to reduce bandwidth used by customers, a caching server is located in the access network.
- The *services network* provides mail, news, and Web services as well as the DNS service that is visible to subscribers and the outside world.
- The *internal network* provides authentication, billing, and network management services — and is the most carefully-protected sub-network in the ISP design.

Access Network

The core of the access network is a router that handles network traffic to and from the Internet over a redundant pair of T1 connections. This Internet connectivity provides an aggregate bandwidth of approximately 3 Mbps, and the two connections ideally attach to the Internet through two different providers. This reduces the chance that a single network failure would cut off

all outside access. To provide additional resiliency to failure, the actual modem pool might be located in two separate telephone company central offices, so that one outage cannot take the entire ISP down.

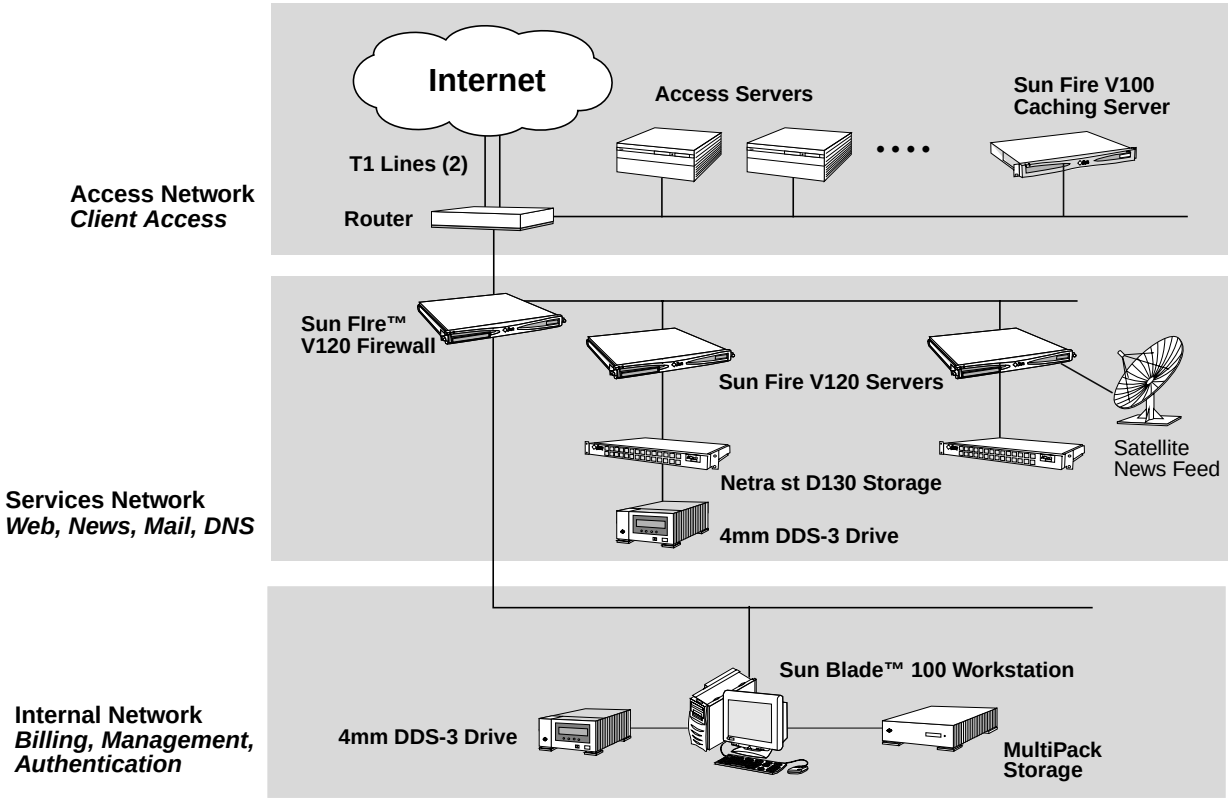

Figure 10-1 Example local ISP architecture

Dial-up access is provided by a set of 10 48-port access servers which provide a total of 480 dial-up connections — sufficient for 4,800 dial-up subscribers at a 10 percent concurrency rate. Note that the access network enters the ISP facility through the core router. This enables packet filtering between both dial-up and Internet users and the services network.

The access network can be scaled to support more dial-up customers by adding more access servers. Reliability can be increased by configuring a redundant pair of routers and an HA firewall configuration so that a failure at this point does not curtail Internet access.

In order to reduce the amount of traffic to the Internet from customers, a Sun Fire V100 server is used as a caching server, running Sun ONE Proxy Server or open source SQUID software. For a commercially-supported solution, the Sun Cobalt CacheRaQ 4 server provides an easy-to-configure appliance solution. Regardless of which approach is used, caching servers can significantly reduce network traffic — and thus the amount of Internet bandwidth an ISP must purchase — by storing local copies of frequently-used Web content.

Address Translation

Because of the shortage of Class B and Class C addresses, some ISPs configure dial-up users with private IP addresses, translating them to public addresses as requests are passed to the Internet. This allows sharing of a large number of private addresses on a small set of official, public IP addresses. This functionality is available with SunScreen Secure Net software, and could be added to this network by connecting the collection of access servers to a separate subnet attached to the firewall.

Switched Ethernet

Note that the network diagram is a logical perspective that does not include the physical wiring details of the ISP. For example, IP switching is required for each independent sub-network, which can be accomplished in a stand-alone ISP through a set of 100 Mbps switches. If the ISP is located in an IDC, larger switches are often used with VLANs supporting switching and isolation of the sub-networks.

Services Network

The services network provides all of the familiar ISP services to subscribers, including mail and Web services. It is populated with two Sun Fire V120 servers, Netra st D130 disk storage, and one StorEdge DDS-3 tape backup device. The two servers are shown with external storage devices that enable ISPs to expand their disk space by adding hot-pluggable drives as needed.

- The mail and Web server is a Sun Fire V120 server configured with 256 MB of memory and 72 GB of external disk space in a Netra st D130 storage unit, scalable up to 108 GB with three 36 GB drives. This disk space allows for more than 5 MB per user for mailboxes, and 5 MB for each hosted 'vanity' site, assuming 10 percent of users host Web pages. Services on this host can be provided with standard sendmail and the open source Apache Web server to minimize cost; commercially-supported products like Sun ONE Messaging Server and Sun ONE Web server can also be used.
- The news server is a Sun Fire V120 server configured with 256 MB of memory, and additional 72 GB of disk space News software is open source Inter-Network News (INN) or commercially-supported products.

  Many small ISPs do not provide netnews services because of the high bandwidth requirements and the declining use of the service. For those that wish to provide the service, some choose to outsource to an upstream provider; the example illustrated shows the ISP subscribing to a satellite download service that can provide news content without burdening the ISP's main Internet connection.
- The ISP's DNS services run on both servers, with one acting as a primary, and one as a secondary.

The allocation of services to these two servers is somewhat arbitrary. The ISP's strategic business goals and expected workload should be used to determine the actual allocation at deployment time. This starting point was chosen because netnews has a heavy, constant bandwidth requirement that can limit the ability of the mail and Web services to handle the peak demands of their services. The 4mm backup drive is configured on the mail/Web server because backups for mail are more important than for netnews.

Scalability

The services network can be scaled up or down depending on the expected number of subscribers and the workload the ISP expects to encounter. Both vertical and horizontal scaling can be used.

Horizontal Scaling

Horizontal scaling is probably the first possibility to consider in the services network. An additional server could be added to allow mail, news, and Web services to be hosted on separate systems. It is possible to scale down to a single server, however the existence of at least two hosts ensures that the failure of one server does not bring down all of the ISP's services.

Vertical Scaling

Vertical scaling is secondary to horizontal scaling in such a small configuration. A more flexible starting point for this network would be to use two Sun Fire 280R servers. As performance requirements dictate, an additional processor could be added to each server as needed.

Reliability

Reliability of the services network could be improved by the addition of more servers, more reliable storage subsystems, and high-availability clusters.

Additional Servers

Adding a third server in the service network would allow each service to be independently partitioned onto its own host. This step would ensure that the failure of a single server would cause the ISP to lose only that function, and users could still access the remainder of the ISP's services in the event that one fails. This approach would increase security, because for example a successful intrusion through a mail server vulnerability could not be used to interfere with Web services. It would also provide resilience against denial-of-service attacks in that an attack on one service does not necessarily impact the other services.

One technique to minimize down time in the event of a server failure is to load all software packages on each server, enabling rapid configuration of one server to take on the workload of a failed server. This technique, of course, requires moving disk storage systems should a manual fail-over be necessary.

Even with an additional server in the services network, the router and firewall present a single point of failure for the ISP. Adding a redundant router and firewall in a fail-over configuration can also significantly increase availability.

Storage System Reliability

With the low cost of external disk storage, and the relatively low disk space requirements of a local ISP, reliability of the storage element could be increased with the use of a redundant set of storage devices configured with host-based mirroring such as that provided with Solaris™ Volume Manager software.

High-Availability Clusters

An approach that leverages both redundant servers and storage is to deploy Internet software compatible with Sun Cluster HA systems (Figure 10-2). The Sun Cluster can be configured with two Sun Fire V120 servers and a pair of cross-connected Netra st D130 disk storage devices. A private pair of network connections allows each server to reliably monitor the status of the other.

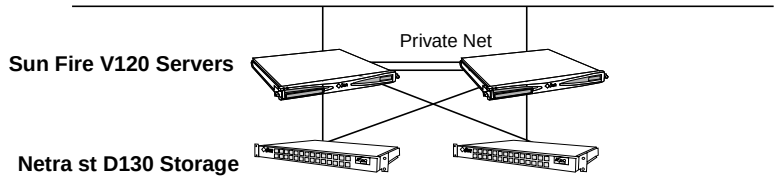

Figure 10-2 Reliability with a Sun Cluster using Sun Fire V120 servers and storage with Sun's HA-ready Internet software.

Internal Network

Business operations for the ISP take place in the internal network. In a local ISP installation, one Sun Blade™ 100 workstation can be used for authentication, billing, logging, network management, and problem tracking. Because this network contains mission-critical data for the ISP, it must be protected accordingly. Some ISPs have additional authentication servers outside of the internal network so that network traffic can be restricted to pass between the internal network and the services network only.

Authentication and Billing

Software for authentication and billing is beyond the scope of this document, however there are many excellent solutions available from Sun's integration affiliates. Some ISPs located in IDC facilities outsource their billing functions to the IDC itself.

Network Management, Backup, Customer Service

There are many choices for network management, backup, and customer service functions. Network management can be provided with tools like Sun™ Management Center and HP Open View. Other third-party tools can be used for backups and for streamlining help desk support.

Scaling

The administration system is a Sun Blade 100 workstation with external storage for billing data and logs, and a 4mm DDS-3 drive for backups. As the ISP subscriber base grows, it is easy to add more workstations and servers to the internal network, each hosting separate functions.

A first step would be to use a separate server for logging so that information on usage and potential security incidents can be stored off the public networks on a server configured only to accept syslog messages. This server would allow for analysis if ISP security is compromised.

Summary

Local Internet service providers can use off-the-shelf products and open source software to configure services and quickly come on-line. Compromises in areas such as performance and security can be made to reduce costs, However even with low-cost configurations such as the one used in the example, additional servers and more powerful CPUs can be added as needed. While scalability is an important consideration for local ISPs, it becomes of paramount concern for regional and national ISPs.

Regional and National ISPs 11

Regional Internet service providers handle roughly 50,000 to 75,000 subscribers. National ISPs begin with around 75,000 subscribers and scale into the millions. Some of the differences that characterize these architectures from local ISPs is increasing redundancy of components for high availability, more layering for security, and specialization of services. Scalability is a paramount concern to these ISPs because most of the problems that can put them on the nightly news are related to their inability to accommodate growth.

Competition between large ISPs is fierce, and their configurations are usually closely-held secrets. As a result, actual regional or national ISP architectures are not presented here, rather this chapter illustrates how the major principles discussed throughout this paper apply. As a rule, these ISPs utilize all of the architectural techniques discussed in chapter 3 — and the way in which these techniques are applied to their networks is discussed here.

Characteristics of Regional ISPs

Based on a 10 percent concurrency level, regional ISPs must service between 5,000 and 7,500 concurrent users. Maintaining security and performance requires larger and more complex networks than local ISPs. These networks are always custom-designed, and vary significantly depending on the kinds of services to be provided. Because of the custom nature of these ISP installations, consultation with Sun Professional Services or Sun's integration affiliates is essential. Some of the characteristics of these ISPs include:

- *Custom Tools.* Most standard UNIX utilities have limited capabilities for handling more than 50,000 subscribers, so specialized packages and home-grown software solutions based on open source software are the rule.
- *Scalability.* A growth rate of even a few percent can result in an additional thousand subscribers, so scalability is especially key for regional ISPs. These networks must support fast and easy configuration of additional servers to maintain performance at a target level. Networks are typically based on 100 Mbps Ethernet, with gigabit Ethernet sometimes used on backbone networks.
- *Reliability.* Reliability is a priority for regional ISPs. Such ISPs are often subsidiaries of larger organizations — quite often telephone companies — who do not want the publicity of network failures to tarnish their reputations in other service areas.
- *Importance of Processes.* Many ISPs observe that as the number of subscribers at an ISP reaches 20,000, administration costs peak. Beyond this point — in the realm of the regional ISPs — it is essential to establish automated operating procedures, including subscriber sign-up and pre-configured tools like Netscape Communicator with settings for their network. The savings in support costs make it possible to provide these packages to subscribers free of charge. Automated help desk software helps to track customer problems and ensure timely problem resolution.
- *Points of Presence.* Regional ISPs must extend their services across many local calling areas, making it essential to have POPs in a wide range of geographical locations. These ISPs usually prefer remote routers and modem banks that connect through high-speed lines (e.g., T1, T3, and ATM) to a single core service center. These POPs must operate in unattended, “lights-out” mode, and require telco-grade equipment like Sun Netra servers.

Characteristics of National ISPs

National ISPs serving more than 150,000 subscribers inherit all of the concerns of regional ISPs, and their large size tends to raise additional issues:

- *Scalability.* Still key to the ISP’s success, national ISPs can provide scalable architectures in two ways. Centralized architectures can be designed with unattended POPs. Alternatively, centralized architectures can replicate regional-sized networks across the country with central control of

administration and billing. Network traffic is such that 100 Mbps Ethernet cannot handle all of the ISP's traffic, and even sub-networks must be carefully designed to ensure that no single segment of 100 Mbps Ethernet becomes overloaded.

- *Reliability.* As failures in these ISP networks frequently receive national news coverage, reliability is a matter of public credibility. National ISPs must have high-availability networks that can dynamically reconfigure around outages in parts of the network, much like the reliability that telephone companies have provided for years.
- *Roaming.* In order to reach more subscribers, national ISPs are beginning to form alliances with other ISPs, allowing subscribers to dial-in using another ISP's points of presence. Using an integrated authentication network, a roaming user can gain access through leased lines back to the national ISP services.
- *Proxy/Caching Servers.* Because of the huge volume of Internet traffic that can be generated by even a small percentage of subscribers accessing the same resources on the World Wide Web, a layer of caching servers is frequently deployed in regional and national ISP networks. These servers cache frequently-accessed Internet data — including that transmitted with the HTTP, FTP, RSTP, and NNTP protocols — and respond to subscriber requests with cached information. Caching hides network latency and can reduce Internet bandwidth requirements by up to 40 percent.

Architectural Overview

The architecture discussed in this chapter illustrates a fully-configured regional Internet service provider network along with some of the features that are found in national ISP configurations (Figure 11-1). This design features mail, news, and Web services decomposed by service, task layer, and by special function. Web services for a small number of virtual sites are illustrated; architectures for outsourced Web services for business customers are discussed in chapter 7.

The decomposition approach illustrated in this example supports scaling up to the national ISP range. Indeed, as the number of subscribers grow, and the ISP targets increasing numbers of business customers, quality of service is a key issue, with customers demanding higher levels of performance and availability.

The architecture described here illustrates the use of five separate security zones supported by five separate sub-networks. Security is enhanced through careful control of traffic between the sub-networks. The illustration shows *logical* connectivity between hosts in a sub-network and security between them. There are a variety of approaches for creating security zones that have been discussed in this paper, including the use of routers, firewalls, load-balancing mechanisms, and VLAN-equipped switching equipment. The details of how to create the separate security zones are not detailed in the diagram.

The five security zones illustrated include:

- The *access network* connecting the ISP to the Internet as well as to residential and business subscribers. There are some local modem pools and connections to points of presence which are unattended installations of modem banks, routers, and high-speed connections to the core installation.
- Redundant pairs of *firewalls* separate the various ISP sub-networks, providing both reliable and secure connectivity. With SunScreen Secure Net software on dual-processor Sun Fire 280R servers, the firewalls have ample capacity for stateful packet filtering and logging, which is an important aspect of firewall operation. Regional and national ISPs use log analysis software to watch for anomalies that might reveal security attacks.
- The access network is attached to the *demilitarized zone* (DMZ). The DMZ contains feed servers that handle incoming mail and news, a proxy/cache server that caches requested Web pages and acts as a proxy for outgoing Internet requests, and a pair of DNS servers, one primary and one secondary.
- The *services network* contains a set of front-end processors, each supporting a single service. Each service can be hosted on a single or multiple hosts with load-balancing routers to distribute the load. Decomposition by special function is used: the dynamic content generation function is split from the Web front-end processor onto a specially-tuned server.
- The *storage network* is connected to the services network by a pair of redundant firewalls. This network contains two servers running high-availability content-management software that provides data to the front-end processors. Illustrated are two Sun Fire V480 servers each having two paths to a Sun StorEdge T3 Array for the Enterprise, yielding a storage solution with no single point of failure. Two firewalls removed from the Internet, this back-end network is highly secure.

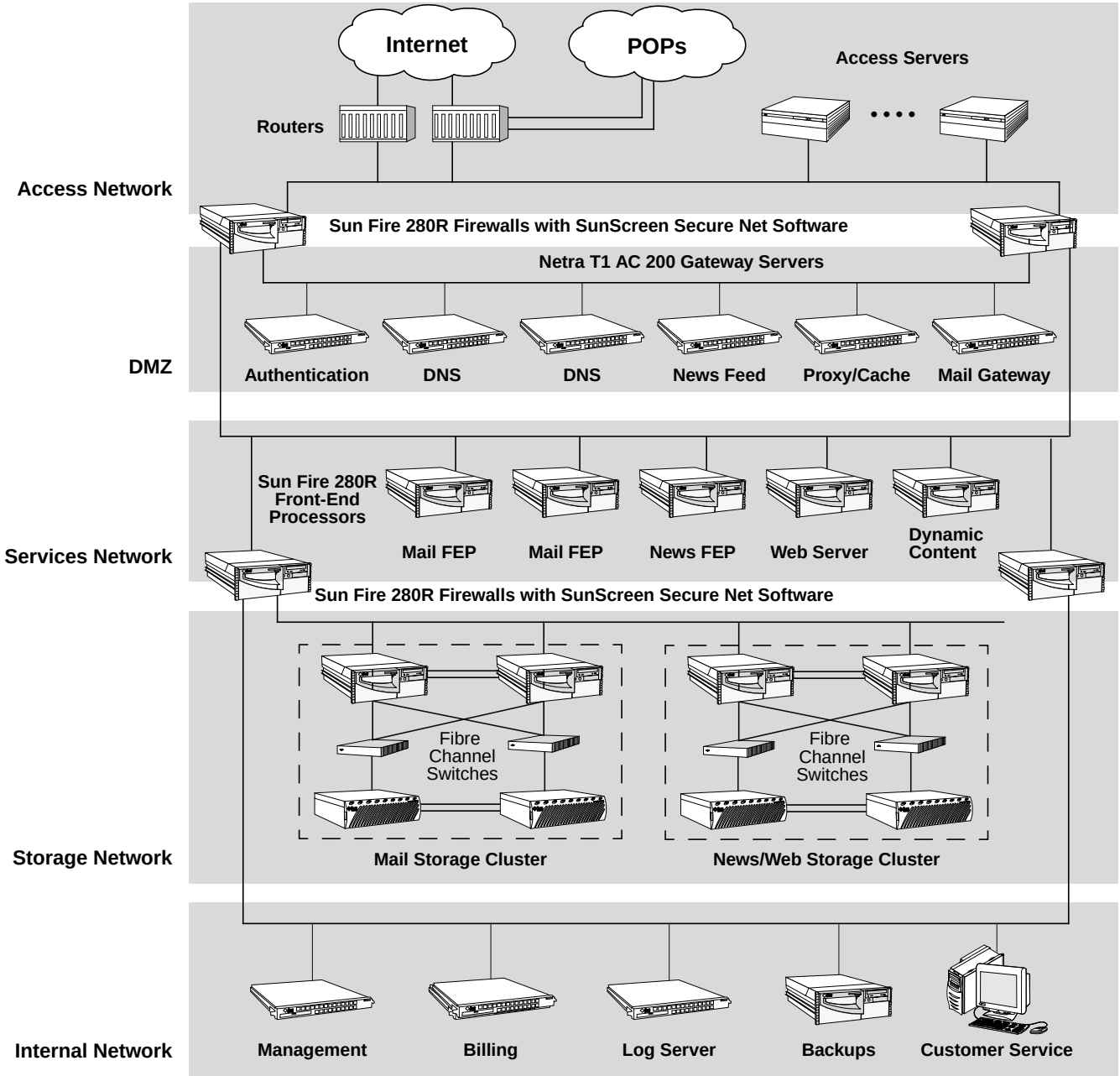

Figure 11-1 Example regional ISP network architecture

- The *internal network* contains several hosts for operations, management, and customer service. A logging server collects log information from each of the firewalls, ensuring that any intrusion attempts are traceable. A backup server uses a tape library (not shown) to back up all ISP content via a set of isolated, separate backup networks (not shown). A Sun Blade 100 workstation provides access to customer service functions, while Sun Netra servers are used to support the remaining internal functions.

Access Network

The access network contains a redundant pair of routers which manage T3 connections to the Internet and high-speed connections to the POPs. Access servers manage modem pools and handle user authentication with the authentication server in the DMZ. This, and all other networks in this ISP architecture, are supported with 100 Mbps Ethernet. The access network can be scaled by adding more basic components.

Demilitarized Zone

The DMZ provides a first line of defense against intrusion and gives low-latency access to services. All of the hosts in this network are Sun Netra T1 servers. The workload on these systems is not as critical as the decomposition of the different services onto separate hosts.

Domain Name Service

Two DNS servers are configured to provide name services for the servers in the ISP. To help ensure uniform response times for DNS requests, they can be configured with load-balancing routers that allocate each request to the most lightly-loaded server. In order to minimize the impact of denial-of-service attacks on DNS servers, a secondary pair of servers would typically be located in another sub-network or even in a location served by different upstream providers. This prevents an attack that saturates one sub-network from making all DNS services unavailable.

Authentication

The authentication server communicates with access servers both locally and at remote points of presence to provide user access to the network. Failure of this single server can bring the ISP down, so an important upgrade to the DMZ would be to integrate multiple servers using a load-balanced configuration. ISPs of all sizes are adopting LDAP as the standard repository for user preferences and authentication information. RADIUS authentication can be integrated with LDAP for authentication at remote POPs.

Mail Gateway

The mail gateway server uses Secure Mail Access Protocol (SMAP) to receive mail and store it to a file for later transfer to the mail storage server using SMTP. Initial acceptance of mail using SMAP is preferable to using the sendmail daemon because the SMAP daemon does not actually deliver the mail, making the service more resistant to penetration. The mail gateway can be scaled by adding a second server and using load-balancing techniques, or by using commercial software which can partition the user workload using configuration information provided by an LDAP server.

News Gateway

The news server receives articles from upstream news feeds and delivers them to the news storage server; additional servers can be added as necessary. Since failure of this server only causes a delay in articles (they are queued on the upstream side) this server is probably the least important one to duplicate.

Proxy/Cache Server

The proxy/cache server filters outgoing Web requests and caches incoming data so that multiple subscribers accessing the same Web page receive the latest information from the cache — significantly reducing the demand for Internet bandwidth while simultaneously improving quality of service for subscribers. The Sun ONE Proxy Server, Inktomi Traffic Server and the open source SQUID packages are popular among ISPs.

Services Network

The services network consists of a set of front-end processors that provide user access to mail, news, and Web services. These processors are configured to run their assigned services, however they contain no user data — they access mail, news, and Web pages from the storage network. This services network contains a dynamic content server to aid in the Web front-end processor's ability to quickly generate dynamic Web pages.

Scalability

The beauty of this architecture is that the front-end processors do not hold any user data, making horizontal scaling simple. Servers can be added by re-configuring DNS to include them in a service group, or by configuring load-balancing mechanisms. Using load-balancing routers, a collection of mail front-end processors are addressed with a single address like *mail.isp.net*, with the router translating the address to the least heavily-loaded server each time that a request is made.

As regional and national ISPs begin to grapple with the problems of deploying large numbers of servers in each sub-network, Sun's 106-processor Sun Fire 15K server presents an interesting option. This high-end server offers reliability, availability, and serviceability (RAS) features such as system-wide environmental sensing, resource management, dynamic reconfiguration, and hot-swap capability of all system components. These features are all critical for ISPs with large physical installations.

The feature that sets the Sun Fire 15K server apart from other large servers is the ability to configure it as multiple independent domains, each acting as an electrically-isolated server. This allows the ISP to exploit the RAS features of a large system while at the same time reaping the benefits of horizontal scalability — all in the same server configuration. Dynamic reconfiguration in the Sun Fire 15K server allows ISPs to respond to workload fluctuations in real time by re-allocating processors, memory, and even I/O devices without re-booting any of the domains.

Storage Network

The storage network contains two sets of Sun Fire V480 servers in clustered configurations running HA-NFS software. Each cluster consists of a pair of up to four-processor Sun Fire V480 servers attached via two separate storage-area network fabrics using Fibre Channel switches. These two networks connect to two StorEdge T3 Arrays for the Enterprise. The storage arrays are configured with interconnected backplanes, so that if one array's controller fails, storage can be accessed through the second array's controller.

Scalability

The storage network is populated with two content clusters — one for mail, and the other for news and Web page storage. The host systems are Sun Fire V480 servers, each with the capability to accommodate additional processors; more vertical scaling can be achieved by upgrading to servers capable of supporting more processors, like the eight-processor Sun Fire V880 server. Horizontal scaling can be accomplished with the addition of separate clusters for each service, and by partitioning some services (such as mail) across multiple clusters.

Internal Network

The internal network for this ISP contains a minimum of three servers, including those for billing, logging, backups, and other management functions such as problem tracking. With Web-based customer service and billing software available from Sun's affiliates, workstations like the Sun Blade 100 system can be used to support customer service representatives.

Summary

Regional and national Internet service providers are faced with using highly-scalable software packages, and developing custom software and unique architectural solutions to the problems of scaling to large numbers of subscribers. Performance, reliability, and security are paramount, and ISPs of this class cannot accept compromises. These configurations are always pushing the state-of-the-art, and utilize combinations of all of the architectural techniques that were discussed in chapter 3.

Conclusion 12

As the Internet service provider market grows, companies are diversifying, offering more unique services to differentiate themselves. ISPs continue to be the on-ramp of the Internet for most subscribers; Network Service Providers specialize in high-bandwidth connections and supporting Internet backbones; Application Service Providers host services ranging from sophisticated Web sites that conduct electronic commerce transactions to Enterprise Resource and Planning services; and finally, many ISPs are covering all niches and acting as Full Service Providers.

Regardless of how the ISP market is segmented, its growth continues unabated, and one immutable fact remains — that it will take ever more computing power to satisfy the growing number of commercial and residential subscribers and their increasingly sophisticated demands for services. Independent of which market segment they serve, ISPs, NSPs, ASPs, and FSPs will have astonishing needs for scalable architectures to support enormous growth.

With the scalable hardware and software technologies available to support ISPs of all sizes, Sun and its affiliates are well-positioned to support the range of services that all providers must deliver. The constantly-changing demands for services and the quickly-changing set of ISP technologies results in every ISP configuration being different. A network designed for one set of customers today may look quite different if designed six months in the future — as the state-of-the-art advances, ISPs can exploit new configuration options to remain competitive in this ever-changing market.

This document has presented an overview of some of the services that ISPs are expected to deliver and the principles behind the network architectures that support them. Examining these principles in actual ISP networks reveals insights into how to configure new ISPs, and how to scale existing ones. This document has drawn on the experiences of Sun Professional Services and Sun's integration affiliates, namely Global One, Omnes, Solect, and Technology Applications, Inc. An ISP can only be as good as the architecture on which it is built, so partnering with an organization that understands the latest technologies — and how to implement them in ISP configurations — is key to the success of Internet service providers.

Acknowledgements 13

This paper was made possible through the contributions of Jeff Allison, Dan Berg, Kirk Brown, Jason Carolan, Vincent Costa, Barb Dijker, Charlie Martin, Helen Robinson, Abe Schoenfeld, Jaap Romers, Ellen Siegel, James Simon, Stu Stern, David Troup, Michael Wallace, Dave Williams, and Joe Zender

About the Authors 14

Bruce Baikie

Bruce Baikie is the driving force behind *Internet Service Provider Configuration Guidelines*. Having conceived of the need for a comprehensive guide to the intricacies of configuring ISPs in 1995, Bruce provided the vision and direction that enabled a wide range of organizations inside and outside of Sun to be represented in this document. He has forged relationships with Sun customers and partners around the globe that have been critical to Sun's success in the ISP marketplace. Bruce is currently focusing on Internet Data Center issues in Sun's Network Service Provider Group.

Steve Gaede

Since its first edition, Steve Gaede has been the principal researcher and author of the ISP Guide. Steve has worked with Sun's ISP affiliates and across Sun's organization to unravel the common threads of ISP configurations everywhere, the ways in which Sun builds ISP infrastructures, and how Sun products are used to deliver ISP services. With a B.S.E. in Computer Engineering from the University of Michigan, and a M.S. in Computer Science from the University of California, Berkeley, Steve has prepared more than fifty research projects, whitepapers, and presentations for Sun. Steve is affiliated with PointSource Communications, Inc. and Lone Eagle Systems Inc.

Jim Baty

Jim Baty has been the architect behind many of the successful ISPs that Sun Professional Services has helped to deploy, and has provided his experience, insight, and vision as input to the ISP guide — as well has having authored key sections of the chapter on Web hosting. Jim holds a Ph.D. from University of Texas, Austin, Jim is currently Chief Architect and a Distinguished Engineer with Sun Professional Services' Internet Architecture Practice.

References 15

Sun Microsystems posts complete information on Sun's hardware and software products and service offerings in the form of data sheets, specifications, and white papers on its Internet Web site at *http://www.sun.com/*.

Sales Offices

Argentina: +5411-4317-5600
Australia: +61-2-9844-5000
Austria: +43-1-60563-0
Belgium: +32-2-704-8000
Brazil: +55-11-5187-2100
Canada: +905-477-6745
Chile: +56-2-3724500.
Colombia: +571-629-2323
Commonwealth of Independent States: +7-502-935-8411
Czech Republic: +420-2-3300-9311
Denmark: +45 4556 5000
Egypt +202-570-9442
Estonia: +372-6-308-900
Finland: +358-9-525-561
France: +33-134-03-00-00
Germany: +49-89-46008-0
Greece: +30-1-618-8111
Hungary: +36-1-489-8900
Iceland: +354-563-3010
India:
    Bangalore: +91-80-2298989 / 2295454
    New Delhi: +91-11-6106000
    Mumbai : +91-22-697-8111
Ireland: +353-1-8055-666
Israel: +972-9-9710500
Italy: +39-02-641511
Japan: +81-3-5717-5000
Kazakhstan: +7-3272-466774
Korea: +822-2193-5114
Latvia: +371-750-3700
Lithuania: +370-729-8468
Luxembourg: +352-49 11 33 1
Malaysia: +603-21161888
Mexico: +52-5-258-6100
The Netherlands: +00-31-33-45-15-000
New Zealand:
    Auckland: +64-9-976-6800
    Wellington: +64-4-462-0780
Norway: +47 23 36 96 00
People's Republic of China:
    Beijing: +86-10-6803-5588
    Chengdu: +86-28-619-9333
    Guangzhou: +86-20-8755-5900
    Shanghai: +86-21-6466-1228
    Hong Kong: +852-2202-6688
Poland: +48-22-8747800
Portugal: +351-21-4134000
Russia: +7-502-935-8411
Saudi Arabia: +9661 273 4567
Singapore: +65-6438-1888
Slovak Republic: +421-2-4342-94-85
South Africa: +27 11 256-6300
Spain: +34-91-596-9900
Sweden: +46-8-631-10-00
Switzerland:
    German: 41-1-908-90-00
    French: 41-22-999-0444
Taiwan: +886-2-8732-9933
Thailand: +662-344-6888
Turkey: +90-212-335-22-00
United Arab Emirates: +9714-3366333
United Kingdom: +44-1-276-20444
United States: +1-800-555-9SUN OR +1-650-960-1300
Venezuela: +58-2-905-3800
Worldwide Headquarters:
    Sun Microsystems, Inc.
    4150 Network Circle
    Santa Clara, CA 95054 USA
    Phone: +1-800-555-9SUN OR +1-650-960-1300
    Internet: www.sun.com