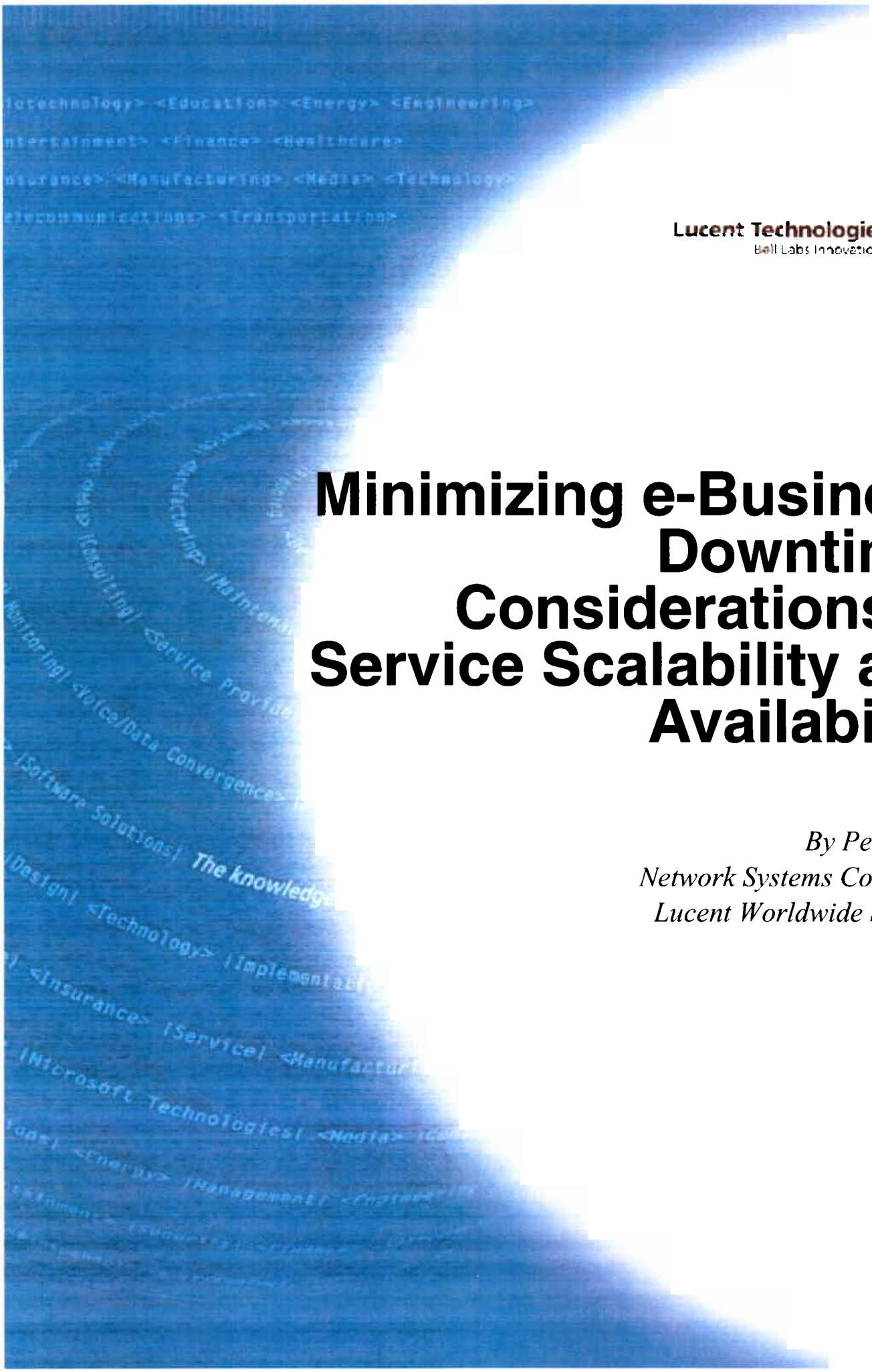

Lucent Technologies
Bell Labs Innovations

Minimizing e-Business Downtime: Considerations in Service Scalability and Availability

By Peter Farr
Network Systems Consultant
Lucent Worldwide Services

Minimizing e-Business Downtime: Considerations in Service Scalability and Availability

By Peter Farr, Senior Network Systems Consultant, Lucent Worldwide Services

Introduction

Your company has decided to make the investment in an e-Business presence. Business processes are being developed to support the new model, and senior management is looking to you to build the infrastructure. The new site will be the public representation of your company. Its stability, integrity and availability are critical, but the growth figures are just estimates based on possibly incomplete data. How do you go about designing and building a reliable and scalable infrastructure to meet the needs of your business? What are the issues that must be addressed?

Too often e-Business sites are designed and implemented as an independent collection of applications, servers and network components. To the end user, your customer however, only the service is important. The user doesn't know or care what components were combined to provide the service - only that it is responsive and useful.

Availability and scalability are two very important aspects of providing a positive and useful service to the end user. In some cases, both of these characteristics may be provided by the same design features while in other cases they may be mutually exclusive. This paper discusses some of the options and issues associated with service availability and scalability. A generic architecture, shown in Figure 1, will be used for discussion in this paper. This idealized architecture breaks an Internet site into three layers:

1. Service or front-end layer
2. Business logic layer
3. Data layer

Each of these layers performs a unique function in the overall service, and each has its own unique challenges and opportunities that will be discussed further in the sections below.

An e-Business site generally consists of one or more Web servers providing HTTP/HTTPS (Hypertext Transfer Protocol/Hypertext Transfer Protocol/Secure) and FTP (File Transfer Protocol) services. Sometimes other services such as e-mail are also offered, but this paper focuses on Web and FTP. Behind the Web server are often database, data warehouse, and transaction servers. Sometimes links to other services such as credit card authorization bureaux or fulfilment systems are also needed.

One concept that must be kept in mind throughout these discussions is called affinity. Affinity means that, once an initial connection or transaction is made between a source and destination, all subsequent transactions must continue to use the same source and destination. A Secure Socket Layer (SSL) connection between a browser and a Web server provides a good example, since SSL is a "stateful" protocol. Suppose there are two load-balanced Web servers at a site, and a user begins an SSL transaction with one of them – lets call it server A. The SSL handshaking mechanism will cause a digital certificate to be exchanged between the browser and server A. Then, both the server and client will store session encryption keys, sequence numbers and other data in their internal tables. If the next packet is sent to server B by the load-balancing product, server B will have no record of the encrypted session and

the session will hang or, more likely, terminate. Load balancing and high availability solutions should be able to handle sessions requiring affinity.

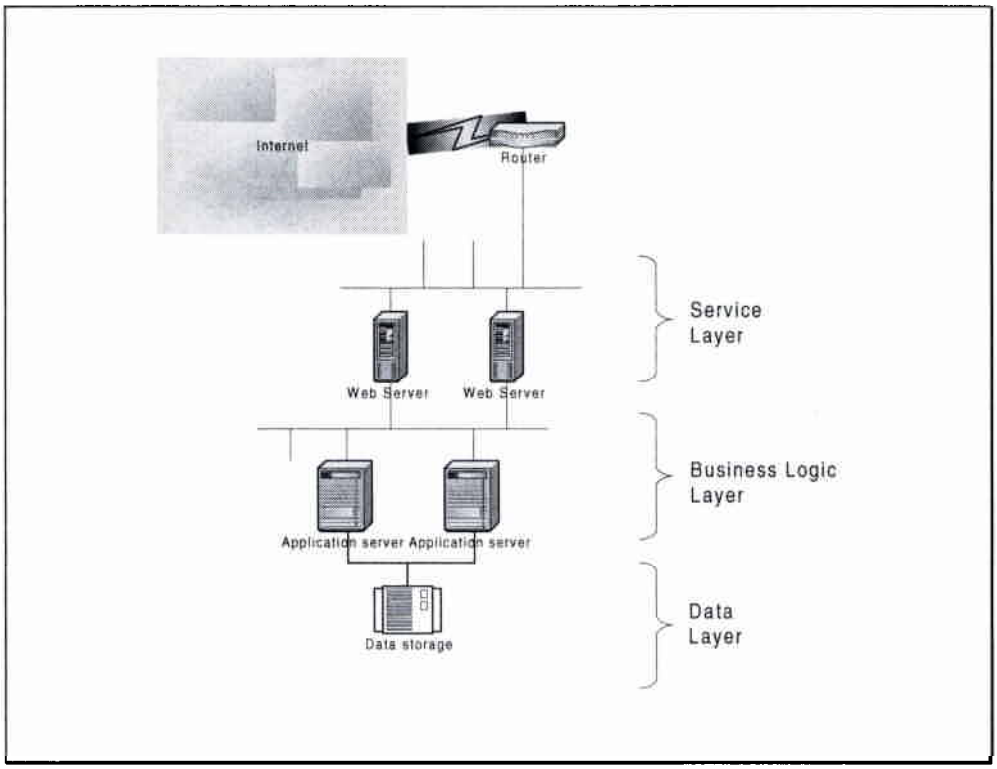

Figure 1 – Multi-tier e-Business Architecture

Service Layer

The Service layer of an e-Business site interacts with the end user's browser through the Internet, generally using HTTP. Sometimes HTTP/S (HTTP/Secure) is used to provide session level encryption for confidential data.

The Service layer presents Web pages and other data in two general forms: static and dynamic. Static data is generated by a content administrator at some earlier time and stored where the front-end server can access it when needed. Dynamic content is generated as needed based on user input or on data that changes frequently. It is not usually stored other than in short-term caches. The amount of static content, such as a simple "home page", vs. dynamic content, such as the results of a database search, may have an impact on your availability options.

There are many load balancing and availability options for the front-end of the service layer, between the end user and the Services systems. These range from simple and inexpensive solutions like round robin DNS to high performance special purpose hardware and software, as discussed in the sections below.

Business Logic Layer

The Business Logic layer contains the applications that service the front-end. These contain the business logic and transaction servers. In some cases, such as smaller sites or legacy applications, the business logic layer is combined with the services layer or the data layer. This may simplify the architecture and the administration of the site, but can reduce the flexibility, availability and scalability of the service.

The Business Logic layer is generally not directly accessible by the end user. Availability and scalability solutions at this level are frequently tied to the application design and are often implemented by application middleware.

Data Layer

The Data layer consists of disk drives, arrays and network storage appliances that provide reliable, high-speed application data to the Business Logic layer. Consolidating the data simplifies data backup, recovery and replication. The opportunities for availability and scalability in the Data layer are also heavily dependent on the capabilities of the application and database software. Some databases come with high availability software either as built in or "bolt on" modules.

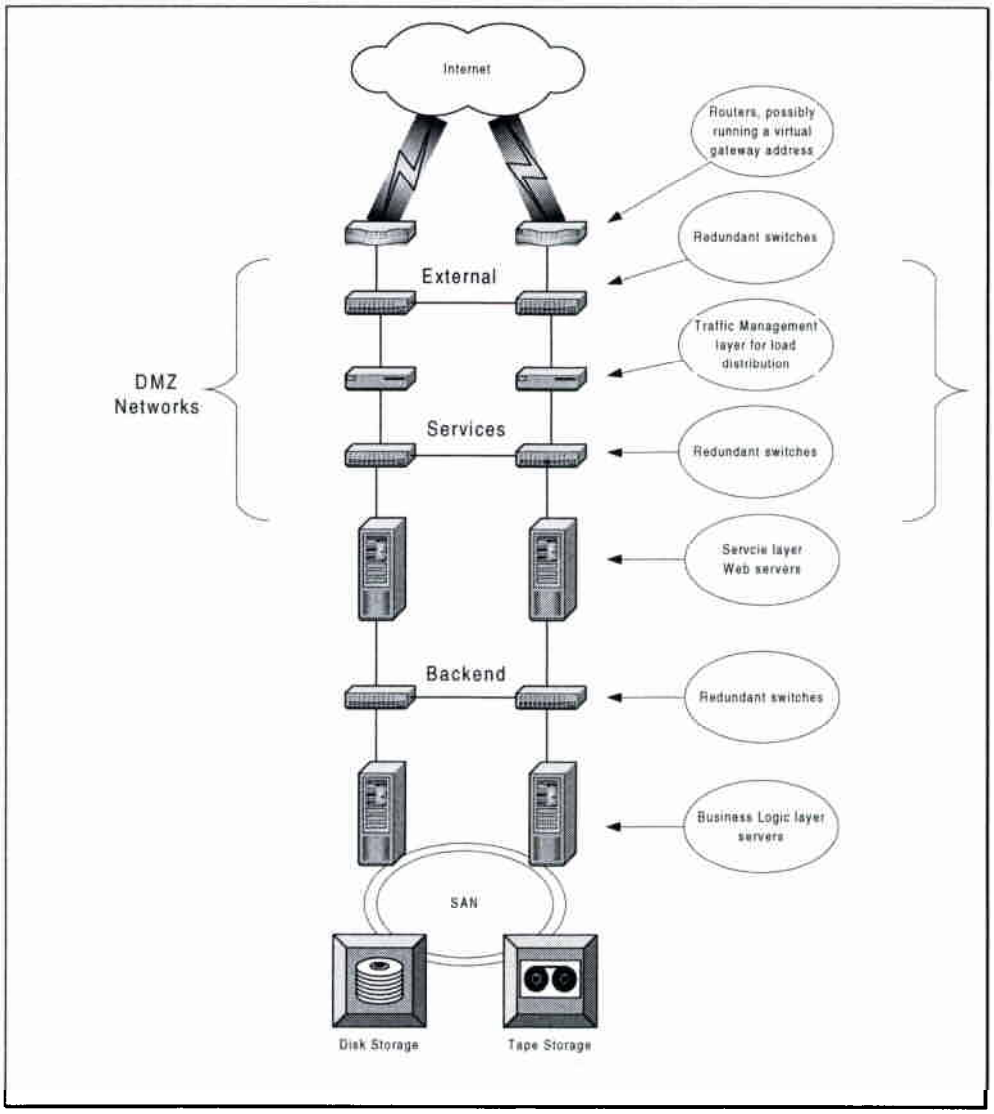

Figure 2 - Fully Redundant e-Business Architecture

Availability

Availability addresses the ability of end users to access the service whenever they need it. This implies end-to-end availability from the user's browser to the e-Business site, and includes all required backend components. It is important for those operating an e-Business site to be able to distinguish between availability problems they can fix, like application or server failures, and those they cannot, like Wide Area Network (WAN) link failures or problems at the end user's ISP (Internet Service Provider).

Redundancy

The first and most common availability option is redundancy. Carefully used, redundancy can greatly increase service availability. Opportunities for redundancy exist at all levels, from Redundant Array of Inexpensive Disks (RAID) storage devices, multiple processors and power supplies, to multiple servers, network devices and Internet connections. It is important to understand, however, that just implementing two of everything does not guarantee availability. Redundancy must be designed into the solution carefully to avoid wasted costs.

Redundancy can be implemented in two flavours: failover and load balancing.

In a failover design, the redundant components normally operate in some sort of standby state, not being used until the primary system fails. Failover can further be divided into one-to-one failover, where each primary system has a dedicated failover "twin", and many-to-one (sometimes called N+1) failover, where one system acts as a backup for two or more primary systems. Sometimes the idle system in a failover configuration can perform other functions like backups, but this often complicates the failover and recovery processes and should be avoided whenever possible. Failover is usually the most application independent form of availability solution.

In a load balancing design, multiple systems are able to simultaneously provide the same service. This is often accomplished through clustering software or some sort of mediation device like a load-balancing switch or network appliance. Often this type of redundancy requires participation at the operating system and sometimes even the application level.

Redundancy can be carried through to the network components as well. In a situation where there are two or more servers, the servers' network interfaces can be spread across redundant switches, with the switches being configured as a single Virtual LAN (VLAN). The switches can be configured with redundant power supplies and backplanes. The WAN links connecting the site to the Internet can also be redundant, even to the point of ordering the circuits from multiple carriers (after ensuring that the carriers use different physical facilities and central offices) and using more than one ISP. Multiple ISP connections can be complex, and are covered in the Lucent white paper "Minimizing E-Business Downtime: Considerations in ISP Multi-homing".

Data Replication

Data that is read only, like static web content or a read-only database table, can be easily replicated to more than one location, and many database packages have the ability to synchronize replicated copies. This way the data can be accessed even when the primary database server is down.

Power and Environmental Systems

e-Business systems cannot run without power. Likewise, if the cooling systems fail and temperatures in the machine room rise too high, servers, routers and switches will shut down. While perhaps not as interesting and high profile as other aspects of e-Business planning, it is very important to pay attention to environmental details. Physically diverse power feeds are as critical as physically diverse telecommunications facilities. Redundant fans and cooling units, each connected to redundant power sources, are as necessary as redundant disk units. Cabinet and server power cords should be equipped with twist-lock power connectors.

Un-interruptible Power Supply/Alternate Power Supply (UPS/APS) systems should be implemented to keep critical components up during power failures. The UPS systems should be tested regularly to ensure that they function correctly, and loads audited regularly to ensure that the power protection is adequate. These systems can be simple stand-alone modules or large-scale modular units. They may be able to protect against spikes and brownouts by filtering power through DC-AC-DC units, from power line to battery and back, before delivering fully conditioned power to the servers. Carefully evaluate what gets plugged in to the UPS/APS system: do you really need to have your consoles connected, or are they just an unnecessary drain?

Backup and Restore

Sound backup, retention and restore processes and procedures will have a direct impact on service availability by minimizing data loss. A well-planned backup schedule should account for the volatility and importance of the data. For example, operating system data that rarely changes may only need to be backed up monthly, while transactional databases may require frequent backups. Temporary data may never need to be backed up since it may be meaningless if restored out of context. It is as important to plan the storage of backups as the schedule. It does not help if the data has been backed up correctly but the tapes cannot be accessed during a 3a.m. emergency recovery.

Traditional tape backup products work well in many circumstances, but may fall short in cases like highly volatile databases with short or non-existent maintenance windows. Some high-end storage products allow for mirrored data volumes that allow the mirror to be split during a backup process and then reattached once the backup has completed. This allows for a snapshot of the database to be backed up, with the tradeoff that the data is not mirrored for a period of time. Some products even have a third mirror volume designed specifically for this purpose.

Taking backups without consideration to the restore process will ultimately handicap the recovery, perhaps fatally. When a business critical service is down and must be restored quickly, the last thing an administrator should be doing is inventing a restore process on the fly. Instead, he or she should be following a well-defined, documented and tested procedure. The more familiar the administrator is with the procedure, the more likely it is that the process will succeed.

Centralized backup solutions should be carefully considered, since they often require special client software to be installed on the distributed servers. In the event of a full system restore, the backup client software must be installed before the backed up data can be restored.

Security

Security plays a vital role in service availability. By now, everyone has read in the press concerning denial of service attacks, where attackers flood a targeted site with traffic to overwhelm it and deny legitimate users access. If a site is not properly secured it is vulnerable to crippling attacks. Another common type of attack replaces authorized content with pornographic, derogatory or otherwise illegitimate content. This not only denies access to the legitimate content, but can also be a public relations nightmare for a corporation.

A thorough security program will also take into account disaster recovery and business resumption planning, including not only security incidents, but as other events such as fire, flood and power failure as well. All of these events have an impact on service availability.

Scalability

Scalability addresses the need to be able to add more resources to an e-Business service as the service grows. A scalable service does not need to be redesigned to increase capacity. Incremental scalability allows for growth in small units. Though planning for this eventuality is no small task.

Internet Traffic Managers

Internet traffic managers (ITMs) take incoming requests for a service and direct them to one of several servers providing what is called, horizontal scalability.

The simplest form of ITM is round-robin DNS (Domain Name Service). DNS allows multiple IP (Internet Protocol) addresses to be associated with a single name. Most recent versions of DNS software will return the list of IP addresses sorted according to a number of simple rules, most commonly a round-robin algorithm. For example, assume that www.lucent.com is associated with three servers, A, B and C. When a system looks up the Web site in DNS, the DNS server will return the addresses of servers A, B, and C in that order. The next lookup will return the same addresses in the order B, C, A. Since most systems will use the first address returned, this distributes the traffic across the three servers. Using a

slightly different configuration, the DNS server can be made to order the addresses so that a "closer" server is returned at the top of the list. For example, a user in Canada could be directed to a Canadian server first, while a user in the U.K. could be directed to one in England. The disadvantage to this approach is that DNS replies are usually cached by intermediate servers and by the client system. Subsequent replies from the caching server will contain the addresses in the same order as before, and this will continue until the cache expires.

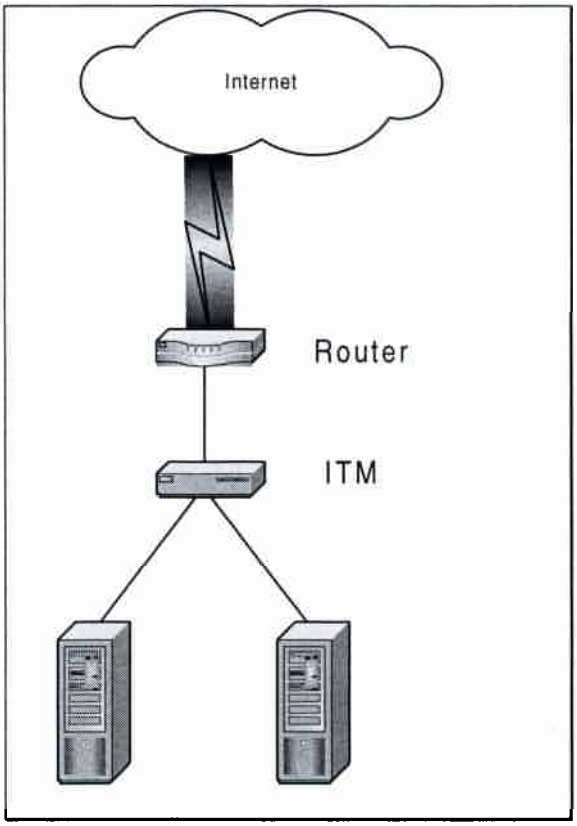

Figure 3 - Network Level ITM

The next form of ITM operates at the packet level. These network devices intercept incoming packets and redirect them to one of the servers as shown in Figure 3. Packet level ITMs frequently operate as bridges or routers, and some are able to understand application layer data. Their primary function is to distribute incoming requests across the servers according to some predefined algorithm. These load distribution schemes can be as simple as a round robin or as complex as measuring the response time and load of the servers in real time and sending traffic to the one most able to handle the load. The servers can usually be pooled by application type, speed or other characteristics and requests distributed across the pools.

A third form of ITM uses agent software installed on the target servers that measures load and other tuneable characteristics. These measurements are collected by the ITM and used in its traffic distribution decisions.

The capabilities of ITMs are many and varied, but it is important to select a product that meets the requirements of the site and the business. ITMs may also improve service availability, since they are often capable of monitoring for failed servers and applications either through internet control message protocol (ICMP) requests, or monitoring transactions for timeouts and application error messages.

ITMs can be used at any layer in the e-Business architecture, but are most commonly found in the Service layer. On the negative side, databases are notorious for not working well with ITM solutions. On

the positive side however, scaling an ITM-based site usually means adding more servers behind the ITM and updating the ITM tables.

Clusters

Clusters are groups of physical servers operating together as a single service unit that usually provides high availability as well as load sharing. There are a variety of clustering options most of which work at the operating systems level. Some use a master/slave relationship where one system is the master and it distributes work to its slaves. In this scheme there is often a backup master in the event that a primary master fails. A cluster will usually present a virtual address that the clients will use to communicate with the cluster. Often there are private network connections used by the cluster elements to communicate with each other. Generally, shared disk arrays are often required.

Some newer cluster solutions use IP broadcasts and multicasts to simulate a single fault-tolerant address to the world. In some cases, applications that are "cluster aware" can take advantage of clusters in ways that normal applications cannot.

Some applications simulate cluster behaviour by distributing work across a set of backend servers at an application level. These are usually capable of recovering from server failures by redirecting work as required.

Cluster solutions are usually well suited to database and data warehouse applications. They can be scaled by adding more cluster elements to some maximum number of processor nodes.

Network Attached Storage

Network attached storage (NAS) products are generally best suited for departmental use, although they can be found in e-Business applications as well. These products are essentially dedicated servers with a network connection and lots of reliable storage. The application servers access the NAS servers across the network using standard access protocols like network file system (NFS), disk file system (DFS) or NetBIOS shares. The primary advantage of a NAS solution is to make the storage pool easier to manage by providing platform independent centralized storage.

NAS servers are not limited to disk storage. They can support optical storage, compact disks, tape and any other type of media desired.

Storage Area Networks

Storage area networks (SAN) are dedicated networks designed specifically for high-speed data connection. Traditionally, data devices have been connected to servers through direct channel connections like SCSI. While this is effective for connecting a disk or array to a single server, or possibly a failover pair of servers, it begins to become problematic since these devices generally do not implement record locking mechanisms, or when two or more systems must write data simultaneously.

Networked file systems such as NFS and DFS can implement data locking and threading. In these schemes, a file server will allow multiple application servers to access its data across a network. If two application servers attempt to write the same record at the same time, the file server can sequence the writes. Unfortunately, the record locking capabilities of these protocols is often crude. In addition, the servers themselves become a bottleneck as the load increases.

SANs surfaced a few years ago in an attempt to solve these and other problems. A typical SAN consists of one or more storage devices, usually with RAID capabilities, connected to multiple servers through a dedicated network. Recently those networks were built with optical fiber, which greatly increases data transfer speeds and reduces distance limitations.

Storage area networks are well suited to high-end database and data warehouse applications, particularly where several systems may need to write data simultaneously. SAN products will usually scale to very large capacities. Some SAN products allow mirrors in remote locations for disaster recovery.

Proxy and Caching Servers

Proxy servers can be used in e-Business applications in two ways. The first places a proxy server between the Internet and the e-Business servers. The proxy server will cache frequently used content and serve it to users offloading this processing from the true servers behind. This technique is really only of value for static content. By its nature, dynamic content does not lend itself well to caching. In addition, the proxy server itself may become a bottleneck.

The second way a proxy may be used is to cache content close to the end users. Some service providers offer distributed caching services. In this case, if the primary e-Business site is located in Canada but there is a large user population in the U.K., caching servers located in the U.K. may improve the response time of the service. They may also be able to service requests when the primary site is down or unreachable. Once again, the effectiveness of this solution is heavily dependent on the amount of static content.

Summary

Architecting the services, systems and networks to provide complementary availability and load balancing measures at each level will contribute to the success of any mission critical services. Before starting, however, it is important to first determine the real business requirements and ensure that the implementation supports them appropriately. Prior to implementation, begin determining the business requirements by answering these simple questions: Is the objective to provide a 24/7 service even at the cost of degraded performance, or is performance the critical attribute? How accurate are the growth projections and how well do they fit the scaling mechanisms in your design? What is your budget, and what is the projected cost of downtime?

Once the business requirements have been established architecture can be developed that will guide the design of the applications, systems and networks. The architecture should provide both mid term and long term goals as well as a roadmap for achieving them. Application, system and network designs, testing and implementation plans will flow from the architecture. Operational processes and procedures can be developed during the design and testing phases that will ensure a reliable and sustainable e-Business site.

Conclusion

Proper planning for all aspects of your e-Business service, including environmental systems, security, storage, connectivity and server hardware, is a necessity to provide true high availability. Carefully consider all of your e-Business requirements and evaluate them against the myriad of configurations and products available to you. Be sure to plan your availability design as a co-operative and integrated solution rather than in isolated pieces, and always implement a solution that fits the needs of the business.

Each of the solutions discussed in this paper has its associated trade-off in costs (both capitol and administrative costs), complexity and effectiveness. There is no "one size fits all" answer. Lucent Technologies Worldwide Services brings a large body of experience in developing cost-effective and appropriate solutions matched to your business needs and technical requirements, and in addressing the many aspects of such a solution, including business requirements, security, network and system performance and availability, storage systems, and operations.

For additional information, visit our web site at www.lucentservices.com or call 1-877-767-2988 in the U.S. or 1-650-318-1020 outside the U.S

Copyright © 2000, Lucent Technologies Worldwide Services

This is an unpublished work protected under the copyright laws. All rights reserved.
WP.EN.EBUS.1200