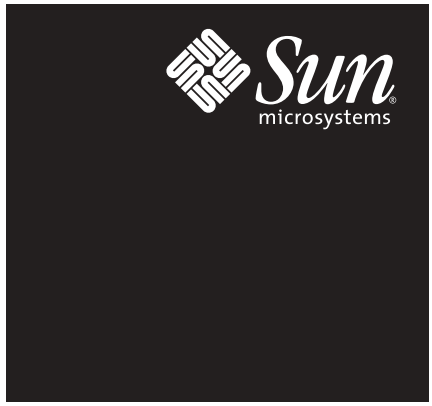
# Sun™ Infrastructure Solution for Enterprise Continuity

Higher Availability at Greater Distances for Reduced Business Loss
April 2003



# Table of Contents

Chapter 1

# Executive Summary

When an application or server is damaged or destroyed, data loss is inevitable. Unless the lost information is recovered seamlessly, system failures costing as much as 6.4 million dollars per hour (source: contingencyplanningresearch.com) may occur. In an effort to minimize the financial impact a system outage has on a company's shareholders, law makers are enacting legislation that will establish more exacting standards for disaster preparedness. For forward-thinking businesses, the challenge is clear: reduce risk while minimizing the cost of compliance.

In the past, companies were forced to rely on inefficient or incomplete disaster recovery strategies that resulted in recovery times as long as seven days (source: Nortel Networks). Even powerful technologies, such as storage clustering, could only partially address this challenge. And, while clusters did improve recovery time by enabling a data center to reconstruct lost information using replicated data stored on redundant machines, the maximum effective distance between clustered devices was limited. As a result, mission-critical data remained vulnerable to major disasters that affected the entire facility.

Recognizing these challenges, Sun began approaching data management in an entirely new way. Rather than relying on any one storage technology operating independently, Sun built an extensible storage solution capable of running several applications (for example, point-in-time copy, remote replication, and remote archiving to tape or disk) in combination. The solution, based on the powerful N1 architecture, reduces enterprise storage complexity by combining multiple data centers into one virtual data services delivery system. And, because N1 is a fully transparent open architecture, information can be shared among multiple devices more efficiently — even when running across heterogeneous systems. If a service interruption does occur, the

virtually connected storage devices can identify, reallocate, and restore vital data at a moment's notice.

As development continued, Sun data management specialists expanded the solution to accommodate businesses that wished to increase data availability over longer distances. Employing Dense Wave Division Multiplexer (DWDM) technology and an active/active data continuance configuration proven to reduce time to recovery, the designers successfully created a powerful new business continuity model that enables companies to locate cluster nodes in separate locations within a wide geographic area (up to 200 km apart) — all without changing the software infrastructure, applications, or data.

The Sun Infrastructure Solution for Enterprise Continuity includes:

- A complete data management model
- Vital infrastructure components
- High-performance servers
- Powerful management software
- Storage, disk, tape, and network solutions from the StorEdge ONE solution portfolio
- Proven best practices and methodologies provided by Sun Services

When disaster strikes, will you be ready? With Enterprise Continuity solutions from Sun, your preparations are underway.

Chapter 2

# Enterprise Continuity Overview

The importance of information technology (IT) is underscored in a disaster situation. Even a seemingly insignificant localized event can damage or destroy mission-critical data, interrupt business operations, and interrupt vital transactions — thereby exposing the affected corporation to significant financial risks. To ensure effective disaster preparedness, companies must implement an open and extensible architecture that is flexible enough to meet today's business needs and support tomorrow's service demands.

## A Short History of Data Storage

Data and application services are constantly being adapted to meet new challenges. The high availability evolution began when storage experts began mirroring data across disks and storage systems to protect against storage hardware and software defects. Later, companies started clustering multiple servers to shelter applications from server failures. More recently, companies have begun focusing on storage and server consolidation to prevent excessive downtime and enable non-destructive growth.

Consolidation of server and storage infrastructures enables businesses to centralize enterprise data while providing common access and scalable growth. For these reasons, disconnected or geographically scattered servers are commonly centralized using a Storage Area Network (SAN). Although this approach helps ease the burden of data management, it does not provide high availability at every level; large high-end servers and storage disk arrays are still vulnerable to physical disasters (earthquakes, fires, and explosions, for example) that can remove access to all of a corporation's data.

To reduce their vulnerability in the event of a disaster, many companies have started implementing backup sites with identically configured stand-by servers. However, to be viable, a backup site must be located 50 km or more from the primary site. At smaller distances, the backup site may be affected by the same event that disabled the primary site.

Until recently, cabling and latency issues have limited real-time data replication and application failover to secondary sites located within 80 km of the primary site. Beyond this limited distance, backup systems have been unable to provide real-time recovery — presenting the possibility of critical data loss. Moreover, the current economic climate is forcing many companies to re-propose secondary site servers. As a result, many formerly passive backup machines are now actively providing services to users, generating additional complexity for potential recovery situations.

## The Problem

When critical applications containing vital data are damaged or destroyed, the business impact is often resounding. The downtime that accompanies a system failure can cost companies as much as 6.4 million dollars per hour (source: contingencyplanningresearch.com). Moreover, in the event of a prolonged outage, customer confidence may be shaken — leading to further revenue loss that can significantly reduce overall stock value.

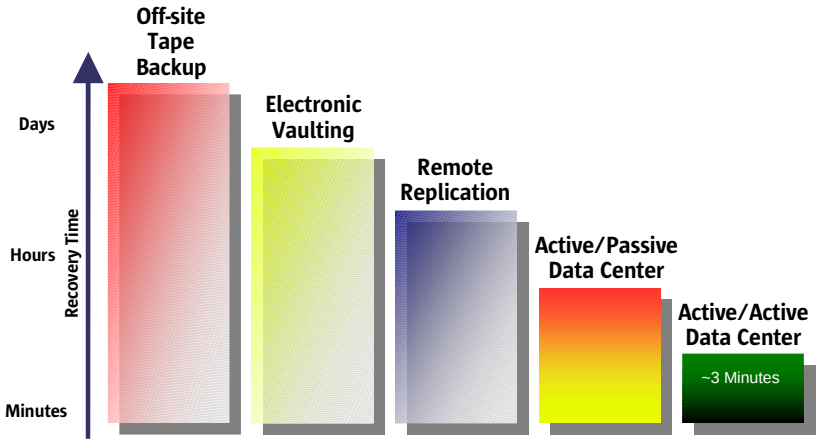
In some business sectors, such as finance, federal regulations require geographically dispersed backup sites to help ensure that certain economy-related functions can continue in any situation. Furthermore, insurance companies will soon raise premiums for companies that do not meet rigorous business continuance standards, making it more crucial than ever for businesses in any industry to implement extended SANs and distance clustering solutions to provide real-time recovery and failover.

## The Solution

Anticipating the introduction of new disaster preparedness requirements and legislation, Sun has developed and validated a complete business continuity infrastructure and support services solution that enables companies to support SANs and clustering to distances up to 200 km apart.

The Enterprise Continuity solution supports an active/active SunPlex™ environment based on the N1 architecture from Sun (see Figure 2-1). Customized Enterprise Continuity solutions may also include other continuity strategies, including off-site tape backup, electronic vaulting, remote replication, and active-passive data centers. Leveraging the flexibility of this infrastructure, multiple data centers are transformed into one highly available data services delivery system.

Figure 2-1: The Enterprise Continuity solution's active-active configuration provides the shortest time to recovery




The Enterprise Continuity solution helps ensure that data is always available at one site in the event of an outage at another site (see Figure 2-2). Multiple sites act as one geographically dispersed SAN, while the SunPlex cluster acts as a unified application cluster. These aspects of the Enterprise Continuity solution help reduce the possibility of unplanned downtime and completely eliminate time-consuming restart procedures.

The Enterprise Continuity solution also helps businesses save money. Using the N1 architecture, businesses can allocate networked storage devices as needed — eliminating the need for additional IT investments in replication hardware and software. And, by allowing clustered devices to be separated over greater distances, the Enterprise Continuity solution enables corporations to locate secondary data centers in less costly real estate markets. Moreover, the solution's optical technologies can reduce costs by as much as 35 percent in the second year of deployment, when compared to older solutions that use leased telephone lines.

**Leveraging Best-in-Class Sun Partners**

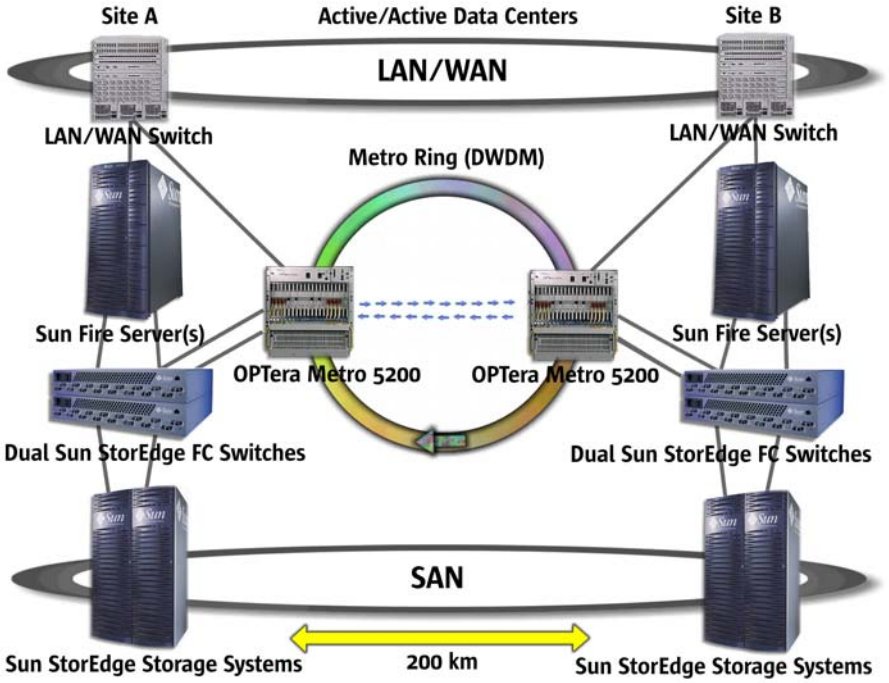
By collaborating with best-in-class partners, Sun delivers powerful end-to-end Enterprise Continuity solutions — including architecting, installation, and management services — capable of meeting the exacting standards being proposed in regulatory circles. For example, Sun is collaborating with SunGard Availability Services, a Sun strategic partner, to create customized Enterprise Continuity solutions for business critical environments in the financial sector.

One of the first companies to achieve SunTone$^{SM}$ certification for disaster recovery services, SunGard runs high availability solutions on Sun systems to help financial institutions reduce information loss, minimize business risks, and maintain revenue streams in the event of an unplanned system outage. In addition, SunGard incorporates Sun StorEdge™ Availability Suite software, Sun StorEdge 9900 TrueCopy technology, and the Enterprise Continuity solution to extend data storage across geographically distant buildings or campuses — enabling compliance with new precautionary regulations.

The benefits of the Sun-SunGard alliance are apparent to SunGard customers as well. Brett Paulson, chief information officer for the Chicago Board of Trade Clearing Corporation, says, “We are very confident that our selection of SunGard as a partner for business continuance will be pivotal to the continued growth of our business. Key to our decision was SunGard's maximum

uptime capabilities coupled with their strong collaboration with Sun Microsystems who already provides a powerful infrastructure for our IT needs. We anticipate that the SunGard and Sun combination will help us further enhance the quality of service to our customers, which translates into higher return-on-investment for our business."

Figure 2-2: Enterprise Continuity using an extended SAN and unified application cluster

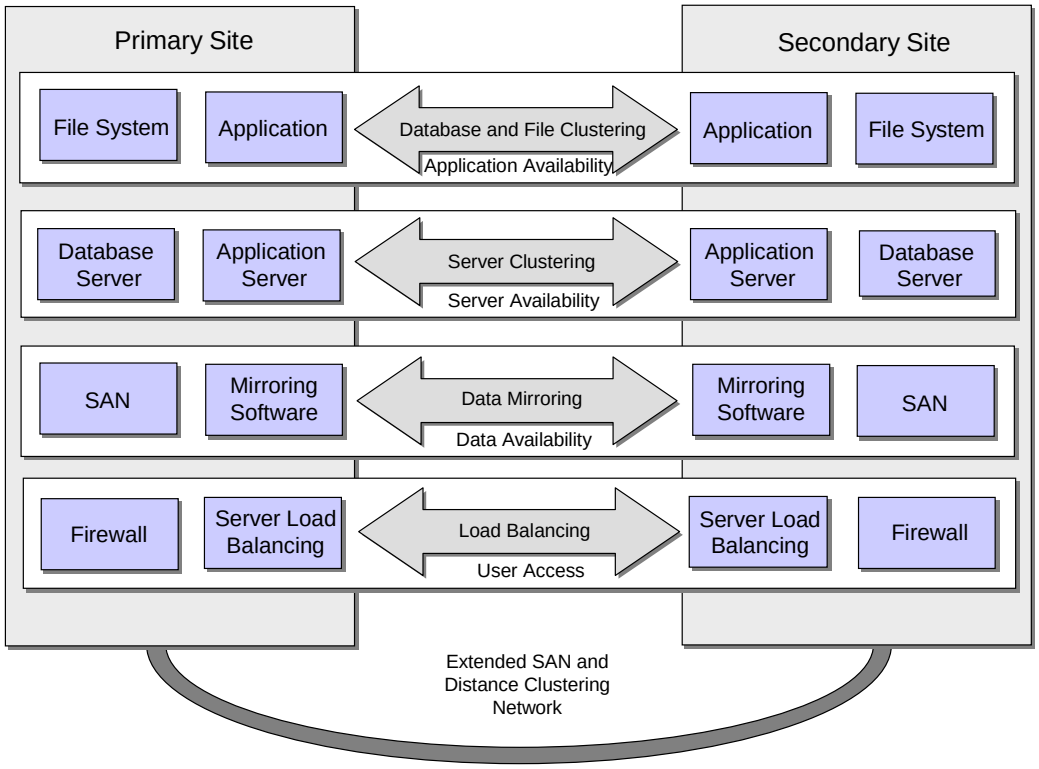



The Enterprise Continuity solution supplies a complete array of products and processes (see Figure 2-3). Primary solution components include:

- *Sun Open SAN environment.* A total SAN solution, the Sun Open SAN environment integrates storage management, storage resource management, host bus adapters (HBAs), interconnect, storage, and backup devices. When implemented in a data continuance environment, the SAN solution helps enable customers to achieve uninterrupted operations using consolidated storage resources, centralized management, faster backups and recovery, and high-availability access to mission-critical data.
- *Sun Cluster 3.0 software*. Sun software delivers high-availability application services to the data center. Sun Cluster 3.0 extends the Solaris™ Operating Environment (Solaris OE) seamlessly across a tightly coupled clustering model to achieve distances of up to 200 km.
- *A DWDM multiservice platform.* Provided in this instance by Nortel Network's OPTera Metro 5200, the DWDM technology enables clustered servers to interact at high speeds from geographically dispersed sites.
- *A complete, pretested solution design.* Sun delivers the base technology, architectural solution design, deployment consultants, and support services to help ensure rapid and thorough deployment.
- *Business continuity services.* Sun offers business continuity services for every phase of the IT life cycle — from deployment to implementation and maintenance.

- *SunVIP℠ program.* The Sun Vendor Integration Program (SunVIP) provides customers with a single point-of-contact when challenges arise with their Enterprise Continuity solution. Sun collaborates with third-party vendors (for example, Nortel Networks) to mutually identify and quickly resolve complex technical problems, saving customers' time and resources.
- *Comprehensive portfolio of learning products.* Sun helps optimize internal IT expertise with customizable consulting and learning path programs.

Figure 2-3: Enterprise Continuity provides high availability at every level




This white paper provides guidelines for architecting and implementing a successful business continuance solution. In the following chapters, we will discuss several topics. Chapter 3 covers the data and storage availability aspects of the architecture. In Chapter 4 we will discuss server and application availability, using Sun Cluster 3.0 software. Networking technologies are examined at length in Chapter 5. Chapter 6 delineates the people and processes necessary for implementing and managing the solution. Finally, Chapter 7 provides technical details discovered during the configuration and tests used to validate the solution.

Chapter 3

# Data Availability and Consistency

In today's global e-commerce environment, businesses rely on computing functions and electronically stored data to assure that day-to-day operations continue profitably and without interruption. Furthermore, the information obtained by interpreting this data can be translated into a competitive advantage. By developing a highly available business continuity solution, any business can protect essential information from corruption and failure while helping to ensure that it remains consistent among sites.

Consolidation of storage resources onto SANs helps make these tasks easier and more cost-effective. SANs completely separate the server/storage network from the LAN. Acting as a separate network behind the server, the SAN is isolated from the LAN and well equipped to transfer data among disks, tape devices, and servers. SANs use Fibre Channel (FC) network transport, and move data to and from disks via the SCSI-3 (serial SCSI) protocol. By combining high-speed Fibre Channel transport with existing SCSI protocols, system architects can enable greater bandwidth, larger capacity, high availability, and enhanced manageability for data storage networks.

Sun StorEdge Open SAN Architecture is an integral part of the Enterprise Continuity solution, as demonstrated in Figure 3-1. The Sun solution provides a complete data management model, including vital infrastructure components, servers, management software, the StorEdge ONE solution portfolio, storage and tape solutions, as well as Sun education and consultancy services. In addition, the architecture supplies the components necessary to protect, consolidate, and share data — helping to ensure consistency between sites.

The following sections provide more detail about the four steps for managing and protecting data in an Enterprise Continuity solution: data protection, storage device consolidation, data

sharing, and data consistency. Each section will outline ways in which businesses can employ Sun solutions to implement a specific step.

Figure 3-1: Sun StorEdge Open SAN Architecture


## Protecting Data from Hardware Failures

The first step in managing data resources for Enterprise Continuity is to protect the data against storage hardware failures and data corruption. Typically, a storage resource failure results in the replacement of that component. Next, the affected data must be restored. And, because recent changes are frequently lost, a time-consuming (and expensive) recovery process may be required for operations to continue.

The best way to protect data from hardware failures is to use highly available hardware and redundant storage of data. Using Redundant Arrays of Independent Disks (RAID), additional parity information can be stored and spread across multiple disks. That way, if any single disk were to fail, the information could be easily reconstructed from the redundant data. With the help of additional disks, companies minimize the window of time occurring between disk failure and data recovery.

Sun offers a complete line of highly available SAN storage systems and Sun StorEdge Traffic Manager software designed for high data availability. The Sun product line includes:

**Sun StorEdge T3 Array**

The Sun StorEdge T3 array is particularly well suited for databases, network file systems (NFS), Web serving, e-mail applications, and companies facing rapidly growing storage capacity on disparate host platforms. Key features of the Sun StorEdge T3 arrays are:

- Complete Fibre Channel architecture
- RAID storage for RAID 0, RAID 1, RAID 5, or a combination of two RAID levels

- High-availability, redundant-failover, dual-controller design
- Three-way scalability providing linear performance, massive capacity, and consistent reliability

**Sun StorEdge 3900 Systems**

The Sun StorEdge 3900 systems are high-performance storage platforms designed for single- or clustered-server environments. A natural step beyond individual and partner pairs of Sun StorEdge T3 arrays, Sun StorEdge 3900 systems can be used to run dedicated applications focused on different kinds of hosts in a homogeneous environment. The key feature and benefits of the Sun StorEdge 3900 systems are:

- Factory integration and pre-testing
- Technology based on Sun StorEdge T3 array
- Pre-configuration for RAID 5 with standby hot spare
- Host-managed multipathing
- Multiple host support
- Optional phone-home capability

**Sun StorEdge 6900 Systems**

The Sun StorEdge 6910 and 6960 systems are easy to order, deploy, and manage. Both systems are pre-tested and offer Sun customers an integrated storage solution for SAN environments that includes:

- Integrated FC switches
- Sun StorEdge T3 arrays
- Centralized configuration and monitoring
- Storage virtualization layer
- Optional phone-home capability

**Sun StorEdge 9900 Systems**

The Sun StorEdge 9970 and 9980 systems are designed to provide high-speed response, continuous data availability, scalable connectivity, and expandable capacity for open and IBM S/390 environments. Intended for use in 24 x 7 data centers, these systems enable high-performance non-stop operations. Moreover, the Sun StorEdge 9970 and 9980 systems are compatible with industry-standard software and support concurrent attachments to multiple host systems and platforms.

Large SANs can be configured to heterogeneous hosts using a powerful tool called “host groups.” Implementing a novel virtualization architecture, the Sun StorEdge 9900 systems can allow multiple host groups to use the same physical port. First, the systems identify the host type by its World Wide Name (WWN) and return the correct SCSI response to each host. Next, host groups, created via Host Storage Domain (HSD), can be assigned volumes to share. Volume access is then controlled by Sun StorEdge 9900 SANtinel software, which allows specified users to restrict access to host groups or a specific host WWN.

The Sun StorEdge 9970 and 9980 systems can operate with multiple-host applications and host clusters and should be employed in the following environments:

- Very large databases
- Data warehousing
- Data mining applications that store and retrieve TBs of data
- Multihost data center applications

**Sun StorEdge Traffic Manager Software**

Sun StorEdge Traffic Manager software (Sun StorEdge TMS) manages devices that are accessed by a host through multiple paths. The application includes multipath configuration management, I/O load balancing across HBAs, failover instructions for interconnect and controller failures, and single-instance multipath devices. With Sun StorEdge TMS, a multipath device can be represented as a single instance, rather than as multiple instances occurring on multiple devices. Sun StorEdge TMS is designed to provide the following benefits:

- Superior SAN-wide failover and load balancing
- Simple, easy-to-administer storage growth

## Protecting Data from Corruption and Destruction

Data corruption or destruction usually results from user error or application failure. These difficulties are compounded in a networked storage environment where information is replicated almost instantaneously because data at secondary sites is often corrupted as well. And, if transaction logs fail to retrieve the original data, IT staff must laboriously retrieve this information from backup media. To help protect against errors, Sun offers tape backup and near-line storage solutions specially designed to help minimize the time spent and money lost restoring data.

**Sun StorEdge L6000 Tape Library**

The Sun StorEdge L6000 tape library is a high-performance device that is designed for data center environments. An extremely scalable solution, it is a perfect complement to Sun Fire™ 15K servers and Sun StorEdge 9900 series storage systems. What's more, the Sun StorEdge L6000 tape library can support up to 144,000 cartridges with 24 silos, enabling extremely rapid systems backup.

**Sun StorEdge Utilization Suite Software**

Sun StorEdge Utilization Suite software provides an additional level of functionality for the Enterprise Continuity solution. Its recovery and archive service acts as an extension to online storage by accessing data and copying it to multiple distributed storage devices simultaneously at or near raw device speeds. Following data distribution, the software automatically protects work in progress and creates duplicate archived files and file systems on multiple servers. In this way, an enterprise can return to full productivity quickly in the event of unplanned downtime. The benefits of the suite are:

- Cost-effective storage resource utilization
- Superior data protection
- Rapid recovery from unplanned downtime
- Protection of IT storage investments
- Increased operational efficiency via sophisticated volume management and policy-based administration
- Increased quality of storage service levels
- Seamless user experience

## Consolidating Storage Resources

Increasingly, companies seeking to reduce the complexity and cost of storage management are turning to storage consolidation solutions (for example, SANs). Consolidation of resources enables enterprise storage systems to become more efficient and cost effective, while making it easier for

businesses to maintain consistent data among remote sites. In addition, storage consolidation can help minimize service disruptions, simplify data sharing, and lower management costs.

Consolidation solutions effectively eliminate the need for sprawling storage networks that are dispersed throughout an entire building or campus. All storage components are housed at a central site, making it easier to administer, maintain, and share. Moreover, consolidated storage systems offer redundancy and data protection methods to help protect the data from hardware failures. WAN users, who experience only minimal changes in response time, benefit from storage consolidation by connecting to the SAN via high-speed 100 Mbps links.

Management of storage as a pool of resources allows companies to better utilize system components. Typically, disks that are attached to single systems are under-utilized — often containing gigabytes of unused space. However, in a consolidated solution, businesses can manage and use resources more effectively, providing or limiting storage for users and applications as needed.

The Enterprise Continuity solution leverages the powerful N1 architecture to provide complete application transparency, enabling companies to refine performance, resource utilization, and overall availability. Using N1, businesses can configure heterogeneous devices from Sun and other vendors to work as a single virtual data services system. Formerly isolated resources are linked, creating a pool of virtual resources that can be rapidly reallocated to match performance requirements.

Using high-speed interconnects, the Enterprise Continuity solution connects SANs from site to site. In this environment, tape library resources can be consolidated and located at a separate site, eliminating the need to ship tapes to remote locations. This way, if an outage occurs, data will already be stored remotely. If necessary, secondary sites can access this information and restore corrupted or destroyed data.

Sun makes consolidation of storage resources easier and more cost-effective by offering large, highly available storage systems and a complete suite of open SAN management software. Businesses that implement a Sun solution can allow more systems to utilize resources, thereby increasing device availability and decreasing total cost of ownership (TCO). What's more, Sun consolidated tape resources require fewer IT personnel for backup data management, providing an additional cost savings.

**Sun StorEdge Enterprise Storage Manager Software**

Sun StorEdge Enterprise Storage Manager (Sun StorEdge ESM) software is the industry's first WBEM/CIM-compliant open SAN management tool. The Web-Based Enterprise Management Common Information Model (WBEM/CIM) is an open-standard, extensible technology for managing heterogeneous devices and SAN environments. Using this technology, businesses benefit from greater adaptability that will enable them to support future storage technologies.

WBEM/CIM compliance is one part of Sun's ongoing Storage ONE initiative. Working under this guiding principle, the Sun StorEdge software family continues to offer open standards-based storage management solutions. These solutions effectively lower TCO and improve application service levels by streamlining operational activities and simplifying the complexities associated with SAN management.

Sun StorEdge ESM software is designed to provide a centralized management platform that consolidates core storage management services through a Web-based common view. It also helps customers increase uptime and improve service levels by delivering a consolidated presentation of three major areas of storage management, including:

- *SAN topology management and reporting services.* By enabling an enterprise-to-enterprise visualization of the SAN infrastructure, Sun StorEdge ESM software provides a central point of access and control for servers and network storage devices, such as storage arrays, fabric switches, and Fibre Channel HBAs.
- *Device configuration management services.* Using a single management console, Sun software provides auto discovery and scalable wizard-based configuration services for Sun StorEdge storage systems and third-party devices.
- *Health monitoring and intelligent diagnostics management services.* Sun StorEdge ESM software supplies proactive health checking, fault isolation, and expert-based advice. (Customers who do not require the full functionality of ESM software are granted free access to the Sun StorEdge Health Diagnostic Expert, a stand-alone diagnostic tool available for download online.)

**Sun StorEdge Resource Management Suite Software**

Sun StorEdge Resource Management Suite software offers a comprehensive array of Web-based management applications designed to provide visibility into the storage infrastructure. In this way, businesses can more effectively manage, implement, and maximize SAN storage investments. Sun StorEdge Resource Management Suite software provides centrally managed reporting and resource monitoring using the following tools:

- Capacity Reporter automates the discovery, monitoring, alerting, and reporting of key storage resource statistics.
- Database Reporter facilitates the discovery, analysis, and management of database storage environments.
- File Reporter generates detailed file-analysis storage reports.
- Global Reporter provides a single management view of global storage resources.

Sun Storage Resource Management Suite software also helps maximize storage network investments, increase operational effectiveness, and increase IT staff efficiency through the following functions:

- Automated network storage discovery
- Historical data analysis and trending
- Simplified charge-back to users and groups
- Resource identification and allocation
- Global storage resource consolidation
- Support for data classification and migration projects

## Sharing Storage Resources and Data

Data sharing technology is redefining storage. By exploiting the power of connectivity, a shared storage environment can exist as a complete storage system in its own right. These enhancements are particularly prominent in traditional systems where storage scope is limited or where management capabilities are distinct from storage systems. In either scenario, shared storage can bring together unbounded collections of interconnected and geographically distant storage resources and management capabilities to provide the following benefits:

- Reduced TCO for storage solutions via more intelligent storage resource management (for example, three servers sharing a common storage pool do not require spare storage to accommodate growth; instead, they share available backup resources)

- Enhanced performance through network-enabled storage management approaches (for example, “server independent” storage management, consolidated storage management, DataMovers, third-party copy)
- More powerful shared storage environment leveraging performance software, highly available storage systems, and open SAN management solutions from Sun

**Sun StorEdge Performance Suite Software**

Sun StorEdge Performance Suite software is a high-performance, massively scalable shared-file system service. It is intended to provide rapid access to shared data — even while processing large files such as satellite pictures, magnetic resonance images (MRIs), check scans, and automotive designs. The software is designed to require less hardware, reduce administrative staff requirements, and provide better data performance. Essential features of Sun StorEdge Performance Suite software are:

- File sharing among multiple readers and writers
- Policy-based administration
- Exceptional performance and scalability
- Rapid recovery
- Integrated and sophisticated volume management
- Heterogeneous host support using third-party application program interfaces (APIs)

## Enabling Data Consistency Across Multiple Storage Devices

Data consistency can be a challenge, especially when information is spread across primary and secondary volumes located on multiple storage subsystems. To ensure data integrity following an outage, it is essential that a business be able to perform a normal database restart. The purpose of the Enterprise Continuity solution is to help make sure that a copy of the data recorded at site one is generated and stored at site two regardless of what happens at the first site.

*Data consistency* comes into play when the secondary data source is created. It is a term used to describe data that has been properly replicated from a disk situated at one site to a disk in a second, separate location. For the second disk to contain consistent data, the copied information must contain all updates that occurred before a specific point in time, but no updates beyond that point.

When the secondary data image is data consistent, applications can be restarted in the secondary location without having to undergo a time-consuming data recovery process. This process can involve restoration of image copies and logs, as well as the execution of forward-recovery utilities to apply updates to the image copies. Because applications only need to be restarted, an installation can be up and running quickly — even if the primary site has been rendered totally unusable.

The Enterprise Continuity solution employs remote mirroring and point-in-time copy to implement data consistency among sites. The solution uses Sun StorEdge Availability software to enable mirroring and point-in-time copy. All subsequent steps, such as allowing volumes to fail over from the primary site to the remote site and then fail back to the primary site once it has been restored, are the domain of the VERITAS Volume Manager software.

**Replicating Data (Mirroring) using Fibre Channel**

High-end storage products offer the capability to replicate data to a remote storage subsystem of the same type using direct connections that do not affect the servers attached to the storage. This

replication usually copies data blocks to a remote site, because maintaining two or more copies of synchronized data provides a basis for fast, automatic failover. However, this approach can be risky because any defect in the original data is synchronously mirrored to the copy —making both the original and the copy useless. To prevent an occurrence of this kind, it is essential that companies create a non-synchronous copy of data, such as those created with backup and restore.

One common tactic is to send logical data packets, such as database logs, to the remote site where they can be applied to the database after a time gap has elapsed. This alternative makes it possible to detect errors — logical administrative errors, as well as hardware-related inconsistencies — and prevent them from being applied to the remote copy. Another option is to replicate database files directly using live data, or indirectly using a snapshot made with point-in-time copy.

**Sun StorEdge Availability Suite Software**

Sun StorEdge Availability Suite software contains robust software that helps create a wide range of protection and recovery solutions for full-scale business continuity. Sun StorEdge Availability Suite remote mirroring software supplies a media independent, IP-based, live data replication software utility that runs on the Solaris OE.

Sun StorEdge Availability Suite remote mirroring software also allows businesses to replicate data volumes over the network from data center sites located anywhere in the world, minimizing the data disruption that results from unplanned downtime. Key features and benefits of the remote mirroring capabilities of this software suite are:

- Synchronous replication of the remote image (updated before the I/O complete signal is sent to the originating host)
- Asynchronous replication of the remote image (updated after the I/O complete signal is sent to the originating host)
- Optimized re-synchronization and change tracking
- Media-independent IP connection
- Storage-independent local and remote servers
- Mutual backup capability for remote sites
- Disaster simulation software
- Remote replication and logging functionality during re-synchronization

The point-in-time capabilities of Sun StorEdge Availability Suite software enable offline backups that help recover corrupted or destroyed data. Using a point-in-time image of online data, backups can be created during online processing using any medium. In addition, point-in-time imaging can also significantly reduce time to recovery when restoring from a backup. Sun StorEdge Availability Suite software makes it easy to take frequent snapshots of data that can then be restored, when required, in minutes rather than hours. Key features of the point-in-time capability are:

- Point-in-time images generated at high speeds
- Master quickly re-synchronized from point-in-time image
- Master rapidly restored from point-in-time image
- Master and point-in-time image span storage subsystems
- Master protected on any supported RAID level
- Point-in-time image protected on any supported RAID level (does not need to match master)
- Resulting image usable for virtually any operation

Chapter 4

# Server and Application Availability

With data separated and protected, the next logical step is to design an Enterprise Continuity solution that can provide highly available servers and application services. Today's globally networked economy demands powerful information systems that are available to customers, business partners, and employees around the clock and around the world. When a company's application is down, the business costs can be enormous. Productivity, sales and profits, customer service, and customer loyalty are all negatively impacted.

## Maximizing Availability with Sun Servers

When advanced systems availability and productivity are critical for competitiveness, highly available applications are required. To provide continuous application availability, the user must be protected against failures occurring in both the server hardware and application software. To protect against individual server failures, two types of redundancy are required:

- Component redundancy — providing fault resilience and tolerance by automatically failing over to redundant components within a single server
- Server redundancy — exchanging a primary server with a secondary machine if the primary server fails

High availability servers from Sun offer a high degree of component redundancy and provide reliability, availability, and scalability (RAS) features that make them excellent for serving highly available applications. Server redundancy is implemented using Sun Cluster 3.0 software.

**Sun Fire 15K Server**

The Sun Fire 15K server is designed to provide rock-solid availability and hot-swappable components. A highly available machine with the flexibility to adapt to application needs, the Sun Fire 15K server is the ideal environment for mainframe application re-hosting, server consolidation, databases, and high performance computing (HPC) applications. Key specifications of the Sun Fire 15K server include:

- Up to 106 UltraSPARC® III Cu 900 MHz processors
- More than 1/2 TB of memory
- Up to 18 fifth-generation Dynamic System Domains (fully configurable while applications are running)
- Hot-swappable CPU/memory boards with Uniboard design (common across the Sun Fire server family)
- Redundant, high-performance Sun Fireplane Interconnect (up to 172.8 gigabits per second of peak bandwidth)
- Fully-redundant power and cooling systems

Equipped with an UltraSPARC III processor, the Sun Fire 15K server has the near-linear scalability necessary to provide high performance for transaction-intensive environments. The server also supports a series of fifth-generation Dynamic System Domains that isolate applications from potential faults for near-continuous application availability. High system availability is further secured by using the Solaris OE and redundant Sun hardware components.

Sun servers help can save businesses money in several ways. The Uniboard CPU and memory boards, interchangeable among Sun Fire mid-frame and 15K servers, help lower total cost of ownership. Further savings can be generated using the server's centralized management capabilities — a feature that can be extended to include other platforms to lower administrative effort and cost. Companies may also protect existing investments in hardware and software using the binary compatibility common in all members of the Sun Enterprise and Sun Fire server families.

**Sun Fire 3800-6800 Servers**

This line of mid-range servers supplies tremendous computing power. Ideal for mission-critical availability and advanced system management tools, the Sun Fire 3800, 4800, and 6800 servers are tailored for databases, business applications, and e-commerce.

**Sun Fire 280R Servers**

The Sun Fire 280R server is a reliable, scalable, and rack-ready server powered by UltraSPARC-III technology, ideal for service provision, e-commerce, financial services, and branch office automation.

**Netra™ Servers**

The Netra server family is designed for demanding users who require Network Equipment Building System (NEBS) compliance. Netra servers are high-availability platforms capable of supplying considerable processing power. Netra servers are ideal for customers deploying Web servers, firewalls, and similar applications.

## Achieving Server Redundancy with SunPlex and Sun Cluster 3.0 Software

Clustering occurs when two or more computers (nodes) or domains are combined into a single, unified resource. Clusters were created to provide continuous access (high availability) to business-critical data and applications by means of server redundancy and application failover.

By clustering two or more redundant servers and related storage arrays, IT managers can implement higher levels of availability in mission-critical or compute-intensive environments. Although the original clusters were often expensive to manage, complex to administer, and difficult to extend, contemporary technology is more affordable and prolific. Now used by enterprises of all types, clustering technology evolved over time to provide much greater flexibility, scalability, and manageability along with increasing levels of service availability.

Today, local clusters (those in which all of the nodes and storage subsystems are placed in the same room) are an intrinsic component of any business continuity solution. But before they can begin providing continuous service availability, clusters must be deployed appropriately. Businesses often use a physical interconnect (usually SCSI) to allow two or more servers to share access to data mirrored across storage subsystems.

**Expanding Storage Clustering Technology**

Using SCSI technology, the maximum distance between clustered devices is limited. In fact, the service radius of many clusters cannot exceed the cable length connecting the server, the shared storage, and the other server. And, while this configuration does provide excellent protection against small disasters (e.g., a node disk crashing), it does not protect against major disasters that might damage or completely destroy the facility site.

Following the invention of Fibre Channel technology, mirrored data can be sent over much greater distances (as much as several hundred kilometers), making clusters a more viable solution for wide-scale disaster protection. Using this technology, enterprises can extend the distance between nodes in a cluster to different buildings or separate locations within a wide geography — all without changing the software infrastructure, applications, or data.

Depending on the distance, companies can use different combinations of technologies and management practices. When planning for wide-area cluster solution deployment, risks must be fully understood and adeptly managed. For example, the primary drawback of the extended cluster is an increase in latency caused by sending data over much longer fibers. As a result, system architects must also consider additional factors such as distance, network backbone connections, application services, manageability, and infrastructure costs.

**Sun Cluster 3.0 Software**

Built around Sun server, storage, and network connectivity products, the Sun Cluster 3.0 solution, and the Solaris OE, the SunPlex environment helps increase business service levels while decreasing the costs and risks of managing complex enterprise networks. Using the SunPlex environment, devices, file systems, and networks are able to operate seamlessly across tightly coupled resources — helping businesses to deploy extended or campus clusters without changing the underlying infrastructure or applications.

Sun Cluster 3.0 software is designed to help protect against single hardware or software failures such as node crashes or service interruptions. The software monitors the status of hardware and software components and initiates appropriate actions if a problem or failure occurs (in the case of a catastrophic disaster, some recovery procedures must be manually initiated).

To improve reliability and performance, Sun Cluster 3.0 software is tightly integrated with the Solaris OE. The application significantly reduces error detection times, providing for a significantly more robust software stack. However, the Sun Cluster 3.0 solution cannot address infrastructures that span extremely long distances (across a continent, for example).

Depending on the failure, Sun Cluster 3.0 software is designed to either fail over the affected services to another node in the cluster or try to restart them. In all cases, the software's highest priorities are maintaining data integrity and minimizing application service downtime. This requirement drives the layout of the infrastructure and all of the algorithms in the product.

Standard monitoring agents are available for most databases and commonly-used applications. Agents for other services can be developed and deployed using either sophisticated APIs or easy-to-use utilities (for example, the SunPlex Agent Builder tool). Furthermore, the Sun Cluster 3.0 software framework and associated algorithms do not change when deployed in a campus cluster or a distributed Enterprise Continuity cluster. The primary goal is to provide service availability with very high data integrity. Not surprisingly, Sun Cluster 3.0 software offers an excellent base for a disaster recovery solution, especially when combined with additional technologies — for example, replication, trained personnel, and intelligent management processes.

**Using Quorums to Enhance Availability**

Quorum devices provide an additional layer of available storage in a disaster situation by determining which nodes are functioning properly and setting them aside for immediate use by the system at large. For example, in a two-site arrangement, the quorum device is contained within one site, making the loss of that data center more catastrophic than the loss of the other. By implementing a three-site infrastructure, the quorum device can be placed in the third site, so that the loss of one site will not affect the majority of quorum votes.

**Understanding Cluster Limitations**

Although extended clusters provide significant disaster protection, they do not offer a complete recovery solution. A cluster that has only one logical copy of data is still vulnerable to inconsistencies that might be introduced by faulty software or hardware — even when the data is mirrored. Common user errors (an erroneously deleted database table, for example) can still cause a major disaster. To protect against these situations, tape backup can prove invaluable for speedy recovery.

## Enhancing Application Availability

Today, application service providers have ramped up their service offerings, forcing e-commerce to evolve in the process. Internet applications are expected to provide 24 x 7 access to critical information — meaning downtime and poor performance are no longer acceptable. Sun Cluster 3.0 software is designed to help businesses meet consumer expectations by helping to ensure that high levels of application availability and scalability are maintained.

As data becomes accessible to multiple hosts, processing should be distributed across more than one server. To accommodate increased activity, applications can be load balanced across a number of specialized servers while database applications are consolidated and maintained separately. The Sun Cluster 3.0 software framework also extends the Solaris OE to core services (for example, devices, file systems, and networks), allowing resources to operate seamlessly across a SunPlex system while maintaining compatibility with existing applications.

Sun Cluster 3.0 software is designed to provide scalable services and Oracle high availability. The application also supplies global services for applications and users by means of global devices (for example, disks, tapes, CD-ROMs, and meta-devices) that are attached via cluster nodes and accessed as if the devices were local. Networking in this manner enables users to connect to a highly available system service with a single address. Moreover, this type of global file service allows file systems to be mounted concurrently and coherently from every node in the cluster.

- *Global Devices*
  The Sun Cluster 3.0 platform automatically detects all storage devices in the SunPlex system, including disk drives, tape drives, and CD-ROMs. When the system boots, each device is assigned a global name and integrated into the SunPlex system by default. (Names can also be added dynamically after system boot-up.) This feature improves ease of use and supplies seamless storage scalability — for better availability and easier administration.
- *Global File Service*
  The Sun Cluster 3.0 platform abstracts data locations from file services, making it unnecessary to attach data to the host server. The application also utilizes a global file service that operates between the kernel and the UNIX® file system (UFS); the entire process is recorded by the same mount point throughout the system. In addition, the global file service helps provide continuous data availability and reduces system administration complexity.
- *Global Network Services*
  The Sun Cluster 3.0 platform abstracts IP services using network interfaces, allowing them to reside in and be shared by any domain on the SunPlex system. To network clients, the entire SunPlex system appears as a single computing resource. IP abstraction is possible because global network services operate under a many-to-many relationship model that connects IP services to IP addresses. This model enables multiple IP addresses to be used by one or more services using a single IP address. Another key feature of IP services is that they can be started or stopped dynamically and migrated from one domain to another — without any service interruption. To help ensure that application connections are seamless, the Sun Cluster 3.0 platform supports network adapter failover capabilities within the same domain.
- *Scalable Services*
  The Sun Cluster 3.0 platform allows one or more applications within a service to run across multiple domains or systems. By adding more domains or systems to the SunPlex system, capacity and continuity can be increased. And, in the event that a planned or unplanned outage occurs, service levels can be maintained.
- *Oracle High Availability*
  Oracle High Availability is a cluster framework provided by Sun that allows an instance of Oracle to be failed over from a primary server to another server in case of planned maintenance or unforeseen challenges.

**Implementing Cluster-aware Applications**

Cluster-aware applications are architected to take advantage of underlying cluster technology. Typically, several application instances run on different cluster nodes, each taking maximum advantage of clustering components. One application agent available for the Sun Cluster 3.0 software is the Oracle 8i Parallel Server (OPS), and Oracle 9i Real Application Clusters (RAC), a single relational database system. Using this configuration, companies can distribute the work load concurrently across several servers.

Other application agents currently available for Sun Cluster 3.0 software are:

- HA DNS
- HA NFS
- HA Oracle
- Oracle Parallel Server (OPS)
- Oracle9i RAC
- HA Sybase
- Scalable SAP
- HA Sun ONE Web Server
- HA Sun ONE Messaging Server
- HA Sun ONE Directory Server
- HA Sun ONE Calendar Server
- HA Netscape Directory Server (LDAP)
- HA Apache Web/Proxy Server
- HA NetBackup
- HA Solstice Backup software
- Scalable Sun ONE Web Server
- Scalable Apache Web/Proxy Server
- Scalable BroadVision One-To-One
- IBM WebSphere MQ — HA Agent for IBM WebSphere MQ that improves the availability of WebSphere deployments
- IBM WebSphere MQ Integrator — HA Agent for IBM WebSphere MQ Integrator that improves the availability of WebSphere deployments
- Samba — HA Agent for Samba, that improves the availability of Samba deployments
- DHCP — HA Agent for DHCP, that improves the availability of DHCP deployments

Agents available through third parties:

- IBM DB2 (EE and EEE)
- HA Informix Dynamic Server
- Sybase ASE

Chapter 5

# Creating an Available, Reliable Network

For an Enterprise Continuity solution to be successful, key system elements must remain in constant communication. Cluster applications, computing resources, and storage devices are linked to one another by a powerful network to help ensure that vital data remains scalable, available, and secure.

Networks can be divided into access and transport segments. The transport segment helps ensure high availability by supplying a means to connect geographically separate data centers. Unfortunately, most transport networks are reliant on the Fibre Channel protocol, which generally limits the distance between data centers to 10 km or less. The challenge, therefore, is clear. New methods for expanding connectivity and extending network range — even across significant distances — must be designed and developed.

## Rerouting Users via the Access Network

The Enterprise Continuity solution separates data from servers and specifies server and application configurations to improve performance during failure situations. If a failure occurs, users expect that applications will automatically redirect client requests to the applicable failover servers. It is both impractical and imprudent to expect users to know how to change their client configurations. Consequently, businesses cannot simply install secondary servers and assume emergency preparations are sufficient. It is necessary to add a mediation service to help ensure that the applications are readily available — even during failover situations.

In localized cluster implementations, applications fail over to a nearby, but physically separate, server. Long-distance application clusters, on the other hand, require public network access across much greater distances. In configurations where the primary and failover servers are located on opposite sides of a metropolitan area network (MAN), it is essential that user requests reach the failover server and that the server returns replies.

The Enterprise Continuity solution solves the problem by implementing several pretested components from Sun and its SunVIP partners. Supported technologies include aggregation and load-balancing devices and Fibre Channel fabric switches such as:

- ***The Sun StorEdge 2 GB FC Switch-16***
  Offering state-of-the-art SAN features, Sun StorEdge switches are used to connect the local storage network to the transport network.
- ***The Brocade SilkWorm family of switches***
  Brocade SilkWorm switches connect servers and storage devices via Fibre Channel SANs. If already deployed in an enterprise, these switches can be used to connect the SAN to the transport network.
- ***The Nortel Networks Passport 8600 Series routing switch modules***
  A hardware-based Layer 2 — Layer 7 routing and traffic classification switch, the Nortel Networks Passport 8600 Series module supplies high redundancy, multiple high-bandwidth connections, wire-speed performance, and quality-of-service (QoS) support. Within the Enterprise Continuity solution, the Passport 8600 is primarily used for aggregation of 10/100 BASE-T and 1000 BASE-SX gigabit Ethernet public network access and cluster server interconnect traffic.
- ***The Nortel Networks BayStack switch family***
  The Nortel Networks BayStack switch family is a series of Ethernet devices used for small-scale Ethernet aggregation of public network access and cluster server interconnect traffic.
- ***The Nortel Networks Alteon Web switching portfolio***
  A powerful line of Layer 4 — Layer 7 switching and application acceleration devices, Nortel Networks products help ensure that traffic is load balanced intelligently among clustered servers. The Alteon portfolio includes stand-alone switches and a Web Switching Module (WSM) for the Passport 8600.

## Improving the Transport Network

Companies that place all data processing equipment at a single site risk total data loss in the event of a disaster. To ensure that enterprise systems are adequately protected, a contingency plan that provides for remote recovery must be in place. SAN inter-networking provides an infrastructure that enables companies to protect data over long distances.

Long-distance networking can take several forms. One solution is to develop synchronous data mirroring capabilities built atop a SAN/DWDM infrastructure. Using this design, a designated remote system can access mirrored data volumes if the primary server fails. Single-mode, long-wavelength Gigabit Interface Converters (GBICs) are capable of transporting data up to 10 km. However, if the network requires connectivity over a greater total area, SANs can utilize WANs and MANs to encompass distant production and recovery sites. DWDM's “extended fabric” feature enables native Fibre Channel connections to 200 km at nearly 100 MB per second.

The Enterprise Continuity solution employs an optical transport network that allows Fibre Channel extension to 200 km. Dense Wave Division Multiplexer (DWDM) technology is used to regulate native transport of different protocols (for example, Fibre Channel and gigabit Ethernet)

across the same fiber — simultaneously maximizing bandwidth and minimizing the number of network devices and connections required.

## Connecting Multiple Sites

There are two types of Wave Division Multiplexing: Coarse Wavelength Division Multiplexing (CWDM) and Dense Wavelength Division Multiplexing (DWDM). The distinguish between the two is the amount of data being transferred across a fiber strand. As the name implies, DWDM packs considerably more information onto each fiber. A smaller enterprise requiring limited bandwidth (generally, multiples of 2.5 gigabits per second of protected traffic) can begin with a more cost-efficient CWDM installation that evolves over time into a DWDM implementation. Larger enterprises can connect smaller satellite locations into a DWDM implementation using a CWDM connection, allowing them to fully utilize the fiber bandwidth.

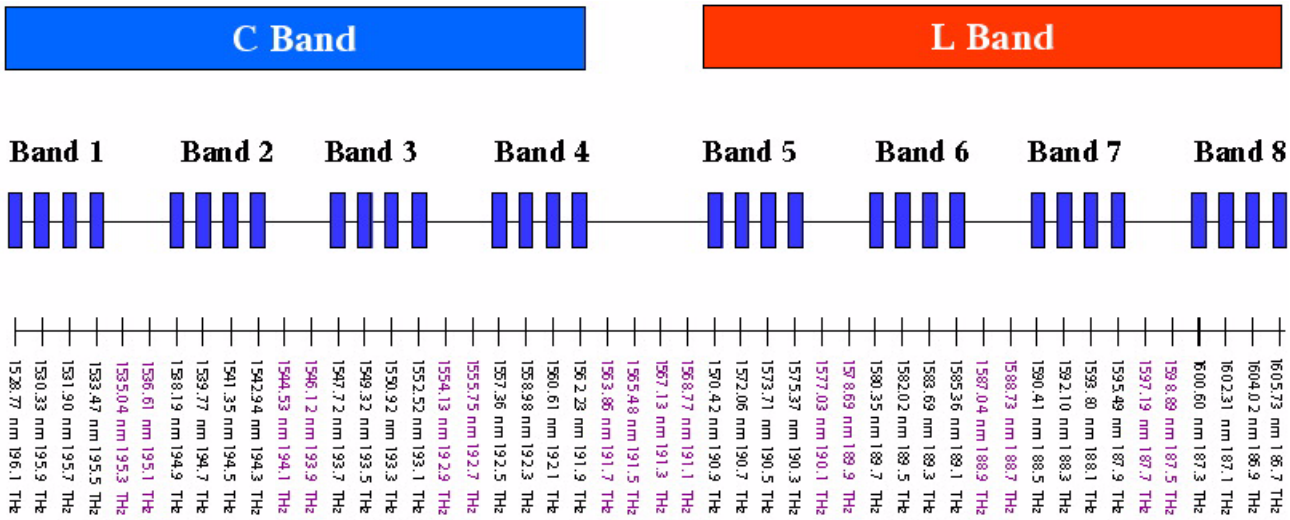
### Bundling Signals onto a Single Fiber

As the term “multiplexing” implies, DWDM technology transmits multiple signals across the same physical fiber connection concurrently. Unlike the time-division multiplexing commonly used by service providers, DWDM uses many different wavelengths of light carried on the same fiber for each incoming signal.

Using DWDM, each signal is allocated a separate color or wavelength. The DWDM device receives and interprets incoming optical signals before converting them into electrical signals. Each signal is then assigned to a specific outgoing wavelength, converted to a corresponding optical signal, and retransmitted on that unique wavelength or channel. Each channel corresponds to a different color that can be measured by its frequency in nanometers (1 nm = 10(-9) m).

Figure 5-1 shows each of the incoming optical signals, the optical-electrical-optical conversion, and the resulting outgoing optical signals. In this example, all of the incoming optical signals use the same frequency, demonstrating the conversion process that results when the outgoing signals are transported using different frequencies.

Figure 5-1: Optical Signals



- This representation is based on an ITU wavelength grid
- Wavelengths are placed 200 GHz apart
- There are a total of 32 wavelengths per fiber (16 in the conventional band and 16 in the long band)
- C and L Band are subdivided into four bands each (eight bands per grid)
- Each of the eight bands contains four channels or wavelengths

**Connecting DWDM Devices**

Three types of sites exist in a DWDM network: terminal sites, optical add-drop multiplexer (OADM) sites, and intermediate sites. Optical signals cannot pass through terminal sites. When a signal is received by a terminal device, all wavelengths used in the network are terminated (de-multiplexed and converted) and directed to access devices. An OADM site, on the other hand, terminates some wavelengths and allows others to pass through optically. Intermediate sites are primarily used for amplification or regeneration of optical signals.

**Building with Dark Fiber**

In the past, carriers built optical fiber networks and assumed responsibility for “lighting” the fiber (attaching lasers and telecommunications equipment) and providing managed service to the customer. Dark fiber shifts the responsibility for lighting the fiber to customers. Contemporary dark fiber networks enable the customer to own and control the actual fiber, allowing them to connect directly to service providers. Available services include telephony, cable TV, and Internet; however, most customer-owned dark fiber deployments are used to support high availability IT solutions.

**Nortel Networks OPTera Metro 5200 for DWDM**

Nortel Networks OPTera Metro 5200 Multiservice platform delivers immense bandwidth (32 wavelengths) using DWDM technology. The platform can provide 10 Gbps per wavelength scalability along with a network modeling tool that simplifies DWDM deployment. The OPTera Metro 5200 relies on an open and modular architecture that supplies excellent network scalability, per-wavelength manageability, bit rate and protocol independence, and ring survivability.

The OPTera Metro 5200 offers several other performance benefits. For example, the platform sets aside one wavelength for use as a service delivery vehicle, dramatically increasing flexibility and scalability. The platform also supports a wide array of protocols and bit rates, allowing service providers to leverage investments in existing networks, differentiate service offerings, and create new revenue opportunities. OPTera Metro 5200's protocol and bit rate independent interfaces enable rapid delivery of Synchronous Optical Network (SONET/SDH) and native data services (for example, ESCON, D1 Video, gigabit Ethernet, 10 gigabit Ethernet, FICON, and Fibre Channel). In addition, customers who wish to relocate can purchase a free-standing cabinet designed for speedy deployment. Key platform features include:

- Carrier-grade availability and scalability
- 32 protected (64 unprotected) wavelengths
- Per-wavelength and per-fiber optional protection switching
- Point-to-point and survivable ring-based optical network
- Performance monitoring and fault sectionalization for SONET/SDH and native data services

- Tunable bandwidth for flexibility
- High-performance service velocity (16 Mbps to 10 Gbps per wavelength)

**Nortel Networks OPTera Metro 5100 for CWDM**

The Nortel Networks OPTera Metro 5100 is a compact Coarse Wavelength Division Multiplexing (CWDM) multiservice platform that delivers eight protected (16 unprotected) wavelengths. Designed to extend optical broadband services to the network edge more cost-effectively, the forecast-tolerant OPTera Metro 5100 enables high connectivity for point-to-point and survivable ring optical networks — making it suitable for both co-location sites and customer premises. Key platform features include:

- High bandwidth delivery of SONET/SDH and native data services
- Protocol and bit rate independent interfaces
- Small footprint and flexible packaging
- Carrier-grade availability and scalability
- Eight protected (16 unprotected) wavelengths
- Per-wavelength optional protection switching
- Point-to-point and survivable ring-based optical network
- Tunable bandwidth

**Implementing Fabric Switches and Buffer Credits**

Buffer credits affect the throughput available to each Fibre Channel switch port. Buffer credits are especially critical to long-distance applications that rely on long runs of optical cable to connect geographically distant hardware devices. If all available buffers are in use, the amount of data that can be transmitted over the cable decreases — negatively impacting performance across the entire Enterprise Continuity solution.

Most FC switches are configured with enough credits per port to handle distances up to 30 km. To support larger service areas (up to 200 km), far more ports are required. The FC protocol, for example, uses 2 KB frames to transmit packets of data. Therefore, it would take five 2 KB frame buffers at 1 Gbps to provide enough buffering for full-bandwidth performance at 10 km. The rule of thumb applied to configuring buffer credits is:

- At 1 Gbps, a frame occupies approximately four km of fiber
- To find the approximate number of buffer credits needed to fully utilize the link, divide the distance in km by two

Sun experts tested both the Fibre Channel switches and the Nortel Networks OPTera Metro 5200 Multiservice platform for the Enterprise Continuity solution. In combination, the device pair provides synchronous mirroring at full Fibre Channel speed for distances up to 200 km. Achieving full Fibre Channel speed — even at distances approaching 200 km — is possible due to the large number of buffers (credits) available at the interswitch ports (E_Ports) on the selected switches.

Chapter 6

# People and Processes

Continuous service availability cannot be achieved with technology alone. To accommodate their company's unique business requirements, IT executives require an business continuity solution that provides a complete retinue of people, processes, and products. For this reason, Sun consultants supply client companies with the field-tested tools, infrastructure, and methodologies necessary to architect, implement, and manage any business continuity solution. Moreover, IT executives who engage Sun Services can access all necessary solution components via a single point-of-contact — effectively reducing project complexity while minimizing financial risks.

## Assessing Business Requirements

Sun Services professionals recognize that application availability is not confined to the server, but rather extends to include the people and processes responsible for supporting and sustaining the environment. In fact, these people and processes are generally responsible for system recovery when a failure does occur. Sun created the SunReady$^{SM}$ Availability Assessment (SRAA) service to help identify gaps in the IT environment. Whether related to the people, processes, or technical architecture involved, the assessment service helps correct potential problems before they affect the level of system availability.

Sun Services professionals evaluate the impact of people, processes, and products on several critical elements of an IT infrastructure, including:

- Service management
- Problem management
- Change management
- Staff management

- Account management
- Program management
- Asset management
- Improvement management
- Implementation management
- Application availability management
- Availability architecture and configuration
- Execution management
- Data center operations, security, and network management

Sun Service reports can provide a comprehensive management-level assessment of IT infrastructure availability, including recommendations for optimizing business services and reducing outage risk. The reports can also compare current operation procedures with industry best practices in order to:

- Identify staff skill gaps that may affect system availability
- Balance needs and costs by determining level of availability required
- Locate high-risk areas within existing IT infrastructure
- Prioritize areas for improvement
- Identify the most effective and cost-efficient means to address prioritized issues

## Providing Expertise Throughout the Project Lifecycle

Enterprise Continuity is an ongoing process. For this reason, Sun business continuity specialists provide expert assistance for every stage of an Enterprise Continuity project. The Sun Availability Services Suite was designed by Sun engineers to help ensure that high service levels are maintained throughout deployment and implementation life cycles. To help IT teams address issues as they occur, Sun Services consultants are available to help internal staff architect, implement, and manage cluster-based solutions for optimal availability. These service offerings are spread over several distinct life cycle phases:

- ***Business Impact Analysis***
  Before work on an Enterprise Continuity project can begin, Sun consultants complete a business impact analysis. This analysis helps reveal time scale and recovery requirements so that the costs associated with a given disaster can be more accurately measured.
- ***Risk Assessment and Management***
  Next, Sun risk management experts carry out a structured review of specific business processes, computer complexes, buildings, or sites to identify risks to long-term business continuance.
- ***Business Continuity Design***
  During this stage, Sun engineers use proven design processes to create long-term disaster provisions, including a complete recovery strategy. Specific installation details — such as system software configurations, internal disk layouts, management tools, security demands, and network configuration — are recorded and documented in the design specification.
- ***Business Continuity Planning***
  Sun consultants collaborate with service providers to support selected recovery strategies and alternatives. Following this step, businesses can identify and document recovery plan components, enabling changes to user procedures and upgrades to existing data processing operations as required.

- *Implementation*
  Throughout the implementation phase, Sun consultants manage and configure the Sun Cluster software. They set up resource groups and logical hosts, configure shared storage, integrate customized applications, and configure the application agent to provide appropriate test failover capability. In addition, Sun experts deliver an on-site management review to help ensure that a company's internal IT staff has a basic understanding of cluster technology operations, upgrades, and patches (complete Sun Cluster Administration training is available from Sun Educational Services).
- *Testing*
  Following implementation, Sun consultants perform a series of basic functional acceptance tests to demonstrate that the Sun Cluster software has been installed and configured successfully. Next, Sun experts orchestrate a walkthrough of technical and non-technical business plans to help identify infrastructure flaws and skills deficiencies (additional technical resources and expertise via its SunSkills™ program).
- *Review and Audit*
  The final phase begins with a formal review of all project activities and deliverables. Particular emphasis is placed on clarifying support arrangements for third-party products. Sun consultants work alongside the service provider team to assess the risks associated with the completed solution and verify that current provisions are well managed.

## Educating Employees

Sun education experts can help companies analyze the skill sets of existing IT staff, build a customized education plan, and certify employee performance by means of comprehensive skills analysis, training, education, and certification offerings. By optimizing internal IT expertise and leveraging customizable Sun Services support plans, companies can maximize uptime and minimize long-term storage challenges. Sun offers a number of Enterprise Continuity-related education services, including:

- Standard role-based pretests
- Learning needs analysis and planning
- Custom course design and development
- eLibraries
- Standard certification paths
- Custom IT management seminars

## Providing a Single Point-of-Contact

Sun implementation experts recognize that IT managers are faced with a bewildering variety of hardware and software systems — all claiming to meet the challenge of system availability and recovery. By engaging Sun Services, these managers can leverage some of the best available people, processes, and technologies from Sun and its industry-leading partners.

Sun consultants supply the overall architecture, assist with implementation, and manage day-to-day interactions with third-party companies — providing a single point-of-contact for the entire Enterprise Continuity solution. For example, the configuration described in this paper uses Nortel Networks DWDM technology to manage Fibre Channel connectivity.

A member of the SunVIP program, Nortel Networks helps provide collaborative support for the Sun servers, storage devices, and software linked to Nortel Network's OPTera Metro 5200

Multiservice platform. Via the SunVIP program, customers can call Sun Services concerning interoperability issues for either company's products to achieve extremely efficient problem resolution. The program is provided, at no additional cost, to customers holding SunSpectrum Gold℠ or SunSpectrum Platinum℠ contracts.

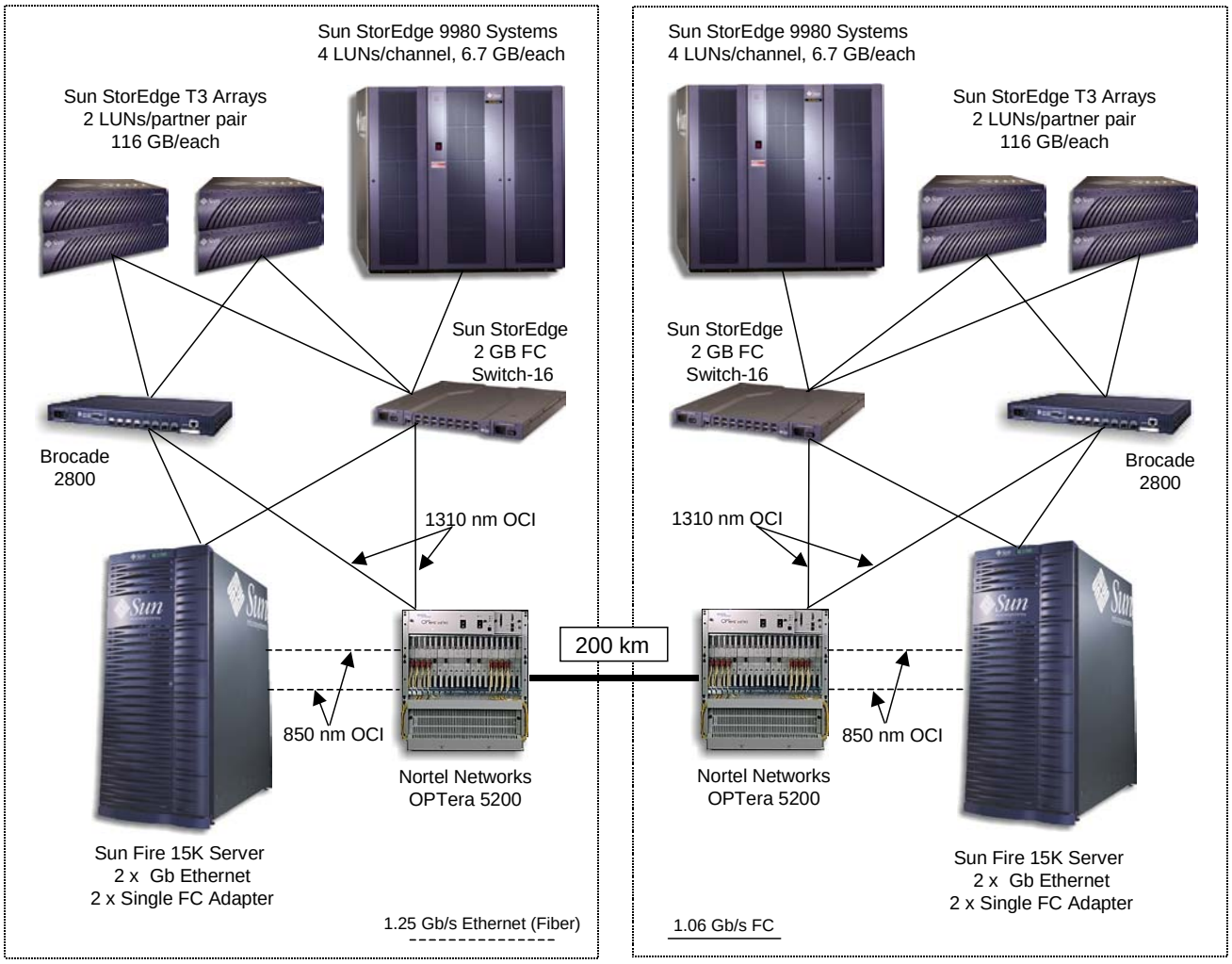
Chapter 7

# Enterprise Continuity Solution Test Configuration

Sun and Nortel Networks performed successful tests to validate one instance of the Enterprise Continuity architecture (using the configuration shown in Figure 7-1). Various components of the test, including configuration, software, hardware revision levels, and test details, are provided as an aid to companies implementing similar environments. The software and version levels used in the test are listed below:

- Solaris 8 OE, update 7, using recommended cluster patches
- Sun Cluster 3.0 software, update 3, using latest cluster patches
- VERITAS VxVm 3.2t, patch 111909-04
- Vdbench 3.02
- GRITS I/O load test suite 1.0E
- Oracle Database 9.0
- HA-NFS
- HA-Oracle
- HA-OPS
- Sun SAN 4.0 software
- SANSurfer 2.08.22, using patch 110696-xx
- Sun StorEdge Traffic Manager software, version 6.0.42-2-1-19
- Sun StorEdge Sbus Dual Fibre Channel Network Adapter software (SUNWcfpl and SUNWcfplx packages)
- Explorer 5.3

Figure 7-1: Enterprise Continuity Test Configuration




The following hardware was used in the test:

- Sun StorEdge 2 GB FC Switch-16, running the latest firmware (4.02.26)
- SANBox4-16-a16M
- Flash 30462, prom software 40500
- Brocade SilkWorm switches, running the latest firmware (v2.6.0c)
  - Model 2400 and 2800
  - Software kernel 5.4 fabric OS version 2.6.0c
- Sun Gigabit Ethernet PCI Adapter, X1141A
- Two Sun StorEdge 9900 system ports, with four 6.7 GB LUNs
- Sun StorEdge T3 and T3B arrays
  - Sun StorEdge T3 partner pairs, running firmware 1.18.00
  - Sun StorEdge T3B partner pair, running firmware 2.01.011.18.00
- Two Sun Fire 15k servers, using one domain each with 16 GB memory, 8 x 900 MHz CPUs, Solaris 8 OE, update 7
- Two Single Fibre Channel Network Adapter X6799A, with firmware ISP2200 FC-AL host adapter driver 1.13 01/03/05

- Two Dual Fibre Channel Network Adapter XX6726A, with firmware ISP2200 FC-AL host adapter driver 1.13 01/03/05
- Five 80 km Single Mode Fiber Network simulators
- Four media converters
- Four long-wave GBICs
- Nortel Networks OPTera Metro 5200 Multiservice platform
  - OCLD ver.4, rev.5; OCI ver.1, rev.0; OCM ver.3, rev.6; SP ver.0, rev.0
  - Firmware: 4.0.53.3

The following tests were performed to prove the functionality of the wide area cluster configuration:

- Basic functionality of the Nortel Networks Optera 5200 systems and Sun hosts
- Basic functionality of the Nortel Networks Optera 5200 systems and Brocade 1 GB switches
- Basic functionality of the Nortel Networks Optera 5200 systems and Sun 1 GB switches

Several failover tests were performed by injecting faults into the networks to test the following scenarios:

- Failover within the Nortel Networks OPTera 5200 systems
- Fail and recover paths to Nortel Networks OPTera 5200 systems
- Failover within the hosts in the cluster setup

In addition, the following features and functions of the configuration were tested in order to provide a fully functional Enterprise Continuity solution:

- PCI-based Sun servers
- Sun SAN 4.0 software
- Sun StorEdge Traffic Manager software (Sun StorEdge TMS), v6.0.2-2-1.19
- Sun StorEdge 2 GB FC Switch-16, flash 30462, prom SW 30462
- Brocade SilkWorm 2250 and 2800 switches: kernel 5.4, fabric OS: v2.6.0c
- Use of a Sun StorEdge 9900 system LUN as a quorum device
- One level of cascading through T-port between Sun switches
- One level of cascading through L-port between Brocade switches
- Basic cluster functionality in SAN environment
- Distance connectivity (200 km), measured throughput, and latency impact

Table 7-1 provides a more detailed description of the tests performed on the Enterprise Continuity configuration.

Table 7-1: Enterprise Continuity Test Cases

| ID-Version | Title | Score (pass/fail) |
|---|---|---|
| 02318-01 | Nortel Networks 24 hour stability CLUSTER | 10/19/02 P (P:1 F:0) |
| 00676-01 | Cluster-node power failure simulations | 10/22/02 P (P:1 F:0) |
| 02312-01 | Nortel Networks FC HBA failure CLUSTER | 10/18/02 P (P:1 F:0) |
| 02314-01 | Nortel Networks OCLD failure (unprotected) CLUSTER | 10/18/02 P (P:1 F:0) |
| 02316-01 | Nortel Networks fast HBA fiber pull CLUSTER | 10/18/02 P (P:1 F:0) |

| | | |
|---|---|---|
| 02319-01 | Nortel Networks 24 hour stability CLUSTER | 10/18/02 P (P:1 F:0) |
| 02320-01 | Nortel Networks line fiber failure (unprotected) CLUSTER | 10/20/02 P (P:1 F:0) |
| 02321-01 | Nortel Networks slow HBA fiber pull CLUSTER | 10/18/02 P (P:1 F:0) |
| 02375-01 | Sun StorEdge T3 array path failure w/ VxVM volumes and Sun StorEdge TMS enabled | 10/21/02 P (P:1 F:0) |
| 03463-01 | Disrupt interlink between 2 Brocade cascaded switches | 10/18/02 P (P:1 F:0) |
| 03464-01 | Reset Brocade switch | 10/18/02 P (P:1 F:0) |
| 03465-01 | POWER OFF Brocade switch | 10/21/02 P (P:1 F:0) |
| 03466-01 | Change a port definition on the Brocade switch (for example F -> SL) | 10/21/02 P (P:1 F:0) |
| 03467-01 | Invalidate path between one the Brocade switch and the node that owns the data service | 10/22/02 P (P:1 F:0) |
| 03468-01 | Disrupt interlink between 2 Sun cascaded switches | 10/18/02 P (P:1 F:0) |
| 03469-01 | Reset Sun switch | 10/18/02 P (P:1 F:0) |
| 03470-01 | POWER OFF Sun switch | 10/21/02 F (P:0 F:1) |
| 03471-01 | Change a port definition on the Sun switch (for example F -> SL) | 10/21/02 P (P:1 F:0) |
| 03472-01 | Invalidate path between one the Sun switch and the node that owns the data service | 10/22/02 P (P:1 F:0) |
| 03474-01 | Power off ALL the equipment (node+switches+storage) on one side of the node | 10/22/02 P (P:1 F:0) |
| 00665-01 | Sun Cluster 3.0 software verify basic cluster function | 10/11/02 P (P:1 F:0) |
| 02310-01 | Nortel Networks Sun Cluster 3.0 software installation with Veritas VxVM 3.2 Sun StorEdge TMS enabled | 10/11/02 P (P:1 F:0) |
| 02295-01 | Nortel Networks Switching from gigabit Ethernet to FC on the Nortel Networks switch | 10/11/02 P (P:1 F:0) |
| 01786-01 | Sun Cluster 3.0 software U2 basic setup — with Veritas file system | 10/11/02 P (P:1 F:0) |

Appendix A

# References

Sun Microsystems posts product information in the form of data sheets, specifications, and white papers on its Web site at: http://www.sun.com/.

Please also refer to the following resources for more information on topics discussed in this paper:

- *Deploying SunPlex Systems, A Technical White Paper*
- *Sun Cluster 3.0 Software Cluster File Systems (CFS): Making the Most of the Global File Service, A Technical White Paper*
- *Campus Clusters Based on Sun Cluster 3.0 Software, A Technical White Paper*
- *Sun Flexes Its SunPlex Muscle as Sun Cluster 3.0 Version 5/02 Adds and Improves Capabilities by D.H. Brown Associates Inc.*
- *Storage Consolidation, How to Reduce IT Costs and Complexity*
- *Building a Data Continuance Environment, An Executive Brief*

Nortel Networks information can be found at: http://www.nortelnetworks.com/.

- *Nortel Networks OPTera Metro 5000-series Multiservice Platform Planning Guide*

Oracle information can be located at: http://www.oracle.com/.

- *High Availability White Paper*

Books:

- *Building SANs with Brocade Fabric Switches, Chris Beauchamp, Syngress Publishing, Inc.*

Appendix B

# Glossary and Terms

**CWDM**
Course Wavelength Division Multiplexing
**DWDM**
Dense Wavelength Division Multiplexing
**ESCON**
Enterprise Systems Connection. A 200 Mbps serial I/O bus used on IBM Corporation's Enterprise System 9000 data center computers. Abbreviated ESCON. Similar to Fibre Channel in many respects, ESCON is based on redundant switches to which computers and storage subsystems connect using serial optical connections.
**ETSI**
European Telecommunications Standardization Institute
**FICON**
Fiber Connectivity. IBM Corporation's implementation of ESCON over Fibre Channel.
**GbE**
Gigabit Ethernet. A group of Ethernet standards in which data is transmitted at 1 Gbit per second.
**IP**
Internet Protocol. A protocol that provides connectionless best effort delivery of datagrams across heterogeneous physical networks.
**LAN**
Local Area Network. A communications infrastructure designed to use dedicated wiring over a limited distance (typically a diameter of less than five kilometers) to connect a large number of intercommunicating nodes. Ethernet and token ring are the two most popular LAN technologies.
**MAN**
Metropolitan Area Network. A network that connects nodes distributed over a metropolitan (city-wide) area as opposed to a local area (campus) or wide area (national or global).
**NEBS**
Network Equipment Building System
**OADM**
Optical Add-Drop Multiplexer

**OCI**
Optical Channel Interface
**QoS**
Quality of Service
**SAN**
Storage Area Network. A storage system consisting of storage elements, storage devices, computer systems, and/or appliances, plus all control software, communicating over a network.
**SONET**
Synchronous Optical Network. A standard for optical network elements. SONET provides modular building blocks, fixed overheads, integrated operations channels, and flexible payload mappings. Basic SONET provides a bandwidth of 51.840 megabits/second. This is known as OC-1. Higher bandwidths that are n times the basic rate are available (known as OC-n). OC-3, OC-12, OC-48, and OC-192 are currently in common use.
**WAN**
Wide Area Network. A a communications network that is geographically dispersed and that includes telecommunications links.
**WSM**
Web Switching Module

**Sun Microsystems, Inc.** 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-800-555-9786 or 1-800-555-9SUN **Web** sun.com


**Sun Worldwide Sales Offices:** Africa (North, West and Central) +33-13-067-4680, 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, 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