

*Sun<sup>TM</sup> Service Point Architecture —  
Delivering Services on Demand for the  
Networked World*

---

*October, 2001*



© 2001 Sun Microsystems, Inc.  
901 San Antonio Road, Palo Alto, California 94303 U.S.A

#### TRADEMARKS

Sun, Sun Microsystems, the Sun logo, The Network is the Computer, Java, J2EE, JavaServer Pages, JSP, Enterprise JavaBeans, EJB, Solstice AdminSuite, Solaris, Sun ONE, Sun Service Point Architecture, iPlanet, Sun Management Center, Sun Trunking, SunTone, SunSpectrum, Sun StorEdge, Sun Enterprise, SunPlex, Sun Cluster, and Sun Fire are service marks, trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and other countries.

UNIX is a registered trademark in the United States and other countries, exclusively licensed through X/Open Company, Ltd.

THIS PUBLICATION IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT.

THIS PUBLICATION COULD INCLUDE TECHNICAL INACCURACIES OR TYPOGRAPHICAL ERRORS. CHANGES ARE PERIODICALLY ADDED TO THE INFORMATION HEREIN; THESE CHANGES WILL BE INCORPORATED IN NEW EDITIONS OF THE PUBLICATION. SUN MICROSYSTEMS, INC. MAY MAKE IMPROVEMENTS AND/OR CHANGES IN THE PRODUCT(S) AND/OR THE PROGRAM(S) DESCRIBED IN THIS PUBLICATION AT ANY TIME.



Please  
Recycle

# *Contents*

---

<b>Intended Audience</b> .....	<b>ii</b>
<b>Executive Summary</b> .....	<b>iii</b>
<b>The Net Effect</b> .....	<b>1</b>
<b>Services on Demand</b> .....	<b>2</b>
<b>Sun Open Network Environment (Sun ONE) Software Platform</b> .....	<b>4</b>
<b>Introducing the Sun Service Point Architecture</b> .....	<b>6</b>
Edge of the Network .....	9
Heart of the Data Center .....	10
<b>Implementing the Sun Service Point Architecture</b> .....	<b>11</b>
Scalability .....	11
Availability .....	13
Making Complexity Manageable .....	16
Security .....	22
Sun Products — Enabling the Service Point Architecture .....	23
Related Programs and Services from Sun .....	24
<b>The Value of the Sun Service Point Architecture</b> .....	<b>25</b>
<b>For More Information</b> .....	<b>27</b>

---

## *Intended Audience*

This paper describes an architecture for effectively implementing application services. The ideas and technologies discussed are geared toward *IT managers and executives* in organizations that are implementing software services or Web-based software.

## *Executive Summary*

---



The last several years have witnessed a profound transition in computing. With users demanding new on-line services, and businesses rushing in to provide them, the Internet is growing explosively. Online services that reach several million users are not uncommon. In order to deliver high quality services on this scale, however, businesses have had to *reinvent the way they develop and deploy information systems*. Initial implementations proved to be complex and costly to manage, so businesses have begun to emphasize consolidation and simplicity. The *Sun™ Service Point Architecture* described in this paper was born out of the need to decrease costs without compromising user service levels in these next generation software implementations.

Application server technology and the Java™ 2 Platform, Enterprise Edition (J2EE™) have fundamentally changed the way that developers write software. J2EE compliant software environments such as the *Sun™ ONE software platform* use open technologies to allow developers to create *reusable software components* and to assemble *application services* from these components. The primary benefits of the application services model are the inherent efficiencies of reusable components and the scalability and availability offered when these components are deployed in a distributed architecture. With developers continuing to migrate to this model, the Gartner research firm estimates that the worldwide market for Web-based software services will grow to \$1.7B by 2002.

However, this new software services model brings a new set of issues for CIOs and IT managers. Deploying large scale Web services and keeping them continuously up and running requires businesses to *reinvent the data center*.



Today's IT managers are looking beyond their previous concerns about scalability, availability, and predictable service levels. They must also deliver a secure environment that enables remote or external users to access systems and they must find ways to reduce the cost of deploying their IT systems.

Application services that support millions of users can require thousands of hardware servers. Managing these servers individually, especially in a heterogeneous environment, is not only time consuming and costly, but also leaves room for administration errors that can threaten network security and reduce service levels.

The industry's answer to these challenges is a combined hardware and software approach. By implementing application services with the Sun ONE software platform and by deploying them using the Sun Service Point Architecture, businesses can deliver high quality services at a reduced cost.

The *Sun Service Point Architecture* is a flexible methodology for deploying software services on a hardware systems infrastructure. Its modular component approach allows software components to be mapped onto hardware platforms that can be easily replicated for scalability and availability. The Sun Service Point Architecture also simplifies the management of complex IT environments by allowing administrators to manage higher-level pools of resources instead of individual hardware servers. Shared pools of resources can also be used to handle spikes in user demand. This reduces capital expense by protecting user service levels without requiring extra hardware resources that are dedicated to each application service.

The combined benefits of using the Sun ONE software platform and the Sun Service Point Architecture include:

- *Reduced cost of operations* through better management of complex IT systems
- *Increased value from legacy software* applications through *Web delivery*
- *Flexibility* to respond quickly to change
- *Lower investment in capital equipment* due to better resource utilization

Sun brings unique value by helping enterprises build solutions that can be efficiently managed and easily adapted to changing business needs. Sun hardware and software solutions allow aggregation of resources and add flexibility to data center operations without giving up reliability. The Sun Service Point Architecture provides a foundation for delivery of software services while protecting investment in legacy information systems and providing a means to regain control over complex IT environments.

# *Sun<sup>TM</sup> Service Point Architecture — Delivering Services on Demand for the Networked World*

---



## *The Net Effect*

Changes in technology are creating new opportunities for the way in which businesses utilize computers. The number of devices accessing the Web continues to increase at an exponential rate as high bandwidth connections are becoming more widely available at lower prices. Content providers are therefore making more information available via the Web and businesses are executing more transactions electronically at a cost of pennies rather than dollars per transaction. This increasing utilization of the Web is driving the need for more devices and greater bandwidth. At Sun we call the tremendous opportunity created by this phenomenon the *Net Effect*.

The Net Effect has created an environment where access to the Internet has become mission critical. And the speed at which new Web-based applications are being rolled out has made it impractical for businesses to invest heavily in training their users. Web applications must be so intuitive that users don't think about using an application, but instead think only about the specific transaction they are executing. Users simply access their email, place an order, or find product information, all on the Web.

The increase in value of the Internet is accompanied by an increase in complexity and is forcing a rethinking and reinvention of computing. Specifically the Net Effect is driving reinvention in four areas:

- The Network
- The Client
- The Application
- The Data Center



In the past, off-the-shelf software applications were loaded onto PCs and servers and then licensed for a specific number of users. This model is giving way to a service-oriented application design. Today, smaller, more modular application components or services can be assembled together over the Net as an application service. Once a service component is deployed, it can easily be reused by other services. Developing a new business service can be as simple as reassembling a few existing service components to deliver new functionality to an end user.

Many of these new application services do not reside on the client, but instead are on the server. This allows a wide variety of new devices to utilize these services as easily as traditional PCs. The result will be new devices such as cell phones and PDAs (personal digital assistants) that take on an important role as client platforms and displace the traditional PC.

Since these new services don't reside on the client, the servers must take up the slack. In this new environment, servers in the data center must support greater activity, constant 24 X 7 access, and unpredictable usage patterns. The required features in these new Internet data centers include massive scalability, continuous real-time availability, and greater agility.

The network itself is also being reinvented. Instead of separate networks for different modes of access such as the corporate WAN, wireless, and home networks, a single unified network will be available to end users. This simplifies the task of providing anytime, anywhere access, from any device.

The network has become an integral part of the data center. Unlike traditional mainframe data centers, the networked data center extends out into the network with its loosely coupled services being distributed across the network. Efficiently managing loosely coupled services requires a new focus on the data center. The Sun Service Point Architecture addresses the need for manageability in complex distributed data center environments.

## *Services on Demand*

The Net Effect trend toward smaller, more modular application components using a service-oriented design is already reshaping the industry. Most of today's new software applications are actually an aggregation of multiple software components that can be deployed in a distributed architecture. The component approach allows businesses to quickly assemble new application services by reusing existing software components.



---

Other industries have long profited from reusable components. For example, a spark plug in a car has a standard interface and can be replaced by a new one made from a number of competing manufacturers. Similarly, application services utilize standard interfaces to allow for interchangeable, reusable software components.

These component services can be distributed across the network using XML (eXtensible Mark-up Language) and Java™ technologies. Communication over the network using technologies such as Enterprise JavaBeans™ (EJB™), Servlets, or the JavaServer Pages™ (JSP™) framework greatly simplifies the development of distributed software components.

Sun's vision of the service driven network combines legacy applications and new component services into macro services that are delivered to end users. These dynamic combinations of macro services are referred to as *services on demand* (figure 1.)

The services on demand software model provides the ability to extend existing applications and allow them to be delivered over the Web. At the same time, it simplifies the task of integrating legacy applications with other new applications services. Most businesses, are not simply throwing away their old software applications and starting anew. On the contrary, maximizing return on investment requires an evolutionary approach that builds on existing information assets and leverages new technologies to extend the life of previous investments.

Sun views the services on demand software model as a logical progression from Web-enabled applications. In order to make the transition to services on demand, businesses need a software platform that provides the ability to monetize existing information assets by delivering them via the Web. The platform must also provide a solid foundation for efficient delivery of next generation application services. The Sun ONE software platform provides a software development and delivery environment that meets these needs.



Figure 1 The dynamic combinations of macro services are referred to as *services on demand*.

## *Sun Open Net Environment (Sun ONE) Software Platform*

Reusable software components are made possible by the open technologies in the Sun Open Net Environment (Sun ONE) software platform. Developers can save valuable implementation time and reduce the risk of unforeseen problems by using this integrated suite of technologies. Businesses can also explore new sources of revenue through personalized information services. And most importantly, the Sun ONE platform provides investment protection by allowing legacy information assets to be made available as Web-based services.

Information assets include all of the information technology investments that are used throughout a business. Sun uses the acronym DART to refer to the four primary types of information assets: *data, applications, reports, and transactions*.

The Sun ONE software platform allows businesses to add value to these assets through Web-based delivery and integration with other application services. Figure 2 shows how DARTS can be connected to a software services environment through the components of the Sun ONE software platform. Many businesses start with their legacy information assets when implementing the Sun ONE platform because it provides a quick return on investment.

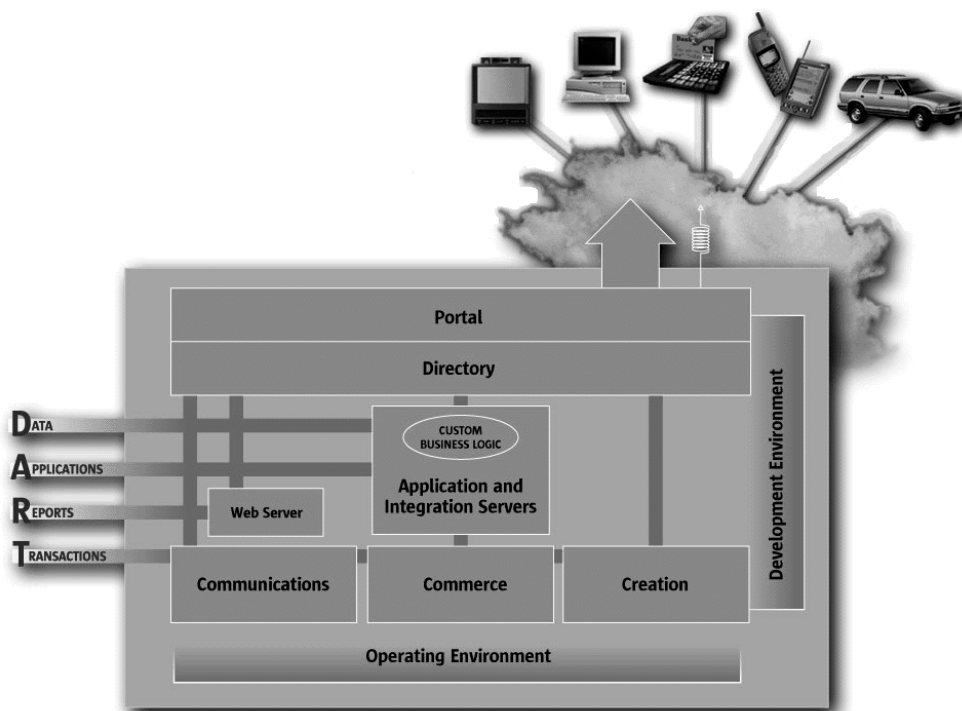


Figure 2 Information assets (DARTS) can be connected to a software services environment through the components of the Sun ONE software platform.

The Sun ONE platform helps ensure that application services are developed using open standards and can seamlessly interoperate. A portal server such as *iPlanet™ Portal Server* handles all of the presentation of data and interaction with end users on various device types. A standard interface is then presented to other software components in the Sun ONE platform to simplify device integration with application services. Similarly, the directory server provides a standards-based management environment for managing user profiles, permissions, and policies.



Integration with existing heterogeneous systems, including legacy applications and databases, is enabled through the use of J2EE technology compliant application and integration servers. J2EE has become a *de facto* industry standard with the top five application servers (ranked by market share according to research firm IDC) all J2EE certified. Third party software products that conform to the J2EE specification can also be used with the Sun ONE software platform. More information about the Sun ONE software platform can be found at <http://www.sun.com/sunone/>.

Delivering scalable application services with the Sun ONE platform also requires a sophisticated infrastructure of hardware servers, storage, and management software. The Sun Service Point Architecture described in the following section is an implementation methodology for delivering application services in an integrated hardware and software environment. It allows software components from the Sun ONE platform to be mapped onto distributed hardware servers.

## Introducing the Sun Service Point Architecture

The Sun Service Point Architecture is a methodology for organizing systems infrastructure to support best practices for implementing services on demand and the Sun ONE software platform. The Sun Service Point Architecture is designed to maximize *scalability, availability, manageability, and security*.

Perhaps most important in this new architecture is the need for manageability. When a service is scaled to support millions of users, the sheer number of hardware servers required and the heterogeneous nature of most IT networks creates a level of complexity that makes it difficult to manage.

The complexity of a modern networked services data center in relation to a traditional client-server data center is illustrated by the comparison in Table 1.

<b>Traditional Client-Server Data Center</b>	<b>Networked Services Data Center</b>
Hundreds or thousands of users per server	Millions of online users
Access by known client systems	Access by any type of device
Scheduled downtime for maintenance	24 X 7 availability
Management of individual hardware components	Management of user service levels



<b>Traditional Client-Server Data Center</b>	<b>Networked Services Data Center</b>
Dedicated hardware servers for specific applications	Multi-tier architecture with many distributed servers that are not necessarily dedicated to a single application
Relatively static system configurations	Hardware resources that are redeployed whenever application demands change

Table 1 Characteristics of traditional client-server and networked services data centers

In traditional client-server environments, IT managers emphasized reliability and delivering predictable levels of service. Scaling or upgrading an application was planned well in advance and care was taken not to create instability. Modern systems are easier to scale and allow varying degrees of availability to be balanced against cost. However, IT managers now require more. New networked data centers require agility, manageability, and increased security without giving up the traditional push for predictability and high availability.

The Sun Service Point Architecture helps address all of these needs. The modular building block approach of the Sun Service Point Architecture allows software and hardware components to be easily replicated wherever additional performance or availability is needed. Manageability is enhanced through aggregation of resources into higher level building blocks and security can be controlled at multiple levels of the architecture.

The foundation of the Sun Service Point Architecture is the notion of multiple points of access, or service points (figure 3.) Service points are made possible by the use of application server technology and open standards described earlier. Application servers execute business logic and transparently manage the connections between software components that comprise an application service. Even a user's method of access (e.g. WAN, mobile device, Internet) is transparent to the application server.

Transparency allows the software components of the application services model to be logically divided into three tiers. There are front-end components that manage interaction with users, a middle tier where the application servers reside, and a back-end where database systems are housed. The Sun Service Point Architecture maps these distinct components of a software architecture onto hardware platforms that reside in the data center.

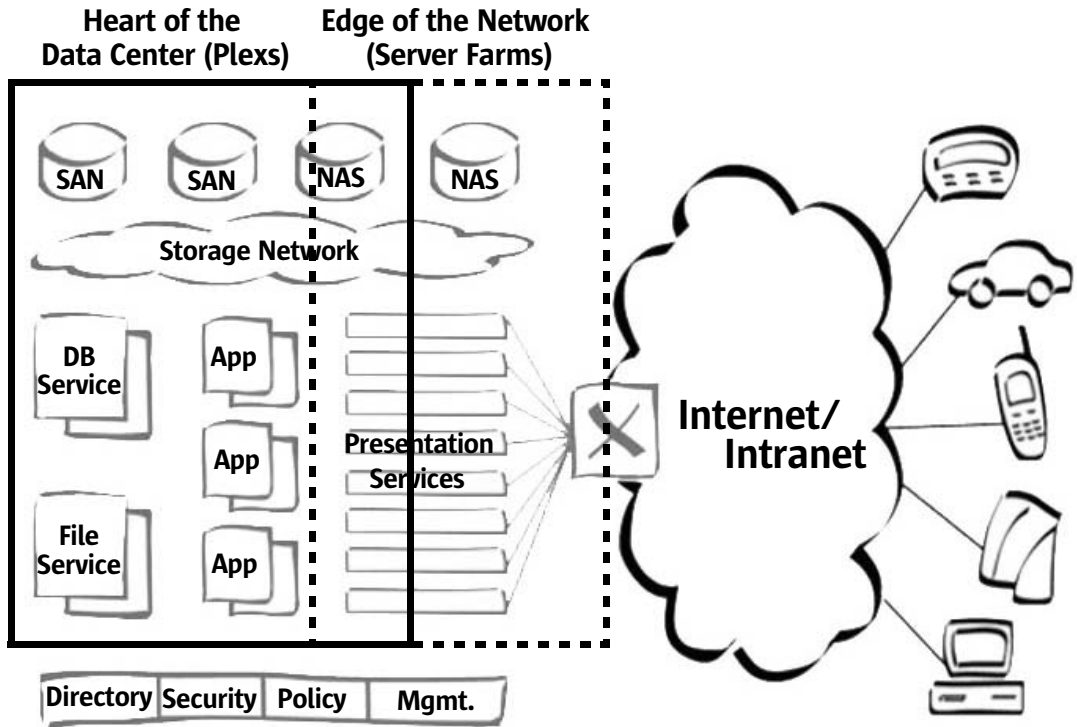


Figure 3 The foundation of the Sun Service Point Architecture is the notion of multiple points of access, or *service points*.

The data center can be thought of as a continuum from the *heart of the data center* to the *edge of the network*. The heart of the data center is where transactions and business logic are executed. It includes both the back-end databases and the application servers. At the other end of the data center continuum is the edge of the network which includes the systems that interact with users and devices.

There are some very distinct patterns in how systems are replicated at opposite ends of the data center. The characteristics and requirements of each are quite different. The following sections explore the patterns in each of these two areas and identify implementation approaches that work best.



---

## *Edge of the Network*

The systems that interact with things outside of the data center are referred to as the *edge of the network*. These systems manage all types of user connections and deliver service requests to the underlying software components in the heart of the data center. Servers at the edge of the network are usually dedicated to specific tasks such as running Web server or portal software. They are generally small front-end servers that utilize local disk and perhaps Network Attached Storage (NAS) to store temporary data. The number of servers at the edge of the network depends upon the volume of user traffic through each mode of access.

A unique characteristic of servers at the edge of the network is that they do not require knowledge of user or application state. They can fail without losing user data.

The recognizable patterns at the edge of the network include:

- *Stateless* — can fail without disruption to the service
- Simple, durable *component servers (blades)*
- Loosely coupled connections between systems (*server farms*)
- Network Attached Storage (NAS)
- *Horizontal scaling* with load balancing— add capacity by adding more blades
- *Service capacity recovery* — reallocate resources to match service level needs
- Remote, *lights out operation* — physically distributed and managed over the network

These patterns have many implications for the way in which servers are deployed and managed at the edge of the network. For instance, to add capacity at the edge of the network, businesses simply add more servers with additional copies of the presentation layer software. Large online services, such as those offered by brokerage firms, can deploy as many as a thousand Web front-end servers, each running a single instance of Web server software. Sun understands that businesses need tools to help them efficiently manage such environments using an aggregated Web front-end service rather than managing a thousand individual servers.

Another common management task at the edge of the network is to install a new server or repurpose an existing server to perform a new task. If this can be done quickly and efficiently, it can help prevent service levels from dropping below critical levels and can reduce the cost of operation. Solutions that simplify or further automate the task of installing or repurposing a server can



help maintain predictable service levels while reducing management costs and minimizing the risk of administrator errors. Tools that simplify lights out operation provide similar service level benefits.

## *Heart of the Data Center*

At the opposite end of the continuum is the *heart of the data center*. This is where application and user state are managed, and is the core infrastructure for processing data intensive transactions. Both back-end database servers and the application servers that execute the business logic are deployed here. Mission-critical data such as sales orders, inventory, product designs, human resources records, etc. must be protected in the heart of the data center to maintain the health of the business.

The ability to recover application and user state are critical in this part of the data center. These systems must be available 24 X 7 and must be designed to handle peak loads without concerns for performance.

Increasing the capacity of systems at the heart of the data center can be accomplished by either vertical or horizontal scaling. Multiple instances of application servers can be deployed along with load balancing to evenly distribute requests from the edge of the network.

Back-end database servers require vertical scaling so that a single instance of the database can manage the data for any one application. The back-end database servers are usually implemented in a centralized location so that resources can be more efficiently shared. Tightly coupled servers that may use private networks in clusters can make it easier to share resources between applications. Shared centralized storage via Storage Area Networks (SANs) can also be used to more efficiently manage storage allocation when application needs change. In addition, data center operations such as backup and restore can be efficiently performed with these shared resources.

The recognizable patterns in the heart of the data center include:

- *Stateful* — failure requires exact recovery
- *Tightly coupled systems* via private networks (plexes)
- *Dynamic allocation of resources* in centralized pools
- Both *horizontal and vertical scaling*



---

Systems deployed at the heart of the data center must be scalable and highly available in order to maintain application service levels. They must also support rapid recovery of user and application state in the event of a failure. The real challenge, however, is maintaining system availability in periods of surging demand. Thus, a critical need in the heart of the data center is the ability to utilize a pool of shared resources to meet peak demands.

Rather than overprovision system resources for each application service, the use of a shared pool from which resources can be quickly and easily reassigned provides the safety margin to maintain service levels without undue capital expense. Unlike at the edge of the network, where complete servers may be repurposed, the requirement in the heart of the data center is that resources within a given server or cluster be dynamically reassigned to an application service upon demand. This must be done without system downtime.

## *Implementing the Sun Service Point Architecture*

Businesses have already begun to implement many elements of the Sun Service Point Architecture using technologies that exist today. Sun is building its competitive edge by designing systems that are tailored to dedicated functions in the Sun Service Point Architecture.

The following sections provide an overview of Sun's technologies and program offerings that address the basic customer needs of scalability, availability, manageability, and security. Additional details are available on the Web sites referenced throughout these sections and the Web references in table 3 at the end of this document.

### *Scalability*

One of the hallmarks of the Internet age and the advent of online services is the rapid and sometimes unpredictable growth of services. In response to rapid growth, businesses must be able to add or reassign resources quickly without disrupting service. To address this need, Sun solutions for the Sun Service Point Architecture allow scalability at both ends of the data center spectrum.



---

## *Scalability in the Heart of the Data Center*

An important consideration in the heart of the data center is getting scalability without requiring downtime and without adding complexity. Sun systems include hot-swappable components that allow the addition of system resources without interrupting application performance. And when it's time to upgrade to an even larger server, businesses can migrate with minimal downtime. With the scalable Solaris™ Operating Environment, applications run unchanged on a single-processor server all the way up to Sun's newest high-end server, the 106-processor Sun™ Fire 15K. Sun's 19 years of engineering investment in the Solaris Operating Environment and in Sun server technology have allowed Sun to maximize vertical scalability in a single system. Properly tuned applications can experience near linear scalability on the Sun Fire 15K server due to its balanced system architecture and efficient utilization of system resources. Sun recently achieved world record benchmark results with the SAP R/3 enterprise software application using a 76-processor Sun Fire 15K server.

In addition to vertical scalability, Sun systems have been designed with network scalability in mind. Features such as high bandwidth, minimal latency, and data coherency are emphasized by Sun. These features are important when multiple servers are used together to tackle problems that require more performance than a single system can provide.

Businesses can further leverage the scalability of the Sun platform by taking advantage of Sun's Capacity on Demand program. In this program, businesses can purchase most Sun servers and Sun™ StorEdge systems already equipped with extra resources such as CPUs, memory, I/O boards, or disk space. Increased capacity is immediately available upon demand, but like purchasing a utility service, only the resources used are paid for.

Pooling resources into tightly coupled clusters can also increase scalability while allowing businesses to more efficiently manage resources. SunPlex™ systems are a combination of Sun™ Cluster software, Sun servers, Sun StorEdge solutions, and a high speed private interconnect, and provide increased scalability as well as increased availability of services.

## *Scalability at the Edge of the Network*

Since the edge of the network includes large numbers of small compact servers, the key to scalability here rests not on scaling within each server, but on quick and easy deployment of additional servers. Sun solutions make the process of



---

installing a server fast and efficient while minimizing the risk of errors in the installation. The Sun tools that provide these capabilities are described in the later section titled, *Replicating Servers at the Edge of the Network*.

## *Availability*

Today, more applications than ever are mission-critical. The industry is learning what administrators already know — 24 x 7 availability is the new standard, with availability often measured by service level agreements that are focused on end-to-end delivery of the application service to the user. To better address this reality, Sun has incorporated a variety of innovations in the Solaris Operating Environment and into its latest generation of servers to increase system availability.

### *Availability in the Heart of the Data Center*

To maintain user service levels, the heart of the data center must be online 24 hours per day. Its systems must be managed to handle peak loads without disruption of service or performance degradation. To this end, availability and service levels can be improved by giving system administrators the power to reallocate resources from one service to another and to add capacity to (upgrade) a service without taking the systems offline.

The following Sun technologies can improve availability in the heart of the data center:

- Highly reliable Sun servers
- SunPlex cluster technology
- Dynamic system domains
- Concurrent maintenance and live upgrades
- Data availability

### *Highly Reliable Sun Servers*

Sun's new generation of Sun Fire servers are based on a strong foundation of technologies that are the result of nearly two decades of innovation from Sun. They are designed to boost availability with redundancy of key system components and an Automatic System Recovery (ASR) feature which prevents faulty components from keeping an entire system down or causing it to fail repeatedly. These systems can be configured so that failure of components such as power supplies, fans, AC power, or even the system clock or system



interconnect can occur without interrupting key services. Even a CPU or a critical ASIC can fail and the system will automatically configure around the failed component at startup.

### *SunPlex Systems and Sun Cluster Software*

SunPlex systems provide increased scalability as well as increased availability of application services. Applications can run unchanged on any node in the SunPlex environment because Sun Cluster software transparently manages shared access to its global file system, global network services, and global devices. Sun Cluster software effectively extends the Solaris Operating Environment, including devices, file systems and networks across cluster nodes while maintaining full software compatibility for existing applications. Application services that are written to the simple-to-use Sun Cluster 3.0 API (Application Programming Interface) can achieve even higher availability with additional failover services.

The health of each individual SunPlex node, including both its hardware and software components, is constantly monitored by Sun Cluster software. Data integrity is maintained by ensuring that only healthy nodes participate in the delivery of services. In the event of a failure, software components can be restarted within the cluster. Following repair, failed nodes may be returned to the SunPlex system without interrupting service. Applications can also be made scalable through replication which increases performance, throughput, and failure resiliency. For maximum availability, the *SunTone™ Cluster Platform 15K/9960* provides a pre-integrated and pre-tested cluster platform that is capable of delivering 99.999 percent availability.

### *Dynamic System Domains*

First introduced on the Sun Enterprise™ 10000 server, Dynamic System Domains are now available across Sun's entire Sun Fire server product line. Dynamic System Domains allow these servers to be divided into separate, virtual systems. Each domain runs its own copy of the Solaris Operating Environment and has exclusive use of its own I/O devices and CPU/Memory boards. With this capability, domains can essentially be regarded as separate systems, each isolated from the effect of errors that occur in other domains.

Domains are important in maintaining flexibility and agility in rapidly changing business conditions. Resources (processors, memory, and I/O devices) can be shifted from one domain to another based on a variety of



---

criteria such as time of day or workload levels. Resource pools can also be set up to automatically replace failed components. Reconfiguration is simple and can be administered remotely through a Web-based interface.

### *Concurrent Maintenance and Live Upgrades*

The Sun Fire midrange server product line and Sun's high end Sun Enterprise 10000 and Sun Fire 15K servers also allow upgrades and replacement of failed or degraded parts while the system continues to operate, thus avoiding downtime. In addition, newly inserted devices can be activated in a running system without the need for a system reboot. With this online upgrade capability, selected resources may be added to a system to increase capacity on an as-needed basis, avoiding downtime when the system is needed the most.

In conjunction with online maintenance for hardware, Sun's Solaris Operating Environment offers parallel capabilities for systems software. Solaris™ Live Upgrade and Dynamic Upgrade to Kernel allow key components of the operating system — including the kernel, to be upgraded while systems continue to run. Solaris Live Upgrade can also be used with Solaris™ Web Start Flash to reduce the time required to update the Solaris Operating Environment on a set of identically configured servers.

### *Data Availability*

By combining appropriate data archival and recovery procedures with the right mix of storage hardware and software components, businesses can minimize unplanned downtime due to data errors.

The new Sun StorEdge™ 9910 and 9960 systems deliver extreme levels of availability and performance for high-end data center environments. They are optimized for the Solaris Operating Environment and are certified for use in SunPlex configurations.

Sun offers a broad range of solutions for managing data availability including:

- Sun StorEdge tape library systems
- “Hot” backup solutions which allow uninterrupted data access
- Sun StorEdge Instant Image software for point-in-time backups
- Sun StorEdge™ Network Data Replicator software to replicate disks between physically separated servers
- Service solutions for establishing best practices as well as remote monitoring



## *Availability at the Edge of the Network*

Because user state is not maintained at the edge of the network, the loss of a server is not as serious as at the heart of the data center. The main concern at the edge of the network is ensuring that enough resources are available to each access method (i.e. WAN, wireless, etc.) so that response times are good and service levels are high.

When service levels are in jeopardy due to a system failure or a rapid surge in user activity, system managers can react quickly by installing additional Sun servers to bring more resources to the problem area. The Sun technologies that provide fast efficient installation of servers and their software are described in the later section titled, *Replicating Servers at the Edge of the Network*.

## *Making Complexity Manageable*

IT infrastructures everywhere have undergone a massive build-out, which has led to unprecedented server sprawl. Businesses are faced with a need to manage hundreds, or in some cases, thousands of servers. Managing these systems is made more difficult in heterogeneous environments where hardware, software and storage solutions come from multiple vendors. Finding qualified system managers is also difficult because system management expertise has not kept pace with the growing need. Businesses are feeling the pain of the increasing costs of managing such complex environments while they also struggle with staffing issues that make it difficult to keep systems running at peak performance.

A new approach to managing the IT infrastructure is needed. Sun is addressing these problems by providing solutions that simplify and automate system management tasks in addition to providing efficiency gains. Sun's own experience in this area is vivid proof, with the efficiency of IT operations at Sun improving dramatically in recent years. Whereas a few years ago a Sun system administrator could manage approximately 20 systems, the ratio today is one system manager for every 50 servers.

## *Managing Server Sprawl*

Sun can help solve the system management problems associated with managing an extensive network of systems by allowing fine-grained control of Sun systems, and by providing the ability to aggregate servers into SunPlex systems that can be more easily managed.



---

*Sun™ Management Center 3.0* software gives IT professionals a single point to proactively manage and monitor remote Sun systems, storage components, the Solaris Operating Environment, and applications. System managers can see their entire network of Sun systems including both the edge of the network and the heart of the data center. This feature-rich next-generation software offers comprehensive knowledge of Sun components, allowing management to the board level in hardware and to the kernel level in the Solaris Operating Environment. A Sun system can be viewed as a server, or as a domain within a server.

The adaptable Sun Management Center also easily integrates with leading third party enterprise management platforms such as Computer Associates' *Unicenter*, BMC *Patrol*, HP *OpenView*, and Tivoli *TES* to give businesses the added flexibility to manage their Sun environment as part of a heterogeneous enterprise management framework.

Resource aggregation using SunPlex systems also allows efficiencies in systems management and can help administrators manage server sprawl. Sun Cluster software provides a single management view of the entire SunPlex system so that all nodes of the cluster can be managed as a single aggregated resource. The system administrator manages the SunPlex cluster and then the SunPlex system automatically manages all of its components. For example, an administrator can install a software component on a single node in the SunPlex system and then run it on any other node in the cluster without having to reinstall it. This saves valuable time and reduces the chance for errors and outages. Administrators can also add or remove nodes while the SunPlex system is on-line. Different models of Sun servers can be intermixed within the SunPlex system to meet specific needs. System management efficiency will increase even further when the current limit of eight nodes per SunPlex system is extended in future releases of Sun Cluster software.

To better manage user service levels, system administrators can use SunPlex systems to provide an aggregated view of all of the software components that make up each Web service running in the Plex. This allows administrators to quickly scan an application service and drill down into more detail if a problem is detected. Sun is continuing to drive its development efforts to provide such aggregated views outside of a Plex. In the future, an entire application service will be viewable even if components of the service are spread across multiple separate Sun systems.



---

## *Addressing Heterogeneous Networks*

In an effort to address the cost of managing heterogeneous networks, Sun has been an active participant in the Distributed Management Task Force (DMTF), an industry group which has developed a new enterprise management standard. The Common Information Model (CIM) Web-Based Enterprise Management (WBEM) model developed by the DMTF unifies the management of enterprise environments with a set of management- and Internet-standard technologies. Through WBEM, companies can develop and deliver a set of standards-based management tools that leverage new technologies, such as the eXtensible Markup Language (XML) and the Common Information Model (CIM), operating over HTTP.

To enable other applications to more easily manage the Solaris Operating Environment, Sun has released *Solaris™ Web-Based Enterprise Management (WBEM) Services* as a core element of the Solaris 8 Operating Environment. The *Sun™ WBEM Software Development Kit (SDK)* also delivers a comprehensive set of Java tools and services for the development of management applications and instrumentation to manage Solaris software as well as other platforms.

Other system management offerings from Sun are described below and additional tools are listed in the summary table end of this section.

## *Manageability at the Heart of the Data Center*

Managing server sprawl helps create efficiencies in the overall management environment. However, the heart of the data center has another issue that is central to its mission. Resource utilization rates must be closely monitored to maintain service levels and minimize capital expenditures on new systems.

System administrators need flexibility to quickly change resource allocations when spikes in user demand occur. To do this, they need visibility and control of resources at a detailed level. Sun Management Center provides detailed control of Sun components both within the heart of the data center and the edge of the network.

*Solaris™ Resource Manager* takes resource control a step further by allowing administrators to allocate system resources to a specific application service. It uses dynamic resource allocation based on policies set by administrators to control how resources are utilized. Combining Sun Management Center



---

software and Solaris Resource Manager software allows businesses to deliver a consistent level of service across multiple application services that are sharing system resources through the same server or the same SunPlex system.

In addition, the heart of the data center is where directory access is controlled. The *iPlanet™ Directory Server* delivers a high performance, highly scalable, user-management infrastructure that helps organizations manage identity, relationships, and risk. By centralizing users, groups, and access controls across multiple applications, iPlanet Directory Server dramatically simplifies administration. When implementations cross company boundaries, iPlanet Directory Server Access Management Edition and Integration Edition work together to unify user-identity account creation, management, and deletion. iPlanet Directory Server provides simplified management and better integration of directory management across company boundaries to deliver lower total cost of ownership.

Businesses can also reduce their high mainframe hardware and software costs by migrating existing CICS applications to the Sun platform. CICS applications run unchanged on the Solaris Operating Environment using technology obtained in Sun's recent acquisition of Critical Path's mainframe rehosting business.

### *Manageability at the Edge of the Network*

The edge of the network is a lights-out management environment where many of the components being managed are likely to be remotely located. Here, the administrator might only want to see data for those servers in which an error condition exists, or in which a threshold utilization rate has been crossed. The healthy nodes can be ignored unless the administrator is looking for an idle resource that can be redeployed to help scale another service.

The administrator may also want to view data for a logical grouping of servers such as those that support Internet or wireless access. Sun Management Center software allows these logical groupings to be established so that the administrator can quickly scan a category of servers to isolate errors or monitor operations.



---

### *Replicating Servers at the Edge of the Network*

To avoid complexity, businesses want exact copies of the software on replicated servers at the edge of the network. The system administrator's task of rapidly deploying or redeploying servers at the edge of the network is made easier by Solaris™ Web Start software.

*Solaris Web Start* software speeds up the process of installing a new server at the edge of the network. The latest edition to the Solaris Web Start product line, *Solaris Web Start Flash*, provides new installation and provisioning functionality that gives system administrators the power to capture a snapshot image of a complete server. This minimizes the risk of errors in the installation by deploying copies of proven and tested software stacks. The Solaris Web Start Flash image includes not only all of the applications installed on the server, but also the specific configuration of the Solaris Operating Environment used on that server. Using this system image, administrators can then replicate server configurations onto multiple (clone) servers for rapid deployment. This provides the ability to repurpose servers on demand.

With Solaris Web Start Flash technology, installation time can be cut significantly. Tests conducted by Sun have shown that a complete server, including the operating system, middleware, and applications, can be set up using the network install feature of Solaris Web Start Flash in under three minutes. Solaris Web Start Flash can also be used with Solaris Live Upgrade to reduce the time required to update the Solaris Operating Environment on a set of identically configured servers.

*Solaris Web Start* software also reduces the cost of administration by eliminating the UNIX® system administration normally associated with software installation and setup of the Solaris Operating Environment. The new Web Wizards give developers the power to simplify the installation, setup, and administration of native Solaris and Java applications, allowing software to be setup by novice administrators safely, easily, and economically.



## *Summary of System Management Tools from Sun*

A summary of key system management tools that can help businesses more efficiently manage their IT systems appears in Table 2, along with links where more information can be found.

<b>Sun Solution</b>	<b>Description and URL</b>
Sun™ Management Center	Single point of management for all Sun systems <a href="http://www.sun.com/sunmanagementcenter/">http://www.sun.com/sunmanagementcenter/</a>
Solaris™ Resource Manager	Control and allocate system resources within Sun systems <a href="http://www.sun.com/software/resourcemgr/">http://www.sun.com/software/resourcemgr/</a>
iPlanet™ Directory Server	User management infrastructure <a href="http://www.iplanet.com/products/iplanet_directory/">http://www.iplanet.com/products/iplanet_directory/</a>
Sun™ Cluster Software	Sun's next generation clustering software <a href="http://www.sun.com/software/cluster/">http://www.sun.com/software/cluster/</a>
Solaris™ WBEM Services	Manage the Solaris Operating Environment from other enterprise management environments <a href="http://www.sun.com/solaris/wbem/">http://www.sun.com/solaris/wbem/</a>
Solaris™ Web Start	Simplify the installation of the Solaris Operating Environment and bundled software <a href="http://www.sun.com/solaris/webstart/">http://www.sun.com/solaris/webstart/</a>
Solaris™ Web Start Flash	Capture a snapshot image of a complete server <a href="http://www.sun.com/solaris/webstartflash/">http://www.sun.com/solaris/webstartflash/</a>
Solaris™ Bandwidth Manager	Control the bandwidth assigned to particular applications, users, and departments <a href="http://www.sun.com/software/bandwidth/">http://www.sun.com/software/bandwidth/</a>
Sun Trunking™	High performance networking using multiple Ethernet connections <a href="http://www.sun.com/products-n-solutions/hw/networking/connectivity/suntrunking/">http://www.sun.com/products-n-solutions/hw/networking/connectivity/suntrunking/</a>
Solstice AdminSuite™	Administer distributed systems <a href="http://www.sun.com/software/adminsuite/">http://www.sun.com/software/adminsuite/</a>
Sun Remote Services	Remote monitoring services <a href="http://www.sun.com/service/support/srs/">http://www.sun.com/service/support/srs/</a>
BigAdmin portal	Community-driven repository for system administrators <a href="http://www.sun.com/bigadmin/">http://www.sun.com/bigadmin/</a>

Table 2 Sun system management solutions



---

## *Security*

Businesses need comprehensive, integrated end-to-end security based on open standards. By combining people, process, and the right technologies, tight security can be achieved.

Sun solutions are based on the robust and secure Solaris 8 Operating Environment with its key security features and an open standards approach that allows third party security solutions to be integrated into a single end-to-end security architecture. The Solaris Operating Environment can easily establish secure network communications from host to host and from site to site using encrypted Internet Protocol Security Architecture (IPSec) connections. It also provides strong public and secret key-based authentication and single network sign-on techniques. Role-based access control also reduces the risk of human error and malicious threats by assigning system manager privileges based on roles. This allows junior system managers to perform certain tasks such as system backups without the risk of inadvertently modifying a system configuration.

User authentication and encryption using a secure portal environment is also critical to end-to-end security. iPlanet Portal Server works well in a heterogeneous environment because it is integrated with a wide variety of authentication methods, such as Microsoft Windows NT domains, UNIX log-in, X.509 certificates, LDAP, RADIUS, and token-based mechanisms. In addition to authentication, the iPlanet Portal Server maintains the status of a user's session and enables single sign-on capabilities with portal applications and services.

The iPlanet Portal Server enables secure remote access to the portal and other Intranet resources through a firewall, making it possible to provide access to all of the confidential content, applications, and services a user might need from anywhere on the Internet using only a Java technology-enabled Web browser.

Service offerings from Sun Service can also help businesses pinpoint possible security holes in their systems architectures. Additional information about Sun security solutions can be found at <http://www.sun.com/security/>.

## Sun Products — Enabling the Service Point Architecture

Sun will continue its efforts to develop solutions that are easy to deploy as components in the Sun Service Point Architecture. Future products will further simplify both implementation and management of systems and will further automate operations across the spectrum from the heart of the data center to the edge of the network.

Figure 4 illustrates many current Sun products that address the needs of both the heart of the data center and the edge of the network. The breadth of Sun's offerings and the level of integration between the various systems and the software components provides unique value by simplifying management of the overall IT environment.

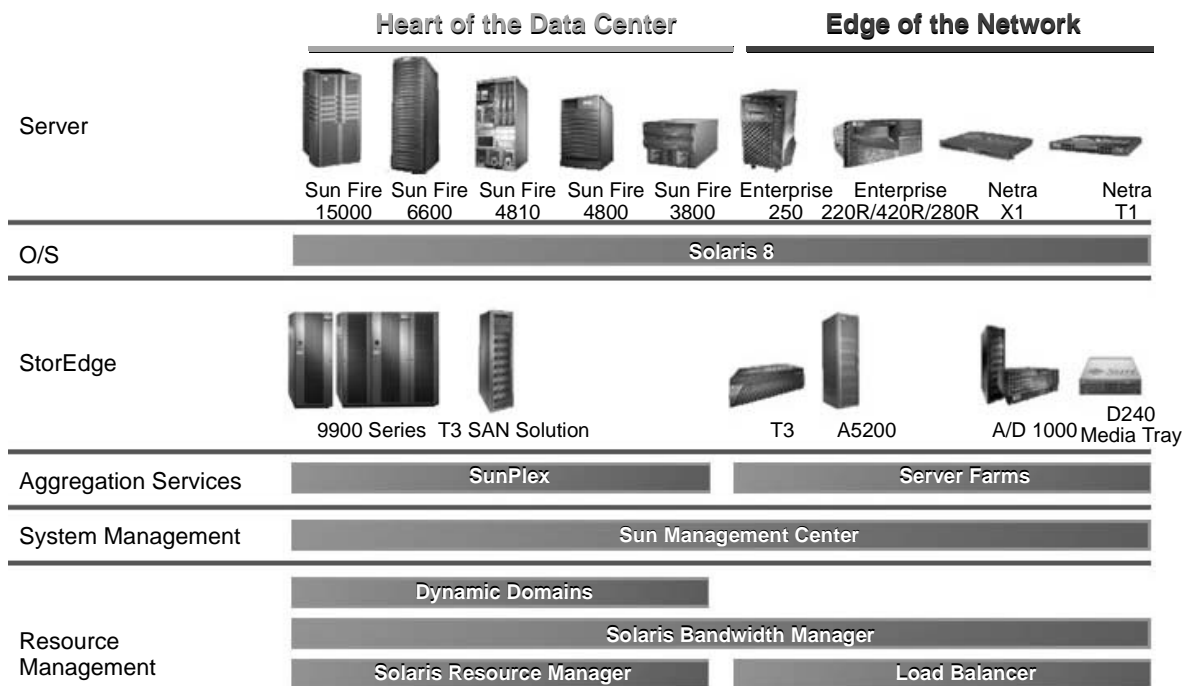


Figure 4 Sun Products in the Sun Service Point Architecture



---

## *Related Programs and Services from Sun*

In the age of services on demand and the Net Effect, availability is of paramount importance. Keeping availability high means relying on people and processes even more than products. Sun helps businesses invest in people and processes through the services and programs described below.

### *Sun Service Offerings*

Sun Service experts can help integrate the methodologies, expertise, products, and services needed to craft the kind of high availability environment required to make services on demand work smoothly. Indeed, Sun's comprehensive portfolio of consulting, training, and support services is one of the many reasons that the majority of Internet traffic runs through Sun servers.

### *Sun Professional Services*

Sun Professional Services provides the technology and business expertise to turn plans into real business solutions. In addition to consulting services for a variety of distributed network computing needs, Sun Professional services offers specialized services that include diagnostic testing, best practices approaches, and documentation to help maximize service levels.

### *Sun Educational Services*

Well-trained staff can help ensure high levels of availability, fast and accurate deployment of new products and services, quick identification of problems, and rapid response to system failures and interruptions. As a leading supplier of training solutions around the world, Sun can help organizations take full advantage of their hardware and software investments. A range of training courses as well as customized assessments and training are available from Sun Educational Services.

### *Support Offerings*

Sun offers a variety of support programs designed to simplify and streamline the installation and operation of modern data centers to keep them running smoothly around the clock.



---

Some key Sun support offerings include:

- Sun RAS Profile, a preemptive service to evaluate the reliability, availability, and serviceability of enterprise computing operations
- Sun Remote Services — remote, real-time monitoring by Sun with alerts when service levels deteriorate
- Sun Installation Services — for accurate configuration and timely completion of installations
- SunSpectrum<sup>SM</sup> Support — a range of support offerings for hardware service

### *The SunTone<sup>SM</sup> Certification and Branding Program*

The SunTone<sup>SM</sup> Certification and Branding Program includes a rigorous set of assessments that focus on the impact of people, products, and processes across multiple dimensions of the service delivery environment. SunTone certification demonstrates Quality of Service and Best Practices for outsourced applications, services and integrator services.

The SunTone Portfolio containing profiles of companies participating in the SunTone initiative is available online at <http://www.suntone.org/sp/suntone/find/>.

## *The Value of the Sun Service Point Architecture*

Sun understands that IT systems which scale to support millions of users can be difficult and costly to manage. Businesses need to maintain high service levels, but are eager to reduce the cost of their IT expenditures. In addition, existing investments in application software to run the business cannot be overlooked.

The Sun Service Point Architecture provides a foundation for delivery of services on demand while protecting investments in legacy information assets. Its modular approach allows software components to be mapped onto hardware platforms which can then be tailored for dedicated functions.

Sun brings unique value by helping enterprises build solutions that can be efficiently managed and easily adapted to changing business needs. Sun solutions allow managers to aggregate resources and add flexibility to data center operations without giving up reliability.



**Service Level**

**Service Level  
Cost**



Figure 5 Value of the Sun Service Point Architecture

Businesses that implement this architecture using Sun technologies will have three major advantages:

- *Greater simplicity* in managing the overall environment through better instrumentation and the ability to aggregate resources
- *More flexibility to reallocate resources* and maintain systems without requiring downtime
- *Reduced risk* by following *industry standards* such as Java, XML, J2EE, and CIM WBEM

For nearly twenty years, Sun has been busy moving our customers toward the Network Age, proclaiming that “The Network is the Computer™”. Today we are helping enterprises take advantage of the Web to build a whole new level of competitive advantage.



## *For More Information*

The following *URLs* contain further information as indicated:

<b>Web site URL</b>	<b>Description</b>
<a href="http://www.sun.com/solaris/">http://www.sun.com/solaris/</a>	Solaris Operating Environment
<a href="http://www.sun.com/servers/">http://www.sun.com/servers/</a>	Sun hardware servers
<a href="http://www.sun.com/sunmanagementcenter/">http://www.sun.com/sunmanagementcenter/</a>	Sun Management Center software
<a href="http://www.sun.com/storage/">http://www.sun.com/storage/</a>	Sun Network Storage solutions
<a href="http://www.sun.com/security/">http://www.sun.com/security/</a>	Sun security solutions
<a href="http://www.sun.com/clusters/">http://www.sun.com/clusters/</a>	Sun Cluster software
<a href="http://www.sun.com/suntone/">http://www.sun.com/suntone/</a>	SunTone Certification and Branding Program
<a href="http://www.sun.com/service/">http://www.sun.com/service/</a>	Sun Service offerings
<a href="http://www.sun.com/sunone/">http://www.sun.com/sunone/</a>	Sun Open Net Environment (Sun ONE)
<a href="http://www.sun.com/products-n-solutions/software/management/">http://www.sun.com/products-n-solutions/software/management/</a>	Sun Systems management software solutions

Table 3 Web links for further information

Contact your local Sun sales representation to learn more about how Sun can help your business build competitive advantage with a computing infrastructure that increases service levels while reducing costs.



## Sales Offices

Africa (North, West and Central): +9714-3366333  
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: +33-01-30-67-50-00  
Germany: +49-89-46008-0  
Greece: +30-1-618-8111  
Hungary: +36-1-202-4415  
Iceland: +354-563-3010  
India: +91-80-5599595  
Ireland: +353-1-8055-666  
Israel: +972-9-9710500  
Italy: +39-039-60551  
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: +603-264-9988  
Mexico: +52-5-258-6100  
The Netherlands: +00-31-33-45-15-000  
New Zealand: +64-4-499-2344  
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  
    Hong Kong: +852-2202-6688  
    Shanghai: +86-21-6466-1228  
Poland: +48-22-8747800  
Portugal: +351-21-4134000  
Russia: +7-502-935-8411  
Singapore: +65-438-1888  
Slovak Republic: +421-7-4342 94 85  
South Africa: +2711-805-4305  
Spain: +34-91-596-9900  
Sweden: +46-8-631-10-00  
Switzerland:  
    German: 41-1-908-90-00  
    French: 41-22-999-0444  
Taiwan: +886-2-2514-0567  
Thailand: +662-636-1555  
Turkey: +90-212-335-22-00  
United Arab Emirates: +9714-3366333  
United Kingdom: +44 0 1252 420000  
United States: +1-800-555-9SUN OR +1-650-960-1300  
Venezuela: +58-2-905-3800  
Worldwide Headquarters:  
    Sun Microsystems, Inc.  
    901 San Antonio Road  
    Palo Alto, CA 94303-4900 USA  
    Phone: 650-960-1300 or 800-555-9SUN  
    Internet: [www.sun.com](http://www.sun.com)