Better By Design — The Solaris™ 9 Operating Environment

A Technical White Paper
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Solaris Design Principles</td>
<td>4</td>
</tr>
<tr>
<td>Designed to Evolve</td>
<td>4</td>
</tr>
<tr>
<td>Processor- and Platform-Specific Code</td>
<td>5</td>
</tr>
<tr>
<td>The Solaris Operating Environment Kernel</td>
<td>5</td>
</tr>
<tr>
<td>Device Drivers</td>
<td>5</td>
</tr>
<tr>
<td>STREAMS</td>
<td>6</td>
</tr>
<tr>
<td>Resource Management</td>
<td>6</td>
</tr>
<tr>
<td>File Systems</td>
<td>6</td>
</tr>
<tr>
<td>Directory Services</td>
<td>6</td>
</tr>
<tr>
<td>UNIX APIs</td>
<td>6</td>
</tr>
<tr>
<td>Shared Libraries</td>
<td>7</td>
</tr>
<tr>
<td>Linux Libraries</td>
<td>7</td>
</tr>
<tr>
<td>Java Technology Virtual Machine</td>
<td>7</td>
</tr>
<tr>
<td>Binary Compatibility</td>
<td>7</td>
</tr>
<tr>
<td>Open Standards-Based Implementation</td>
<td>8</td>
</tr>
<tr>
<td>UNIX Platform Open Standards</td>
<td>8</td>
</tr>
<tr>
<td>Linux Interoperability and Compatibility</td>
<td>8</td>
</tr>
<tr>
<td>Interoperability Open Standards</td>
<td>8</td>
</tr>
<tr>
<td>Java Technology Interoperability With IPv6 Networking</td>
<td>8</td>
</tr>
<tr>
<td>Committed to Excel</td>
<td>9</td>
</tr>
<tr>
<td>Reliability, Availability, and Serviceability</td>
<td>9</td>
</tr>
<tr>
<td>Performance</td>
<td>9</td>
</tr>
<tr>
<td>Quality of Service</td>
<td>9</td>
</tr>
<tr>
<td>Security</td>
<td>9</td>
</tr>
<tr>
<td>Scalability to Support Growth</td>
<td>10</td>
</tr>
<tr>
<td>Committed to Scalability</td>
<td>11</td>
</tr>
<tr>
<td>Multiple Page Size Support</td>
<td>11</td>
</tr>
<tr>
<td>Memory Placement Optimization</td>
<td>11</td>
</tr>
<tr>
<td>Multithreading Support</td>
<td>12</td>
</tr>
<tr>
<td>Full 64-bit Environment</td>
<td>12</td>
</tr>
<tr>
<td>64-bit Solaris Operating Environment</td>
<td>12</td>
</tr>
<tr>
<td>64-bit Java Virtual Machine</td>
<td>12</td>
</tr>
<tr>
<td>Network Cache and Accelerator</td>
<td>12</td>
</tr>
<tr>
<td>Enterprise DHCP</td>
<td>13</td>
</tr>
<tr>
<td>Scalability Measurements</td>
<td>13</td>
</tr>
<tr>
<td>Performance on Database-Intensive Operations</td>
<td>13</td>
</tr>
<tr>
<td>Oracle9i Data Warehouse Scalability</td>
<td>14</td>
</tr>
<tr>
<td>Designed to Be Reliable, Available, and Serviceable</td>
<td>15</td>
</tr>
<tr>
<td>Robust</td>
<td>16</td>
</tr>
<tr>
<td>Dynamic System Domains</td>
<td>16</td>
</tr>
<tr>
<td>Kernel and Device Driver Hardening</td>
<td>16</td>
</tr>
<tr>
<td>Memory Scrubber</td>
<td>17</td>
</tr>
<tr>
<td>Kernel and User Mode Separation</td>
<td>17</td>
</tr>
<tr>
<td>Recoverable</td>
<td>17</td>
</tr>
<tr>
<td>File System Journaling</td>
<td>17</td>
</tr>
<tr>
<td>NFS Failover</td>
<td>17</td>
</tr>
</tbody>
</table>
IP Network Multipathing .............................................. 17
Sun StorEdge Traffic Manager .......................................... 18
Cluster Failover ......................................................... 18
Manageable .................................................................. 18
  Dynamic Reconfiguration ................................................ 18
  Solaris Live Upgrade .................................................... 18
  Configurable Dump Device .......................................... 19
  Device Hot-Plug Management ...................................... 19
Observable ..................................................................... 19
  Syslog ....................................................................... 19
  Kernel Analysis Tools .................................................. 19
  Modular Debugger Framework ...................................... 19
Compatible ..................................................................... 20
  Application Binary Interface ......................................... 20
  AppCert Program ......................................................... 20
Protected ........................................................................ 20
  Component Quality ...................................................... 20
  Security Advisories ...................................................... 20
  Role-Based Access Control ......................................... 21
Enhancing Availability ..................................................... 21
  Availability With Sun Cluster Software .......................... 21
  Availability With Sun Management Center ..................... 21
  Availability With Solaris 9 Resource Manager ................ 22
Designed for Manageability ............................................. 23
Data Management ........................................................ 23
  Data Management through Solaris Volume Manager ....... 23
  Soft Disk Partitions ...................................................... 23
  Snapshot Backups ....................................................... 24
  mkafs Performance ....................................................... 24
Resource Management .................................................... 24
  Support for Service Containers ..................................... 25
System Management ...................................................... 25
  Installation and Provisioning ......................................... 26
    Solaris Flash and Solaris Live Upgrade ......................... 26
    Sun Management Center Change Manager .................... 26
Administration ............................................................ 26
  System Administration Facilities ................................... 26
  Patch Management ....................................................... 27
Monitoring and Management .......................................... 27
  Sun Management Center .............................................. 27
  Standards-Based Management ..................................... 27
User Management ........................................................ 28
Designed for Security ..................................................... 29
Network Security .......................................................... 29
  Integrated Kerberos V5 .................................................. 29
  Kerberos-Secure NFS .................................................... 30
  Solaris Secure Shell ...................................................... 30
  IPSec/IKE for IPv4 and IPSec for IPv6 ......................... 30
  SunScreen Security Technology .................................. 31
Platform Security ........................................................ 31
  Modular Software Packaging ....................................... 31
  Extended Auditing ....................................................... 31
  Pluggable Authentication Modules (PAM) ....................... 31
  Secure LDAP Client ..................................................... 32
  Role-Based Access Control (RBAC) .............................. 32
Chapter 1

Introduction

The Internet has dramatically changed the way in which companies provide services. The days of selling multiple copies of desktop PC or client-server software are yielding to a Services on Demand model where companies develop a single copy of their software and deploy it to hundreds, thousands, and millions of users accessing the Internet from any IP-enabled device, anywhere, at any time. The shift towards a Services on Demand model moves the burden of computing infrastructure from the end user and their PCs to the organization providing the services.

The Solaris™ Operating Environment value proposition is to help organizations increase service levels while simultaneously reducing service cost and risk. The key to reaching these goals is using systems that perform on the fundamentals, namely those that are scalable, available, manageable, and secure:

- **Scalable.** Companies offering Web-based services have lost control over how many customers might visit their sites — successful companies are the ones that can grow to support millions of customers almost overnight.
- **Available.** With businesses operating around the clock and around the globe, organizations have lost control of when their customers might demand their services, making availability of paramount importance.
- **Manageable.** With business models changing rapidly and service-delivery platforms growing by orders of magnitude as a matter of course, deploying and redeploying servers is a daily event. Facilities to manage growth are a must, but today's companies must also reduce risk by reducing complexity from using fewer hardware and software components to using servers that can be repaired or upgraded with no downtime.
Secure. As the sophistication of hackers increases and the threat of cyberterrorism looms, providing services to millions of users requires bulletproof security to protect both data and the servers that manipulate them.

Figure 1-1: Providing Services on Demand requires operating environment support that is scalable, available, manageable, and secure.

With the need for rock-solid scalability, availability, manageability, and security, it is no surprise that so many successful organizations are using Sun’s Solaris Operating Environment:

- Gartner Dataquest (February 2002) reports that Sun continues its leadership of the RISC/UNIX® server market, with more than two out of every three servers shipped in 2001 coming from Sun and running the Solaris Operating Environment.
- A Netcraft Web server survey (May 2001) indicates that the majority of the top 50 Web sites are running Solaris software.
- An Infonetics Research report (2000) notes that 89 percent of the UNIX servers at Tier-1 service providers run the Solaris Operating Environment.

While Sun continues to capture market share because of its focus on a single operating environment for its entire line of binary-compatible servers, the competition is increasingly left behind:

- As IBM pursues a fragmented strategy that requires them to support no fewer than 14 different operating environments (www.ibm.com), Sun earned D.H. Brown Associates #1 UNIX operating system award (2001 UNIX Function Review, March 2001). D.H. Brown cited Sun’s total focus on UNIX systems and its support of a “single line of binary-compatible platforms that span the range from very price-competitive low-end servers up to clusters of high-end, mainframe-like SMP servers.”
- Microsoft’s backpedaling due to customer dissatisfaction over its new licensing schemes, as well as its inability to deliver secure, robust Web server software has Gartner Group (September 2001) recommending that those using Microsoft’s IIS switch to Sun™ ONE Web Server (formerly iPlanet™ Web Server) or open-source Apache — both products that ship with the Solaris Operating Environment.
As the competition struggles with their own flawed business models, Sun’s delivery on the fundamentals is recognized with Network Computing’s “Best Server Operating System 2001” award, declaring that “if Sun Microsystems has proven one thing, it’s that its operating system can stand the test of time and still beat the competition.”

Sun’s Solaris Operating Environment delivers on the fundamentals because it is better by design, enabling it to more effectively support the high-level infrastructure that organizations need to provide services on the Internet. Because it is scalable, available, manageable, and secure, the core Solaris Operating Environment provides optimal support for the wide range of services necessary to support both enterprise and Internet-based Services on Demand environments.

The Solaris Operating Environment supports mainframe-quality database servers and secure, flexible, front-end Web servers. It supports software components, including the Java™ 2 Platform, Standard Edition (“J2SE”) and the Java 2 Platform, Enterprise Edition (“J2EE”) environments, and service-delivery platforms, including the Sun ONE Web Server, Sun ONE Application Server (formerly iPlanet Application Server), and Sun ONE Portal Server (formerly iPlanet Portal Server). And because it is based on open standards, Sun invests its efforts in creating superior implementations rather than scheming vendor lock-in strategies. For customers wishing other competing implementations of services, there are more than twelve thousand third-party applications running on the Solaris Operating Environment, giving customers greater choice.

This white paper discusses how the Solaris Operating Environment is better by design, and how it provides superior scalability, availability, manageability, and security.
Chapter 2

Solaris Design Principles

Sun designed the Solaris Operating Environment for performance and reliability; to scale and evolve as business needs change, and to readily support new hardware architectures as they become available. The Solaris Operating Environment is better by design, resulting in a system that grows and extends in an evolutionary fashion that doesn’t require major, error-prone rewrites from one release to the next. This chapter discusses the design principles and business practices that enable the Solaris Operating Environment to provide the scalability, availability, manageability, and security that users, networks, and mission-critical and enterprise applications demand.

Designed to Evolve
The Solaris Operating Environment is designed in a modular fashion, so that it can adapt to new hardware platforms, incorporate new features, and change configurations without the need for rebooting. (Figure 2-2). The Solaris kernel is dynamic, composed of a core system that is always resident in memory, with services beyond the core loaded as needed. What this means for customers is that Solaris software can evolve to accommodate new devices and services without even rebooting — resulting in less downtime and greater agility to meet changing business needs.
Processor- and Platform-Specific Code

At the layer closest to the hardware is platform- and processor-specific code that enables Solaris software to easily port to different processors and system architectures. This code comprises less than five percent of the operating system kernel, resulting in an agile, adaptable code base that gives Sun the ability to rapidly bring new SPARC® processors and servers to market as those technologies evolve. Whether it is a brand-new UltraSPARC® processor or a server with more processors than ever imagined, the Solaris Operating Environment can rapidly adapt.

The Solaris Operating Environment Kernel

One level up from the platform- and processor-specific code is the Solaris Operating Environment kernel, a compact code base that is built to be extended. Unlike other operating environments where the line between the operating system and applications is confused, Sun designed the Solaris Operating Environment so that the line is very clear. This clarity enables a smaller code base that results in a more reliable and secure operating environment. It is nearly impossible, for example, for a Web browser to crash the operating system because it is clear to Solaris software what is an application and what is an operating system feature.

Device Drivers

Device drivers plug into the Solaris kernel and provide access to I/O devices, including disks, tapes, CD drives, serial ports, and networks. Above these device drivers, software is layered to support device-specific functions. For example, the TCP/IP stack uses networking devices; volume management and file system software use disk device drivers. In the Solaris Operating Environment, new device drivers can be added dynamically, enabling the kernel to be reconfigured without rebooting.

One very useful way in which drivers are used in the Solaris Operating Environment is to manage large numbers of disks as a single volume. Using disk management software like Solaris Volume Manager, software RAID arrays as well as disk mirrors, striping, and even the capability to add disk volumes to existing logical file systems can be created. Organizations can layer volume management under whatever file system software they wish because the existence of the underlying driver is transparent.
Device drivers for the Solaris Operating Environment are written using published, stable interfaces to the software platform. These interfaces do not change from release to release, so when a vendor delivers a driver for one version of Solaris software, they will continue to work in the next release. Sun is careful to avoid any changes that might undermine the work of its internal and third-party device driver developers.

STREAMS
STREAMS provide a software backplane into which modules can be plugged to dynamically build higher-level services. STREAMS increase flexibility in the use of devices so that new and custom protocols and line disciplines can be configured using the same underlying physical device. They allow various modules to dynamically be pushed onto a stack. An underlying device driver passes data (using standard interfaces) to the next module in the stack, which performs its unique processing and in turn passes data along to the next module and ultimately to a user process.

High-level services can be dynamically configured with STREAMS without rebooting the kernel. For example, a point-to-point (PPP) protocol stack can be built using a serial device driver and layering in modules to support PPP network control protocols and the TCP/IP protocol itself. No major operating environment updates or changes to the code are necessary.

Resource Management
Resource management facilities enable fine-grained control of which processes and tasks get resources. This software includes a fair-share scheduler that supports task- and project-based scheduling and accounting priorities, processor sets that enable specific named processes to be bound to specific processors, and resource controls that regulate the rate at which a task (which can be a collection of processes) consumes resources, including network bandwidth, number of processes, and threads per process.

File Systems
File systems provide another way in which the capabilities of the Solaris Operating Environment can be dynamically configured and extended to meet business needs such as managing disk volumes and files. Any third party who writes software that uses the Virtual File System (VFS) interfaces can provide file systems that integrate with the Solaris environment. Examples of the supported extensions include the high-performance UNIX file system (UFS), the PC file system (PCFS), the ISO 9660 CD-ROM file system known as the High Sierra File System (HSFS), and the VERITAS file system (VxFS).

Directory Services
The Sun ONE Directory Server (formerly iPlanet Directory Server) is integrated into the Solaris 9 Operating Environment, creating a Lightweight Directory Access Protocol (LDAP)-based environment for comprehensive information assets management, including user names and attributes such as passwords, as well as group, project, and task definitions.

UNIX APIs
The Solaris software kernel provides a set of well-documented, public APIs that are used by the writers of the Java virtual machine, shared libraries, and additional libraries to support Linux commands. In contrast to Microsoft’s proprietary, constantly shifting operating system interfaces, Sun’s open APIs invite others to compete for the best product implementation, improving product quality and resulting in greater customer choice.
Shared Libraries
The Solaris Operating Environment supports extensions through its shared library mechanism. This technology allows multiple applications to share large amounts of system software. Clearly separated from the kernel address space, the Solaris Operating Environment supports multiple versions of the same shared libraries so that old and new applications requiring different libraries can run side-by-side. Organizations can be free to consolidate multiple functions on the same server — a key feature for companies under pressure to reduce costs. This capability stands in stark contrast to Microsoft Windows operating systems, where applications often install new versions of Dynamically-Linked Libraries (DLLs) over existing ones, creating incompatibilities that result in the new application’s DLL causing existing applications to stop working.

Because Sun develops both the processors and the platforms on which the Solaris Operating Environment runs, it is free to make processor and architecture-specific software optimizations. In the case of shared libraries, Sun overlays different shared libraries at install time depending on the underlying system. Transparent to application software, these libraries help to fully exploit the capabilities of the underlying hardware.

Linux Libraries
Sun encourages the sharing of software between Linux and Solaris environments, and facilitates developers using the Solaris Operating Environment to write Linux applications by including six key Linux libraries with Solaris software. Glib supports useful data types, macros, type conversions, and a lexical scanner. Gtk+ supports graphical user interfaces, including the GNOME desktop. The JPEG, PNG, and TIFF libraries support creating compressed image files in the corresponding formats. And the XML2 library supports structured XML documents through a C language interface. Through support of these libraries, many Linux applications can compile and run virtually unchanged in the Solaris Operating Environment.

Java Technology Virtual Machine
Sun’s Java virtual machine (JVM™) software is the key technology supporting Web-based services. Java technology enables much of the software that drives the Web today, from Java Servlet and JavaServer Pages™ (JSP™) software to the business logic in application servers. Because the Java run-time environment is so important for organizations providing Services on Demand, Sun has invested heavily in making its JVM machine the best it can be.

Sun builds both the hardware and software platform, which enables a tighter impedance matching between the two. With improvements in the Solaris kernel’s support of threads, optimizations for the UltraSPARC III processor pipeline, and a new parallel garbage collection algorithm, Sun sets the standard against which other vendors compete.

Binary Compatibility
The Solaris Operating Environment supports binary compatibility across the product line. This means that any service, once compiled, runs on every SPARC platform, from entry-level workstations to the 106-processor Sun Fire™ 15K server, offering unprecedented flexibility in deploying services. Because Sun maintains a constant focus on a single operating environment, companies can scale their services horizontally (by adding more servers) or vertically (by adding more computing resources to a single server) without worrying about whether their application software will run. Contrast this to the competition, where moving up the product line means changing server architectures and operating environments — all costly propositions while customers are demanding increasingly higher levels of service.
Open Standards-Based Implementation
Sun promotes and adheres to open standards, believing that healthy competition in an open marketplace fosters superior implementations, higher quality, and lower costs. Indeed, even with innovations that originate at Sun — ranging from the Network File System (NFS) to Java technology — Sun’s philosophy is to make interfaces public and stable, and successfully compete with other vendors by providing the best implementation, rather than locking out rivals by obscure or constantly changing interfaces.

UNIX Platform Open Standards
Portability of applications between UNIX platforms is enhanced by Sun’s endorsement and support of open standards. Sun adheres to X/Open®, POSIX, UNIX 98, and POSIX thread standards for interoperability between UNIX implementations from competing vendors.

Linux Interoperability and Compatibility
Linux systems interoperate easily with Solaris Operating Environment-based servers and workstations because they share many of the same TCP/IP-based network communication mechanisms, including Telnet, FTP, NFS, and Samba.

Linux commands compile and run in the Solaris Operating Environment virtually unchanged, and the Linux Compatibility Assurance Toolkit (LinCat) further simplifies the process of assuring that Linux applications will run on Sun workstations and servers.

With hundreds of Linux applications and libraries provided with the Solaris Operating Environment, Linux users will find almost all of the tools and commands they are familiar with, right down to the GNOME desktop environment. Because of the popularity of GNOME, Sun plans over time to make it the standard desktop environment for the Solaris Operating Environment.

Interoperability Open Standards
Open standards supporting interoperability between systems is key to today’s networked environments. For example, Sun’s implementation of the J2EE environment is compelling to developers because of the wide range of open standards that the software embraces, including CORBA, JDBC®, RMI, SOAP, UDDI, XML, and XML RPC specifications.

Java Technology Interoperability With IPv6 Networking
To make it easier for organizations to use IPv6 to alleviate addressing issues with large networks, Sun’s Java virtual machine has been extended to support the 64-bit IP addresses required for IPv6 interoperability.
Committed to Excel

Reliability, Availability, and Serviceability

The degree to which the Solaris Operating Environment’s reliability, availability, and serviceability (RAS) exceeds the competition has earned it first place in D.H. Brown’s analysis of UNIX operating systems. To advance even beyond this prestigious ranking, Sun has set its sights on becoming number one among the larger, entire server market that also includes mainframes. To take this next step, Sun is integrating RAS into every step of the development process with formal methods for improving what is already rated the best UNIX operating system on the market. RAS is such an important facet of the Solaris Operating Environment — and such an important need of Sun customers — that an entire chapter in this white paper is devoted to the topic.

Performance

Sun is continually improving the Solaris Operating Environment so that it can squeeze every last bit of performance out of Sun’s potent symmetric multiprocessor systems. One step that Sun takes to ensure continually improving performance is the execution of more than 50 performance benchmarks on every operating environment build.

Quality of Service

Sun’s Solaris 9 Operating Environment has an unprecedented number of resource management facilities that enable service providers and enterprise data centers to manage service levels for their services and their networks. Depending on the underlying server capabilities, Dynamic System Domains can assign an entire electrically isolated operating environment to a single service, with multiple domains possible on a single server. Within a single server domain, multiple services can run in separate “containers,” each of which has fine-grained control over the services’ resource consumption and performance. This quality-of-service management capability saves companies money because they can consolidate many services onto a single server that costs less to maintain.

Security

There’s often much press devoted to operating systems being unsecure, but not the Solaris Operating Environment, because it is more secure. Security is one reason why 89 percent of Tier-1 Internet service providers use Sun (Infonetics Research, 2000). Solaris security is far beyond the competition because Sun has built systems with the network in mind since 1982, long before the Internet took its current form. With more security issues hammered out of the system with the test of time, Sun has gone on to build innovative security software like SunScreen™ technology, whose stealth-mode protection protects networks without detection, and whose technology is shipped with every copy of the Solaris 9 Operating Environment.

When Sun develops revolutionary software like Java technology, security is built in from day one. For example, the Java technology sandbox security model requires obtaining user permission before a downloaded applet can access resources on the local machine. Imagine how many viruses would have been impossible to develop if desktop PC software were written with this level of security in mind.
Scalability to Support Growth

Scalability is critical to companies needing to make the most cost-effective use of their computing resources, gracefully handle peak workloads, and grow their computing environment as rapidly as their business grows. There are three aspects to scalability:

• How the system responds to increasing workloads: what performance it provides, and as the workload increases, whether it crashes or allows performance to gracefully degrade.
• How easy it is to add processors, CPUs, storage, and I/O resources to a system or network in order to serve increasing demands from users.
• Whether the same environment can support applications as they grow from low-end systems to mid-range servers and mainframe-class systems.

The Solaris Operating Environment delivers superior capabilities in all aspects of scalability, providing rock-solid reliability as workloads increase and offering the most processing power available for applications. It provides the ability to add or reconfigure resources such as CPUs, memory, and I/O devices to supported systems without requiring a system shutdown — enabling rapid responses to changing workloads. As business needs grow, Sun’s binary compatibility means that applications can move up the product line from desktop workstations to the most powerful servers that Sun sells — all without touching the application software or changing operating environments. Binary compatibility across the product line, plus unparalleled ability to scale, lifts the constraints on how customers deploy their applications and provides investment protection and reduced administration and training costs because only one operating environment needs to be supported.
Committed to Scalability

Nowhere is Sun's commitment to scalability more apparent than in the Solaris Operating Environment's support for massive systems and applications, including:

- Up to 128 CPUs in a single system and 848 CPUs in a clustered environment
- Configurations with up to 576 GB of main memory
- On 64-bit processors, a 64-bit process address space, supporting extremely large applications with enormous amounts of data
- One million processes
- One terabyte native file systems, and up to 252 TB file systems using Sun StorEdge™ QFS software
- Four billion open sockets, providing practically limitless scalability for network applications
- IPv6 support, enabling a 128-bit IP address space and lifting the limits faced by large organizations and global service providers
- 64-bit Java virtual machine, enabling very large Java technology applications and seamless integration with IPv6 environments

In addition to enabling applications and their data to grow to almost arbitrary size, the Solaris Operating Environment also provides facilities for applications to run faster by optimizing their memory utilization.

Multiple Page Size Support

Many applications — such as database management systems — use a small amount of code to process a large amount of data, resulting in a small locality of reference for instructions and a large locality of reference for data. In order to improve processor cache behavior for these situations, the Solaris Operating Environment enables applications to have multiple page sizes for different memory segments, ranging from 8 KB to 4 MB. Using multiple page sizes can increase performance by reducing Translation-Lookaside Buffer (TLB) misses. Sun’s implementation allows multiple page sizes to be specified while leaving the binary untouched, changing page size dynamically as the application executes.

Memory Placement Optimization

As both processor and system technologies progress, an increasing number of systems have multiple layers of memory. For example, as chip fabrication techniques improve, more cache memory can be colocated with the processor itself, helping to reduce instruction and data fetch times. The Solaris Operating Environment is designed to be memory hierarchy-aware, enabling it to optimize memory management to best match an application’s memory access patterns.

Rather than forcing recompilation for different memory hierarchies, Sun’s philosophy of binary compatibility across the product line dictates a mechanism whereby the application can give hints as to its data access patterns, and the Solaris Operating Environment then manages the memory hierarchy in the way that best suits the particular server on which it is running.
Multithreading Support

The Solaris Operating Environment achieves superior scalability by using threads as the fundamental unit on which to allocate processors. A thread is an independent sequence of program instructions that can be executed by a processor. Many threads can be executed in parallel by separate processors in multiprocessor systems. The Solaris Operating Environment responds to interrupts, performs driver and background activities, and handles application requests using threads, and the number of possible threads is limited only by the amount of available memory.

Synchronization and locking delays are minimized by designing fine-grained locks to protect kernel data structures. The fully preemptive Solaris kernel also implements priority inheritance for locks, allowing threads running in different scheduling classes and with different priorities to interact with locks without compromising priority or correctness. Delays due to synchronization and locking are minimized by designing locks on fine-grained data objects. Deadlocks are avoided by designing the operating environment to avoid holding locks when the flow of control moves from one kernel subsystem to another.

By using the same mechanism for scheduling both kernel activities and user processes, improvements in threads improve all aspects of performance for applications built to utilize multiple processors. With significant performance enhancements designed into the Solaris 9 Operating Environment threads library, applications automatically receive performance improvement when using either a multithreaded or a multiprocess architecture.

Full 64-bit Environment

With the price of memory continuing to tumble, organizations are reaping the performance advantages of keeping more application data in main memory rather than on disk. Although few applications require an entire 64-bit address space, the impact of more than 32 bits of address space benefits database servers, Web search engines and Web server caches, mechanical computer-aided design, electronic design automation, and high-performance computing applications.

64-bit Solaris Operating Environment

Sun has designed the Solaris Operating Environment to support both 32- and 64-bit applications on the same system, enabling users to benefit from increased memory space for those applications that can utilize it — while providing equal support for 32-bit applications.

64-bit Java Virtual Machine

Complementing the Solaris Operating Environment is Sun's 64-bit Java virtual machine. This design enables both larger Java technology applications and larger amounts of data through the use of 64-bit object pointers. In addition, the Java Native Interface API supports full connectivity with the 64-bit Solaris Operating Environment along with full support of IPv6 connectivity.

Network Cache and Accelerator

A performance enhancement specifically designed to improve scalability of Web servers is the Network Cache and Accelerator (NCA). Using an in-kernel cache, NCA uses kernel memory to increase the performance of HTTP requests handled by Web servers. By responding to requests for cached content, Solaris NCA eliminates the multiple context switch overhead necessary for Web server interaction, speeding performance by up to a factor of five. Solaris NCA sockets are currently used in the Apache Web server, Sun ONE Web Server, and Zeus. By using an interposition library provided by Sun, NCA socket calls can be utilized without modifying the Web server itself.
Enterprise DHCP
Sun’s Enterprise Dynamic Host Configuration Protocol (DHCP) server enables host systems to receive IP addresses and network configuration at boot time. Built to support hundreds of thousands of hosts, Enterprise DHCP is essential for organizations wishing to deploy networks on a massive scale. Enhancements that enable an even greater number of hosts include:

- A multithreaded implementation, enabling the server to respond to multiple requests simultaneously
- A new data store that is more efficient than ASCII files or NIS+ repositories
- A data access architecture that enables third-party modules to store DHCP information into any repository
- Dynamic Domain Name Service (DNS) updates, enabling the DHCP server to automatically update a DNS server when client systems are provided with their preferred host names

Scalability Measurements
Sun’s commitment to scalability is demonstrated with objective, real-world performance measurements. A key measure of scalability is speedup — how much additional processing power is provided to applications with the addition of each new CPU. In an ideal world, each CPU added to a symmetric multiprocessing server would contribute 100 percent of its processing power to additional application performance. For example, adding a second processor would result in an application speedup of 2, and adding a sixty-fourth processor would result in a speedup of 64. This section documents the near-linear scalability on Sun servers executing compute-intensive database operations.

Performance on Database-Intensive Operations
Database management systems are often the largest applications that an organization runs because they manipulate huge amounts of data and require enormous amounts of processing power. This area is where the efficiencies of the Solaris Operating Environment help to unleash the full power of Sun’s symmetric multiprocessing servers.

The first set of data (Figure 3-3) giving testimony to this powerful combination are speedup measurements from a Sun Enterprise™ 10000 server configured with 1, 16, 32, 48, and 64 processors executing various database-intensive operations, including:

- A four-way join on a large decision-support system running an Oracle database
- A TmC data-mining operation
- Amdocs telephone company billing system executing online transaction processing operations
- Scan and aggregation on a decision-support system running DB2
- An ad hoc query against an insurance company online transaction processing system running an Oracle database
The benchmark results show near-linear speedup with up to 64 processors. The best result is the Amdocs telco billing test, with a speedup of 62 out of a possible 64. The worst result in this series is the insurance company ad hoc query, which still demonstrates a speedup of more than 52.

**Oracle9i Data Warehouse Scalability**

A second set of data illustrating the potent combination of the Solaris Operating Environment and Sun servers uses a Sun Fire 15K server to demonstrate speedup of a single query when equipped with 8, 16, 32, 64, and 72 processors and 576 GB of memory (Figure 3-4).

The single query test executed a query on a 1 TB database. The test query performed 8 aggregations per row on table with 819 GB of data, and demonstrated a speedup of 71.25 out of a possible 72 — astonishingly linear speedup for such complex operations.

These benchmarks demonstrate to Sun customers the fact that very large numbers of CPUs can be brought to bear on real business problems without involving the complexity of clustering just for the sake of scalability. Because of Sun's ability to deliver near-linear performance on database queries, customers can use vertical scaling as the primary means of increasing performance with little performance penalty — delivering high quality of service while reducing management costs.
Designed to Be Reliable, Available, and Serviceable

Chapter 4

The Solaris Operating Environment provides a level of reliability that gives users confidence that their long-running and resource-intensive applications will execute without interruption. There is no secret to the source of reliability in the Solaris Operating Environment — Solaris software is reliable by design. Sun designed the Solaris Operating Environment with a small, compact kernel that limits the exposure to errors that can bring a system down, as well as a clear distinction between the kernel, shared libraries, and applications, which further limits the impact of application failures. And when a higher level of reliability and availability is needed, Sun Cluster software can be deployed to provide an integrated, high-availability environment where application-level failover can keep services running with continuous uptime.

Sun understands the value of putting more effort into reliability before a product is shipped to increase customer satisfaction and reduce development costs. To the customer, this means less planned downtime for equipment upgrades and repairs, and also less unplanned downtime due to hardware or software failures.

This solid design and execution puts Sun at the top of the list of UNIX operating system vendors in D.H. Brown’s 2001 study. With Sun already at the top of the reliability, availability, and serviceability (RAS) category, where is there to go? Sun’s goal is to be number one among all operating system vendors by delivering an operating environment with mainframe reliability and Internet agility. Sun plans to succeed with this goal by increasing its commitment to RAS and further integrating RAS into every aspect of the development process.
Sun's process audits the following characteristics for each software component it develops:

- **Robustness** against user failures that could cause kernel failures
- **Recoverability**, including data protection and failover
- **Manageability** from the administrator's perspective and integration with Sun management tools
- **Observability** with respect to debugging and testing
- **Compatibility** with established APIs
- **Protection** with respect to bad input handling, memory leaks, and security advisories

For example, a new disk driver would have to be *robust* enough to accept a kernel crash dump in the event of an operating environment failure. It would have to be *recoverable* by being integrated into Sun's multiplexed I/O framework, which allows multiple paths to storage devices. It would have to be *manageable* by supporting hot-plug functions. Its *observability* would have to be supported with a driver-specific kernel debugger module and software enabling users to view driver characteristics. It would have to be *compatible* with published interfaces so that it can seamlessly integrate with third-party volume managers and file systems. And finally, it would have to be *protected* against memory leaks as well as protecting the kernel so that a disk failure would not cause a system failure.

These six characteristics are not only intrinsic to each software component that makes up the Solaris Operating Environment, they are also supported by specific features that have bottom-line impacts on its users.

**Robust**

Robust support for critical applications is delivered by features including dynamic system domains, kernel and device driver hardening, a proactive memory scrubber, and a clear separation between user and kernel software.

**Dynamic System Domains**

Dynamic System Domains provide high-level control over allocation of processors, memory and I/O devices to separate, and electrically isolated domains on Sun Fire 3800, 4800, 4810, 6800, 12K, and 15K as well as Sun Enterprise 10000 servers. Unlike static domain support from IBM, the Solaris Operating Environment supports Dynamic System Domains, where system resources can be reallocated on the fly to rapidly adapt to changing workloads or on a scheduled basis so that the balance of resources can be shifted, for example, to support online transaction processing during the day and long data warehouse queries at night.

**Kernel and Device Driver Hardening**

Sun's philosophy is that when any software component is checked into the main source code tree, it must be of production quality. The kernel and device drivers are further hardened by identifying and eliminating panics and kernel memory leaks. Users can review code to eliminate past errors and the potential for future crashes. Sun's enhanced device driver testing framework enables Sun engineers and third-party developers to stress drivers and simulate hardware failures, further validating that the drivers can handle unforeseen circumstances.
Memory Scrubber

Sun’s servers are built with ECC memory throughout, enabling them to correct any single-bit error and detect double-bit errors. Given that it is possible for transient single-bit errors to be introduced by external sources of radiation, it is important to detect and correct them early, reducing the possibility of a single-bit error later becoming a double-bit error. The cache and memory scrubbers built into the Solaris Operating Environment scan for undetected single-bit errors, writing back correct data so that subsequent program references succeed without error. The memory scrubbers note when memory errors are persistent, enabling defective memory modules to be replaced before they can cause a system outage.

Kernel and User Mode Separation

Sun designed the Solaris Operating Environment with a clear separation between processor kernel and user modes. With the kernel occupying its own protected address space, and user-level shared libraries and applications occupying separate user process address spaces, it is extremely difficult for a user-level error to cause a system failure. When compared to the competition, where even Web browsers are inseparable parts of the core operating system, Sun’s RAS advantages are clear.

Recoverable

The Solaris Operating Environment’s ability to quickly recover from errors is due in part to its file system journaling, NFS failover, Network Multipathing, I/O multiplexing, and cluster failover.

File System Journaling

Sun’s file system journaling is integrated into the Solaris 9 Operating Environment, increasing recoverability by reducing file system check times when the operating environment is rebooted. Easily enabled via a configuration option for the UNIX file system (UFS), journaling writes every file system transaction to a journal before writing it to the actual file system structure. This approach protects against corruption of file system metadata, and makes recovery from unexpected shutdowns faster because the operating environment only needs to write journaled transactions to the file system structure at reboot time, rather than having to verify the integrity of the entire file system.

NFS Failover

Network File System (NFS) failover enables client systems to retry network file accesses on an alternate server in the event that the primary server fails. Integrated with the automounter, this feature enables a client system running the Solaris Operating Environment to continue executing applications even when its NFS server goes down.

IP Network Multipathing

Network Multipathing software in the Solaris Operating Environment — called Solaris IP Multipathing (IPMP) — supports multiple paths for IP networks. Under normal operations, network traffic is balanced across multiple devices. In the event that a failed network interface card or an upstream switch or router makes a path unresponsive, Solaris IPMP redirects network activity to the devices that remain operational. When the problem is repaired, Solaris IPMP software restores traffic on the interface. This feature enables the Solaris Operating Environment to automatically recover from single-point failures in a network infrastructure.
**Sun StorEdge Traffic Manager**

Sun StorEdge Traffic Manager software in the Solaris Operating Environment supports multiple paths for I/O devices, such as SCSI and Fibre Channel-accessible storage. Not only does I/O multiplexing balance the workload across multiple devices, it increases reliability by redirecting requests from a failed interface card or storage device to the one that remains operational.

**Cluster Failover**

The previously mentioned features support recoverability of a single operating environment; Sun Cluster 3.0 software supports recoverability across multiple servers as well as applications themselves. When a server in a Sun Cluster configuration fails, the cluster mechanism automatically transfers the workload from the failed server to the remaining server(s) in the network so that clients continue operation uninterrupted.

**Manageable**

The Solaris Operating Environment is highly manageable, exemplified through its support of Dynamic Reconfiguration, installation through Solaris Live Upgrade, configurable dump devices, and support for hot-plugging devices.

**Dynamic Reconfiguration**

As workloads change rapidly over time, server capacities inevitably must be adjusted through addition and removal of processors, memory, and I/O devices. Likewise, when tools like Sun Management Center software suggest proactive replacement of marginally performing components, server configurations must also be modified. With Dynamic Reconfiguration, processors, memory, and I/O interfaces can be hot-plugged into Sun Fire V880, 3800, 4800, 4810, 6800, 12K, and 15K as well as Sun Enterprise 10000 servers and dynamically configured into or out of the operating environment’s domain, enabling such critical components as CPUs to be replaced with no downtime.

**Solaris Live Upgrade**

Upgrading operating environment software for critical servers must be done as carefully as possible, with avenues for backing out changes if errors are found or if there are inconsistencies between operating environment or application software versions. Solaris Live Upgrade software significantly reduces the downtime associated with an operating environment upgrade or maintenance. It enables the existing version of the operating environment to continue to run while a system administrator upgrades the system to a newer version of the operating environment, applies patches, or does routine maintenance on an inactive or duplicate boot environment. When the process is finished, administrators can simply reboot the system to run the updated operating environment. Solaris Live Upgrade also supports fallback in case of upgrade failure by rebooting back to the original version of the operating environment.
Configurable Dump Device
In the event of an operating environment failure, it is important for Sun service technicians to be able to analyze crash dumps that the kernel provides as it shuts down. In order to facilitate this process, the device to which a dump is saved can be configured on a server-by-server basis, and a repository of dump images can be maintained. In the Solaris 9 Operating Environment, additional information is included with system dumps, including saving memory pages from the application running on the CPU that triggered the dump. This facilitates fault diagnosis by providing a comprehensive picture of the system environment at the time the error occurred.

Device Hot-Plug Management
The Solaris Operating Environment supports hot-pluggable disks, CPUs, memory, and I/O boards that can be connected to a variety of high-capacity storage devices and cluster interfaces. Sun understands that when data centers dynamically reconfigure servers, they must have the ability to hot unplug as well as hot plug their devices — another feature of the Solaris environment that supports Dynamic Reconfiguration.

Observable
Observability comes in many forms in the Solaris Operating Environment, including its syslog logging facilities, on-the-fly kernel analysis tools, and a modular debugger framework that helps to rapidly identify errors.

Syslog
The Solaris Operating Environment’s syslog facility enables both remote and local logging of errors, as well as fine-grained control over how each level of log entry is reported to users and system administrators. Syslog has become a de facto industry standard, supported by all UNIX operating systems and many other devices such as routers and switches.

Kernel Analysis Tools
The Solaris Operating Environment provides a comprehensive set of on-the-fly kernel analysis tools, enabling administrators, developers, and Sun support staff to observe the state of the kernel in real time. Tools like kstat, lockstat, prstat, and cpustat give statistics and parameters for the kernel, file locks, process status, and CPU utilization, respectively. Tools like truss can be used to “wrap” an application so that all of its system calls can be externally observed.

Modular Debugger Framework
Sun’s Modular Debugger (MDB) is an extensible utility for low-level debugging and editing of the live operating system, crash dumps, user processes, core files, and object files. The framework integrates a set of debugger modules, each of which is written to have specific knowledge of a kernel component or device. This makes it easier to have symbolic access to kernel information with component-specific commands. The result is a framework that speeds problem resolution through effective management of information. In the Solaris 9 Operating Environment, MDB provides new symbolic debugging support for the kernel, new kernel debugger commands, new features for examining and controlling running user processes, and the ability to examine raw disk files and devices.
Compatible
Compatibility is one of the hallmarks of software using the Solaris Operating Environment, and is a key feature enabling customers to move up the product line without ever having to port or recompile their applications.

Application Binary Interface
Software is one of the most costly investments that an IT organization makes. Operating environments that force frequent application software upgrades due to operating system interface changes can result in skyrocketing costs. Sun supports a public application binary interface (ABI) that is designed to enable properly written applications to run on all Sun systems without modification — from desktop workstations to the 106-processor Sun Fire 15K server. This implicit contract with developers helps ensure binary compatibility throughout the product line while enabling Sun to make improvements in the Solaris Operating Environment so long as the improvements conform to the ABI.

Sun also offers the Solaris Application Guarantee, which enables developers to test their application for ABI conformance. Applications passing the Sun Guarantee Test Suite are then guaranteed to run on supported forward versions of the Solaris Operating Environment.

AppCert Program
A key component of the Guarantee Test Suite is the appcert program, which developers can use to help ensure that the applications they are writing will run unmodified on future releases of the Solaris Operating Environment. It evaluates how an application uses Solaris APIs and warns developers about the use of potentially unstable interfaces as well as the use of static linking to system libraries, either of which may cause compatibility problems with future Solaris Operating Environment releases. Applications that pass appcert’s evaluation are highly likely to run on forward releases without modification.

Protected
Protection against possible failures is just as important as recovering from errors that have occurred. The Solaris Operating Environment protects against failures with a high level of software component quality, rapid response to security advisories, and role-based access control, which limits the scope of administrator activities.

Component Quality
Every component in the Solaris Operating Environment is tested for memory leaks. Leveraging the kernel modular debugging framework, Sun’s memory-leak checking is performed proactively throughout the development software, helping to uphold the standard that any component, once checked in, is of production quality.

Security Advisories
Sun provides e-mail security advisories and operating environment patches, helping all customers to stay up-to-date with security advisories and keep their systems safe from hackers. Sun works with the Computer Emergency Response Team (CERT) to identify problems and quickly prepare patches to rectify them.
Role-Based Access Control

One of the most serious threats to system integrity is not from physical disk or server failures, but from human intervention and administrator error. A study by Oracle Research finds that while 25 percent of server downtime is due to physical failures, 34 percent is from human error. One way to minimize human error is to limit the scope of changes that different administrators can make. Using role-based access control (RBAC) in the Solaris 9 Operating Environment, the UNIX superuser can delegate various levels of authority to different classes of administrators through an easy-to-use GUI. This role-based control can, for example, give database administrators full control over the database environment while preventing them from making changes to the operating environment.

Enhancing Availability

Several features of the Solaris Operating Environment help deliver higher levels of application service delivery.

Availability With Sun Cluster Software

Many organizations require applications that are always available, and Sun Cluster software provides a highly available computing environment for running general-purpose applications as well as those specifically tailored for the high-availability environment. Sun Cluster software enables two or more servers to work together as a single entity to provide high-availability applications, system resources, and data to users. The Global File Service enables all elements of the cluster to share the same exact file system image, making it transparent to applications which physical servers they are executing on. Sun Cluster software integrates servers using redundant high-speed interconnects, redundant networking, and redundant connections to mirrored disk systems so that no single failure can interfere with application availability (Figure 4-5).

Availability With Sun Management Center

Sun Management Center software provides an extensible environment for managing Sun hardware platforms, storage components, and the Solaris Operating Environment. Sun Management Center can optimize performance, enhance application availability, and simplify management of IT environments, all while easily integrating with leading third-party enterprise management platforms including HP OpenView and BMC Patrol products. Sun Management Center increases availability by establishing an agent on each managed server, enabling it to evaluate both operating environment parameters (such as CPU utilization) and hardware parameters (such as power supply temperatures). This allows the tool to do predictive fault analysis, monitor performance, and isolate hardware and software faults — all increasing server availability.
Availability With Solaris 9 Resource Manager

A server that is operating normally, but whose applications are resource-starved, is as good as down. Solaris 9 Resource Manager software enhances application availability by helping to ensure that available resources — including network bandwidth, CPU utilization, and memory consumption — are allocated so that one application cannot impair another’s ability to provide services. This helps availability through fair allocation of resources, and also reduces costs by enabling organizations to consolidate many applications onto large, cost-effective servers from Sun.
Data Management

Data integrity is one of the foremost concerns of administrators, and the Solaris 9 Operating Environment now includes features to help manage the very data that gives a company its competitive edge.

Data Management through Solaris Volume Manager
The first step in maintaining data integrity is to use reliable volume management software to concatenate multiple disk volumes — which themselves might be large disk arrays — into single, large, logical volumes. This capability, as well as software RAID, is integrated into the Solaris 9 Operating Environment as Solaris Volume Manager. The task of managing disk volumes is further simplified in the Solaris 9 Operating Environment with a new graphical user interface (GUI).

Soft Disk Partitions
Using built-in Solaris Volume Manager software, up to 8,192 “soft” disk partitions can be created per disk volume, enabling each user in a shared environment to have a personal, private virtual disk. In addition to providing privacy, soft partitions create a firewall between users and make it easy to establish different backup policies for each partition. Soft disk partitions can be expanded as long as free disk space exists in the underlying partitions.
Snapshot Backups
The Solaris 9 Operating Environment eliminates the need to take file systems offline for backups. Through UNIX file system (UFS) snapshots, a temporary copy of a file system can be created so that a consistent point-in-time image of the file system can be backed up to tape.

MKFS Performance
Growth in database tables means growth in disk capacity, and for database management systems storing their tables in UNIX file systems, creating a file system structure on a large disk array can be a time-consuming operation. With improvements to the Solaris 9 mkfs command, a UNIX file system can now be created 96 times faster than before.

Resource Management
A wide range of resource controls are provided by the Solaris Operating Environment. These facilities are controlled by user-level commands and APIs, including Solaris 9 Resource Manager and Sun Management Center.

- A fair-share scheduler enables the creation of scheduling classes for groups of processes, one of several scheduler choices available.
- Processor sets can partition a system's processors so that they can be used only by a process or group of processes. In the Solaris 9 Operating Environment, greater administrative control is provided, enabling naming and automatic creation and deletion of processor sets. In addition, processor sets remain defined across reboots.
- Resource controls are available through Solaris 9 Resource Manager software. Solaris 9 Resource Manager can regulate the number of threads in a process, with hard and soft limits that can be used to contain the impact of denial-of-service attacks. Signals can be sent to administrative entities when limits are exceeded so that immediate action can be taken in the event that a hard limit is exceeded, and so that service providers can take proactive measures when soft limits are exceeded.
- Solaris 9 Resource Manager bandwidth management features facilitate the support of multiple services on a server, all of which compete for system resources. Through tasks, projects, and accounting, the number of processes in a task can be regulated, directly supporting how many processes a particular service can utilize. Finally, a wide range of specific resources can be controlled, including CPU time used, memory allocated, and file descriptors opened.
- Tasks, Projects, and Accounting supports management and accounting based on tasks and projects, enabling greater control over quality of service and enabling chargebacks based on actual resource utilization. Projects are defined to be a collection of tasks, which may in turn be made up of multiple processes. A user login, for example, is a task associated with a particular project, and the activation of the project can invoke resource controls such as the creation of a processor set in which to contain the project. This level of resource management and accounting is critical to both IT organizations (supporting large numbers of cost centers) and service providers (hosting applications for many customers). Using this mechanism, application service providers (ASPs) are now empowered to make customer chargebacks based on resource consumption.
• Dynamic System Domains provide high-level control over allocation of processors, memory, and I/O devices to separate electrically isolated domains on supported servers. The Solaris Operating Environment supports Dynamic System Domains, where system resources can be reallocated on the fly or on a scheduled basis to rapidly adapt to changing workloads.

Support for Service Containers
The fine-grained resource controls provided by the Solaris Operating Environment support the creation of service containers, an abstraction that enables managing performance, availability, and security of services in a Web-based architecture. Services can be managed within the context of service containers, and then the containers can be deployed onto the appropriate instance of the Solaris Operating Environment whose underlying hardware bests supports the quality-of-service characteristics needed by the service.

By providing the fine-grained resource controls necessary to support multiple service containers on the same server, the Solaris Operating Environment offers an unprecedented opportunity for organizations to consolidate workloads onto larger servers. This makes more cost-effective use of resources and reduces cost of ownership by reducing the total number of servers to support (Figure 5-6). This technique can be used everywhere from running multiple instances of Oracle database management system software on a single server, regulating resource consumption to support specific quality-of-service levels, to using the same multitier server infrastructure to support multiple Services on Demand software stacks. Resource controls can be used to establish walls between customers in every context — from running timesharing environments to services provided by ASPs.

System Management
Organizations serving large numbers of customers — either across the enterprise or over the Internet — are faced with managing a large number of systems. Sun helps its customers reduce management costs through two avenues.
• Administrator efficiency is increased with automated installation and deployment tools that reduce the time required to execute routine tasks.
• Efficiency is further increased by providing server monitoring and management tools that can automatically respond to a wide range of faults, reducing the number of events to which administrators must respond.
Installation and Provisioning
Many organizations have large numbers of servers all delivering the same services, however, each requires its own similar yet unique software configuration. For example, one Sun customer provides Internet search services using 250 Sun Enterprise 4500 servers, each one configured with up to 14 processors and loaded with similar software so that the entire server complex can support millions of customers.

Managing this number of servers can be costly in terms of administrator time without automated tools. New features in the Solaris 9 Operating Environment make it easier to manage large numbers of servers in this type of server farm.

Solaris Flash and Solaris Live Upgrade
Solaris Flash makes it easy to provision large numbers of servers having similar configurations. It enables administrators to create a single reference installation of the entire system software stack (including the Solaris Operating Environment and applications) and replicate the installation on a number of servers. Solaris Flash reduces installation time, configuration complexity, and simplifies the process of redeployment to support different service levels as customer demands change.

Solaris Live Upgrade is integrated to work with Solaris Flash technology. This functionality provides a mechanism to install the Solaris Flash archive on an inactive boot environment while the active boot environment is fully functional and unaffected by the installation. When the process is finished, a simple reboot will migrate the system to the updated environment. The system can be quickly rolled back to its original software configuration with another reboot in the event that the upgrade has errors or is otherwise unsatisfactory.

Sun Management Center Change Manager
Sun Management Center Change Manager software, available at an additional cost, helps enterprises manage configurations created with Solaris Flash. It provides a drag-and-drop graphical interface and version control over the operating environment, patches, and application software, and enables organizations to manage and provision entire software configurations on hundreds of servers as a single, integrated software stack.

Administration
The Solaris Operating Environment provides a wide range of administration tools that assist with both user administration and system administration tasks.

System Administration Facilities
The Solaris Operating Environment provides both command-line tools and graphical user interfaces for managing users, resources, and disk storage. Command-line tools can be invoked locally or remotely using a secure Telnet session. GUI-based tools can also be invoked locally or remotely.
**Patch Management**

Patch management features in the Solaris 9 Operating Environment are the most comprehensive ever offered for Solaris administrators — and are designed to help reduce costs through increased efficiency. With these patch management functions, administrators can easily analyze the patch state of a system and automatically download the recommended patches. The administrator is provided with the install order necessary to accommodate patch dependencies. Administrators can use the patch management tools on local and remote systems. All patches delivered through Solaris Patch Manager are digitally signed, helping ensure that the patches have come from Sun and have not been altered in transmission.

**Monitoring and Management**

Sun's comprehensive monitoring and management facilities provide efficient and proactive management of Sun servers, resulting in increased availability and lower operating costs due to reduced administrator costs. The immediate, automated fault correction provided by Sun management tools improves availability while reducing costs.

**Sun Management Center**

Sun Management Center software provides a powerful, easy-to-use single management point for all Sun servers and storage, independent of geographic location. System administrators can perform remote system configuration, performance monitoring, and isolate hardware and software faults through a single interface. Sun Management Center easily integrates with enterprise management frameworks, and provides a central facility for managing events and alarms, automated responses, and diagnostics.

- All events and alarms come to a single console, enabling a rapid response to problems as they occur. Sun Management Center is the perfect complement to enterprise management systems. Whereas products such as HP OpenView can give visibility into CPU utilization and process statistics, Sun Management Center has extremely fine-grained insight into the operation of Sun servers, and can provide a detailed analysis of any situation and escalate notification up to higher-level enterprise management systems.
- Sun Management Center software can automatically respond to a variety of situations using built-in event/action automation. It can also be augmented with site-specific scripts that provide rapid, automated responses to events such as hardware or software thresholds being exceeded.
- Sun Management Center can direct the execution of server diagnostics and take proactive steps to prevent problems. For example, it can notify an administrator when diagnostics indicate an imminent failure of a CPU or disk drive, or when power supply operational parameters go out of range. Identifying problems before they result in a server shutdown helps increase uptime.

**Standards-Based Management**

For organizations managing large numbers of servers in a heterogeneous environment, standards-based management tools are a must. The Solaris Operating Environment supports management through SNMP as well as through Web-based enterprise management (WBEM). Standards-based management gives administrators the choice of using Sun management products or those available from competitors.
User Management

For user management, Sun ONE Directory Server is integrated with the Solaris 9 Operating Environment, creating a Lightweight Directory Access Protocol (LDAP)-based environment that can scale to millions of users. The Sun ONE Directory Server provides comprehensive information asset management, including user names and attributes (such as passwords) as well as group, project, and task definitions. LDAP can unify a complex, fragmented, and heterogeneous infrastructure by managing all of its network resources from a central, interoperable server. For those developing Web-based services, the Sun ONE Directory Server central repository provides the basis for storing user preferences, context, and attributes necessary for delivering Services on Demand.

For those using the Sun ONE Directory Server for user management in the Solaris Operating Environment, Sun has made it easy to migrate to LDAP by providing conversion and synchronization tools to migrate from NIS+ to LDAP-based management. The secure LDAP client has been enhanced to support various encryption mechanisms, including DIGEST-MD5 and SSL, enabling secure password management through the Sun ONE Directory Server.
Chapter 6

Designed for Security

From the time users log into a server and access files and applications on the system, to when they venture out onto the Internet, they can be assured of one of the industry’s strongest security solutions when they use the Solaris Operating Environment. Because Sun designed the Solaris Operating Environment with the belief that all systems should be networked, it has been built to securely interconnect with other systems and also to be more secure from the viruses and worms that so often plague software designed without security and networking in mind. These aspects of security — network and platform security — are critical whether building an enterprise network or providing services to millions of users over the Internet.

Network Security
The network security provided by the Solaris Operating Environment helps ensure secure authentication over the network, secure remote access, secure network connections that can’t be snooped, and protection from network-based attacks.

Integrated Kerberos V5
Kerberos is a standards-based software solution for providing a secure, network single sign-on environment useful for both operating environment and application authentication. Once authenticated, network traffic can be encrypted to help ensure privacy of user data. In a Kerberos environment, users log on only once, and subsequent accesses to servers and applications are authorized by tickets granted by the Kerberos Key Distribution Center (KDC).
Because it is standards-based, Sun’s Kerberos implementation interoperares with MIT Kerberos and Microsoft’s Active Directory, simplifying deployment across heterogeneous environments. Sun has additionally improved interoperability for changing passwords from its KDC implementation to those from MIT and Microsoft.

Beginning with the Solaris 9 Operating Environment, the Kerberos V5 Key Distribution Center can be configured at install time and subsequently administered through a convenient graphical user interface. Kerberos-ready client software (including telnet, rlogin, rsh, and rcp) is available as a free Web download.

**Kerberos-Secured NFS**
Privacy for NFS network connections is enhanced with Kerberos extensions that authenticate NFS access on a per-user, rather than a per-client, basis. This enables only the authorized user, not an entire server, to access their files over NFS, and supports user-specific encryption of NFS traffic.

**Solaris Secure Shell**
For secure access from outside a protected network, Solaris Secure Shell software provides strong authentication and encryption for remote access using the SSH v1 and SSH v2 protocols. Solaris Secure Shell is integrated and supported in the Solaris Operating Environment, simplifying installation and enhancing standard utilities such as login and accounting.

In addition, the Solaris 9 Operating Environment integrates a large array of precompiled open-source utilities, providing safe, secure replacements for Telnet, FTP, rlogin, rsh, and rcp, as well as enabling tunneling of X Window System and other network protocols.

**IPSec/IKE for IPv4 and IPSec for IPv6**
The Solaris Operating Environment now includes the Internet Protocol Security architecture (IPSec) with Internet Key Exchange (IKE), which allows digital certificates to be dynamically created and used as an authentication mechanism for authenticated or encrypted traffic. IPSec is a highly effective tool for securing network traffic because it does not require applications to be modified in order to use it. Because of its transparency, existing customers may move to IPSec in the network without incurring additional costs of modifying applications. Using IPSec, security policies can restrict inbound and outbound communication as well as public access on a port-by-port basis. Similarly, authenticated hosts can be offered different levels of access depending on security policy.

The IPSec Authentication Header (AH) can be required to ensure the authenticity of a remote server or workstation. Network communication can be encrypted using the Encapsulated Security Payload (ESP). Encryption using 128-bit keys is provided with the Solaris Operating Environment, with even stronger algorithms available through Web downloads. These two mechanisms facilitate a wide range of security options, and since IPSec is implemented at the IP level, it can be deployed so that it is transparent to applications, forming the basis of a true, end-to-end security solution. In fact, since IPSec is implemented in the lowest levels of the protocol stacks, only trusted, authenticated parties are able to reach the applications servicing IPSec-restricted ports.

The IPSec implementation provided with the Solaris Operating Environment follows Internet Engineering Task Force (IETF) specifications, and as a result will interoperate with all other vendors supporting the standard. Various levels of encryption are available depending on export restrictions.
**SunScreen Security Technology**

SunScreen 3.2 software is included with the Solaris 9 Operating Environment at no extra charge. This high-speed, stateful packet-filtering firewall offers advanced features that protect a single system or an entire network of servers. Based on Sun's many years of SunScreen product development, SunScreen 3.2 software offers privacy features such as network address translation (NAT), IPsec/IKE and SunScreen SKIP virtual private network (VPN) client support, and secure, remote administration through any Java technology-enabled Web browser.

Proxies with virus scanning and user authentication are available for FTP, HTTP, SMTP, and Telnet protocols. In addition to traditional routing interfaces, SunScreen 3.2 software can filter packets in stealth mode without any visible IP address on the network, making it very difficult to attack the firewall directly. Sun is providing this sophisticated firewall product to enable customers to protect their machines without needing to purchase expensive add-on security products.

**Platform Security**

The platform security features of the Solaris 9 Operating Environment support server and cluster hardening through administration, system-level auditing, pluggable authentication modules that support smart cards, role-based access control, and increased security through nonexecutable program stacks.

**Modular Software Packaging**

Where the competition adds features without balancing the need for security, Sun knows how to remove software to simplify and secure the operating environment. The Solaris Operating Environment’s fine-grained software packaging allows systems to be installed with the minimum number of software components necessary to do the job, enhancing security because it is impossible to mount security attacks on software that is not installed. Also, the Solaris Security Toolkit — included with the Solaris 9 Operating Environment — supports server minimization by removing packages not necessary for a server’s specific function and changing default configurations to further tighten security.

**Extended Auditing**

A key security principle is accountability, and the Solaris 9 Operating Environment provides the ability to trace actions taken by a user that could have malicious intent and thus compromise a system’s security. The SunSHIELD™ Basic Security Module (BSM) provides a mechanism for fine-grained logging and auditing of user activities. It creates an audit trail that can be used for cause/effect analysis, user activity history, suspicious activity detection, and forensic analysis of a system that has come under attack.

**Pluggable Authentication Modules (PAM)**

The Solaris Operating Environment provides a framework for pluggable authentication modules (PAM), an innovation from Sun that can be used to integrate a variety of authentication mechanisms with third-party applications and the operating environment itself. Beginning with the Solaris 9 Operating Environment, smart cards and iButton reader devices are supported, and are integrated with the Common Desktop Environment (CDE) login itself.
Using a set of public APIs, developers can easily construct and configure PAM modules to support site- and application-specific authentication requirements. For example, a PAM module can be developed to interface with biometric scanning devices such as palm geometry scanners. Simply by modifying a configuration file, all facilities that utilize PAM for authentication subsequently require a palm scan at user login time. For even higher security, PAM configurations can require combinations of authentication mechanisms, such as smart card authentication and successful palm scan.

PAM has been internally restructured in the Solaris 9 Operating environment to make it even easier for developers to write new modules. For example, PAM modules can be written to handle only authentication without also integrating with name services. Separation from the underlying name service means that developers can concentrate on the task at hand rather than concerning themselves with the naming service (e.g., files, NIS, NIS+, or LDAP) in use.

Secure LDAP Client
In the Solaris 9 Operating Environment, the LDAP naming services client has been enhanced with Secure Sockets Layer (SSL) support for encryption between the client and LDAP server. Solaris software uses LDAP as a first-class directory service, just as NIS or NIS+ have historically been used as a first-class naming service. Future enhancements will include password management with centralized policy storage on the LDAP directory server.

Role-Based Access Control (RBAC)
Even with state-of-the-art securing of systems and networks using IPSec and strong user authentication using mechanisms such as smart cards and Kerberos, enterprise computing systems have a remaining vulnerability: human error. The best system administrators and most trusted users are still fallible, and the most frequent source of system failures is operator error.

Role-based access control (RBAC), provided with the Solaris Operating Environment, is an alternative to the all-or-nothing model of root access to UNIX systems. Role-based access control enables assigning rights to perform specific operations. This allows individual trusted users to assume a higher privilege for performing limited administration functions, and can be used to partition root privileges among a group of administrators. By separating powers and controlling the delegation of privileged operations to individual users, RBAC minimizes the chance that any user will go beyond their realm of expertise and inadvertently — or intentionally — make a change that results in a system failure. For even greater security, RBAC is integrated with the audit mechanism, enabling administrators to trace back to the original user that assumed a role to execute a malicious action.

Role-based access control is implemented using roles and attributes; once a user is logged in through a preconfigured PAM module, they can assume roles that have been made available to them through the use of role-aware shells. By ensuring that users are first authenticated before any role is assumed, all privileged activities can be logged and associated with a person.

Nonexecutable Program Stacks
Because Sun develops the entire hardware/software stack, including processors, systems, and operating environment, it has the flexibility to create solutions that require more than just software to implement. The Solaris Operating Environment can prevent execution of user program stacks, reducing the possibility that a programming error can cause malicious code to be executed on the stack through a buffer overflow. This feature is available through a linker option that protects all 64-bit programs by default and can be manually specified for 32-bit programs.
To enhance security even more, Sun has built all core operating system and networking components using nonexecutable stacks, and has made additional effort to eliminate other programming errors that can result in buffer-overflow vulnerabilities. To further protect against buffer overflow failures in third-party software, a simple kernel configuration parameter can be used to help prevent stack-based buffer overflows from executing code in any process on the entire system.

*Kernel Pseudo Random Number Generator*

Secure systems often depend on high-quality random number generation. For example, encryption keys of all kinds depend on random numbers to reduce the possibility of compromised keys by understanding the random number generation process. IP security depends on random selection of TCP/IP sequence numbers to prevent session hijacking.

The Solaris 9 Operating Environment creates a kernel-based entropy “pool” that is used in random number generation provided by the `/dev/random` device. The entropy pool is populated with nonrepeatable physical measurements — such as disk latencies temperatures — as well as rapidly changing parameters that only the kernel can observe. Because the sources of entropy and the state of the kernel entropy pool are hidden from outside observation, the kernel and applications can generate strong cryptographic keys and use random numbers that prevent even sophisticated attacks.
With many organizations providing Web-based services both inside and outside of their company boundaries, even more challenging demands are placed on the operating environment. Rather than serving a single user on a personal computer, or a limited collection of users on a workgroup server, organizations now must be prepared to serve millions of users over the Internet. Companies and service providers are no longer in control of how, when, and from what devices their customers access their applications, increasing the importance of managing service levels through scalability, availability, manageability, and security.

Supporting Services on Demand requires more than just these essential characteristics — it requires innovations like Java technology, and bulletproof applications like the Apache and Sun ONE Web Server products.

Java 2 Platform, Enterprise Edition
Java 2 Platform, Enterprise Edition (J2EE) software leverages features of the Java 2 Platform, Standard Edition, including Write Once, Run Anywhere™ portability, the JDBC specification for database access, CORBA technology to facilitate integration with existing legacy resources, and a security model strong enough for Services on Demand applications. Adding to this base, J2EE technology also provides support for Enterprise JavaBeans™ components, Java Servlet, JavaServer Pages, and XML technology.
The J2EE platform supports portability and scalability for long-term application viability because it is supported by all of the major application server platforms, including those from BEA, ATB, Oracle, and Sybase. And for the difficult chore of integrating legacy systems and software into a Services on Demand environment, J2EE software wraps and embraces these systems into a unified, component-based application model.

Further extending the capabilities of Java technologies and XML, the Java XML Pack (JAX) enables Java technology developers to build Services on Demand that incorporate XML RPC, Simple Object Access Protocol (SOAP), as well as XML messaging, processing, and registries.

The Foundation for Sun ONE Software
The Solaris Operating Environment supports more than 12,000 commercial applications and even more custom applications, such as those developed to provide Services on Demand through the Sun Open Net Environment (Sun ONE) architecture. Some key applications are — or will be — integrated directly into the Solaris Operating Environment, including:

- **Sun ONE Application Server, Platform Edition**, providing a robust environment for running business logic functions encapsulated in Enterprise JavaBeans components
- **Sun ONE Directory Server**, for user authentication and storing customer-related information

Additional applications are provided with the Solaris Operating Environment media kit. These applications are being provided with full developer licenses, with additional licensing costs necessary only when applications are deployed:

- **Sun ONE Integration Server (formerly iPlanet Integration Server)**, enabling legacy custom packages and Java technology applications to be made available through Web services as well as business process management
- **Sun ONE Web Server** to support Web presentation layers
- **Sun ONE Portal Server**, securely delivering services to both IP and wireless-enabled devices
- **Sun ONE Message Queue (formerly iPlanet Message Queue for Java)**, providing port-to-port communication using Java Message Services (JMS)

In addition, an evaluation license is provided for:

- **Oracle Database Management System**, providing a world-class repository for storing data for all aspects of a company’s operation, from product catalogs and pricing to Web site user preferences

Finally, a wide range of open-source software components are tested, integrated, and fully supported within the Solaris Operating Environment, including:

- **Apache Web server**, providing an open-source tool for Web hosting
- **Perl**, a powerful programming environment optimized for rapid creation of utilities and applications
- **Samba**, widely recognized as a vital tool for interoperability with PC networking protocols
Conclusion

Sun’s continuous improvement of the scalable, available, manageable, and secure Solaris Operating Environment has earned it first place on D.H. Brown’s list of UNIX operating systems, and has won it Network Computing’s “Best Server Operating System” award for 2001. While Sun maintains its single focus on bringing the full capabilities of its hardware to the user, both Microsoft and IBM struggle with flawed business models that keep them behind. Microsoft must develop to the least-common-denominator platform, preventing it from making the integrated improvements for which Sun has long been known. With an abysmal record for security, Microsoft software is simply not ready for prime time. Among Sun’s UNIX operating system competitors, IBM, with its support of no fewer than 14 operating environments, must attempt to support different hardware architectures and operating systems for each tier in a service point architecture, spreading staff and resources thin.

It costs less to use one operating environment for all aspects of a company’s operation, so why not use the best? The reliability, availability, and serviceability of the Solaris Operating Environment puts it at the top of UNIX operating system vendors, lowering costs and making Web-based applications truly available anywhere, anytime, and from any IP-enabled device.
The scalability provided by the Solaris Operating Environment and the most flexible line of binary-compatible servers anywhere gives organizations flexibility to start small and grow with the speed of their business — without ever having even to recompile their applications. Manageability is key for organizations focused on containing costs, and the management features of the Solaris Operating Environment support local user and resource configuration as well as large server farms with hundreds of servers all performing similar functions. Finally, Sun excels due to its superior management capabilities, its ability to minimize the operating environment footprint, and its formidable security through time-tested security technologies.

For organizations having to cut costs and focus on the fundamentals, using the Solaris Operating Environment means server consolidation and fewer, more cost-effective servers, more efficient use of administration staff, and the most scalability, availability, manageability, and security available. For organizations taking and maintaining a competitive edge by delivering Services on Demand, the Solaris Operating Environment is ahead of the competition with services companies need today. And with through its constant drive for quality and excellence, Sun will continue lead in the future, and deliver the best-of-breed solutions customers have come to expect.