The Solaris™ 10 Operating System—
Breakthrough Technology for the Telecom Market

A Technical White Paper
January 2005

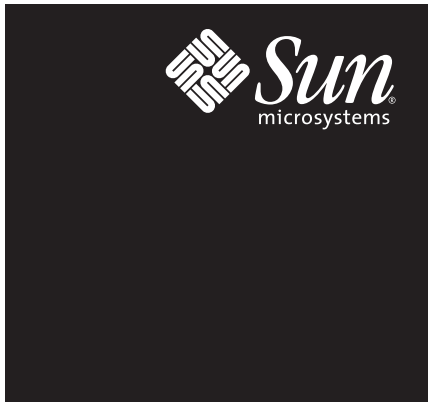

Table of Contents

Chapter 1

Executive Summary

Telecom companies today face an unusually difficult situation to remain competitive. Overall, there is pressure to reduce costs and expand services. Some are facing flat revenues and subscriber growth in their core markets. There is relentless pricing pressure, while increasing competition and uncertain government regulations continue to create a difficult environment for carriers. This is driving carriers, network equipment providers, and service providers to reduce costs. At the same time, converging networks—legacy circuit switch with next-generation packet switch—are forcing new investment and creating opportunity. This is expected to create an infrastructure for services that can generate new, more profitable services to both business and residential customers.

Sun has a long and successful history with the telecom industry, helping companies move away from customized systems to one based on a commercially available operating system—the Solaris™ Operating System (OS). The Solaris OS continues to provide many compelling capabilities to the telecom industry, offering numerous improvements in performance, availability, and security, increasing service levels.

- **Performance**: High performance means maximizing the number of users per system, decreasing response times for Web or data access which results in improved customer satisfaction.
  - New tools, such as the Solaris Dynamic Tracing (DTrace), help developers and administrators uncover and reduce performance bottlenecks.
  - A streamlined networking stack that increases the throughput of virtually all applications.

- **Availability**: Telecom companies must meet or exceed the industry benchmark of "five nines" availability and incorporate multiple safeguards to ensure uptime. The Solaris 10 OS is the culmination of years of technology innovation that help carriers and service providers keep their systems up and running.
  - Predictive self-healing features are designed to automatically identify and recover from system faults as they occur.
  - Solaris Containers allocate additional resources to applications as needed to meet peak demand, providing predictable service levels even when multiple applications are on the same system.

- **Security**: With the advent of IP-based infrastructure such as soft switches, media gateways, and Session Initiation Protocol (SIP) servers, the network has become increasingly vulnerable to application failures and Denial of Service attacks. The security enhancements in Solaris 10 can help speed development and deployment of commerce-related services while increasing the load factor in security-related servers. Specifically, Solaris 10 delivers the following security enhancements that can reduce the likelihood of network downtime:
  - Buffer overflow protection at both the server and application level.
  - Process rights all but eliminate root requirements for applications and users; instead, users and processes are granted the rights needed to accomplish their tasks.
  - Solaris Containers provide a private execution environment for each application or service on a system.
  - A new Cryptographic Framework speeds up security-related tasks. In addition, a common environment for both users and data improves the performance and uptime of key services such as virtual private network (VPN) and secure sockets layer (SSL).

The Solaris 10 OS helps service providers and carriers reduce costs by improving:

- **Administration and Management**: The Solaris 10 OS helps telecom companies manage services, not processes. In addition to predictive self-healing capabilities, new management features can help reduce system administrative costs. For example, Solaris Flash Archives quickly and easily install new OS and application software on servers over the WAN or LAN, reducing the time and expertise required to redeploy servers. In effect, this creates a field replaceable unit (FRU) for the entire software environment of a given server type.
- **Compatibility**: The Solaris OS is a UNIX® operating system, and has long supported interfaces from many worldwide standards organizations. Open, consistent, and standardized interfaces mean that telecom companies have investment protection, interoperability with a wide variety of middleware, and support for next-generation services such as Web conferencing, unified messaging, and other presence-based services.
  - Solaris 10 provides in-kernel support for emerging standards such as Stream Control Transmission Protocol (SCTP) and SIP.
  - The Solaris Application Binary Interface (ABI) guarantees compatibility across previous versions of the Solaris OS, ensuring a smooth migration of existing software assets as well as leveraging the existing investment to speed deployment and reduce migration costs.

- **Platform Choice**: Network equipment providers (NEPs) and carriers now have a real choice when delivering Solaris 10 applications and services because Solaris 10 supports SPARC®, Intel x86, and AMD Opteron hardware platforms.
- **Maintenance and Support**: Solaris 10 continues Sun's long history of stability and support for carriers, NEPs, and service providers. A worldwide presence—more than 170 countries—means that Sun offers support when and where you need it. A professional service organization is ready to deliver end-to-end architectures and solutions.

More than a collection of innovative features, Solaris 10 offers capabilities that combine to improve telecom network operations environments. The features are not freestanding or isolated, but work together. For example, Solaris Containers automatically allocate system resources. Predictive self-healing can be configured to shut down a server or application instance and move sessions to another Solaris Container. Self-healing capabilities can be dynamically modified through open interfaces—starting at the CPU, to the memory, and working up through the file system, applications, and services. The Solaris Cryptographic Framework adds transparent, application-level protection while automatically load balancing encryption hardware, such as SSL accelerators. The Solaris 10 IPv6 capabilities support the latest Java™ SDK 1.5, which is also IPv6 aware. All this means that Solaris 10 helps to provide end-to-end improvements in performance, reliability, availability, and manageability for all telecom providers, resulting in increased service levels and reduced costs while delivering a strong foundation for new markets and services.

Chapter 2

Improving Service Levels

The features, performance, and capabilities of the Solaris 10 OS extend Sun's commitment to the telecom market. Solaris 10 offers new capabilities that are designed to increase performance and raise service levels by:

- **Improving Performance**: DTrace enables developers and administrators to maximize end-to-end performance of applications and services with pinpoint accuracy. Bottlenecks in performance often mean poor service and lost revenue. DTrace provides a means to find a problem quickly and get the network up and running at top speed. Performance improvements have been shown to range as high as 30 times faster.
- **Increasing Availability and Reliability**: A new architecture for building and deploying telecom systems and services that are capable of predictive self-healing. Self-healing technology enables Solaris systems and services to reduce application and system downtime, significantly increasing availability.
- **Enhancing Security**: A stronger security framework that incorporates least privileges and process rights capabilities. Those processes that previously required superuser capabilities now require process rights. Process rights management uses privileges to restrict processes to only those rights that are required to perform the task.

New Performance Levels Increase Load Factor

Rolling out new services, such as unified messaging, is challenging in the telecom environment due to the sheer scale involved. Even a subset of the subscriber base is a large number, demanding significant investment in development and hardware. Carriers and service providers must maximize capital asset productivity and return on investment (ROI). High performance—fast

processing and high throughput—means getting the most out of every piece of hardware and software.

Solaris 10 delivers new levels of performance—faster processing and higher throughput.

DTrace Uncovers Bottlenecks

A telecom provider rolls out a new operations server, but begins to notice occasional delays in service. The delays don't show up in the test bed, and are too elusive to be picked up by traditional diagnostic tools. Solaris 10 is brought in, and DTrace is used to discover a configuration error that is causing a process to periodically tie up the system with disk I/O. The configuration problem is rectified and the delays disappear—the problem is solved in minutes or hours, instead of weeks or months.

Applications—especially real-time applications used in the telecom market—have become more complex. Real-time debuggers are difficult and impractical in a production environment. Understanding what an application is doing during both development and deployment—how it flows through the system, what resources it is consuming, what calls it is waiting on—has been nearly impossible until now.

DTrace is a comprehensive dynamic tracing facility that gives telecom users, administrators, and developers a new level of observability into the kernel and user processes. DTrace provides performance and operating data on everything that goes on in Solaris hardware servers—including the SPARC, Intel x86, and AMD Opteron platforms— and enables developers and administrators to pinpoint exactly where the bottlenecks are by providing comprehensive, real-time observability into the kernel and user processes such as system and application calls. DTrace enables administrators and developers to:

- Examine the behavior of user programs and the Solaris OS to identify the root causes of system and application bottlenecks.
- Highlight trends and patterns to tune the system for best performance.
- Track down performance problems across many layers of software.
- Locate the cause of aberrant behavior.

DTrace helps you understand your system by permitting dynamic instrumentation of the operating system kernel and user processes to record data specified at locations of interest, called probes. Probes are like small programmable sensors—there are more than 30,000 probes included in the Solaris 10 OS. A probe is a location or activity to which DTrace can bind a request to perform a set of actions, like recording a stack trace, timestamp, or argument to a function. Each probe can be associated with custom programs that are written in the new D programming language. This association enables access to system data by using ANSI C types and expressions that can easily capture stack traces, record timestamps, build histograms, and more.

All of DTrace's instrumentation is entirely dynamic and available for use on a production system. The performance impact is variable. When DTrace is off, there is no performance cost. If only a few probes are enabled, the impact is relatively small, while enabling all 30,000 probes creates a noticeable impact—though the system is still functional and operational. The performance impact of tracing is limited to only those probes and actions that are enabled. In addition, DTrace is safe—it cannot damage the production system because DTrace has security, complete safety, and error checking at the core of its design. These features enable developers and

administrators to use DTrace with confidence on production systems in real time whenever problems need investigation.

Across the Board Improvement with New Network Stack Architecture

As IP becomes dominant in the telecom environment, raising the performance of IP network interfaces becomes critical. The streamlined networking stack of the Solaris 10 OS delivers performance improvement for Web and other TCP services, as well as core interfaces such as 10-Gbps Ethernet. Higher network performance improves customer satisfaction and enables next-generation network services on a carrier-grade scale.

Solaris 10 implements a rearchitecture of the networking stack to improve both single CPU performance and scalability across a large number of CPUs, delivering extremely high performance while improving scalability. For telecom providers, this means improved performance for TCP services such as Web services at the front end, and at the network core with support for very high-speed interfaces.

This new architecture—code named FireEngine—supports 10-Gbps Ethernet networks, efficient resource utilization, and various types of offload technologies. It also speeds up TCP/IP connection setup and teardown, and is transparent to applications. Almost all applications that use networking can benefit.

The new networking stack represents the next evolution in multithreaded design. Instead of a large number of threads contending for resources across all CPUs, FireEngine imposes order over chaos by partitioning workloads across CPUs and within a CPU, letting a thread finish a reasonable amount of work before it relinquishes control. This innovation enables Solaris networking to vertically partition the workload by using an IP classifier-based, lockless design. This design reduces synchronization overhead and reduces cross-communication between CPUs.

The improvements are visible across all networking workloads and transparent to users. A rearchitected TCP/IP stack delivers from 30- to 50-percent improvement in network throughput, doing so with a 10- to 15-percent lower CPU load while still maintaining the reliability for which the Solaris OS is known.

- Internal testing shows that this stack delivers a 14- to 38-percent performance improvement[1] on Apache Web servers compared with other UNIX distributions.
- On the Java System Web Server, the Solaris 10 OS on the AMD Opteron processor-based systems improves performance by nearly 45 percent on the same hardware, as compared to the Solaris 9 OS.

High Availability and Reliability Increase Service Levels and Subscriber Satisfaction

Massive scalability can quickly swing demand from one end to the other. Hot stand-bys are not cost effective—carriers and service providers must efficiently and quickly adjust to rapid changes in demand. System resources must be available virtually all the time, with no planned or unplanned downtime from a customer perspective. The Solaris 10 OS offers innovative technology to meet these demands.

1. See “The Solaris Performance Advantage” at www.sun.com/2004-1012/feature/ for information and details.

Predictive Self-Healing

A Solaris 10 server detects a small, nonrepeatable RAM error. This event is logged. When the soft error rates exceed a threshold, Solaris 10 predictive self-healing features isolate the memory from active processing, and alert the system administrator to this potential problem. Subscribers and administrators see no interruption in service.

In deployments with large numbers of users and service-level agreements, system reliability and availability are essential. A small memory glitch can affect hundreds or thousands of subscribers. The Solaris 10 OS features predictive self-healing capabilities, significantly reducing downtime or unplanned outages. Self-healing technology enables Sun™ systems and services to maximize availability when software and hardware faults occur. Examples of self-healing activities include administrator messaging, isolation or deactivation of faulty components, and guided repair. Self-healing technology facilitates a simpler and more effective end-to-end experience for system administrators and service providers, thereby reducing costs.

Predictive self-healing systems include a simplified administration model. Traditional error messages are replaced by telemetry events that are consumed by software components. Error events and other data that can be gathered to facilitate self-healing are dispatched to software components called *diagnosis engines*, which are designed to pinpoint the underlying problems corresponding to these symptoms. Diagnosis engines run in the background, consuming telemetry until a diagnosis can be completed or a fault can be predicted. This produces another event, called a *fault event*, which is broadcast to any agents deployed on the system that know how to respond.

The Solaris 10 OS includes self-healing agents that can dynamically take processors, regions of physical memory, and I/O devices offline. Through these reconfiguration agents, a self-healing Solaris system can take proactive and immediate action to isolate and disable a faulty component and continue providing service—even before human administrators know there is a problem. Solaris reconfiguration agents are integrated with other Solaris features, such as Solaris Containers. They provide a consistent administrative experience and are transparent to application programmers.

A new software component, called Fault Manager, manages the telemetry, log files, and components. When appropriate, the Fault Manager sends a message to notify an administrator that a problem has been detected. The message directs administrators to a knowledge article on Sun's new message Web site, www.sun.com/msg/, which explains more about the problem's impact, appropriate responses, and repair actions.

Any number of additional agents may be deployed on a self-healing system to act in parallel with reconfiguration, including agents to provide local or remote messaging or other connections to higher level management software.

Once system components have been converted to properly handle errors and produce telemetric events, diagnosis software can be developed, improved, and deployed in parallel without inducing further system downtime by requiring operating system kernel patches.

Solaris Service Manager

The Solaris Service Manager software is a core element of the predictive self-healing implementation in the Solaris 10 OS. It turns application services into first-class objects that administrators can observe and manage in a uniform way, and implements the capability of automatically restarting and managing them.

The Solaris Service Manager software can restart services if they are accidentally terminated by an administrator, aborted as the result of a software programming error, or innocently affected by an underlying hardware problem.

Additionally, the Solaris Service Manager software simplifies and secures common administrative tasks, such as disabling services and changing properties. The Solaris Service Manager software also speeds system boot by starting services in parallel according to their dependencies. The undo feature helps safeguard against human errors by permitting easy rollback of changes.

Roll Out Service Level Agreements with IPQoS

As IP-based services gain acceptability, carriers and service providers need a way to maximize revenues based on service level agreements, or according to application types. For example, service providers may need to prioritize voice, data, and hosted (VPN) traffic according to service level agreements, and in some cases bill according to service metrics.

Carriers are often required to provide their customers with guaranteed levels of service, which are called Service Level Agreements (SLAs). IP Quality of Service (IPQoS) enables the Differentiated Services (DiffServ) architecture that is defined by the Differentiated Services Working Group of the Internet Engineering Task Force (IETF). DiffServ is enabled by the Type of Service (ToS) field, and is typically processed by routers. In addition, IPQoS utilizes Class of Service (CoS) 802.11d VLAN tags. IPQoS is implemented at the IP level of the TCP/IP protocol stack. This architecture offers a range of different classes of service to different customers.

By enabling IPQoS, service providers can provide various levels of network service for selected customers and applications. These differentiated services can be based on a structure of service levels that are offered to customers. You can also provide differentiated services that are based on the priorities set for applications or users on your network.

In addition, IPQoS can be used to prioritize among applications, so that critical ones get a higher quality of service than less-critical applications. For example, in a hosted service, a service provider could offer top priority to all users accessing a voice mail program and lower priority to those using FTP.

Using IPQoS, system administrators can establish varying levels of service based on a price structure. IPQoS incorporates an accounting module in the DiffServ model. You can use it to take statistics on traffic flow, and bill customers in agreement with their SLAs. Flow accounting is also useful for capacity planning and system monitoring.

IPQoS contains features that can help you make network performance more efficient as you implement quality of service. When computer networks expand, the need also increases for managing network traffic that is generated by increasing numbers of users and more powerful processors. Some symptoms of an overused network include lost data and traffic congestion that results in slow response times.

In the past, system administrators handled network traffic problems by adding more bandwidth. Often the level of traffic on the links varied widely. With IPQoS, you can manage traffic on the existing network and help assess where, and whether, expansion is necessary.

Strong Security Increases Availability and Reliability

A Web server comes under attack from multiple IP addresses. While the traffic approaches CPU saturation, specific HTTP requests are sent. This causes a buffer overflow, and in the fraction of a second the Web server is unprotected, a new shell account is created.

It is difficult to overestimate the value and benefits of security in a telecom environment. Public and private networks of all types are under constant attack, with hackers probing for any weaknesses while trying to determine the effectiveness of their latest tools. These attacks are becoming more sophisticated, and if successful, the losses are severe in both revenues and customer satisfaction.

Solaris 10 continues to enhance Sun's 20-year commitment to security with many new features and capabilities. These include:

- Key technologies that offer both system-wide and application-level protection against buffer overflow attacks launched from the network.
- Solaris Process Rights Management constrains processes to just the resources they need.
- Solaris User Rights Management delegates administration and assigns rights to users, preventing security violations.
- Solaris Cryptographic Framework adds a common API for system-wide cryptographic routines.
- User-selectable security profiles at the time the operating system is installed create a secure building block for a customized system.

Enhanced security profiles are due in a future update of the Solaris 10 OS, which will install with minimized, protected network services and protective firewall services enabled. These features save many hours of configuration and training costs. Effective security reduces downtime, raises quality of service, and keeps costs low.

Buffer Overflow Protection

Stack buffer overflow attacks open up gateway systems and enable hackers to execute unauthorized code, such as a root shell. The Solaris 10 OS protects the entire server against this type of attack—all Solaris 10 privileged utilities have this protection. Stack buffer overflow protection can be applied at the application level for more flexibility and compatibility—carrier and service provider developers can link into designated libraries, protecting their applications. The result is a more reliable and available environment, delivering improved customer satisfaction.

Process Rights Management

Process rights management employs privileges to restrict processes at the command, user, role, or system level. A privilege is a discrete right that is required by a process to perform an operation—the system restricts processes to only those privileges that are required to perform the current task. Therefore, fewer root processes are vulnerable to exploitation. In Solaris 10, the number of `setuid` programs has been greatly reduced.

This means more control over the IP portion of the network and applications that are constrained to their expected tasks. This greatly reduces the chances of key network applications, such as VoIP and unified messaging, being hijacked by hackers.

Devices and system interfaces, which were previously protected with UNIX permissions, are also protected with a security policy that is enforced with privileges. Therefore, the permissions on

a device file do not fully determine the device's availability. Privileges might also be required to operate the device. To further protect system resources, access to all devices at system boot is restricted until the `devfsadm` command runs during the boot sequence. The policy prevents all users except the superuser from initiating connections.

Solaris process rights information is centrally managed to reduce administration costs and increase flexibility for rapidly changing business requirements. As installed, the Software Express releases and the Solaris 10 release are completely compatible with previous releases of the Solaris OS in terms of privilege enhancements. Unmodified programs that run as root run with all privileges.

User Rights Management

Solaris 10 features User Rights Management (referred to as Role-Based Access Control, or RBAC, in earlier releases of the Solaris OS). This enables administrators to restrict access to applications and other Solaris 10 OS functions on a per-user or per-group basis. Administrative privileges can be assigned according to job roles and training levels. This increases security by reducing the chances of administrative errors, as well as accidental or malicious use of IT resources. With User Rights Management, users are granted just the capabilities they need to run a select number of commands, rather than being granted full superuser access to the system.

Solaris Cryptographic Framework Provides Common Security

The Solaris Cryptographic Framework provides a common API for system-wide cryptographic routines. It allows a single point of administration and uniform access to hardware-accelerated cryptographic functions for cryptographically aware applications, and can load balance across accelerators, increasing encrypted network traffic throughput. This framework includes support for:

- IPsec, which provides strong security for TCP data communication, and Internet Key Exchange (IKE), which manages the required encryption/decryption keys. IPsec/IKE can secure almost any protocol without changing the application in both IPv4 and IPv6 environments. Strong encryption is supplied by DES, 3DES, AES, and Blowfish, with support for X.509 certificates.
- Solaris Secure Shell, with enhanced encryption support and integration with Kerberos authentication for single sign-on.
- OpenSSL, an open source set of libraries for secure Web transactions, is integrated into the framework. There is transparent support for hardware acceleration, as well as for encrypted Web pages from an Apache Web server.
- TCP wrapper support, which enables administrators to grant access to specific services based on a domain name. For example, NFS file sharing can be granted to engineering and development, but denied to customer support.
- Simple Authentication and Security Layer (SASL), which provides developers of applications and shared libraries with interfaces for adding authentication, data integrity checking, and encryption to connection-based protocols. SASL enables the developer to write to a generic API without having to be concerned about the details of security mechanisms.

A comprehensive and tightly integrated security framework means a more secure environment with less integration time, strengthening overall security while lowering integration and deployment costs.

In addition to providing a comprehensive security infrastructure, the Solaris Cryptographic Framework delivers an across-the-board performance improvement—from 15 percent to 130 percent on cryptographic routines. Higher performance raises service levels, which means more users per server in e-commerce, or more traffic per server in VPN deployments.

Integrated Firewall Offers Multiple Security Layers

For maximum protection, carriers and service providers use multiple domains within their infrastructures, protecting regional, application, and functional assets. Firewalls commonly protect the gateways between networks and subnets. While firewall capabilities are readily available from many sources, they are not always fully supported by professional services and management infrastructure.

A built-in, stateful IP Filter firewall controls the interaction of services on the network. The open source IP Filter can control access to IP services not only at the gateway, but also to systems and networks inside the firewall. IP Filter is fully supported by Sun Services and can be deployed to provide protection and services to small and branch offices over the Internet, with Network Address Translation capabilities as well as Web, e-mail, and file sharing functionality.

Chapter 3

Reducing Costs by Lowering Capex, Opex

Telecom carriers and providers are under relentless cost pressures. Reducing the number of servers required for a given service lowers capital expenditures; reducing maintenance and support costs lowers operating expenses. The Solaris 10 OS helps telecom providers achieve these goals by:

- **Optimizing Utilization**: Solaris Containers maximize server utilization by running more independent applications on the same hardware. Solaris Containers are used to virtualize operating system services and provide an isolated, secure, and individually manageable environment for running applications.
- **Simplifying Administration**: A simplified and unified installation process can be consistently, reliably, and dynamically applied across LANs and WANs without local intervention. A new and highly innovative file system, ZFS, can dramatically improve data system management, lowering costs and providing a superior environment for new data services.
- **Leveraging Existing Assets**: A strict adherence to open standards and consistent application binary interfaces ensures that carriers can continue to deploy their existing network infrastructures, keeping their capital and deployment expenditures to a minimum.

Maximizing System Resources Lowers Overhead Costs

SLAs mean that service providers maintain hot stand-bys in case of component failure, or if there is a spike in demand elsewhere on the network. Hot stand-bys exist for each service or application, resulting in excess capital expenditures and maintenance costs.

Fault- and security-isolated IT environments protect users and help maintain overall system reliability and availability in telecom environments. The Solaris 10 OS delivers on this core requirement with the introduction of Solaris Containers, which extend the concept of containers. Solaris Containers were first introduced with the Solaris 9 OS to provide full resource containment and control for more predictable service levels. With the release of the Solaris 10 OS, Solaris Containers extend this paradigm to create a private execution environment for applications and services within a single instance of the Solaris OS. Solaris Containers add security to prevent unauthorized access and isolation from faults, minimizing fault propagation and unplanned downtime. Each virtual environment can have separate IP addresses, storage, and root access. They can be dynamically or automatically reconfigured, providing unparalleled resource utilization.

Telecom providers have traditionally obtained additional system availability or scalability by adding hardware. However, when more components are replicated throughout an IT architecture to give it greater resiliency and throughput, the result is a sprawling, complex network of systems that is inefficient and difficult to manage. Capacity planning techniques generally include allocating extra capacity to handle occasional high peak loads. For example, service providers often design their systems to handle traffic volumes far above those that have ever been experienced in order to ensure service availability should a spike in demand occur. This means that normal activity levels may require only a fraction of the total available capacity of the infrastructure—so for much of the time, a percentage of the IT infrastructure is not utilized.

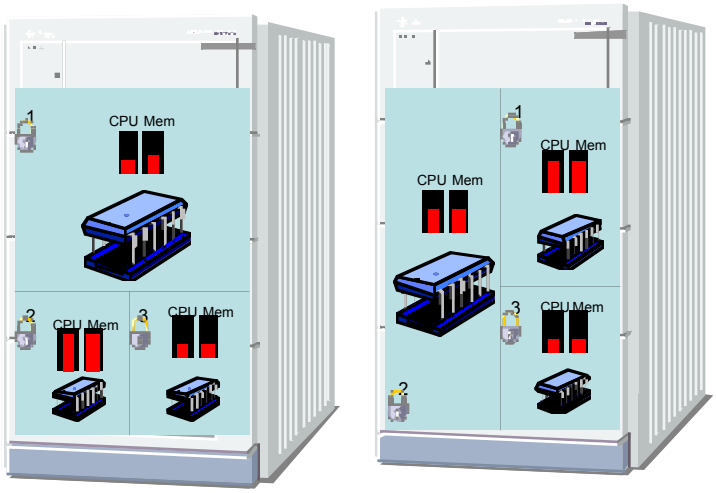

Figure 3-1: A single instance of the Solaris OS with separate Solaris Container environments. Each environment is isolated from the others, and can be reconfigured to adapt to changing loads.

Solaris Containers enable other applications to share resources with the systems when those resources are underutilized—a much more cost-effective implementation. In the event of a spike in traffic volume, resources can be dynamically reallocated to the appropriate systems while the less-critical applications receive fewer resources, ensuring that the primary customer systems function as needed. Sharing resources in this way reduces total capital investment costs and lowers system management expenses by reducing the total number of systems required. For example, it may be possible to consolidate eight two-CPU servers with two single-CPU servers configured with eight containers.

In order for server consolidation to be truly effective, telecom providers must still be able to manage each application independently. This requires the ability to control IT resource utilization, isolate applications from each other, and efficiently manage multiple applications on the same

server. In other words, it requires establishing easily managed virtual server boundaries within a single system on which to provision applications.

A Solaris Container can be thought of as a box. One or more applications can run in this box without affecting the rest of the system. This isolation prevents processes that are running in one container from monitoring or interfering with processes that are running in other containers. Even a process with superuser credentials that is running inside a container cannot view or affect activity in other containers. Containers can also be administered by individual zone administrators. The privileges of a zone administrator are confined to a unique zone. However, system-related administrative tasks are performed for the entire system, saving time and money.

Unique containers provide isolation at almost any level of granularity required. A container does not need a dedicated CPU, physical device, or portion of physical memory. These resources can either be multiplexed across a number of containers running within a single domain or system, or allocated on a per-container basis utilizing the resource management features available in the operating system. To achieve process isolation, a process can see or signal only those processes that exist in the same container.

Overhead is very low—even a small, single-processor system can support a number of containers running simultaneously. Typically, four-way CPUs can support hundreds of containers, and thousands of containers can be supported in larger systems. Applications can be loaded to a unique container or to the global environment.

Disk and network interfaces are managed centrally, but are leveraged by all containers—since all containers share a single file system, less storage infrastructure is needed. Basic communication between containers is provided by giving each container at least one logical network interface. Applications running in different containers on the same system can bind to the same network port by using the distinct IP addresses associated with each container or by using the wildcard address. An application running in one container cannot observe the network traffic of another container. This isolation is maintained even though the respective streams of packets travel through the same physical interface.

Each container is given a portion of the file system hierarchy. Because each container is confined to its subtree of the file system hierarchy, a workload running in a particular container cannot access the on-disk data of another workload running in another container. Files used by naming services reside within a container's own root file system view. Thus, naming services in different containers are isolated from one other and can be configured differently.

Security is provided by Solaris least privileges capabilities. Each container has a limited set of capabilities assigned, as required. For example, `sendmail` may be run as root, but the ability to fork an `exec` process can be removed. This helps prevent, among other things, a buffer overflow attack. Least privileges are fully described later in this paper.

Minimize Costs and Reduce Time to Service—Test and Deploy in Identical Environments

In addition to maximizing resource utilization, Solaris Containers offer telecom providers a unique ability to provision and deploy applications. Operations can employ testing and deployment environments on the same hardware, with the same instance of the Solaris OS identically configured. Many NEPs have lab systems that support different release views for developers to use in testing any changes. Solaris Containers can reduce the amount of time spent switching between these views. Any issues with testing are isolated to the preproduction environment. Once the application is certified, it can be deployed with the assurance that it was tested in an identical environment. New or updated applications can be rolled out as self-contained environments,

further minimizing administration costs. Service providers can be assured that the production environment is identical to the testing and developer environments, minimizing deployment risks and reducing the number of systems required.

Managing Services, Not Systems

Solaris Containers enable system administrators to describe the service levels, dependencies, and relationships between the various software components that comprise a business service. Solaris Containers manage the end-to-end availability of the business service as a whole. By enabling IT staff to manage complete services rather than the individual components that make up those services, Solaris Containers reduce management complexity and increase system utilization, lowering operating expenses and TCO for the organization.

Solaris Containers also include enhanced resource usage accounting. This highly granular and extensive resource tracking capability can support advanced client billing models employed throughout the telecom industry.

Solaris Containers work across all Solaris 10 supported hardware platforms—SPARC, x86, and AMD Opteron single- and multi-CPU systems.

Advanced Administration and Management Reduce Overhead

Carriers and service providers need the ability to manage large or distributed networks, securely and efficiently provisioning, reconfiguring, and redeploying IT resources. The ability to quickly and inexpensively provision and retask servers is an important part of reducing maintenance and overhead in a telecom environment. With potentially hundreds or thousands of systems to manage, automation is essential. Reducing TCO depends on enabling a small administrative staff to manage a large number of systems. Solaris 10 offers a comprehensive set of tools and utilities for this purpose.

Solaris Flash Automates Upgrades and Redeployment

> *Cloning disks has long been an accepted practice for distributing disk images. But while cloning can be useful, updating and redeploying servers often calls for incremental upgrades, such as service packs, patches, or host-dependent files. When updating, cloning often requires servers to be offline for a significant amount of time while they are reconfigured—reducing overall efficiency.*

The Solaris Flash feature provides new installation and provisioning functionality. System administrators can capture a snapshot image of a complete server—including the Solaris OS, applications stack, and system configuration—into a new Flash Archive format. Using this system image, administrators can then replicate reference server configurations onto multiple (clone) servers. Solaris Flash images can be deployed via standard media or over the network via HTTP and NFS. The Solaris Flash images can be installed using custom Solaris JumpStart™ scripts, the Solaris Web Start graphical interface, or Solaris interactive installation.

Once the master is created, it can be replicated out to any number of clone systems. Solaris Flash installations can overwrite the existing clone OS, or use the update facility to install only the differences between master and clone systems. A differential update changes only the files that are specified, and is restricted to systems that contain software consistent with the old master image.

Solaris Flash installations are very customizable. System administrators can:

- Install a Solaris Flash archive with additional packages. For example, install the same base archive on two machines, but add a different set of packages to each machine. These packages do not have to be part of the Solaris OS distribution.
- Install a customized list of patches.
- Identify nonclonable, host-dependent data, which enables the Flash archive to be host independent. Host independence is enabled by modifying such data or excluding the data from the archive. An example of host-dependent data is a log file.
- Validate software integrity in the archive during creation.
- Validate the installation on the clone system.
- The Solaris software now enables booting and installing software over a WAN using HTTP. The WAN boot installation method allows administrators to install the Solaris release on systems over a large public network where the network infrastructure might be untrustworthy. WAN boot can be used with new security features to protect data confidentiality and installation image integrity.

Solaris Flash technology can also be employed to store existing server configurations, thus making them an FRU, where reinstallation can be achieved within a few minutes in the event of a complete system failure. Solaris Flash is ideal for the replication and deployment of server farms, with features specifically designed for service providers and others that require rapid deployment and the ability to reprovision servers on demand. For example, carriers and service providers can easily reconfigure a cache server to a Web server to reprovision their services by redeploying, thereby increasing service levels. This can be accomplished by deploying a Flash Archive containing the Solaris OS and files to run a Web server.

Solaris JumpStart Software

A standard component of the Solaris 10 OS, Solaris JumpStart software enables fully automatic and remote operating system installation based on custom profiles and scripts.

Solaris Live Upgrade Software

Solaris Live Upgrade software allows the creation, management, upgrade, comparison, and activation of multiple boot environments on systems running the Solaris OS. In particular, Solaris Live Upgrade software enables systems to run uninterrupted while a system administrator installs a Solaris Flash archive or upgrades to a new version of the Solaris OS. As a result, downtime for an upgrade can be reduced to the time needed for a simple reboot.

Reduce File System Management and Overhead with ZFS

When reviewing the design of an upcoming voicemail service, a system architect comments that it requires yet another file system. A discussion ensues regarding the relative trade-offs of a file system optimized for voice, compared to the additional training and deployment costs. In the end, the team approves the specialized file system, but notes that additional data clustering and backup resources may be needed.

Solaris 10 delivers a new model of a storage file system, ZFS, which replaces and significantly improves upon traditional volume managers and file systems. For telecom providers and carriers, ZFS provides numerous enhancements, including automating many common tasks usually performed by system administrators, such as protecting data from corruption, while providing

virtually unlimited scalability and interoperability. ZFS offers high availability, scalability, and performance, increasing service levels for virtually all application and service data.

Data Protection

Data can disappear in many ways—system errors, power outages, and so on. ZFS uses many mechanisms to protect data, keeping it available for all users at all times. All operations are copy-on-write, writing a new block before changing data pointers and committing the write. This ensures always-valid on-disk states, provides consistent, reliable backups, and enables data rollbacks to a known point in time. Data is protected by 64-bit checksums for 21 nine's reliability in error detection and correction. Copy–on-write operations eliminate the need for `fsck`.
As part of a self-healing systems initiative, ZFS can self-heal data in a mirrored or RAID configuration. When one copy is damaged, ZFS detects it via the checksum and uses another copy to repair it.

Interoperability

ZFS is POSIX compliant, making it transparent to virtually all applications. It supports both SPARC and x86 platforms, easily moving disks between the two hardware environments. Neither architecture suffers byte-swapping performance issues due to Sun's patent-pending *adaptive endian-ness* technology. Sun continues to support the UFS file system.

Scalability and Performance

ZFS is presented to administrators as a pooled storage mode. This eliminates the legacy concept of volumes, as well as all of the related partition management, provisioning, and file system sizing problems. Multiple block sizes can be automatically chosen to match application and workload requirements. Thousands of file systems can draw from its common storage pool, each consuming only as much space as it needs—service providers can deploy unique file systems for each service, such as video, voice, text messaging, and so on.

The combined I/O bandwidth of all of the devices in that storage pool is always available to each file system. Administrators can set explicit I/O priority with deadline scheduling, and create multiple, independent prefetch streams with automatic length and stride detection.

A 128-bit file system, ZFS offers 16 billion *billion* times the capacity of 32- or 64-bit file systems. ZFS offers exceptional performance—it is based on a transactional object model that removes most of the traditional constraints on the order of issuing I/Os, which results in huge performance gains. Service providers and carriers, who by definition manage millions of subscribers, can leverage ZFS to offer their customers new services with maximum performance and availability.

Naming and Directory Services Enhance Security, Lower Costs

New subscriber services increasingly rely on user identity. Protecting user identities is essential in an e-commerce environment where subscriber information is securely requested and delivered among multiple entities. In addition, interoperability and management among multiple carrier-grade directories are also requirements.
Solaris 10 contains a number of directory service improvements, especially in the area of LDAP.

- The Solaris installation programs now support LDAP Version 2 profiles. These profiles enable administrators to configure systems to use a proxy credential level.

- A NIS-to-LDAP transition service can be used to switch a network transition from NIS to LDAP as the primary naming service. With this transition service, administrators can utilize the Java System Directory Server, which works with LDAP naming service clients.
- Several LDAP commands are updated to include full SSL support and extended support for SASL. The modifications also provide support for managing smart referrals, using virtual list views (VLVs), and establishing stronger authentication when binding to the LDAP server.
- A password management feature enables tracking password aging and expiration, prevents users from choosing trivial or previously used passwords, warns users if their passwords are about to expire, and locks out users after repeated login failures.

The BIND version 9.2 name server and tools are available, in addition to BIND version 8.4.2. BIND 9 offers full IPv4 and IPv6 functionality.

Multiprotocol Routing Optimizes Network Traffic

The Solaris 10 OS now includes the Zebra multiprotocol routing suite. This enables system administrators to use well-known open source routing protocols such as Routing Information Protocol (RIP), Border Gateway Protocol 4 (BGP-4), and Open Shortest Path First-Multipathing (OSPF-MP) for administering their Solaris based networks. BGP-4 and OSPF make it easier to administer complex routing policies. The OSPF daemon can also be used for high network availability on multihomed servers. This provides end-to-end Layer 3 redundancy, facilitating greater protection from network failures. Multiprotocol routing helps service providers create more efficient and manageable networks, lowering costs and exerting more control over SLAs.

Compatibility and Interoperability Lower Deployment Costs

Using a commercially available UNIX platform, a service provider successfully developed and tested a call center application. Before deployment, the provider considered upgrading to a later OS release to take advantage of newer features. Unfortunately, enough of the binary interfaces had changed that the development team determined the program would need to be retested and recertified for the newer OS. The provider opted to stay with the older release.

Carriers and service providers make substantial investments in their IT capital assets, and want to continue using them in production environments as long as possible—often longer than in enterprise environments. Solaris 10 continues Sun's long-term support for interoperability and open standards, including traditional standards and more recent innovations.

The Solaris Application Binary Interface (ABI) is a set of supported, run-time interfaces available for an application to use on the Solaris OS. The ABI toolset determines whether or not a binary built on one release of Solaris is able to run on subsequent releases. This release-to-release binary compatibility is important to telecom service providers, because it means that their investment in applications can be preserved across Solaris OS upgrades.

Source-code compatibility alone is not sufficient for most service providers for two reasons: they may not have the source code for their applications and hence cannot recompile, or they may have the source code but find that the cost of recompilation, revalidation, certification, and redistribution is prohibitively high in terms of time and resources.

Consider a large telecom operation with literally hundreds of purchased or in-house applications running the business. Even if source code recompilation was an option, the process of

recompiling, revalidating, and distributing the result to thousands of computers would be extremely expensive and time-consuming.

Linux Compatibility

The Solaris 10 OS kernel includes a handler for Linux system calls for x86 architecture platforms. On top of the Solaris kernel is a set of UNIX APIs that provide public interfaces to the Solaris kernel. Sun recently expanded its set of APIs to include x86 Linux Native Libraries, allowing Linux applications to run unchanged on the Solaris OS when coupled with a Linux distribution.

Solaris 10 offers Linux interoperability with a new feature, the Linux Application Environment, for x86 platforms including AMD Opteron and Intel Xeon-based systems. This kernel-based handler manages system calls issued by Linux applications, and dispatches the equivalent Solaris kernel functions to handle the requests. The module takes care of device-specific requests, so that even device-dependent Linux software can run unchanged. In addition, Sun recently expanded its set of APIs to include x86 Linux Native Libraries, allowing Linux applications to run unchanged on the Solaris OS when coupled with a Linux distribution. This enables businesses to take advantage of innovations in the Solaris 10 OS without sacrificing investments in existing Linux applications. The API is designed for 100-percent compatibility with Red Hat Enterprise Linux 3.

Linux compatibility is kernel-integrated and supported as an operating system feature. This means it can take advantage of other Solaris 10 features, such as Solaris Containers, security features, and ZFS file systems.

Chapter 4

Leveraging the Next-Generation Network

The next growth phase of the Internet will be driven not by computer users, but rather by telecommunication users. With billions of people accessing the Internet via mobile phones, kiosks, PDAs, and other devices, an entirely new communication infrastructure is required. Solaris 10 provides an excellent foundation for the delivery of reliable services in support of this emerging, next-generation network.

New IP-based Services with SCTP and SIP

Internet technologies are driving new telecom business initiatives. The rapid adoption of IP-enabled mobile phones, combined with the carrier's need to increase revenue, is driving the creation of new services such as Voice over IP (VoIP), location-based services, Web conferencing, instant messaging, push-over-cellular, and presence-based services that take advantage of the network in innovative ways. For example, next-generation services include credit card authorization via SMS, retail outlet recommendations based on preferences and location, ringback tones, buddy finder programs, TV to the handset, and more. Solaris 10 enables the creation of these new services with the introduction of two important protocols: Stream Control Transmission Protocol (SCTP) and Session Initiation Protocol (SIP).

- Designed by the IETF SIGTRAN working group, SCTP will carry SS7 messages over IP networks. SCTP’s robustness stems from its capacity to maintain several data streams (multihoming) within a single connection. This makes SCTP ideally suited for connecting and monitoring wireless devices, cell phones, and Internet appliances.

- SIP is an application-layer protocol that establishes, modifies, and terminates interactive sessions, including VoIP sessions—the Internet age equivalent of SS7. While typically used for streaming audio and video applications, incorporating SIP into the Solaris OS offers the potential for next-generation applications and devices.

In-kernel support for SCTP and SIP makes Solaris 10 an ideal development and deployment platform for VoIP and other telephony applications.

Unlimited Addressing with IPv6 Opens New Markets

Faced with exponential growth, the international community came together to define a new standard that could support the billions of devices likely to populate global networks. That standard, IPv6, creates a virtually unlimited number of unique IP addresses so that the world can continue its rapid adoption of the Internet. The importance of this technology is best demonstrated by China, which launched the China Next Generation Internet (CNGI) in 2003, signaling a concrete development period for IPv6 commercialization. By the end of 2005, China will have one of the largest IPv6 networks in the world.

IPv6 is a fundamental technology that enables service providers to create next-generation services. It strengthens security, improves network efficiency, integrates mobile devices, and improves end-user and back-end configuration.

Solaris 10 fully supports the current IPv6 specifications and APIs. The Solaris 10 IPv6 implementation offers updated APIs, security and privacy extensions, and transition mechanisms. The IP address size increases from 32 bits in IPv4 to 128 bits in IPv6, resulting in a virtually unlimited number of addresses. IPv6 addresses are assigned to interfaces, rather than nodes, in recognition that a node can have more than one interface. In addition, more than one IPv6 address can be assigned to an interface.

The Solaris IPv6 implementation fully integrates with Solaris IPsec, including IKE, DNS, and NFS. Sun offers comprehensive support for IPv6 in the Java Enterprise System, including mail, directory, Web, and application servers. The Java Developer Kit (JDK™) 1.5 also supports IPv6. This means that applications written in the Java programming language can work transparently with either IPv4 or IPv6, and run on a Solaris 10 environment.

Solaris 10 offers a strong migration path from IPv4, including dual network stacks and tunneling tools. For most carriers and telecom providers, the introduction of IPv6 to an existing IPv4 network must occur on a gradual, step-by-step basis. The Solaris dual-stack network environment supports both IPv4 and IPv6 functionality. Because most networks use the IPv4 protocol, IPv6 networks currently require a way to communicate outside their borders. IPv6 networks use tunnels for this purpose.

In most IPv6 tunneling scenarios, the outbound IPv6 packet is encapsulated inside an IPv4 packet. The boundary router of the IPv6 network sets up a point-to-point tunnel over various IPv4 networks to the boundary router of the destination IPv6 network. The packet travels over the tunnel to the destination network's boundary router, which decapsulates the packet. Then, the router forwards the separate IPv6 packet to the destination node.

Platform Choice

Solaris 10 supports a broad range of platforms – 270 and growing[2].

Sun and the Solaris OS have long demonstrated commitment to the exacting hardware requirements of the telecom industry. This includes support for Certified NEBS Level 3 hardware, including automatic restart, component redundancy, and hot swap.

Solaris 10 will run on a forthcoming family of servers that will comply with the AdvancedTCA standard. AdvancedTCA incorporates the latest trends in high-speed interconnect technologies and next-generation processors, as well as improved reliability, manageability, and serviceability, resulting in a new blade (board) and chassis (shelf) form factor that is optimized for communication. AdvancedTCA provides a standardized platform architecture for carrier-grade telecommunication applications, with support for carrier-grade features such as NEBS, ETSI, and 99.999-percent availability.

For many network equipment manufacturers, the switching architecture and chassis was a core technology traditionally designed and manufactured in house. The introduction of AdvancedTCA chassis, blades, and switch cards makes this standardized approach a viable option for even high-performance systems, leveraging the COTS model to bring down operating costs.

2. See www.sun.com/bigadmin/hcl for an up-to-date hardware compatibility list.

Chapter 5

Summary

The rapid growth of Internet-enabled devices is driving the development of new network applications and services, throughput-oriented infrastructure, and secure connections. This growth creates both opportunities and challenges for carriers, NEPs, and service providers. While it represents a dramatic opportunity for new sources of revenue, managing the convergence of legacy and IP-based networks poses a significant challenge. The pressure to reduce costs across the board as these new IP-based applications and services come online remains a top priority.

Solaris 10 represents a breakthrough in operating system technology that can help carriers, NEPs, and service providers meet these challenges and realize the opportunities by:

- **Improving service levels**: The Solaris 10 OS offers performance-enhancing improvements and capabilities throughout the system—network stack, OS, applications, and more. Innovative features such as DTrace for real-time debugging, predictive self-healing, and IPQoS improve performance, availability and reliability. These new capabilities help drive higher service levels and increase customer satisfaction.
- **Reducing costs**: With Solaris Containers, utilization of development and production systems can be increased—in many cases reducing the number of systems required. Highly automated tools facilitate the efficient administration management of complex systems. New, self-healing file system technology—ZFS—can help reduce data management overhead.
- **Leveraging next-generation networks**: The Solaris 10 OS supports new and emerging telecom standards, such as SIP, SCTP, and AdvancedTCA, which represent Sun's continuing investment in the telecom market. NEPs and carriers can optimize throughput and cost-effectiveness for new 3G network applications such as push-to-talk, push-to-video, multiparty gaming, presence services, video conferencing, and many more.

The largest telecommunication companies in the world have come to rely on Sun for carrier-grade systems that can stand up to the high reliability and scalability demands of supporting subscriber bases in the millions. The Solaris OS has been there from the start. Now, with the Solaris 10 OS, the telecom industry stands to gain more than ever from Sun's delivery of a set of truly revolutionary features and capabilities that work in concert to improve their business. The Solaris 10 OS is the right choice for telecom.

Chapter 6

More Information

To learn more about the Solaris 10 OS, visit the Web sites below.

Sun in Telecom	www.sun.com/solutions/landing/industry/telecom.xml
Solaris 10 OS	sun.com/software/solaris/10/index.jsp
FireEngine	sun.com/bigadmin/xperts/sessions/11_fireengine/
DTrace	sun.com/software/solaris/10/ds/dtrace.jsp
Predictive Self-Healing	sun.com/software/solaris/10/ds/self_healing.jsp
Network Performance	sun.com/software/solaris/10/ds/network_performance.jsp
Solaris Containers	sun.com/software/solaris/10/ds/containers.jsp
Security	sun.com/software/solaris/10/ds/security.jsp
Linux Interoperability	sun.com/software/solaris/10/ds/linux_interop.jsp
ZFS	sun.com/software/solaris/10/ds/zfs.jsp
Flash Archives	sun.com/software/solaris/webstartflash/
Java SDK 1.5/J2SE 5.0	java.sun.com/j2se/1.5.0/download.jsp
NEBS	sun.com/products-n-solutions/hw/networking/

Learn More

Get the inside story on the trends and technologies shaping the future of computing by signing up for the Sun Inner Circle program. You'll receive a monthly newsletter packed with information on the latest innovations, plus access to a wealth of resources. Register today to join the Sun Inner Circle Program at sun.com/joinic.

To receive additional information on Sun software, products, programs, and solutions, visit sun.com/software.

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA Phone 800 786-7638 or +1 512 434-1577 Web sun.com

Sun microsystems
We make the net work.

Sun Worldwide Sales Offices: Africa (North, West and Central) +33-13-067-4680, Argentina +5411-4317-5600, Australia +61-2-9844-5000, Austria +43-1-60563-0, Belgium +32-2-704-8000, Brazil +55-11-5187-2100, Canada +905-477-6745, Chile +56-2-3724500, Colombia +571-629-2323, Commonwealth of Independent States +7-502-935-8411, Czech Republic +420-2-3300-9311, Denmark +45 4556 5000, Egypt +202-570-9442, Estonia +372-6-308-900, Finland +358-9-525-561, France +33-134-03-00-00, Germany +49-89-46008-0, Greece +30-1-618-8111, Hungary +36-1-489-8900, Iceland +354-563-3010, India–Bangalore +91-80-2298989/2295454; New Delhi +91-11-6106000; Mumbai +91-22-697-8111, Ireland +353-1-8055-666, Israel +972-9-9710500, Italy +39-02-641511, Japan +81-3-5717-5000, Kazakhstan +7-3272-466774, Korea +822-2193-5114, Latvia +371-750-3700, Lithuania +370-729-8468, Luxembourg +352-49 11 33 1, Malaysia +603-21161888, Mexico +52-5-258-6100, The Netherlands +00-31-33-45-15-000, New Zealand–Auckland +64-9-976-6800; Wellington +64-4-462-0780, Norway +47 23 36 96 00, People's Republic of China–Beijing +86-10-6803-5588; Chengdu +86-28-619-9333; Guangzhou +86-20-8755-5900; Shanghai +86-21-6466-1228; Hong Kong +852-2202-6688, Poland +48-22-8747800, Portugal +351-21-4134000, Russia +7-502-935-8411, Singapore +65-6438-1888, Slovak Republic +421-2-4342-94-85, South Africa +27 11 256-6300, Spain +34-91-596-9900, Sweden +46-8-631-10-00, Switzerland–German 41-1-908-90-00; French 41-22-999-0444, Taiwan +886-2-8732-9933, Thailand +662-344-6888, Turkey +90-212-335-22-00, United Arab Emirates +9714-3366333, United Kingdom +44-1-276-20444, United States +1-800-555-9SUN or +1-650-960-1300, Venezuela +58-2-905-3800