PROTECTING INVESTMENTS THROUGH TECHNOLOGY ADVANCEMENTS

Brian Down, Client Solutions Organization—Global Data Center Practice

Sun BluePrints™ OnLine—October 2005

Part No 819-3931-10
Revision 1.0, 9/26/05
Edition: October 2005

Please
Recycle

Table of Contents

Protecting Investments Through Technology Advancements

Introduction

Many businesses depend on their IT infrastructure, and in fact, use it to gain competitive advantage. Toward this end, IT organizations are constantly seeking new ways to implement these vital assets in a cost-effective manner that supports business goals. In addition, budget pressures are pushing organizations to find ways to protect investments. Doing so requires an IT architecture that is designed, implemented and deployed with components that help you manage risk and increase both efficiency and agility so that the business operates optimally.

As with all investments, investors want to know they are making a sound investment in their IT infrastructure—an investment that will provide good value over time. The good news is the IT industry has matured, and most IT investments are well understood in terms of measurement and efficiency. This Sun BluePrints™ article explains what it means to protect IT investments, and what you need to consider when protecting them. It also illustrates how Sun's platform of UltraSPARC® processor-based servers running the Solaris™ Operating System can be used to build an infrastructure with investment protection built-in.

Investments That Provide Value

Because IT assets depreciate, it is important they provide *value*—business flexibility, agility, and efficiency—for as long as possible, and be easy to replace when the time comes. This concept is an essential component of determining in what technology to invest, as well as deciding when and how to make that investment. What was once leading edge technology can—and will—quickly become out-of-date. A million dollar investment in IT infrastructure can have a book value of only a few thousand dollars, or less, within a decade. As the value of this investment deteriorates, the competitive advantage it initially offered also begins to wane as that technology becomes ubiquitous. The challenge is to invest in an IT infrastructure that not only provides an immediate benefit to the enterprise, but to also ensure this technology has the flexibility and longevity to provide good value over time.

What Constitutes an IT Infrastructure?

Companies tend to have lots of computing equipment spread throughout the enterprise. An IT infrastructure is more than just a collection of computer systems. Consider the typical enterprise stack—a construct that represents, in a simplified manner, three distinct types of architectures that comprise an IT environment: the business, execution, and management architectures (Figure 1).

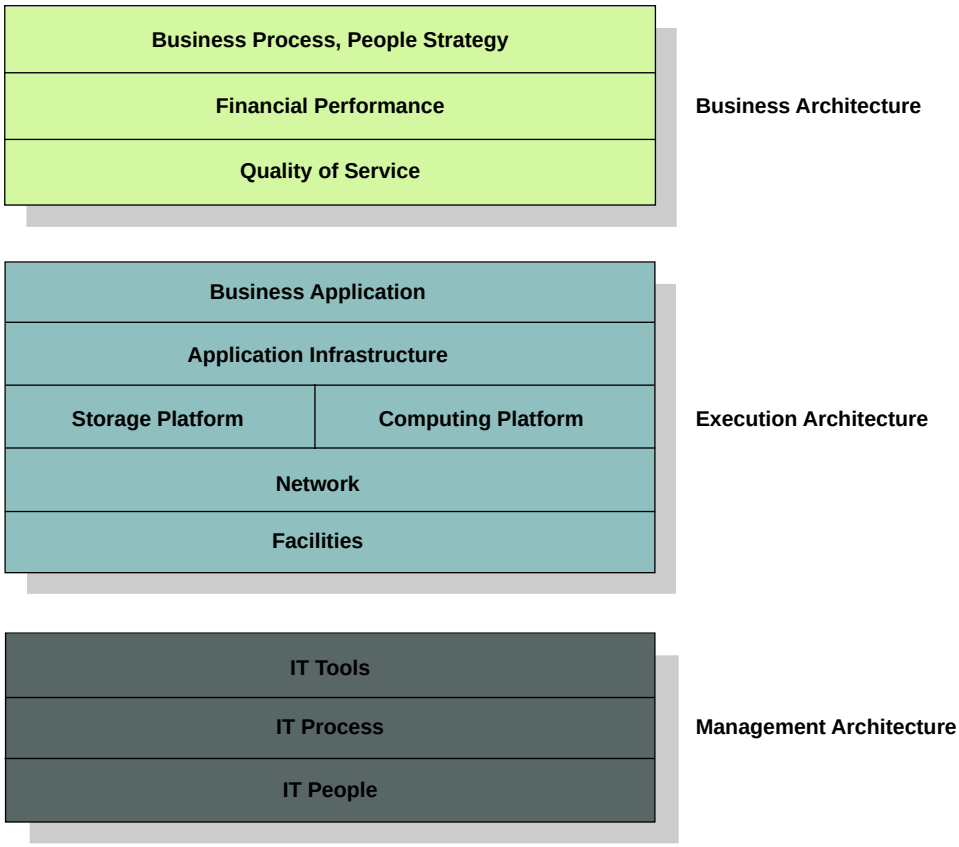

Figure 1. A typical enterprise stack

The *business layer* includes tasks related to the overall business, such as setting corporate direction, investment strategies, improving business processes, quality of service (QoS) metrics, partner relationships, financial goals, and more. It is this level of the organization that makes decisions about business execution and the technology to be used to execute that mandate. It must be able to respond to competitive pressures and adapt to outside forces, such as evolving regulatory requirements, that can force change upon an organization. Understanding the technological trends that exist in competitive enterprises, and knowing what technology is currently available, is critical to making a sound investment in an IT infrastructure that will support these endeavors.

The *management laye*r includes the people, processes, and tools used to manage the execution layer and meet the QoS metrics and goals defined by the business layer. Comprising a significant portion of an IT budget, IT staff create and manage the IT infrastructure, and ensure the requirements of the business layer are met. Processes are used to operate and change the IT infrastructure. Adopting a proven management methodology helps streamline data center operation by ensuring problems are reported to the correct entity, service level agreements (SLAs) are monitored and met, resources are deployed in a dynamic fashion to meet projected peaks, changes are documented and rolled out with minimal impact on the production environment, and so on. Implementing the correct management policies helps ensure costly service interruptions are kept to a minimum. In addition, the sheer number of applications, services,

compute platforms, network segments, and switches has increased to such a point that it is impractical for many installations to monitor them manually. Tools that monitor utilization and capacity, record measurements in a central location, and analyze them, may remediate problems in an automated fashion or provide an alert for manual intervention.

While the business and management layers play an important role, the *execution layer* is the heart of the IT infrastructure and serves as the engine room of the IT organization. Typically thought of as *the investment*, the execution architecture is composed of items found in the glass room, or data center. These components include applications that implement the business rules for the enterprise, the software environment that supports these applications, the underlying computing platform and storage devices, networking to tie it all together, and the facilities that house it all. Understanding how best to analyze IT infrastructure choices and protect investments in these components is essential to optimizing the data center and meeting business objectives.

When evaluating the efficiency of, or the effect of change to an IT infrastructure, it is very important that the impact to the entire enterprise stack, not just the execution architecture, is taken into account. For example, a change to the business architecture may lead to the acquisition of a large compute grid that will provide financial or competitive benefit to the enterprise. However, the increase in complexity to the execution and management architectures must also be considered, in terms of operational and facilities overhead, and the increased number of operating system instances that must be maintained and managed. The common reaction to addressing tactical short-term computing requirements—acquiring hardware platforms in an ad hoc manner—often leads to the deployment of an infrastructure that is difficult to maintain, significantly reducing the value of that IT investment.

Protecting IT Infrastructure Investments

Protecting an investment in IT infrastructure is built on the principle of managing risk without impacting—indeed while increasing—agility and efficiency.

- *Managing business risk—minimizing unpredictable costs and ensuring solution longevity*

  Pressure to deliver results to the bottom line is forcing IT managers to find ways to drive costs out of their operations and align IT infrastructures with business priorities. As companies look for ways to address these concerns, they must be mindful of the risks change of any kind can pose to the business. Spending should be—must be—predictable. Changes in one part of the data center should not cause a cascading effect throughout the entire IT organization. Computing infrastructures must be able to scale both horizontally and vertically to keep pace with changing business demand and ensure business opportunities are not missed. At the same time, other business requirements must also be considered. Systems must be available and ensure service level agreements are met—without creating complex environments that cannot be controlled or managed. Finally, security that protects businesses and identities during times of heightened global tensions is becoming more important every day as businesses go online.

  How is risk managed? Through the adoption of the right innovative, yet stable technologies; a commitment to compatibility that ensures existing hardware and software assets are not wasted and can

be put to use where they are needed most; flexible systems that adopt and take advantage of these advancements; consistency across systems; and a systemic approach to implementing security.

- *Increasing agility—adapting to change and meeting business demands*

  Change is everywhere, and it is inevitable. Technology has continually changed and advanced—and that trend is likely to continue for the foreseeable future, given the proven accuracy of Moore's Law. This continual change often results in complex solutions comprised of archaic systems for which components are no longer available. Compounding these difficulties: it can be difficult or cost prohibitive to get or retain expertise for these older technologies. Solutions are superseded by better, more cost-effective ones that must be integrated with the existing environment. In addition, hardware wears out. It may experience increased failures, or simply run to the end of its lease or depreciation cycle. Either way, it is time to replace equipment. Furthermore, consolidation within the computing industry is forcing IT executives to make difficult decisions. Compaq's acquisition of DEC and its subsequent merger with Hewlett Packard has resulted in the end-of-life (EOL) of a number of operating systems (Tru64, eMPE) and processors (Alpha, PA-RISC). Clearly, investments in these sorts of assets provides limited protection.

  Change does not just happen in hardware. New business pressures fuel changes to the enterprise. These demands may include increased and dynamically changing workloads, additional applications to support new (internal or external) services, and more. Changes to the infrastructure occur all the time, including fresh approaches to resource utilization and accounting, identity management, and security. New operational requirements may present themselves, such as a need to consolidate lab facilities or support disaster recovery. To be effective, the data center must be agile and able to respond quickly to these changing business demands. To faciliate change and flexibility, the IT environment should not be “hardwired”. By creating a flexible environment—one in which applications are not tied to specific operating systems, operating systems are not tied to specific compute platforms, and compute platforms are not directly attached to specific storage platforms—enables flexibility and agility and helps ensure the infrastructure can be provisioned in a dynamic fashion to meet cyclical business demands.

  Not only is change a given, but in many industries it seems the rate of change is actually accelerating. This is particularly relevant to IT organizations. Indeed, change affects IT from the top down and the bottom up. Top down change can be thought of a change that originates in the business architecture, and can be something as exciting as a new product launch or business direction, or something as mundane as the introduction of a new regulatory framework. Bottom up change arises from software or hardware innovation that fundamentally alters the way things are done, and typically arises as a result of the introduction of a new technology. For example, the microprocessor made it possible for an individual business (or just an individual) to own their own computer and move away from the centralized mainframe computing model. Other examples of bottom up change include the introduction of a new paradigm like the client-server model, which changed how applications are written; or the creation of the World Wide Web, which has changed how people store and retrieve information and conduct electronic commerce.

Regardless of the agent of change, the IT infrastructure must adapt to accommodate these new requirements. Compatibility is key. By utilizing systems that ensure instruction set compatibility and application programming interface (API) consistency, as well as enable systems and their components to be reused, repurposed, and mixed and matched, organizations are better able to adapt to evolving requirements without suffering a cascading effect throughout the IT infrastructure.

- *Increasing efficiency—getting more out of currently available and future resources*

  Do more with less. It is an old adage that is particularly relevant today. Finding a way to increase efficiency and get more out of available resources is one of the best ways to protect significant IT infrastructure investments. These efficiencies can be realized by improving system utilization, decreasing licensing costs, reducing costs associated with facilities and maintenance, and reducing IT staff replication.

Virtualization—A Cornerstone of Investment Protection

With the escalating costs of managing large networks of servers and software components, companies are seeking new ways to reduce IT infrastructure costs and better manage end user service levels. Sun's vision is that businesses can accomplish this through a technique called *virtualization*. Just as the operating system virtualizes the capabilites of the microprocessor hardware to facilite ease of use, creating an IT environment based on virtualized compute, network, and storage resources aggregated as services helps ease business reconfiguration, deployment, and management.

Consider a typical data center. The IT infrastructure mimics the business structure with each different business unit owning its own compute resources. This silo-like structure lends itself to duplication with applications tightly-coupled to the compute platform, a replication of IT management processes and an ad-hoc IT roadmap. Resources cannot be easily moved from one silo to another because of mismatches between operating system of software versions. Frequently, it is difficult to measure and monitor resource utilization because of the different tools required by different hardware and software vendors. Storage and networking capabilites are all hardwired to specific platforms, making reconfiguration an extremely difficult task. Not surprisingly, this hardwired, silo approach to building IT infrastructures leads to low utilization levels on individual components as the infrastructure cannot be easily reconfigured so that spare resources from one business unit can be used by another that is running out of capacity.

Security and application isolation are other concerns. One business unit may not want the sensitive data stored on its infrastructure to be visible to others within the organization. Additionally, to ensure operations are not effected by events elsewhere in the enterprise, one business unit may not wish to share its resources with another, fearing that increased utilization levels on those machines will impact its ability to meet sevice level agreements.

At the same time, economic realities and changing business requirements are forcing enterprises to consolidate applications onto fewer servers. Security and resource availability still require isolation be maintained between applications. Virtualization allows large servers to be flexibly partitioned into independent execution environments that provide total isolation within the same server. It allows data

center resources to be visualized and managed as a fabric of interconnected computing resources rather than as a room filled with individual systems.

Virtualizing the environment requires the deployment of products and technologies focused on this goal—from microprocessors, to systems and operating environments, to middleware and applications—and leads to greater IT flexibility.

Processor Compatibility Is Important

When an incompatibility occurs in the upgrade cycle:

- *Planning costs increase*
- *Costly operating system and application software upgrades are likely required*
- *Additional hardware upgrades may be forced on the organization*
- Applications do not run, and often require recompilation

When the incompatibility occurs at the heart of the computing platform—in the microprocessor—it can have a far reaching impact. Eliminating these unnecessary and costly events is a key factor in creating a predictable (both in terms of functionality and cost) computing infrastructure. By ensuring stability and compatibility from one processor version to the next, new systems are able to immediately fit into the enterprise without disrupting the software stack which comprises business-critical or business support applications.

Open Availability and Longevity Are Essential

Unlike financial investments that we hope will appreciate over time, capital investments in IT infrastructure are expected to depreciate. Indeed, improved design and fabrication capabilities ensure the next generation of processors and systems will be faster and cheaper than their predecessors. New and different architectures can hold promise—if they are adopted by hardware and software vendors. Furthermore, as hardware vendors retire processors and operating systems, enterprises that utilize these technologies are forced to undergo costly transformations within their IT infrastructure. When the future of an operating system and processor line is unclear, further investment in that technology is at risk.

Standards help ease these concerns. Standards may come from formal acceptance processes or through the de facto adoption of a technology by industry and the enterprise. They provide compatibility, interchangeability and stability, leading to open availability and a greater selection of products that should work together. The open availability of components, systems and solutions based on standards help companies deal with technology refresh. While it is understood that hardware and software advances dictate that IT infrastructure be refreshed, from an investment protection perspective the longer the existing assets can meet business needs but be effective in terms of availability, scalability, TCO, and more, the better. Standards and open availability enable IT infrastructures to be refreshed in an incremental fashion, avoiding the "big bang" approach, which is costly, complex and error prone.

Ensuring the IT infrastructure is open helps the enterprise avoid being locked into one particular hardware or software vendor. By creating an environment in which new hardware is compatibile with old systems, and software can be easily moved between existing resources, the enterprise supports flexibility and better

utilization. It also creates a low barrier to exit—the ease with which infrastructure can be moved from one vendor to another should the enterprise feel it is not happy with the efficiency, security or availability of a particular vendor.

Matching the Right System to the Right Job

Failure to utilize the right kind and size of systems can impact the effectiveness of an IT infrastructure and the bottom line. When systems are underutilized, a portion of the money spent on hardware acquisition is lost, as well as the opportunity costs that go unrealized when that lost money is not invested in another asset. When systems are overutilized, they have very little headroom and companies run the risk of running out of capacity at a critical time. This can have a devastating impact on the business if systems fail to meet services levels. As a result, companies should adopt the business engineering practice of matching the right resources to the job.

Partitioning the IT Infrastructure for Agility

One of the ways to create an agile IT infrastructure and get the most value out of hardware investments is to use a single system for multiple tasks. *Partitioning* is the ability to take a single system and divide it into multiple smaller systems, each separate from the other in terms of security and resource consumption. This ability to have multiple systems within a single server can help save valuable floor space, localize and simplify system control, and increase resource management and availability.

There are two types of partitioning. *Hard partitions* divide the computing system into flexible, fault isolated segments, enabling multiple copies of the operating system to run on a single server. In contrast, *soft partitions* create barriers in the operating environment, enabling multiple execution environments to exist within a single instance of the operating system. Using partitioning technology, IT organizations are better able to:

- *Create, delete and resize partitions based on workloads*
- *Isolate workloads and workgroups by putting them in their own partitions*
- Give applications dedicated hardware resources

Creating Workload Optimized Systems

No two organizations have the same kind of workload or use system resources in the same way. Some may utilize batch compute servers and application or database servers, while others may employ complex timesharing systems or implement online transaction processing. Different types of applications also place different demands on a computing platform. Some workloads are essentially serial in nature and do not benefit from multiprocessing. Others are floating-point intensive and require platforms that can support these requirements, while yet other processes require integer performance. Some applications have limited compute requirements but move huge amounts of data.

Regardless of the environment and application mix, a vast amount of potential computing capacity typically remains untapped. Users are continually searching for more computing resources to help solve problems, resulting in systems that are alternately over-loaded or under-utilized. At the same time, a lack of trust

hinders the sharing of computing resources—users are afraid their computing resources will be overtaken by other applications. System administrators must find ways to gain control and establish isolation mechanisms in order to improve resource utilization. With a finer granularity of control, administrators can ensure all workloads have access to an appropriate amount of resources, and that no workload consumes the entire system.

Consider two business units: human resources (HR) and accounting. HR supports queries by employees during the hours of 8:00 am and 6:00 pm, and contains sensitive employee information. The accounting department runs a nightime batch job to reconcile accounts. Rather than support two separate computing environments for both business units, organizations can support each application in a virtualized, partitioned environment that provides needed security for sensitive data. The resources used by each department can also be limited, ensuring the batch job can complete overnight and free the computing resources in time for the HR departments daily routine.

Providing Consistency Across Systems

The complexity and lifecycle management of many software infrastructures is an impediment to many businesses. Driving down these integration costs is important to many organizations. When all key software components are available on all systems, software delivery becomes more predictable. Furthermore, when these components are built into the operating environment, companies can focus on how to allocate services—not how to integrate them.

Using systems that provide the same functionality—even if those systems are based on disparate computing technologies—helps manage cost and complexity. When these tools are integrated directly into the platform, integration costs are reduced. Such a methodology identifies and transfers integration from the customer production environment to the vendor's front-end development process, reducing deployment risk and driving down total software ownership costs.

Planning for Change—Scalability

Even when systems are matched to the right job, business changes can result in a demand for more processing power. In an ideal world, changes to the IT infrastructure can be performed without downtime or added system complexity. Two techniques are available to help this effort: horizontal and vertical scalability.

- *Horizontal scaling*

  Horizontal scaling utilizes multiple servers in a tier. Changes to servers, and the addition of servers, can be deployed incrementally, enabling applications, services and updates to be replicated quickly to multiple systems. If a server ceases to function or is taken offline, the remaining servers in the horizontal set of servers continue to provide service to users. As a result, horizontal scaling provides resiliency and is ideally suited for applications that perform independent, stateless transactions, such as those serving up Web pages.

  Horizontal scaling is perhaps best illustrated by the concept of a grid. In such a deployment, many small processing units are networked together to create a larger entity. The aggregate can be a huge

deployment, with thousands of nodes providing significant compute capacity. While grid environments are trivial to scale (just add more processors) they can be considerably more complex to manage and maintain because of the number of operating system instances needed. Similarly, workload management can become an issue as the work must be decomposed into smaller units that can be dispatched to grid elements. Once the work of a single processor is complete, the results must be synchronized and integreated into the overall solution.

- *Vertical scaling*

  Some applications—particularly those with a dependency on a centralized resource, inter-process communication or synchronization—are better suited to run on a single machine in which the processors and memory are more tightly-coupled. For this class of applications, vertical scaling is more appropriate. With vertical scaling, resources are scaled within the system—CPUs,memory and storage are incrementally added to the server over time to increase scalability. Larger systems, or those with increased capacity, enable multiple applications to be consolidated onto a single server and affect better resource utilization. As demand grows, systems can be scaled to accommodate it. Vertically scaled solutions tend to be less complex, making them easier to manage and monitor.

Keeping Systems, Applications and Services Running—Availability

The availability of business- and mission-critical applications is an important factor for organizational success. Data centers are often required to be on-line 24 hours a day, with systems, applications and services able to handle peak loads without disruption or service degradation. In this kind of environment it is critical that system components not be allowed to fail and create an outage. This requires the detection of performance degradation, and the ability to reconfigure resources as needed. Component degredation must be flagged to the operating system and system administrators so the affected component can be taken offline. In addition, the system should support dynamic reconfiguration, enabling failed or failing components to be hot-swapped without bringing the system down for maintenance. This is particularly important in mission-critical environments, where systems cannot be shutdown to swap failed components.

Indeed, the overall reliability of the operating system is important when considering an IT investment. Platforms that suffer from performance degredation over time, or become unresponsive, often require a daily reboot as a matter of standard operating procedure. The overhead required for this sort of outage or maintenance schedule can quickly eliminate any savings that may have been realized by purchasing a proprietary operating system that runs on commodity hardware.

However, accidents and catastrophic componet failures are a fact of life. Truly mission-critical applications require highly available (HA) environments. High availability can be provided in several ways:

- *Fault tolerance*. Some vendors use a proprietary, costly architecture of redundant processors that compare the results of calculations to ensure all faults are identified and flaged.
- *Clustering*. Rather than providing redundancy at the processor level, clustering provides redundancy at the system or application level. When an application becomes unresponsive due to software, hardware

or environmental factors, the application can switch or failover to another system in the campus cluster. This form of redundancy not only ensures high availability, it also allows for the physical separation of compute environments, supporting disaster recovery.

Securing Information

Many enterprises protect their IT infrastructure investments with physical security for their computing environment, including security forces, card locks, and video monitoring. However, another great theat to security comes from connecting the IT infrastructure to an intranet or extranet. One of the biggest risks to any enterprise is the security of the information it uses and manages. The media is rife with stories about information loss or unauthorized access by either external hackers, internal employees or contractors who were able to gain unauthorized access to information due to an inadequate security policy or the use of software that provides administrative access to the operating system.

To help this effort, IT organizations can take advantage of several technologies that aim to help secure the IT infrastructure and the information it contains.

- *Role-based access controls*. These controls enable administrators to assign specific access rights to programs and commands for each user. This reduces the chance of administrative errors or accidental or malicious use of IT resources. When these controls can be centrally managed, organizations are better able to reduce costs and increase flexibility.
- *Stack buffer overflow protection*. Stack buffer overflow conditions enable many types of common attacks. System should be designed to be able to defend the system from such attacks, helping prevent malicious code from granting unauthorized access or permissions.

Achieving a High Level of Operational Management Capability

Deploying an IT infrastructure that protects investments is only the first step. Indeed, the work is not done after the execution architecture is built—IT organizations must be able to maintain it. Once an IT execution environment is defined, it must be monitored and manged to ensure service level agreements (SLAs) can be achieved, problems can be effectively and accurately reported, and change can be agreed upon and executed through a non-disruptive process. In order to achieve these goals, the people, processes, and tools of the organization must be coordinated.

The complexity of this process in a large, multi-national organization can be significant. While many organizations attempt to solve these problems using technology alone, experience shows this is often impossible. While organizations spend a great deal of money on technology to ensure applications are highly available, study after study shows the majority of outages are caused by operator (or human) intervention. Many companies struggle to manage the IT infrastructure in a way that balances the trade-offs between return on investment (ROI), accountability, predictability, reliability, productivity and real utility and functionality for users. This typically results in customer dissatisfaction and increasing cost as the IT organization spends more on resouces in an attempt to fix the symptom rather than the problem.

The use of a methodology to achieve higher levels of operations management allows the organization to move in an incremental, evolutionary fashion towards higher levels of availability, scalability and reliability in a manner that is aligned with existing business drivers. This in turn drives down costs and increases service levels. A significant amount of work has already been done to document and support these frameworks. Perhaps the most popular of these is the Information Technology Information Library (ITIL) created by the British Standards Institute (BSI). Other frameworks have extended the basic concepts and procedures defined by ITIL. The common thread in these frameworks addresses the interaction between the people who manage the environment (their roles, skills and capabilities), the tools they deploy to manage and monitor the environment (capacity management, fault detection), and the process used to maintain the infrastructure (error reporting and change management).

Taking a Look at How Technologies Help Protect IT Infrastructure Investments

Table 1 summarizes the technologies discussed and how they can impact a company's ability to protect IT infrastructure investments by minimizing risk, and increasing agility and efficiency.

Table 1. Protecting investments with technology

	Minimize Risk	Increase Agility	Increase Efficiency
Microprocessor Compatibility	•	•	
Open Standards	•	•	
Open Availability	•	•	
Longevity	•		•
Matching Systems to Jobs	•	•	•
Partitioning	•	•	•
Virtualizing the Environment	•	•	•
Workload Optimization		•	•
Decoupling Applications from Platforms	•	•	•
Consistency Across Systems	•	•	•
Scalability	•	•	•
Availability	•	•	
Security	•		
Operational Maturity	•		•

The UltraSPARC® and Solaris™ OS Platform—A Foundation For Investment Protection

Sun has designed its entire product line with virtualization in mind, resulting in a breadth of hardware and software products that can be used to create a more agile and efficient IT infrastructure that supports business goals and helps minimize business risk. At the foundation of this product line are UltraSPARC processor-based servers running the Solaris OS. This platform is tailored to meet a range of industry requirements, as well as computing and economic demands. As a result, Sun's servers range from general

purpose, entry-level servers, to massively scalable servers designed for mission-critical network computing, to NEBS certified, carrier-grade systems, to racks of systems designed to work together to create compute grids (Table 2). This broad product offering gives companies the flexibility to choose the right platform for the job at hand.

Table 2. Sun server family

Server Category	Description	Designed To...
Compute Grids	• Collection of entry-level servers • Management software	• Reduce deployment time • Ease scalability
Carrier-Grade Servers	• Carrier-grade systems • Open standards-based • NEBS Level 3 certified	• Accelerate time-to-market • Reduce development costs • Accelerate development of telecommunications services • Deliver high service levels
High End Servers	• Massively scalable • Highly available • Chip Multithreading technology • More throughput in a smaller footprint • 36 to 72 UltraSPARC® IV or IV+ processors • Execute up to 144 concurrent threads • Dynamic reconfiguration	• Support mission-critical network computing • Help businesses deliver innovative new applications with less risk • Deliver high availability • Deliver manageability
Mid-Range Servers	• 4 to 24 UltraSPARC processors • Execute up to 48 concurrent threads	• Deliver enterprise-class price/performance • Deliver high availability • Deliver manageability
Entry-Level Servers	• 1 to 4 UltraSPARC processors • Widely scalable • AMD Opteron™ processors also available	• Optimize price/performance

All Sun systems can run the Solaris Operating System (Solaris OS). The Solaris OS combines several computing elements—operating system, networking and user environment—into a foundation organizations can use to develop, deliver and manage business- and mission-critical computing solutions. The Solaris OS also includes several capabilities to help IT organizations create more effective environments and protect investments in both hardware and software. Support for binary compatibility and an Application Compatibility Guarantee combine with strong independent software vendor (ISV) support to ensure applications run on the systems of yesterday, today, and tomorrow. Innovative technologies, such as partitioning, Solaris™ Containers, and fine-grained resource management controls, give organizations the ability to get more out of available computing resources yet ensure isolation between critical—and sensitive—applications. In addition, the Solaris OS includes many security features, including role-based access control and stack buffer overflow handling, giving IT organizations an extra measure of confidence that systems, applications, and services are better protected from malicious attack.

Providing a Foundation for Virtualization

The processor forms the basis of a number of elements that comprise a computing environment. Programming at the machine instruction level is difficult and prone to error. To help this effort, the Solaris OS abstracts the hardware layer by providing a set of application programming interfaces (APIs) that simplify the programming process (Figure 2). Each API has a specific syntax and set of semantics associated with it that are well understood by application software developers focused on providing business-critical support. These APIs can be thought of as a set of contracts that are provided to developers.

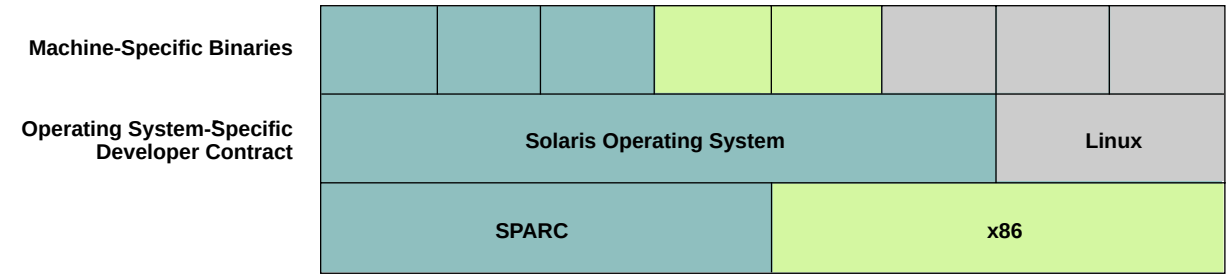

Figure 2. The Solaris OS provides a set of APIs to developers to help simplify the programming process

Operating systems change for two main reasons: the processor they abstract may provide new hardware features to improve application performance, or new algorithms may be developed to better virtualize or operate the hardware platform. Sun designed the Solaris OS such that these changes are completely transparent to the application developer, and consequently the application software. The APIs that virtualize the hardware platform are fixed, both in terms of the parameters they accept and return, as well as the semantics associated with their operation. New hardware features and improvements are either adapted to this static set of APIs, or new APIs are created to support new techonolgy. This ensures old applications can run unchanged but still take advantage of improvements in the underlying hardware and operating system (Figure 3). Introduced in the Solaris™ 7 Operating System, the Applications Binary Interface (ABI) ensures applications developed on older Solaris OS versions run—without change or recompilation—on newer versions of the Solaris OS.

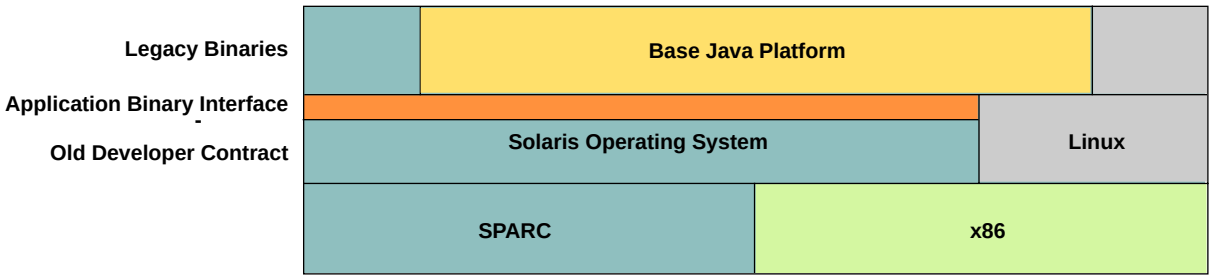

Figure 3. The Solaris OS includes an Application Binary Interface that ensures applications developed on older operating system versions run without modification on newer versions

This hardware and software compatibility has a significant impact on the risk of an IT investment. It means that an application binary that was developed, compiled, and executed on the the Solaris 7 OS for an UltraSPARC® II processor will run with out change or recompilation on an UltraSPARC® IV processor running the Solaris™ 10 Operating System. There is no need modify the code or recompile it. By decoupling the hardware, operating system, and applications, companies can now mix and match their processor base and operating system without having to change applications, substantially protecting software investments.

Build In Investment Protection with the UltraSPARC and Solaris OS Platform

The UltraSPARC and Solaris OS platform incorporates several technologies that can be used to build an IT infrastructure with an eye toward protecting investments in hardware and software, and providing value to the business over time.

Based on Microprocessor Compatibility

Based on SPARC® technology, Sun's UltraSPARC microprocessor family is binary compatible—from one processor generation to another. As a result, there is no need to port, migrate, or recompile applications when a new processor is released. Furthermore, there is no need to retrain system administrators, as the same architecture is used throughout the processor family.

Open Standards for Open Availability

Open standards ease the transition when moving IT environments between platforms from different vendors. While the benefits of standards are clear, the process by which they are created, and their variations and abundance, can make selecting the right ones hard. Sun knows their importance, and is engaged in helping to establish useful, broadly applicable standards that can be used to create open systems and benefit customers. Sun has been instrumental in developing many standards, including UNIX®, TCP/IP, X Windows, NFS, Ethernet, ATM, FDDI, ISDN, Java™ technology, and more. Sun's past and present efforts include leadership in the Fast Ethernet Alliance, the ATM Forum, CommerceNet, the Object Management Group, the IEEE, IETF, and several other standards organizations and consortia.

Supporting Agility Through Built-In Partitioning Technology

Sun's Dynamic System Domain technology can be used to physically partition systems and create a more agile data center. Dynamic System Domains enable a single system to be logically divided into multiple stand-alone domains, each running its own instance of the operating environment. Failures in one domain do not affect applications running in other domains, increasing availability and providing a reliable, secure platform for running multiple applications simultaneously. In addition, these partitions are electrically isolted from one another to help improve availability.

Solaris™ Containers—A Foundation of Virtualization

To effectively implement server virtualization, organizations must be able to manage applications independently, control resource utilization, isolate faults, and ensure security between multiple applications. Sun's approach to server virtualization centers on a technology called Solaris Containers. Solaris Containers isolate software applications or services using flexible, software-defined boundaries.

Applications can be managed independently, even while running in the same instance of the Solaris OS. Solaris Containers create an execution environment within a single instance of the Solaris OS and provide:

- *Full resource containment and control, enabling more predictable service levels*
- *Fault isolation, minimizing fault propagation and unplanned downtime*
- Security isolation, preventing unauthorized access and unintentional intrusions

Create Workload Optimized Systems With Built-In Technology

Optimizing systems for varying workloads requires advancements in both hardware and software. To help this effort, Sun supports several hardware and software technologies that work together to accommodate changing workload requirements:

- *Do more work in parallel through software implementation*

  To increase performance, applications can be architected to execute many tasks simultaneously. To make this possible, the Solaris OS provides support for concurrent processing, such as multiple threads, shared memory, and asynchronous I/O. At the application level, the Solaris OS user-threads library provides a mechanism for exploiting multitasking. Using user-level threads, applications can be decomposed into subtasks to take advantage of multiple processors. The Solaris OS provides POSIX and Solaris threads at three levels:

  – The kernel uses one thread to handle each system call and interrupt, allowing multiple CPUs to accelerate kernel tasks.
  – The user code to kernel interface uses multiple lightweight processes (LWPs)—a software representation of processors—allowing a single user process to have multiple outstanding system calls and be scheduled onto multiple CPUs.
  – A user mode thread library allows very lightweight thread switching and large numbers of threads, using LWPs as virtual processors to run threads.

  The Solaris OS allows an application to control how many threads and LWPs it needs to maximize performance and minimize kernel overhead. Threads are the primary interface for application parallelism. Multithreaded applications can be written without regard for the number of CPUs configured on the target machine—the thread library detects whether there are insufficient LWPs to run the program efficiently, and can automatically start more LWPs as needed.

- *Get more work done on the chip*

  Chip Multi-Threading (CMT) technology is a technique that focuses on how much work a processor gets done rather than pure clock speed. CMT integrates the power of symmetric multiprocessing (SMP) onto a single chip, allowing a single processor to execute several software threads simultaneously. Available on new Sun systems, CMT increases hardware density while simplifying overall design. As a result, CMT reduces complexity and increases efficiency. In addition, these new innovations can be rapidly adopted in the data center because of their backwards compatibility with previous UltraSPARC product offerings.

- *Accommodate changing workload conditions*

  While server virtualization enables enterprises to respond rapidly to change, it also helps increase IT effectiveness and efficiency. One of the main objections business units have to running under a shared services model is their belief that other business units will impact available compute resources. Frequently, they size their IT environment to accommodate peak loads, a scheme which can be very inefficient and expensive. Solaris Containers enables business units to create, size, and manage a virtualized IT environment that can be dynamically reconfigured to provide needed resources on demand. Because multiple independent, virtualized operating system instances appear as a single, ordinarily operating system instance, traditional system administrator skills are still relevant.

- *Dynamically adjust to seasonal and peak requirements*

  The Solaris OS enables system resources to be controlled through the creation of a *resource pool*, a collection of resources, such as CPUs, physical memory, or network I/O bandwidth, that are reserved for exclusive use by an application or set of applications. Resource pools enable system administrators to partition a system into a set of smaller virtual environments, each providing resources for a fixed workload consisting of one or more applications. These partitions provide fixed boundaries between workloads, ensuring each has access to a consistent set of resources regardless of resource usage on the rest of the machine. As a result, workloads can be separated to eliminate competition for resources, helping achieve predictable application and system performance, as well as user trust.

Decouple Applications from Platforms with Java™ Technology

The drawbacks of having an application tied to a particular machine are clear. The lack of flexibilty and difficulties associated with increasing the capacity and availability of that platform reduces the flexibility of the IT infrastructure. In an ideal world, an application would not even be tied to a particular operating system. If possible, the creation of such an environment would enable a truly low barrier to exit—IT organizations would be able to move the infrastructure to a new platform, or a combinatioin of any vendor's platforms, without excessive disruption or cost.

With the introduction of Java technology, Sun extended the decoupling of the application from not only different hardware and software provided by a single vendor, but from all operating systems on all processor architectures provided by all vendors that support the Java environment. Just as the traditional operating system provides an abstraction of the underlying processor, the Java™ Virtual Machine provides an abstraction of the operating system. This provides true investment protection as the application is not tied to any particular vendor product line. Applications written in the Java programming language run on any Java technology-compliant system, regardless of the operating system or processor architecture.

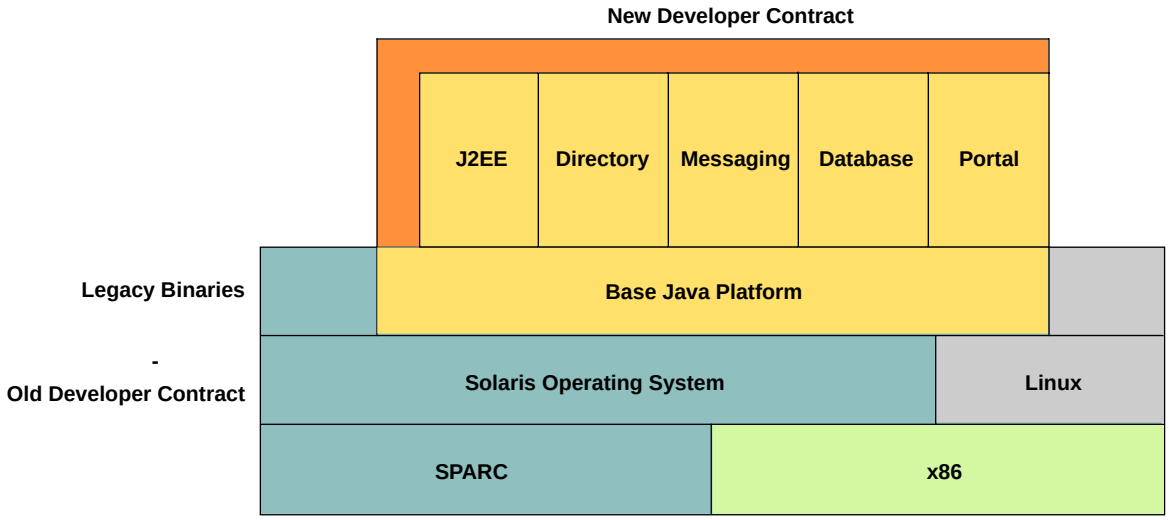

Figure 4. Java technology abstracts the operating system, enabling applications to run on any Java technology-compliant system

Gain Consistent Functionality Across Systems

As a result of the platform independence provided by Java technology, ISVs and large enterprises have invested significantly in the software design and implementation process. Why? They know they can reuse this software across vendor offerings. Software reuse removes the burden of reinventing the wheel for programmers. Rather than expending the time, effort, and expense of creating applications from scratch, developers can leverage a significant body of middleware that provides a number of proven implementations of common logic required for enterprise applications.

This middleware presents a new set of APIs to developers (Figure 4). It masks the complexity of operating system functionality with application servers, identity management, portal creation, database management, and other key capabilities. Using systems that provide the same functionality—even if those systems are based on disparate computing technologies—helps manage cost and complexity. However, getting middle ware from different vendors to interoperate can be a challenge as these middleware APIs have different semantics and signatures.

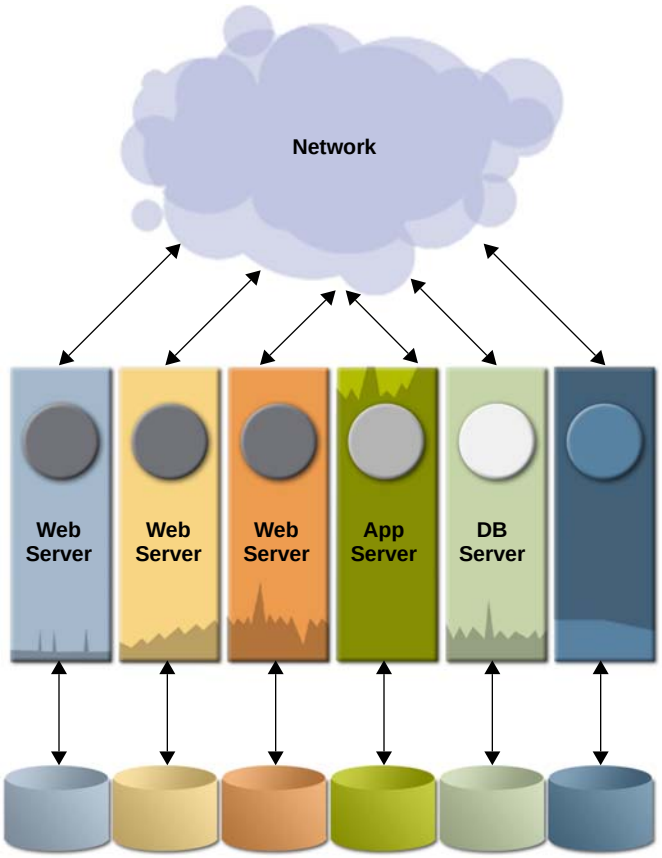

Figure 5. Middleware presents a new set of APIs and functionality to application developers

To further reduce the risk of application development and integration, Sun has implemented a complete Sun Java™ Enterprise System and integrated it into its platforms. These middleware components are fully integrated and tested to work together. This methodology identifies and transfers integration from the customer production environment to Sun's front-end development process, reducing deployment risk and driving down total software ownership costs.

Adapt to Change and Optimize Resource Utilization

Sun designed its systems to be able to accommodate planned and unplanned changes to the environment.

- *Dynamic Reconfiguration*

  Dynamic Reconfiguration (DR) provides the capability to reconfigure the operating environment without service interruption. Application execution continues uninterrupted when adding or removing components, and transaction and data integrity are guaranteed as in normal operation. Organizations can utilize dynamic reconfiguration to allocate additional resources on-the-fly during periods of peak demand, which can commonly occur at regular cyclic intervals. Dynamic reconfiguration works equally well for unexpected changes in demand, enabling administrators to respond and reallocate resources without interrupting system, application, or service operation.

- *Mixing and matching of boards*

  In Sun systems, boards can be mixed and matched regardless of processor speeds. One of the reasons clients choose to adopt a vertical scaling program is the ease with which resources can be added to a system. With Sun's Dynamic Reconfiguration capabilities, an existing chassis can be populated with new boards to scale machine resources. Unlike other vendors, Sun allows new generations of the processor (such as UltraSPARC IV or UltraSPARC® IV+) to be mixed with previous processor generations (such as UltraSPARC® III) in the same chassis. These new boards operate at their own clock speed and do not have to slow down to interoperate with exisiting IT assets, further protecting hardware investments.

- *Provisioning software*

  Software is often installed in an ad hoc fashion. At best, a runbook may exist that provides guidelines for software installation and configuration. However, these documents require updating and maintenance, and the actual installation is often still performed by an operator who has his own experiences and preferences. Whether you wish to repartition a large symmetric multiprocessing system to better manage changing workloads, or re-image several hundred blade servers that are providing horizontal scaling, Sun's N1™ Provisioning Service automates the process. This automation not only helps improve agility by reducing the amout of time required, it also improves reliability and availablity by eliminating any variance that may be introduced based on the preferences of the installer. As we will see in later discussions, the creation of a standard for software deployment and configruation reduces operation costs, which in turn helps protect that aspect of IT investments.

Secure Information

The base level security in the Solaris 10 OS delivers one of the highest levels of security found in any system. The Solaris 10 OS now provides several security features previously only found in Sun's military-grade Trusted Solaris OS. User and Process Rights Management work in conjunction with Solaris Containers to let organizations securely host a large number of applications and multiple customers on the same system. For example, the concept of isolated containers is implemented based on process rights inheritance, limiting access to what a process or user can do on a system. Furthermore, when operating inside a Solaris Container, access to raw devices (DLPI layer) and file systems (`/proc`) is restricted. In addition, security administrators can minimize and harden the Solaris OS to help create a secure foundation for deploying services. Additional extended security features include authentication, data integrity, data privacy, and single sign-on capabilities so that tampering, snooping, and eavesdropping do not compromise data or associated transactions. Integrity of systems can also be protected using Solaris Secure Execution for file and data verification purposes.

Customers needing even higher levels of security can take advantage of the Trusted Solaris operating system. Trusted Solaris has been used for many years in industries with sensitive information, including government, intelligence, and security agencies, as well as finance, healthcare, and retail where high security and auditing capabilities are a must.

Achieve Operational Management Capabilities

The final step in the quality of any solution is putting in place the operational management practices to meet service level agreements. Indeed, disciplined operations management is critical to service quality and availability. This becomes apparent as organizations strive to handle conflicting business requirements, such as increasing sevice levels while reducing costs, or doing more with less resources. However, improving service quality for its own sake—without tying it to a specific business driver—can increase complexity and reduce the value of IT assets. Deploying new technology simply for technology's sake is not cost-effective. Similarly, minimizing costs without understanding the impact these changes will have on existing and future business services and processes can lead to lost opportunities and service outages.

To address these issues, the SunTone Service Excellence Model helps organizations focus on a key principal—optimizing IT resources and infrastructure to bring the appropriate level of quality to contracted services. It provides best practices for IT service level, availability, capacity, and data center management, and provides guidance for both service architecture and continuity (disaster recovery) services. Adopting this methodology can help customers streamline operational practices and drive efficiencies up front, while addressing people and process requirements to meet on-going service delivery and management quality goals, such as availability, reliability, response times, and security. SunTone helps companies:

- Define and meet service level agreements
- Utilize existing IT resources more effectively through the adoption of streamlined processes
- Build and maintain staff competencies to pre-empt people and process errors that result in unplanned downtime
- Enhance service security, performance, manageability, and disaster recovery capabilities

The SunTone[SM] Service Excellence Program consists of several key components, including a SunTone Service Delivery Specification, SunTone quality enablers and reference architectures, and certification audits.

- *The SunTone Service Delivery Specification*

  The SunTone Service Delivery Specification provides a benchmark to measure quality service delivery and management. The SunTone specification builds on and complements the frameworks of other industry standards such as ITIL, SysTrust, COBIT and ISO20000 (formerly BS15000). It provides a comprehensive technical guide covering 12 assessment areas under the categories of service delivery and service support that includes: service level management, service architecture, availability management, capacity management, security management, service continuity, data center management, configuration management, service desk, incident and problem management, change management, and release management.

  Backed by an open community process and years of practical industry experience, the SunTone Service Excellence Model and the underlying SunTone Service Delivery Specification provide a comprehensive listing of service delivery and management requirements, including architecture, policies, operations, and procedures. The specification is the result of strong industry collaboration led by the SunTone Council, a group of subject matter experts and senior technical experts from Sun, and its customers and

partners. Far from static, the SunTone Service Delivery Specification is updated frequently to reflect real-world best practices from some of the world's largest IT organizations.

- *SunTone Quality Enablers and Reference Architectures*

  SunTone Quality Enablers are pre-qualified tools and offerings for ensuring SunTone quality standards throughout the service delivery lifecycle. These pre-qualified building blocks—reference architectures, consulting and managed services, as well as SunTone certified applications, allow for easy adoption and implementation while helping to reduce complexity, costs, and risks of new solutions.

- *The SunTone Certification Audit*

  Through an assessment process and on-site audit provided by Client Solutions or an authorized partner, SunTone certification validates the quality of ongoing service delivery and management practices. It also verifies the solution's ability to meet customer SLAs, whether operated and managed in-house or outsourced.

Operational management capabilities are achieved by understanding the relationship between the people, processes, and tools that comprise the IT management infrastructure. By defining service levels, designing clear management processes, engaging trained people who can implement these processes, and ensuring tools are deployed to monitor infrastructure performance against the service levels, IT management can improve the organization in an iterative fashion (Figure 6).

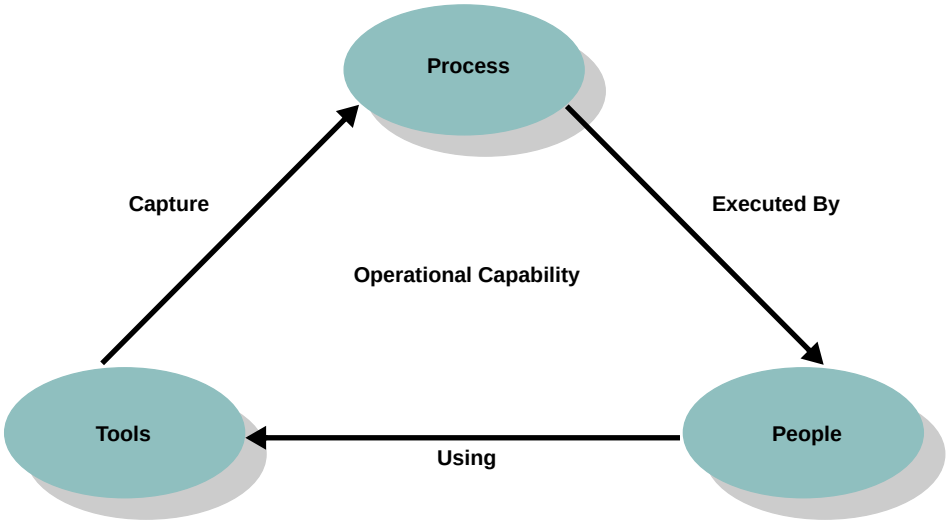

Figure 6. Sun's operational management model focuses on people, processes, and tools

The SunTone Service Excellence Program aims to tie IT infrastructure creation and management to underlying business process requirements. Using an evolutionary approach, SunTone enables organizations to define the level of governance they wish to achieve and move towards that goal. Instead of operating in a reactive, crisis-control fashion, IT organizations can choose to manage the service at the component, operations, or service level based on business requirements and the processes and tools deployed (Figure 7).

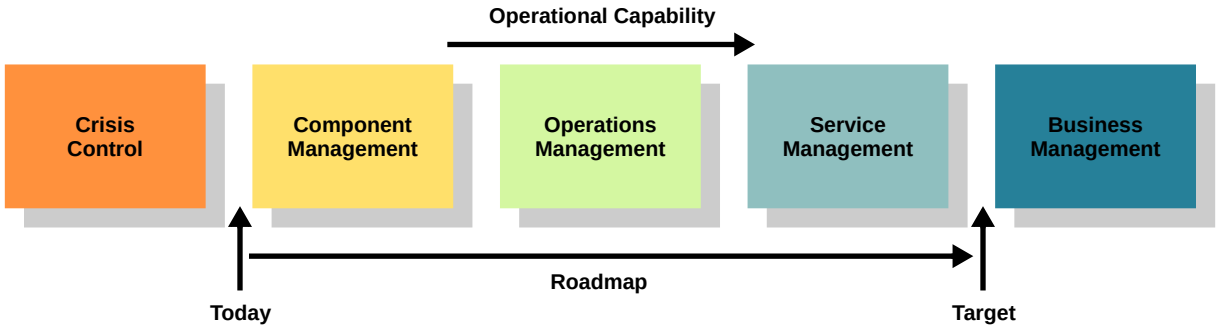

Figure 7. Sun uses a model that targets key areas for improvement

The adoption of a process like that used in the SunTone methodology ensures enterprises are better able to manage and grow the IT environment in a predictable fashion. Unlike investments in some of the hardware and software components of the IT infrastructure previously discussed, investments in management methodologies does not depreciate over time. In fact, the opposite is true—an investment in the people, processes, and tools required to implement a management framework only serves to improve operational capabilities and introduce efficiencies into the IT infrastructure.

Create A Service-Optimized Data Center

The beginning of this article outlined a number of problems and realities that conspire to minimize the protection afforded to IT investments, and subsequently discussed a number of innovations Sun has introduced to enable the creation of an IT infrastructure with built-in investment protection. This section introduces a framework that pulls these innovations together to enable an infrastructure that is flexible enough to meet business demands, yet support the reuse of the capital assets most business loathe to replenish, as well as new hardware and software technologies that are just over the horizon.

A *service-optimized data center* is an evolving deployment of process and technology that allows IT organizations to create a data center that is aligned with business goals, IT drivers, and technology roadmaps in such a way that it increases IT agility and maximizes IT efficiency while minimizing risk. It includes both the technical architectures used to create and reconfigure services, as well as IT management services to manage, maintain, and evolve the infrastructure.

Increasing agility helps enterprises respond to changing markets and business opportunities without having to recreate the infrastructure. IT capabilities can become a competitive advantage, extending the shelf life of assets while supporting rapid service deployment with minimal disruption. Virtualization of the environment increases flexibility, allowing the enterprise to rapidly respond to different workloads and keep utilization high throughout the entire organization rather than just a business unit, increasing IT efficiency. Finally, the ability to rapidly and efficiently reconfigure the enterprise in a predictable fashion, using a methodology focused on the alignment of business processes and IT infrastructure, can significantly reduce risk, which in turn helps protect IT resource investments.

One way to build a service-optimized data center is to use a standardized operating environment as the basis of the IT infrastructure. Standardized offerings reduce variability and configuation complexity, increasing efficiency and reducing cost. By having fewer different elements to manage, provision, and

secure, the IT environment can reduce costs while increasing availability and security. Sun has created a number off tools and technologies that automate the creation, distribution, and provisiong of software on Sun and heterogenous environments, including Live Upgrade, JumpStart, JET, N1 SPS, and more. It is this automation that supports agility and efficiency.

Once standardized operating environments are specified, attention can turn to how they are implemented or instantiated. The underlying compability of Sun's hardware and software offerings ensures old software can run on new plaforms, and that new plaforms are compatible with old software. This dramatically reduces the complexity of data center reconfiguration, and extends the life of the existing IT infrastructure. Regardless of the workload, Sun's broad range of products ensures the correct compute platform can be found for the job. Sun offers a wide variety of system products to scale horizontally (blade, x64, UltraSPARC), diagonally (four to eight processor-based systems) or vertically (massivelly parallel SMP UltraSPARC systems.) The advent of improved capabilities within Sun systems featuring the AMD Opteron™ processor, as well as emerging UltraSPARC processors with CMT technology, ensures this flexibility will continue to exist within the product line. A building block for Solaris Containers, resource pools ensure applications are isolated from one another both in terms of information privacy and competition for resources, such as CPUs, network bandwith, and disk space. By deploying the correct tools and implementing appropriate management processes with trained people, the IT infrastructure can be managed at the service or business level rather than lurching from one crisis to another.

Once standardized operating environments are deployed in a dynamic, virtualized environment, they can be used to create the services the business delivers. These services can be optimized, based on the their individual characteristics. Typically, services fall into one of the following network service architectures:

- *Service-Oriented Architecture (SOA)*. By decomposing a service in reuable components, business service flexibility is increased while time-to-market is reduced. By deploying proven service components that implement business processes, organizations protect investments in that technology. Sun's approach to a service-oriented architecture—Pragmatic SOA—emphasizes incremental projects, quick return on investment, and a phased build process. Reusing existing software components reduces the likelihood of errors and eliminates the need to rip out and replace components, minimizing the disruption created by wholesale change and reducing IT investment risk.
- *Service Delivery Networks*. This is can be thought of as creating a networking fabric within the enterprise. Just as Sun has virtualized the compute platform, operating system, and application, the network can also be virtualized. By partitioning network traffic into secure enclaves based on the networking protocol used, Sun encourages higher utilization of existing resouces along with ease of reconfiguration, creating a client *service* as opposed to a client-server model.
- *Shared Network Services*. Standardized services, such as identity, LDAP, and password services, can be created that leverage economies of scale. Rather than implemeting these services on a per application or business unit level, they can be created and managed at an enterprise level. Ideally, these services would be available on a utility basis.

Investment in a data center that implements these concepts will appreciate over time as more operational efficiencies are realized through the reuse of software, infrastructure, and hardware components.

What Will the Future Bring?

Imagine not having to worry about protecting IT investments—because you do not have any. Imagine a scenario in which the IT services required are provided by someone else's IT infrastructure, and instead of focusing on patch management and resource utilization, your IT staff is able to focus on the value-added services that improve business processes.

This is exactly what Sun is proposing with its subscription based utility computing model. Organizations can purchase common services, such as HTTP, identity, LDAP, and portal services, from a service provider! In such a model, organizations do not have to worry about the type of compute platform or operating system on which their applications run. Instead, enterprises provide the content, and the service is delivered per a contract! The capital costs—and hence your investment—is considerably less than in a traditional IT service deployment environment. This is analagous to purchasing power or water from the utility company. Organizations do not have to worry about electric generators, septic systems, or water quality and testing because the utility company leverages the economy of scale and does all of this for you. However, for all of this to work, it is critical that people create open standards that are rigorously followed. Imagine the difficulties you would have connecting to the electricity grid if your applicance used 150 volts or required an eightpronged plug. Openness and standardization, two policies in which Sun is an acknowledged leader, leads to continued, best-of-breed improvement and ubiquitous adoption.

While we are not there yet when it comes to some of the more advanced services, Sun's concept of utility pricing for CPU ($1/cpu/hour) and storage ($1/GB/day) resources are leading to not having to worry about investment protection at all. Sun is also moving toward a subscription model for middleware products. Today, the Java Enterprise System (JES) is also available for a per-user licensing fee.

Conclusion

This article introduced the concept of investment protection as it applies to IT infrastructure, illustrating there is more to an IT infrastructure than just the compute platforms, software, and networking resources found in the execution architecture of the typical enterprise stack. IT infrastructure extends from the definition of business processes and determination of critical to quality (CTQ) metrics to measure performance, down to the people, processes and tools used to maintain and manage the data center.

Investments in hardware resources will always depreciate—the fast pace of innovation ensures faster processors that consume less power and are easier to cool will always be right around the corner. The issue is how to ensure that compute resouces, once acquired, have a meaningful shelf life. Investment protection is achieved by adhering to open, extensible architectures built out of reusable components. Many vendors require that a new versions of the operating system and application software be installed when a new or upgraded processor is added to the product line. The complexity associated with maintaining multiple versions of the operating system and software stack can make it prohibitive to keep old hardware around, especially after a newly released product takes hold in the marketplace.

Sun helps enterprises protect capital IT investments by ensuring application software is reusable when a new version of the operating system is installed. Hardware assets within the data center can be refreshed and operating systems upgraded without disturbing the software application base. Additionally, Sun has

defined a number of service-based initiatives that focus on data center operation and optimization. These initiatives are designed to help increase and protect equity within the IT infrastructure. Adopting an extensible data center design and a forward-looking operations management methodology by investing in the right kind of staff training helps significantly reduce costs. Table 3 summarizes the technologies available from Sun, and describes how they can be used to create an infrastructure with built-in investment protection.

Table 3. Protecting investments with Sun products technology

		Minimize Risk	Increase Agility	Increase Efficiency
Microprocessor Compatibility	• UltraSPARC Technology	•	•	
Open Standards	• Adherence to Standards	•	•	
Open Availability and Longevity	• Adherence to Standards	•	•	
Matching the Right Systems to the Right Job	• Broad, Flexible Product Line	•	•	•
Partitioning the IT Infrastructure for Agility	• Dynamic System Domains	•	•	•
	• Solaris Containers	•	•	•
Virtualizing the Environment	• Resource Pools	•	•	•
	• Solaris Containers	•	•	•
	• Application Binary Interface (ABI)	•		•
	• Application Programming Interfaces (APIs)	•		•
Creating Workload Optimized Systems	• Chip Multithreading Technology			•
	• Solaris Operating System	•	•	•
	• Threading			•
Adapting to Change and Better Utilizing Resources	• Dynamic Reconfiguration	•	•	•
	• Flexible Product Line	•	•	•
	• Mix-and-Match Boards	•	•	•
Decoupling Applications from Platforms	• Java Technology	•	•	•
Consistency Across Systems	• Java Enterprise System	•	•	•
	• ISV Adoption	•		•
Securing Information	• Solaris 10 OS	•		
	• Trusted Solaris	•		
Operational Maturity	• SunTone	•	•	•

About the Author

Brian Down is currently a member of Sun's Global Data Center Practice. Previously, Brian served as CTO for the Enterprise Consolidation and Migration sub-practice within the Global CSO organization. Before that, Brian was the Chief Architect for Enterprise Migration for the Americas Data Center Practice, and was a Senior Staff Engineer in the Americas Customer Engineering (ACE) and OPCOM groups. Prior to joining Sun in 1998, Brian held senior positions within a number of technology start-up firms, and was a research associate with the Department of Computer Science at the University of Toronto. Brian is also an author of *Migrating to the Solaris Operating System: The Discipline of UNIX-to-UNIX Migrations*.

References

Collier-Brown, David "Creating Self-Balancing Solutions with Solaris Containers," *Sun BluePrints OnLine*, June 2005. To access this article online, go to:
`http://www.sun.com/blueprints/0605/819-2888.pdf`

Lageman, Menno "Solaris Containers—What They Are and How to Use Them," *Sun BluePrints OnLine*, May 2005. To access this article online, go to:
`http://www.sun.com/blueprints/0505/819-2679.pdf`

Ken Pepple, Brian Down and David Levy, *Migrating to the Solaris Operating System: The Discipline of UNIX-to-UNIX Migrations*, ISBN 0-13-150263-8. To order this book online, go to:
`http://www.sun.com/books/blueprints.series.html`

Related Resources

Consolidation through Virtualization with Sun's x64 Servers:
`http://www.sun.com/amd/briefs/consolidation-sol-bf.pdf`

Operations Management Capability Model (OMCM) Consultation Services:
`http://www.sun.com/products-n-solutions/sodc/SODCwp.pdf`

SODC:
`http://www.sun.com/products-n-solutions/sodc/`

Sun's Software Compatibility Between Processors and Operating Systems:
`http://www.sun.com/software/solaris/programs/abi/documentation/index.xml`

Throughput Computing:
`http://www.sun.com/processors/throughput/`

Ordering Sun Documents

The SunDocs℠ program provides more than 250 manuals from Sun Microsystems, Inc. If you live in the United States, Canada, Europe, or Japan, you can purchase documentation sets or individual manuals through this program.

Accessing Sun Documentation Online

The `docs.sun.com` web site enables you to access Sun technical documentation online. You can browse the `docs.sun.com` archive or search for a specific book title or subject. The URL is `http://docs.sun.com/`

To reference Sun BluePrints OnLine articles, visit the Sun BluePrints OnLine Web site at: `http://www.sun.com/blueprints/online.html`

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-650-960-1300 or 1-800-555-9SUN **Web** sun.com