Linux 2.6 and Solaris 10

An Analysis of Two Strategies for Enterprise Operating Systems

By Evan Bauer
Principal Research Fellow
Robert Frances Group

February 2005

IBM Corp. and Red Hat, Inc. sponsored this study and analysis. This document exclusively reflects the analysis and opinions of the Robert Frances Group (RFG) author, who has final control of its content.

22 Crescent Road
Westport, CT 06880
http://www.rfgonline.com

Table of Contents

Executive Summary

With the release of the Solaris 10 operating system, Sun Microsystems, Inc. has made a fundamental decision to re-position Solaris as an open source alternative to Linux and to invest in Solaris on x86-based platforms. This raises a variety of questions and choices for enterprises with large Unix infrastructure investments and workloads. IT executives need to evaluate their organizations' operating system strategy, and ask the crucial question – is it better for an IT architecture to be derived from work controlled by a single corporation or by a community?

This question is of particular interest to three classes of organizations:
1. Those organizations with large investments in Sun's SPARC hardware and the Solaris operating system, many of whom have been making or contemplating a strategic move towards Linux and away from Sun.
2. Organizations without a significant Sun investment selecting a strategic operating system platform for enterprise applications and workloads.
3. Organizations with a large number of servers deployed planning to consolidate to a smaller number of more powerful systems.

RFG examined this question across three dimensions:
1. Vendor and community strategies and interests
2. Comparative capabilities of the hardware/software platforms
3. User experiences in the Linux and Sun Solaris worlds

RFG found that Solaris 10 is a significant improvement over previous Solaris releases and that Solaris remains a market-leading Unix offering. However, RFG continues to believe that Linux, not Unix, is the compelling option for the majority of future investments.

Research Premise: Comparing Solaris to Linux and Not to a Single Distribution

Among the possible approaches RFG could take to this subject would be to compare Solaris, the product of a single company, to a single vendor's distribution of Linux. RFG chose to compare Solaris with Linux 2.6 as the basis for our analysis because it matches how users and independent software vendors (ISVs) initially assess platform decisions. In the vast majority of user inquiries RFG has received on this subject, our clients make a Solaris versus Linux decision first, and then select one or more distributions to deploy and support Linux within their organizations. Unlike the various Unix operating systems that fragmented into related but incompatible product lines, the Linux community has kept from forking the kernel and requires a very high degree of compatibility between distributions that can legally call themselves Linux. One result of the approach taken by RFG, however, is that the specific strengths and weaknesses of any particular Linux distribution fall outside the scope of this analysis.

Factors driving the rapid adoption of Linux for a growing range of workloads and applications include the unanimity of positive results and experiences of enterprises across industries and the ubiquity of Linux expertise. In addition, the wide range of contributors to not only the kernel but also the ecosystem around Linux and the leading and constantly improving economics of Linux-based platforms have driven Linux adoption.

Enterprises evaluating and selecting alternative operating systems to meet general purpose application and infrastructure requirements need to understand the capabilities provided by the operating system environments and the business strategies and organizations that manage them. These factors contribute not only to the current technical capabilities, but also to the directional options, integration, quality, and support that IT departments will have for future implementations. All of these elements affect the cost, quality, and risk of delivering information services to the enterprise. The change in Sun's strategy for Solaris provides an opportunity to compare these two enterprise operating system offerings, as well as the environments, implications, and strategies that create, support, and come with them.

IT management needs to look beyond the marketing messages and feature lists to understand what the choice of Solaris or Linux means for the enterprise. RFG believes that a better Solaris has long been needed by Sun's installed base.

Solaris 10 offers many technical enhancements and addresses long-standing weaknesses, and is being offered under a new licensing and pricing model.

Over the last six years, RFG has conducted primary research on the uses and adoption of Linux and open source technologies within the enterprise. In a late-2004 RFG survey of leading financial institutions, Linux deployment and migration appeared on the radar screen of major IT initiatives. Roughly 25 percent of the companies interviewed had significant budget allocations for Linux. While trailing other initiatives such as compliance and application development, Linux has become an important component of today's computing infrastructure, and was the only technology called out by name among the initiatives cited. In a different study, users across industries surveyed by RFG in Q4 2004 and Q1 2005 see the current 2.6 version of the Linux kernel capabilities as meeting the needs of most organizations for a wide range of workloads.

In the last few years, there has been a significant shift in the enterprise IT executive's perception of the viability and applicability of Linux. Most compellingly, over the long term, is that the extraordinary diverse global community that creates, enables, enhances, and builds upon Linux is greater than the resources and innovative potential of any single corporation.

Enterprise Operating System Context

The selection of an operating system for most major enterprises is not made in a "green field" setting. Selection needs to be seen in the combined context of the following:

- An organization's hardware and software installed base.
- IT direction and future requirements of the business and the underlying technical resources needed to support that business.
- IT direction and opportunities for control points of the firm's infrastructure and application architecture.
- Capabilities and resources of the vendors and communities needed to support the technical elements of that architecture.
- Costs, opportunities, and risks associated with architecture change.
- Strength and commitment of original equipment manufacturer (OEM) partnerships, the ISV ecosystem, and the level of maturity of open source communities.

Within this context, most leading firms are in the process of moving away from proprietary, tightly linked, hardware and software platforms to the combination of hardware-portable operating systems running on a combination of specialized and commodity hardware platforms. Microsoft Corp. Windows and Linux are the only two operating system platforms showing substantial growth since CY 2000. During that period, spending on Linux installations grew more than 350 percent, while spending on proprietary Unix servers shrank by more than 18 percent. This is particularly true in the market of the under $100,000 per server segment, where both shifts are magnified by another 50 percent.[1]

Price, Performance, and Price-Performance

Lower price, price performance, and the lack of dependence on any single vendor are the most frequently cited reasons for this shift from proprietary Unix to Linux.

The consolidation and migration of server platforms has been moving major technology consumers away from proprietary Unix platforms towards Linux on commodity chip sets (chip designs used in applications for the consumer, embedded systems, and business markets). This is primarily 32-bit x86, but increasingly includes the Advanced Micro Devices, Inc. (AMD) 64-bit x86 Opteron, IBM's POWER-based systems, and Intel Corp.'s EM64T. Sun's Q1 2005 release of the new Solaris 10 Unix operating system for SPARC x86, and Opteron is being positioned as a revolutionary product and an alternative open source operating system. Sun released the operating system under a newly created open source license.

Over the last six months, RFG has conducted in-depth interviews of more than a dozen major Sun customers. Most are uncertain of the long-term impact of this pricing change and Sun's Enterprise Java package offerings on their overall IT spend. Users are evaluating carefully the near-term moves Sun will be making in pricing.

At the same time, for those customers that chose to keep their applications on a Unix platform, Sun and Solaris remained their preferred investment choice.

[1] Summary of Industry Analyst Surveys by CSFB Equity Research, 2004.

The Operating System Decision

Operating systems are only of interest because they enable the enterprise to run the applications it needs. The characteristics of the operating system, however, can greatly affect the behavior of those applications and the costs and risks associated with the organization‘s technical infrastructure.

The following factors are involved in the enabling and controlling role of the operating system within the enterprise architecture, and are (at least in part) a function of the operating system that is adopted.

- Compiler and Performance
- Reliability, Availability, and Serviceability
- Security
- Hardware Support
- Operating System Support
- ISV Support
- Licensing
- Standards, Openness, and Control

Although no one needs an operating system by itself, the decision impacts the way in which the firm will execute its applications and manage its data. In addition, it will impact from whom and for what cost it will purchase its hardware, the cost and quality of system support, and who controls the future of that investment. As we compare Solaris 10 with Linux 2.6, we do so using this framework.

Sun‘s Position in Today‘s Market

Sun has been the dominant Unix-based server supplier over the last five years, particularly within the telecommunications and financial services industries – the two largest technology-consuming sectors in the global economy. RFG user surveys conducted in 2003 and 2004 found that the actual price performance of a Linux-based server was between a two- and nine-times improvement over that provided by Solaris on SPARC systems. As a result of the price-performance gap, Sun has been losing both its market share and revenue to Linux and Windows. This loss of market share reduces both Sun‘s profitability and its ability to control the direction of the enterprise software market as the ubiquitous platform of choice.

Sun has instituted a number of changes to its strategy over the past several years. A large portion of Sun‘s intellectual property (IP) is in software: Forte (now Sun Studio) compilers and development tools, Java, N1/iPlanet (the former Netscape server products), and the StarOffice suite. Sun‘s Solaris-based SunRay thin-client environment has been marketed as an alternative to the enterprise desktop PC. None of these, however, has produced revenue streams or profitability to replace Sun‘s dominant results during the era when its marketing position was that of the ”dot in .com.” Sun still has $7 billion in cash, short-term investments, inventory, and accounts receivable, and has demonstrated a return to break-even performance in the last two quarters.

SPARC, by not gaining traction outside of the server market, does not have the economic advantages of commodity production volumes to offset the continually rising and expensive R&D and fabrication costs associated with each new generation of chip technology. RFG sees a downward spiral in the SPARC chip technology, with performance, power consumption, and heat generation all being issues. Whether Sun‘s quad-core Niagara chip (scheduled for release in 2006) can break that trend remains to be seen. A plus for Sun has been a strategic relationship with AMD, making it the number one OEM of Opteron-based servers in 2004. In addition, Fujitsu Ltd remains a partner with Sun in its SPARC server business.

The Linux Community and Marketplace

The Linux operating system is the product of a community. Originated and still overseen by Finnish computer scientist Linus Torvalds, the Linux operating environment now counts Computer Associates International, Inc. (CA), Dell Inc., Hewlett-Packard Co. (HP), IBM Corp., Novell Inc., Oracle Corp., and Red Hat among its major enterprise contributors. In addition, the European Union (through grants to universities and research institutes), the GNU Project, and the U.S. National Security Agency (NSA) contribute to Linux's development. Many individuals are working on not just the kernel but also hundreds of open source projects (and an increasing number of commercial products) that provide the complete operating environment. Linux is developed within an operating system model that is noted for being a meritocracy. Linux developers are only as good as the code they contribute, the best ideas and code are adopted and extended; problems and opportunities lead to new projects and products, and competitors co-exist for the long haul.

One example of this diverse meritocracy is the range of available distributions (installation packages with update mechanisms and surrounding software) for Linux. Linux distributions prosper with a variety of approaches to the commercialization of Linux. More than 95 percent of the RFG survey respondents (primarily global 1000 companies, government agencies, and their suppliers in North America) make use of distributions from one or both of Red Hat or Novell, with more than 70 percent running a distribution from Red Hat. These, however, are not the only choices for the Linux user. France's Mandrakesoft and Xandros Inc. (the former Corel distribution) focus on the desktop. Debian and Gentoo remain entirely non-commercial, and Asianux, Conectiva, and Turbolinux, Inc. are just a few of the many regional distributions that enjoy government and commercial support in Asia and South America. Linux has benefited from the sharing of IP under the General Public License (GPL) and constant improvement from the largest group of contributors associated with any operating system. The Linux Standards Base (LSB) provides for compatibility of application software across distributions. It is RFG's experiences that user applications developed on one distribution will run without modification on others.

RFG's in-depth survey findings are confirmed by other independent studies with larger samples. In the Ziff-Davis 2004 CIO Insight Vendor Value Research Study of 1,050 IT executives, Red Hat took the number one position, with five other global hardware and software vendors with major Linux strategies all ranking higher than Sun and Fujitsu (at positions 19 and 18 respectively)[2]. Despite marketing messaging that positions a Sun and Solaris platform versus a Red Hat on Intel platform, users are faced with a choice between product sets with very different types of sponsors. In a distinct difference, from reliance on the financial condition, intentions, and marketing strategy of a single firm, the future of Linux adopters comes from the strengths and options provided by a community.

Licensing and Open Source

Sun has announced the release of the Solaris source code under a newly created license structure, as a fulfillment of its commitment to "open source Solaris." In understanding and evaluating Sun's challenge to Linux's role in the technology ecosystem, it is necessary to understand the licensing underpinnings of both Solaris 10 and Linux. The open source development process is very analogous to that of the scientific community. Researchers do not keep findings secret, but publish their results for everyone to verify, use, and extend, continually increasing knowledge and making possible new discoveries.[3]

Linux and the GPL

Linux has always been released under the Free Software Foundation's General Public License[4] (the GPL) – a licensing strategy that has been a key element in Linux's extraordinary ongoing extension and speed of innovation that has driven its unprecedented adoption. Software licensed under the GPL is "contributed," and the copyright stays with the developer. However, the copyright is available for inclusion, use, and redistribution as part of derivative works, as long as those products do not restrict anyone else's future use and redistribution of that software and preserve the original

[2] CIO Insight Magazine, December 2004, pages 53-56.

[3] For a discussion of the GPL and other open source licensing schemes, see the Open Source Initiative website at http://www.opensource.org/ particularly Bruce Perens' Open Source definition at http://www.opensource.org/docs/definition.php

[4] http://www.gnu.org/licenses/licenses.html

copyright. The result of GPL licensing has been the encouragement of the best new ideas and approaches in software to both come to and to move out from the Linux code base. The GPL is not the only license used by the open source community, but it is the most widely adopted. An intended result of GPL licensing is if the source code is distributed that the best new ideas and code are available to all participants in a community of developers who are building projects that benefit all individuals and organizations.

Open Source Licensing and Solaris 10

Sun has taken the unusual step of creating a new open source license for Solaris 10. The Common Development and Distribution License (CDDL) is a new license designed by Sun based on the Mozilla Public License (MPL). The reaction of the open source community to Sun's announcement has been muted. The issues and concerns expressed to date with Sun's approach to the licensing of Solaris include the following:

1. The CDDL Contributor's Agreement has yet to be released, and the responsibilities and indemnifications associated with it have not been defined. Without it, there is no legal basis for developers not employed by Sun to participate in the development of Solaris.
2. Sun's patent cross-licensing agreement with Microsoft protects Sun from litigation by Microsoft, but the provisions that would have protected other contributors to Solaris had the MPL been used, have been removed from the derivative CDDL. Sun is protected from Microsoft litigation, while other contributors are not. Sun has announced the release of 1,600 of its patents into the open source community, but apparently for Solaris and CDDL projects only, with Sun retaining the right to sue participants in other projects.
3. All code released by Sun under the CDDL is precluded from inclusion in any GPL-licensed projects[5]. In a letter accompanying the submission of the license to the Open Source Initiative, Sun stated "like the Mozilla Public License, the CDDL is not expected to be compatible with the GPL, since it contains requirements that are not in the GPL. Thus, it is likely that files released under the CDDL will not be able to be combined with files released under the GPL to create a larger program."[6]
4. RFG contacted the Free Software Foundation, the maintainers of the GPL, for their understanding of the relationship between the two licenses. According to the Free Software Foundation, "This is a free software license which is not a strong copyleft; it has some complex restrictions that make it incompatible with the GNU GPL. That is, a module covered by the GPL and a module covered by the CDDL cannot legally be linked together."

The net result of this is to separate Solaris from the larger open source community that supports Linux and other open source operating system projects. Sun does participate in many open source projects and includes over 180 open source packages in its operating system release, making Sun part of a larger open source community in peripheral areas but not for the core operating system.

RFG believes that unless Sun's licensing approach changes, innovation in Solaris will continue to be almost fully dependent on the efforts, resources, and interests of Sun and Fujitsu. Solaris' larger user community will have the opportunity for greater study and understanding of the operating system. However, academic, government, and other researchers and developers writing in response to the CDDL release indicate that they will steer clear of Solaris, so as not to taint future submissions to other projects by participating in a CDDL-based project. Whereas Linux participates in the mainstream efforts of the open source and operating system research community, Sun has chosen to isolate Solaris' contribution to those individuals with interests in no other projects – precluding contributions from other hardware and system software vendors. Sun's approach is in some ways a variant on that used by MySQL for its open source database products, providing users with free access to software while retaining overall product control within a single vendor organization. Sun has differed in that it is using a single license approach with the CDDL, versus MySQL's use of separate commercial and open source licenses. Time will tell what the impact of Sun's vendor-controlled approach to an open source Solaris will have on capability and cost of ownership for its user community.

---

[5] From section 3.1 of the CDDL: "Any Covered Software that You distribute or otherwise make available in Executable form must also be made available in Source Code form and that Source Code form must be distributed only under the terms of this License."

[6] eWeek, December 13, 2004

Comparing Major New Features of Solaris 10 with Linux 2.6

The next section compares the newly released features of the Solaris 10 operating system with the capabilities of the most recent Linux distributions, with a view of their impact on end-user organizations.

Application Performance

3GL Compilation and Tuning
Performance is a function of both hardware capability and the efficiency with which software can exploit that capability in executing user workloads. In today's complex, multi-tiered application environment, the ability to understand what the system is doing and to examine the software causes of that behavior are critical elements in tuning performance-sensitive applications and relieving performance bottlenecks.

Solaris 10
Sun's compilation and tuning offering for Solaris is based on its re-branded Studio IDE and compilers, as well as the new DTrace tool. (Note that many Sun ISVs developing across-vendor platforms opt for the GNU project's GNU Compiler Collection (GCC) compiler and related tools for portability and reduction of support costs, in lieu of Sun's offerings.)

DTrace is "a dynamic tracing framework for troubleshooting systemic problems in real time on production systems." DTrace provides a level of transparency and ease-of-use not previously available in the Unix/Linux worlds, providing the ability to follow a thread of execution from the application into the operating system and back out again. These are the types of performance and runtime analytics that the proprietary IBM zOS and DEC/Compaq/HP OpenVMS systems administrators and application developers have long relied upon to get the most out of their production intensive environments. Augmenting existing tools like lockstat, DTrace extends the state-of-the-practice in providing performance information on running processes to developers, lowering the effort to tune both application code and environmental configurations.

Developers of performance-sensitive C and C++ applications running on older 32-bit SPARC Solaris versions interested in moving to 64-bit SPARC systems running Solaris 10 or Opteron-based systems running Solaris 10 or Linux will need a tuning process to make efficient use of the various x86-derived hardware architectures. Areas to be watched and covered in the test cycle include:

- Application programming interface (API) differences, not for system calls but for all the third-party libraries invoked by the application
- Cache handling and access times for parallel execution
- Data type differences (big-endian, little-endian integer representations), as well as the use of multiple datatype models and the need to move to system-derived or specific types rather generic ones (e.g. "long" that have alternate implementations)
- Data alignment and pointer alignment issues
- Memory-move instructions that can have very different performance characteristics on the different hardware architectures, even when they are executed correctly

These porting issues are common to both Unix-to-Unix, Linux-to-Linux, and Unix-to-Linux ports across hardware architectures, and are best handled with source code profiling (see below) and then require parallel testing of the ported applications. DTrace provides a means of observing performance and behavior of the ported application environments, but addresses only a portion of the steps involved. End-user organizations working with Sun's Opteron-based servers told RFG that ports of high performance computing application stacks were "more than just a recompile" and "no easier than porting to Linux on an x86 machine." Solaris users also have access to the complement of Sun tools and third-party code profilers like IBM's Rational Purify for Linux and Unix.

Linux 2.6

3GL application developers have a choice of the GNU GCC compilers (across all Linux hardware platforms), as well as the high-performance compilers for x86 hardware from Intel and PathScale, Inc. A variety of IDEs are available. These include leading-edge offerings from Metrowerks, which provides a more comprehensive commercial alternative to the standard Unix/Linux GNU debugger, and eclipse.org, an open source IDE bootstrapped by IBM with the contribution of WebSphere studio workbench as a foundation layer. Eclipse is quickly becoming the integration point for a range of developer tools based on its open plug-in architecture. With the release of Sun Studio 10 development environment on Linux, this full-featured toolset is available to Linux users and may be especially attractive for developers building software for both Linux and Solaris.

There is no direct equivalent to DTrace currently available for Linux users. However, for both system software and performance-critical applications, there are two major performance profiling tool sets currently available, with a third on the way.

1. For users of both the Intel C++ Compiler for Linux (ICC) and the open source GCC compiler, the Intel VTune tool provides runtime performance information that details hardware utilization and identifies performance bottlenecks within compiled application code on x86 hardware. VTune now comes with an eclipse plug-in for integration with a wide-range of compilers and debuggers popular among Linux developers.
2. In-depth performance tuning for Linux on multiple hardware platforms has been accomplished using a diverse set of tools, including the following:
    - kprobes (for kernel performance information) have been ”accepted upstream” (by the Linux community), and are integrated into the standard kernel starting with version 2.6
    - strace project provides details on system call utilization
    - SGI-Lockmeter provides kernel spinlock information
3. The SystemTap project is a new open source performance tool project starting with code contributions from IBM, Intel, North Carolina State University, Red Hat, and Stanford University. Explicitly looking to provide system tracing and performance monitoring capabilities comparable to those in DTrace[7], SystemTap uses IBM‘s kprobe technology, and extends it from the kernel into user space. (Thereby, replacing the dprobe option that had fragmented information available to Linux developers.) SystemTap provides a library of ”probe handlers” to allow for end-to-end tuning without requiring kernel programming and probe skills on the part of the user. A graphical user interface (GUI) plug-in for oprofile and a scripting interface are already available; a plug-in that will integrate SystemTap into the popular eclipse development environment is under development. Kprobes are enabled in the 2.6 kernel and shared libraries, and are targeted for inclusion in RHEL4 Update 2 and are being back-ported to RHEL3. The initial release is targeted to support IA32, x86-64 (Intel and AMD), and PowerPC hardware.

Performance Benchmarks

One way to compare the overall performance of hardware and operating system platforms without a user organization benchmarking its own applications (a best practice for performance-critical functions) is to compare results for competing platforms on industry standard benchmarks. Enterprise executives should look at vendor-submitted benchmark results with a critical eye. In particular, they should look at the following:

- Configurations – try to compare apples-to-apples
- Date submitted – vendors and technologies will leapfrog each other, but benchmarks with close submission dates provide a reasonable basis for comparison
- Length of leadership periods for any one benchmark – can provide a comparative view of performance differences across incremental hardware and software upgrades

In addition, enterprise executives should look not only at the result but also the cost per unit of performance in getting that result and at the platform across a number of scales for the benchmark.

[7] http://sourceware.org/systemtap/

Over the last 12 months, Linux platforms have shown both best in class and overall leadership benchmark results across a number of standard performance tests. The AMD-based Sun Fire systems have also reported results showing leadership in some categories; but as of the publication date of this report, the total number of current leadership results using Linux outnumber those for any other operating system. To verify the state of the performance competition when making their own decisions, IT professionals should consult the following results publication sites:

- http://www.tpc.org
- http://www.spec.org

Java Development and Tuning

In 2004, the majority of new application development for enterprise server platforms within RFG's survey base moved from 3GLs (C, C++, and COBOL) to Java. Much of this started with the Web-based delivery of interactive applications. It has now been accelerated by the widespread adoption of service-oriented architectures (SOAs) that benefit strongly from the wide range of facilities provided by Apache/Tomcat/JBoss, Sun Java System Application Server, WebLogic, and WebSphere. The Java language and these server environments are designed to be operating system- and hardware-independent (though RFG recommends a "write once, test everywhere" approach pre-deployment). For these applications, the operating system becomes more of a commodity. Both the DTrace and SystemTap development teams have Java Run-time Environment (JRE) performance trace as a project goal. The most frequently used Java-specific tools are Quest Software's JProbe for the developer and Wily Technology's IntraScope for production environments.

What does seem to matter for performance-sensitive Java applications is the availability of competitive Java Virtual Machines (JVMs) on a given hardware/operating system platform. The only supported JVM on Solaris is Sun's HotSpot JVM. However, on Linux x86 platforms, Sun's JVM has competition from BEA Systems, Inc. (JRockit), IBM, and even the GNU Complier for the Java Programming Language (GCJ). Published benchmark performance rankings change with new hardware and software releases. However, performance and price performance, both in-the-wild (as reported to RFG by multiple end-user organizations) and in-the-lab as reported by vendors to SPEC (http://www.spec.org/), are better for Linux on x86 with the BEA and IBM JREs and Linux with the IBM JRE on POWER versus Sun's JRE on its SPARC or AMD platforms.

Security

System and network security is an increasing expense for IT organizations, and remains a top five concern among CIOs in a 2004 RFG survey of senior IT executives. Operating system features cannot provide complete protection for an enterprise's applications and data, but they are the basis for much of that protection. Operating system security faults, like those continually reported for the Windows operating systems, are often exploited in the most expensive of security breaches. Effective security postures are based on a combination of hardened resources (applications, layered software, operating systems, and network devices), security policies on how those resources are used, and tools to implement, test, enforce, and monitor activity against those policies.

Security Capabilities and Implementations – More Commonalities than Differences

Both Solaris and Linux have deservedly good reputations for their security capabilities, and have significantly fewer security incidents per year than Windows systems running comparable classes of applications. There are new features in both Solaris 10 and Linux 2.6 worth noting and taking advantage of, as organizations further harden their IT assets out of necessity.

The algorithms, facilities, and implementations of security-critical elements of the Unix and Linux platforms have more in common than in areas that differentiate them. As examples, both make use of:

- Blowfish, MD5, and RC2 cryptography to protect data and connections
- Kerberos as the default token passing mechanism for remote authentication
- OpenSSL and SSH implementations for secure remote server connections
- Internet Protocol security (IPsec) to provide secure virtual private network (VPN) connections through cryptographic tunneling at the network protocol layer on both IPv4 and v6

- Installation-time securing of new operating system instances based on scripted policies, including JumpStart for Solaris and RHN/Kickstart for Red Hat Enterprise Linux
- Security auditing and reporting tools

At the same time, many security areas, such as firewall software and multi-factor authentication device support, are best augmented with additional commercial or open source packages. Most of these packages are supported in binary code today on Linux on x86 and Solaris on SPARC. The range of specialized security offerings narrows as users move off the primary hardware platforms for each operating system[8].

Military- and Intelligence-Grade Secure Operating Systems

The U.S. military and intelligence communities have long required more secure versions of commercial operating systems to meet the needs of computing in sensitive environments, formerly defined by the "Orange Book" and now specified through TTAP and Common Criteria. Both Solaris and Linux have increased security offerings to meet the needs of classified environments. To meet these needs operating systems must:

- Require levels of security classification (i.e., confidential, secret, top secret) for both users and resources (data, processes, and devices)
- Provide additional discretionary controls on access to those resources
- Require auditing of those accesses within the classification constraints

Trusted Solaris

Trusted Solaris has met those needs for government agencies and (increasingly) for a commercial user base. Both user and process rights management combined with extended access controls are the major visible differences between Trusted Solaris and standard Unix implementations. Since the release of Trusted Solaris 8, Sun has had a good reputation for the usability of Trusted Solaris, Trusted Solaris 8 is evaluated at the Common Criteria LSPP/RBACPP/CAPP/EAL4+ level, the highest level available for general purpose computing systems. Solaris 9 is evaluated at RBACPP/CAPP/EAL4+. With Solaris 10, a number of the underpinnings have been moved into the base product, and a ""Trusted Solaris Extensions" package will provide the additional facilities needed for the LSPP element of the Common Criteria evaluation, primarily labeling and multi-level network and desktop support. Sun told RFG that those extensions are en route for delivery by the end of the year; evaluation of Solaris 10 is underway now. This is clearly an area where Sun has the leading product offerings.

Security Enhanced Linux

The U.S. National Security Agency, which has often been unofficially involved with the open source community, combined the results of some of its internal and funded research projects to create and release Security Enhanced Linux (SELinux). To quote the NSA, "This version of Linux has a strong, flexible mandatory access control architecture incorporated into the major subsystems of the kernel. The system provides a mechanism to enforce the separation of information based on confidentiality and integrity requirements. This allows threats of tampering and bypassing of application security mechanisms to be addressed and enables the confinement of damage that can be caused by malicious or flawed applications."

On the SELinux Web site at http://www.nsa.gov/selinux/, the NSA points out that Linux was selected not only because of its strengths as a platform, but also because the NSA wanted to contribute this work back to the public. In so doing, the NSA would help create more secure platforms and encourage research to improve overall information system security. SELinux is becoming widely adopted with the U.S. Department of Defense (DoD) and intelligence communities.

New Security Extensions to Solaris 10

The major extensions to user security in Solaris 10 are based on the move of process rights management capabilities originally developed for the Trusted Solaris product into the base Solaris 10 offering. In addition, the Solaris Container capability is positioned as a major security feature of Solaris on SPARC. The process rights management capability

[8] Though the vast majority of the open source security packages will compile from source for Linux on PPC hardware and receive community support on these platforms. Most of these packages can also be compiled for Solaris on x86 as well.

makes it possible to provide system administrators with the ability to administer systems without having the unlimited power provided by the standard Unix/Linux superuser (root) account. In commercial environments where sensitive financial and personal data are being stored and manipulated, this provides for better security protection. However, organizations need to change their systems management and security practices to take advantage of the facility.

The Solaris Container capability is based on Zones and a Resource Manager. Zones isolate application components from each other, even though they share a single Solaris instance of Solaris 10 running on a single computer. (See the virtualization section below.) The security benefit here is that separate application and middleware functions (databases, firewalls, mail servers, and Web servers) can run at separate IP addresses and as if they were on separate machines. This makes it substantially more difficult for an intruder or piece of malware that compromised one such function to compromise other functions running on the same piece of hardware.

Since Solaris 2.6, the system call mechanism in Solaris has been designed to block buffer overflow attacks at the system level. These facilities have been strengthened and extended with each additional release. This closes off a wide range of vulnerabilities in code that has been relinked on the Solaris 9 platform. Users of earlier Solaris releases will want to rebuild their applications to take advantage of this protection[9]. Solaris 10 has added both process rights management and a secure execution facility based on signed executable binary files that authenticate images to the operating system. This should prevent the execution of the preponderance of unauthorized code. Sun has also replaced its old firewall software with a new, much stronger kernel-based capability and an associated configuration tool taken from the open source BSD community.

New Security Extensions to Linux 2.6

Linux security has been enhanced in the 2.6 kernel in several areas:

- The entirety of kernel-based security has been modularized to allow for partitioning of superuser privileges, making the necessity of a root account with access to all facilities and data to be optional, not required. However, as of this time, other standard security models need to be developed for practical application of this facility. This facility and the security policy abstraction capability have been moved into the 2.6 kernel from the NSA's SELinux distribution.
- Binary modules can no longer overload (redefine) system calls and have access to the system call table. ExecShield run-time technology and the gcc capability to create position independent executables (PIE) both provide increased protection from buffer overflow and code insertion exploits.
- Linux now has access to hardware-based random number generators, not just the entropy pool mechanism used in previous kernel releases, to provide system-specific random number seeding essential to cryptographic algorithms.
- EMC Corp.'s VMware products for x86 architectures and micro-partitioning capabilities on IBM's zSeries and POWER architecture products provide for enhanced security through complete isolation of operating system images on their respective hardware platforms. (See the Operating System Virtualization discussion below.)

Linux has had an improved kernel-based firewall capability since the replacement of IPchains with the BSD-derived IPtables facility in the 2.4 kernel. This is just one example of a situation where IP originally developed in one open source project has been adopted by another to the benefit of a large user community.

Novell SUSE LINUX Enterprise Server 8 and Red Hat Enterprise Linux Release 3 were certified to the CAPP/EAL3+ level in 2004. Both Red Hat and Novell SUSE are certifying their 2.6-based enterprise distributions to CAPP/EAL4+ (the highest CAPP level for which certification testing is available) in 2005. Novell certified with the LAuS audit tool against the full IBM eServer product line, that certification is complete. Red Hat's certification is in progress with hardware from multiple vendors, including IBM, and is using the "audit" tool accepted by the upstream Linux community.

[9] Although vendors often offer binary compatibility across releases, it is a best practice to recompile and relink all in-house applications from sources at the time of each operating system upgrade. This provides certainty that in the event of an application error requiring repair that the application can, in fact, be built on current systems.

Hardware Support

This is an area where the differences between the choices and flexibility offered by Solaris and Linux vary greatly. The nature and range of hardware support is an area where the differences between a vendor offering and a community-supported open source system become readily apparent.

Solaris 10
Solaris 10 supports most of Sun's installed base of SPARC64, UltraSPARC-2, UltraSPARC-3 and UltraSPARC-4 hardware, as well as provides renewed support for 32-bit x86 and AMD Opteron-powered x86-64 hardware and EM64T. With the release of free Solaris under the CDDL, Sun's mechanism for monetizing its investment in Solaris is through the sale of hardware and the licensing of its Java Server and Java Desktop environments (as well as Solaris licensing for large-scale SMP servers) as well as support contracts for Solaris itself.

- Sun has a rich RISC product line, and its Enterprise 15K and 25K servers have set the standard for the very high end of the RISC Unix vertically scaled market. For very large transactional databases and enterprise resource planning (ERP)/customer relationship management (CRM) workloads, these have been the scale-up system of choice for most major enterprises.
- After choosing not to release Solaris 9 for generic x86 offerings in 2001, Sun has returned to that market in 2002, this time with its own hardware offerings. The Sun Fire server line began with one and two rack unit server offerings, and has since been expanded to include a new generation of AMD Opteron-powered server offerings. Sun offers not only Solaris, but also the most popular Linux distributions on this hardware. Sun supports 64 bit Solaris, Red Hat, and Novell SUSE.

Many of the strengths of the Sun user experience have come from the tight linking of Solaris with SPARC and AMD Opteron hardware offerings. In particular, Sun has provided hardware fault detection and predictive self-healing capabilities based on Solaris' access to features designed into the underlying hardware. Some of the other availability offerings of the Solaris Services Manager, such as Intuitive Shutdown and Manual Configuration Rollback, are hardware independent and available to all Solaris systems.

User Experiences and Intentions

- Enterprises surveyed by RFG (all of whom had substantial numbers of SPARC Solaris systems installed) were nearly unanimous in that they had no Sun Xeon-based systems installed and no intention of making Solaris-based Opteron investments in 2005.
- Comments from senior IT executives were often variations on this one: "SPARC Solaris is the one constant element in Sun's strategy. Everything else has changed too often for [my firm] to even place a bet."
- More than half of the firms that had existing large-scale SMP servers (32-way and 64-way) told RFG that the high-end of Sun's Enterprise server line would continue to be a part of their ongoing system architectures.
- The small number of firms that had tested Solaris 10 on Opteron made the point that "the hardware differences are real, it is not the same support experience."
- None of the end users surveyed found any indications of support or "marketing push" from other x86 or AMD-powered vendors for Solaris as an operating system alternative.
- A major financial services company told RFG: "The effort and cost to port to Solaris and AMD is no lower than porting to Linux. There is no advantage to staying with Solaris when moving hardware architectures."

The consensus among the leading firms surveyed was that Solaris remains a RISC Unix operating system to them, not an alternative to Linux on open platforms. This is why the majority of Sun's Opteron server sales to date are deployed with Linux, not Solaris. Over time, the awareness of Solaris on Opteron is likely to increase, and AMD has reported that Sun is the leader in AMD Opteron shipments.

Linux 2.6
The dozens of Linux distributions support the widest range of hardware manufactured by the largest number of vendors of any operating system yet created. The range extends from Motorola, Inc. smart phones through Sharp Electronics Corp. industrial handhelds; to PC laptops and desktops; and to servers from low-end Pentiums, through Alpha, SPARC,

and POWERPC RISC servers all the way to IBM's zSeries mainframes. At all of these levels, there are multiple competitive vendors for both the Linux distribution and the class of hardware.

The Linux community, out of the interests of its users, provides support for the complete server stack on every major available platform. This range of offerings provides ongoing protection for the enterprise investment in operating system expertise and application support, a five- to 10-year investment that long outlasts any hardware vendor's recent leadership of innovation or benchmarks. For the first time, open systems can break the use of operating system dependence as a means to co-opt user organization architectures, and locking customers in to product lines long after their leadership status is past. In RFG's recent surveys of 30 major financial institutions worldwide, flexibility and control in infrastructure selection was second only to the combination of performance and price performance in users' decisions to standardize on Linux as their primary new server operating system. RFG believes it is the level playing field provided by Linux and the competition that makes it possible that have made the platform's price performance so compelling.[10]

Operating System Selection and Enterprise Hardware Strategies

The firms RFG surveyed were remarkably unanimous in their endorsement of Linux on Intel and AMD platforms as a successful platform in their infrastructure strategies, with all continuing to make increasing investments in 2005. 64-bit Linux strategies differed, with the following responses:

- Most planned purchases for Linux still focus on the 32-bit Intel Xeon hardware, but
- AMD Opteron-based systems showed the most future interest, with
- Intel EM64T next, followed by
- IBM POWER5.
- There was only minimal interest in Itanium as a platform, most citing cost, heat, performance, and power concerns with the chip. The cost of porting applications to the Itanium system was also cited, along with a comparative lack of ISV software support[11].
- A minority of respondents had implemented Linux on zSeries mainframe systems, primarily to provide Web application access to zOS hosted databases.
- Firms agreed strongly that Linux was their primary new application deployment platform. There was a range of variation, however, on whether they were actively porting existing Unix (AIX, HP-UX, or Solaris) and Windows applications to Linux. Some large international financial services firms were porting everything possible with a goal of no Unix by 2006, mostly moving to Linux and Windows on x86 platforms. Others were making decisions based on a range of financial and technical criteria, including availability of third-party applications and software libraries. What RFG found was diversity in approach driven by differing requirements, priorities, and maturity/comfort levels with Linux across the surveyed population.

[10] Though clearly sharing the x86 hardware platform with PC operating systems (Windows, Netware, UnixWare) provided for volume pricing, as Linux's popularity grew during the period from 1992-2000 it became a volume platform in its own right.

[11] The sample lacked the very high-end of scientific and technical workloads (primarily in physics and fluid dynamics) that require floating point performance as their ultimate concern.

File Systems

Both Solaris 10 and Linux provide core operating system upgrades to improve input/output (I/O) performance and supported file system sizes. In both operating system environments, there are multiple file system offerings available. IT professionals should examine these offerings in the context of application requirements for:

- Performance
- Recovery and integrity
- Clustering
- Volume Management.

For many users of lower-end servers, network attached storage (NAS) storage has preempted most of the operating system-level file and volume management requirements, since these capabilities are supported within the storage appliance. The Network File System (NFS), created by Sun and always intended to be network-independent, provides the standard means of remotely serving file systems to both Unix and Linux servers and workstations. Multiple NFS implementations exist, and the core technology was released to the open source community by Sun under the specially created Industry Standards Source License (ISSL) in 2000.

Solaris 10

Most commercial users (versus scientific) have long discounted the Unix File System (UFS) included in previous versions of Solaris. With availability, manageability, performance, and size limitations, users looking for the full power of Sun's Unix for database and other data-centric applications have turned elsewhere. For many Sybase Inc. and Oracle users, raw partitions were allocated for full management by the RDBMS; this high-performance approach has the down side of limiting manageability. For most of the large Sun shops surveyed by RFG, the Veritas Software Corp. Foundation Suite (VFS), which includes the VxFS file system and the VxVM volume manager, has been the basis for enterprise storage strategies incorporating storage area networks (SANs) and providing multi-site business continuity strategies.

File System Name	Source	Key Features	Open Source
UFS	Sun	General-purpose file system volume management tailored for small, reusable files and scaling up to a few of data	No (though CDDL release is possible with OpenSolaris)
ZFS	Sun	Self-healing, self-managing file management capability designed to protect data from corruption offer 128-bit address space	No(TBA with OpenSolaris)
VxFS	Veritas /Symantec	Part of VFS with VxVM. Includes support for Veritas database-specific back-up facilities, mirroring, and cluster products. Compatible across other operating systems.	No
QFS	Sun	SAN file sharing and distribution, linked to the StorEdge suite and SAM-FS policy manager.	No

One to Watch: Solaris 10 ZetaByte File System (ZFS)

Sun has announced that Solaris 10 will, by the end of 2005, include the new ZFS file system. The key characteristics of ZFS are:

- Combination of file system and volume manager into a single subsystem with a single management syntax
- 64-bit checksums for increased data integrity
- Copy-on-write to minimize recovery times after failures
- 128-bit addressing

ZFS addresses a number of the historical weaknesses in UFS. If successful, it will give Sun a high-performance journaled file systems with addressing capabilities beyond those of the major Linux distributions. All server environments have

recently needed to expand beyond 32-bit file systems to meet the 40 to60 percent per year increases in data storage in major enterprises. As the first 128-bit file system anywhere, ZFS is making a dramatic move toward conceptually managing storage volumes that would far exceeds storage volume requirements for the foreseeable future. In fact, RFG believes 64-bit file systems should be able to manage far more storage than is needed for any currently extant or planned enterprise workload for many years to come. The combination of file and volume management into a single facility should both simplify and accelerate a range of file management activities.

Once released, ZFS should be a very attractive alternative for Solaris users; but as an entirely new file system, it is a technology that end-user organizations will need to test carefully before trusting it with critical enterprise data. As a result, RFG believes that the Veritas Foundation Suite (VFS) will remain the default enterprise file and volume management choice for enterprise Solaris users, particularly for those seeking a single cross-operating system solution for the near to intermediate term. Note that VFS is supported on Linux as well as a variety of other Unix platforms.

Linux 2.6 File Systems and I/O Enhancements

The 2.6 kernel provides the greatest number of I/O and file system support enhancements seen in any Linux release. They include:

- Kernel support for 64-bit file system implementations
- Increased I/O throughput
- Improved threading support with the Native Posix Thread Library
- A standard asynchronous I/O implementation (replacing multiple competing patches to the 2.4 kernel)
- Improved kernel HBA support for SAN connectivity
- TCP/IP stack improvements that speed NAS I/O access

The native 64-bit support is implemented in all of the current releases of the major Linux file systems. I/O improvements in the 2.6 kernel have made it a tier-one platform for all of the major portable database vendors, including CA, IBM, MySQL, Oracle, and Sybase. In the Linux environment, user organizations have a range of high quality file system choices. The table below lists these major choices in a rough order of popularity within RFG's survey population.

File System Name	Source	Key Features	Open Source
Ext3	Red Hat	Journaled for fast writes and short recovery times. Block reservations and hash tree directories improve performance.	Yes
Reiser	Novell	Journaled for fast rights and short recovery times.	Yes
VxFS	Veritas /Symantec	Part of VFS with VxVM. Includes support for Veritas database-specific back-up facilities, mirroring, and cluster products. Compatible across other operating systems.	No
XFS	Silicon Graphics, Inc. (SGI)	High-performance journaling file system with full 64-bit addressing and scalable structures and algorithms.	Yes

Cluster Technology and File System Support

Getting server applications beyond 99.9 percent (three-nines) availability has long required clustering facilities. The Unix and Linux operating systems have long lagged behind proprietary commercial operating systems, particularly HP/DEC OpenVMS, HP/Tandem NonStop, and IBM zOS Parallel Sysplex, in providing these capabilities.

Clustering for Solaris

In the Sun world, this has primarily been met through ISV products from Veritas and EMC Legato, with native Sun Cluster technology slowly taking hold. There are currently no clustering facilities available for Solaris 10. Sun told RFG that Sun Clusters would support Solaris 10 with ZFS by the end of this year. An operating system/file system-based

distributed lock manager accessible to data management software would have been welcome, but is apparently still missing in the planned ZFS for Solaris 10 release.

Clustering for Linux

One result of Linux's early adoption on PC and low-end server platforms for use in horizontally scalable applications has been the concentration of much of the academic and commercial operating system clustering research on Linux. In addition to strictly scientific clustering packages like Beowulf (initially developed at NASA and now maintained as a commercial open source offering by Penguin Computing Inc.), more general-purpose high availability and scalability-oriented clusters for commercial workloads abound. Some of the more widely used options are listed in the table below.

Clustered File System	Source	Key Features	Open Source
GFS	Red Hat (Sistina acquisition)	Clustered, distributed lock manager.	Yes
Matrix Server	PolyServe, Inc.	Clustered, distributed lock manager.	No
OCFS	Oracle	Clustered file system developed originally to support Oracle database tablespaces.	Yes
VCS	Veritas	Multi-platform, including Solaris and other Unix offerings. Supports companion remote mirroring products.	No

These choices,with the exception of the Oracle Clustered File System, have all been in production use for at least three years, and provide Linux users with a range of powerful options. It should be noted that there are more choices available (such as Lustre, xCFS, and GPFS) than those listed, and the list is growing. RFG expects that clustering and clustered file systems will substantially grow in popularity with Linux users. Steps have begun to create an open source commercial clustering solution for Linux. Led by Red Hat engineers with DEC Tru64 Unix cluster experience, with input from IBM, Oracle, and other vendors, this may be the most important initiative for server technology over the next several years.

Operating System Virtualization

Operating system virtualization is a term generically applied to breaking the link between a single running operating system instance and a single physical computer. There is a wide range of solutions available to provide this capability, and these solutions need to be evaluated in light of the business goals for the virtualized system(s). The desired benefits of virtualized operating system instances include:

- Server consolidation to increase utilization of hardware assets while maintaining operational independence of consolidated software stacks
- Isolation of incompatible software stacks
- Isolation of environments to simplify change control and upgrades
- Use of a single piece of hardware to run multiple different operating system types or versions of operating systems

Operating system virtualization technology can be done at multiple levels, starting at the bare iron and working its way up. The options available are a function of the following:

1. Hardware
2. Base or host operating system
3. Guest or functional operating system actually running applications and user workloads

The original model for system virtualization is IBM's Virtual Machine (VM) operating system, originally created for the System 370, and now providing the virtualization facility in the modern zSeries machine. In this case, VM boots onto the underlying hardware, and functions solely as an efficient host for the guest operating systems used to accomplish both user and administrative functions. Each guest operating system functions as if it had all the resources of a physical machine to itself, everything from memory to CPU compute resources are timesliced among the guest operating system instances based on a weighted, fair-use schedule. The virtual machine capabilities of modern platforms are all variations on this model, providing for different server node names and network addresses for the virtual systems running inside

of a single physical machine. In understanding these options, enterprise executives need to start with the hardware-based offerings first, and then those related to the operating system software itself.

Hardware-Based Virtualization

VMware
EMC's VMware product line is the primary commercial virtual system product line for x86-type hardware. The VMware product line is built on two underlying virtual machine technologies:

1. VMware ESX Server that runs as the native operating system on server hardware to host guest operating systems
2. VMware GSX Server that runs as an application on top of either a Windows or Linux operating system to then host the guest operating system products

VMware GSX is primarily used as a development and test platform, while ESX provides the more robust environment for running production workloads. ESX, however, is available only on a limited number of hardware platforms, and supports neither AMD Opteron platforms in 64-bit mode[12] nor Sun Solaris as a guest operating system. AMD, EMC, and HP announced in January 2005 an agreement to port VMware ESX to HP's Opteron-based servers. VMware ESX does provide a tested means of running multiple, entirely independent, production Linux and Windows virtual machines on a single physical SMP server. The VMware Suite provides a comprehensive set of tools for creating, migrating, and managing virtual machine instances among VMware-equipped servers.

VMware GSX is supported on AMD Athlon-, AMD Opteron-, EM64T-, Pentium-, and Xeon-based hardware. VMware GSX requires a Linux or Windows Server host operating system. GSX fully supports hosting of a range of Linux, Netware, and Windows guest operating systems. GSX has experimental support for Solaris 10 as a guest operating system only.

One to Watch: The Xen Project
The Xen project is an open source virtual machine monitor for the x86 platform with a focus on performance and reliability using paravirtualization (modifying the guest operating system to minimize expensive hardware emulations). As it is a GPL project and requires guest operating systems be ported on to the Xen monitor, it is only available for open source operating systems. Currently, 32-bit 2.4 and 2.6 kernel-based Linux distributions are running on Xen (along with the NetBSD open source Unix operating system). With the enhanced hardware virtualization support coming in the next generation of Intel and AMD chips, many of the modifications to the guest operating system needed for paravirtualization will no longer be required. Sun told RFG that Xen support for Solaris 10 is planned. Support of Windows 2000/2003 is an announced intention. Support for Solaris should be possible but has not been announced. Xen still is a work-in-progress, and is a technology for commercial enterprises to track and test. It is not yet ready for adoption.

POWER5 Hardware Virtualization Technology
The IBM eServer pSeries and OpenPOWER hardware platforms come equipped with a virtualization facility at the hardware/firmware level that provide mainframe-style logical partitioning (LPAR) capability. The LPAR capability allows each partition to be booted separately, each potentially running a different operating system image. This allows, for example, the ability to run both IBM AIX Unix partitions and Red Hat Linux partitions simultaneously on the same pSeries server. One seemingly simple but operationally critical feature is the ability to independently patch and upgrade LPARs running different software stacks. This both lowers the risk of upgrade induced failures, and allows patching to correct critical operational and security flaws that may affect one software stack and not another. This allows the critical patching of a Web server with a patch not yet tested for a database server on the same hardware, but not exposed to the same vulnerabilities.

The eServer OpenPOWER server line differs from pSeries machines in that it is engineered to run Linux only, and

[12] EMC, AMD, and HP announced in January of 2005 an agreement to port VMware ESX to HP's Opteron-based servers. There is currently no target release date or discussion of Sun hardware support.

priced to compete with commodity x86 hardware rather than being offered at RISC Unix price points. New micro-partitioning technology available to all POWER5 servers provides the ability to have more active partitions than the physical CPU count in a given machine, allowing optimal resource utilization in server consolidation configuration. Dynamic logical partitioning (DLPAR) capability was added with the p5 generation of hardware, and allows reallocation of resources without the rebooting of LPARs. Although dynamic memory reallocation is available only for AIX today, Linux kernel enhancements (possibly available as a patch) that would be required for dynamic memory upgrades are being developed.

Hardware-Independent Virtualization

Solaris 10 Containers
The Solaris 10 Container virtualization technology runs above the hardware level as a configuration facility within the operating system. It is a separate capability from the hardware system domain facility for the Enterprise servers (Enterprise 3800 and up) that isolates physical components within a chassis for fault independence. Unlike either VMware or IBM LPARs, Solaris Containers provide for the creation and isolation of separate application and user stack environments on top of a single running operating system image. The result is lower memory requirements and simpler configuration (no need to install or clone an operating system for each virtual machine or LPAR). The net benefit of this is the automation of the server and service isolation that knowledgeable system administrators have created for years using the change root (chroot) function of the Unix operating system. Individual application stacks cannot corrupt each other, especially if compromised by a hacker, and software failures in one container are isolated to that container.

The limitation of the container approach is that only a single image of the operating system is running across all containers, and any upgrade or patch applied to the base Solaris version has to be compatible will the full software stacks running in all containers. This increases the management costs of server consolidation projects by requiring coordination of all application and support groups for an upgrade whenever any one application or stack requires a patch for operational or security reasons. Even if no one has to change a line of code, everyone has to test – and do it on a separate physical machine. Operating system upgrades will only be possible when all layered products, including ISV offerings, have been tested and certified for the new operating system version.

User Mode Linux
In addition to the hardware platform-based virtualization options available to the Linux user from VMware on x86 and LPARs on POWER5, there is an additional option available at the operating system level. User Mode Linux (http://usermodelinux.org) describes itself as a patch for the Linux kernel, which allows an executable binary to be compiled and executed on a host Linux machine. The kernel can be assigned virtual resources, including a root file system and swap space, and can have a hardware configuration entirely separated from that of the host. This allows multiple different versions and distributions of Linux to run on a single machine. User Mode Linux is currently running on 32- and 64-bit x86 platforms as well as 32-bit PowerPC; other platform ports including 64-bit POWER5 are in process.

RFG Conclusions and Recommendations

Both Solaris 10 and the enterprise distributions based on the Linux 2.6 kernel are complete and capable server operating systems. The roles that each best play in an organization's enterprise architecture is a function of application requirements, existing hardware and software installed base, and future directions. In making operating system technology investments, IT executives need to do so in the context of these three elements.

With the release of Solaris 10, Sun has brought out the largest group of new features for a proprietary Unix release in more than five years. Enhanced security features, a new code trace debugging and tuning tool, a new file system, performance improvements, and support for a Sun server line based on AMD Opteron, make this the most robust and exciting Unix release in many years. Solaris 10 adds additional lifetime to the investments that user organizations have made in their SPARC Solaris applications and infrastructure, and provides the basis for a strong, vendor-directed Unix platform. In RFG's view, the CDDL licensing of Solaris 10 makes it legally open source; however, it does not provide the basis for Solaris users to benefit from the broad open source community. It particularly will not benefit from contributions from a larger hardware vendor community (the source of billions of dollars of Linux investment), which is discouraged from seriously investing in a product that will remain controlled by a single direct competitor.

With the major server distributions based on the 2.6 Linux kernel, Linux has grown up to become a first-tier server operating system – not just a low-cost alternative for utility computing and static Web serving. SMP support has now extended to the 8-way to 16-way scale, although the preponderance of server hardware shipments for all operating systems are 2-way and 4-way, and has improved threading and scheduling speed complex workloads. In addition, asynchronous I/O is now a core kernel functionality and not a patch needed to run specific relational database management system (RDBMS) products. Domain-based security schemes and process isolation have moved from the NSA's Security Enhanced Linux into the standard kernel. Most critically, the ecosystem of open source contributors, Linux distribution vendors, hardware vendors, layered software vendors, and user organizations continue to propel Linux development and improvements.

The Support Question

Operating systems are a key element of an enterprise's server platform architecture, and support options need to be evaluated in that light. User organizations have long been used to the proprietary server model, with hardware and operating system support coming from a single vendor, providing the user with "one throat to choke" in the event of a problem or concern. Sun has had a leading support lifecycle process for several years, providing a 10-year roadmap to end-of-life for its operating system customers.

A best-of-breed approach to the hardware and software stack has become the standard for successful IT departments in large enterprises since the mid-1990s. Having the flexibility to license an operating system independent of the manufacturer of the hardware it runs on has certainly driven the PC revolution. Moreover, the benefits of that commodity-driven pricing have now been brought from the desktop to the server room and now to the data center. Linux, as a broadly supported community-developed open source product, provides additional support options that even the PC community has not had. Operating system support is available from the distribution vendor, the hardware vendor, services vendors, and even from the development community over the Internet. User experiences conveyed during RFG interviews with Linux adopters show that the level of service available for enterprises dealing with top-tier providers is now at a level equal to or better than that available from a single system vendor. Support lifecycles vary for Linux, depending on the choice of operating system support vendors; Red Hat, as one good example, provides a 7-year support lifecycle for purchasers of its enterprise offerings. Customers who were strongly interested in having multiple Linux hardware vendors tended to purchase support from a distribution vendor; those with a stronger interest in running multiple distributions tended to purchase Linux support from their primary hardware vendor.

Findings and Recommendations

The following points summarize RFG's analysis and research findings for platform selection.

- For the vast majority of applications, user organizations report superior results at low costs by deploying new

applications using enterprise Linux running on 2-way and 4-way blade or rack mount servers.

- The easiest applications to port to a Linux platform are those running inside of virtual machines or other types of ”containers,” where the code is already operating system- and hardware-independent. This includes Web applications with displays in HTML and logic in ColdFusion, Javascript, Perl, PHP, or similar facilities. Linux makes an excellent Java applications platform, whether Java 2 Platform, Enterprise Edition (J2EE),including JBoss, Jonas, Oracle AS, WebLogic, or WebSphere, JSP, servlet, or stand-alone server side Java. Report writing tools fall into this category as well.

- Linux has become a preferred platform for the major database vendors, including IBM, Oracle, and Sybase. This is a major cost reduction and performance opportunity for future and upgraded database deployments to 2-way and 4-way servers. This is an effective platform for a range of small to medium-scale transactional applications, especially given the performance of the most recent Opteron and POWER5 4-way servers.

- Oracle RAC (9i and 10g) as well as DB2 ICE horizontally scaled clustered database servers provide a very cost effective alternative to high-end Unix SMP servers for decision support database applications.

- High-end online transaction processing (OLTP) database workloads have tended to stay on large-scale SMP Unix servers for availability. Firms that have these on Solaris may want to keep them there at least for the intermediate term, unless the economic incentive to migrate to clustered Linux database servers is substantial.

- Applications that benefit strongly from very high-end SMP configurations (32-way and above), where seamless shared memory is critical to the architecture, should remain on Unix. Sun users will want to upgrade these to Solaris 10.

- Applications that benefit strongly from horizontal scaling and require proven cluster technology will get price/performance advantages from the wide range of competitive, proven, Linux-based cluster technologies.

- There are widely installed and well supported versions of most middleware software packages on both Solaris/SPARC and Linux/x86 (all the various chipsets).That level of software support is not currently available on Solaris/x86.

- Applications written in 3GLs, particularly C, C++, and FORTRAN, require code inspection and specific cost-benefit analysis to determine the ROI of a port to a commodity hardware platform. A key factor in determining the benefits of such migrations will be the availability of third-party layered software and libraries on the target platform. If migration or upgrade requires newer versions with changed APIs, then the cost to migrate can escalate rapidly. Note that user experiences reported to RFG with migrating 3GL applications from Solaris 2.5, 6, and 7 to either Linux on x86 or to the pre-release versions of Solaris 10 on Opteron showed roughly equivalent levels of effort.

- As stated in the beginning of this paper, there are many factors that users must consider in making a final operating system selection, including support, company stability, and technology innovation, in addition to application requirements and third-party software availability. For many users, the Linux-versus-Solaris decision is only the first step in this process.

In summary, RFG believes that a better Solaris has long been needed by Sun‘s installed base. While several of Sun‘s ideas and approaches for Solaris 10 extend the state of the practice, whether this will have a positive effect on its position in the market remains to be seen. The advanced capabilities in the current (2.6) Linux kernel more than meet the needs of most organizations for an increasingly wide range of workloads. Moreover, the diverse global community that creates, enables, enhances, and builds upon Linux has demonstrated its ability to devote more resources and create more innovation than may ever be possible for Sun and its nascent OpenSolaris community to match.