



## Operating System Virtualization Simplifies Server Resource Management

Tony Iams, VP and Senior Analyst, [tonyi@ideasinternational.com](mailto:tonyi@ideasinternational.com)  
July 22, 2005

### SERVICE NAME

- » System Software

### CORE TOPICS

- » Server Partitions and Virtual Machines
- » Resource Management Software

### KEY ISSUES

- » Which vendors will provide robust UNIX-based resource management solutions?
- » Who will be the successful providers of partitioning and virtual machine solutions?
- » How do partitions and virtual machines compare for managing workloads?
- » How are partitions/virtual machines used in practice?
- » How can partitions/virtual machines help with server consolidation?
- » How will Sun Solaris evolve?

### RECOMMENDATIONS FOR USERS

- » Assess the dependency of applications on specific kernel releases or patch levels. If multiple applications are independent of these factors, they are potentially suitable for consolidation using operating system virtualization.
- » Explore the use of operating system virtualization for specific workloads that have a high degree of commonality across instances, such as ISP applications and secure remote management consoles.
- » Determine the sensitivity of applications to resource reallocation and granularity. If it is necessary to start or stop applications quickly, or provide CPU, memory, and I/O resources to them in very precise allotments, operating system virtualization may offer some advantages.

### Summary

As virtualization gains momentum, users are becoming increasingly concerned with managing the burden of installing and updating large numbers of operating system instances. For appropriate workloads, operating system virtualization tools such as Solaris Containers and SWsoft Virtuozzo provide many of the same benefits as virtual machines, while avoiding much of the overhead associated with managing each virtual server.

### Increasing Server Utilization through Virtualization

Advocates of virtualization often promise that it will improve server utilization. As the story goes, many organizations have become accustomed to deploying one workload per server. While expedient, this approach leads to relatively low utilization of server resources, with many servers today using only 8% to 15% of their available capacity. As a result of such suboptimal usage, organizations tend to deploy far more servers than necessary to handle their workloads. Virtualization helps to increase server utilization by enabling multiple resource-intensive applications to run simultaneously on a single server, reducing the "white space" of unused resources that might be incurred by standalone servers. Running multiple critical applications on a single machine requires mechanisms to prevent them from interfering with each other's operation, so that even if one application fails catastrophically or its consumption of resources surges dramatically, the remaining applications continue to deliver their expected levels of performance.

Traditionally, the options for running multiple workloads on a single server have included the following tools:

- » *Resource management software* allocates system resources such as CPU, memory, and some forms of I/O in very precise increments to individual applications based on flexible scheduling policies. Resource management software generally manages workloads running within a single instance of the operating system, doling out resources to individual operating system processes, or groups of them. It effectively overrides the function of the standard operating system scheduler, instead assigning resources such as CPU cycles, memory, and I/O bandwidth according to business-driven policies that consider the relative priority of different applications. Resource management software has long been available on mainframes and UNIX systems, with offerings such as Solaris Resource Manager, IBM Workload Manager, HP Process Resource Manager (HP PRM), and HP-UX Workload Manager (HP-UX WLM). Microsoft introduced native resource management functions to Windows Server 2003 with its Windows System Resource Manager (WSRM), and SUSE Linux Enterprise Server 9 (SLES 9) includes an early implementation of resource management software called Class-based Kernel Resource Manager (CKRM). Third-party

resource management tools such as Aurema ARMTech support deployment on different operating systems.

- » *Virtual machines* involve creating an entire computer system in software. Because the computer system itself becomes an application, it is possible to treat it like any other piece of software, which is isolated from the underlying hardware server it runs on. This capability allows managers to host multiple virtual computers on a single physical server, and to precisely control the resources assigned to each system. Each application runs as if it were on a real server with its own operating system, and thus does not have to be adapted for virtualization. Virtual machines have been successfully adopted by many users, first on mainframes and more recently on industry-standard hardware through software options such as VMware and Microsoft Virtual Server.
- » *Partitions* also allow multiple instances of an operating system to be hosted on a single physical server. Unlike virtual machines though, partitions receive some assistance from hardware to provide electrical isolation between the different environments. The “bulletproof” isolation that this provides between partitions lends itself well to the consolidation of ultracritical applications, which carry the utmost concerns for integrity and security. However, partitions often have relatively coarse granularity, which limits the precision at which resources can be assigned to workloads. For example, many partitioning options need to be configured with a minimum of four processors and can only be expanded or contracted in four-processor increments. Still, partitions are relatively mature technology. They have long been available on mainframes, and they have been available on UNIX systems for some time, with options such as Sun’s Dynamic Domains, IBM’s Dynamic LPARs, and HP’s nPars. On industry-standard hardware, partitions are supported on products such as the Unisys ES7000 and IBM’s x440.

Other variations of these technologies exist, including HP’s vPars and IBM’s p5 Micro-Partitions. Like virtual machines and partitions, these options allow multiple instances of an operating system to run on a single server, but their implementation is somewhat different from the other mechanisms. HP vPars allow instances of the HP-UX operating system to run on as little as one processor.<sup>1</sup> IBM’s shared-processor LPARs for its POWER5-based pSeries servers allow multiple partitions to simultaneously use individual processors, memory modules, and I/O adapters.

## Operating System Virtualization Addresses Management Burden

Resource management tools, virtual machines, and partitions all enable multiple workloads to be consolidated onto a single server, and thus help administrators increase the utilization of their servers. However, these options do not necessarily help address another, more fundamental problem, which is the burden of managing the environments that are needed to host the consolidated workloads. By increasing the utilization of servers that are deployed, fewer physical servers may be required to host an organization’s workloads, resulting in lower overall hardware costs. However, the relative impact of hardware cost is gradually becoming less significant as server prices continue to drop. By contrast, the burden of installing and updating large numbers of operating system instances is becoming an increasingly significant component of user costs.

Virtual machines and partitions both require every instance of a virtualized server to have its own instance of an operating system, which must be installed and maintained. This requirement maximizes the ability for the virtualized server to deliver an operating environment that is specifically tailored for the application that it has to host, but it does not fundamentally alter the management costs that would be associated with hosting those applications on standalone servers. With virtual machines and partitions, the absolute number of operating systems that have to be managed remains largely unchanged, even as the number of physical servers that must be managed is reduced.

Resource management software potentially helps address this issue by enabling multiple dominant workloads to coexist within a single operating system instance. For workloads that are suitable for consolidation via resource management software, administrators can implement management procedures with significantly improved economies of scale, since they can increase the number of workloads that can be hosted without increasing management overhead of additional operating systems. However, traditional resource management software has proven unsuitable for consolidating many classes of applications, because the isolation between workloads is not terribly strong. With most existing resource management solutions, workloads share access to the same system resources, such as files, devices, and network connections. Thus, applications have relatively little protection from each other in terms of security. Further, there may be far more opportunities for applications to interfere with each other than there is conflicting demand on CPU, memory, and I/O resources. Windows applications have proven to be particularly unsuitable for consolidation with resource management software. They typically do not “play well” with each other, often demanding specific instances of shared resources such as Registry entries and Dynamically Loadable Libraries (DLLs), which causes mutual interference when administrators attempt to host different applications on the same server.

Users are showing increased interest in virtualization solutions that help address the cost of managing multiple instances of operating systems (along with their updates and patches). As such, a new type of virtualization function is starting to gain visibility that is designed to combine the best aspects of resource management and virtual machines. These *operating system virtualization* functions implement a hybrid of virtual machines and resource management in a single mechanism. Like partitions and virtual machines, operating system virtualization provides workloads with private sets of resources that are protected from each other. But while partitions and virtual machines involve hosting a separate copy of the operating system for each virtualized workload, OS virtualization uses only a single copy of the operating system that is shared by all virtual operating system instances – just like traditional resource management software. The virtual operating system instances create the illusion that workloads are running virtually on their own private machines by making a distinct set of OS resources visible to the applications, including:

- » File namespaces, which provide virtualized applications with access to their own logical file system
- » User namespaces and services, which enable personalized instances of network functions such as bind and sendmail

- » IP addresses and hostnames, which provide virtualized applications with a private network connection so that they appear as distinct servers on the network
- » Root passwords, which can restrict access by managers to control individual virtualized operating systems
- » Security profiles, which shield access to the resources of applications from applications running in other virtualized operating systems or the host operating system

Several solutions for implementing operating system virtualization are now available, each with slightly different design points. Sun's Solaris 10 operating system, which began shipping at the beginning of this year on both Sun's SPARC systems and industry-standard X86 hardware, introduced a function called Solaris Containers, which resemble, in concept, the "Jails" mechanism that was originally developed for the FreeBSD operating system. Sun's approach involves constructing kernel-level "zones" to isolate resources and to contain faults between subsets of resources. Solaris Containers are built on top of zones, preventing any user-level action from compromising other Containers or the host operating system; Solaris Process Rights Management functions are employed to maintain security privileges between Containers. Each Container appears to users and applications as if it were a private server running Solaris.

SWsoft offers an operating system virtualization product called Virtuozzo that supports multiple instances of Virtual Private Servers (VPS) on a single physical server. Each VPS behaves as if it were a standalone server from the standpoint of users and applications. A VPS can be rebooted independently, and has its own root access, users, IP addresses, memory, processes, files, applications, system libraries, and configuration files. Virtuozzo is built on top of standard operating systems; the Linux edition was launched in 2001, and the Windows edition was just recently released. Users first install either Linux or Windows as a host operating system, and then install Virtuozzo to virtualize the host operating system. The Virtuozzo management tools have been evolving since the product launched five years ago, and have become comprehensive and quite sophisticated. In addition to supporting 32-bit Intel X86 platforms, SWsoft recently announced 64-bit support for the Itanium platform, and it will add Opteron support later this year. In a sign that it is gaining traction in the enterprise market, SWsoft recently announced a reseller agreement with the system vendor Bull, which will sell and support Virtuozzo on its NovaScale blade servers, as well as on larger Itanium-based SMP servers and Xeon-based EXPRESS5800 servers.

Table 1. Operating System Virtualization Options

Function	Operating Systems Supported	Hardware Platforms Supported
Sun Solaris Containers	Solaris	SPARC, Intel X86, AMD-64, EM64T
SWsoft Virtuozzo	Linux, Windows	Intel X86, Itanium, AMD-64 (32-bit now; full 64-bit support in August 2005)

Operating system virtualization functions potentially have several advantages over other server virtualization options such as virtual machines and partitions. These benefits of the operating system virtualization approach include the following:

- » *Lower management costs* – Administrators do not have to install and maintain a separate instance of a complete operating system for each workload. Clearly some aspects of the virtual operating system will have to be customized, but most operating system virtualization solutions provide tools to help automate this customization. Having fewer operating system instances potentially lowers the cost of management per workload, since far more workloads can be managed on a single operating system instance.
- » *Rapid response/better performance* – Workloads running in virtual operating system instances can be started and stopped rapidly, since it is not necessary to boot an entire operating system to initiate a task. Also, because workloads consume almost only the resources needed by their applications, resources can be allocated to them more precisely based on changing business needs. Without multiple layers of operating systems, the virtual system's performance is similar to the performance on native server systems. The recovery in disaster recovery situations is much faster as well.
- » *Superior utilization* – The relatively small footprint of virtual operating system instances allows far more workloads to be virtualized on a single server. For example, in testing, Sun has been able to host as many as 4,000 Solaris Containers per system. SWsoft's Virtuozzo for Linux has hosted up to 5,000 VPSs on x86 32-bit systems, and over 10,000 on the Itanium edition.
- » *Limited system software license costs* – When operating systems that incur licensing fees are virtualized, users do not incur increased costs for additional workload instances due to licensing. For example, when operating systems such as Windows are deployed in virtual machines, or Red Hat Linux is deployed on IBM p5 partitions, each new instance requires the purchase of an additional license. With operating system virtualization, only one OS is actually installed per server, regardless of the number of virtual operating system instances that are created, limiting the licensing fees incurred for system software. Note, though, that there is currently little consistency among ISVs with regard to licensing terms for running higher-level software products, such as middleware, applications, and database servers in virtualized environments.<sup>2</sup> Indeed, the licenses for third-party software packages may vary considerably in terms of how their providers treat virtualized operating system instances. For example, Oracle treats Solaris Containers as if they were partitions; therefore, users would only have to license usage of Oracle's software based on the number of CPUs that were actually utilized by a Container, rather than the capacity of the entire server on which the Containers are hosted.
- » *Fine-grained resource control* – Because operating system virtualization is essentially a variation of resource management, with additional mechanisms to isolate workloads from each other, it benefits from the same fine-grained control over resource allocation as traditional resource management software. Such control enables metering tools to be used to precisely allocate resources to virtual operating system instances, making it easier for users to build cost-driven models for provisioning resources.

Note that operating system virtualization also has some limitations compared to virtualization approaches that involve separate operating system instances, such as partitions and virtual machines. For example, if an application requires any customizations to the operating system environment beyond those that can be captured in the virtual operating system instance (e.g., a unique device driver or a kernel patch, or a dependency on different versions of the operating system), it may not be suitable for virtualization using this approach. Also, as with resource management tools and virtual machines, all workloads deployed in virtualized operating systems depend on the availability of a host operating environment, which carries the potential to become a single point of failure. If one workload suffers a failure that is catastrophic enough to compromise the host, all workloads are potentially put at risk.

As a result, developers of virtualized operating system solutions generally try to protect the integrity of the host operating system as much as possible. For example, SWsoft has taken steps to minimize the host's exposure to failure by inserting a kernel abstraction layer, which mediates all traffic between the virtualized operating systems and the host's software kernel. Sun continues to deliver on its long-standing commitment to drive mainframe-grade reliability functions into Solaris 10 that help maintain uptime. For example, a new "predictive self-healing" architecture in Solaris 10 is designed to detect and respond to hardware and software failures with a minimum of downtime. Based on a new centralized engine for managing system services and system errors, the architecture uses kernel telemetry, the Solaris Fault Manager, the Solaris Services Manager, and enhanced error reporting to minimize downtime from failures in hardware components and software by quickly migrating workloads away from areas of failure.

Table 2 summarizes the tradeoffs between operating system virtualization and other server virtualization approaches.

Table 2: Tradeoffs between Operating System Virtualization and Other Server Virtualization Techniques

	Resource Mgmt Software	Virtual Machines	Partitions	OS Virtualization
Private File Space per Workload	No	Yes	Yes	Yes
Private Network Address per Workload	No	Yes	Yes	Yes
Private Security Profile and Root Password per Workload	No	Yes	Yes	Yes
Distinct OS Kernel Version per Workload	No	Yes	Yes	No
Distinct Kernel Patches per Workload	No	Yes	Yes	No
Granularity of Resource Allocation	Fine	Moderate	Coarse	Fine
Number of Complete Operating Systems to Manage	Single	Multiple	Multiple	Single
System Software Licenses Required	Single*	Multiple*	Multiple*	Single*
Host OS Is Single Point if Failure	Yes	Yes	No	Yes
Startup Time for Virtual Environment	Fast	Moderate**	Slow	Fast
Isolation between Workloads	Weak	Strong	Very Strong	Strong

\* For operating systems that incur commercial licensing fees, i.e. Windows. Solaris 10 no longer has licensing fees, and costs of Linux distributions in virtualized environments vary by product and platform.

\*\* Startup times for virtual machines can be reduced by maintaining and customizing images of running systems.

Based on these tradeoffs, operating system virtualization is most suitable in the following situations:

- » A high degree of commonality exists across instances (i.e. workloads are similar enough that they have the exact same requirements in terms of kernel release and patch levels).
- » Administrators want to maximize their control over resource allocation; they are concerned with the precision at which they can allocate resources to workloads, monitoring consumption of those resources, and maximizing the flexibility to adjust resources on a real-time basis in response to demand fluctuations.
- » Administrators want to maximize the utilization of hardware resources, allocating as many instances as possible on a single server.
- » Administrators want to minimize the cost of managing the overall environment by avoiding as much overhead as possible for managing each instance.
- » Administrators want to minimize the impact of any relevant system software licensing costs.

Perhaps the most obvious candidates for operating system virtualization are the workloads of Internet Service Providers (ISPs) that offer websites for consumers and small businesses. These users typically require only basic services such as e-mail and low-traffic content, which are largely undifferentiated and work well in the kind of environment that virtual operating system instances provide. ISPs are also notoriously cost-sensitive, and would value the ability to exploit the operational efficiency and software licensing benefits of operating system virtualization in order to drive up their margins. But as users in other industry segments are becoming aware of the differentiation that operating system virtualization offers over alternatives such as virtual machines, new kinds of applications are starting to emerge, including some that target more sophisticated enterprise requirements.

For example, one scenario that has emerged for the use of operating system virtualization relates to the deployment of secure remote management consoles. Server applications increasingly depend on web-based front ends for managing and monitoring their operations. The intention is that these applications should be able to be managed from any location using a standard web browser. However, users that require high levels of security, such as government or financial services organizations, cannot afford the risk of granting remote access to these management interfaces from any workstation or notebook computer on the network. Instead, administrators can set up a virtual browser session running within a virtual operating system for each application that has to be managed remotely (which may run in the hundreds or thousands). Secure access to these virtual sessions can be granted through the remote graphics capabilities and native security mechanisms built into the UNIX and Linux operating systems. This scenario is particularly compelling for Solaris Containers, given the presence that Sun has long established with users in government and financial services environments, and the extensive investments it is making in enhancing the security capabilities in Solaris 10.

## The IDEAS Bottom Line

Among the tools that are currently available for implementing virtualization, virtual machines clearly retain the lead in terms of deployment maturity. Virtual machines have penetrated the mainstream, and they are now widely used for a variety of applications in both test/development and production environments. However, in spite of their popularity, virtual machines do not necessarily solve a problem that users are becoming increasingly concerned with, which is the burden of installing and updating large numbers of operating system instances. Although virtual machines have the ability to reduce the number of physical servers that users have to manage, they do not generally reduce the absolute number of operating systems that must be managed. By contrast, operating system virtualization tools such as Solaris Containers and SWsoft Virtuozzo provide many of the same benefits as virtual machines, but avoid much of the overhead associated with managing each virtual server. For appropriate workloads, operating system virtualization tools can help reduce the cost and complexity of managing the overall environment in which multiple consolidated workloads are hosted.

### Americas

Ideas International, Inc.  
222 Grace Church Street  
Port Chester, NY 10552  
USA  
Tel + 1 914 937 4302  
Fax +1 914 937 2485

### Asia/Pacific and Worldwide Headquarters

Ideas International Limited  
Level 3  
20 George Street  
Hornsby, NSW, 2077  
Australia  
Tel +61 2 9472 7777  
Fax +61 2 9472 7788

### Europe, Middle East, Africa

Ideas International Europe  
1 Deanes Close  
Steventon  
Oxon OX13 6SZ  
United Kingdom  
Tel +44 (0) 1235 437 850  
Fax +44 (0) 1235 437 851

[www.ideasinternational.com](http://www.ideasinternational.com)



- 
- <sup>1</sup> This year, HP plans to introduce virtual machine technology for its Integrity servers that will allow multiple partitions to be deployed per CPU, increasing the granularity at which resources can be assigned to partitions.
  - <sup>2</sup> See *Technology Trends*, "IBM Seizes Initiative on Utility Pricing for Software," Ideas International, Inc., June 14 2005.