# Xen™: Enterprise Grade Open Source Virtualization

## Inside Xen™ 3.2

## A Xen™ White Paper

Virtualization has become a key requirement for the enterprise. This results from an urgent need to focus on reduced total cost of ownership (TCO) for enterprise computing infrastructure. In spite of – or indeed because of – the widespread adoption of relatively cheap, industry standard x86-based servers, enterprises have seen costs and complexity escalate rapidly. In addition to the capital expenditures to purchase them, each server in the data center costs an enterprise about an additional $10,000 per year to run, with costs including provisioning, maintenance, administration, power, real-estate, and hardware and software licenses.

At the same time, most servers today run at less than 15% utilization, meaning that most server capacity is wasted. Operating system (OS) virtualization, a concept pioneered by IBM in 1972 on the System 360, has become a key requirement to alleviate this problem because it enables increased utilization, allowing multiple operating system and application images to share each server, cutting both per-server capital cost and the overall lifetime operational costs. If every physical server can host multiple virtual servers, the overall required number of servers is reduced, and with that reduction come increased utilization, reduced complexity, and lower total cost of ownership.

Virtualization also offers many additional benefits – including development, staging and testing, dynamic provisioning, reduced patching, no-downtime maintenance, high availability and load balancing. Further, virtualization opens the door for enterprises to realize the promise of utility computing. But today's virtualization offerings are crippled by poor performance, lack of security and scalability, and an inability to offer the fine-grained resource guarantees that are required to provide true application level SLAs and support high availability. To get high performance virtualization on today's modern hardware takes a new kind of software: a low level, minimal, secure and high performance virtualization software layer called a hypervisor.

The Xen open source hypervisor, created by the XenSource founders and now developed collaboratively by over 20 major enterprises working in the Xen project, is leading the industry in high performance enterprise virtualization. Moreover the Xen hypervisor is the industry's first virtualization software to fully support Intel's VT-x hardware virtualization support.

This white paper provides an overview of the Xen architecture, and describes the features of the latest version, Xen 3.2.

### Virtualization: The New Infrastructure Requirement

The need for OS level virtualization has arisen as a result of a strange coincidence of market forces. First, enterprise software application architectures have become complex, multi-threaded, multi-process and multi-tiered systems that are difficult to provision, configure and manage. Second, the adoption of so-called "scale-out" computing infrastructure based on inexpensive industry-standard servers has led to server sprawl in the data center. Frequently IT staff provision one application per server, because that is the easiest way to ensure that the application and its configuration state can be isolated from other applications in the data center. Moreover, it provides a simple model for dealing with reliability and servicing – if the server fails, only the single application it hosts will fail, and if the application must be protected against downtime during server maintenance or from faults, then it is relatively straightforward to ‘clone’ the entire state of a server, and copy it to an identical machine that can be brought into service to replace the system that went offline. Finally, provisioning resources at the server level provides a way to identify the true resource needs of an application. If multiple applications share a single server it is difficult to determine the real resource needs of each, and to provision additional resources as needed.

Of course, serious drawbacks result from the apparent convenience of tying applications to the physical infrastructure. First, since most applications demand less than the full capacity of the server, the CIO will quickly find that most servers are severely under-utilized (typically today, with the powerful capabilities of modern 2- or 4-way servers, utilization figures are about 10-15% per server) . Of course, each server consumes full rack space and a full power load, and therefore requires cooling to match. And every year it also costs about five times the initial hardware cost to run and maintain. The net result: proliferation of under-utilized and expensive servers. Finally, the true benefits of scale-out computing are placed firmly out of reach: easy maintenance, “dial-up/dial-down” provisioning of additional resources in response to the dynamically changing resource requirements of different applications, support for high availability and remote standby and handoff, and an ability to easily develop, test, stage and rapidly provision new applications across distributed data centers are all impossible without the help of OS virtualization.

**What Virtualization Enables**

OS virtualization is achieved by inserting a layer of software between the OS and the underlying server hardware. This layer is responsible for allowing multiple OS images (and their running applications) to share the resources of a single server. Each OS believes that it has the resources of the entire machine under its control, but beneath its feet, the virtualization layer, or hypervisor, transparently ensures that resources are properly shared between different OS images and their applications.

In OS virtualization, the hypervisor must manage all hardware structures, such as the MMU, I/O devices, DMA controllers and the like, to ensure that each OS, when running, has a consistent view of the underlying hardware. There are several ways to achieve this. The simplest method, with the worst performance, is to provide a software emulation layer of the underlying chipset. This imposes very severe performance overheads, restricts the technique to a single chipset architecture (such as x86), and involves patching the kernel of a running operating system dynamically, to prevent it from modifying the state of the hardware. The additional overhead that is required to manage the hardware state for the OS, and to present to it an emulated chipset abstraction causes a significant performance overhead, frequently as much as 30-50% of the overall system resources. This binary patching technique is the core technology used by most virtualization software available today.

Virtualization software vendors today charge a hefty premium (multiples of the server cost) for their software, to which must be added the usual OS and application costs. But while today’s virtualization products have allowed enterprises to realize significant benefits in the development, testing and QA of n-tier applications, a very high performance hypervisor is a requirement for data-center wide, production-grade server consolidation and to realize the promise of a more dynamic IT infrastructure. The award winning Xen open source hypervisor, created and maintained by XenSource, is now delivering the benefits that enterprises demand from virtualization software, because it outperforms existing hypervisors by an order of magnitude while providing guaranteed service levels to each guest OS. Moreover, Xen technology is secure by design, runs across chipset architectures from IBM, Intel, AMD and Sun, supports all operating systems, and is freely available as open source software. Consequently the Xen hypervisor is being strongly endorsed by major industry players as the right way to achieve the benefits of enterprise virtualization.

**What the Xen hypervisor is:**

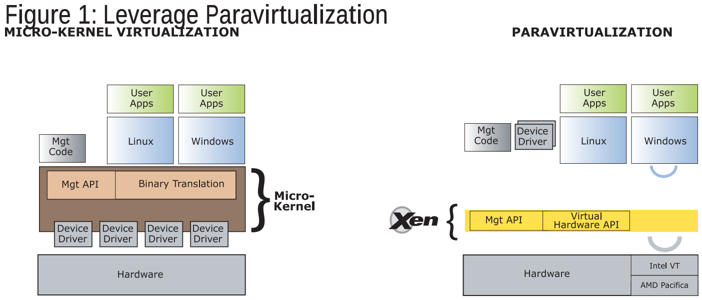
- The industry’s fastest and most secure hypervisor.
- Open source software, collaboratively developed by over 20 of the world’s leading enterprise infrastructure vendors, yet backed by an enterprise quality and service focused vendor, Citrix, which ensures the quality and reliability enterprises demand from their virtualization vendor.
- A common, open industry standard code base that supports all operating systems with high performance and security.

- Open for innovation with the addition of additional value added features by Xen ecosystem vendors, and for distribution by OEMs and operating system vendors alike.
- Available free for download at www.getxen.org and www.xensource.com.

### Xen technology: The Best in Virtualization, for Free

The Xen hypervisor uses a new and powerful virtualization technology invented by the Xen technology creators, called paravirtualization. In paravirtualization, the guest OS and hypervisor collaborate closely to achieve optimal performance. (See figure 1.)

Figure 1: Leverage Paravirtualization




The paravirtualized Xen hypervisor provides the best approach to virtualization for

- CPUs (and indeed multiple cores per device)
- MMU and memory
- I/O devices
- Asynchronous events, such as interrupts

### *CPU & Memory Virtualization*

In Xen paravirtualization, virtualization of CPU, memory, and low level hardware interrupts are provided by a low level, extremely efficient hypervisor layer that is implemented in under 50,000 lines of code. When the OS updates hardware data structures, such as the page table, or initiates a DMA operation, it collaborates with the hypervisor by making calls into an API that is offered by the hypervisor. This, in turn, allows the hypervisor to keep track of all changes made by the OS, and to optimally decide how to manage the state of hardware data structures on context switches. The hypervisor is mapped into the address space of each guest OS, meaning that there is no context switch overhead between any OS and the hypervisor. Finally, by co-operatively working with the guest OSes, the hypervisor gains additional insight into the intentions of the OS, and can make the OS aware of the fact that it has been virtualized. This can be a great advantage to the guest OS – for example the hypervisor can tell the guest that real time has passed between its last run, and its present run, permitting it to make smarter re-scheduling decisions to appropriately respond to a rapidly changing environment.

A recent innovation in hardware, namely hardware support for virtualization of CPU, MMU and memory, has enabled bare-metal performance for operating systems that cannot use the Xen hypercall API to virtualize CPU and memory. Microsoft Windows, legacy Linux and other operating systems can still benefit from the Xen hypervisor virtualization on Intel VT-x hardware virtualization support, which is now shipping on client and server systems. AMD's Pacifica hardware virtualization support adds similar capabilities to the AMD processor line, in a way that is compatible with the Xen technology. The support for hardware virtualization enables the Xen hypervisor to virtualize all operating systems "on the bare metal" but does not provide high performance I/O virtualization. This is achieved across all operating systems in a common way, by extending Xen's paravirtualization into the guest operating systems' device drivers.

### *I/O Virtualization*

Paravirtualization provides significant benefits in terms of I/O virtualization. In the Xen product, I/O is virtualized by using only a single set of drivers for the entire system (across all guests and the hypervisor), unlike emulated virtualization in which each guest has drivers of its own, and the hypervisor has yet another set of drivers for its own use.

In each Xen hypervisor guest, simple paravirtualizing device drivers replace hardware-specific drivers for the physical platform. Paravirtualizing drivers are independent of all physical hardware, but represent each type of device (e.g.: block I/O, Ethernet, USB). These drivers enable very high performance, virtualization safe I/O to be accomplished by transferring control of the I/O to the hypervisor, without any complexity in the guest. Moreover, in the Xen architecture the drivers run outside the base hypervisor, at a lower level of protection than the core of the hypervisor itself. In this way the hypervisor can be protected from bugs and crashes in device drivers (they cannot crash Xen technology), and can make use of any device drivers available on the market. Also, the virtualized OS image is much more portable across hardware, since the low levels of the driver and hardware management are modules that run under control of the hypervisor.

The net result is that the Xen product offers superb performance – typically over an order of magnitude faster than any hypervisor on the market.

### Virtualization with Xen 3.2 and the Promise of Utility Computing

For all the potential benefits of virtualization and utility computing, few enterprises have yet managed to achieve the levels of performance and support for a broad range of software and hardware that they desire. Xen 3.2 fulfills that need.

With a high performance hypervisor, it is possible to deliver on many of the key demands of major enterprises for an adaptive, responsive IT architecture. Xen 3.2 supports a very wide range of hardware platforms and processor architectures, including x86, IA64 and Power 5. As a result, Xen guest OSes can run on a wide variety of hardware. Xen 3.2 supports up to 32-way SMP guest operating systems, a key requirement for applications that require multiple CPUs to deliver enterprise performance. Xen 3.2 supports 32 bit, 64 bit and Physical Address Extensions (PAE) addressing modes. Xen technology also offers a capability for live VM migration, in which a running guest OS, in its virtual machine, is moved to a second machine in a very short time. While existing products on the market today claim live migration as a feature, they typically cause the migrated application to be unresponsive for tens of seconds while it is moved. Using the Xen product, with a feature that enables "copy on write" for guest OS pages, the "downtime" is typically 30-60ms, orders of magnitude faster than available using other vendors' approaches today.

With these raw capabilities, Xen technology is ideally positioned to allow major enterprises to realize the promise of utility computing. The Xen hypervisor moves the level of infrastructure up above the basic hardware, by providing a common, low-level, high speed set of execution primitives that can be used to provide a dynamic and responsive computing environment.

### The Need for an Open Source Industry Standard Hypervisor

Today, several hypervisors are available on the market. None are free, and all are closed and tied into expensive, proprietary software stacks. To fully take advantage of current and future virtualization features, the best technology should be widely adopted by the market. In addition, major enterprises want a virtualization layer that is not tied to any one OS, and that offers the best performance. The Xen hypervisor fulfills the need for an unencumbered virtualization standard, and offers an opportunity to all players to take advantage of the massive trend towards dynamic datacenter management.

The Xen product is an open source project, run under the open source community rules. By virtue of its availability, and because it offers the best virtualization technology available, it is a natural candidate as the broadly adopted "standard" hypervisor. The open source community has embraced the Xen hypervisor as offering both the right technology – through its paravirtualization approach and extremely high performance – and lack of bias towards any chip architecture, operating system or application vendor.

***About XenSource***

*XenSource, Inc. a Citrix company, develops enterprise grade virtualization infrastructure solutions based on the industry's highest-performance virtualization technology, the open source Xen hypervisor. Founded and run by the original Xen development team, XenSource products allow enterprises to realize the TCO savings that result from server consolidation, increased utilization, and reduced complexity in the datacenter. For more information, visit* ***www.xensource.com****.*

Delivering the Power of Xen

***www.xen.org***
*xen-community@lists.xensource.com*