WHITE PAPER

Understanding Full Virtualization, Paravirtualization, and Hardware Assist

Contents

Introduction

In 1998, VMware figured out how to virtualize the x86 platform, once thought to be impossible, and created the market for x86 virtualization. The solution was a combination of binary translation and direct execution on the processor that allowed multiple guest OSes to run in full isolation on the same computer with readily affordable virtualization overhead.

The savings that tens of thousands of companies have generated from the deployment of this technology is further driving the rapid adoption of virtualized computing from the desktop to the data center. As new vendors enter the space and attempt to differentiate their products, many are creating confusion with their marketing claims and terminology. For example, while hardware assist is a valuable technique that will mature and expand the envelope of workloads that can be virtualized, paravirtualization is not an entirely new technology that offers an "order of magnitude" greater performance.

While this is a complex and rapidly evolving space, the technologies employed can be readily explained to help companies understand their options and choose a path forward. This white paper attempts to clarify the various techniques used to virtualize x86 hardware, the strengths and weaknesses of each, and VMware's community approach to develop and employ the most effective of the emerging virtualization techniques. Figure 1 provides a summary timeline of x86 virtualization technologies from VMware's binary translation to the recent application of kernel paravirtualization and hardware-assisted virtualization.

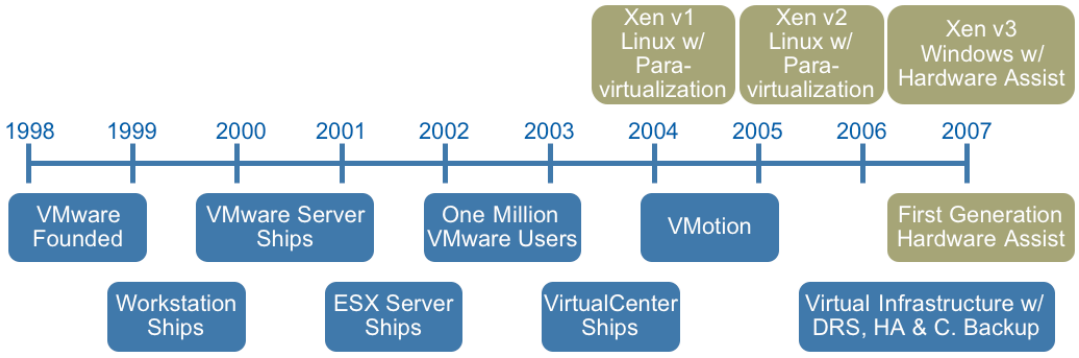

Figure 1 – Summary timeline of x86 virtualization technologies

Overview of x86 Virtualization

The term virtualization broadly describes the separation of a service request from the underlying physical delivery of that service. With x86 computer virtualization, a virtualization layer is added between the hardware and operating system as seen in Figure 2. This virtualization layer allows multiple operating system instances to run concurrently within virtual machines on a single computer, dynamically partitioning and sharing the available physical resources such as CPU, storage, memory and I/O devices.

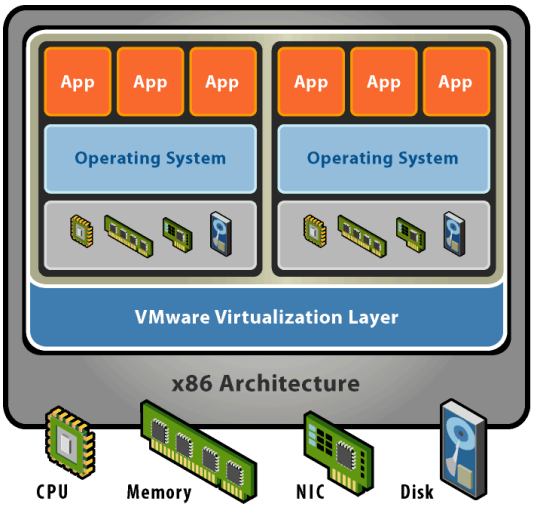

Figure 2 – x86 virtualization layer

As desktop and server processing capacity has consistently increased year after year, virtualization has proved to be a powerful technology to simplify software development and testing, to enable server consolidation, and to enhance data center agility and business continuity. As it turns out, fully abstracting the operating system and applications from the hardware and encapsulating them into portable virtual machines has enabled virtual infrastructure features simply not possible with hardware alone. For example, servers can now run in extremely fault tolerant configurations on virtual infrastructure 24x7x365 with no downtime needed for backups or hardware maintenance. VMware has customers with production servers that have been running without downtime for over three years.

For industry standard x86 systems, virtualization approaches use either a hosted or a hypervisor architecture. A hosted architecture installs and runs the virtualization layer as an application on top of an operating system and supports the broadest range of hardware configurations. In contrast, a hypervisor (bare-metal) architecture installs the virtualization layer directly on a clean x86-based system. Since it has direct access to the hardware resources rather than going through an operating system, a hypervisor is more efficient than a hosted architecture and delivers greater scalability, robustness and performance. VMware Player, ACE, Workstation and Server employ a hosted architecture for flexibility, while ESX Server employs a hypervisor architecture on certified hardware for data center class performance.

To better understand the techniques employed for x86 virtualization, a brief background on the component parts is useful. The virtualization layer is the software responsible for hosting and managing all virtual machines on virtual machine monitors (VMMs). As depicted in Figure 3Figure 3, the virtualization layer is a hypervisor running directly on

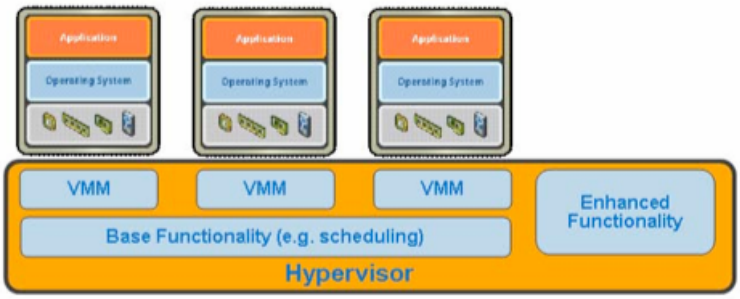

Figure 3 – The hypervisor manages virtual machine monitors that host virtual machines

the hardware. The functionality of the hypervisor varies greatly based on architecture and implementation. Each VMM running on the hypervisor implements the virtual machine hardware abstraction and is responsible for running a guest OS. Each VMM has to partition and share the CPU, memory and I/O devices to successfully virtualize the system.

CPU Virtualization

The Challenges of x86 Hardware Virtualization

X86 operating systems are designed to run directly on the bare-metal hardware, so they naturally assume they fully 'own' the computer hardware. As shown in Figure 4, the x86 architecture offers four levels of privilege known as Ring 0, 1, 2 and 3 to operating systems and applications to manage access to the computer hardware. While user level applications typically run in Ring 3, the operating system needs to have direct access to the memory and hardware and must execute its privileged instructions in Ring 0. Virtualizing the x86 architecture requires placing a virtualization layer under the operating system (which expects to be in the most privileged Ring 0) to create and manage the virtual machines that deliver shared resources. Further complicating the situation, some sensitive instructions can't effectively be virtualized as they have different semantics when they are not executed in Ring 0. The difficulty in trapping and translating these sensitive and privileged instruction requests at runtime was the challenge that originally made x86 architecture virtualization look impossible.

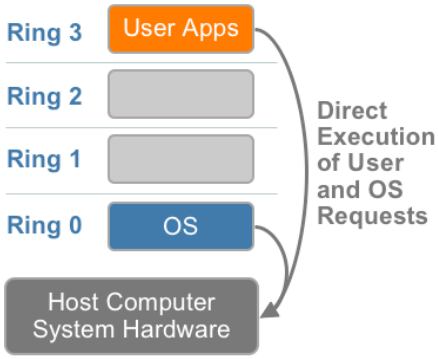

Figure 4 – x86 privilege level architecture without virtualization

VMware resolved the challenge in 1998, developing binary translation techniques that allow the VMM to run in Ring 0 for isolation and performance, while moving the operating system to a user level ring with greater privilege than applications in Ring 3 but less privilege than the virtual machine monitor in Ring 0. While VMware's full virtualization approach using binary translation is the de facto standard today based on VMware's 20,000 customer installed base and large partner ecosystem, the industry as a whole has not yet agreed on open standards to define and manage virtualization. Each company developing virtualization solutions is free to interpret the technical challenges and develop solutions with varying strengths and weaknesses.

As clarified below, three alternative techniques now exist for handling sensitive and privileged instructions to virtualize the CPU on the x86 architecture:

- Full virtualization using binary translation
- OS assisted virtualization or paravirtualization
- Hardware assisted virtualization (first generation)

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.