# WHITE PAPER

## Server Refresh: Meeting the Changing Needs of Enterprise IT with Hardware/Software Optimization

Sponsored by: Intel Corporation and Oracle

Jean S. Bozman Katherine Broderick

July 2010

## EXECUTIVE SUMMARY

The installed base of server systems has grown rapidly over the past 10 to 12 years, driven primarily by x86-based system growth. According to IDC's Worldwide Quarterly Server Tracker, by 2009, x86-based servers accounted for more than 95% of the server unit shipments and for more than 50% of the revenue in the server market.

Following a year of project delays and deferrals in 2009, many IT organizations are planning technology refresh for their aging x86 servers — especially those that are more than three years old. Many net-new technologies have become available in the marketplace over the past three years — and IT organizations will benefit in terms of IT efficiency, power/cooling, performance, and scalability that simply were not shipping in 2007 and earlier. New capabilities in x86 processors will provide support for more demanding enterprise workloads and for larger groups of end users.

There are many drivers for the refresh wave that is expected worldwide; for example, more powerful multicore processors as well as denser configurations of servers with more memory, higher I/O, and faster onboard bandwidth to speed overall system throughput. These technical improvements support workload consolidation — allowing sites to gather up workloads that had been running on large numbers of older x86 servers. To the extent that consolidation takes place, power/cooling costs can be reduced because more powerful workloads are running on fewer individual server footprints.

There are many reasons to refresh servers, including replacing systems that are no longer functional, meeting the compute demands required by new business requirements, or saving energy and cutting costs with new, more efficient systems, or a combination of all of the reasons. The technology available in servers has changed rapidly, giving end users increased ability to build more flexible and efficient environments. The hardware is more powerful and efficient, and a series of capabilities on x86 processors, such as enhanced support for virtualization, have improved users' ability to maximize the use and flexibility of their servers and to consolidate workloads on them. Some of the drivers for change are underutilized systems, space and power constraints, operating expenses, and business growth. Details are as follows:

- Systems that are underutilized are being looked at to understand how to make better use of server investments. One solution is to increase usage through virtualization and consolidation. This is frequently accomplished by refreshing with more capable systems while consolidating on fewer systems. This results in fewer systems with a smaller footprint and lower energy costs.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

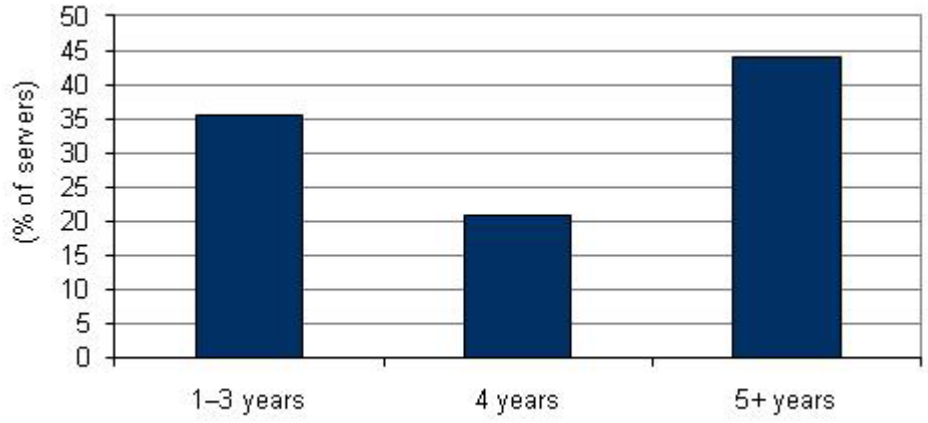
- Space and power issues may be driven by limited availability of power for the datacenter, budgetary constraints, or a plan to slow the growth of operational costs.
- Operating expenses may be the result of proliferation of systems or system types, or both. Server consolidation helps to reduce the overall server count by reducing the number of physical server "footprints" while continuing to host all of the workloads via virtual machines (VMs). Typically, virtualization/consolidation projects are planned as a means to reduce operational costs for the business.
- Business growth and competitive pressures are creating a need for more compute resources to support increased usage and new applications.

## SITUATION OVERVIEW

IDC believes that more than 32 million servers are installed worldwide today and expects that number to grow to more than 41 million servers by 2013 for all processor types. The age and usefulness of these systems vary widely. According to a recent IDC survey, 35% of installed systems are one to three years old, 21% are four years old, and 44% are five years or older (see Figure 1).

**FIGURE 1**

Length of Time Servers Are Deployed




Source: IDC, 2010

### Why Refresh?

The top questions that IT organizations face when thinking about their installed base of servers are "when" and "how" to refresh those servers. IT organizations are constantly evaluating their ability to meet the needs of their users and balancing their existing infrastructure against those needs. The demand for compute resources is

constantly increasing in most businesses today as the number of Web-enabled workloads increases, as applications are being accessed by broader populations of end users, and as databases grow over time.

There is a long list of potential business issues that may need to be addressed. For example:

- Update/retire obsolete systems
  - Need to replace old systems that are no longer supported and/or servers that lack the performance needed to run newer applications
- Workload growth
  - Support new applications
  - Increased workload on existing applications exceeds the capabilities of the current systems
  - Consolidated workloads
- Operating efficiency improvement
  - Improve operating efficiencies — such as ease of manageability, scalability, sustainability, and technology change management — to increase output and control operating expenses
  - Need for higher reliability to meet end-user SLA requirements
  - Decrease time to deploy applications and application services
- Environmental control
  - Reduce or control environmental issues such as power, cooling, and space limitations
- Cost control
  - Reduce capital expenditure costs

## Common Approaches

When demand increases, or when changes occur in the structure of a given business, IT organizations may use several different approaches to meet the new requirements. Common tactics include replacing or upgrading existing systems, adding new systems, consolidating or restructuring application usage, or outsourcing applications, or a combination of these choices. The numerous choices available to IT organizations result in a multitude of options to consider when organizations decide to transform their systems and operations and to update or refresh the server technologies that support the business.

### *Replacement/Refreshed Servers*

Servers are replaced or refreshed for many different reasons. Some of the reasons, or drivers, for server refresh that are frequently cited by respondents to IDC surveys are:

- Servers are no longer viable or desirable for upgrades or need additional capability that is not available through an upgrade.
- Existing system is being retired because of nonsupport or maintenance issues.
- The system is unable to support the growth of the current application and will be repurposed to run another application.
- To meet consolidation requirements, a server with higher performance, more memory capacity, and increased I/O capabilities is required.
- Restructuring, using virtualization to gain better utilization and flexibility, exceeds the system's capabilities.
- Power and cooling constraints and/or goals require more efficient systems.
- Space constraints require servers with "smaller footprints."
- Operational efficiency requires more efficient server systems.

### *Upgrades*

The following list highlights some of the criteria that IT organizations should consider when deciding whether or not to upgrade rather than purchase an additional server or a replacement server. An upgrade is a viable option when all of the following are true:

- The system will have sufficient life span to support new and/or additional workloads with an upgrade.
- There is available space for the additional processors, memory, and I/O required for an upgrade.
- The upgraded system will be able to support the added workload requirements.
- New workloads are compatible with existing workloads already running on the server either with or without an upgrade to the existing system.
- There will be sufficient headroom and flexibility to handle potential usage spikes and changes.

### *Outsourcing Applications*

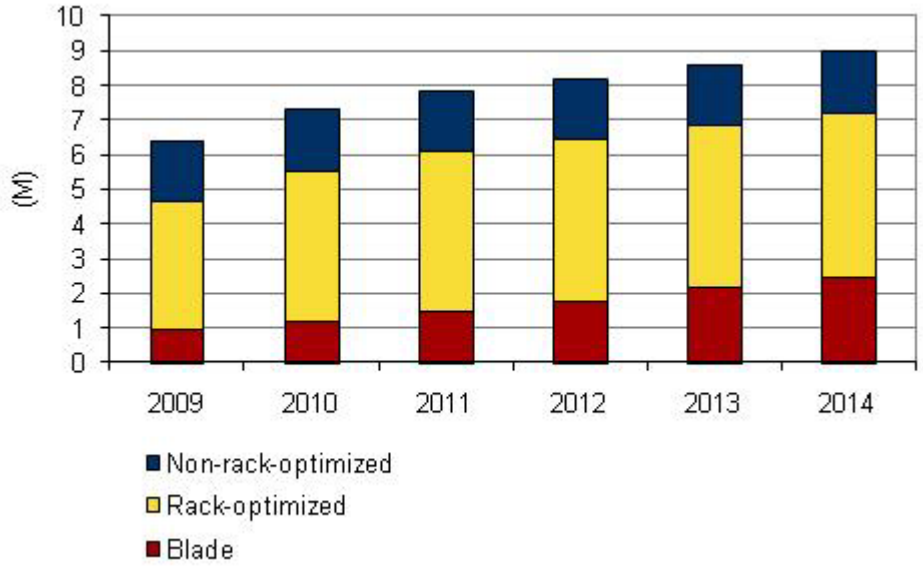
Outsourcing runs from single applications to complete IT operations. Examples of two commonly outsourced applications are payroll and HR. This paper focuses on decision making for internal IT infrastructure and as such does not attempt to answer questions relative to outsourcing decisions.

## The x86 Server Market

This paper focuses on Oracle's x86-based Sun server line using Intel Xeon processors. The x86 server market is expected to increase from 2009 to 2014 at a 20.4% CAGR in terms of unit shipments. According to IDC's Worldwide Quarterly Server Forecast, the year-over-year growth rate for x86 servers from 2009 to 2010 is projected to increase dramatically, with 100.35% growth over that one year alone. Much of this growth can be attributed to the aftereffects of the worldwide economic recession of 2009, but a fair amount of the expected growth is a reflection of the changing requirements in the market. The general growth indicates an increase in computing requirements. The growth in smaller, more compact form factors points to the change in space, power, and manageability requirements. An examination of form factors shows that blades are expected to increase at a 20.8% CAGR from 2009 to 2014, which is almost three times the total market growth over the same period. Figure 2 shows that rack-optimized and non-rack-optimized servers are growing slightly. In essence, blades represent most of the unit growth in the x86 server market.

**FIGURE 2**

x86 Server Forecast, 2009–2014




Source: IDC's Worldwide Quarterly Server Forecast, June 2010

### *Addressing Operational Problems*

It's an easy decision to retire an outdated, underperforming, and unsupported system. The more important decision is what the replacement should be and what criteria should be used to select the new system. In some cases, the systems may not meet those requirements, making the decision easy, but they may still need to be replaced with higher-performing, more efficient and flexible up-to-date systems. The following are some of the potentially key reasons to refresh a server (sample problems and solutions are also noted):

- Boost application performance and responsiveness
  - **Problem:** Borderline or unacceptable application speed and responsiveness
  - **Solution:** A new server will typically be multiple times faster than an older system. Server technology continues to advance rapidly and has increased server performance more than four times over the past three to four years while maintaining the same cost structure.
- Contain or lower costs plus reduce server footprint
  - **Problem:** Higher energy consumption and cost, plus decrease datacenter space requirements; or increase compute capacity in the same space while maintaining or lowering energy requirements.
  - **Solution:** New, more energy-efficient systems in smaller form factors utilizing lower-power multicore processors provide more performance in a smaller space while decreasing the power required to perform an operation.
- Create a flexible, manageable environment for consolidation and better operations
  - **Problem:** Maximize flexibility, increase utilization, and reduce management overhead.
  - **Solution:** Virtualization capabilities provide a platform for getting more flexibility and utilization out of servers. New servers are generally designed to be virtualized with virtualization hardware acceleration built into the Intel silicon. In addition, they have a more balanced system to eliminate potential bottlenecks while running highly utilized systems with multiple applications operating simultaneously. Virtualization can be used to better control applications on a single server and to make multiple servers look alike for transparent operations. This allows flexibility in where applications run, thus making a more flexible environment.

These are a few examples of why and how server refresh should be done. The best choice is always to plan ahead in a proactive fashion. The goal should be to make changes as soon as possible to meet business needs and maximize resources.

### *An End-User Example*

When a large hosting company purchased several smaller companies, it acquired multiple datacenters with a variety of equipment. After analyzing its newly acquired assets, it decided on a plan for one of the datacenters to consolidate over 800 servers and 80 racks to a group of four Sun Blade System chassis in two racks. The company was able to virtualize the servers and load the applications onto fewer, more powerful servers. This move significantly reduced space, power, and cooling requirements and costs associated with maintenance and management complexity. This consolidation was accomplished while simultaneously increasing the workload and maintaining desired headroom for the business' changing requirements. This is an excellent example of proactively looking at resources and planning to be more efficient and cost-effective.

## ORACLE AND INTEL SOLUTIONS

In 2010, Intel launched two new high-end Xeon processors for the enterprise — the Xeon 5600 processor (formerly code-named Westmere-EP) and Xeon 7500 processor (formerly code-named Nehalem-EX) — both of which are used as the engines for Oracle's Sun line of x86 servers. This latest generation of Intel Xeon technology has been incorporated into the Sun product line and then combined with Oracle's use of flash memory, flash memory acceleration, and large storage capacity to support data-intensive workloads for the enterprise.

The Xeon 5600 was designed for energy efficiency and performance. This is particularly important in a rack-optimized server deployment, where dense packing of multiple servers had been driving up heat dissipation in the mid-2000s and increasing power costs to cool the densely populated racks of servers.

The Xeon 7500 was specifically designed to support highly demanding workloads. It includes more than 20 reliability, availability, and serviceability (RAS) features. Servers based on Xeon 7500 processors will be aimed at scalable application workloads, consolidated workloads, support for virtualization and VMs, and database processing.

### Three Years of Optimization Work for Solaris on x86

On January 22, 2007, Intel announced a formal agreement with Sun Microsystems to optimize the Solaris operating system for use on the Xeon processor product line; that product is now called Oracle Solaris. The result of that work has been to provide powerful innovation for customers that deliver compelling value to the datacenter.

That work drove toward the release of the latest generation of Intel Xeon processors, the Xeon 5600 and 7500 processors, allowing Oracle Solaris to improve performance on Xeon-based systems and to leverage all onboard features of these Xeon multicore processors. Oracle's Sun servers, therefore, will get the full benefit of all of the new RAS and performance features that are built into the Xeon processors.

The optimizations for and enabling of Oracle Solaris on Intel Xeon 7500 and 5600 processors focus on four key areas:

- **Improved performance.** Building on a proven track record, Oracle Solaris is ready to take advantage of the performance capabilities of the Intel Xeon processor 7500 and 5600 series for enterprise computing.
- **Efficient power management.** Sun (now Oracle) and Intel have developed ways to lower power utilization on Intel Xeon processor–based systems by optimizing Oracle Solaris to take advantage of Intel's built-in power management functions.
- **Increased RAS.** The Oracle Solaris Fault Management Architecture (FMA) software is being enhanced to adapt to the Intel Xeon processor's RAS features to provide a more reliable enterprise compute solution.
- **Cost-effective virtualization.** Oracle and Intel have worked together to enhance Oracle Solaris to take advantage of the latest Intel Virtualization Technology (Intel VT) features to enable the highest consolidation ratios.

Since that time, Oracle has built a family of Intel Xeon–based servers as a part of its x86-based line of servers. These systems offer end users a choice of operating system (Oracle Solaris, Linux, or Windows) on Intel Xeon processor–based systems, shipping in both rack and blade form factors. These servers will support a variety of virtualization technologies, including VMware vSphere, Microsoft Hyper-V, and Oracle Virtualization Manager (OVM). One advantage of this approach is that customers have the choice of deploying mixed workloads into a heterogeneous environment running on x86 servers. This approach allows them to manage virtualized hosts and guests, and to move them across the virtualized hardware as needed — and as business needs change.

## A Variety of Deployment Models

IDC end-user research shows that servers are used in a wide number of ways based on different usage models. For example, databases of different sizes and purposes run on every size system.

Large complex databases typically run on large servers, but they can also be run across a cluster of smaller systems. Oracle has fully recognized the need for a family of servers that meet a range of enterprise computing requirements. This document highlights Oracle's Sun line of Xeon-based servers, which represent various sizes and types of systems to meet the needs of enterprise organizations. They cover usage models from cloud-optimized entry servers to application and database servers and the need for high density versus high capacity within the same product line.

## Oracle Solaris Optimized for Intel Xeon 7500 Processors

Oracle's plan for the Sun servers is to optimize them for high performance and high availability. As x86 servers, they are being built to support the most demanding workloads in business, including enterprise applications and scalable databases. These

servers combine flash memory, large onboard storage, and large memory capacity to support data-intensive workloads. The combination itself is a form of innovation by bringing together components that, when used with Oracle software, speed performance and improve reliability for applications running on x86 servers. This approach to the market is particularly important in view of Oracle's strengths in providing infrastructure software, application software, middleware, and database software. Oracle offers tightly integrated x86 clustered systems engineered across server, storage, networking, operating system, database, middleware, business applications, and enterprise management, all of which can be managed and supported as a single system, from the application to the disk. It is an approach that will allow customers to deploy scalable x86 servers that already include the software stacks for enterprise computing.

IDC notes that, prior to the time when Oracle acquired Sun Microsystems, much work had already been done by Sun in the areas of flash memory — including the development of Sun Flash Memory devices and Sun Flash Accelerator boards. These hardware components are now being included in the Oracle x86 server solutions. By combining these components with the multicore processing capabilities of Intel Xeon processors, customers will find a combination of capabilities in a software environment that has been optimized to support demanding workloads.

The joint work that was done by Intel and Sun regarding Oracle Solaris, which officially began in 2007, is now resulting in high-performance x86 server systems based on the Intel Xeon 7500 processors. This optimized software stack provides performance and reliability enhancements, and strengthens virtualization capabilities, with support for dozens of VMs per server for the midrange models and up to hundreds of VMs per server for the high-end, eight-socket Sun x4800 x86 server model.

### *New Intel Processors for Enterprise Computing*

During the 2009–2010 time frame, Intel announced two new processors for enterprise computing: the Intel Xeon 7500 and the Intel Xeon 5600. Both of these processors are now shipping in volume for inclusion in x86 servers in the enterprise.

The 7500 series is intended for use in supporting mission-critical workloads and has more than 20 RAS features. The Xeon 7500 replaces older Intel MP designs. It is built on a 45nm process and will provide three to five times the performance of earlier Xeon processors, according to Intel. Its QuickPath Interconnect (QPI) speeds on-chip communications and throughput, while its Machine Check Architecture (MCA) isolates faults so that processing can continue.

The Xeon 5600 was designed to provide improved security for applications and improved energy efficiency, compared with earlier generations of Xeon processors. The 5600 supports a wide range of workloads including virtualization/consolidation projects in IT sites that are looking to reduce server footprints within the datacenter — and to be the engine for one-socket and two-socket systems that will replace x86 servers that are aging in place. It also has built-in AES encryption for application security, an enhanced L3 cache for each socket, and Intel Turbo Boost to ramp up clock speed. Based on a 32nm processor design, the Xeon 5600 is expected to provide performance levels up to 60% greater than those of the earlier 45nm Xeon 5500 series processors.

### *Oracle's New Sun x86 Server Models*

The following sections contain descriptions of the new models in the Oracle Sun x86 server product line.

#### Sun Blade X6270 M2

With its large memory footprint and high I/O bandwidth, Oracle's Sun Blade X6270 M2 server module is a highly capable virtualization blade for business applications and enterprise collaboration workloads. Powered by the highest-speed six-core Intel Xeon Processor 5600 series, this balanced blade leverages the unique I/O architecture of the Sun Blade 6000 chassis, making it easy to deploy and upgrade. When combined with Oracle's blade-specific storage and advanced networking products, the Sun Blade X6270 M2 server module offers high levels of performance and energy efficiency while reducing the cost and complexity of today's datacenter.

#### Sun Blade Storage Module M2

Oracle's Sun Blade Storage Module M2 is a second-generation direct-attached storage blade that incorporates SAS-2 technology to offer better manageability and performance than the previous-generation storage module. This storage module features eight front-accessible and hot-pluggable drives (with up to 2.4TB of capacity) that can be grouped together with the server modules to set up multiple zones. It leverages the highly available power, cooling, and I/O infrastructure that is provided by the Sun Blade 6000 chassis.

#### Oracle's Sun Fire X2270 M2 Server

This two-socket server is an entry-level server for network-facing Web tiers supporting enterprise and cloud computing. It houses up to two Xeon 5600 processors, and it is outfitted with up to 96GB of memory and up to four internal 3.5in. SATA drives for internal storage. It can support up to 16 PCIe I/O slots and two 1GbE connections. The server supports up to 176GB of flash memory and is managed by Oracle Integrated Lights Out Manager (ILOM) software.

#### Oracle's Sun Fire X4170 M2 Server

This two-socket server hosts infrastructure workloads, including collaboration and email workloads, and the Oracle Beehive suite of store-and-forward message software. It houses up to two Intel Xeon 5600 processors, with up to 144GB of memory, and up to eight 2.5in. SAS-2 or SATA drives for internal storage. This server has up to two flash modules, with a maximum flash capacity of 368GB, and up to four SSDs. By leveraging Oracle Solaris and the ZFS file system, the Sun Fire X4170 M2 server can store, and manage, the large amounts of data for the collaborative applications that are being processed by the system. The server is managed by Oracle ILOM software, which comes standard and integrates with Oracle's Ops Center/Enterprise Manager via plug-in software agents.

#### Oracle's Sun Fire X4270 M2 Server

This two-socket server is designed to support business applications — and to support Oracle Real Application Clusters configurations spanning multiple X4270 M2 servers. Through the use of flash memory and support for up to six PCIe I/O slots, the server

can rapidly move transactional data among multiple servers. This server houses up to two Xeon 5600 processors and up to 144GB of memory. It can house up to 12 of the 3.5in. SAS-2 or SATA drives or up to 24 of the 2.5in. SAS-2 or SATA drives. Regarding flash memory, the system has two Sun Flash modules and up to three Sun Flash Accelerator F20 PCIe cards, which combine to support Flash memory on board, speeding data transfer from outside the system for further processing. This server supports more than 1.1TB of flash, which has the result of accelerating applications performance. The server is managed by Oracle ILOM software, which comes standard with the server.

### Oracle's Sun Fire X4470 Server

This four-socket server is designed to support OLTP, transactional workloads, and application server workloads — and it is intended for use as a platform for workload consolidation. It is a 3U system that supports up to four Xeon 7500 processors. It supports up to 512GB of memory, and has 10 PCIe I/O slots, in order to provide high performance for data-intensive transactional workloads and line-of-business workloads from Oracle's Siebel CRM and Oracle E-Business Suite ERP applications. Onboard, internal storage is supported by six 2.5in. SAS-2 or SATA drives. This server makes extensive use of flash memory, with four Sun Flash modules on board, providing a maximum flash capacity of 672GB. It also has up to four Sun Flash Accelerator F20 PCIe cards. This server has hot-swappable components for high availability and reliability and is managed by Oracle ILOM software, which comes standard with the server.

### Oracle's Sun Fire X4800 Server

This eight-socket server supports scalable database workloads and workload consolidation leveraging virtualization software that is running on the x86 server. With eight Intel Xeon 7500 processors — each supporting six or eight cores per processor — this server supports up to 64 processor cores. It is also outfitted with up to 1TB of memory, and it supports eight PCIe I/O slots and eight 2.5in. SAS-2 or SATA drives. It can support up to eight 10GbE links and up to eight 1GbE links. This server has hot-swappable, hot-pluggable components for high reliability and availability and is managed by Oracle ILOM software, which comes standard with the server. This is the most scalable x86 server that Oracle will be shipping, based on its first announcement following the Sun acquisition. It is intended for use with memory-intensive applications, in-memory databases, enterprise high-performance computing (HPC), and business application consolidation (when many applications are consolidated onto a more scalable server).

As discussed earlier in this paper, these Sun x86 server systems from Oracle offer a broad set of choices and capabilities in hardware and software, including support for multiple operating systems and virtualization technologies. They can be assembled into arrays, grids, or clusters of x86 servers — allowing workloads to scale up when used in conjunction with Oracle system-level infrastructure software and systems management software.

## CHALLENGES/OPPORTUNITIES

Oracle's Sun servers will have substantial opportunity in the x86 server market because they can support demanding workloads for the enterprise and have been optimized for those computing requirements. However, because the x86 server market is a competitive one, with several top vendors, Oracle will face the challenges of price competition, frequent refresh of onboard technologies for x86 server hardware, and support for a wide variety of workloads within the enterprise.

To answer those challenges, Oracle's Sun server group has developed a comprehensive line of x86-based servers, scaling from two sockets to eight sockets. These systems will support technology refresh on an ongoing basis to allow customers to take advantage of the latest system-level innovations in hardware and software.

Because Sun was known as a technology company that was focused on Solaris (now Oracle Solaris) and SPARC processors, Oracle will need to educate the market about its strong commitment to shipping multiple operating systems on its x86 server platforms. New technology is an opportunity for IT organizations to improve the performance and services supplied to their company. Just as importantly, as a key supplier of IT systems for the enterprise, Oracle will need to explain the business value being provided by combining optimized software stacks, flash memory, and accelerators in terms that both business managers and IT managers can understand and appreciate.

## CONCLUSION

IDC estimates that the server market has grown to an installed base of more than 32 million systems worldwide. This large installed base represents an opportunity for vendors to supply the systems required to replace these systems as they need to be refreshed. Importantly, many IT projects were delayed during the economic downturn that began in late 2008 and continued through 2009. Now, new workloads are entering the server space, including workload consolidation on virtualized servers, intensive Web-based applications, collaborative workloads, and demanding applications and databases for the enterprise. All of these workloads will create new demand in the x86 server market as they come online.

The x86 server market is the largest and fastest-growing segment of the market. These Sun Intel Xeon processor–based systems are targeted directly at the market — and they will support both Oracle Solaris and Linux operating systems, along with Windows. Their form factor is focused on the rack-optimized and blade server segments, which represent the growth within the x86 server market, as seen in IDC's quarterly and annual server market reports and in IDC forecasts, by form factor.

End users have an opportunity to improve their environment with wise choices that will give them better performance, lower operating costs, better efficiency, and more flexibility. The challenge associated with this process of technology refresh is to understand their site requirements and to discover what options are available. This process should involve working with trusted partners to find technology solutions that map to their business needs.

Oracle's portfolio of x86 systems is entering the marketplace as a series of options for technology refresh when refreshing customers' systems or adding additional systems. Following on the recovery from the economic downturn, the number of systems that need to be refreshed will continue to grow over the forecast horizon. Ultimately, the presence and the availability of new hardware and software technology, when combined and deployed together at customer sites, will drive high-performing, efficient, and cost-effective systems to be deployed into the enterprise to meet real-world business needs.