IT@Intel White Paper

Intel Information Technology
Enterprise Resource Planning
March 2010

An ERP Platform Strategy Based on Industry-standard Servers

Four-socket servers provide the performance, memory capacity, I/O expandability, and proven technology required to run our larger production ERP instances.

Sudip Chahal
Principal Engineer, Intel IT

Karl Mailman
ERP Infrastructure Architect, Intel IT

Executive Overview

Many large organizations have a centralized enterprise resource planning (ERP) environment based on proprietary mainframes or RISC-based systems. In contrast, Intel IT has successfully implemented a decentralized ERP environment that is based on industry-standard servers and supports more than 10,000 active users. We have found that this approach offers several advantages, including lower server acquisition costs and increased flexibility and agility.

Four-socket servers perform essential roles within this environment, providing the performance, memory capacity, I/O expandability, and reliability required to run our larger production ERP instances. These servers also provide enough headroom for anticipated growth of large production ERP databases over the depreciation cycle of the server. We use two-socket servers for smaller production instances and a variety of non-production uses.

Compared with two-socket servers based on Intel® Xeon® processor 5600 series, four-socket servers based on Intel® Xeon® processor 7500 series have key characteristics that enable them to support more demanding ERP requirements:

- Greater performance headroom to support workload growth, demand spikes, and failover situations
- Larger memory capacity to support large ERP workloads and failover
- I/O expandability
- Additional reliability, availability, and serviceability (RAS) features to support mission-critical workloads, such as the new machine check architecture recovery (MCA recovery) capability

We therefore expect to continue to use four-socket servers to run our most demanding ERP application production instances, as well as instances used for benchmarking and disaster recovery. Because four-socket servers are well suited to act as virtualization hosts, we may also use them in the future to host vitualized ERP instances used for development, testing, and application support.

Contents

IT@INTEL

IT@Intel is a resource that enables IT professionals, managers, and executives to engage with peers in the Intel IT organization—and with thousands of other industry IT leaders—so you can gain insights into the tools, methods, strategies, and best practices that are proving most successful in addressing today's tough IT challenges. Visit us today at www.intel.com/IT or contact your local Intel representative if you'd like to learn more.

BACKGROUND

There are many large organizations that run their enterprise resource planning (ERP) applications on a small number of proprietary mainframes or RISC-based systems. This is primarily because ERP applications have historically been marketed and implemented as integrated, company-wide solutions. The perception that ERP requires a large proprietary system remains common today.

Intel has taken a different approach, eschewing a monolithic, company-wide ERP implementation. Instead, different business groups implemented ERP individually to support their own operations. Each of these decentralized implementations runs on industry-standard servers located in corporate data centers. This approach offers several advantages, including lower acquisition costs and increased flexibility and agility.

At Intel, these implementations add up to a very large ERP environment. Overall, there are about 10,000 active users, and the environment runs on approximately 250 servers.

Today, ERP suppliers increasingly deliver products as a suite of individual applications rather than as a single, integrated solution as in the past. These applications typically can run on industry-standard servers. This suggests that industry-standard servers will continue to be a good fit for ERP.

INTEL'S DECENTRALIZED ERP STRATEGY

At Intel, the sales and marketing group was the first to implement ERP to support its own operations. Other business groups followed similar approaches, and today Intel has ERP implementations for finance, materials management, warehouse management, and other functions.

This approach means that each implementation is optimized for the unique needs of the group it serves. We use middleware from the ERP supplier to share and synchronize data between these implementations.

Intel IT's Rationale for Decentralized ERP on Industry-standard Servers

Decentralized ERP on industry-standard servers offers some significant advantages to Intel, compared with a centralized approach based on large proprietary mainframes or RISC-based systems.

- **Lower server acquisition cost.** A large, centralized proprietary system has a high acquisition cost—particularly since the system must include enough headroom for growth across all the company's operations. Acquisition costs for industry-standard servers are a fraction of this.
- **Forecasting and flexible scalability.** With a decentralized design, forecasting and capacity planning can be much simpler

relative to a monolithic design, since the needs of only one business unit have to be comprehended. In contrast, monolithic designs have to contend with the requirements of multiple business units.

- **Faster development of new ERP capabilities.** At Intel, each business group has its own development, test, and production systems. This enables each group to develop new capabilities more quickly and to release them on an independent schedule. They can inexpensively add servers to create test environments that closely resemble production environments.
- **High availability through clustering.** With industry-standard servers, it is relatively inexpensive to add systems to achieve high availability through clustering. In contrast, it is very expensive to add additional large, centralized proprietary systems.
- **Reduced support costs through standardization.** Managing applications on a large number of decentralized servers can be complex compared with managing a few large multi-processor systems. Conversely, because we use standard servers, we can take advantage of the same management tools, platform engineering, and support structure used throughout the rest of Intel's environment. A centralized system used only for ERP requires dedicated, specialized management tools and support.
- **Budgeting.** With our decentralized approach, business groups pay only for their own implementation. This fits Intel's decentralized budgeting model.

ERP Environment Background

Two key concepts significantly affect the way an ERP environment is organized: ERP pipelines supporting the path to production and ERP components.

ERP INSTANCES AND PIPELINES

For each Intel business group ERP implementation, there is a pipeline of ERP application instances. Each instance supports a specific function in the life cycle of ERP releases along the path from development to production. Additional instances are used for production support and disaster recovery, as shown in Figure 1. Each instance may be implemented on one or more dedicated servers.

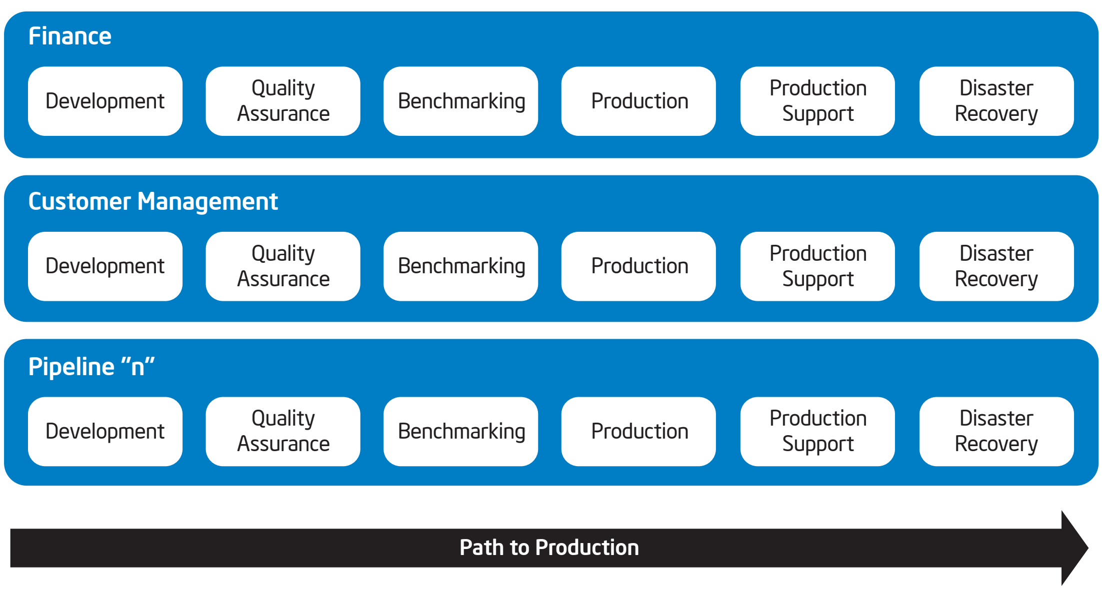

Figure 1. Examples of enterprise resource planning (ERP) pipelines at Intel. Separate instances support each stage in the path from development to production and support. Each instance is implemented on one or more servers.

The primary instances and their uses are:

- **Development.** Creating the initial pipeline configuration, populating master data, and creating and unit testing ERP application modifications.
- **Quality assurance.** Integrated testing of project changes and system testing to assess the impact across the environment.
- **Benchmarking.** Performance and scalability testing prior to production.
- **Production.** Executing production transactions.
- **Production support.** Rapid testing and validation of fixes to the production environment.
- **Disaster recovery.** A remote copy of the production instance for executing production transactions in the event of a disaster.

ERP COMPONENTS

An ERP instance in an ERP pipeline is typically comprised of multiple components. Primary ERP components include:

- **Database server.** This often requires the most server resources.
- **Primary application server.** This application server performs important coordinating functions, such as load balancing user connections and lock management.
- **Secondary application server.** One or more application servers within an instance typically execute ERP source code written in the supplier's programming language.

SERVER PLATFORM STRATEGY

We have standardized on a small set of server platform reference designs based on industry-standard four-socket and two-socket servers. We use these reference designs to support most of the ERP instances in our environment, selecting four-socket or two-socket designs for each instance depending on the requirements.

Standardization

Platform standardization is a fundamental part of our strategy. The cost and effort of validating a new platform is significant. It typically takes about three months to qualify a new platform and six months if substantial changes to the software stack are introduced at the same time.

The primary validation steps include identifying and testing the server hardware and each feature. We also need to identify and test the exact stack of drivers, firmware, OS, and third-party software required to support our needs. The entire reference design then needs to be tested to make sure that it is capable of running production instances and that it can be supported. We conduct further tests to verify correct operation in failover and disaster recovery situations, and we verify that the server can be monitored in our environment and then make any necessary changes to help ensure manageability.

Because this process is expensive and time-consuming, there is a significant cost advantage in validating a small set of reference designs and then using these designs throughout our ERP environment. When selecting a server configuration for a specific ERP instance, it is much faster and more cost-effective to choose an existing reference design than to qualify a new platform.

It is also typically more cost-effective to select a larger server that initially has more capacity than required, rather than qualify a new server in an attempt to find the perfect fit for a specific instance. This also helps to reduce support costs.

Four-socket and Two-socket Server Positioning Framework

Our reference designs are based on industry-standard four-socket and two-socket servers. In the past, four-socket servers have been the largest-capacity servers available on the market in industry-standard designs and price bands, although industry-standard eight-socket designs are becoming available with the Intel Xeon processor 7500 series.

The scalability differences between four-socket and two-socket platforms vary over time as new generations of each platform are released. However, the general distinctions remain. Table 1 summarizes the scalability differences between current four-socket ERP platforms based on Intel Xeon processor 7500 series and current two-socket platforms based on Intel Xeon processor 5600 series.

MAPPING ERP INSTANCES TO SERVER PLATFORMS

The various instances in our ERP pipelines require different levels of capacity, performance, and availability depending on factors such as the ERP components they include, the anticipated loading levels, and the usage. A production instance has considerably higher availability and performance requirements than a development system, for example.

We determine whether to use four-socket or two-socket reference designs by mapping the requirements of the instance to the capabilities of each platform. Our current framework for mapping instances to servers is shown in Figure 2.

Four-socket Server Usages

Intel IT has standardized on four-socket servers for larger ERP production instances

Table 1. Comparison of Current Two-socket and Four-socket Enterprise Resource Planning (ERP) Platforms

	Two-socket Server Based on Intel® Xeon® Processor 5600 Series	Four-socket Server Based on Intel® Xeon® Processor 7500 Series	Benefits of Four-socket Servers over Two-socket Servers[1]
Number of cores	8 to 12	16 to 32	Performance scaling: Up to 2.3x
Number of threads	16 to 24	32 to 64	
I/O slots	4 to 6	8 to 10	Greater I/O expandability: Typically 2x
Memory slots	Up to 18	Up to 64	Greater memory capacity: Up to 3.55x
Maximum memory capacity with 4-GB DIMMs	72 GB	256 GB	
Reliability features	Standard	Advanced features including machine check architecture recovery (MCA recovery)	Greater uptime for mission-critical workloads

[1] Scalability differences depend on specific server configurations.

and supporting database environments. This is because four-socket servers provide enough headroom for anticipated growth of a large production database and instance over the depreciation cycle of a server. We also use four-socket servers for benchmark and disaster recovery instances because they need to reflect production, and we are investigating their use as virtualization hosts.

LARGER PRODUCTION ERP INSTANCES

The ERP production environment is mission critical; maximizing application availability is paramount. The need to minimize scheduled and unscheduled downtimes greatly outweighs the requirement to optimize hardware acquisition costs. An additional important factor is that the cost of industry-standard server platforms represents only a few percent of the total project cost when development costs, software costs, and other expenses are tallied.

Therefore, allowing significant headroom for unanticipated needs—at a very modest overall impact on total cost of ownership (TCO)—is almost always greatly preferable to trying

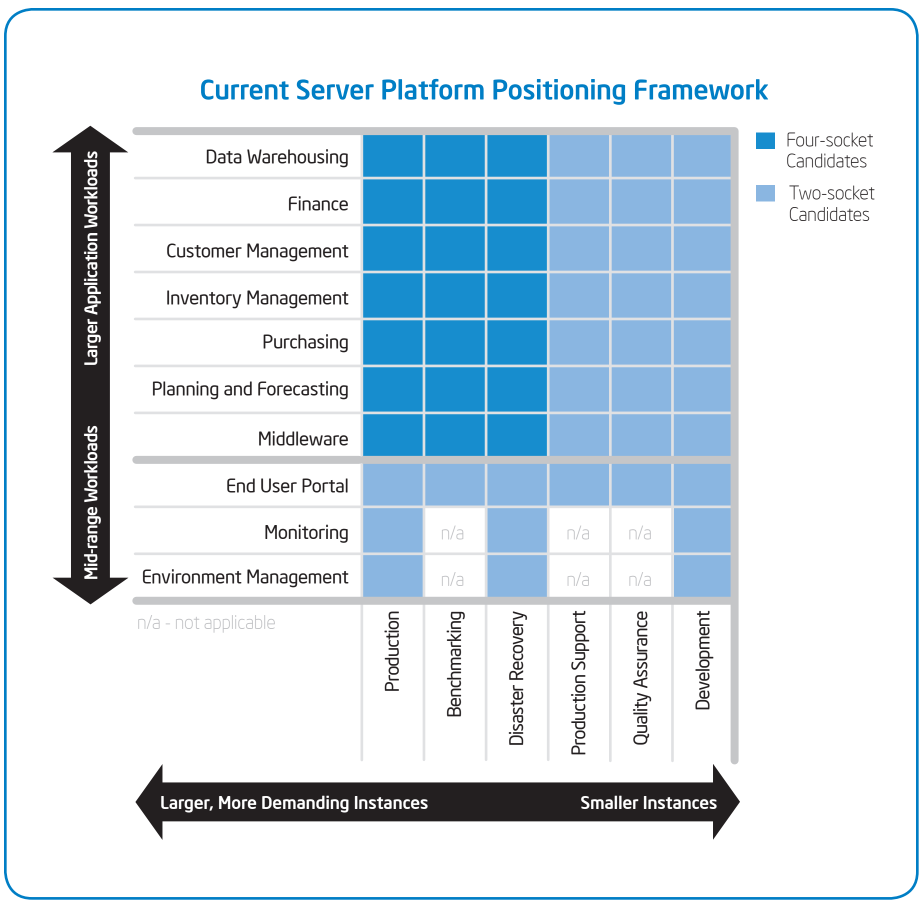

Figure 2. Current server positioning framework for Intel's enterprise resource planning (ERP) environment.

to optimize hardware costs at the risk of impacting application availability.

In general, downtimes must be scheduled far in advance and typically entail complex negotiations with Intel business groups. The data center operations team does not have the option of incurring unscheduled downtime by bringing down the application in order to resize the server to cope with increases in demand.

It is much better to modestly overprovision upfront than to go back to the business group to request scheduled downtimes for hardware upgrades. For example, it would simply not do to delay quarter-end financial processing because the underlying hardware platform had not originally been sized with adequate headroom for "unanticipated" workload growth.

Therefore, a production ERP server platform must include enough performance headroom, memory capacity, and I/O scalability to accommodate increases in the workload.

Our standard configuration for supporting larger production instances is a cluster of two four-socket servers. This provides a high-availability solution for critical ERP components with just two servers, as shown in Figure 3. In normal operations, the database and the primary and secondary application server software components are distributed across the two physical servers for performance. In failover mode, all these components run on the single surviving server.

DETERMINING WHETHER PRODUCTION INSTANCES REQUIRE FOUR-SOCKET SERVERS

We determine whether a production instance is large enough to require a four-socket server by comparing the anticipated peak workload requirements with the capacity of a cluster based on four-socket and two-socket reference designs.

If peak requirements of an ERP database workload are expected to exceed 40 percent of the performance, memory, or I/O resources of a cluster based on two-socket servers, we standardize on a two-node four-socket server cluster. This is because in a failover situation, the surviving server must be able to run the entire workload that was previously supported by both servers and continue to deliver good performance while doing so.

Performance

In general, the performance scalability and headroom of a given server platform is a function of the overall platform design encompassing the processor, cache, memory, and I/O subsystems as a whole.

Testing performance with actual workloads. While published benchmark results can convey some information on the performance and scaling potential of a given server platform, it is critical to measure performance using actual mission-critical ERP production application workloads.

Published benchmark data can also be a good starting point for comparing the performance scaling potential of different server platforms—especially platforms with similar designs, such as two four-socket servers—but it does not provide enough information to make platform selection and configuration decisions.

In particular, extrapolations from benchmark results obtained on server platforms with widely differing expandability and scalability attributes, for example two-socket and four-socket servers, to production workload performance projections is fraught with many risks and must be approached with great care for a number of reasons.

First, the benchmark workload cache, memory, and I/O resource consumption profile may differ materially from that of the production workload. For example, the benchmark workload may fit comfortably in the memory configuration of a two-socket server, while the production workload may require the much greater memory capacity of a four-socket server to handle peaks in demand and/or failover modes. In these situations, the benchmark results are unlikely to correlate

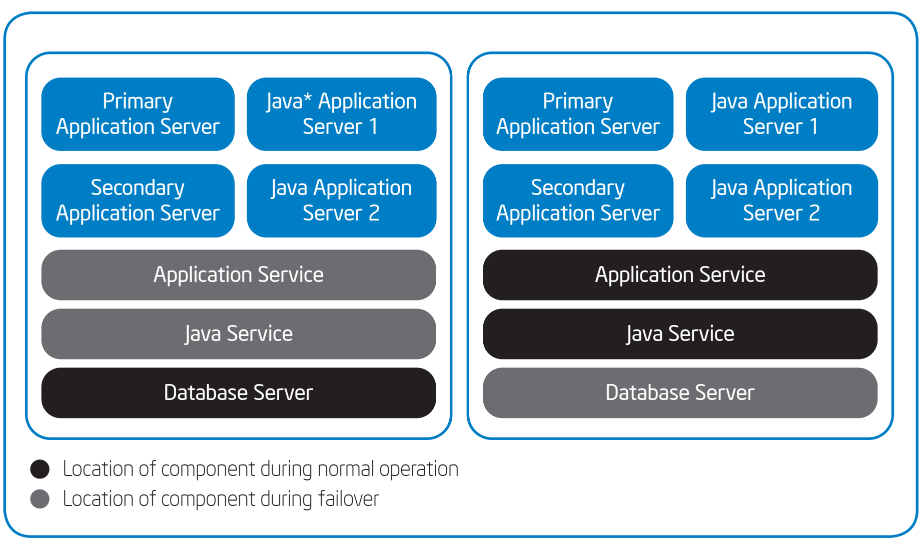

Figure 3. Standard clustered production instance.

well with production workload performance results. The reason is that attempting to run the production workload in a memory-constrained two-socket configuration may sustain severe paging-related performance slowdowns that were not a factor in the published benchmark results due to the much smaller memory footprint of the benchmark workload. Similar considerations can apply to I/O scaling requirements. For example, if the benchmark workload I/O profile is significantly less intensive than the production workload I/O profile, the applicability of the benchmark comparisons to the production environment will be limited.

Second, there may be significant configuration differences between benchmark and planned production configurations, both at the platform level and at an infrastructure level. It is very possible, even likely, that published benchmark results are based on very expensive, high-density RAM subsystems and I/O subsystems—configurations that would not be practical or cost-effective for implementing a large-scale ERP pipeline. Likewise, the supporting benchmark infrastructure may feature high-speed fabric interconnects and high-end storage subsystems that may not be available in the targeted enterprise infrastructure—and the infrastructure may be limited by the presence of significant legacy elements. It is entirely possible that a benchmark configuration with more mainstream (lower capacity, speed) memory and I/O subsystems and infrastructure may have resulted in significantly lower performance than the published benchmark metrics.

Factors influencing performance headroom requirements. Four-socket servers currently have up to four eight-core processors; as a result, they typically have 1.9x to 2.3x the performance headroom compared with two-socket servers available at the time of publication. This headroom can be very important for the following reasons:

- **Uncertainty about future workload volumes.** Despite the use of sophisticated benchmarking and planning discipline to forecast growth in workload average volumes, there can be considerable differences between the expected average volumes and what we actually observe.
- **Uncertainty about workload spikes.** Long-term workload averages may not reflect brief periodic spikes in demand. To support business requirements, the ERP platform must maintain a responsive environment under all loading levels including workload spikes, avoiding response time service-level agreement (SLA) excursions if at all possible.
- **Uncertainty in forecasting future workload placement decisions.** In some situations it may be desirable—for cost, performance, or architecture reasons—to land new applications on existing servers rather than installing new servers. The headroom provided by four-socket servers can give us additional operational flexibility to do this.
- **Failover mode.** When failover occurs, all critical ERP components move to the surviving node in the cluster. The surviving node must still continue to deliver excellent responsiveness while maintaining sufficient headroom for unanticipated workload spikes.

Some of these concepts are shown in Figure 4.

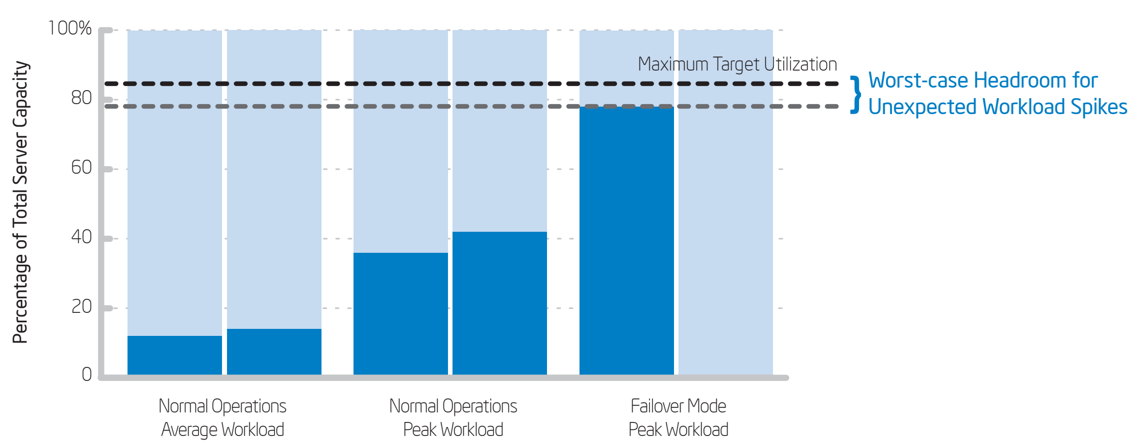

Figure 4. Performance headroom planning.

Memory scalability

ERP workloads can have demanding peak memory requirements. The much greater memory capacity of four-socket servers—typically more than 3.5x the capacity of two-socket servers at the time of writing—is often a decisive factor in server selection. Considerations include:

- Large ERP workloads. Total memory requirements during normal operation may be too high to be safely accommodated within the smaller RAM capacity of a two-socket server.
- Failover. Even with smaller ERP workloads, memory requirements may be too high in a failover situation to be safely accommodated on a single two-socket server.
- Memory redundancy. A four-socket server makes the use of memory redundancy schemes much more feasible. Memory redundancy may become more desirable with increasing memory sizes due to the need to minimize soft-error rates. Using memory redundancy means that twice as much physical memory may be needed to achieve the same effective memory capacity.

I/O scalability

Four-socket servers typically have approximately twice as many Peripheral Component Interconnect Express* (PCIe) expansion slots as two-socket servers, providing the ability to support ERP workloads with much more demanding I/O requirements. Even if workloads do not require this I/O scalability immediately, they may need it in the future to host more users, back up larger databases, or support more applications.

Our current four-socket designs typically have eight slots, compared with four in our two-socket reference designs. This enables us to add network interface cards (NICs) and host bus adapters (HBAs) to increase performance and redundancy; for example, we include two dual-port NICs in a typical production ERP server.

Without an adequate number of I/O expansion slots, platform architects may need to make undesirable trade-offs between performance, scalability, redundancy, and standardization. These interrelated considerations include:

- Data center LAN and SAN connectivity constraints. A straightforward upgrade to respond to growing I/O demand would be to swap out existing LAN or SAN cards with higher speed NICs or HBAs. However, if the data center LAN infrastructure cannot support 10 gigabit Ethernet (GbE), then swapping out existing 1-GbE NICs for 10-GbE NICs will not solve server I/O bottlenecks. Similarly, if the data center SAN fabric is limited to 2 GB per second, upgrading the HBA to 4 GB per second will not help performance.
- Limitations of multi-port NICS and HBAs. Our reference designs use NICs and HBAs with no more than two ports each. To increase the network I/O capacity of four-socket servers, we add more NICs. An alternative way to increase I/O capacity in a server with a limited number of slots is to use quad-port NICs and HBAs. However, this creates a single point of failure: With only one multi-port NIC or HBA per server, a hardware problem on the multi-port card or the slot itself can leave the server with no LAN or SAN connectivity. Additionally, a design with multi-port LAN and SAN connections may not scale as effectively because the ports share common hardware resources and because drivers may limit performance.
- Platform standardization. The presence of additional connectivity slots allows an architect to simply add standard, proven building blocks such as pre-qualified NICs or HBAs to the platform as necessitated by demand signals. Adding non-standard solutions such as multi-port HBAs to solve a specific problem can create substantial additional cost because of the need to qualify and support each new component.

Reliability, availability, and serviceability

In the ERP production environment, high application availability is paramount. To maximize uptime of mission-critical applications, four-socket servers include a range of RAS features that are not currently available in two-socket servers.

For example, four-socket servers based on Intel Xeon processor 7500 series include more than 20 new RAS features. These include MCA recovery, available for the first time in servers based on Intel Xeon processors.

With MCA recovery, the hardware works with the OS or virtual machine monitor (VMM) to increase system availability by allowing servers to recover from uncorrectable memory errors that would otherwise cause a system crash. MCA recovery detects the error, and then enables the OS to determine the best course of action. This can keep the server running when the error affects a non-critical process, or does not affect any processes. For example, if the error affects a non-critical process, the OS may terminate and restart the process while keeping the server running.

VIRTUALIZATION HOSTS

Beyond existing deployments, we are investigating the possibility of virtualizing many non-production ERP instances. Our goals are to reduce cost and increase agility in servicing requests to host new ERP instances.

We are investigating the use of four-socket servers for this due to a number of factors:

- **Large number of cores and threads.** A four-socket platform based on Intel Xeon processor 7500 series has up to 2.7x as many processor cores and available threads as current two-socket servers based on Intel Xeon processor 5600 series. This confers a considerable advantage in serving as a virtualization host (see sidebar for more detail). We expect to configure ERP virtual machines (VMs) with two to four virtual CPUs each (or in the future, eight virtual CPUs in some cases), depending on the virtualized workload. Given the importance of ERP applications and the need to help ensure responsiveness, we are evaluating the possibility of a balanced approach in which the number of virtual CPUs does not exceed the number of physical cores in the server. In a balanced configuration, we can run 2.7x as many VMs on a four-socket server as we can on a two-socket server.
- **Memory capacity.** Non-production ERP VMs have large RAM requirements—typically 8 GB or more per VM. Four-socket servers, with their large memory capacity, can accommodate this.
- **MCA recovery.** MCA recovery, which allows servers to recover from uncorrectable memory errors, is particularly valuable in a virtualized environment where multiple applications are consolidated into VMs on each server. When the hardware detects an uncorrectable memory error, MCA recovery enables the VMM to determine the best course of action. For example, the VMM may shut down and restart the VM directly affected by the memory error, while all other VMs on the server are able to continue to run unaffected.
- **I/O slots.** Virtualization hosts have demanding I/O requirements. Our current virtualization host platform reference design includes seven ports for LAN connectivity. These are used for functions including production networks, live migration, management, and backup and restore. The design also includes two SAN connectivity functions. The eight slots typically available in four-socket platforms translate into much greater I/O headroom and flexibility. In the future, it is possible that high performance requirements may necessitate dedicating NICs or other I/O devices to a given VM. The additional I/O slots in four-socket servers make them well suited to support this need.

Two-socket Server Usages

While four-socket servers are targeted for higher-end applications and more mission-critical functions, two-socket servers are the workhorses for a vast number of mainstream uses.

Newer technologies are often incorporated into two-socket server designs first, where they prove themselves in less-demanding scenarios before making their way into four-socket server designs where there is less room for relatively unproven technologies or concepts. Conversely, four-socket servers tend to rely on more mature and proven technologies and to have longer design and validation cycles.

Within Intel's ERP environment, two-socket servers are used predominantly in cases that do not require the scalability of a large production instance: smaller production instances and a range of non-production uses such as development, production support, and QA. They often run the same ERP components as four-socket servers, but because they have fewer resources, they provide less scalability, support fewer users, or support a less stringent SLA in terms of performance or system availability. Production instances may run on a two-node cluster, while many non-production instances are supported using an "all-in-one" configuration, as shown in Figure 5.

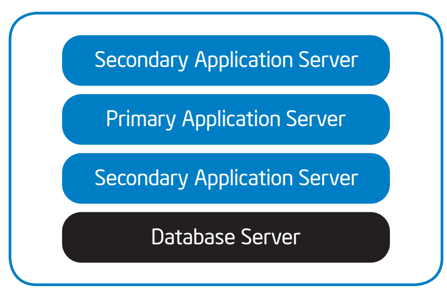

Figure 5. All-in-one non-production enterprise resource planning (ERP) instance running on a two-socket server.

Usages Combining Four-socket and Two-socket Servers: Large Cluster Designs

Two-socket servers running additional application server modules can be used to add incremental scale-out capacity to an existing large instance hosted on a clustered pair of four-socket servers, as shown in Figure 6. Application servers that are part of a large ERP instance can become capacity constrained due to workload growth or the addition of uses or users. We add incremental capacity using industry-standard servers to scale out application processing as opposed to qualifying and standardizing larger proprietary servers.

The scale-out design is also very important when the business processes include batch

Core Count Considerations in Virtualization Hosts

Virtualization software emulates hardware as virtual machines (VMs) that share the resources belonging to the underlying virtualization host server, providing a way to efficiently run multiple OSs and applications concurrently on the same physical server.

Depending on the processing requirements of an application, the VM hosting the application may be configured with one or more virtual CPUs. In order to closely emulate the behavior of a multi-processor computer, the virtualization software typically assigns each virtual CPU to a physical core on the host server, or to a thread in the case of hosts that support hyper-threading. On a virtualization host running multiple VMs, virtual CPU oversubscription occurs when the number of virtual CPUs required to concurrently run all the VMs exceeds the number of physical cores (or threads) available. In these oversubscribed environments, VMs take turns sharing the available cores. The scheduling and associated context-switching overhead, such as saving and restoring VM states and flushing and refilling caches, can sometimes result in performance degradation. The performance impact may differ depending on workload characteristics. In a significantly oversubscribed environment, workloads that are easier for the virtual machine monitor (VMM) to schedule because they are hosted in VMs requiring only a few CPUs, or that are more compute-intensive and therefore can more efficiently use the CPU time allocated to them, may be less impacted than workloads requiring a large number of virtual CPUs.

The decision about how many VMs to run on a virtualization host may depend on whether the priority is to maximize consolidation ratios or to achieve predictable high performance and minimize performance-related variability. Oversubscription allows us to consolidate more virtualized workloads onto a single virtualization host, improving overall resource utilization. However, as we add more VMs and oversubscribe the available cores (or threads), there may be an increasing possibility of performance excursions that are intermittent and difficult to reproduce. In contrast, a balanced configuration in which the number of requested virtual CPUs equals the number of physical cores (or threads) available on the platform may help deliver relatively more predictable performance for all VMs.

For mission-critical applications where predictable performance is critical, such as ERP environments, IT organizations may choose the balanced approach. In these situations, the substantially greater core count of four-socket servers provides a key advantage, because it allows us to run many more VMs per host without oversubscription.

For example, a VM might be assigned two virtual CPUs. A two-socket server based on Intel® Xeon® processor 5600 series, with six cores per processor, has a total of 12 cores and 24 threads and can host twelve VMs without oversubscription. A four-socket server based on the Intel® Xeon® processor 7500 series, with eight cores per processor, has a total of 32 cores and 64 threads and can host 32 VMs without oversubscription—about 2.7x times as many as on the two-socket server. This scalability may be an important factor favoring the selection of four-socket servers as virtualization hosts.

To read more about Intel IT and virtualization, go to www.intel.com/it/virtualization.htm processing, programs that interface between systems, or parallel processing of order documents. This creates a place to isolate the impact of these very resource-intensive workloads from the rest of the system usage. ERP software with support for log on workload balancing and other workload directing schemes makes this isolation scheme possible.

If the database server becomes a bottleneck, then the instance can be repartitioned by moving some of the application servers off the four-socket server running the database service to an additional two-socket host.

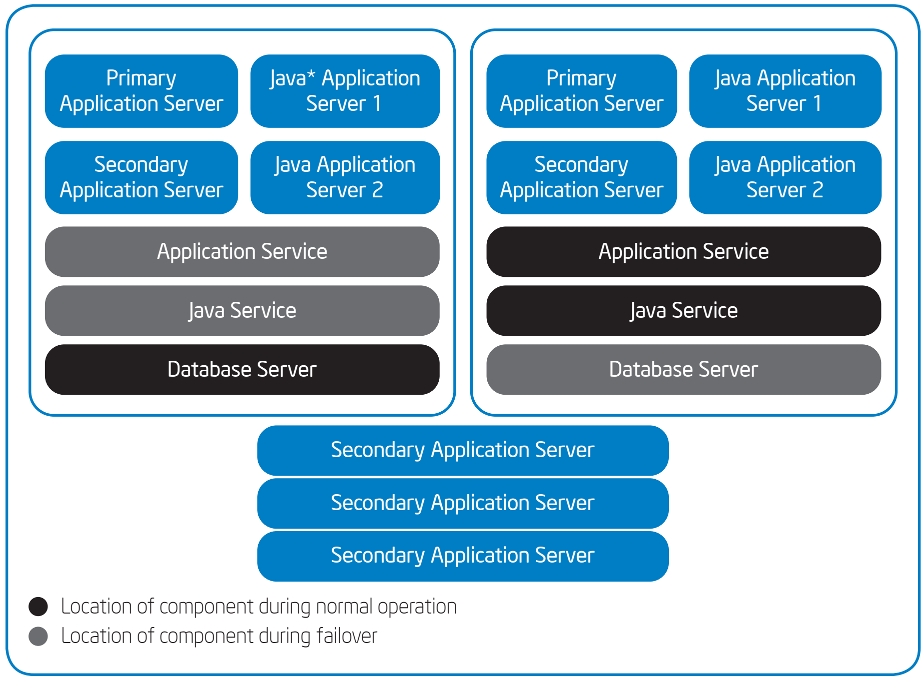

Figure 6. Two-socket servers running additional application server modules can be used to add incremental scale-out capacity to an existing large instance hosted on a clustered pair of four-socket servers.

Increasing Use of Virtualization

Over time, we expect steady adoption of virtualization in the ERP environment, starting with the lowest-risk instances and progressing to more critical server roles as virtualization delivers proven stability, reliability, and performance.

The possible changes to our future server platform positioning framework due to virtualization are shown in Figure 7. In this scenario, many smaller non-production ERP instances are virtualized and run on four-socket virtualization hosts. Large production instances continue to run on four-socket servers as they do today, because they require the additional performance, memory, and I/O scalability that these servers provide.

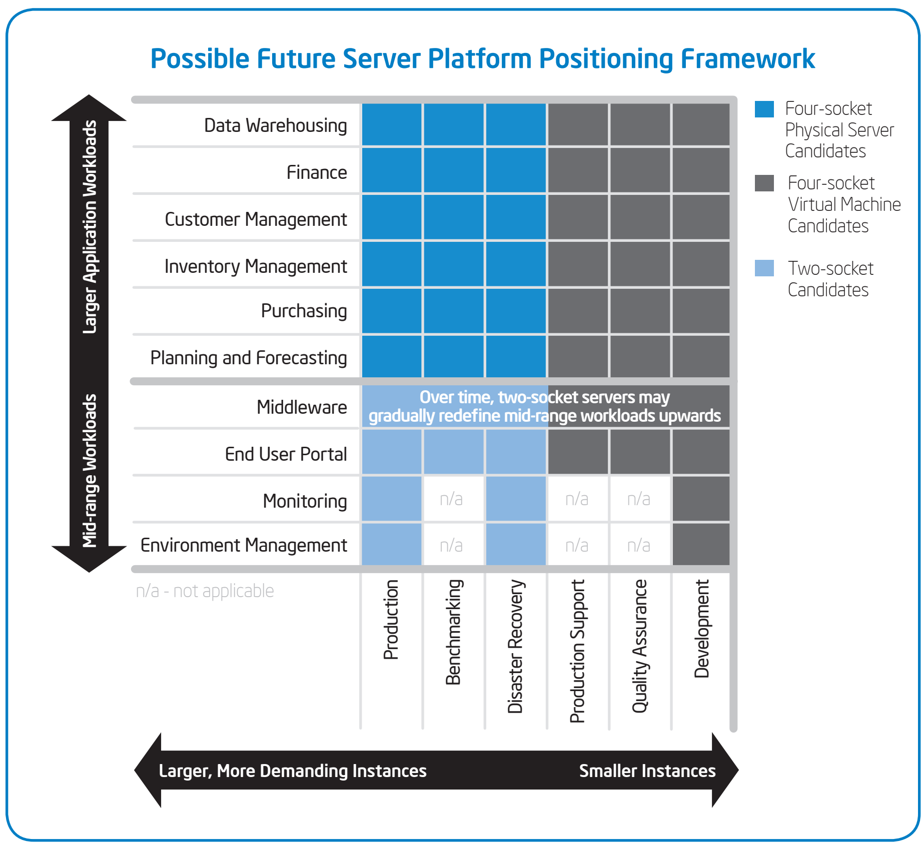

Figure 7. Possible future server positioning framework for Intel's enterprise resource planning (ERP) environment.

CONCLUSION

Intel's ERP strategy has enabled Intel IT to effectively support more than 10,000 active users with industry-standard servers.

Four-socket servers play an essential role in this strategy. They provide the performance, memory capacity, I/O expandability, and proven technology required to run our larger production ERP instances. They also have the headroom to continue delivering good performance while accommodating workload growth, spikes in demand, and failover situations. Their scalability also makes four-socket servers well suited to act as virtualization hosts, and we are investigating using them as hosts for virtualized non-production ERP instances.

For more straight talk on current topics from Intel's IT leaders, visit www.intel.com/it.

ACRONYMS

ERP	enterprise resource planning
GbE	gigabit Ethernet
HBA	host bus adapter
MCA recovery	machine check architecture recovery
NIC	network interface card
PCIe	Peripheral Component Interconnect Express
RAS	reliability, availability, and serviceability
SLA	service-level agreement
TCO	total cost of ownership
VM	virtual machine
VMM	virtual machine monitor

Printed in USA
0310/KAR/KC/PDF

Please Recycle
322692-001US