Cisco Unified Computing System Extended Memory Technology

Cost-Effective Support for Large-Memory Workloads

At-A-Glance

Eliminating Memory Bottlenecks

In both good and bad economic times, the pressure to maintain required service levels while holding expenses in check continues unabated. In response to this need, organizations look for ways to increase server computing platform capacity and efficiency through consolidation and virtualization. As the workload on line-of-business (LOB) applications, corporate databases, online transaction processing (OLTP), and web server applications increases, the memory capacity of server systems plays an increasingly important role in overall system performance. More memory is needed to support increased core counts and clock rates in today's multicore processors. More memory is needed to support more virtual machines on the same server. And more memory is needed for a class of enterprise applications that can improve performance by caching more of their data in main memory.

There are three approaches to resolving these challenges to eliminating memory bottlenecks:

- Increase the number of servers—typically two socket x86 servers.
- Purchase larger and more expensive four-socket-systems for increased memory capacity.
- Increase the memory capacity per server.

Increasing the number of servers drastically increases operating expenses (OpEx), management, and maintenance costs due to increased power, cooling, and infrastructure needs. This option results in a high total cost of ownership (TCO) and may not be the optimal solution for today's businesses.

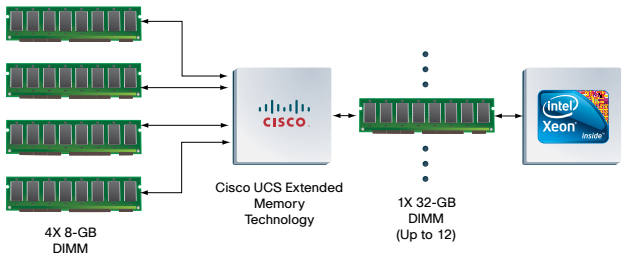
Purchasing more expensive four-socket systems increases the amount of memory available, as each socket offers more addressable memory, but four-socket servers require more expensive processors that may not be needed if memory is the bottleneck. Increasing the number of sockets may also incur additional software licensing costs in addition to increasing capital expenditures (CapEx) and OpEx. Again, the result is higher TCO.

In contrast, increasing the memory capacity per server is a potent, cost-effective, and scalable approach that can deliver a better return on investment (ROI) and lower TCO compared to other approaches, enabling IT departments to accomplish more with less.

Cisco Unified Computing System Extended Memory Technology

Cisco® Extended Memory Technology is available on the Cisco UCS B250 M1 and UCS B250 M2 Extended Memory Blade Servers and the Cisco UCS C250 M1 and UCS 250 M2 Extended Memory Rack-Mount Servers. The technology maps four physically distinct DIMMs to a single logical DIMM as seen on the processor's memory channel (Figure 1). This mapping supports extended memory servers with 48 DIMM slots in which traditional servers and blade systems using the same processors can have only up to 12 slots at full performance, or 18 slots at reduced performance.

Figure 1. Cisco Extended Memory Technology Makes Four Physical DIMMs Appear to the CPU as a Single, Large, Logical DIMM



The 48 DIMM slots in a Cisco extended memory server can be populated with 4- or 8-GB DIMMs, equipping data center operators with two powerful new options for balancing memory capacity and performance.

Low-Cost Option

A low-cost option delivers a memory footprint of up to 192 GB using low-cost 4-GB DIMMs rather than more expensive 8-GB DIMMs. This option enables data center operators to save up to 60 percent on memory costs compared to traditional two-socket servers. These savings come without the performance penalty incurred when traditional systems use 18 rather than 12 DIMM slots (Table 1). In addition, Cisco Extended Memory Technology can support up to 192 GB of memory using 4-GB DIMMs—48 GB more than traditional servers using more expensive 8-GB DIMMS.

Virtualized environments can have large fixed memory allocations for each virtual machine without the need for two layers of paging. Database management systems, business logic applications, and electronic design automation (EDA) simulations can be improved by allowing entire data sets to be loaded into main memory at one time. Likewise, the performance of memory-based workloads, including in-memory databases and name and content caching servers, can be improved by memory-based retrievals.

Large-Footprint Option

A large-footprint option can accommodate selected extremely memory-intensive workloads. With up to 384 GB of memory available using 8-GB DIMMS, the Cisco UCS B250 M1, UCS B250 M2, UCS 250 M1 and UCS C250 M2 servers deliver the largest memory footprint available in any two-socket sever using Intel® Xeon® 5500 or 5600 series processors. This capacity rivals that of current four-socket x86-architecture servers and provides an economical two-socket solution alternative to more expensive and larger four-socket servers.

Cisco Unified Computing System Extended Memory Technology

Cost-Effective Support for Large-Memory Workloads

At-A-Glance

The result is memory scalability that is decoupled from traditional costs. With reduced costs and larger than ever memory footprints, IT departments can now consolidate more applications and virtual machines more economically.

Table 1. Savings with Cisco Extended Memory Technology Based on Publicly Advertised Prices Obtained August 2009

Memory Capacity	Typical System Memory Cost (US$)	Cisco Unified Computing System™ Memory Cost (US$)	Cost Savings (US$)	Savings (Percent)
96	11,880	4,800	7,080	60
144	17,820	7,200	10,620	60
192	Not available	9,600		
384	Not available	47,520		

Memory Latency Characteristics

Intel Xeon 5500 and 5600 series processors have a built-in memory controller that supports three memory channels per processor. In a two-socket server, the two processors are interconnected through the Intel QuickPath Interconnect (QPI). Memory latency is characterized by local access (to memory directly attached to the processor) and remote access (to memory connected to the other processor, requiring traversal of the interconnect). With Cisco Extended Memory Technology, latency to local memory is marginally higher than in a system without memory extension, and it is significantly lower than the latency for remote memory access. This feature can result in significant performance improvements based on memory access times in addition to the performance improvements that can be achieved because of a larger memory capacity in the server.

Business Benefits of Extended Memory Technology

Cisco Extended Memory Technology provides cost savings and improved performance in addressing large server workloads through:

- Increased flexibility to balance server price and performance; 4-GB DIMMs are less than half the cost per gigabyte of 8-GB DIMMs, allowing large memory footprints to be established at a lower cost
- Reduced capital, operational, power, cooling, and maintenance costs through deployment of fewer servers
- Reduced software license costs, enabled by increased consolidation and use of two-socket servers
- Lower system capital expenditures through more economical memory costs

Cisco Unified Computing Services

Using a unified view of data center resources, Cisco and our industry-leading partners deliver services that accelerate your transition to a unified computing environment. Cisco Unified Computing Services help you quickly deploy your data center resources and optimize ongoing operations to better meet your business needs. For more information about these and other Cisco Data Center Services, visit http://www.cisco.com/go/dcservices.

Why Cisco?

The Cisco Unified Computing System continues Cisco's long history of innovation in delivering integrated systems for improved business results based on industry standards and using the network as the platform. Recent examples include IP telephony, LAN switching, unified communications, and unified I/O. Cisco began the unified computing phase of our Data Center 3.0 strategy several years ago by assembling an experienced team from the computing and virtualization industries to augment our own networking and storage access expertise. As a result, Cisco delivered foundational technologies, including the Cisco Nexus® Family, supporting unified fabric and server virtualization. The Cisco Unified Computing System completes this phase, delivering innovation in architecture, technology, partnerships, and services. Cisco is well-positioned to deliver this innovation by taking a systems approach to computing that unifies network intelligence and scalability with innovative application specific integrated circuits (ASICs), integrated

For More Information

Please visit http://www.cisco.com/go/ucs.

C45-555038-03 02/10