White Paper

Cisco Unified Computing System BIOS Settings

Tune for Performance

Guide for Cisco UCS B200 M4 Blade Servers and C220 M4 and C240 M4 Rack Servers

What You Will Learn

Setting performance options in your system BIOS can be a daunting and confusing task given some of the obscure options you can choose. For most options, you must choose between optimizing a server for power savings or for performance. This document provides some general guidelines and suggestions to help you achieve optimal performance from your Cisco UCS® B200 M4 Blade Server and Cisco UCS C220 M4 and C240 M4 Rack Servers.

Processor Configuration: Intel SpeedStep and Turbo Boost

Intel SpeedStep Technology is designed to save energy by adjusting the CPU clock frequency up or down depending on how busy the system is. Intel Turbo Boost Technology provides the capability for the CPU to adjust itself to run higher than its stated clock speed if it has enough power to do so. One new feature in the Intel Xeon processor E5-2600 v3 CPUs is the capability for each core to run at a different speed, using Intel SpeedStep. In prior generations, all cores on a chip ran at the same speed.

Intel Turbo Boost depends on Intel SpeedStep: if you want to enable Intel Turbo Boost, you must enable Intel SpeedStep first. If you disable Intel SpeedStep, you lose the ability to use Intel Turbo Boost.

Intel Turbo Boost is especially useful for latency-sensitive applications and for scenarios in which the system is nearing saturation and would benefit by temporarily maximizing the CPU speed. If your system is not running at this saturation level and you want the best performance at a utilization rate of less than 90 percent, you should disable Intel SpeedStep to help ensure that the system is running at its stated clock speed at all times.

Processor C3 and C6 States

C3 and C6 are power-saving halt and sleep states that a CPU can enter when it is not busy. Unfortunately, it can take some time for the CPU to leave these states and return to a running condition. If you are concerned about performance (for all but latency-sensitive single-threaded applications), and if you have the option, disable anything related to C states.

CPU Hyperthreading

You should test the CPU hyperthreading option both enabled and disabled in your specific environment. If you are running a single-threaded application, you should disable hyperthreading.

Core Multiprocessing and Latency-Sensitive Single-Threaded Applications

The core multiprocessing option is designed to give the user the capability to disable cores. This option may affect the pricing of certain software packages that are licensed by the core. You should consult your software license and software vendor about whether disabling cores qualifies you for any particular pricing policies. Set core multiprocessing to All if pricing policy is not an issue for you. For latency-sensitive single-threaded applications, you can optimize performance by disabling unnecessary cores, disabling hyperthreading, enabling all C states, enabling Intel SpeedStep, and enabling Intel Turbo Boost. With this configuration, the remaining cores often will benefit from higher turbo speeds and better use of the shared Layer 3 cache.

Energy or Performance Bias

You can use the power-saving mode to reduce system power consumption when the turbo mode is enabled. The mode can be set to Maximum Performance, Balanced Performance, Balanced Power, or Power Saver. Testing has shown that most applications run best with the Balanced Performance setting.

Power Technology Setting

For best performance, always set the power technology option to Custom. If it is not set to Custom, the individual settings for Intel SpeedStep and Turbo Boost and the C6 power state are ignored.

CPU Prefetcher Settings

Intel Xeon processors have several layers of cache. Each core has a tiny Layer 1 cache sometimes referred to as a data cache unit (DCU) that has 32 KB for instructions and 32 KB for data. Slightly bigger is the Layer 2 cache, with 256 KB shared between data and instructions per core. In addition, all cores on a chip share a much larger Layer 3 cache, which is about 10 to 45 MB in size (depending on the processor model and number of cores). The prefetcher settings provided by Intel primarily affect the Layer 1 and Layer 2 caches on a processor core (Table 1). You will likely need to perform some testing with your individual workload to find the combination that works best for you. Testing on the Intel Xeon processor E5-2600 v3 CPUs has shown that most applications run best with all prefetchers enabled. See Table 2 for guidance.

Table 1. CPU Performance and Prefetch Options from Intel

Performance Option	Cache Affected
Hardware prefetcher	Layer 2
Adjacent-cache-line prefetcher	Layer 2
DCU prefetcher	Layer 1
DCU instruction pointer (DCU-IP) prefetcher	Layer 1

Hardware Prefetcher

The hardware prefetcher prefetches additional streams of instructions and data into the Layer 2 cache upon detection of an access stride. This behavior is more likely to occur during operations that sort through sequential data, such as database table scans or clustered index scans, or that run a tight loop in code.

Adjacent-Cache-Line Prefetcher (Buddy Fetch)

The adjacent-cache-line prefetcher always prefetches the next cache line. Although this approach works well when data is accessed sequentially in memory, it can quickly litter the small Layer 2 cache with unneeded instructions and data if the system is not accessing data sequentially, causing frequently accessed instructions and code to leave the cache to make room for the “buddy” data or instructions.

DCU Prefetcher

Like the hardware prefetcher, the DCU prefetcher prefetches additional streams of instructions or data upon detection of an access stride; however, it stores the streams in the tiny Layer 1 cache instead of the Layer 2 cache.

DCU-IP Prefetcher

The DCU-IP prefetcher predictably prefetches data into the Layer 1 cache on the basis of the recent instruction pointer load instruction history.

Table 2. Cisco UCS CPU Prefetcher Options and Target Benchmarks and Workloads

Prefetchers	Target Benchmarks and Workloads
All enabled	High-performance computing (HPC) benchmarks, webserver, SAP Application Server, virtualization, and TPC-E
DCU-IP enabled; all others disabled	SPECjbb2005 and certain server-side Java application server applications

Memory Performance Settings

Memory Reliability, Availability, and Serviceability Configuration

Always set the memory reliability, availability, and serviceability (RAS) configuration to Maximum Performance for systems that require the highest performance and do not require memory fault-tolerance options.

Nonuniform Memory Access

Most modern operating systems, particularly virtualization hypervisors, support nonuniform memory access (NUMA) because in the latest server designs, a processor is attached to a memory controller: meaning that half the memory belongs to one processor, and half belongs to the other processor. If a core needs to access memory that resides in another processor, a longer latency period is needed to access that part of memory. Operating systems and hypervisors recognize this architecture and are designed to reduce such trips. For hypervisors such as those from VMware and for modern applications designed for NUMA, keep this option enabled.

Isochronous Mode

Enabling the Iscochronous (ISOC) Mode option reduces the credits available for memory traffic. For memory requests, this option reduces latency at the expense of throughput under heavy loads.

CPU Snoop Settings

When a system has more than one CPU, the chipset must work to maintain data consistency so that CPUs don't simultaneously modify a data value and thus create inconsistent data. For instance, if a CPU is tasked with modifying a data value, it will first want to check with the other CPUs in the system to verify that they are not using the data value. CPU snoop settings dictate the way that a system maintains this data consistency between two or more processors. For the Cisco UCS B200 M4, C240 M4, and C220 M4, three potential snoop settings are available: Early Snoop, Home Snoop, and Cluster on Die. The value for best performance depends on software support, the Intel Xeon processor E5-2600 v3 CPU that is installed, the specific workload, and NUMA settings.

Cluster on Die Snoop

Cluster on Die (CoD) snoop is available on Intel Xeon processor E5-2600 v3 CPUs that have 10 or more cores. Note that some software packages do not support this setting: for example, at the time of writing, VMware vSphere does not support this setting. CoD snoop is the best setting to use when NUMA is enabled and the system is running a well-behaved NUMA application. A well-behaved NUMA application is one that generally accesses only memory attached to the local CPU. CoD snoop provides the best overall latency and bandwidth performance for memory access to the local CPU. For access to remote CPUs, however, this setting results in higher latency and lower bandwidth. This snoop mode is not advised when NUMA is disabled.

Early Snoop

Early Snoop (ES) is available on all Intel Xeon processor E5-2600 v3 CPUs. It provides good local memory latency and bandwidth, but with a severe penalty in bandwidth for remote CPU access. ES should be used for latency-sensitive applications that do not require high remote bandwidth: for example, certain online transaction processing (OLTP) database workloads.

Home Snoop

Home Snoop (HS) is also available on all Intel Xeon processor E5-2600 v3 CPUs and is excellent for NUMA applications that need to access a remote CPU on a regular basis. Of the snoop modes, HS provides the best remote CPU bandwidth and latency, but with the penalty of slightly higher local latency, and is the best choice for

memory and bandwidth-intensive applications: for example, certain decision support system (DSS) database workloads.

Conclusion

When tuning system BIOS settings for performance, you need to consider a number of processor and memory options. If performance is your goal, be sure to choose options that optimize for performance in preference to power savings, and experiment with other options such as CPU prefetchers, snoop settings, and CPU hyperthreading.

For More Information

- Cisco UCS B200 M4 Blade Server: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b200-m4-blade-server/index.html
- Cisco UCS C220 M4 Rack Server: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-c220-m4-rack-server/index.html
- Cisco UCS C240 M4 Rack Server: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-c240-m4-rack-server/index.html

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-733130-00 10/14