Best Practices for Setting BIOS Parameters for Performance

Cisco UCS E5-based M3 Servers

May 2013

Data Center of the Future

Contents

Purpose and Scope

The Basic Input Output System (BIOS) tests and initializes the hardware components of a system and boots the operating system from a storage device. In a typical computational system, there are several BIOS settings that control the system's behavior. Some of these settings are directly related to the performance of the system.

This paper discusses the BIOS options that can be selected for various workload types on Cisco Unified Computing System® (UCS) M3 Intel E5-based servers. Understanding the BIOS options helps to select appropriate values to achieve an optimal system performance.

It is beyond the scope of this paper to discuss the BIOS options for a specific firmware release of the Cisco UCS server. The settings demonstrated are generic.

Note: In certain firmware versions, some of the BIOS options are invisible in a server type, and would upgrades. In this case, check the proper firmware release notes and perform the required upgrades. These BIOS settings are ONLY applicable to Cisco UCS M3 servers that use Intel E5- series processors.

What Will You Learn

Determining performance options to set in your system BIOS can be challenging. For most options, you must choose between optimizing a server for power savings or for server performance. This paper provides general guidelines and best practices for achieving the best performance on the Cisco UCS E5-based M3 B-Series blade and C-Series rack-mount servers.

In most situations the default BIOS settings work fine. However, some BIOS settings can have a significant performance impact, depending on the workload and the applications that are run on the system. BIOS settings are critical for the following:

- Performance tuning
- Time-critical systems
- Real-time systems

BIOS Settings for Workload Types

In this paper we will discuss BIOS settings for the following types of workloads:

- Online Transaction Processing (OLTP)
- High-Performance Computing (HPC)
- Java Enterprise Edition Application Server
- Decision Support Systems
- Virtualization

## Online Transaction Processing (OLTP)

OLTP systems contain the operational data to control and run important transactional business tasks. These systems are characterized by their ability to complete various concurrent database transactions and process real-time data. They are designed to provide optimal data processing speed.

OLTP systems are often decentralized to avoid single points of failure. Spreading the work over multiple servers can also maximize transaction processing volume and minimize response times.

Processor Settings for OLTP Workloads

Obtaining peak performance requires some system-level tuning. Figure 1 and Figure 2 illustrate the optimum selections that are recommended for OLTP workloads.

Figure 1.   Processor Settings for OLTP Workloads in Cisco UCS E5-based M3 B-Series Servers
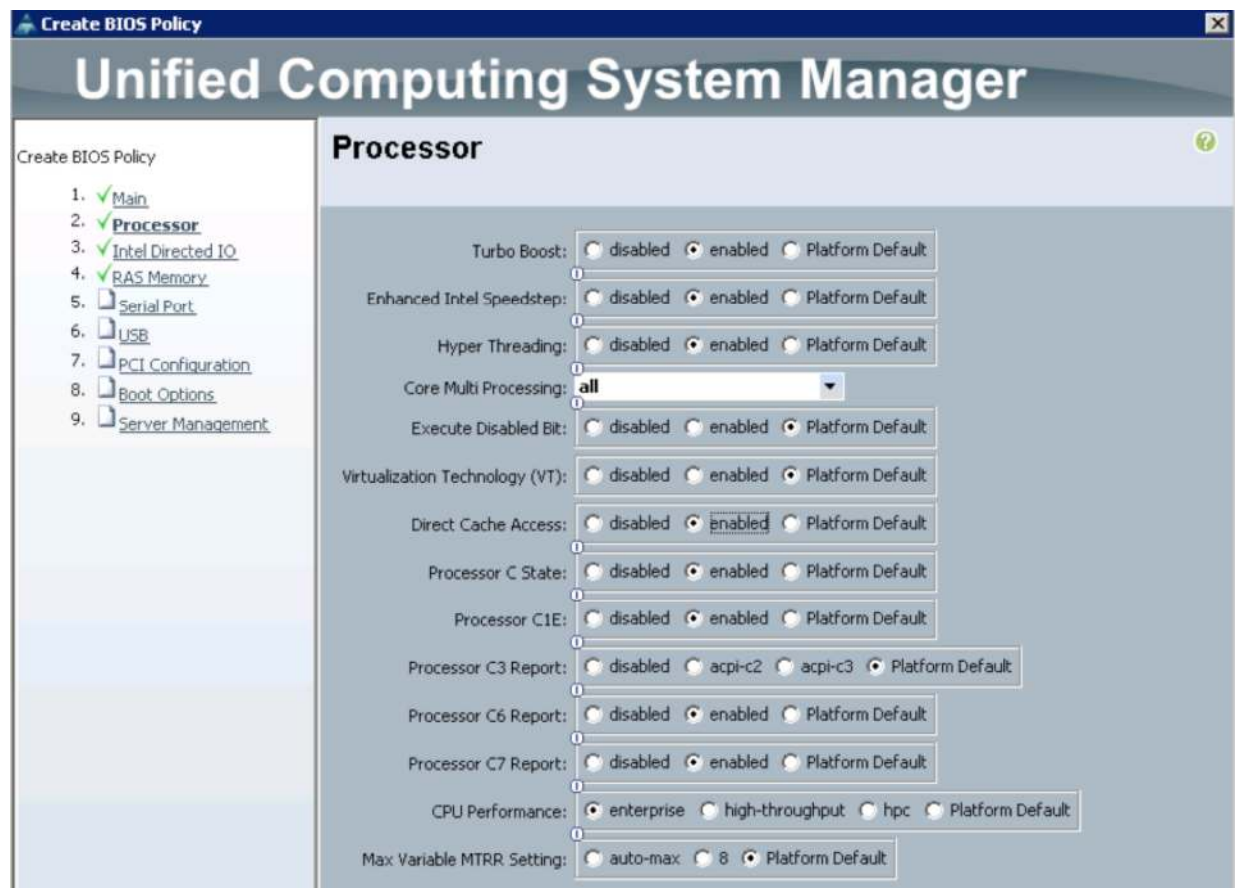
Figure 2.   Processor Settings for OLTP Workloads in Cisco UCS E5-based M3 C-Series Rack Servers


```
Processor Configuration
         Intel(R) Hyper-Threading Technology:    Enabled
                    Number of Enabled Cores:    All
                           Execute Disable:    Enabled
                              Intel(R) VT:    Enabled
                            Intel(R) VT-d:    Enabled
            Intel(R) VT-d Coherency Support:    Disabled
                 Intel(R) VT-d ATS Support:    Enabled
                          CPU Performance:    Enterprise
                        Hardware Prefetcher:    Enabled
               Adjacent Cache Line Prefetcher:    Enabled
                       DCU Streamer Prefetch:    Enabled
                          DCU IP Prefetcher:    Enabled
                  Direct Cache Access Support:    Enabled
                          Power Technology:    Custom
        Enhanced Intel Speedstep(R) Technology:    Enabled
              Intel(R) Turbo Boost Technology:    Enabled
                    Processor Power state C6:    Enabled
            Processor Power state C1 Enhanced:    Enabled
                     Frequency Floor Override:    Enabled
                       P-STATE Coordination:    HW ALL
                        Energy Performance:    Performance
```

The Intel Turbo Boost and Speedstep technology are powerful management features that adjust the CPU voltage or the frequency to optimize performance and power consumption dynamically. During periods of low CPU consumption, Speedstep can reduce the CPU frequency by reducing power consumption. The Turbo Boost boosts the execution processing speeds to accommodate higher demand in situations where CPU utilization is extremely high. Each core will have 20 to 30 percent more processing capability when Turbo Boost is enabled. For example, if Turbo Boost is enabled on the Cisco UCS CPU E5 2690, the system can achieve frequencies as high as 3.8 GHz.

When you tune for consistent performance for OLTP applications on a system that does not run at near 100 percent CPU utilization, we suggest that you enable Intel Speedstep and Turbo Boost, and enable C States. While foregoing power savings during idle times, this will keep all CPU cores running at a consistent speed and will deliver the most consistent and predictable performance.

Enabling Hyper-Threading helps the OLTP systems to handle the heavy I/O workloads by allowing the execution of multiple threads per CPU core. The OLTP applications are typically multithreaded, with each thread doing a short amount of work that may include I/O operations. Large number of threads results in heavy context switching, but with Hyper-Threading, the effect of context switching is reduced. When Direct Cache Access (See Figure 1 and Figure 2) is enabled, the I/O controller places data directly into the CPU cache to minimize the cache misses while processing OLTP workloads. This results in improved application performance.

Figure 3.   Intel Directed I/O Settings for OLTP Workloads in Cisco UCS E5-based M3 B-series Blade Servers
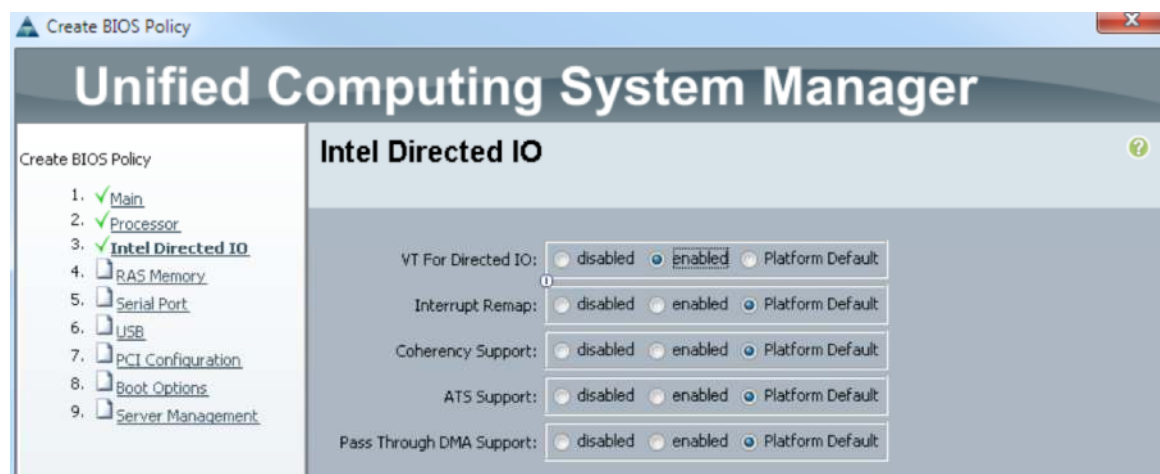
If you are deploying the system in a virtualized environment and the OLTP application uses a directed I/O path, make sure to enable the VT for Directed IO option. With Cisco® VM-FEX technology, virtual machines can now directly write to the virtual network interface cards (NICs) when directed I/O is enabled at the BIOS level.

Note: This is only applicable if the OLTP system is running in a virtualized environment.

Memory Settings for OLTP Workloads

Figure 4 and Figure 5 show memory settings for the OLTP workloads.

Figure 4.   Memory Settings for OLTP Workloads in Cisco UCS E5-based M3 B-Series Blade Servers
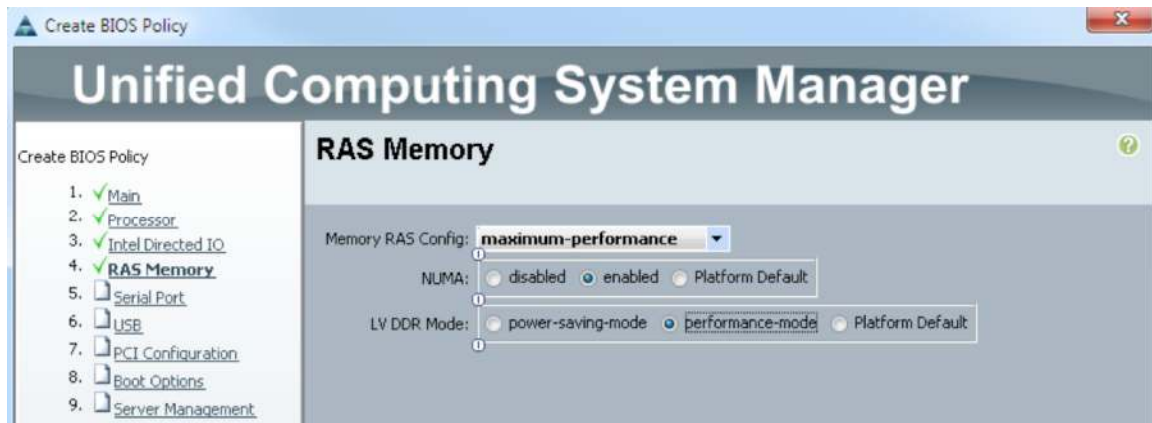
Figure 5.   Memory Settings for OLTP Workloads in Cisco UCS E5-based M3 C-Series Rack Servers
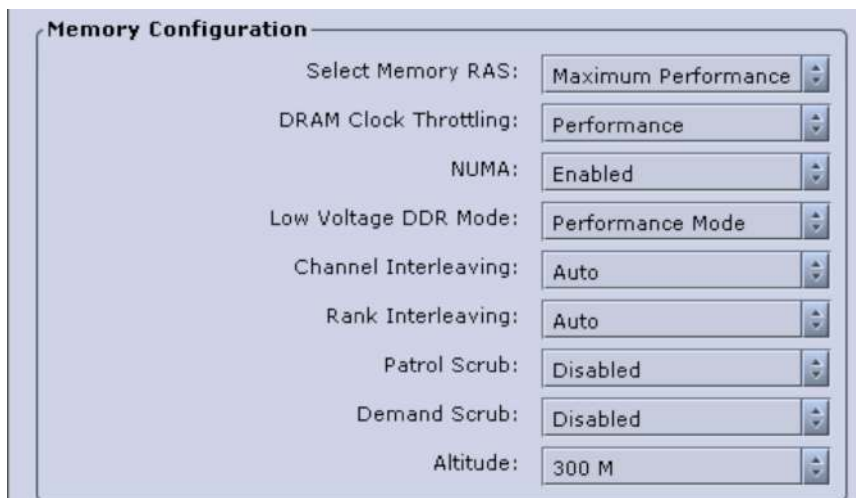
The OLTP applications have a random memory access pattern and it benefits greatly from larger and faster memory, therefore, Cisco recommends setting Memory RAS features to the "maximum performance" for the optimal system performance. Also, DDR Mode should be set at performance mode so that the DIMMS run at the highest frequency for the installed memory/CPU combination. In OLTP transactions, if these modes are enabled, I/O operations will be serviced at the maximum frequency and will have reduced memory latency.

Note: If the DIMM pairs in the server have identical type, size and organization and are populated across the Scalable Memory Interconnect (SMI) channels, you can enable lockstep (option under "Select Memory RAS") mode to minimize memory access latency and achieve better performance.

High-Performance Computing

High-performance computing (HPC) refers to cluster-based computing that utilizes various individual nodes that are connected and that work in parallel in order to reduce the amount of time required to process large data sets that would otherwise take exponentially longer to run on any one system. HPC workloads are compute and typically network I/O intensive. The HPC workloads require high-quality CPU components and high-speed, low-latency network fabrics for their Message Passing Interface (MPI) connections.

The compute clusters consist of a head-node that provides a single point of operation to administer, deploy, monitor, and manage the cluster, as well as an internal workload management component, known as the scheduler that manages all incoming work items (referred to as "jobs"). Typically, HPC workloads require large numbers of nodes with non-blocking MPI networks in order to scale. Scalability of nodes is the single biggest factor in determining the realized usable performance of a cluster.

Processor Settings for HPC Workloads

Figure 6 and Figure 7 show processor settings for HPC workloads.

Figure 6.   Processor Settings for HPC Workloads in Cisco UCS E5-based M3 B-Series Blade Servers

Figure 7.   Processor Settings for HPC Workloads in Cisco UCS E5-based M3 C-Series Rack Servers

The Turbo Boost technology should be enabled for HPC workloads to increase the compute power. When Turbo Boost is enabled, each core provides a higher computing frequency potential such that a greater number of parallel requests can be executed efficiently.

The Intel Speedstep technology is enabled because it is required for Turbo Boost to function. The HPC workloads typically do not benefit from Hyper-Threading- additional threads only serve to create resource contention within the micro-architecture of the CPU. Generally, Hyper-Threading has the greatest impact on workloads where threads are forced to wait on the completion of backend I/O requests that reduce the thread contention for CPU resources.

Enabling the Processor Power state C6 helps in saving the power when the CPU is idle. Since HPC is high on computing, the chance of the CPU going into an idle state is limited. However, enabling C states will save the CPU power in case there are any inactive requests.

The CPU performance should be set to HPC mode to handle more random, parallel requests by the HPC applications. When HPC performs more in-memory processing (for example, video data), it is best to have Prefetcher options enabled so that they can handle multiple parallel requests. This will also help in retaining some hot data in the Layer 2 cache, and it also improves HPC performance (CPU performance).

The HPC requires a high-bandwidth I/O network. When you enable Direct Cache Access Support, network packets will fall directly into the Layer 3 processor cache instead of the main memory. This will reduce the number of HPC I/O cycles generated by HPC workloads when certain Ethernet adapters are used, which in turn increases the system performance.

The Energy Performance mode can be set to Maximum Performance, Balanced Performance, Balanced Power, or Power Saver (See Figure 7.). Test results demonstrate that most applications run best with the Balanced Performance setting. Applications that are highly I/O sensitive perform best when the Energy Performance option is set to Maximum Performance.

Memory Settings for HPC Workloads

Figure 8 and Figure 9 show the memory settings for HPC workloads.

Figure 8. Memory Settings for HPC Workloads in Cisco UCS E5-based M3 B-Series Blade Servers
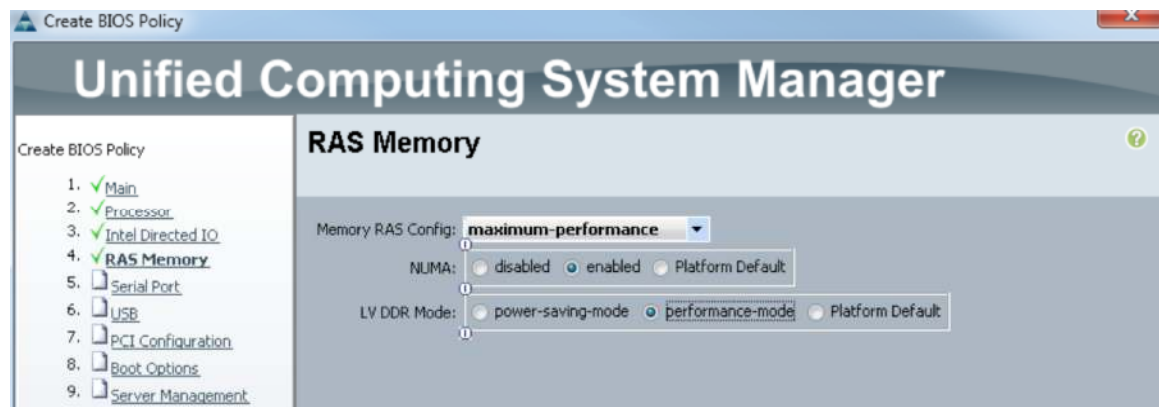
Figure 9. Memory Settings for HPC Workloads in Cisco UCS E5-based M3 C-Series Rack Servers
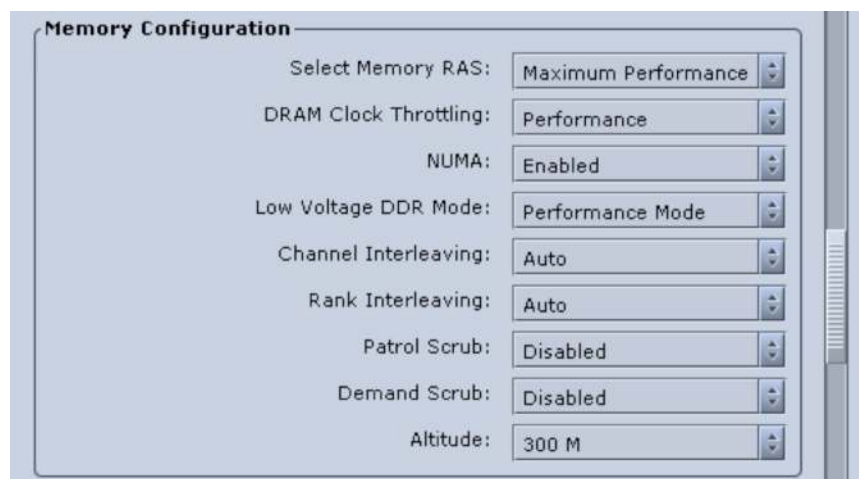
The Non-Uniform Memory Access (NUMA) option should be enabled for HPC workloads so that NUMA can decide on the memory allocations for each thread executed by the HPC applications.

Since the HPC workloads do mostly in-memory processing, it is best to have DIMMs running at the highest available frequency to process the data faster.

Java Enterprise Edition Application Server

Java Platform, Enterprise Edition, or Java EE (earlier referred to as the J2EE) defines the core set of APIs and features of Java application servers. Usually, Java EE applications are of client-server or server-side type, which requires a Java EE application server.

In the industry, Java EE application servers are distinguished as the following:

- Fully compliant application servers that implement the full Java EE stack specification with features like Enterprise JBoss. Examples of the fully compliant application servers are, Apache Geronimo and JBoss Application Server.
- Web application servers, which support only the web-tier of Java EE, including the servlet, and so on. Examples of the fully compliant application servers are Apache Tomcat and Jetty.

Processor Settings for Java Application Servers

Figure 10 and Figure 11 show processor settings for Java EE application servers.

Figure 10. Processor Settings for Java EE Applications in Cisco UCS E5-based M3 B-Series Blade Servers

Figure 11. Processor Settings for Java EE Applications in Cisco UCS E5-based M3 C-Series Rack Servers

The Turbo Boost technology enables higher CPU frequency, which helps in faster processing of application requests. This helps in lower end-user response time. Business scenarios such as batch processes, executed at a certain time of the day, benefit from the Turbo Boost Technology, which enables the CPU cores to clock at higher frequency. This helps lower the batch execution time, thereby, helping in the faster completion of business reports generation.

All the C States should be enabled. This helps in reduced power consumption as only active cores are processing the requests during non-peak hours. If the application demands more CPU cores, the inactive cores will become active and in turn help in higher throughput.

The CPU Performance should be set to Enterprise. When a web server needs to process a large amount of data in a system, the data access pattern is predictable (mostly sequential or adjacent lines are accessed). In this situation, it is desirable to enable the Prefetchers (Mid-Level Cache (MLC) and Data Cache Unit (DCU)) by setting the CPU Performance to Enterprise to reduce the access latency for the memory-bound operations.

Memory Settings for Java EE Application Servers

Figure 12 and Figure 13 show the recommended memory settings.

Figure 12.   Memory Settings for Java EE Applications in Cisco UCS E5-based M3 B-Series Blade Servers
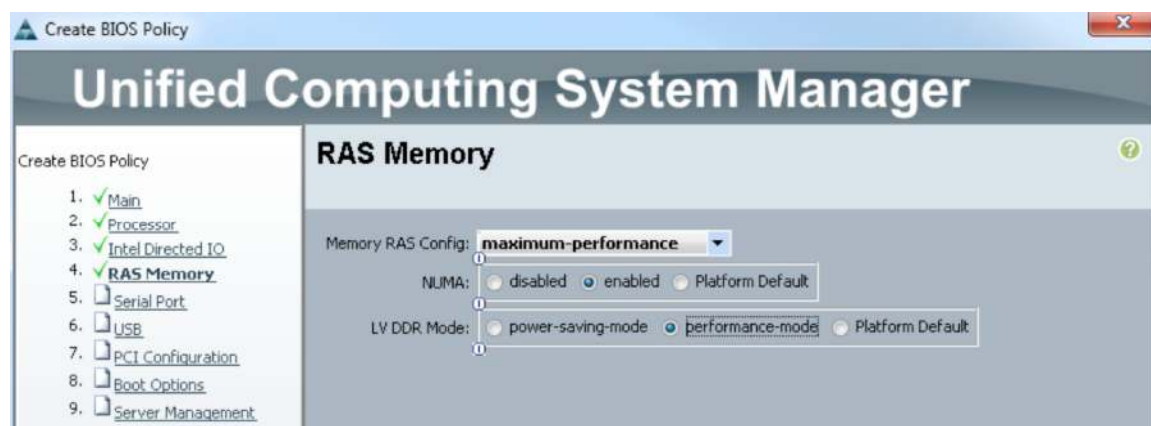
Figure 13.   Memory Settings for Java EE Applications in Cisco UCS E5-based M3 C-Series Rack Servers

The DDR mode should be set to Performance Mode so that the DIMMs work at the highest available frequency for the installed memory/CPU combination. The in-memory enterprise applications such as Terracotta and Ehcache benefit from the high memory speed. If these modes are enabled in web server workloads, the I/O operations will be serviced at maximum frequency and will have reduced memory latency.

Decision Support Systems

The decision support system (DSS) applications are designed to help make decisions based on the data that is picked from a wide range of sources. The DSS applications are not a single information resource, such as a database or a program that graphically represents sales figures, but the combination of integrated resources working together.

Processor Settings for DSS Workloads

Figure 14 and Figure 15 show the recommended processor settings for DSS workloads

Figure 14.   Processor Settings for DSS Workloads in Cisco UCS E5-based M3 B-Series Blade Servers

Figure 15.   Processor Settings for DSS Workloads in Cisco UCS E5-based M3 C-Series Rack Servers

The Intel processors provide a unique combination of enhanced performance capabilities and lower CPU utilization. These processors provide the processing power required to meet the demands of the DSS applications today, as well as the headroom needed to accommodate even more demanding applications in future.

The DSS applications perform operations through parallel processing, and therefore it is essential to enable Turbo Boost. This ensures that each core gets a higher computing power so that multiple parallel requests can be executed efficiently. The DSS systems have longer running queries than OLTP workloads, so Hyper-Threading should be disabled to avoid high query wait time.

In the DSS workloads, the I/O throughput is more important than the IOPS performance. To derive more I/O bandwidth, set the CPU Performance to Enterprise mode. For the C-Series servers, Cisco recommends enabling all the Prefetcher options, including MLC and DCU, to reduce latency for memory-bound operations. For

demanding applications such as DSS, Cisco recommends disabling all the C States to ensure that all the cores are always active so that they can service multiple parallel requests.
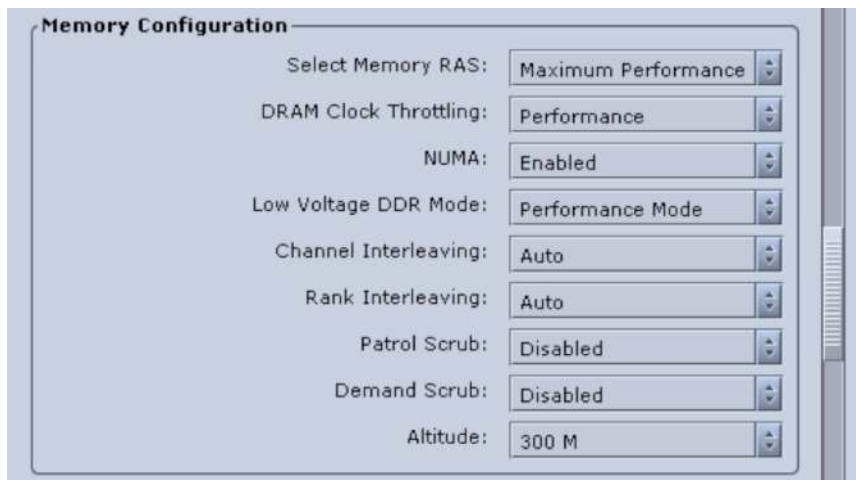
Memory Settings for DSS Workloads

Figure 16 and Figure 17 show the recommended memory settings for DSS workloads.

Figure 16.   Memory Settings for DSS Workloads in Cisco UCS E5-based M3 B-Series Blade Servers
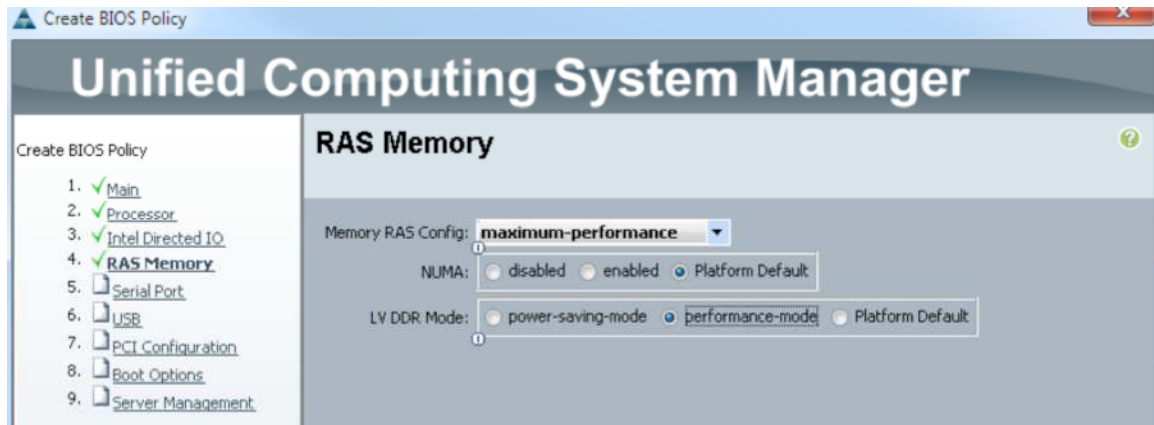
Figure 17.   Memory Settings for DSS Workloads in Cisco UCS E5-based M3 C-Series Rack Servers

The DSS applications benefit greatly from larger and faster memory, therefore, Cisco recommends setting the DDR Mode to Performance Mode so that the DIMMs run at the highest available frequency based on the installed memory/CPU combination. In DSS, the transactions are sequential in nature. If these modes are enabled, the I/O operations will be serviced at the maximum frequency and memory latency will be reduced.
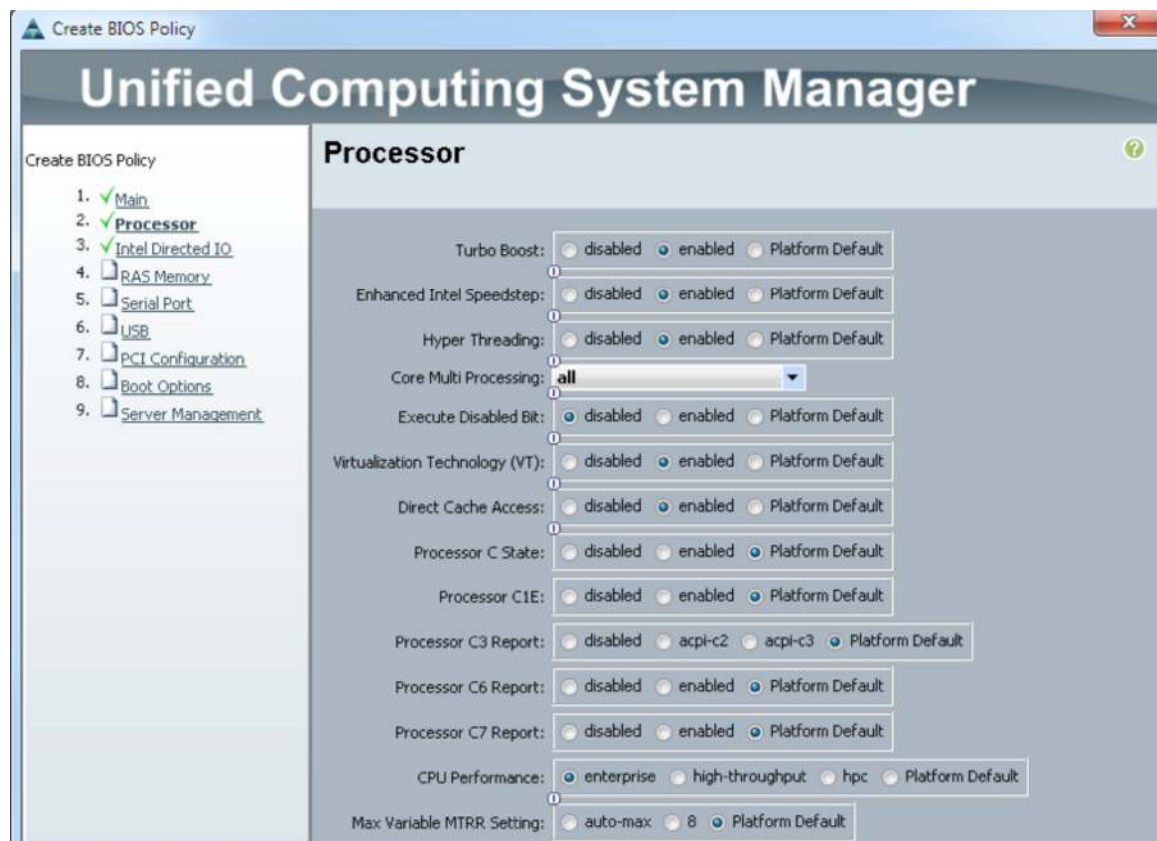
Virtualization

The Intel virtualization technology provides manageability, security, and flexibility in IT environments that use software-based virtualization solutions. With this technique, a single server can be partitioned and can be projected as several independent servers, allowing the server to run different applications on the operating systems simultaneously.

Processor Settings for Virtualization Workloads

Figure 18 and Figure 19 show the recommended processor settings for virtualized workloads.

Figure 18.   Processor Settings for Virtualized Workloads in Cisco UCS E5-based M3 B-Series Blade Servers

Figure 19. Processor Settings for Virtualized Workloads in Cisco UCS E5-based M3 C-Series Rack Servers

Most of the CPU and memory settings for virtualized workloads are the same as discussed in the OLTP workloads section. It is important to enable virtualization technology in BIOS to support the virtualization workloads. Make sure that the Intel VT-d options are enabled.

The CPUs that support hardware virtualization allow the processor to run multiple operating systems in the virtual machines. This involves some overhead because a virtual operating systems performance is comparatively slower than the native OS. To enhance performance, make sure to enable Turbo Boost and Hyper-Threading for the processors.

The Cache Prefetching mechanisms (sometimes called DPL Prefetch, Hardware Prefetchers, L2 Streaming Prefetch, or Adjacent Cache Line Prefetch) usually help systems performance, especially when memory access patterns are regular. As virtualized workloads have random memory access it is recommended to disable the Prefetchers.

Intel Directed I/O for Virtualized Workloads

Figure 20 shows the recommended Directed I/O settings for virtualized workloads.

Figure 20.   Intel Directed I/O Settings for Virtualized Workloads in Cisco UCS E5-based M3 B-Series Blade Servers
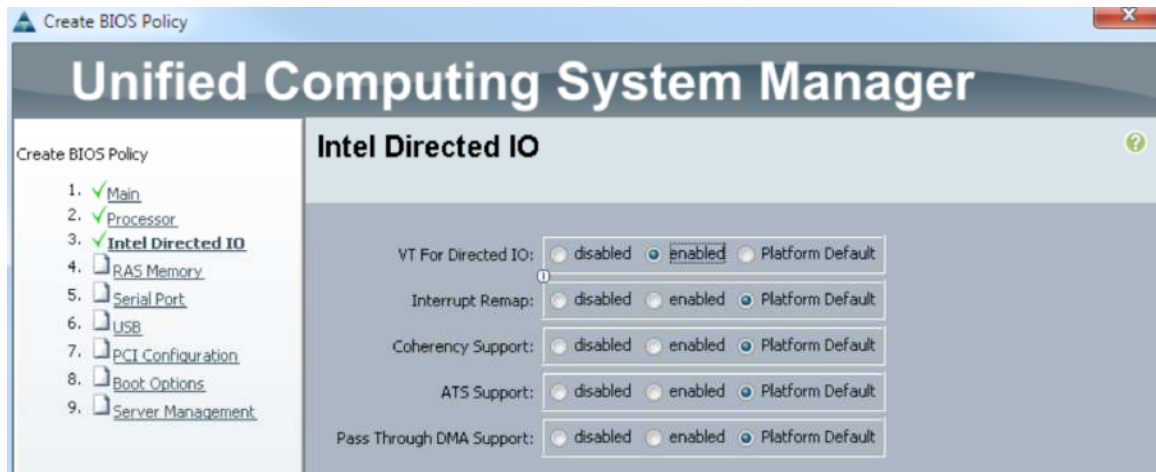

With the Cisco VM-FEX technology, virtual machines can now directly write to the virtual network interface cards (NICs) when the directed I/O option is enabled at the BIOS level.
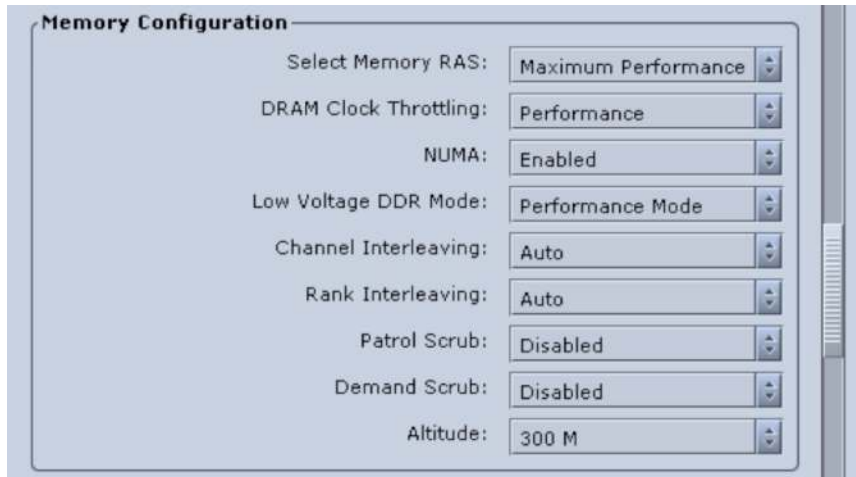
Memory Settings for Virtualized Workloads

Figure 21 and Figure 22 show the recommended memory settings for virtualized workloads

Figure 21.   Memory Settings for Virtualized Workloads in Cisco UCS E5-based M3 B-Series Blade Servers
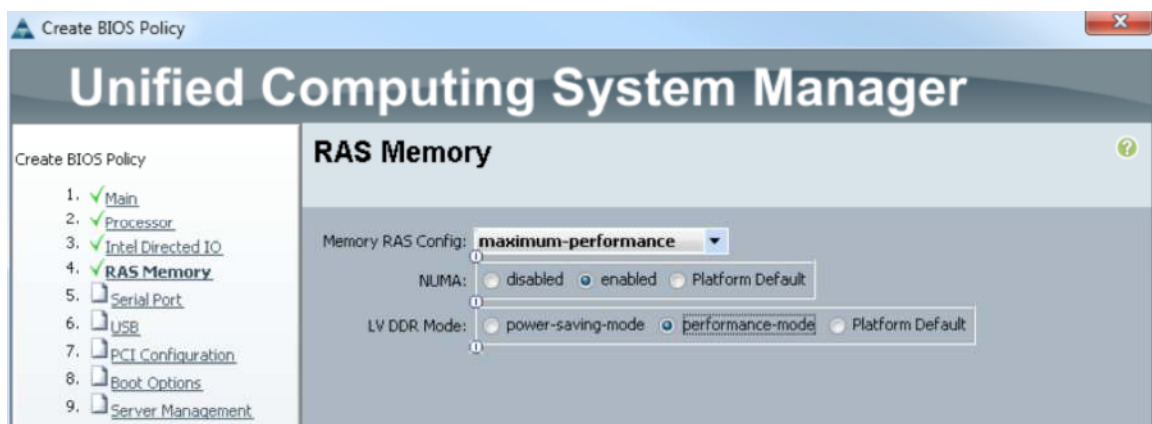
Figure 22.   Memory Settings for Virtualized Workloads in Rack Servers

When running the applications that access memory randomly, enable the RAS memory to maximum performance. This setting results in improved performance for optimal system performance. In virtualized environments, run the memory at the highest frequency to avoid memory latency.

Conclusion

There are several processors, directed I/O, and memory options in BIOS settings that can significantly impact the system performance. When optimizing for system performance, select the BIOS settings that enhance the system performance over those that contribute to power savings.

By enabling or disabling the BIOS options such as the CPU Prefetchers and CPU Hyper-Threading, the system performance can be improved effectively and also optimize the power consumption. When tuning system BIOS settings for performance, it is imperative to consider the various processor and memory options. If performance is the goal, make sure to choose options that optimize for performance in preference to power savings, and experiment with other options such as CPU Prefetchers and CPU Hyper-Threading.

Printed in USA

C11-727827-00   05/13