



# Best Practices When Deploying Linux on the HP ProLiant DL580 Gen8 Server

## Table of contents

Executive summary .....	3
ProLiant DL580 Gen8 server overview.....	3
ProLiant DL580 Gen8 server architecture .....	4
ProLiant DL580 Gen8 server system ROM.....	4
ProLiant DL580 Gen8 server management .....	5
Configuring the ProLiant DL580 Gen8 server .....	5
Pre-configuration tasks.....	5
Memory configuration.....	5
Boot adapter location.....	5
UEFI boot order.....	6
Power Management .....	6
HP Power Regulator .....	7
General purpose workload – OS control.....	7
Maximum Performance power consumption setting .....	8
Enable/disable Collaborative Power Control .....	9
System ROM modes overview.....	10
CPU mappings.....	11
PCI Express slot mappings.....	11
Non-Uniform Memory Access overview .....	12
Displaying node distances .....	12
Displaying task information.....	13
Displaying memory allocation statistics .....	13
Controlling tasks and memory object placement.....	14
Moving tasks to other NUMA nodes .....	15
Determining task memory allocation location .....	15
OS support matrix for the ProLiant DL580 Gen8 server.....	15
Test Linux OS workloads with Hyper-Threading (HT) .....	16
Red Hat Enterprise Linux 5.x.....	17
Disable x2APIC with Intel Xeon E-7 series processors .....	17
Verify Legacy BIOS boot mode before OS installation.....	18

Use NOOP in the I/O scheduler .....	19
Increase crash kernel size .....	19
Red Hat Enterprise Linux 6.4 and SUSE Linux Enterprise Server 11 SP3 best practices .....	20
Single Root I/O Virtualization Support.....	20
Secure Boot .....	20
Test Red Hat Enterprise Linux 6.4 (and later) and SUSE Linux Enterprise server 11 SP3 workloads with Hyper-Threading (HT).....	21
Increase crash kernel size .....	21
x2APIC support .....	22
Performance and tuning.....	22
Additional software components .....	23
HP SPP .....	23
HP System Management Homepage.....	24
Resources .....	26

## Executive summary

The purpose of this white paper is to facilitate the successful deployment and configuration of Linux operating systems on HP ProLiant DL580 Gen8 Servers. It includes a description of recommended HP software components for the DL580 Gen8 server, and best practices you can use when deploying Linux on the DL580 Gen8 server. This document specifically discusses Red Hat® Enterprise Linux and SUSE Linux Enterprise server, although the information may apply to other similar Linux distributions.

## ProLiant DL580 Gen8 server overview

The DL580 Gen8 server is a four-socket (4S) scale-up platform designed for performance, efficiency, and reliability. To meet the needs of your environment and to utilize the Linux server operating system to its full potential, the DL580 Gen8 server offers many benefits.

Figure 1. HP ProLiant DL580 Gen8 Server



Featuring Intel® Xeon® E7-4800/8800 v2 processors, the DL580 Gen8 server delivers blazing fast results with enhanced processor performance, more memory slots, greater I/O bandwidth, and increased storage performance. It also offers security and data protection features for system resiliency that your business can depend on. With intelligent manageability through HP OneView, HP Integrated Lights-Out 4 (iLO 4), and user-inspired features, you can achieve faster, lower cost infrastructure management.

The DL580 Gen8 server is well suited for data-intensive workloads where performance and scalability are primary concerns. These environments include large online transaction processing and decision support databases; large-scale consolidation and virtualization; and resource-intensive applications, such as customer relationship management (CRM), capacity requirements planning (CRP), and enterprise resource planning (ERP).

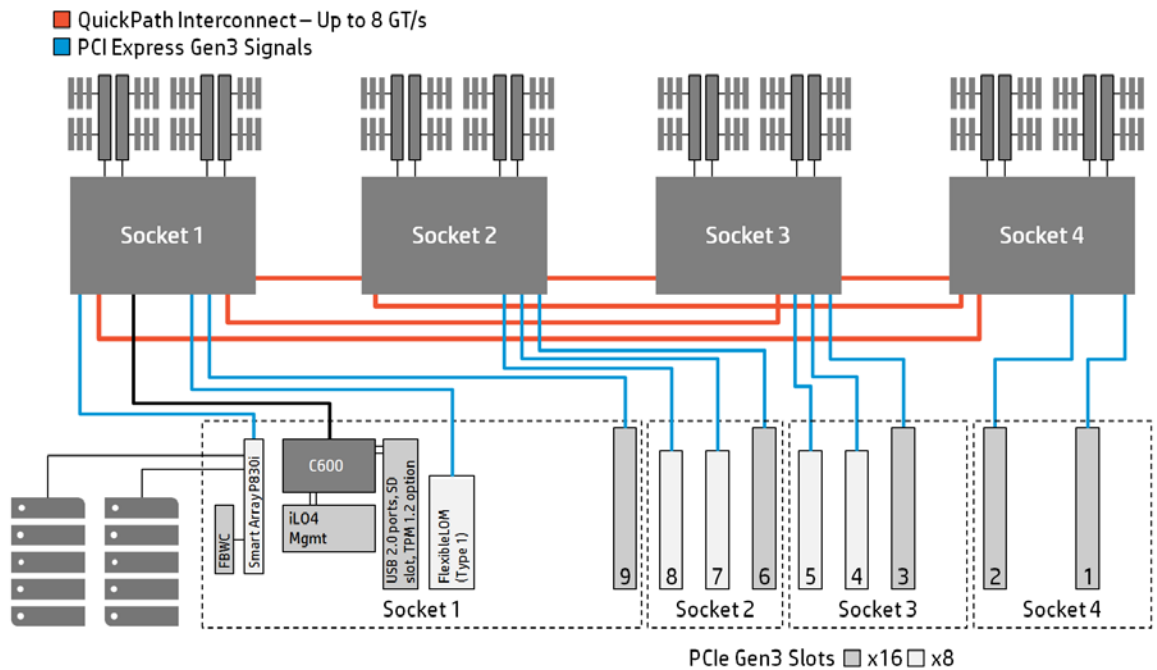
For more information and technical specifications for this server, see the DL580 Gen8 server web page at [hp.com/servers/dl580-gen8](http://hp.com/servers/dl580-gen8).

## ProLiant DL580 Gen8 server architecture

The DL580 Gen8 server architecture lets you configure your server to meet the needs of demanding, data-intensive workload environments. The basis of this architecture for the DL580 Gen8 server is the Non-Uniform Memory Architecture (NUMA) Intel® processors. Each processor has its own set of memory and is directly connected to a set of peripherals. As illustrated in Figure 2, key features of the DL580 Gen8 server architecture include:

- Processors: Having up to four processor sockets, you can achieve the performance that you need by configuring your DL580 Gen8 server with up to four Intel® Xeon® E7-4800/8800 version 2 (V2) processors. These processors have up to 6, 8, 10, 12, or 15 cores per processor.
- Memory: Having 96 DDR3 DIMM slots, you can achieve the maximum memory configuration of 3 TB by populating all 96 DIMM slots with 32 GB DDR3 DIMMs. As larger DIMMs become available, this maximum memory configuration will increase.
- I/O expansion: Nine slots give you the flexibility to add I/O adapters to meet your environment’s needs:
  - Five, PCIe Gen3, 225 Watt, x16 slots
  - Four, PCIe Gen3, 300 Watt, x8 slots for adapters that require a higher wattage, such as Graphic Processing Units (GPUs)

Figure 2. HP ProLiant DL580 Gen8 Server architecture



## ProLiant DL580 Gen8 server system ROM

The DL580 Gen8 server lets you choose the System ROM mode that best suits your environment. You can choose between the Legacy BIOS Mode and Unified Extensible Firmware Interface (UEFI). The Legacy BIOS Mode is compatible with all supported Linux distributions while UEFI is only compatible with RHEL6.4 and SLES11 SP3. For more information about UEFI and the DL580 Gen8 server, see the HP documentation library at [hp.com/go/uefi/docs](http://hp.com/go/uefi/docs).

The DL580 Gen8 server ships with Legacy mode enabled. If you choose to enable UEFI mode, do so through the RBSU. For information about Legacy mode and UEFI, see the “[System ROM modes overview](#)” section of this paper.

## ProLiant DL580 Gen8 server management

The ProLiant DL580 Gen8 server includes the complete set of embedded management features that support the lifecycle of ProLiant servers, from initial deployment, through ongoing management, to service alerting. These embedded management features include:

- HP OneView, which provides integrated management.
- HP Intelligent Provisioning, which reduces system deployment time.
- iLO Management engine with HP Agentless Management, which provides embedded intelligence for managing your server.
- Smart Update Solution, which reduces the time required for system maintenance and updates.

---

### Note

The following pages contain numerous examples. The text displayed in the examples may differ from what you see in your specific environment, depending on the configuration of your system and the OS version used.

---

## Configuring the ProLiant DL580 Gen8 server

This section provides recommendations for running any of the supported operating systems on the ProLiant DL580 Gen8 server. Before installing the operating system, make sure that you have considered these recommendations, and implement them as necessary for your environment.

### Pre-configuration tasks

Before configuring the ProLiant DL580 Gen8 server and installing any operating system, HP recommends that you make sure that your server has the latest firmware, drivers, and HP software. For optimal performance and functionality, complete the following tasks:

- Review the HP ProLiant DL580 Gen8 server User Guide.  
The user guide provides additional information to help you configure your server for optimal performance, availability, and manageability. It also discusses software that can help in the deployment, management, and support of the DL580 Gen8 server. System ROM updates are available from the HP Support Center at [hp.com/go/support](http://hp.com/go/support).
- Determine whether you will use Legacy mode or UEFI mode. The DL580 Gen8 server ships with Legacy mode enabled. If you choose to use UEFI mode, use the ROM-Based Setup Utility (RBSU) to change the configuration. Changing between Legacy mode and UEFI mode after OS installation will require the OS to be re-installed.
- Use the HP Service Pack for ProLiant (HP SPP) to make sure that you have the latest firmware, drivers, and HP software installed on your server. The HP SPP is a comprehensive system software and firmware solution delivered as a single ISO file. Using the HP Smart Update Manager (HP SUM) as the deployment tool, the HP SPP installs the latest HP firmware, drivers, and software. The DL580 Gen8 server requires HP SPP Version 2014.02.0 (or later). You can download the HP SPP at [hp.com/go/spp](http://hp.com/go/spp) or obtain the HP SPP CD on the HP Support page at [hp.com/support](http://hp.com/support).

You can read more about the HP firmware and software mentioned in the “[Additional software components](#)” section of this document.

### Memory configuration

Dual and quad rank DIMMs should be used in high performance implementations. In addition, for proper and optimal DIMM installation guidelines, consult the “HP ProLiant DL580 Gen8 server User Guide” at: [hp.com/servers/dl580-gen8](http://hp.com/servers/dl580-gen8).

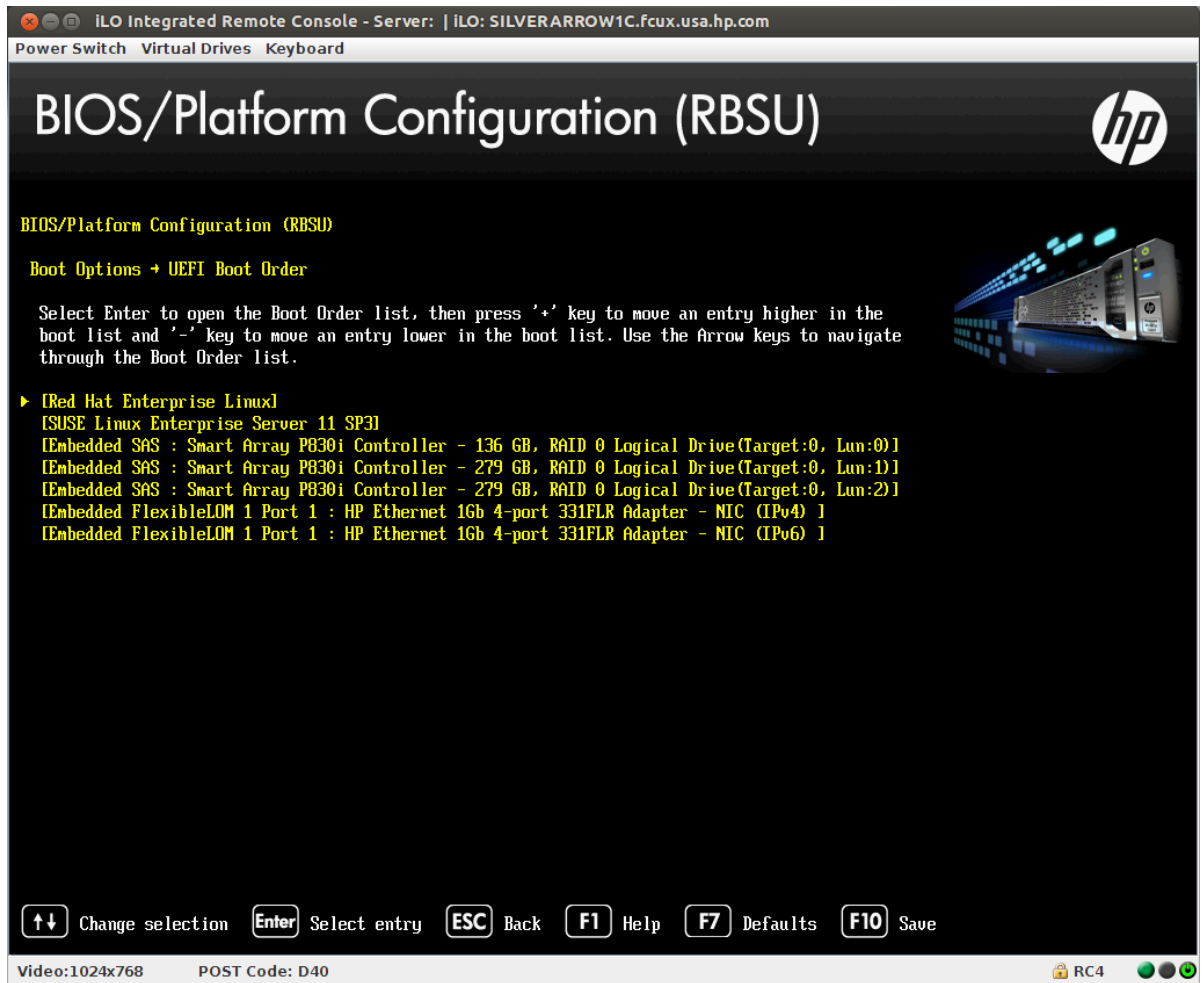
### Boot adapter location

On the DL580 Gen8, slots 7 through 11 are the first ones scanned during the boot phase. Therefore, the boot adapter must be installed in one of these slots in order for it to be enumerated first. Once that is accomplished, the boot order is prioritized by the BIOS settings, as usual.

## UEFI boot order

You can change the UEFI boot order by following the indications in Figure 3.

Figure 3. UEFI boot order



You change the boot order from RBSU "BIOS/Platform Configuration (RBSU)"

--> Boot Options --> UEFI Boot Order menu

## Power Management

Linux and the HP ProLiant servers together use processor-based features to achieve better power efficiency for processors. The processor-based features include:

- Performance states (P-states) define a set of fixed operating frequencies and voltages, where P0 represents the highest operating frequency and voltage. You can save power by entering P-states with lower frequency and voltage levels. Either the platform firmware or the operating system controls the P-states.
- Power states (C-states), excluding the C0 state, represent idle states and determine the power consumed when a processor is idle. C0 is a non-idle state with higher C-states representing idle conditions with increasing power savings. The operating system controls the C-states.
- Throttle states (T-states) define a set of fixed frequency percentages, which can be used to regulate the power consumption and the thermal properties of the processor. ProLiant systems may reserve the use of T-states for the system firmware.

In addition, ProLiant servers are also capable of using the various processor states to support innovative power management features that are operating system independent and are implemented in the hardware and firmware:

- HP Power Regulator provides a facility to control processor power usage and performance efficiently, either statically or dynamically depending on the mode selected.
- HP Power Capping allows an administrator to limit the power consumed by a server.
- HP Dynamic Power Capping has the additional feature of ensuring that the power limit set by an administrator is maintained by reacting to a spike in server workload more rapidly than basic HP Power Capping.

The Power Regulator and Power Capping technologies are designed to work in conjunction with each other. To make the operating system aware of Power Capping, HP provides the Collaborative Power Control technology. This is a two-way communication mechanism established between the operating system and platform firmware, and can be used by the operating system and hardware collaboratively to choose the appropriate performance level for the server. Support for this technology is present in both Linux and on ProLiant Gen8 servers.

## HP Power Regulator

HP Power Regulator is a configurable processor power usage feature, which allows you to choose from several options for the server to manage P-states or to delegate control of regulating P-states to the operating system. HP Power Regulator is implemented within the firmware on both Intel-based and AMD-based ProLiant servers. ProLiant servers provide the HP Power Regulator modes shown in Table 1, selectable from the ROM Based Setup Utility (RBSU) or through HP iLO 4.

**Table 1.** HP Power Regulator modes

HP Power Regulator Modes	Mode function
HP Dynamic	The firmware is capable of managing the P-states. However, when the Collaborative Power Control (CPC) setting is enabled in RBSU, the OS and the firmware collaborate to attain the desired frequency for a processor. When CPC is disabled, this mode allows the firmware to exclusively control the P-states of a processor to match the server load. On HP ProLiant Gen8 servers, HP Dynamic is the default mode with the CPC setting enabled.
HP Static Low	The firmware controls the P-states. The P-state of the processor is static and it is set to the P-state, which corresponds to the lowest operating frequency supported by the processor.
HP Static High	The firmware controls the P-states. The P-state of the processor is static and it is set to P0, which corresponds to the highest operating frequency supported by the processor.
OS Control	The operating system controls the P-states and it manages the P-states according to the policy set by the administrator via the OS.

## General purpose workload – OS control

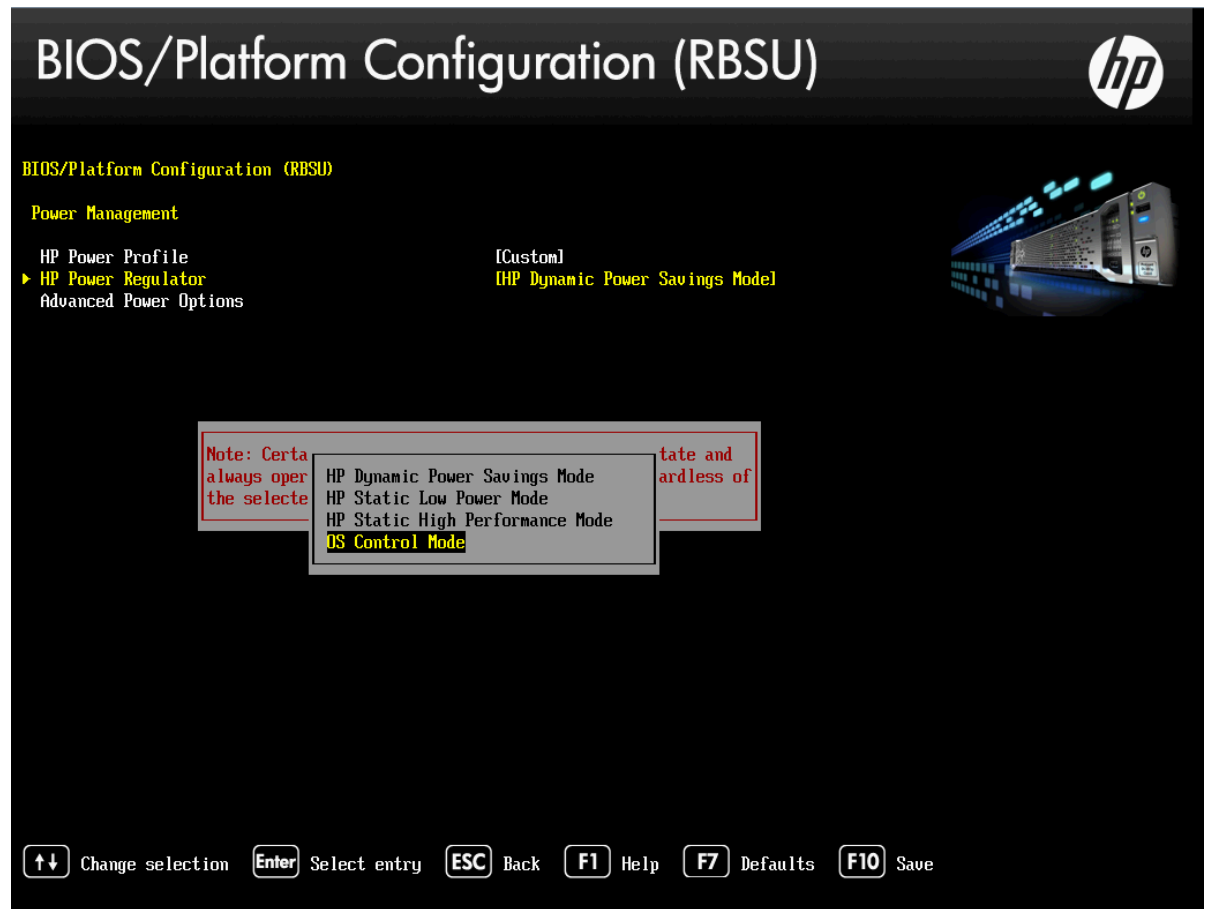
For general purpose workloads where a balance of high performance and power efficiency is desired, use OS control mode. To take advantage of the Linux operating system’s ability to adjust server power consumption (and reduce costs) based on workload, follow these steps:

1. During system boot, press **F9** to run the System Utilities
2. Select System Configuration
3. Select BIOS/Platform Configuration (RBSU)
4. Select **Power Management**.
5. Select **HP Power Regulator**
6. Select **OS Control Mode** (see Figure 4)
7. Exit all RBSU menu screens using the **Esc** key
8. Press the **‘Y’** key to confirm your changes

### Note

OS Control mode is not recommended for implementations requiring predictable high performance.

Figure 4. HP Power Regulator OS Control Mode



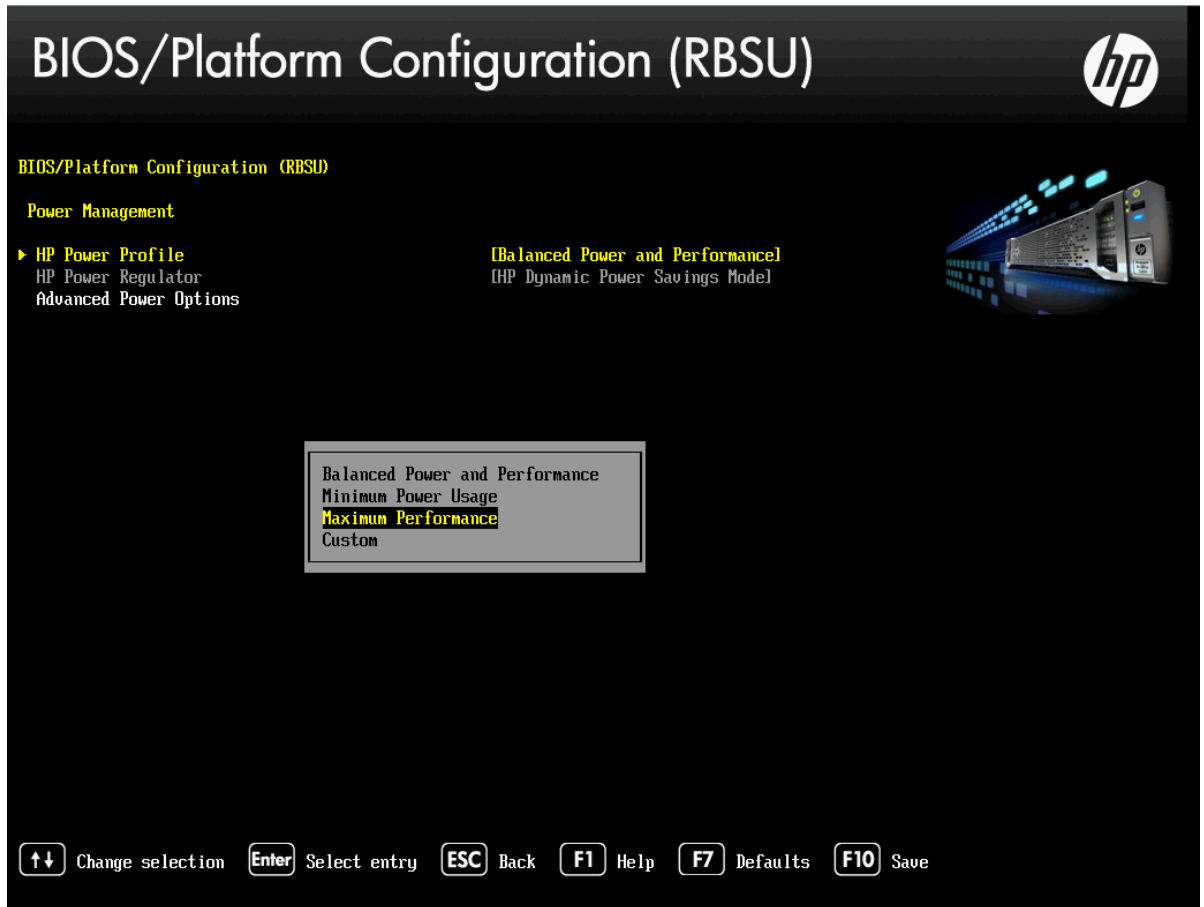
## Maximum Performance power consumption setting

To obtain maximum performance and the most predictable latency, disable all power saving modes by following these steps:

1. During system boot, press **F9** to run the System Utilities.
2. Select **System Configuration**.
3. Select **BIOS/Platform Configuration (RBSU)**.
4. Select **Power Management**.
5. Select **HP Power Profile**.
6. Select **Maximum Performance** (see Figure 5).
7. Exit all RBSU menu screens using the **Esc** key.
8. Press the **Y** key to confirm your changes.



Figure 5. HP Power Profile Maximum Performance



When adding I/O adapters in the PCIe Gen3 expansion slots, note that the number of processors installed in the DL580 Gen8 server determines the PCIe Gen3 x8 or x16 slots that are available for optional I/O cards adapters.

Slot to socket mappings is:

- Socket 1 - slot(s) 9
- Socket 2 - slot(s) 6,7,8
- Socket 3 - slots(s) 3,4,5
- Socket 4 - slots(s) 1,2

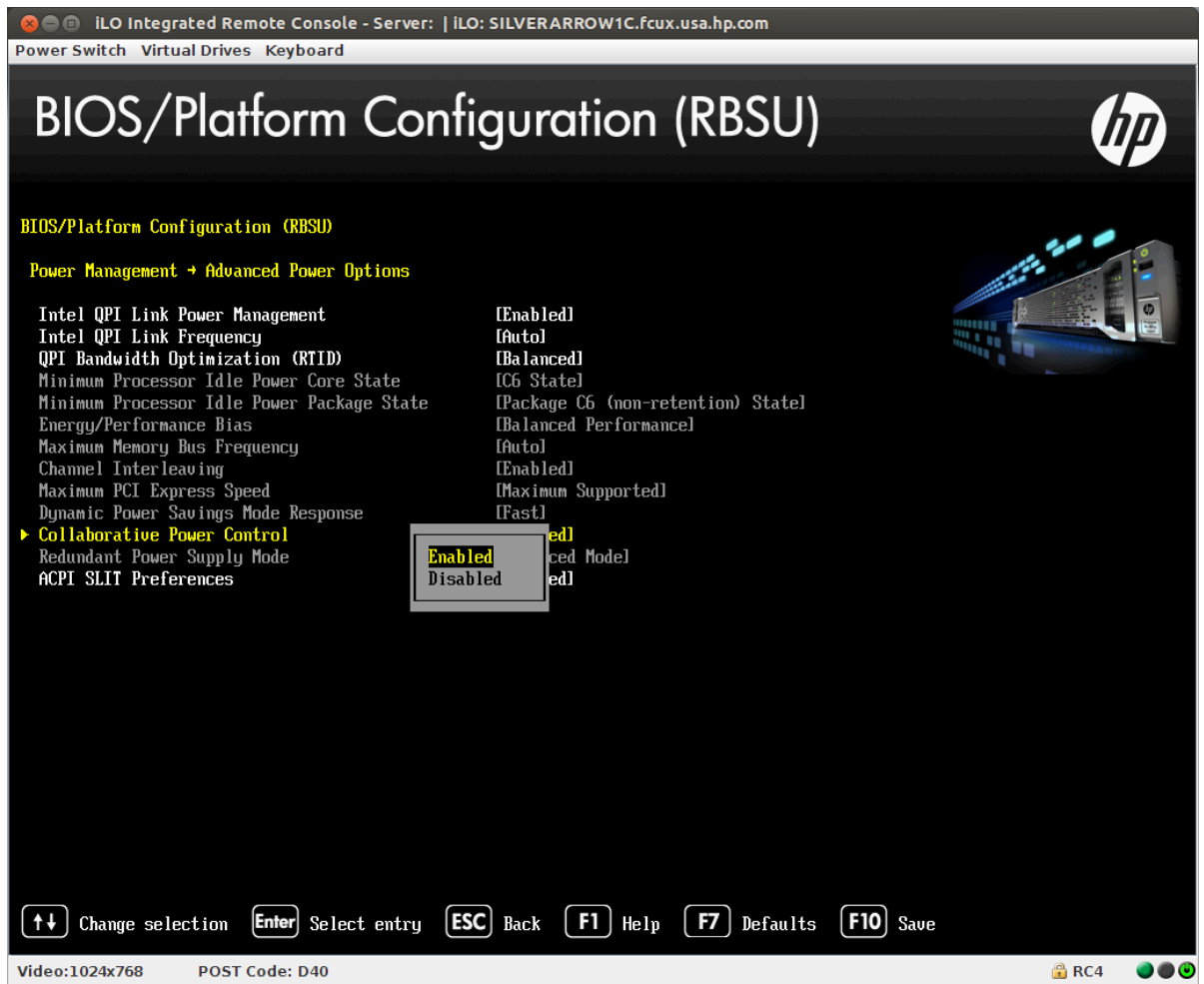
It is important to remember that if a socket is not populated, its corresponding slots are not usable.

### Enable/disable Collaborative Power Control

The RBSU menu shown in Figure 6 will allow you to enable and disable Collaborative Power Control (CPC) as required.

1. During system boot, press **F9** to run the System Utilities
2. Select **System Configuration**.
3. Select **BIOS/Platform Configuration (RBSU)**.
4. Select **Power Management**.
5. Select **Advanced Power Options**.
6. Select **Collaborative Power Control**, and then choose **Disabled** (see Figure 6).
7. Exit all RBSU menu screens using the **Esc** key.
8. Press the **Y** key to confirm your changes

Figure 6. Power Management Collaborative Power Control



## System ROM modes overview

The DL580 Gen8 server gives you the choice of configuring the System ROM for Legacy mode or UEFI mode. Legacy mode and UEFI mode have different schemes for partitioning and booting the operating system. The Legacy mode necessitates a Master Boot Record (MBR) boot device, whereas UEFI mode requires a Globally Unique Identifier (GUID) Partition Table boot device. UEFI mode offers a standard firmware interface that eases configuration and troubleshooting and allows support of large boot devices.

After configuring your server for either Legacy BIOS Mode or UEFI mode, you must install the operating system to match the mode chosen for your server configuration.

For more information about UEFI and the DL580 Gen8 server, see the HP documentation library at [hp.com/go/uefi/docs](http://hp.com/go/uefi/docs) and the DL580 Gen8 server home page at [hp.com/servers/dl580-gen8](http://hp.com/servers/dl580-gen8).

## CPU mappings

Linux uses the “/proc/cpuinfo” file to enumerate logical processors. As an example, Table 2 shows the logical processor assignments for the first processor of a 4 socket, 15 core DL580 Gen8 server with Hyper Threading enabled. The Physical ID corresponds to the physical processor socket. Core ID is the processing core on a particular physical processor. Thus, logical processor numbers 0 and 60 are the two threads on the first core of the first physical processor. If Hyper Threading is not enabled, only cores 0-14 would be listed.

**Table 2.** CPU mappings

Logical Processor	Physical ID	Core ID	Logical Processor	Physical ID	Core ID
0	0	0	60	0	0
1	0	1	61	0	1
2	0	2	62	0	2
3	0	3	63	0	3
4	0	4	64	0	4
5	0	5	65	0	5
6	0	6	66	0	6
7	0	7	67	0	7
8	0	8	68	0	8
9	0	9	69	0	9
10	0	10	70	0	10
11	0	11	71	0	11
12	0	12	72	0	12
13	0	13	73	0	13
14	0	14	74	0	14

When using the **numactl** and **taskset** commands for process and memory object placement, refer to the information in the “/proc/cpuinfo” file. For more information about these commands, see the numactl (8) and taskset (1) manpages. These commands are discussed in the following two sections, respectively: [“Controlling tasks and memory object placement”](#) and [“Moving tasks to other NUMA nodes.”](#)

## PCI Express slot mappings

PCI slot bus numbers are documented in the “HP ProLiant DL580 Gen8 server” QuickSpecs. You can use the document to map a particular I/O slot to its corresponding PCI device listed by the lspci command. The following example shows a dual-port Fibre Channel HBA using bus 0x06, which maps to slot 9:

```
# lspci | grep -e "06:00"
06:00.0 Fibre Channel: Emulex Corporation Lancer-X: LightPulse Fibre Channel Host Adapter (rev 30)
06:00.1 Fibre Channel: Emulex Corporation Lancer-X: LightPulse Fibre Channel Host Adapter (rev 30)
```

## Non-Uniform Memory Access overview

The latest Intel® Xeon® processors provide Socket Local Memory, which consists of embedded memory controllers that access memory DIMMS connected to the socket. A processor and the associated Socket Local Memory are considered a node. Each node can be thought of as a separate Symmetric Multi-Processing (SMP) system where accesses to the Socket Local Memory have the same latency. However, memory accesses from a processor to the physical memory (or cached memory) of another node have greater latencies. The resulting architecture of a system having different latencies for accessing memory is known as a Non-Uniform Memory Access (NUMA) architecture.

### Displaying node distances

The distance between the local and remote node affects memory access latency. The greater the distance, the higher are the memory access latencies. To display node memory statistics and relative distances between nodes, use the `numactl` command (for more information, see the `numactl(8)` manpage). For example:

```
# numactl --hardware
available: 4 nodes (0-3)

node 0 cpus: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 60 61 62 63 64 65 66 67 68 69 70
71 72 73 74
node 0 size: 32665 MB
node 0 free: 31088 MB

node 1 cpus: 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 75 76 77 78 79 80 81 82
83 84 85 86 87 88 89
node 1 size: 32768 MB
node 1 free: 30374 MB

node 2 cpus: 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 90 91 92 93 94 95 96 97
98 99 100 101 102 103 104
node 2 size: 32768 MB
node 2 free: 31853 MB

node 3 cpus: 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 105 106 107 108 109 110
111 112 113 114 115 116 117 118 119
node 3 size: 32768 MB
node 3 free: 31778 MB

node distances:
node 0 1 2 3
0: 10 21 21 21
1: 21 10 21 21
2: 21 21 10 21
3: 21 21 21 10
```

## Displaying task information

You can also use **numactl** command to display the memory allocation policy and processor affinity for the current shell. The following example shows the default NUMA task policy and CPU bindings:

```
# numactl --show policy: default
policy: default
preferred node: current
physcpubind: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52
53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79
80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104
105 106 107 108 109 110 111 112 113 114 115 116 117 118 119
cpubind: 0 1 2 3
nodebind: 0 1 2 3
membind: 0 1 2 3
```

In this example, there is no specific binding because the current shell can be placed anywhere in the system and use any memory location. For an example showing the use of the **numactl** command for performing an explicit binding, see the [“Controlling tasks and memory object placement”](#) section. For more examples, see the `numactl(8)` manpage.

In addition to displaying CPU and memory bindings, the **numactl** command can control NUMA policies for shared memory. For more information, see the `numactl(8)` manpage.

---

### Note

The “--show policy” option does not allow specification of a particular process for viewing.

---

You can read more about NUMA in the “Red Hat Linux NUMA Support for HP ProLiant servers” white paper at [hp.com/portal/site/hpsc](http://hp.com/portal/site/hpsc).

## Displaying memory allocation statistics

To display memory allocation statistics per NUMA node, use the **numastat** command. For example:

```
# numastat
          node0          node1          node2          node3
numa_hit      444225      436242      624763      548726
numa_miss          0          0          0          0
numa_foreign    0          0          0          0
interleave_hit  18955      18959      18953      18962
local_node     413387      389672      591172      506741
other_node     30838      46570      33591      41985
```

Where:

- `numa_hit` is the number of successful allocations for a node.
- `numa_miss` is the number of failed allocations for a node because of low memory (allocations are successful on another node).
- `numa_foreign` is the number of allocations that were intended for another node but ended up on the local node.
- `interleave_hit` is the number of interleave policy allocations that were intended for a specific node and succeeded there.
- `local_node` is incremented when a process running on the local node has allocated memory on the same node.
- `other_node` is incremented when a process running on another node allocated memory on that node.

## Controlling tasks and memory object placement

Best performance is obtained when tasks and their associated memory objects are close together, preferably in the same node. To control the placement of memory objects and task threads, use either the `numactl` command or the `cpuset` command.

### Using the numactl command

For example, the following command binds a shell, its children processes, and associated memory objects to node 1:

```
# numactl --membind=1 --cpunodebind=1 bash
# numactl --show policy: bind
preferred node: 4
physcpubind: 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 45 46 47 48 49 50 51 52
53 54 55 56 57 58 59 cpubind: 1
nodebind: 1
membind: 1
```

The `numactl` command uses the NUMA policy library (`libnuma`) interface for policy retrieval and control. For more information, see the `numactl` (8) and `numa` (3) manpages. You can find the original “A NUMA API for Linux” white paper is at: [halobates.de/numaapi3.pdf](http://halobates.de/numaapi3.pdf).

### Using the cpuset command

The interface to cpusets is the `cpuset` pseudo-filesystem, which was first introduced on Red Hat Enterprise Linux 5 and SUSE Linux Enterprise server 10 and remains available today. Once the pseudo-file system is mounted, you can use the `cpuset` command to partition system resources by creating subdirectories under the mount point. These subdirectories, or cpusets, contain pseudo-files and are used for specifying resources such as the logical CPUs allowed and the memory nodes.

In the following example, the first command verifies that the kernel is aware of the `cpuset` subsystem. The next commands mount the subsystem and list the pseudo-files automatically created in a subdirectory called “mycpuset.” The `echo` commands place logical processors 2 and 3 and memory node 1 into this `cpuset`, and attach the current shell to it.

```
# grep cpuset /proc/filesystems
nodev cpuset
# mkdir /dev/cpusets
# mount -t cpuset none /dev/cpusets
# mkdir /dev/cpusets/mycpuset
# cd /dev/cpusets/mycpuset
# ls
cgroup.clone_children    cpuset.memory_pressure
cgroup.event_control    cpuset.memory_spread_page
cgroup.procs             cpuset.memory_spread_slab
cpuset.cpu_exclusive    cpuset.mems
cpuset.cpus              cpuset.sched_load_balance
cpuset.mem_exclusive    cpuset.sched_relax_domain_level
cpuset.mem_hardwall     notify_on_release
cpuset.memory_migrate   tasks
```

For more information about the `cpuset` command, see the `cpuset` (7) manpage.

Red Hat Enterprise Linux 6 and SUSE Linux Enterprise server 11 introduced Linux Control Groups or `cgroups`, which use the `cpuset` subsystem. For more information, see the “Red Hat Enterprise Linux 6 Resource Management Guide” at: [redhat.com/site/documentation/en-US/Red\\_Hat\\_Enterprise\\_Linux/6/html/Resource\\_Management\\_Guide/index.html](http://redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Resource_Management_Guide/index.html).

## Moving tasks to other NUMA nodes

To move a task from one NUMA node to another by changing the CPU affinity, use the standard **taskset** command. The **taskset** command changes the CPU affinity only. It does not move the physical memory pages allocated by the task. To move the physical pages to the new NUMA node, use the **migratepages (8)** command. For more information about these commands, see the **taskset (1)** and **migratepages (8)** manpages.

In the following example, the first **taskset** command returns the CPU affinity mask for PID 6371, representing logical processors 0 to 79. The second **taskset** command sets a new affinity to logical processors 10 to 19 for this PID. To minimize memory latency, the third command (**migratepages**), migrates the associated memory pages from node 0 to node 1.

```
# taskset -p 6371
pid 6371's current affinity mask: ffffffffffffffffffffffff
# taskset -pc 10-19 6371
pid 6371's current affinity list: 0-79
pid 6371's new affinity list: 8-15
# migratepages 6371 0 1
```

## Determining task memory allocation location

To determine a task’s memory allocation location, examine the “/proc/<pid>/numa\_maps” file. This file displays each memory object for a particular task. The following example shows the entries for the current shell’s heap and stack:

```
# grep -e heap -e stack /proc/$$/numa_maps
7f830068f000 default heap anon=268 dirty=268 active=257 N0=40 N1=191 N2=16 N3=21
7fff64f03000 default stack anon=7 dirty=7 N0=2 N1=1 N2=1 N3=3
```

The first field of each row is the start of the Virtual Memory Address (VMA) range (for example, 7f830068f000 or 7fff64f03000).

The second field is the memory allocation policy (for example, default, which is the System Default Policy). The third field is the path to the mapped file or the use of the shared memory segment (in this example, heap or stack).

The anon= and dirty= show the number of pages. The N<node>= shows the number of pages allocated from each <node>. For a complete description of these fields, see the numa(7) manpage.

## OS support matrix for the ProLiant DL580 Gen8 server

To run Linux and Linux applications optimally, you must understand how to leverage the capabilities and resources of the DL580 Gen8 server. A good starting point is to review the selected OS support matrix in Table 3.

**Table 3. OS support matrix**

Operating System	Max CPUs HP-supported/ OS limit	Max Memory (RAM) HP-supported/OS limit	Notes
Red Hat Enterprise Linux 5.9 (or later) x86-64	160 / 255	1 / 1 TB	No x2APIC support
Red Hat Enterprise Linux 6.4 (or later) x86-64	160 / 4096	3 / 64 TB	
SUSE Linux Enterprise Server 11 Service Pack 3 (or later) x86-64	160 / 4096	3 / 64 TB	

For more information, refer to the following documentation:

[redhat.com/rhel/compare](http://redhat.com/rhel/compare)

[suse.com/products/server/technical-information/](http://suse.com/products/server/technical-information/)

## Test Linux OS workloads with Hyper-Threading (HT)

Hyper-Threading (HT) refers to Intel's proprietary technology for increasing parallel computational power (processor multi-tasking). With HT, the entire pipeline of the processor core is set up to recognize two separate streams of instructions (one for each hardware thread). All the resources in the pipeline are either shared or duplicated between the two hardware threads. Various events cause the pipeline to switch from decoding/dispatching one stream to decoding/dispatching the other stream. For example, when one thread is waiting for data from a memory read, the pipeline is stalled and idle; HT takes advantage of these idle situations to get more work done by switching the pipeline to execute instructions from the other thread.

The benefit from HT depends highly on the application workload and server configuration. While some applications see an overall performance boost, some applications see no or negative performance improvement. HP highly recommends testing HT with your particular workload.

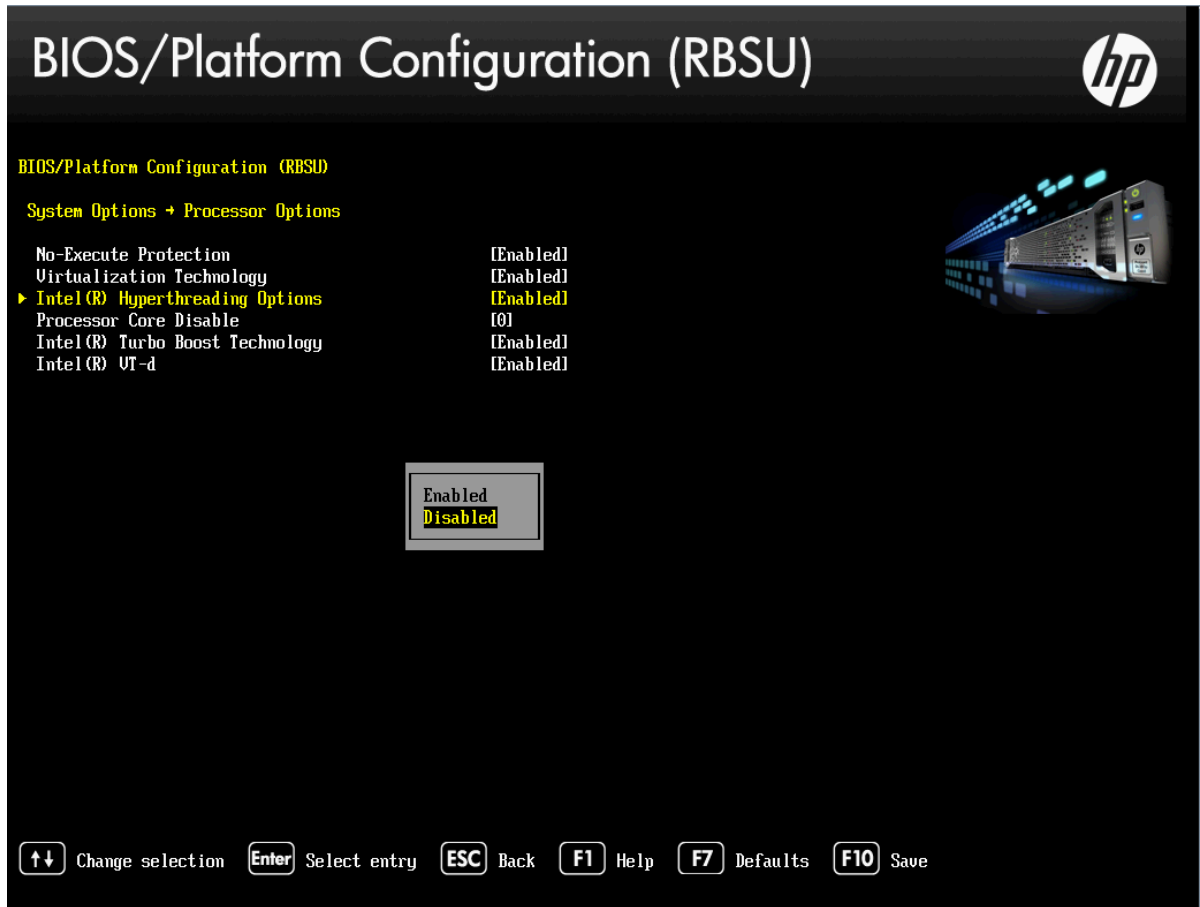
While Linux can support over 160 logical processors, it is not optimized for these high logical processor counts. This can result in increased kernel CPU consumption when attempting to schedule tasks across the available logical processors. Unless your workload is proven to benefit from Hyper-Threading, HT should be disabled to reduce this system overhead. HT can always be enabled later if additional CPU capacity is required.

Hyper Threading is controlled by the system BIOS and is enabled by default on the DL580 Gen8. To enable or disable Hyper-Threading, follow these steps:

1. During system boot, press **F9** to run the System Utilities.
2. Select System Configuration.
3. Select BIOS/Platform Configuration (RBSU).
4. Select **System Options**.
5. Select **Processor Options**.
6. Select **Intel(R) Hyperthreading Options** (see Figure 7).
7. Select **Disabled** to turn Hyper-Threading OFF or Select **Enabled** to turn Hyper-Threading ON.
8. Exit all RBSU menu screens using the **Esc** key.
9. Press the **'Y'** key to confirm your changes.



Figure 7. Hyper Threading options



For more information about HP’s RBSU, see the “HP ROM-Based Setup Utility User Guide” at: [hp.com/support/rbsu](http://hp.com/support/rbsu).

## Red Hat Enterprise Linux 5.x

This section describes several recommended best practices for running Red Hat Enterprise Linux 5.9 (or later) on the DL580 Gen8 server.

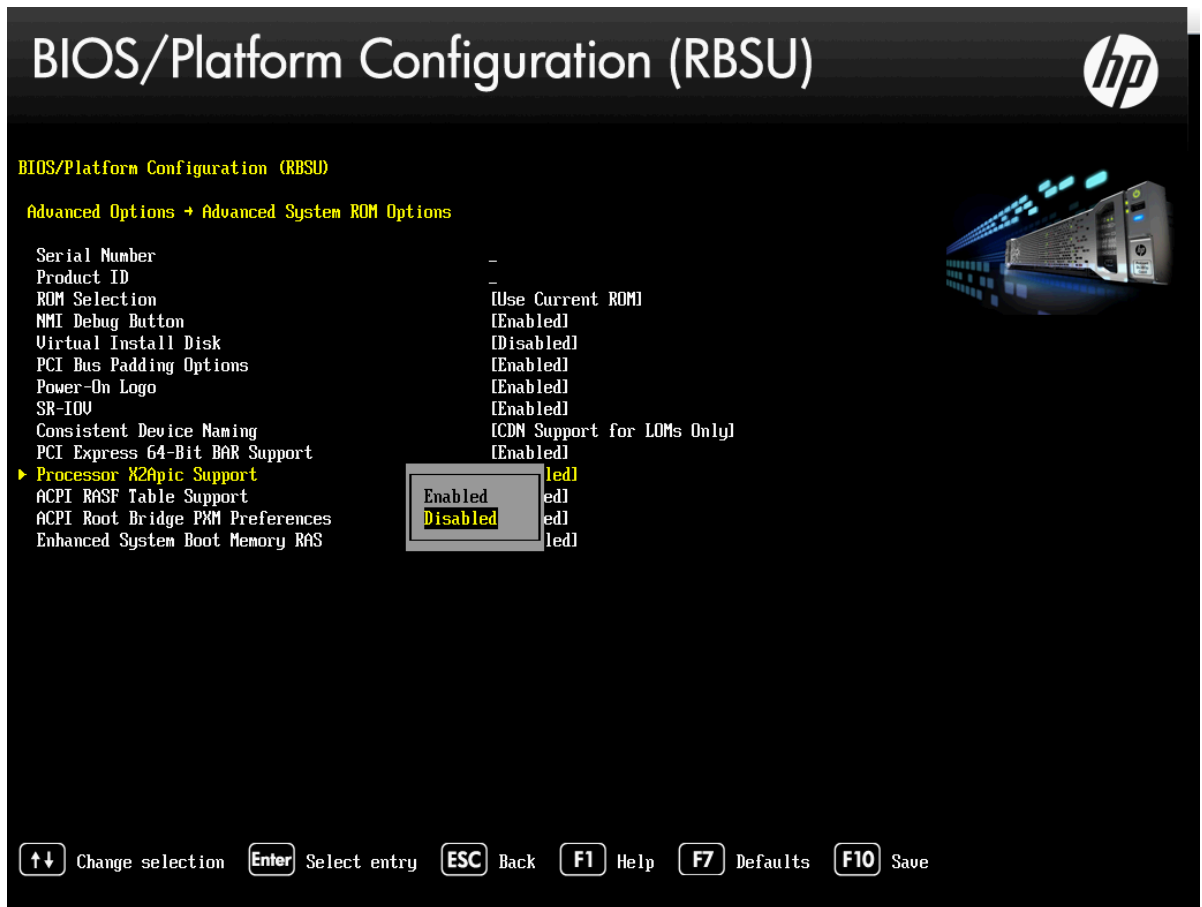
### Disable x2APIC with Intel Xeon E-7 series processors

As noted in the “OS support matrix for the HP ProLiant DL580 Gen8 server” section on page 11, Red Hat Enterprise Linux 5.X does not support the Intel x2 Advanced Programmable Interrupt Controller (x2APIC) architecture. Systems with Intel Xeon E-7 series processors (Ivy-Bridge) require disablement of x2APIC support prior to the installation. By default, x2APIC is disabled on the DL580 Gen8.

x2APIC support is controlled by the system BIOS. To access the x2APIC settings and verify disablement, follow these steps:

1. During system boot, press **F9** to run the System Utilities.
2. Select System Configuration.
3. Select BIOS/Platform Configuration (RBSU).
4. Select **Advanced Options**.
5. Select **Advanced System ROM Options**.
6. Select **Processor X2APIC Support** (see Figure 8).
7. If necessary, select **Disabled** to turn x2APIC OFF.
8. Exit all RBSU menu screens using the **Esc** key.
9. Press the ‘**Y**’ key to confirm your changes.

Figure 8. Processor x2APIC support

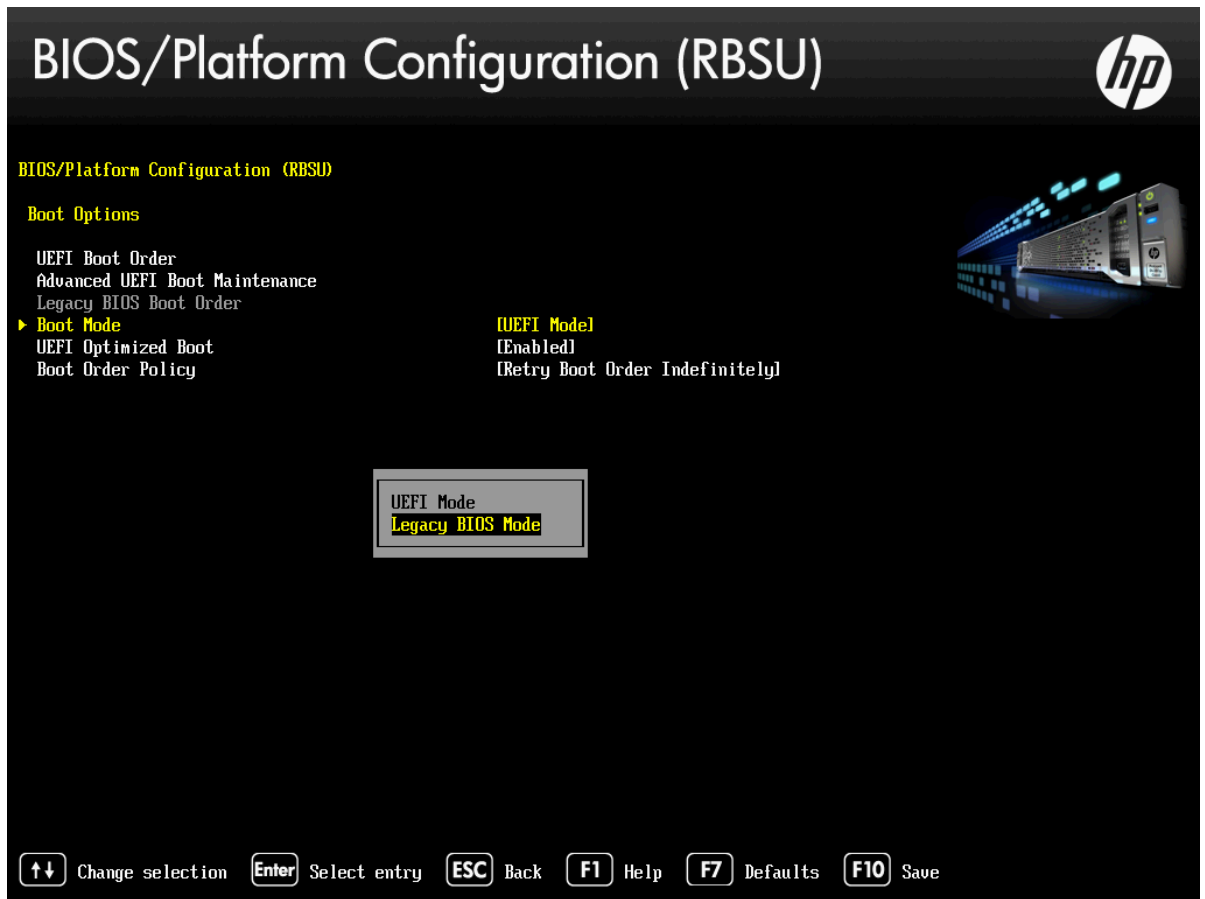


### Verify Legacy BIOS boot mode before OS installation

Red Hat Enterprise Linux 5.X does not support UEFI boot mode. During the boot process, confirm that the system is in legacy BIOS boot mode in the RBSU. If the Boot Mode is set to UEFI, follow these steps to set it to Legacy mode:

1. During system boot, press **F9** to run the System Utilities.
2. Select **System Configuration**.
3. Select **BIOS/Platform Configuration (RBSU)**.
4. Select **Boot Options**.
5. Select **Boot Mode**.
6. Select **Legacy BIOS Mode** (see Figure 9).
7. Exit all RBSU menu screens using the **Esc** key.
8. Press the **Y** key to confirm your changes and reboot the system.

Figure 9. Boot mode options



## Use NOOP in the I/O scheduler

The default I/O scheduler is Completely Fair Queuing (CFQ), which attempts to balance throughput and latency. However, today's intelligent storage subsystems usually perform best by using the NOOP I/O scheduler, which disables all I/O scheduling operations in the operating system. This I/O scheduler may be configured system-wide using the kernel boot option `elevator=noop` or on a per-device basis using the following command:

```
echo noop > /sys/block/<device>/queue/scheduler
```

Increase crash kernel size

## Increase crash kernel size

The HP ProLiant DL 580 Gen8 running Red Hat Enterprise Linux 5.9 supports up to 1 TB of memory, so you should use the kernel boot option `crashkernel` to adjust the crash kernel size accordingly. RHEL 5.9 does not include some recent enhancements to crash utilities, and on large memory configurations, it is recommended to make the crash kernel as large as possible, typically around 850 MB. Keep in mind that the Linux kernel has a maximum limit of 896 MB for crash kernel size, but the configurable crash kernel size will be less due to kernel overhead.

## Red Hat Enterprise Linux 6.4 and SUSE Linux Enterprise Server 11 SP3 best practices

This section describes several recommended best practices for running Red Hat Enterprise Linux 6.4 (and later) or SUSE Linux Enterprise server 11 SP3 on the DL580 Gen8 server.

### Single Root I/O Virtualization Support

If the Single Root I/O Virtualization (SRIOV) feature is enabled in BIOS (using RBSU) and interface cards are being used that support virtual functions, Red Hat Enterprise Linux 6.4 requires the Linux kernel boot parameter “pci=norom” to ensure the BARs associated with virtual functions are mapped properly. If the BARs are not mapped properly, the virtual functions will not be enumerated.

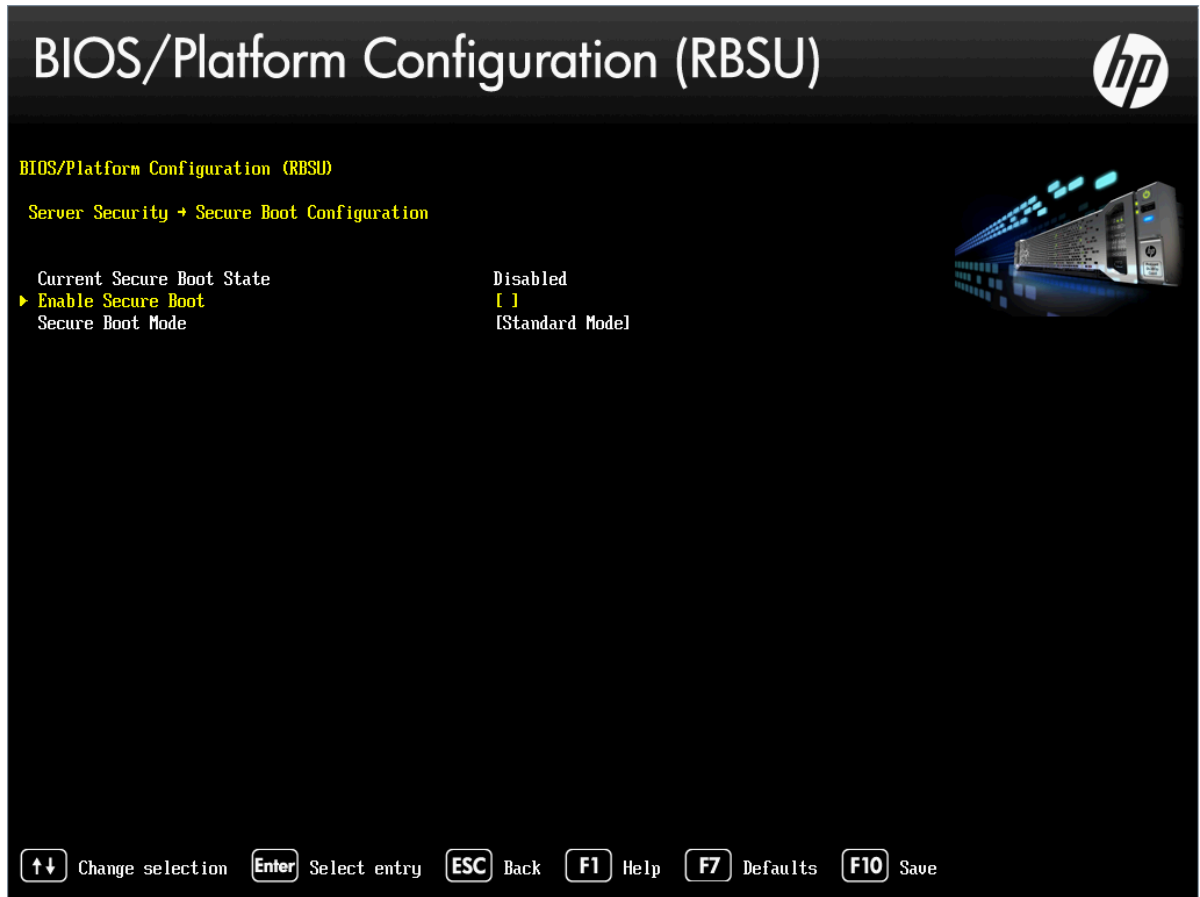
### Secure Boot

Secure Boot is disabled by default on the DL580 Gen8 server and requires UEFI Boot Mode to be enabled. Secure boot is supported on SUSE Linux Enterprise server 11 SP3. To change the state of secure boot on the DL580 Gen8:

1. During system boot, press **F9** to run the System Utilities.
2. Select **System Configuration**.
3. Select **BIOS/Platform Configuration (RBSU)**.
4. Select **Server Security**.
5. Select **Secure Boot Configuration**.
6. Select **Enable Secure Boot** to toggle the state on or off (see Figure 10).
7. Exit all RBSU menu screens using the **Esc** key.

For more information, read to the “HP UEFI System Utilities User Guide” at: [hp.com/support/UEFI\\_UG\\_en](http://hp.com/support/UEFI_UG_en) and the “SUSE Linux Enterprise server 11 SP3 Administrator Guide” at: [suse.com/documentation/sles11/singlehtml/book\\_sle\\_admin/book\\_sle\\_admin.html](http://suse.com/documentation/sles11/singlehtml/book_sle_admin/book_sle_admin.html).

Figure 10. Secure Boot Configuration



### Test Red Hat Enterprise Linux 6.4 (and later) and SUSE Linux Enterprise server 11 SP3 workloads with Hyper-Threading (HT)

The benefits of HT have already been described on page 16. Again, the benefits depend highly on the application workload and server configuration. While some applications see an overall performance boost, some applications see no or negative performance improvement. HP highly recommends testing HT with your particular workload.

As noted in the OS Support Matrix, Red Hat Enterprise Linux 6.4 and SUSE Linux Enterprise server 11 SP3 support up to 4096 logical processors. The supplied kernels include improved scheduling functionality, but managing a high number of logical processors brings higher overhead. Unless your workload is proven to benefit from Hyper-Threading, HT should be disabled in order to reduce this system overhead. You can always enable HT later if additional CPU capacity is required.

Hyper-Threading is controlled by the system BIOS and is enabled by default on the DL580 Gen8. To enable or disable Hyper-Threading, follow the steps described in "Test Red Hat Enterprise Linux 5.9 (or later) workloads with Hyper-Threading (HT)" on page 16.

### Increase crash kernel size

The HP ProLiant DL 580 Gen8 supports up to 3 TB of memory so you should use the kernel boot option `crashkernel` to adjust the crash kernel size accordingly. It is recommended to set the crash kernel size to at least 512 MB on large memory configurations. Keep in mind that the Linux kernel has a maximum limit of 896 MB for crash kernel size, but the configurable crash kernel size will be less due to kernel overhead.

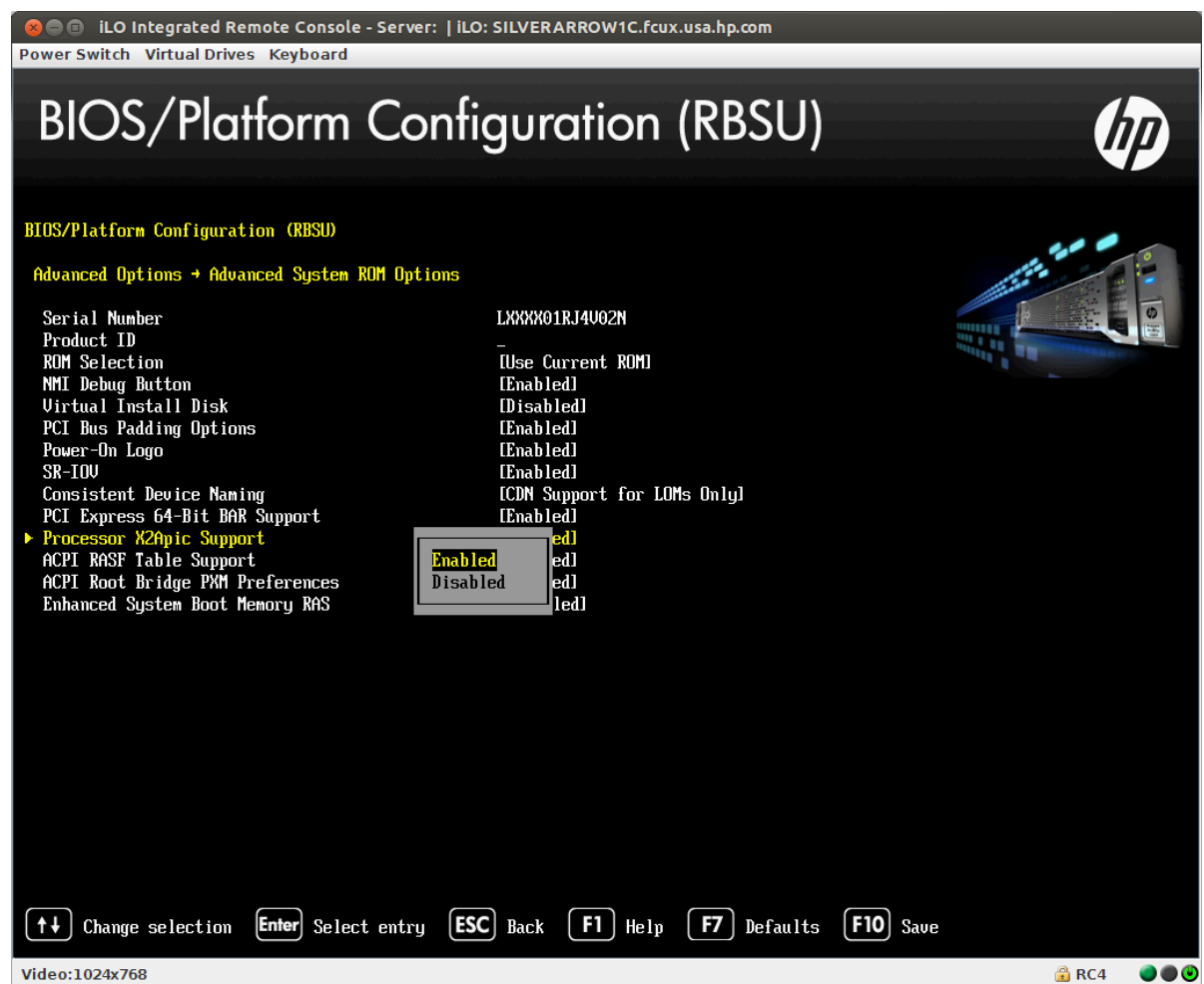
## x2APIC support

RHEL6.4 (and later) and SLES11SP3 support x2apic, and it can be enabled in BIOS as shown in Figure 11.

You can enable x2APIC with Intel Xeon E-7 series processors. x2APIC support is controlled by the system BIOS. To access the x2APIC settings and verify enablement, follow these steps:

1. During system boot, press **F9** to run the System Utilities,
2. Select **System Configuration**.
3. Select **BIOS/Platform Configuration (RBSU)**.
4. Select **Advanced Options**.
5. Select **Advanced System ROM Options**.
6. Select **Processor X2APIC Support** (see Figure 11).
7. If necessary, select **Enabled** to turn x2APIC ON.
8. Exit all RBSU menu screens using the **Esc** key.
9. Press the **Y** key to confirm your changes.

Figure 11. Processor x2APIC support



## Performance and tuning

Although they are not enabled by default, performance and tuning tools such as “Tuned” and “NUMA Affinity Management Daemon” (NUMAD) are important considerations. You can read more about these tools in the “Red Hat Enterprise Linux Performance Tuning Guide” found at: [redhat.com/site/documentation/en-US/Red\\_Hat\\_Enterprise\\_Linux/6/html-single/Performance\\_Tuning\\_Guide/index.html](http://redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux/6/html-single/Performance_Tuning_Guide/index.html).

## Additional software components

This section includes additional recommended software components when running Linux on the DL580 Gen8 server.

### HP SPP

HP SPP is a comprehensive systems software and firmware solution delivered as a single ISO. This solution uses HP Smart Update Manager (HP SUM) as the deployment tool and is tested on all supported HP ProLiant servers including the new HP ProLiant Gen8 servers.

An HP SPP release is delivered as a family of deliverables with a full ISO and six subset ISOs differentiated by operating system and HP ProLiant server type. Each subset version of the full ISO includes the software and firmware smart components required for its specific audience and allows for a customized solution and faster download. The ISOs available are as follows:

- HP SPP
- ProLiant ML/DL/SL Red Hat Enterprise Linux Pack
- ProLiant ML/DL/SL SUSE Linux Enterprise server Pack

### HP SPP benefits

HP SPP simplifies software maintenance in several ways:

- Leverages the power of HP SUM for consolidated management and distribution to multiple servers
- Provides broad portfolio support for HP ProLiant servers, BladeSystem enclosures, and the HP CloudSystem Matrix
- Enables simplified updates with a single-step installation process that packages firmware, systems software, and drivers. The installation provides:
  - A consolidated, single solution for all supported HP ProLiant servers, with interdependency testing of drivers and firmware
  - Convenient access via web delivery
  - PXE bootable ISO images that reduce:
    - Qualification cycles
    - Resource usage
    - Maintenance duration
    - Downtime
- Serves as an integral part of the HP Server Experience that:
  - Increases ease of server management — with enhancements to HP SUM and consolidated sets of tested firmware and software
  - Enhances IT productivity while reducing downtime
  - Provides reliable configuration by:
    - Assisting with server deployment, integrating server support software and firmware, and minimizing configuration and update issues
    - Ensuring consistency and reliability and maximizing system uptime
  - Streamlines delivery of HP ProLiant software and firmware

The SPP is available at the HP SPP Download page at: [hp.com/go/spp/download](http://hp.com/go/spp/download)

### HP SPP minimum requirements for installation on Linux servers

Several package dependencies exist for proper HP SUM execution on Linux. As the list of required packages can change over time depending on new HP SUM features/functionality, you should refer to version-specific HP SPP Release Notes at the HP SPP documentation site ([hp.com/go/spp/documentation](http://hp.com/go/spp/documentation)) for platform-specific requirements for a given HP SPP version.

---

#### Note

Package dependencies for running HP SUM differ depending on whether it is run in “console mode” or “graphical mode.” Required dependencies may also differ depending on the host Linux distribution and version. In addition, pay attention to any specifically required 32-bit versions of packages. These various requirements are documented in the HP SPP Release Notes.

---

**Installing HP SPP**

To update firmware, update drivers, and add any HP value-add software on the DL580 Gen8 server, use one of the two ProLiant ML/DL/SL Server Pack subset ISOs for Linux listed in bold on page 12. From the SPP download page, download the zip file containing the ISO for the target Linux distribution, and mount the ISO as a loop back device for access. The HP SUM utility used for installation is located in `/hp/swpackages/`. If a graphics terminal or a remote desktop program (like VNC) is available, HP SUM can be launched in graphic mode by running `./hpsum` from this directory. Otherwise, the utility can be launched in console mode by running `./hpsum/s`.

**Note**

Although full-featured and robust, HP SUM run in “console mode” is not interactive and thus requires attention to proper command line options and target bundle selections. Also, note that the HP SUM executable must be run in the `/hp/swpackages` directory to avoid installation problems.

Each HP SPP release contains driver and firmware components that, while qualified for compatibility on a given Linux/DL580 Gen8 combination, may or may not be appropriate in all situations. Consideration should be given to determine which included packages are necessary for your server. HP works with the Linux community to incorporate HP drivers directly into the upstream kernel. This ensures native functionality for the vast majority of HP devices using in-box drivers from Linux distribution vendors. For Red Hat and SUSE Linux distributions, HP works with the vendors to provide updates for major and minor releases.

While driver and firmware updates should be reviewed for applicability, HP encourages the installation of included value-add components to optimize user experience with the HP ProLiant server. These include:

- HP ProLiant component health agents
- SNMP agents
- The Systems Management Homepage (SMH), which acts as a GUI for monitoring and managing your DL580 Gen8 server (see the “HP System Management Homepage” section)

**HP Software Delivery Repository (SDR)**

The SDR is a network repository that you can access for driver and software updates. It contains the same Linux components as the SPP. Upon release of a new SPP, we update the SDR to include the drivers, utilities, management agents, and firmware (currently as \*.scexe files, but will be transitioning to RPMs) from the SPP release. This ensures that SPP and SDR contents are identical.

Read the “Linux best practices using HP Service Pack for ProLiant (SPP) and Software Delivery Repository (SDR)” at: [hp.com/V2/GetDocument.aspx?docname=4AA5-0831ENW&cc=us&lc=en](http://hp.com/V2/GetDocument.aspx?docname=4AA5-0831ENW&cc=us&lc=en).

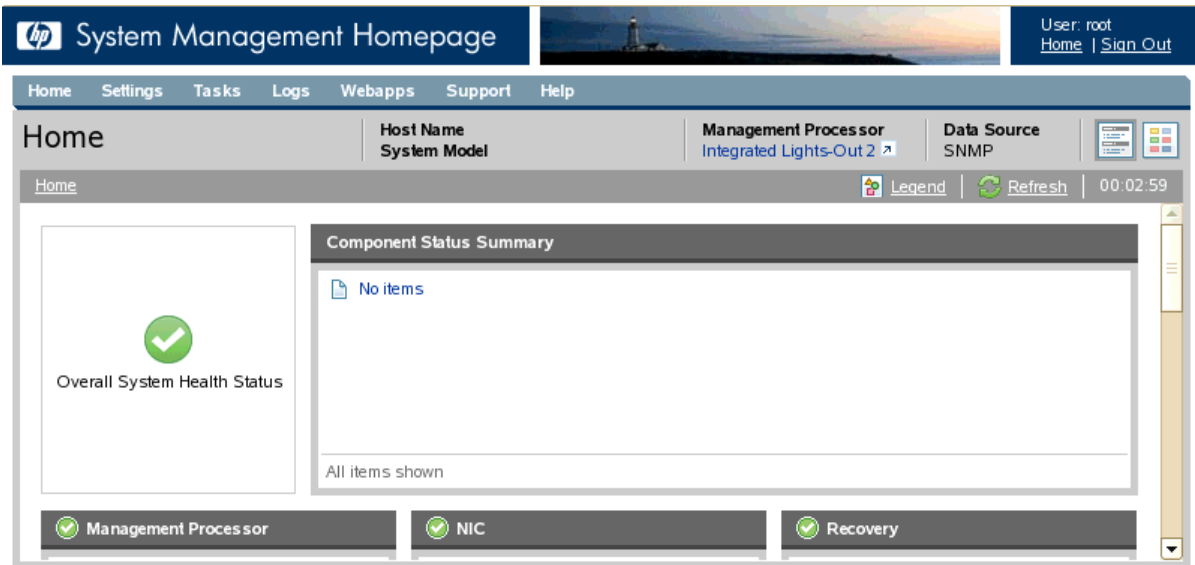
**HP System Management Homepage**

HP System Management Homepage (SMH) is a web-based interface that consolidates and simplifies single system management for HP servers running Linux operating systems. Use it to view in-depth hardware configuration and status data, performance metrics, system thresholds, diagnostics, and software version control information using a single intuitive interface (Figure 12). HP SMH aggregates and displays data from agents and HP Web-enabled systems management software, including the following:

- HP Insight Diagnostics
- Array Configuration Utility
- HP Software Version Control Agents



Figure 12.



### HP SMH features

HP SMH provides streamlined operations for HP servers running Linux operating systems with the following security enhancements and features:

- Browser access using secure OS-based, Secure Sockets Layer (SSL) authentication
- Common HTTP and HTTPS service for HP Insight Management Agents and utilities, for reduced complexity and system resource requirements.
- Simplified architecture for implementing HTTP security and HP management updates
- Access control through Network Interface Card (NIC) binding and advanced configuration features for individuals and user groups
- Broad OS and browser support
- A command line interface (`smhconfig`) that provides those with administrator rights the ability to set values using the command line

---

### Note

For more information, see the "HP System Management Homepage Installation Guide: [2.hp.com/portal/site/hpsc/](http://2.hp.com/portal/site/hpsc/)

---

## Resources

HP ProLiant DL580 Gen8 Server overview

[hp.com/servers/dl580-gen8](http://hp.com/servers/dl580-gen8)

Recommended software and driver downloads

[hp.com/support/dl580q8](http://hp.com/support/dl580q8)

Complete library of documentation supporting the DL580 Gen8 server

[hp.com/servers/dl580](http://hp.com/servers/dl580)

Linux best practices using HP SPP and Software Delivery Repository (SDR)

[hp.com/V2/GetDocument.aspx?docname=4AA5-0831ENW&cc=us&lc=en](http://hp.com/V2/GetDocument.aspx?docname=4AA5-0831ENW&cc=us&lc=en)

Red Hat Enterprise Linux Minimum Operating System Requirements

[hp.com/us/en/enterprise/servers/supportmatrix/exceptions/rhel\\_exceptions.aspx#.UvqTJrRbv4c](http://hp.com/us/en/enterprise/servers/supportmatrix/exceptions/rhel_exceptions.aspx#.UvqTJrRbv4c)

SUSE SLES Certification and Support

[hp.com/us/en/enterprise/servers/supportmatrix/suse\\_linux.aspx#.UvqtnbRbv4c](http://hp.com/us/en/enterprise/servers/supportmatrix/suse_linux.aspx#.UvqtnbRbv4c)

HP ProLiant Server power management for Red Hat Enterprise Linux

[hp.com/portal/site/hpsc/public/kb/docDisplay?docId=emr\\_na-c03231802](http://hp.com/portal/site/hpsc/public/kb/docDisplay?docId=emr_na-c03231802)

Red Hat Linux NUMA support for HP ProLiant Servers

[hp.com/portal/site/hpsc/public/kb/docDisplay?docId=emr\\_na-c03261871](http://hp.com/portal/site/hpsc/public/kb/docDisplay?docId=emr_na-c03261871)

Intel Xeon E-7 processor series

[intel.com/itcenter/products/xeon/E7/index.htm](http://intel.com/itcenter/products/xeon/E7/index.htm)

**Sign up for updates**

[hp.com/go/getupdated](http://hp.com/go/getupdated)



Share with colleagues



Rate this document

---

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel and Xeon are trademarks of Intel Corporation in the U.S. and other countries. Red Hat is a registered trademark of Red Hat, Inc. in the United States and other countries.

4AA5-1115ENW, February 2014

