Lenovo

Energy Efficiency Features of Lenovo System x Servers

Describes how System x servers are designed to be energy efficient

Introduces key efficiency features

Explains ways you can configure your server to be most efficient

Introduces energy efficiency beyond the server

Robert R. Wolford

LENOVO PRESS

Abstract

Several years ago, little was given to how much power servers used. With plentiful energy resources and low electrical rates, server characteristics, such as low cost, high performance, and high reliability were considered more important than improving the power efficiency of servers. That all changed when the price of electricity began to rise dramatically during 2004 and 2005. From 2004 to 2008, the price jumped by another 30%. Even after 2008, the dramatic jump leveled off but did not drop back to the previous levels. It seems that higher electrical rates are here to stay. This dramatic increase in a short period of time sparked the need of data center managers to develop more efficient computing solutions.

Designing high power efficiency into a server requires a balanced approach. If peak power efficiency were the only goal, a server can be designed with considerable efficiency. However, the server would be large, have slow absolute performance and high latency, little expansion capability, and be expensive. It also would have few reliability, availability, and serviceability (RAS) features. Designing a server that is wanted requires the consideration of efficiency, performance, capability, and cost.

This paper is for clients who want understand energy efficiency and the ways that Lenovo servers use components and features to maximize energy efficiency.

At Lenovo Press, we bring together experts to produce technical publications around topics of importance to you, providing information and best practices for using Lenovo products and solutions to solve IT challenges.

For more information about our most recent publications, see this website:

http://lenovopress.com

Contents

Introduction

Several years ago, little thought was given to how much power servers used. With plentiful energy resources and low electrical rates, server characteristics, such as low cost, high performance, and high reliability were considered more important than improving the power efficiency of servers.

That all changed when the price of electricity rose dramatically during 2004 and 2005, as shown in Figure 1. From 2004 to 2008, the price jumped by another 30%. Even after 2008, the dramatic jump leveled off but did not drop back to the previous levels. It seems that higher electrical rates are here to stay.

This dramatic increase in a short period sparked the need of data center managers to develop more efficient computing solutions.

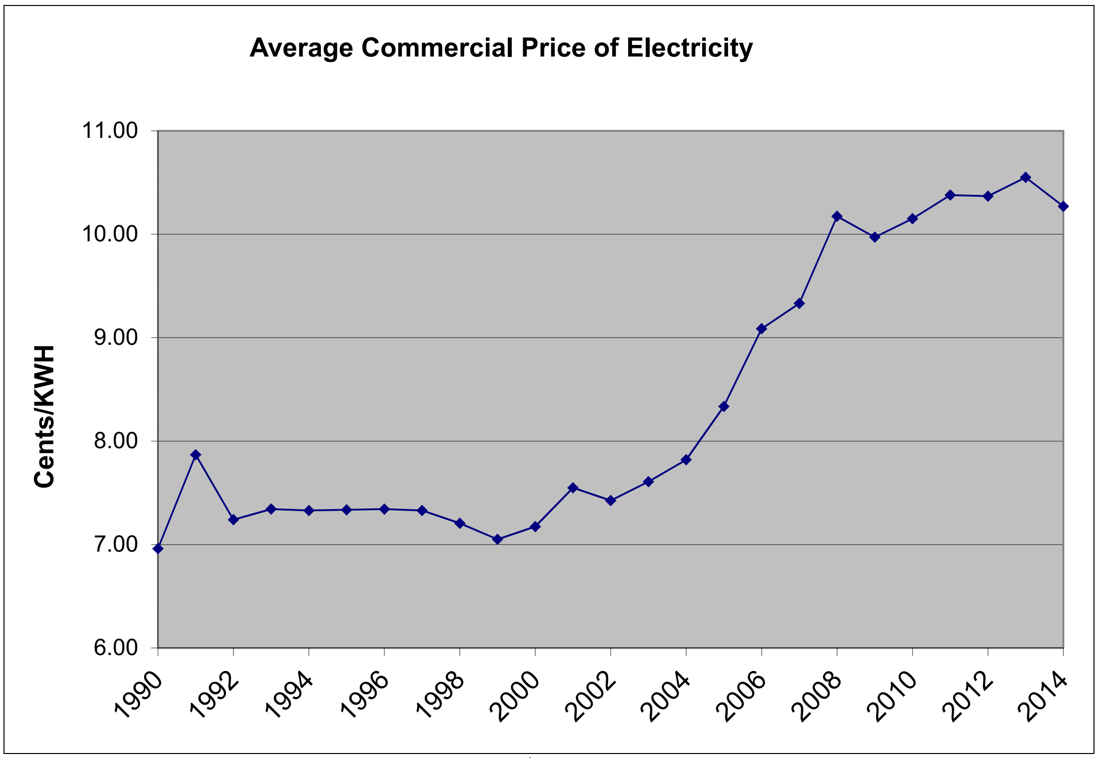


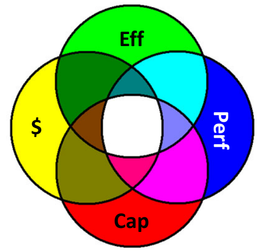
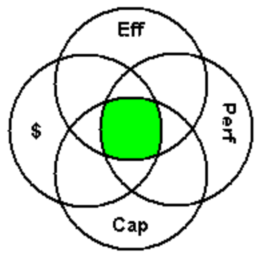
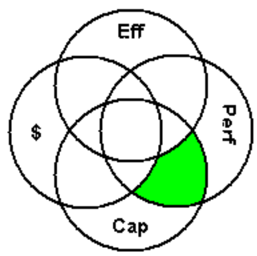
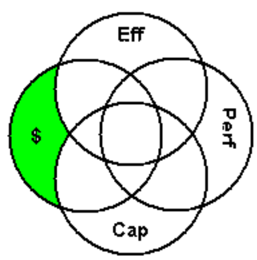
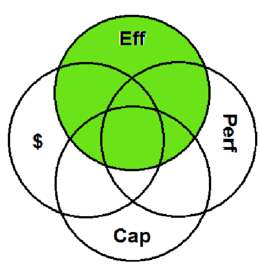
Figure 1 Average commercial price of electricity[1]

A balanced approach

Designing high-power efficiency into a server requires a balanced approach. If peak power efficiency was the only goal, a server can be designed with considerable efficiency. However, the server would be large, have slow absolute performance and high latency, little expansion capability, be expensive, and have few reliability, availability, and serviceability (RAS) features. Designing a server that is wanted requires efficiency, performance, capability, and cost considerations, as listed in Table 1 on page 4.

[1] US Energy Information Administration (EIA): Current issues and trends: http://www.eia.gov/electricity/data/state/avgprice_annual.xls

Table 1 A balanced approach

Legend: ► $ Cost ► Eff Efficiency ► Perf Performance ► Cap Capabilities	
Some server offerings strike a balance among all four areas	
Others focus on a combination of two areas	
Still others focus on cost optimization	
The focus of this paper is to describe the energy efficiency features built into System x® servers.	

Relative influence of power features

Increases in power savings and efficiency at the system level are a combined effect of many individual features, as shown in Figure 2. The benefit of power-saving features varies depending on the usage of the server.

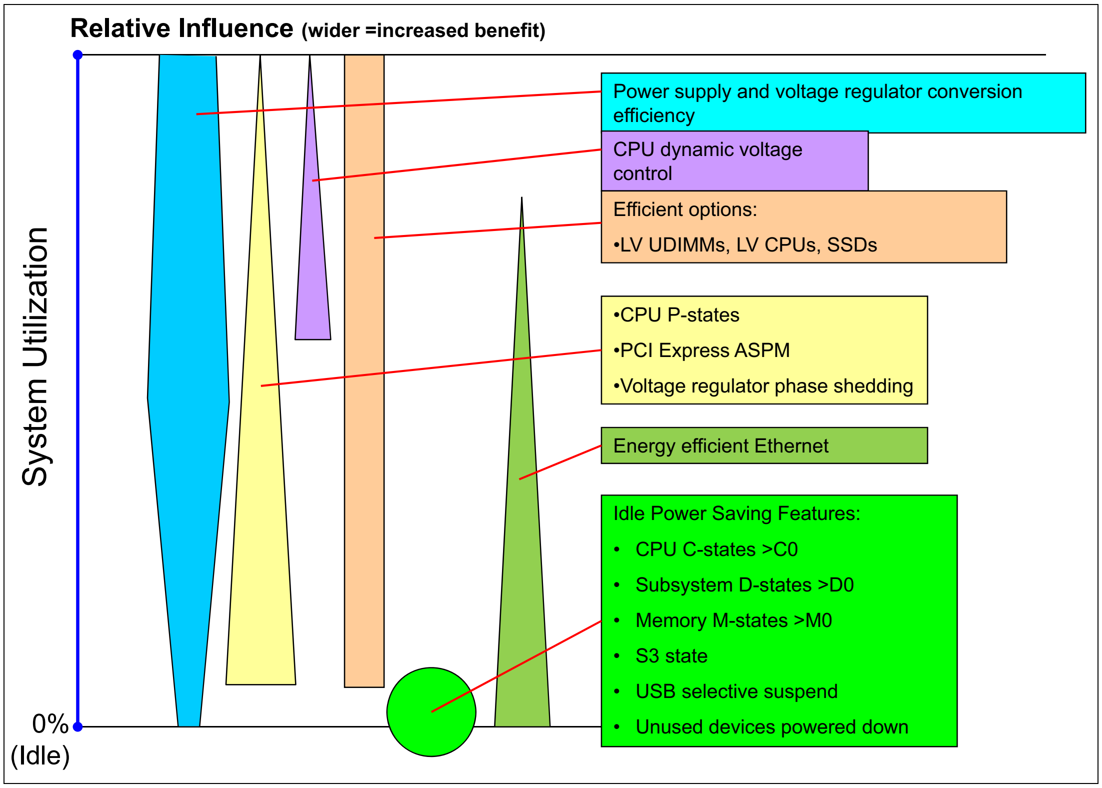


Figure 2 Effects of power savings and efficiency features

Figure 2 shows the relative influence of many power-saving features. The vertical axis represents system utilization ranging from 0% (idle) to 100% (maximum utilization). The width of each polygon at any utilization level represents the relative benefit of each group of power-saving features.

For example, at 50% utilization, the power supply and voltage regulator device (VRD) efficiency has a large influence on overall system efficiency because the blue (leftmost) polygon is wide at the 50% utilization point. In contrast is the green (rightmost) energy efficient Ethernet, which has little benefit at 50% utilization. As another example, the idle (green circle) power-saving features have no benefit at 50% utilization.

It is important to understand the portion of the utilization curve where the server operates. In this manner, it is possible to understand which power-saving features are influencing the overall performance/watt efficiency of the server for the target workload. For example, if a server is running VMWare and spends 95% of the time at 60 - 80% utilization, the features that save power at less than 60% or greater than 80% are not as important, as shown in Figure 3 on page 6.

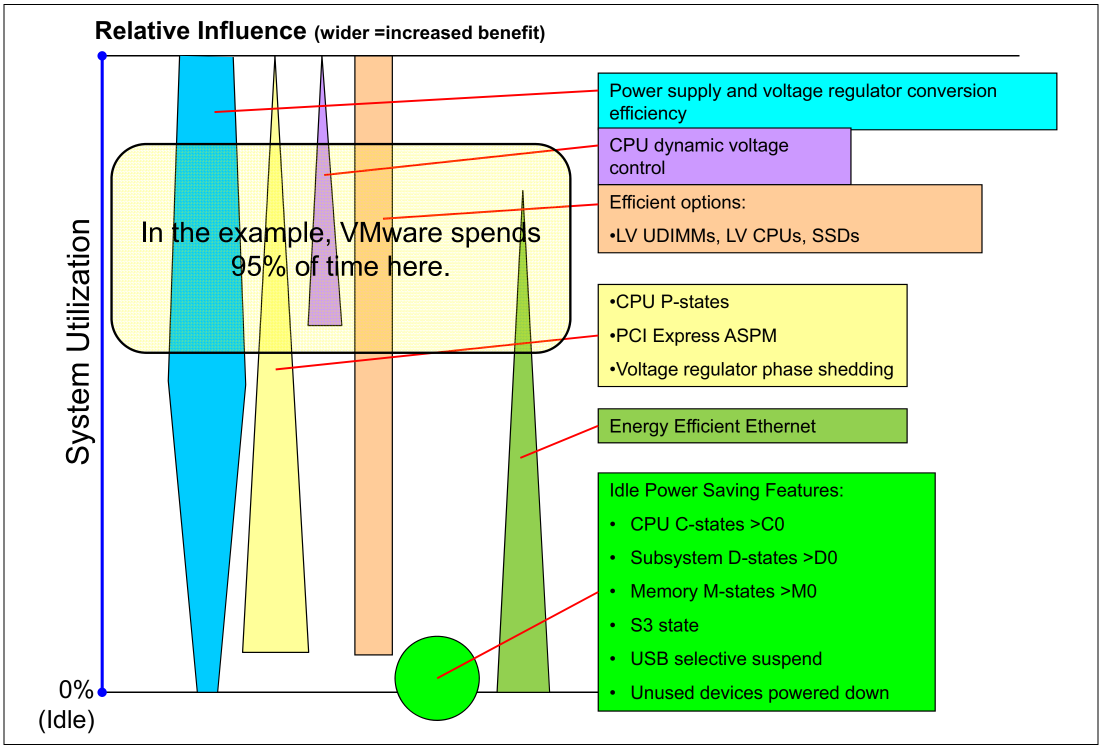


Figure 3 VMware workload example

In the following sections, we describe the individual power features that are shown in Figure 2 on page 5 and Figure 3.

Server features for power efficiency

System x servers include the following features that are designed to save power, improve efficiency, monitor power, and impose power limits where necessary:

- “80 PLUS Titanium server power supply” on page 7
- “Active or standby power supplies” on page 8
- “Voltage regulator efficiency” on page 9
- “Efficient options” on page 9
- “PCI Express Active State Power Management” on page 10
- “Energy-Efficient Ethernet” on page 11
- “S3 State” on page 11
- “USB selective suspend” on page 12
- “240 VA Protection optimized for efficiency” on page 13
- “Unused devices” on page 13
- “Energy efficient thermal design” on page 14
- “Energy efficient turbo mode” on page 14
- “CPU uncore frequency scaling” on page 15
- “Power bias and performance bias” on page 15
- “Lenovo efficiency mode” on page 15
- “Power metering” on page 16
- “Power capping” on page 19
- “Energy Star certification” on page 20
- “Higher temperature limits” on page 21
- “PDUs and power distribution” on page 21

80 PLUS Titanium server power supply

Higher efficiency in the bulk power supply of the server means less heat output from the server. Any AC power that is not directly converted to bulk 12 V power is dissipated as heat or is used as part of the AC-to-DC conversion process. Such efficiency can have a dramatic effect on overall power usage of the server and the data center.

80 PLUS is an industry organization that rates power supply efficiencies and grades them based on the percentage conversion at specific power loads. Ratings include 80 PLUS Platinum and 80 PLUS Titanium. For more information (including the requirements for each level), see this website:

http://www.plugloadsolutions.com/80PlusPowerSupplies.aspx

Consider the following example:

A server contains a 750 W power supply unit (PSU) and the server uses 375 W DC. The PSU operates at 50% load (375 W or 750 W). As listed in Table 2, if the PSU is 80 PLUS certified, the conversion efficiency is 80%. That percentage means that the power draw at the AC power cord is 469 W. In total, 375 W is delivered to the server components and 94 W is dissipated as heat.

That 94 W of heat must be removed from the server. This removal typically is done with a combination of cooling fans in the server and the cooling infrastructure in the data center by the use of computer room air conditioners (CRACs), heat exchangers, air handlers, and so on. The data center overhead (that is, the power that is delivered to the data center that is not used by the IT equipment) can easily reach double the power that is dissipated as heat from the power supply. Therefore, for only one server with an 80 PLUS rated PSU, the total power overhead because of PSU efficiency can reach 188 W.

One of the best ways to reduce the power overhead is to improve the efficiency of the server PSU, as this improvement has a cascading effect up through the data center for power and cooling. In this example, if an 80 PLUS Titanium PSU is used instead of the 80 PLUS power supply, the 50% efficiency jumps to 96% (see Table 2). Working through the same calculations, the input power to the server is lowered to 391 W of which 16 W is dissipated as heat. At the data center level, the 188 W is reduced to 32 W, which is a reduction of 156 W.

Table 2 80 Plus minimum efficiencies[2]

Power supply load (Percent of rated load)	**80 PLUS Titanium Minimum efficiency**	**80 PLUS Minimum efficiency (Industry standard)**
10% load	90%	Not rated
20% load	94%	80%
50% load	96%	80%
100% load	91%	80%

[2] Source: http://www.plugloadsolutions.com/80PlusPowerSupplies.aspx

Active or standby power supplies

System x servers support up to six PSUs that are installed in a server or chassis. If a PSU supply fails, the remaining PSUs provide the power and prevent the server from powering down unexpectedly. When all installed PSUs are operating with no errors, the total power load is evenly distributed among each installed PSU. For example, with two PSUs installed, each PSU provides half of the power that the server uses.

Installing two PSUs is beneficial for fault tolerance. However, there is one disadvantage. When both PSUs are delivering power, the efficiency of each PSU can be reduced because the typical PSU efficiency versus load curve is bell-shaped.

Consider the following example:

If a server is using half of the rated power of the PSU for which the efficiency curve is shown in the blue solid line in Figure 4, the PSU operates at the peak of its efficiency curve (point 1 on the efficiency curve). If a second PSU is installed for redundancy, the total server power load is divided equally between the two PSUs. In that case, the load on each PSU is reduced to 25% (point 2 on the efficiency curve). At that point, the efficiency is lower than at the 50% load point. This concept is represented by the orange boxes (boxes 1 and 2) and lines in the efficiency curve plot in Figure 4.

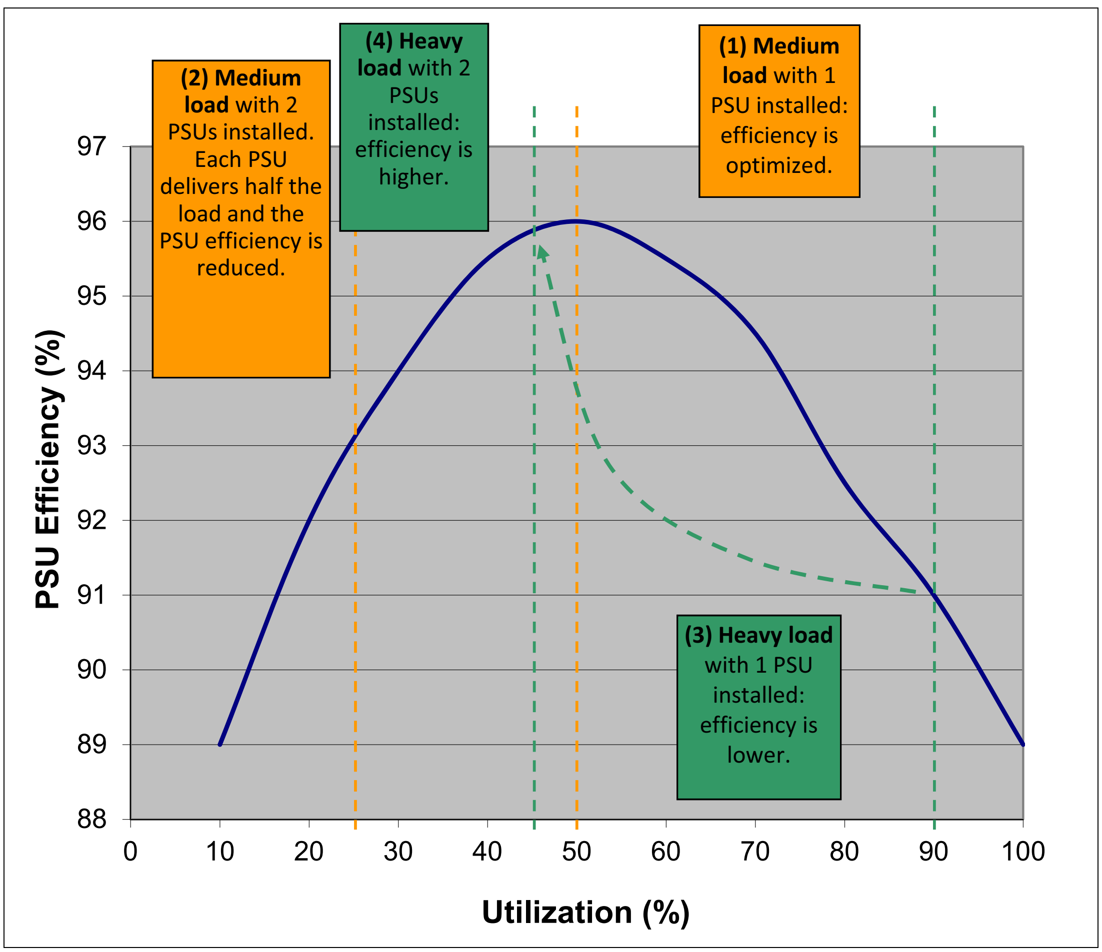


Figure 4 Operating points on the power supply efficiency curve

To combat this problem, System x servers are designed to indicate to the PSUs when the server power load is low enough such that one of the installed PSUs can be placed into a low power, standby state. In this manner, the remaining power supply delivers the entire load and efficiency is boosted. Under heavy loads, it is more beneficial to leave both PSUs in an active state.

With heavy loads, dividing the load equally between both PSUs results in an efficiency increase because the load point on the PSU efficiency curve shifts from the extreme right to the middle portion of the curve, as indicated by the green boxes (points 3 and 4) and lines in the efficiency curve plot in Figure 4 on page 8. Overall, by intelligently managing the power that each PSU provides, the PSU can operate in a more efficient region of its efficiency curve.

Voltage regulator efficiency

There are several VRDs in the system that convert 12 V to the target voltage that is needed for each subsystem. As with the system power supplies, the goal is to make each VRD highly efficient so that minimal power is lost and dissipated as heat. The VRDs incorporate the following features to minimize power loss and maximize efficiency:

- Portions of the VRD are dynamically turned on and off based on the load.
- The frequency in which each VRD operates is optimized to reduce switching losses.
- The use of inefficient linear regulators is minimized.
- The motherboard layout is optimized to reduce power losses between the VRD and the component it is driving.
- VRD components are selected that exhibit minimal inherent power losses.

Efficient options

Some options that can be installed in servers are inherently more efficient than others. They can contain more efficient components, have solid-state media versus spinning media, or contain fewer overall parts that use power. System x servers offer various options with which a user to choose between minimum power usage, maximum performance, and a balance between the two options. The following energy efficient options are available on System x M5 servers:

- LV CPUs: Low voltage (LV) CPUs generally operate at lower power for a specific performance level.
- DDR4 DIMMs: DDR4 DIMMs can save up to 25% power that is compared to predecessor DDR3 DIMMs.
- SSDs: Having no spinning parts helps solid-state drives (SSDs) to achieve lower power and quicker engagement of power management that is compared to traditional spinning hard disk drives (HDDs).
- Lower power I/O adapters: Not every customer wants 10 Gb Ethernet or 16 Gb Fibre Channel connections. Many data centers cannot accommodate the higher power that is associated with high-bandwidth adapters. Lenovo® offers various I/O options to configure the server for a good balance between performance and power usage.

In addition to the energy efficient options that are described here, System x is designed so that higher power subsystems are not soldered down on the motherboard. For example, every user might not need a RAID 10 storage controller. The base storage controller on the motherboard is embedded in the core chipset. In this manner, no extra power is wasted on a higher feature storage controller if that function is not needed.

PCI Express Active State Power Management

Most of the subsystems in a server are connected with PCI Express buses. The PCI Express bus is capable of high bandwidths, but that high bandwidth is not needed all of the time. Running the PCI Express bus at full capacity when it is not needed wastes power. To avoid this, PCI Express Active State Power Management (ASPM) is used by the operating system (OS) to place the physical layer of PCI Express buses into a low-power state when they are idle or lightly used. The level of ASPM is controlled through OS power options. An example of ASPM control under Windows Server 2012 is shown in Figure 5. This window is accessible in the Power Options selection in the Control Panel.

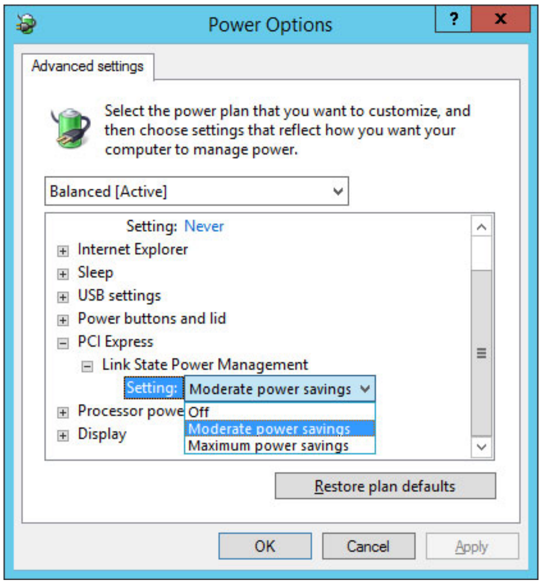


Figure 5   Windows ASPM control panel

Linux distributions offer similar ASPM driver kernel parameters, as listed in Table 3.

Table 3   Linux ASPM parameters

Setting	Operation
pcie_aspm=off	ASPM is disabled.
pcie_aspm=default	Use the default firmware configuration as set in the PCI Express Capabilities list item with ID 0x10. The `pcie_aspm` variable is changed in the Linux kernel configuration file. For example, in RHEL 6.3, the parameter goes in the grub config file at `/boot/grub/grub.conf`. Each adapter or subsystem that supports ASPM starts to a default state. When pcie_aspm=default, the OS does not override the adapter's default setting.
pcie_aspm=performance	Disables ASPM and clock power management.
pcie_aspm=powersave	Enable ASPM and clock power management. Higher power-saving modes in ASPM might incur a longer exit latency, depending on the server workload.

Energy-Efficient Ethernet

Energy-Efficient Ethernet (EEE)[3] is similar to ASPM (see Table 3 on page 10). The difference is that EEE operates on the physical layer of Ethernet transmitters. When EEE is enabled and no data is being sent over the Ethernet link, the link is placed into a low-power sleep state. A low power idle (LPI) indication signal is sent periodically to the transmit chip, which instructs it to turn off for a specified period. If data is ready to be transmitted, a normal idle signal is sent to the transmitter to wake it up before the data is sent. When EEE is used, the Ethernet receive link always remains active, even when the transmit link is in sleep mode.

Figure 6 shows the configuration panel for selecting the EEE option under Windows for a Broadcom Ethernet controller.

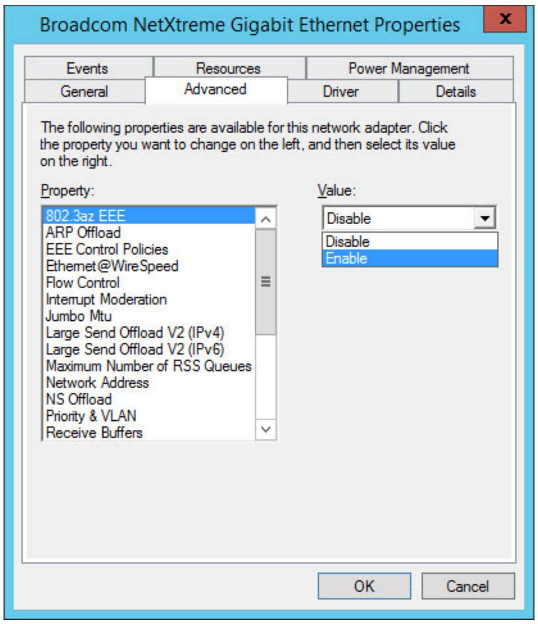


Figure 6 EEE configuration panel for a Broadcom controller in Windows Server 2012 R2

S3 State

In a high performance computing (HPC) environment, many servers work in parallel to solve a particular problem. However, not all problems require the same amount of server resources. For example, consider a cluster of servers with 1000 compute nodes. A compute-intensive problem might require all 1000 nodes, whereas a simpler problem might require only 500 nodes. To save power, 500 unused nodes can be powered down while the simpler problem is being solved. But what happens if another problem request arrives in the queue while the first problem is still being resolved? More compute nodes must be powered on and initialized to accommodate the new request. There is a large time penalty for those extra server nodes to get through the power-on-self-test (POST) and for the OS to start and get to a ready state. Too high of a time penalty can lead to service level agreement (SLA) violations.

An alternative is to leave the unused nodes sitting idle at the OS desktop. In that manner, the OS can get to a ready state quickly. However, the trade-off here is the power that is used is much higher compared to powering the nodes down. A mode that has low-power usage is needed that allows the OS to be restored to the ready state quickly.

[3] The EEE defines higher Ethernet networking standards to reduce power consumption. Defined by the Institute of Electrical and Electronics Engineers (IEEE).

S3 (sleep mode) is supported on select configurations of Lenovo NeXtScale™ System servers. S3 mode provides low power consumption and allows the OS to be restored to the ready state quickly. Figure 7 shows a typical comparison.

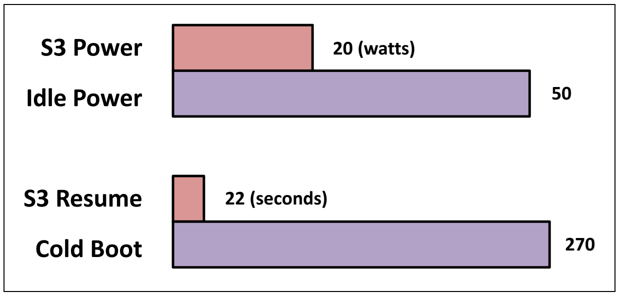


Figure 7 S3 low power and low latency

When a NeXtScale System® server enters S3 state, it essentially goes through the same steps that a notebook completes when the user closes the lid. Computer activity is acquiesced and the computer's state is saved to system memory. Then, when the notebook is reopened, its state is restored from system memory and it picks up where it left off before the lid was closed. The difference between notebooks and NeXtScale System servers is, on the servers, S3 is used on a massive scale on many compute nodes. Control tower software intelligently decides when each server node is placed into S3 or exited from S3. This decision is based on built-in workload algorithms and user configuration options.

For more information about S3 and other system S3 states, see "System power states" on page 23.

USB selective suspend

When USB selective suspend is enabled, a USB hub can suspend operation of individual USB ports based on the amount of activity on a port. This ability can be useful for USB devices that are used intermittently, such as keyboards, mice, and printers. When unused USB devices are suspended, the CPU cores spend more time in the deep power-saving states (for example CPU C3 and C6) when the OS is idle. This state occurs because the CPU cores no longer must intermittently process the transfer schedule from the USB hub. With today's modern CPUs, the savings from USB selective suspend can be 10 W when a server is idle.

Figure 8 on page 13 shows the Control Panel Power Options window for enabling and disabling USB selective suspend under a typical Windows OS.

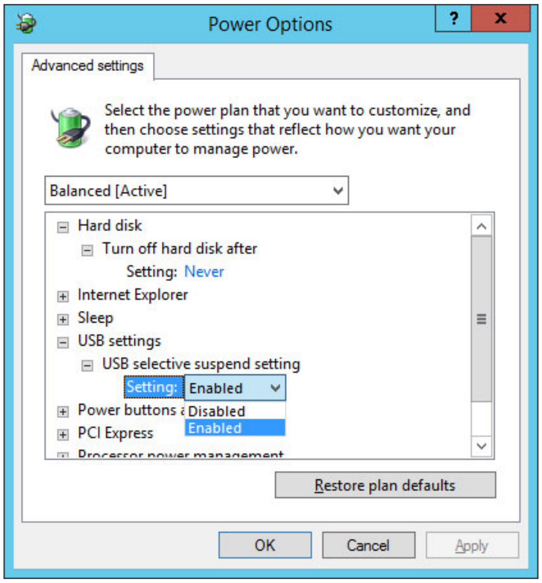


Figure 8 USB selective suspend setting in the Control Panel Power Options window

240 VA Protection optimized for efficiency

Server PSUs provide bulk 12 V DC power that is delivered to the various components in the server. The rated power of the bulk 12 V rail can reach thousands of watts and hundreds of amps. Without adequate protection, this high power is a safety risk to the user and a risk to the server hardware if a short circuit occurs.

Some industry standard servers protect the user by offering mechanical protection against the high power on the bulk 12 V rail. That is, plates or covers are used in an attempt to prevent access to the portion of the server that contains the high-power 12 V rail. Other servers offer no protection and only present warning labels to the user to indicate the presence of high power.

System x servers protect the user and the server hardware by splitting the bulk 12 V power rail into several smaller 12 V power rails and actively monitoring each of the smaller 12 V power rails. Each of the smaller 12 V rails is limited to 240 W for any rail that can be made available to service personal. The current on each of the 240 W power rails is constantly monitored. If the current exceeds a safe limit, the rail is quickly powered down, which protects the user from electric shock and the hardware from short circuits.

While monitoring the 240 W power rails, some small power losses are incurred. However, System x M5 servers minimized these power losses by the use of advanced circuitry that is optimized for light and heavy loads. System x M5 products offer superior user safety and electrical protection while maintaining optimal electrical efficiency.

Unused devices

Unused devices that are left powered on in a server waste power and act as small heaters. The data center is tasked with removing waste heat from those devices that are doing no useful work.

Essentially, any unused device that is doing no work has 0% efficiency and adds to the power overhead of the entire server. To combat this issue, unused devices that are embedded in System x servers are powered down or placed into a low-power state. This process is done automatically during the POST or dynamically at run time. The following devices are intelligently managed for power:

- CPU cores
- Memory channels and DIMMs
- PCI Express ports
- QPI links
- SATA and SAS storage controllers
- Network controllers
- Serial ports
- USB controllers
- VRDs

Energy efficient thermal design

Earlier versions of system level thermal control algorithms were fairly rudimentary, as they used an ambient inlet temperature sensor as the only proxy to fan speed control methods. More recently, there were sophisticated advances in the algorithms that are used and the efficient partitioning of the varying subsystems, which allows System x to reach near perfect optimization.

To better appreciate these advances, it is important to understand the need. Cooling on server designs in the early 2000s used as much as 10+% of total system power, which is a large part of the power budget. With industry focus on energy consumption, continued efforts were made to implement more sophisticated controls, which now yielded solutions that use as little as 2% of total system power while running applications. The added intelligence in the use of power sense circuits, ambient temperature, thermal sensors, configuration data, and the zoned cooling approach gave rise to these significant system level improvements.

In addition to system level advances, there is a growing movement to improve algorithms and controls beyond those of the server, which targets a more holistic solution at the rack and room level integration strategies. Some of the newer System x flagship products now offer environmental health awareness consoles, where hot air recirculation can be detected, and thus optimized at the rack level. Along with these feedback mechanisms, improved user feedback and data center integration controls are available.

Energy efficient turbo mode

Historically, turbo operation on Intel CPUs was a binary function. The maximum available turbo frequency was engaged immediately when the OS requested more performance. The downside to the binary approach is that the CPU can thrash in and out of turbo mode, often over a short period. The transitions into and out of turbo mode use some incremental power. Constantly switching in and out of turbo mode was not energy efficient.

Later generations of CPUs included a power optimized turbo mode, in which a short delay transpired before the CPU was granted maximum turbo frequency. In this manner, short and sporadic turbo requests from the OS were filtered out and turbo ran more efficiently.

The latest generation of Intel CPUs takes turbo efficiency one step further. System x M5 products use this new mode called energy efficient turbo. When energy efficient turbo mode is enabled, the OS might request the maximum turbo frequency. However, the actual turbo frequency that the CPU is set to is proportionally adjusted based on the duration of the turbo request. In addition, the memory usage of the OS is also monitored. If the OS uses memory heavily and the CPU core performance is limited by the available memory resources, turbo frequency is reduced until more memory load dissipates and more memory resources become available.

CPU uncore frequency scaling

CPU P-states are used to run the CPU cores at different frequencies based on workload demand (for more information, see “P-states” on page 28). Before the release of System x M5 products, all of the CPU cores in a CPU package ran at the same frequency, regardless of whether a workload was running on one CPU core or all CPU cores. Also, the CPU package *uncore* (non-core related, for example, QPI links and miscellaneous logic) ran at a fixed ratio of the CPU frequency.

Beginning with the System x M5 products, the uncore now runs at a speed that is independent of the CPU cores. This ability can save power for workloads where the ratio between CPU core demand and uncore demand is not always a fixed ratio. For example, a workload that runs on CPU cores and memory in the same package might require high CPU core frequencies, but low QPI link frequencies. Conversely, a moderate workload that uses memory on multiple CPU packages requires higher QPI frequencies when the workload is accessing memory on a remote CPU socket.

Power bias and performance bias

The power and performance bias setting controls how aggressively the CPU is power managed and placed into turbo. As the bias is adjusted towards harnessing performance, turbo is engaged more quickly and the power management features are engaged less. This configuration has the overall effect of increasing performance and decreasing latency. However, power is also increased.

Conversely, adjusting the bias towards power causes turbo to be engaged less and the power management features to be engaged more. Performance and latency can increase, but power savings also increases. Figure 9 shows the influence of the bias setting.

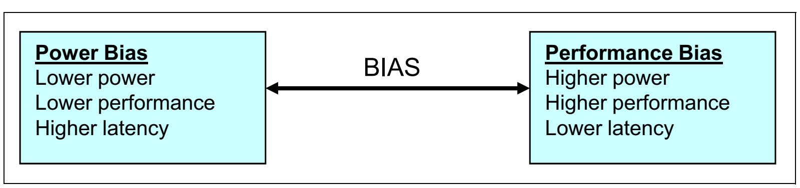


Figure 9 Influence of the bias setting

Lenovo efficiency mode

Efficiency mode works with the OS to fine-tune the operating efficiency of the server. It ensures that enough CPU performance is provided to the current tasks without over-speeding the CPU core. The algorithm optimally tunes the CPU operating frequencies based on the OS scheduler and by monitoring independent CPU performance meters in hardware.

To enable efficiency mode, enter setup (press F1 at POST), then choose **System Settings** → **Power**. As an alternative, the setting can be changed with the advanced settings utility (ASU).

Power metering

System x power metering offers powerful and detailed data collection and analysis capability without the need for cumbersome external power metering that is outfitted in PDUs or with an in-series power device that is added to the rack or upstream power delivery infrastructure. Power metering is available for AC input power, CPU power, and memory power.

Power meter data can be obtained in one of the following ways on System x M5 servers:

- Integrated Management Module (IMM): Power meter date can be read by using the GUI or the command-line interface (CLI) of the IMM. Examples of the IMM web interface are shown in Figure 10, Figure 11 on page 17, and Figure 12 on page 17.

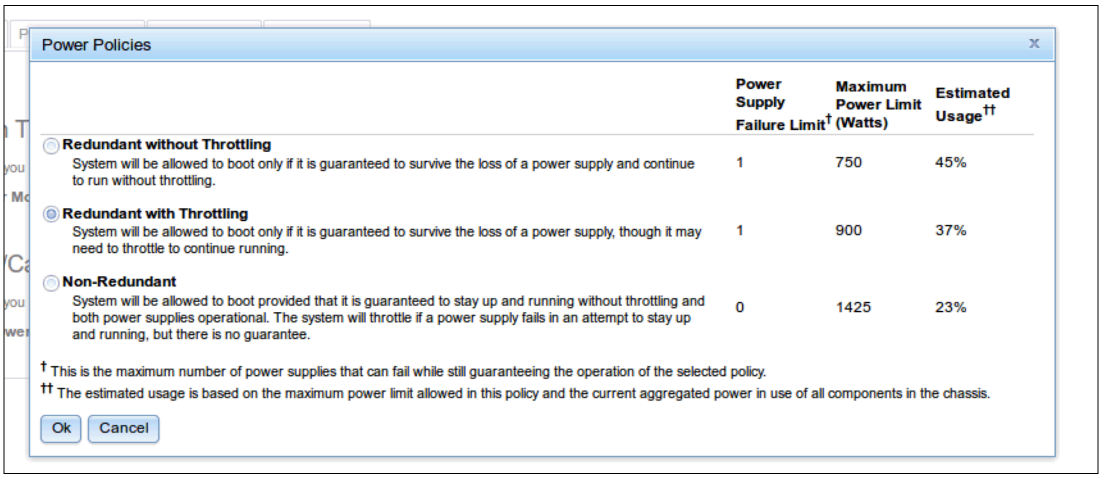


Figure 10 Power policies

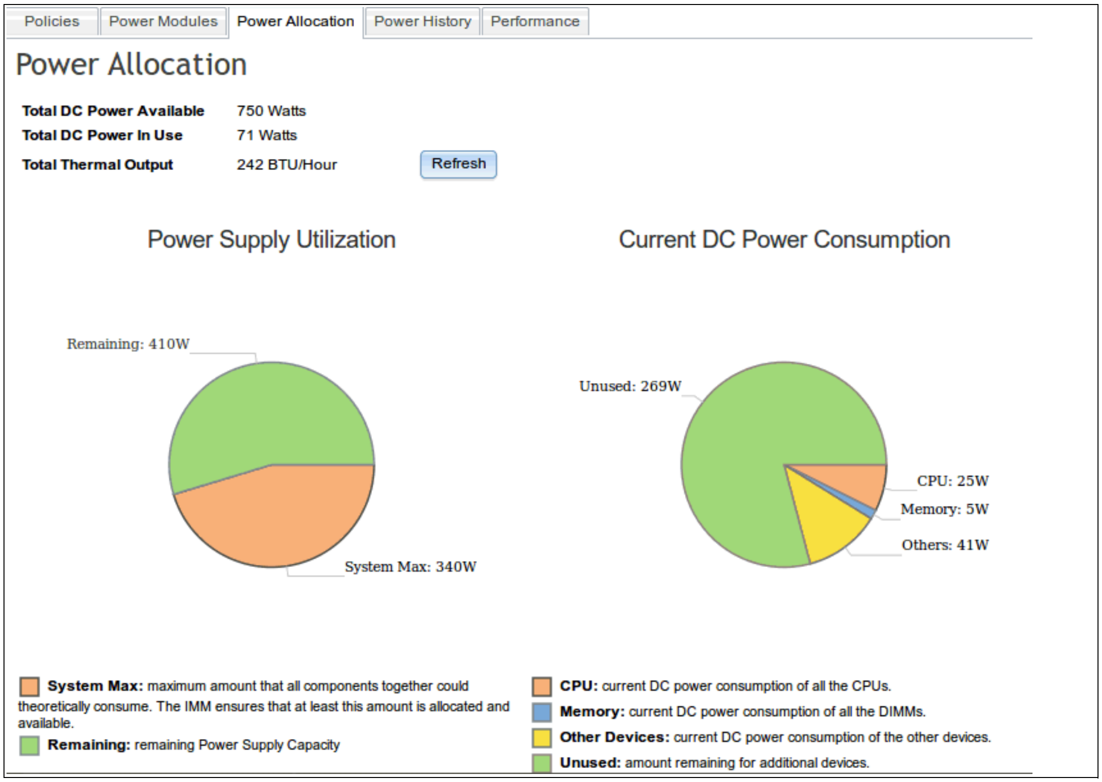


Figure 11 Power consumption metrics

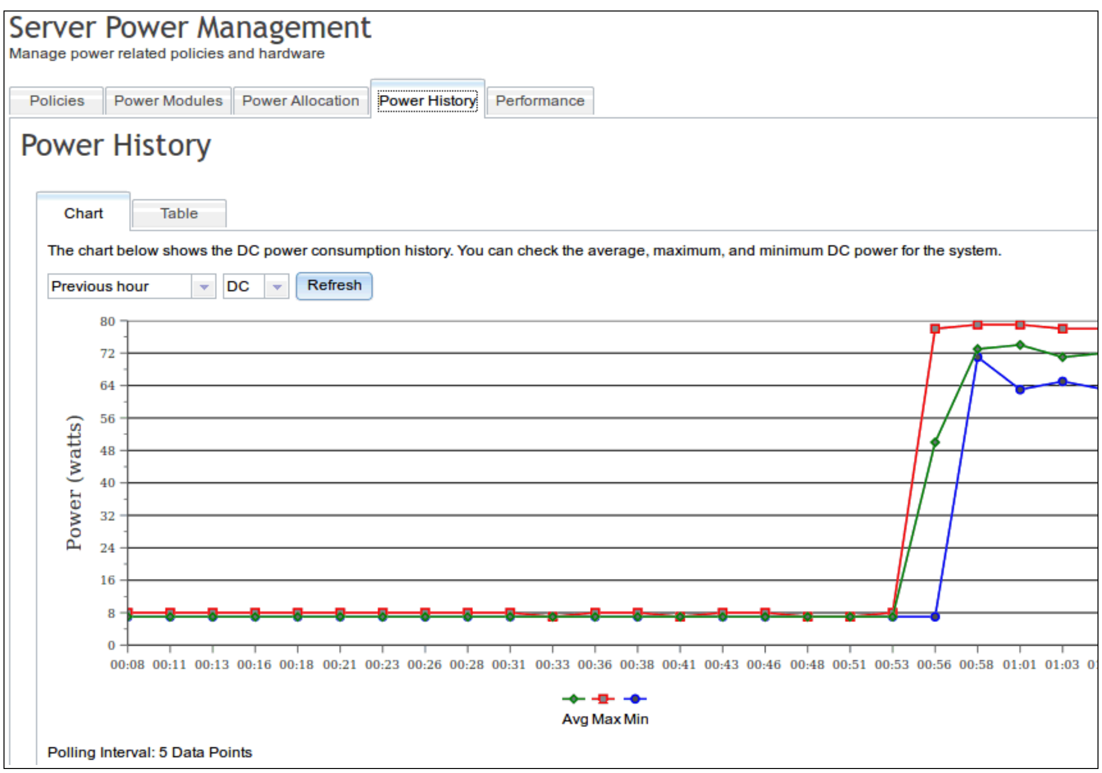


Figure 12 Power consumption history

- Node Manager 3.0 Intelligent Platform Management Interface (IPMI): By using this interface, users can issue Node Manager IPMI commands through the IMM systems management controller on the server, as shown in the following example (for more information, see the user guide for your server and the latest Node Manager specification):

  Get Node Manager Statistics (command code 0xc8)

  Example - Get Global System Power Statistics:

  Request:

  ```
  ipmitool -H $IMM_IP -U USERID -P PASSWORD -b 0x00 -t 0x2c raw 0x2E 0xC8 0x57
  0x01 0x00 0x01 0x00 0x00
  ```

  Response:

  ```
  57 01 00 38 00 04 00 41 00 39 00 ec 56 f7 53 5a 86 00 00 50
  ```

  ```
  System wattage =0x38 =56W AC (min =4W , max=65W, avg=57W)
  ```

- Data Center Manageability Interface (DCMI): DCMI is an industry standard interface specification that is management software neutral, which provides monitoring and control functions that might otherwise be made available through standard management software interfaces. For more information, see the IMM user guide for your server and the latest DCMI specification. Consider the following example:

  DCMI Get Power Reading Command Example

  Request:

  ```
  ipmitool -H $IMM_IP -U USERID -P PASSWORD raw 0x2c 0x02 0xdc 0x01 0x00 0x00
  ```

  Response:

  ```
  dc 39 00 38 00 3b 00 39 00 e3 6f 0a 39 e8 03 00 00 40
  ```

  ```
  System wattage =0x39 =57W AC (min=56W, max=59W, avg=57W)
  ```

- Upward Integration Modules (UIM): Power meter data can also be read by using the UIM for Microsoft System Center and VMWare vSphere, as shown in Figure 13 and Figure 14 on page 19.

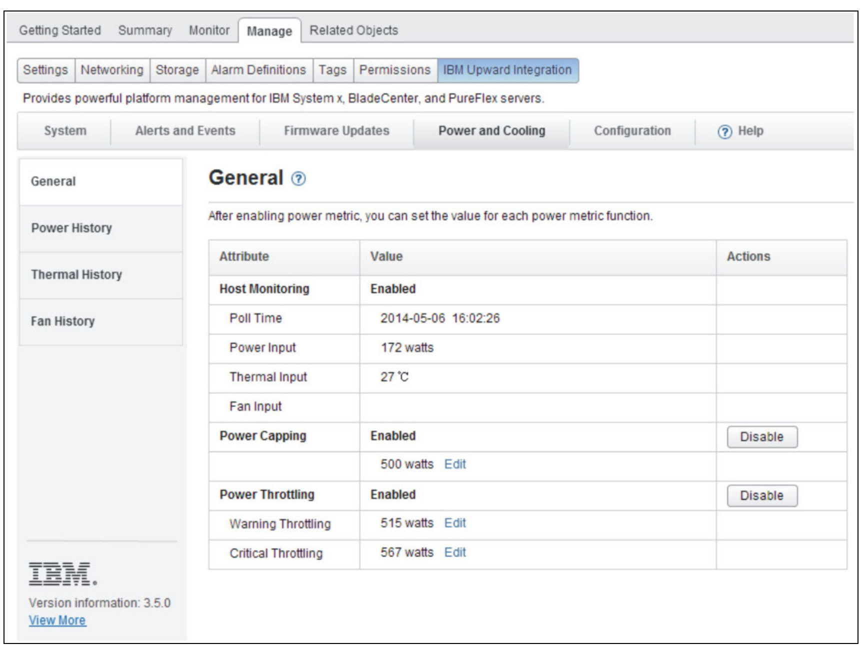


Figure 13 UIM power statistics and settings

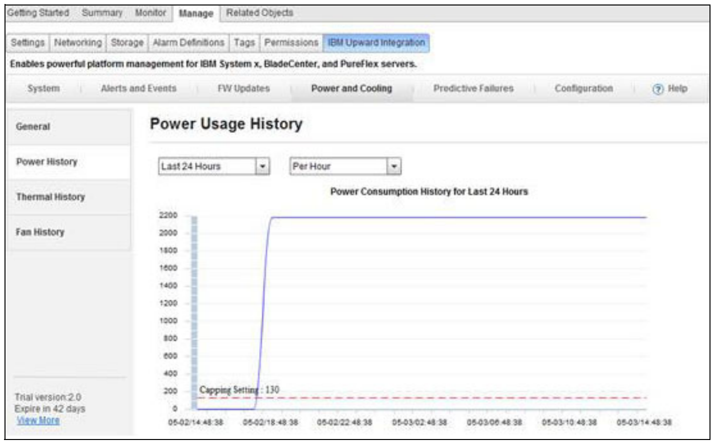


Figure 14 UIM power trending

For more information about the UIMs, see these websites:

- http://ibm.com/support/entry/portal/docdisplay?lndocid=LNVO-MANAGE
- http://ibm.com/support/entry/portal/docdisplay?lndocid=LNVO-VMWARE

Power capping

Similar to power metering capabilities, System x M5 servers also support power capping. Power capping can be used to limit the maximum power that a server uses. It is beneficial in curbing random spikes or surges in power, which allows rack and data center power limits to be maintained.

Server power capping can be controlled in many ways, including the use of the GUI or the CLI in the following interfaces:

- IMM: Figure 15 shows the power capping GUI interface that uses IMM.

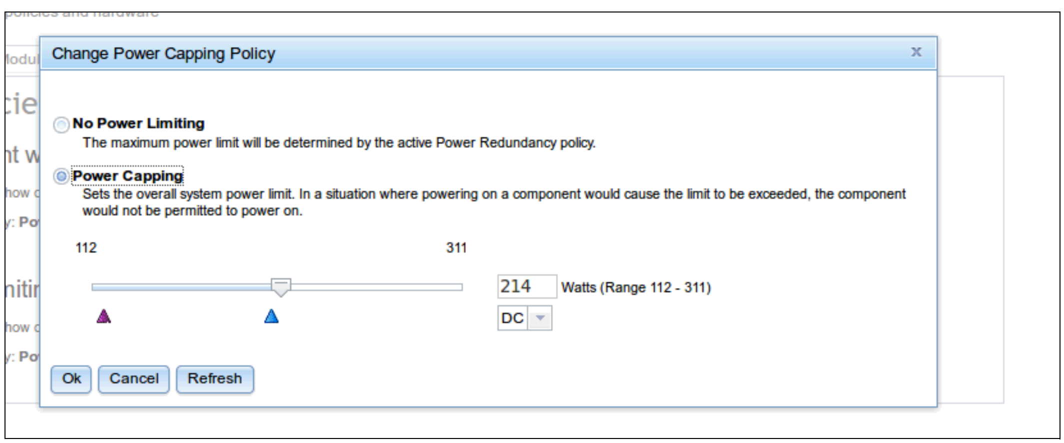


Figure 15 IMM GUI for power capping

- Node Manager 3.0 IPMI: The following example shows how to set a power cap by using the Node Manager 3.0 IPMI interface:

  Set NM Power Cap Command Example

  Set Node Manager Policy (command code 0xc1)

  Example – Set CPU domain power limit =100W with a policy ID=0x60

Request:

```
ipmitool -H $IMM_IP -U USERID -P PASSWORD -b 0x00 -t 0x2c raw 0x2E 0xC1 0x57
0x01 0x00 0x11 0x60 0x10 0x00 0x64 0x00 0xE8 0x03 0x00 0x00 0x00 0x00 0x05 0x00
```

Response:

```
57 01 00
```

- DCMI: A power cap can be set by using the DCMI IPMI interface, as shown in the following example:

  Set DCMI Power Cap Command Example

  Set AC power limit =100W

  Request:

```
ipmitool -H $IMM_IP -U USERID -P PASSWORD raw 0x2c 0x04 0xdc 0x00 0x00 0x00 0x00
0x64 0x00 0xe8 0x03 0x00 0x00 0x00 0x00 0x05 0x00
```

  Response:

```
dc
```

- UIM: Figure 13 on page 18 and Figure 16 show how to set power capping capabilities by using UIM for Microsoft System Center and VMware vSphere.

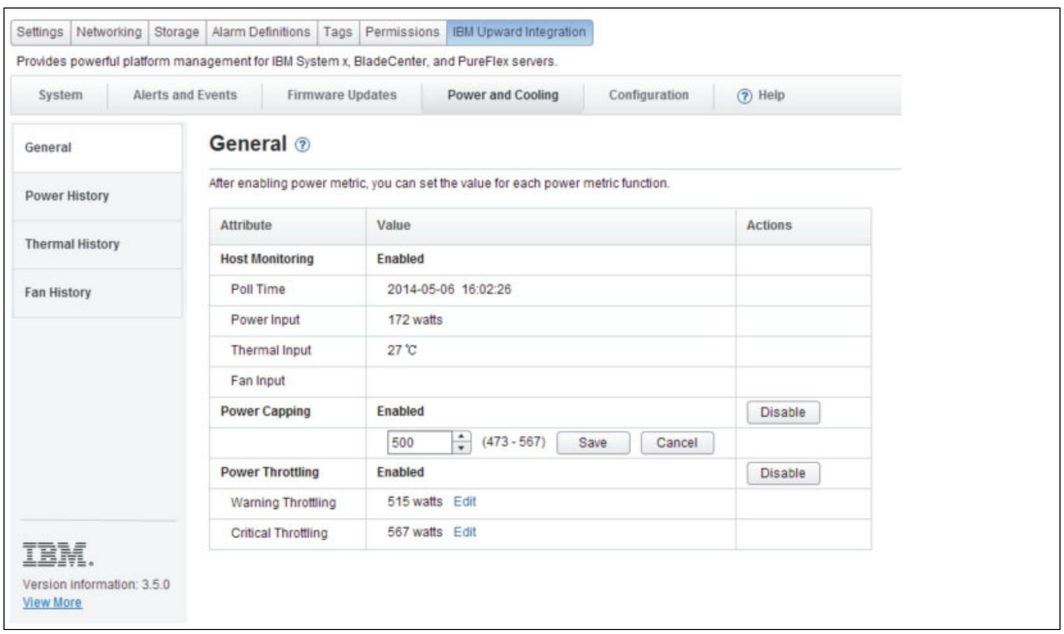


Figure 16 UIM power capping

For more information, see the user guide for your server and the latest UIM documentation.

Energy Star certification

Energy Star sets a minimum set of standards that are required to declare that a server is energy efficient and certified to be Energy Star compliant. Lenovo makes design choices to ensure that many System x servers are Energy Star compliant and certified.

For more information, see the following websites:

- http://www.lenovo.com/social_responsibility/us/en/energy/
- http://www.energystar.gov/productfinder/product/certified-enterprise-servers/results

Higher temperature limits

In 2012, the American Society of Heating, Refrigeration, and Air Conditioning Engineers (ASHRAE)[4] expanded IT equipment server design guidelines to include two more classes that support up to 40 °C (103 °F) and 45 °C (114 °F). These classes were specified in response to regulatory pressures in parts of Europe and Asia, which allows the server hardware to operate at higher temperatures, effectively allowing clients to use free-air cooling schemes that do not require chillers.

Aside from free air cooling schemes and the inherent benefits of those schemes, typical data centers use air handlers and chillers to remove heat. Even in these environments, it was demonstrated that raising the inlet temperature to IT equipment can reduce overall data center power consumption because it allows the air handlers and chillers to run at lower and more optimal set points. The savings can be significant, although there is an inflection point where the IT power draw exceeds the savings at the room level, which might adversely affect true net savings. Data from previous deployments suggests that the optimal temperature is most likely 25 °C - 27 °C for most solutions.

In concert with these industry changes, the ASHRAE book, *Thermal Guidelines for Data Processing Environments*[5], specifically addresses the questions and concerns about this topic. Lenovo System x recognizes the benefits regarding this shift in the industry and responded by supporting the new ASHRAE Class A3 standard that supports up to a 40 °C (103 °F) ambient inlet, which lessens or eliminates the need for chillers.

PDUs and power distribution

Advancements in System x PDUs improved usability and fault tolerance features, yielded higher overall efficiency, and improved power control capabilities. System x switched and monitored PDUs support extensive metering, data logging, and threshold alarms. Individual outlet power control also is supported. Unneeded equipment can be powered down to eliminate phantom loads. System x PDUs typically are rated at 95% or higher efficiency under normal conditions. Finally, System x PDUs protect against many power delivery problems, including the problems that are listed in Table 4.

Table 4 PDU power protection features

Problem	Description
Power failure	Also known as *blackout*. Voltage is turned off for an extended time period
Power sag	Also known as *brownout*. Voltage is reduced for a relatively short time.
Power surge	Also called *power spikes*. These spikes often last for a brief period and can be caused by lightning, motor start, or other sources of noise on the same line.

[4] American Society of Heating, Refrigeration, and Air Conditioning Engineers (ASHRAE). For more information, see this website: http://www.ashrae.org

[5] *Thermal Guidelines for Data Processing Environments, from ASHRAE*, which is available at this website: https://www.ashrae.org/resources--publications/publication-updates

Problem	Description
Under-voltage	Under voltage occurs when the power company delivers less voltage than specified. For example, nominal US voltage is 120 V. However, there are times when the power company might deliver only 100 V. The under-voltage is sometimes evident when lights dim.
Over-voltage	Over-voltage occurs when the power company delivers more voltage than specified.
Harmonic distortion	When harmonic distortion is present, the waveform appears as a non-ideal sine wave on an oscilloscope. Harmonic distortion can be generated from a load.
Line noise	Line noise appears as extra static that is riding on top of the base waveform. Double conversion is used to overcome line noise.
Switching transients	Switching transients appear as glitches or fast radical changes on the smooth sine wave. They occur in the voltage and current waveforms. A good example is a grocery store where many refrigeration units are used. Double conversion is used to overcome switching transients.
Frequency variation	Deviations from the base line frequency (for example, 60 Hz in US). Double conversion addresses this issue.

Summary: Role of efficiency in total cost of ownership

All of the power management and efficiency features that are described in this paper are designed to lower the total cost of ownership (TCO) for the server solution in the data center, specifically the operating costs. Of the following solutions that are described in this section, each contributes to lowering the total power consumption, which makes it easier to control the effect of power consumption in your data center:

- 80 PLUS Titanium power supplies
- System x advanced liquid cooling
- Support for higher operating temperatures
- Intelligent power management
- Ease-of-use attributes

In assessing the total cost of ownership and the benefits of an energy efficient solution, the following areas must be considered:

- Hardware and software costs
- The duration of time that each server node is powered on and off
- Server wattage when powered on and running the customer workload
- Server wattage when powered off
- Available workload capacity of each server node
- Electrical rate at the data center location
- Power factor surcharges and penalties
- Total heat load for the data center servers operating at full workload capacity
- Cooling infrastructure topology and efficiency
- Administrative costs of a standard solution versus an intelligent solution

Lenovo has many resources available to help our clients understand the TCO that is associated with System x solutions. For more information, contact a Lenovo System x representative.

For more information about data center energy efficiency and assessments, see the following resources:

- *IBM Data Center and Facilities Strategy Services – data center energy efficiency assessment*:

  http://www.ibm.com/services/multimedia/DC_energy_efficiency.pdf

- *Resiliency Services - Data center energy efficiency assessment*:

  http://www-935.ibm.com/services/us/en/it-services/data-center/data-center-energy-efficiency-assessment/

- Product energy efficiency:

  http://www.lenovo.com/social_responsibility/us/en/energy/

- Data center energy efficiency:

  http://www.ibm.com/ibm/environment/climate/datacenter.shtml

- Product energy efficiency:

  http://www.ibm.com/ibm/environment/climate/products.shtml

Other resources

This section offers more information about the following topics that are related to power and energy efficiency:

- “System power states”
- “Software considerations” on page 31
- “System x Power Configurator” on page 32
- “Efficiency definitions” on page 32

System power states

Several power states are in System x servers. The entire server or individual subsystems in the server can be placed into different power states to reduce power consumption and optimize efficiency. Some of the power state transitions are started by the user. Other transitions are started automatically if they are enabled. Some power states are used to save power when the server is idle, while others are used to increase efficiency when it is running.

The following sections describe each group of power states that are used on System x servers. The relationship and hierarchy of the power states also are described.

G-states

G-states are *global server states* that define the operational state of the entire server. As the number of the G-state increases, more power is saved. However, as shown in Table 5, the latency to move back to G0 state also increases.

The user typically starts G-state transitions. For example, when an OS is shut down, the server is moved from G0 state to G2 state.

Table 5 G-states

G-State	Can Applications Run?	Relative Wake up Latency	OS Restart Required	Comments
G3	No	Longest	Yes	AC power is removed. OS is not loaded. Server is receiving power only from backup batteries for RTC, CMOS, and possibly RAID data.
G2	No	Long	Yes	System is in a soft-off state. For example, the power switch was pressed. System draws power from AUX rail of power delivery circuit. Power supply might or might not be switched off.
G1	No	Short to medium	No	Standby or hibernate mode under Windows. For more information, see “S-states” on page 25.
G0	Yes	None	No	System is fully on but some components might be in a power savings state.

<<<------- Higher Power -------

-------- Higher Latency --------->>

S-states

S-states define the *sleep state* of the entire server. Table 6 lists the various sleep states. S-states can be started by the user, by using an inactivity timer on the OS, or with a higher-level workload management software.

> **Note:** Although the ACPI specification defines the S2 and S4 states, these states are not enabled in System x servers.

Table 6 S-states

S-State	G-State	BIOS restart required	OS restart required	Relative power	Relative latency	Comments
S0	G0	No	No	6X	0	System is fully on but some components might be in a power savings state.
S1	G1	No	No	2.5X	1%	Also known as *Idle*, *Standby*; if S3 not supported. Typically, when the OS is idle, it halts the CPU and blanks the monitor to save power. No power rails are switched off. This state might be removed from future servers.
S2	G1No	No	No	N/A	N/A	CPU caches are powered down. No known server or OS supports this state.
S3	G1	No	No	1.1X	10%	Also known as “Standby” or “Suspend-to-RAM”. The state of the chipset registers is saved to system memory and memory is placed in a low-power self-refresh state. To preserve the memory contents, power is supplied to the DRAMs in S3 state.
S4	G1	Yes	No	X	90%	Also known as “Hibernate” or “Suspend-to-disk”. The state of the OS (all memory contents and chip registers) is saved to a file on the HDD and the server is placed in a soft-off state. This state is not used by System x servers.
S5	G2	Yes	Yes	X	100%	Server is in a soft off state. When turned back on, the server must completely restart with POST and the OS.

------- Higher Latency ------->>

<<<------ Higher Power ------

As with G-states, higher numbered sleep states save more power, but there is more latency when the system transitions back to S0 state. The middle state, S3, offers a good compromise between power savings and latency.

C-states

C-states are CPU idle power-saving states. C-states higher than C0 become active only when a CPU core is idle for a period. If a process is running on a CPU core, the core is always in C0 state. If hyper threading is enabled, the C-state resolves down to the physical core. For example, if one hyper thread is active and another hyper thread is idle on the same core, the core remains in C0 state.

C-states can operate on each core separately or the entire CPU package. The CPU package is the same physical chip of the CPU cores. It includes the CPU cores, caches, memory controllers, PCI Express interfaces, and miscellaneous logic. The non-CPU core hardware inside the package is commonly referred to as the *uncore*.

Core C-states transitions are driven by interrupts or the OS scheduler with MWAIT commands. The number of cores in C3 or C6 also affects the maximum turbo frequency that is available. Enable all of the CPU C-states if maximum peak performance is wanted.

Package C-state transitions are autonomous. No OS awareness or user intervention is required. The package C-state is equal to the lowest numbered C-state that any of the CPU cores is in at that point. More logic inside the CPU package monitors all of the CPU cores and places the package into the appropriate C-state.

CPU C-states do not directly map to ACPI C-states for historical reasons. ACPI C-states are C0 - C3. When they were defined, there were no CPUs that supported the C6 state. Therefore, the mapping was 1:1 (ACPI C0 =CPUC0, ACPI C1=CPU C1, and so on).

However, newer CPUs support the C6 state. Depending on the class of CPU that is used, the ACPI-to-CPU C-state mapping can vary above the ACPI C1 state. Typically, to get the maximum power savings, the highest numbered ACPI state map to the highest numbered CPU C-state. Some examples are listed in Table 7.

Table 7 C-state mapping

ACPI C-state	CPU C-State for Intel E5-2600 and E5-2600 v2	CPU C-State for Intel E5-2600 v3
C0	C0	C0
C1	C1	C1
C2	C3	C6
C3	C6	Not used

Table 8 lists each core and package C-state.

Table 8 G-states

C-state	CPU core state	CPU core power / latency approximation[a]	CPU package state	CPU package power and latency approximation[b]
C0	► Core is fully on and running code. ► L1 cache is coherent. ► Core power is on.	100% at Pn / 0nS	At least one core is in C0 state.	100% / 0nS
C1	► Core is halted. ► L1 cache is coherent. ► Core power is on.	30% / 5uS	NA –core only state	NA
C1E	NA –package only state	NA	► At least on core is in C1 state and all others are in C1 or a higher numbered C-state. ► All cores are running at lowest frequency. ► VRD 2 switches to minimal voltage state. ► PLL is on. ► CPU package processes bus snoops.	50% / ~5uS
C3	► Core is halted. ► L1 cache is flushed to last level cache. ► All core clocks stopped ► Core power is on.	10% / 50uS	► At least one core is in C3 state and all others are in C3 or a higher numbered C-state. ► VRD 2 is in minimal voltage state. ► PLL is off. ► Memory is placed in self-refresh. ► L3 shared cache retains context but is inaccessible. ► CPU package is not snoopable.	25% / ~50uS
C6	► L1 cache is flushed to LLC. ► Core power is off	0% / 100uS	► All cores are in C6 state. ► Same power-saving features as package C3 plus some additional uncore savings.	10% / ~100uS

------- Higher Latency ------->>

<<<------- Higher Power ------

a. The number of C-states and the specific power savings that are associated with each C-state depends on the specific type and SKU of the installed CPU.

b. Voltage regulator device (VRD).

P-states

P-states are defined as the *CPU performance states*. Each CPU core supports multiple P-states and each P-state corresponds to a specific frequency (see Table 9). P0 is the fastest frequency. P*n* (where *n* is the maximum numbered P-state for the installed CPU) is the slowest frequency. P0 can run above the rated frequency for short periods of time if turbo mode is enabled. The exact turbo frequency for P0 and the amount of time the core runs at the turbo frequency is controlled autonomously in hardware.

Table 9 P-states

P-state	CPU Frequency Approximation (100% =rated CPU frequency)	Description
P0	100 to ~130% (with turbo)	CPU can run at the rated frequency indefinitely or at a turbo frequency greater than the rated frequency for short periods of time. Turbo is an opportunistic feature. The turbo frequency and the time that turbo can be sustained is not guaranteed.
P1	~90 - 95%	Intermediate P-state.
:		
:		
P*n*-1	~45 - 60%	Intermediate P-state.
P*n*	1200 MHz	Minimum frequency that CPU core can run code.

As with core C-states, P-states are controlled by the OS scheduler. The OS scheduler places a CPU core in a specific P-state, depending on the amount of performance that is needed to complete the current task. For example, if a 2 GHz CPU core needs to run at 1 GHz only to complete a task, the OS scheduler places the CPU into a higher-numbered P-state (slower frequency).

Each CPU core can be placed in a different P-state. Multiple threads on one core (for example, hyper threading) are resolved to a single P-state. P-states are valid only when the CPU core is in the C0 state. P-states are sometimes referred to as *dynamic voltage and frequency scaling* (DVFS) or *enhanced Intel speedstep technology* (EIST).

The exact frequency breakdown for the P-states varies with the rated frequency and power of the specific CPU SKU that is used.

In addition to controlling the core frequency, P-states indirectly control the voltage level of the VRD that is supplying power to the CPU cores. As the core frequency is reduced from its maximum value, the VRD voltage is automatically reduced down to a certain point. Eventually, the VRD operates at the minimum voltage that the CPU cores can tolerate. If the core frequency is lowered beyond this point, the VRD voltage remains at the minimum voltage, as shown in Figure 17 on page 29.

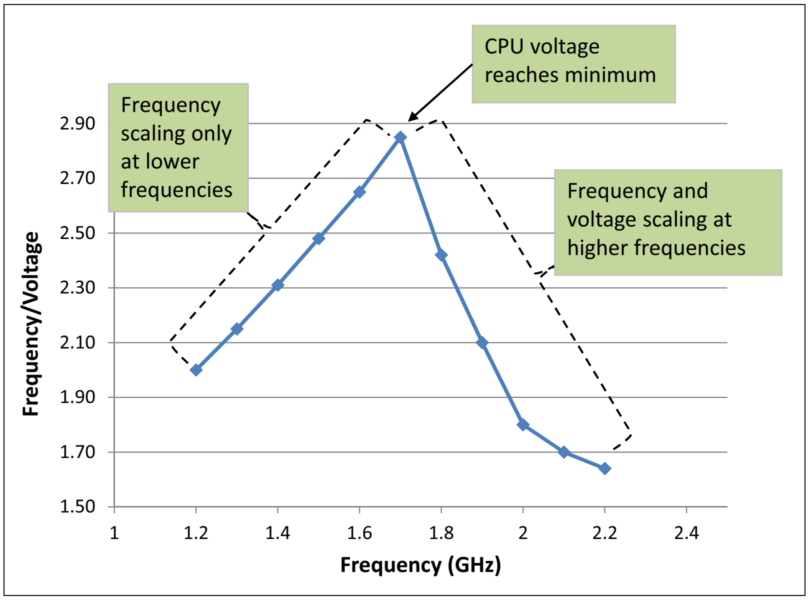


Figure 17 Effect of VRD voltage

Typically, the most efficient operating point is at the peak of the curve that is shown in Figure 17.

D-states

D-states are subsystem power-saving states. They are applicable to devices, such as LAN, SAS, and USB. The OS can automatically transition to different D-states after a period or when requested by a device driver. All D-states occur when the server is in S0 state. Table 10 lists each D-state.

Table 10 D-states

D-state	Device power	Device context	Description
D0	On	Active	Device is fully on. All devices support D0 by default even if they do not implement the PCI Power Management specification.
D1	On	Active	Immediate power state. Lower power consumption than D0. Exact power-saving details are device-specific.
D2	On	Active	Immediate power state. Lower power consumption than D1. Exact power-saving details are device-specific.
D3 hot (ACPI D2)	On	Lost	Power to device is left on but the device is placed in a low-power state. Device is unresponsive to bus requests.
D3 cold (ACPI D3)	Off	Lost	Power to device is removed. All devices support D3 by default even if they do not implement the PCI Power Management specification.

<<------- Higher Power -------

------- Higher Latency ------->>

M-states

M-states control the memory power savings. The memory controller automatically transitions memory to the M1 or M2 state when the memory is idle for a period. M-states are defined when the server is in S0 state only. Table 11 lists each M-state and the relative power and latency that is associated with each.

Table 11 M-states

M-state	Power / latency approximation	Description
M0	100% at idle / 0	Normal mode of operation
M1	80% / 30nS	Lower power CKE mode. Rank power down.
M2	30% / 10uS	Self-refresh. Operates on all DIMMs that are connected to a memory channel in a CPU package.

Relationships among the power states

Figure 18 shows an overview of the relationship among the power states of the server.

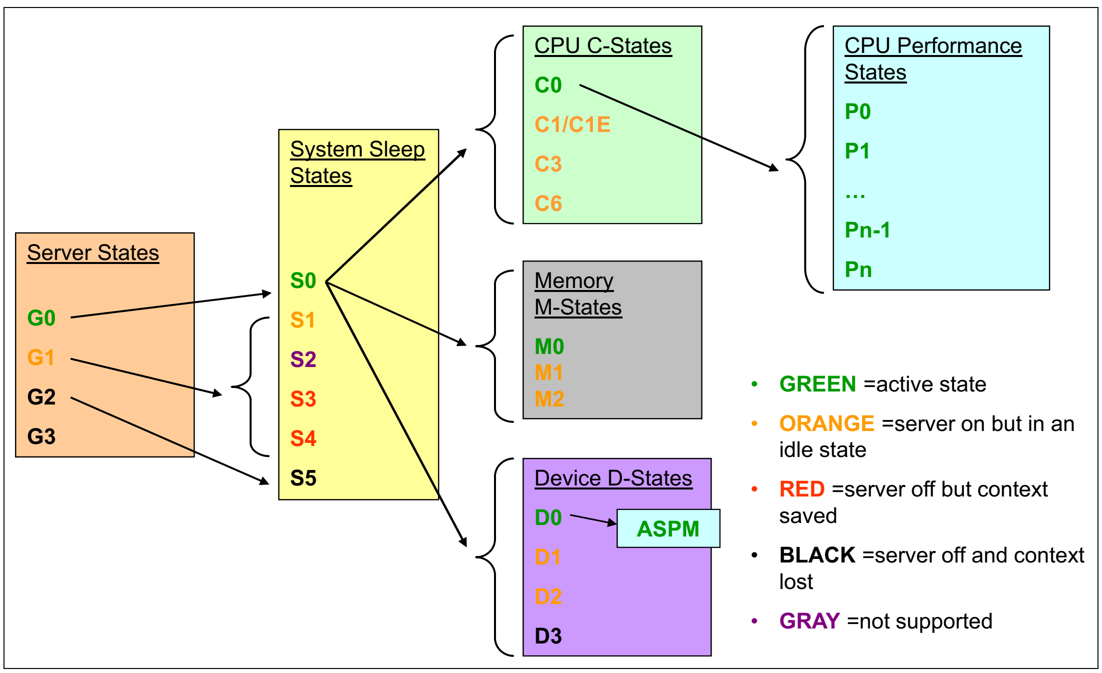


Figure 18 Power state interaction

There is a hierarchy among the power states. At the highest level, the G-states represent the overall state of the server. The G-states map to the S-states (system sleep states). Progressing to the right side of Figure 18, there are subsystem power states that represent the current state of the CPU, memory, and subsystem devices. As shown by the arrows in Figure 18, certain power states cannot be entered if higher-level power states are inactive. For example, for a CPU core to be in P1 state, the CPU core also must be in C0 state, the system must be in S0 system sleep state, and the overall server must be in G0 state.

With any of the power savings states, there is a trade-off between power savings and latency. For example, enabling the CPU C6 state allows CPU cores to be turned off, which saves power. However, because the CPU cores are powered down, it takes more time to restore their state when they transition back to the C0 state.

If maximum overall performance is wanted, all power-saving states can be disabled. Although this configuration minimize latencies to transition into and out of the power states, power is increased dramatically. At the other extreme, if power settings are optimized for maximum power savings, performance can suffer because of long latencies. For most applications, the default system settings offer a good balance between performance and efficiency. If necessary, the defaults can be changed if increased performance or power savings are wanted.

For more information about system power states, see the following ACPI Advanced Configuration and Power Interface website:

http://www.acpi.info/

Software considerations

This section describes the following methods that can be used to adjust the power usage in your environment:

- Unified Extensible Firmware Interface (UEFI) and Advanced Settings Utility™ (ASU)
- OS profiles
- Steady state turbo mode

UEFI and ASU

System x servers contain many UEFI settings that can be changed by entering setup (press F1 at POST) or with ASU. Carefully consider making changes to these settings for achieving optimal efficiency. For more information about each UEFI setting, see the User's Guide for the System x server of interest. We suggest that users who are unfamiliar with the low-level settings chose one of the following preset operating modes:

- Minimal Power mode

  This mode minimizes the absolute power consumption of the system while the system is operating. The trade-off is that performance might be reduced, depending on the application that is running.

- Efficiency – Favor Power mode

  This mode maximizes the performance/watt efficiency with a bias towards power savings. It provides the best features for reducing power and increasing performance in applications where maximum bus speeds are not critical.

- Efficiency – Favor Performance mode

  This default mode optimizes the performance/watt efficiency with a bias towards performance; speeds are derated.

- Maximum Performance mode

  This mode maximize the absolute performance of the system with little regard to power consumption. However, fan speed, heat output of the system, and power consumption might increase. Efficiency of the system might decrease in this mode, but the absolute performance can increase depending on the benchmark that is run.

Custom settings can be determined, with which you can individually modify any of the low-level settings that are preset in these four preset modes.

For more information, see this website:

https://ibm.com/support/entry/portal/docdisplay?lndocid=lnvo-asu

OS power profiles

Similar to UEFI settings, OSs also have profiles and settings with which users can adjust for maximum power savings, efficient operation, or maximum performance. Changing low-level settings for a plan is recommended for advanced users only who are familiar with each of the low-level settings. For most users, the defaults for the preset mode are acceptable. For more information, see the documentation for your specific OS.

An example of a power plan selection pane under typical Microsoft Windows OSs is shown in Figure 19.

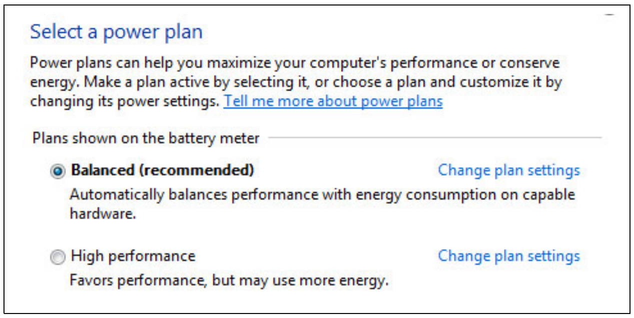


Figure 19 Windows OS power plans

When the UEFI and OS power management and efficiency settings are compared, consider the UEFI settings as controlling the outer limits of the power management and efficiency settings. The OS operates within the limits set by UEFI. In addition, the OS can place more restrictions on the power management and efficiency of the server.

Steady state turbo mode on System x servers with Intel E5 processors

For more information about the trade-offs of maximum performance versus efficiency, see *Steady State Turbo Mode on System x Servers with Intel E5 Family Server CPUs*, which is available at this website:

http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=SA&subtype=WH&htmlfid=XSW03142USEN

System x Power Configurator

By using the Power Configurator, the worst-case power consumption for a specific server configuration can be determined before a purchase is made. For more information, see this website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=LNVO-PWRCONF

Efficiency definitions

This section describes the three methods that can be used to measure the efficiency of a server.

Electrical conversion efficiency (ECE) measures how much power is lost to convert one power level to another (for example, AC-to-DC or DC-to-DC conversion). If a power supply converts 220 V AC to 12 V DC and it is 95% efficient for a 500 W load, 5% of the input power is converted to heat and typically is dissipated with a fan built into the power supply.

In this example, 526 W AC is required, 500 W is delivered to the load, and 26W is dissipated as heat. Although power supply and VRD efficiency improved dramatically in recent years, no electrical circuit is ideal and some power is always dissipated.

The following methods can be used to measure the efficiency of a server:

- ECE =Power out / Power In

  Power usage effectiveness (PUE) measures how much power is lost in the data center to relative to actual IT equipment power. The overall PUE depends on how close to the true compute power load that the power out measurement is taken and what ancillary loads are included in the calculation (for example, lights, humidification, UPS, CRACs, and chillers).

- PUE =Total Facility Power / IT Equipment Power =1 / data center Efficiency

  Performance/watt efficiency (P/W E) is defined as how much performance can be achieved for every watt of power that is consumed.

- P/W E = Σ Performance / Σ Power

  P/W E focuses on the server, chassis, and rack efficiency. By comparison, ECE, or PUE can be extended to the data center level or power station level.

About the Author

Robert (Bob) R. Wolford is a Senior Engineering Staff Member for power management and efficiency at Lenovo. He covers all technical aspects that are associated with power metering, management, and efficiency. Previously, he was a lead systems engineer for workstation products, video subsystems building block owner for System x, System p, and desktops, signal quality and timing analysis engineer, gate array designer, and product engineering. In addition, he was a Technical Sales Specialist for System x. Bob has 11 issued patents and three pending patent applications. He holds a Bachelor of Science degree in Electrical Engineering with Distinction from the Pennsylvania State University.

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service.

Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
1009 Think Place - Building One
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary.

Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk.

Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This document REDP-5259-00 was created or updated on August 24, 2015.

Send us your comments in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:
  ibm.com/redbooks
- Send your comments in an email to:
  redbooks@us.ibm.com

Trademarks