



# Power Regulator for ProLiant servers

technology brief, 4th edition

Abstract.....	2
Introduction.....	2
Processor P-states.....	2
Power Regulator technology.....	3
Power Regulator modes.....	3
HP Static High Performance mode.....	4
HP Static Low Power mode.....	4
HP Dynamic Power Savings mode.....	4
Configuring Ultra Low Power state.....	4
OS Control mode.....	4
Power Regulator and Dynamic Power Savings.....	5
Measuring CPU utilization.....	5
Processor granularity.....	5
Power consumption and system performance.....	6
Power management and reporting.....	10
Insight Power Manager.....	10
Integrated Lights-Out (iLO) and Integrated Lights-Out 2 (iLO 2).....	12
Management.....	12
Reporting.....	12
Typical uses for Power Regulator.....	13
Reducing power and cooling cost.....	13
Increasing server density and performance per watt.....	15
Ensuring uptime.....	15
Summary.....	16
For more information.....	17
Call to action.....	17

# Abstract

HP Power Regulator is an innovative, operating system-independent, power management feature of HP ProLiant servers. Using Power Regulator, a system administrator can manage processor power consumption and system performance to meet critical business needs. This paper describes Power Regulator technology, how it can be used, and its relationship to other power management tools. It is assumed that the reader is familiar with ProLiant servers and the Integrated Lights-Out (iLO) management processor.

## Introduction

Enterprise server systems place large demands on facility power and cooling resources. Expanding these systems challenges IT organizations to maximize existing facility resources that may not be as scalable as the server system. Facility power and cooling limitations can, in fact, delay the expansion of existing server systems.

HP Power Regulator is a standard feature on HP ProLiant servers 200-series and above. It leverages processor power state registers exposed by processor vendors. Power Regulator enables control of processor power usage and performance to minimize power consumption, maintain desired performance levels, and maximize facility resources.

HP Power Regulator is implemented in firmware and is therefore not affected by operating system or application upgrades. The Power Regulator firmware monitors the application load of the processor and, depending on the power mode selected, controls the power state of the processor to provide efficient use of processor power. Power Regulator includes support for dynamic and static power saving modes on multiple ProLiant server models. It increases system efficiency by freeing up system and cooling power for use when and where it is most necessary.

## Processor P-states

Server processors from Intel and AMD have power state registers that are available to programmers. With the appropriate interface, these hardware registers can be used to switch a processor between different performance states or P-states<sup>1</sup>.

Changing the performance state (that is, the processor frequency and voltage) enables processors to operate at different power levels and thus allows some control over total system power consumption. Today's processors may have as many as five defined P-states. However, the difference in power consumption between some P-states is too small to make a significant contribution to overall power savings. Tables 1 and 2 show the maximum and minimum P-states on examples of Intel® Xeon™ and AMD Opteron processors.

**Table 1.** P-states of the Quad-core Intel Xeon E5420 3.00-GHz processor

P-state	Description	Core Frequency
Pmax	Maximum performance	3.0 GHz
Pmin	Minimum power	2.0 GHz

<sup>1</sup> **P-states**—The ACPI body defines P-states as processor performance states. For Intel® and AMD processors, a P-state is defined by a fixed operating frequency and voltage.

**Table 2.** P-states of the Quad-core AMD Opteron 2356 2.3-GHz processor

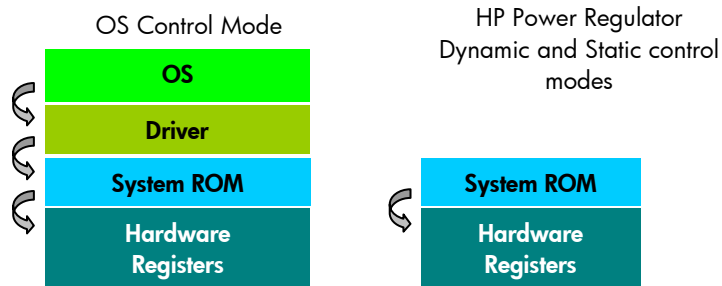
P-state	Description	Core Frequency
Pmax	Maximum performance	2.3 GHz
Pmin	Minimum power	1.15 GHz

## Power Regulator technology

HP Power Regulator functionality is a powerful tool that uses processor P-states to control the power consumption and performance profiles of servers. Using Power Regulator, system administrators can configure a server to maximize power savings or to dynamically match processor power consumption with system load to maintain an optimal balance of performance and power utilization under all operating conditions.

Processor P-states can be controlled either through the operating system or directly by the system ROM firmware (Figure 1). HP Power Regulator functionality is implemented completely in the system firmware, providing a consistent mechanism for optimizing processor power use in all servers, regardless of the operating system in use.

**Figure 1.** Methods for controlling processor P-states



Power Regulator can be configured immediately on a single system using any of several methods, including ROM Based Setup Utility (RBSU), Integrated Lights-Out (iLO) browser, or the iLO command line. Multi-server setup is supported using the iLO scripting interface with batch processes, the Rapid Deployment Pack, or HP Insight Power Manager.

## Power Regulator modes

System administrators can configure Power Regulator for any of four modes of operation:

- HP Static High Performance mode
- HP Static Low Power mode
- HP Dynamic Power Savings mode
- OS Control mode

Because Power Regulator runs from system firmware, changes between HP modes take effect immediately. A system reboot is required when changing into or out of OS Control mode.

Not all Power Regulator modes are supported on all ProLiant servers and processor models. Beginning with G5 servers, however, all Power Regulator modes are supported on all ProLiant 200-series and above server models. For a detailed list of processors supported by Power Regulator, consult the Power Regulator website at <http://www.hp.com/servers/power-regulator>

## HP Static High Performance mode

In Static High Performance mode, the system's processors operate continuously at the highest power and performance state (Pmax). In this mode, neither the system's firmware nor the operating system will ever program the processors to run in a lower power or performance state. This mode is useful for benchmarking a server's power consumption with the power savings technology of HP Power Regulator disabled.

## HP Static Low Power mode

In Static Low Power mode, the system's processors operate continuously at the lowest power state (Pmin). In this mode, Power Regulator programs the processors to run continuously in the low-power state to allow the server to run at a guaranteed lower maximum power level. This feature is useful for environments where power availability is constrained and it is critical that the server's maximum power use be lowered.

## HP Dynamic Power Savings mode

The Dynamic Power Savings mode of Power Regulator is similar in concept to Intel's Demand Based Switching with Enhanced Intel SpeedStep® Technology and to AMD's PowerNow!, but it adds a significant HP innovation. In HP Dynamic Power Savings mode, Power Regulator uses a ROM-based algorithm developed by HP to monitor processor activity. It adjusts processor power use to match performance levels to the application load. Dynamic Power Savings mode can be supported whether or not an operating system supports Intel's Demand Based Switching or AMD's PowerNow and regardless of which operating system the server is running. Current HP ProLiant servers enable HP Dynamic Power Savings mode by default, so they provide significant power savings out-of-the-box without affecting system performance.

## Configuring Ultra Low Power state

The Ultra Low Power state is a Power Regulator sub-mode that is unique to AMD-based G5 ProLiant servers and later. It is used to control which processor P-state is used for the low power state in both HP Static Low Power Mode and HP Dynamic Power Savings mode. When Ultra Low Power state is enabled, Power Regulator uses the lowest defined processor P-state for the low power state. When it is disabled, the next-lowest P-state is used. Ultra Low Power state is enabled by default and can be configured through the ROM Based Setup Utility.

This functionality was implemented to provide maximum flexibility when configuring Power Regulator on AMD-based ProLiant servers, since the lowest P-states for the AMD processors are somewhat lower than those of Intel processors. Testing has shown that for most processor utilization profiles, enabling Ultra Low Power State provides the most power savings with little or no performance degradation.

## OS Control mode

In OS Control mode, HP Power Regulator functionality is not directly active. Dynamic power management for the system is managed instead by the operating system through its policy mechanism. In this mode, the HP system BIOS ROM creates the required ACPI tables to enable the operating system to support Intel's Demand Based Switching or AMD's PowerNow! power

management feature. For this mode, system administrators must configure the operating system to activate the OS-based power management feature. If the operating system does not support dynamic power management, or if the feature has not been configured through the operating system, the processor will always run in its highest power and performance state.

---

**NOTE:**

Some earlier ProLiant servers refer to this option as Power Regulator Disabled.

---

## Power Regulator and Dynamic Power Savings

In Dynamic Power Savings mode, Power Regulator determines the amount of time each processor in the system is spending in the operating system's idle loop. When the ratio of time spent in the idle loop to the time performing useful work is high, the algorithm instructs the processor to set its power state to Pmin (the lowest power and performance mode of the processor). Conversely, when the algorithm detects a low ratio indicating a high application load, the processor is switched in real time to Pmax (the highest power and performance mode of the processor). Each processor in a system is monitored and adjusted independently. Dynamic Power Savings mode allows the processors to operate in a low power state when high processor performance is not needed and in a high power state when high processor performance is needed. Many customers find benefit in simply enabling the Dynamic Power Savings mode and allowing it to run continuously.

When in HP Dynamic Power Savings mode, the Power Regulator algorithm continuously monitors application and processor loading up to eight times a second and adjusts the P-state accordingly. This continuous monitoring results in optimized P-state transitions.

### Measuring CPU utilization

CPU utilization is determined by reading a performance event counter residing within the processor that is programmed to collect NON-HALTED clock cycles. This is an important event to harvest because modern operating systems execute a HLT instruction when idle, instead of spinning in an idle loop looking for work to do. Executing the HLT instruction during idle automatically brings the processor down to a low-power state, and stops incrementing the event counter of NON-HALTED clocks. The event counter is programmed to count kernel and user mode NON-HALTED clock cycles to properly account for processor utilization attributed to system software. Therefore, all processor activity (kernel-mode and user-mode) is accounted for in making power-state decisions. An interrupt pulls the processor out of the HLT instruction and restores the previous power state.

### Processor granularity

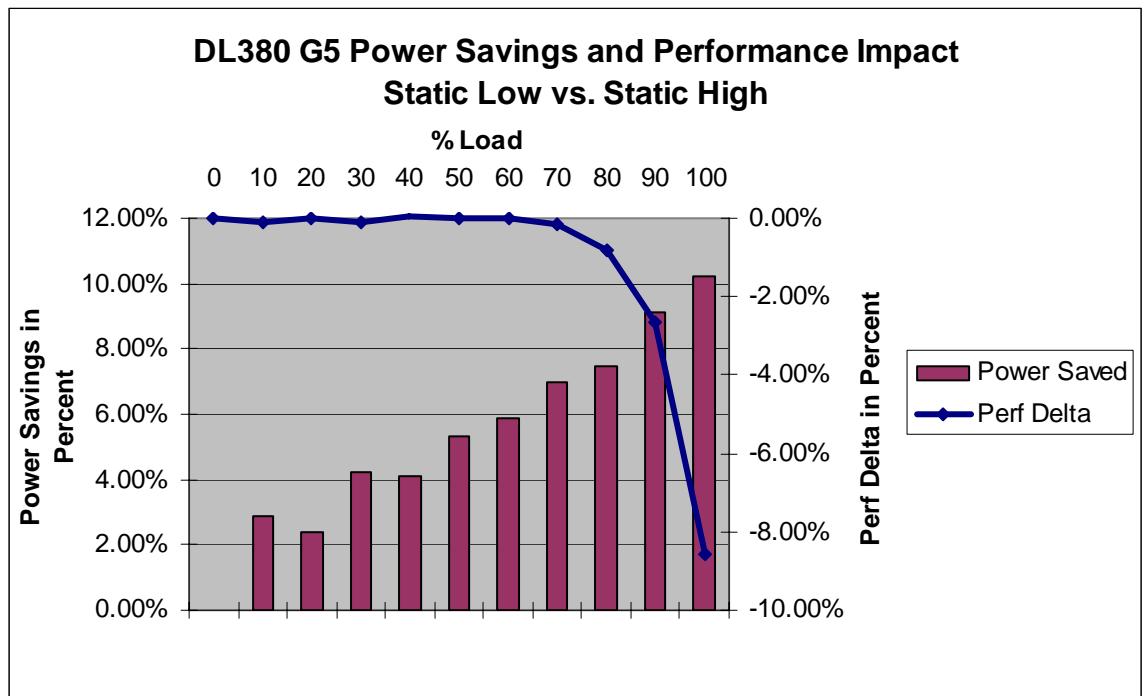
The ROM-based dynamic power algorithm executes on a per-core basis, regardless of processor socket. More specifically it executes on every logical processor, including each core and hyper-thread, and makes requests of the processor hardware to set the P-state of each logical processor based on its current utilization. However, the physical processors may require that some or all logical processors within the processor package be at the same P-state. As a result, while the power algorithm only requests the higher P-state for logical processors that require it, other logical processors within the same physical processor may be made to run at the higher P-state due to these architectural dependencies. This approach also ensures that no process running on the system has its performance adversely affected by the setting of a low P-state on a different logical processor.

## Power consumption and system performance

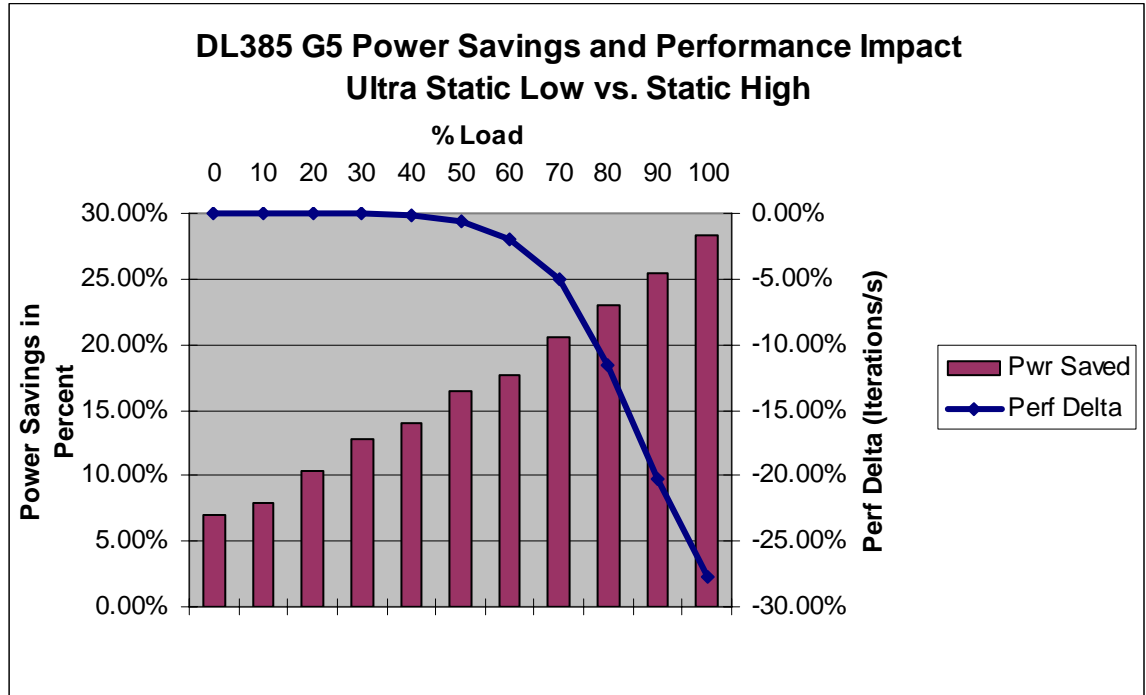
HP Power Regulator gives system administrators options for managing server power consumption and system performance to meet critical business needs. One of the key underlying characteristics of this technology is the ability of processors to operate at lower power levels without sacrificing performance at relatively high levels of application load.

HP engineers have performed extensive tests to characterize performance on systems operating in different Power Regulator modes. As an example, Figures 2 and 3 display the power and performance results for both an HP ProLiant DL380 G5 server and an HP ProLiant DL385 G5 under different levels of application load while in the Static Low Power mode. These charts illustrate that Static Low Power mode can provide significant power savings with little or no degradation in system performance when processor utilization remains below 60 to 70 percent. At greater than 70 percent utilization in the Static Low Power mode, performance begins to drop significantly because holding the processors in low power state limits their frequency.

**Figure 2.** Power savings and performance comparison of Static Low Power mode vs. Static High Performance mode on a DL380 G5 server configured with two Quad-core Intel Xeon E5450 3 GHz processors

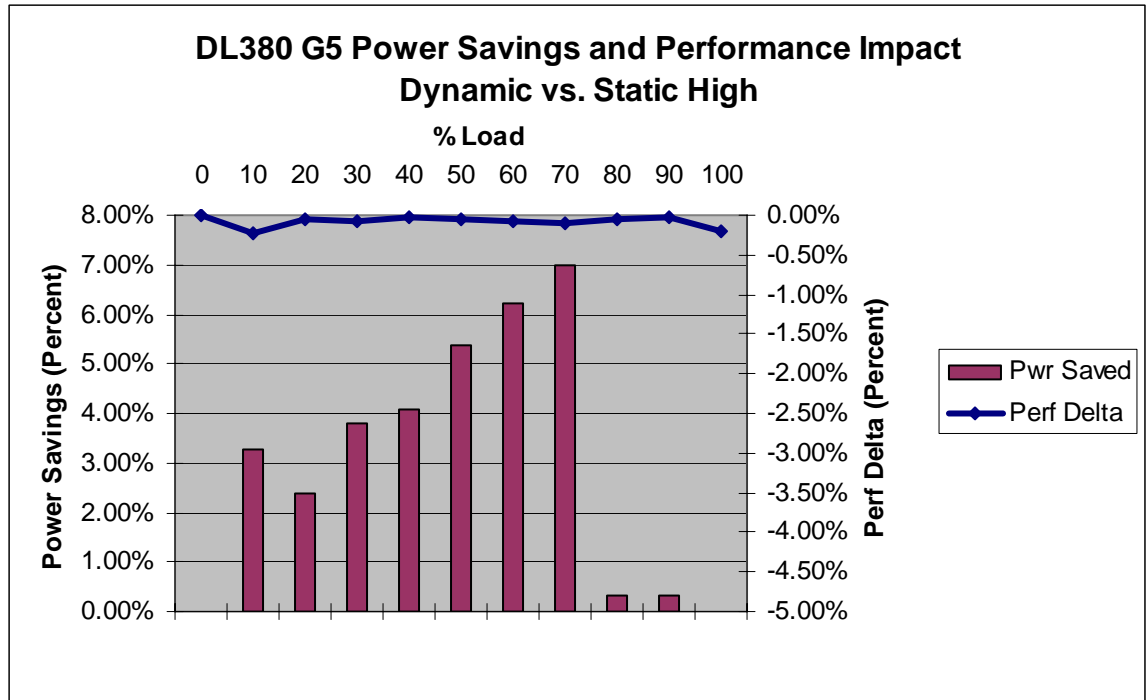


**Figure 3.** Power savings and performance comparison of Static Low Power mode vs. Static High Performance mode on a DL385 G5 server configured with two AMD 2356 2.3 GHz processors

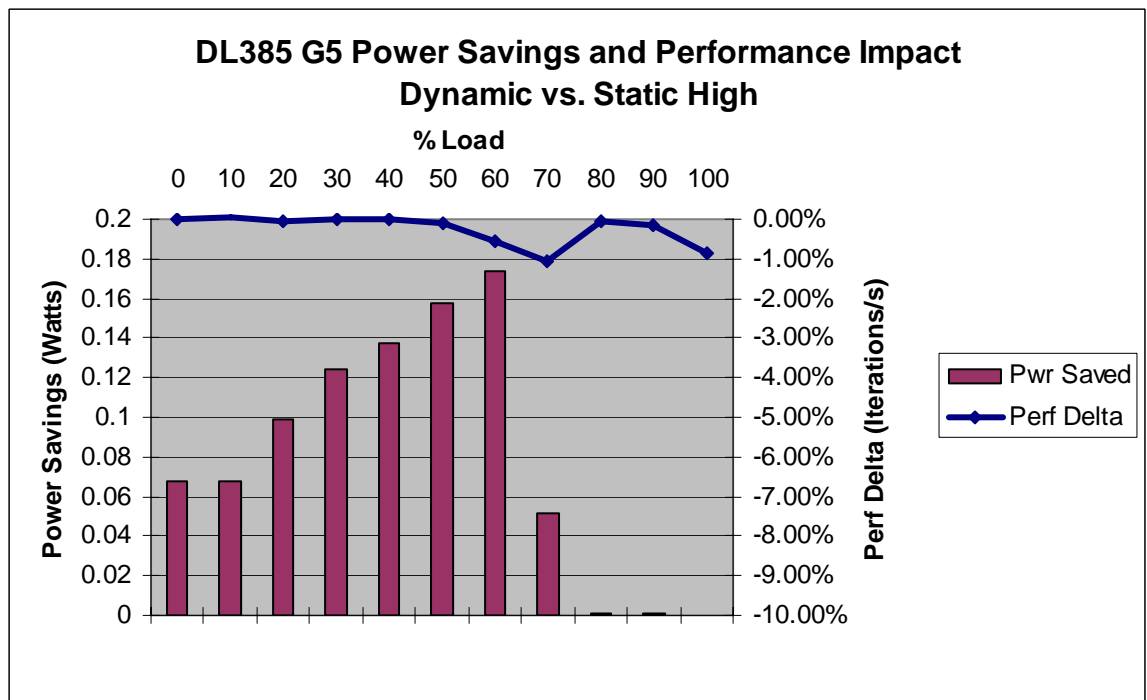


Figures 4 and 5 represent the same power and performance results when Power Regulator is configured for Dynamic Power Savings mode. In Dynamic mode, processors are kept in the lowest (Pmin) P-state when utilization remains below 70 percent, providing the same power savings as in Static Low Power mode. At greater than 70 percent utilization, the Dynamic Power Savings mode begins to run the processors at the higher power level to maintain optimal system performance.

**Figure 4.** Power savings and performance comparison of Dynamic Power Savings modes vs. Static High Performance mode on a DL380 G5 server configured with two Quad-core Intel Xeon E5420 3 GHz processors



**Figure 5.** Power savings and performance comparison of Dynamic Power Savings mode vs. Static High Performance mode on a DL385 G5 server configured with two Quad-core AMD Opteron 2356 2.3 GHz processors



These test results indicate that there is a broad range of conditions under which servers can operate at low power states without affecting system performance. Since most systems operate with average application loads well below 70 percent, many systems should realize power savings without sacrificing performance.

HP Power Regulator results will vary depending on the system configuration, utilization, and application. Power Regulator directly affects only the power consumed by the processors: however, this has secondary effects on memory power consumption as well as on the efficiencies of power supplies and voltage regulators. Thus the total power saved in a particular system is dependent on the number of processors, DIMMs, I/O cards, and drives installed, as well as the processor utilization.

The maximum potential performance loss in the Static Low Power mode is dependent on the application. For example, if the application data set is small enough to fit in the processor cache, the maximum performance loss could approach 33 percent at full load based on a processor with a Pmax of 3.0 GHz and a Pmin of 2.0 GHz. If the data set significantly exceeds the cache size of the processor, for example with Online Transaction Processing (OLTP), web server, and Exchange applications, then overall application performance is governed more by the memory subsystem and less by processor speed. With these memory-bound applications, performance loss could be limited to five percent or less at full load. This is illustrated in Figure 6, which shows the power and performance characteristics for a ProLiant DL380 G5 running a large memory-bound application in both Static High Performance and Static Low Power modes. In this particular case, using Static Low Power mode provides the expected power savings without impacting the overall application performance, even at processor loads above 70 percent.

**Figure 6.** Power savings and performance comparison of Static High Performance vs. Static Low Power modes running a memory bound application on ProLiant DL380 server

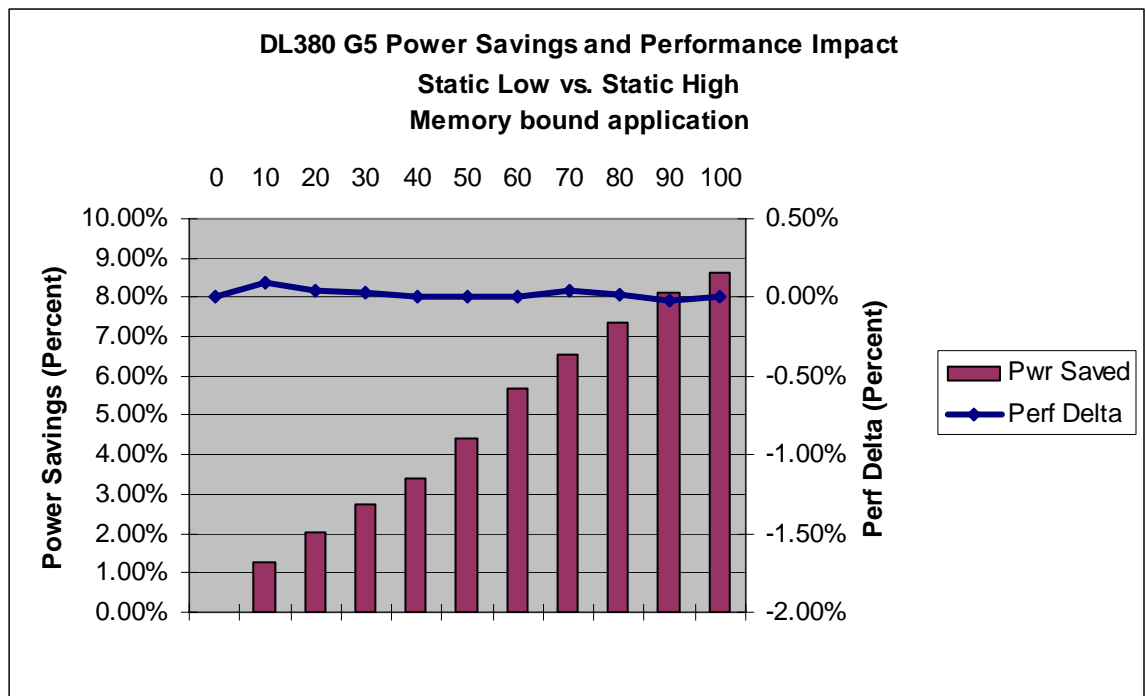
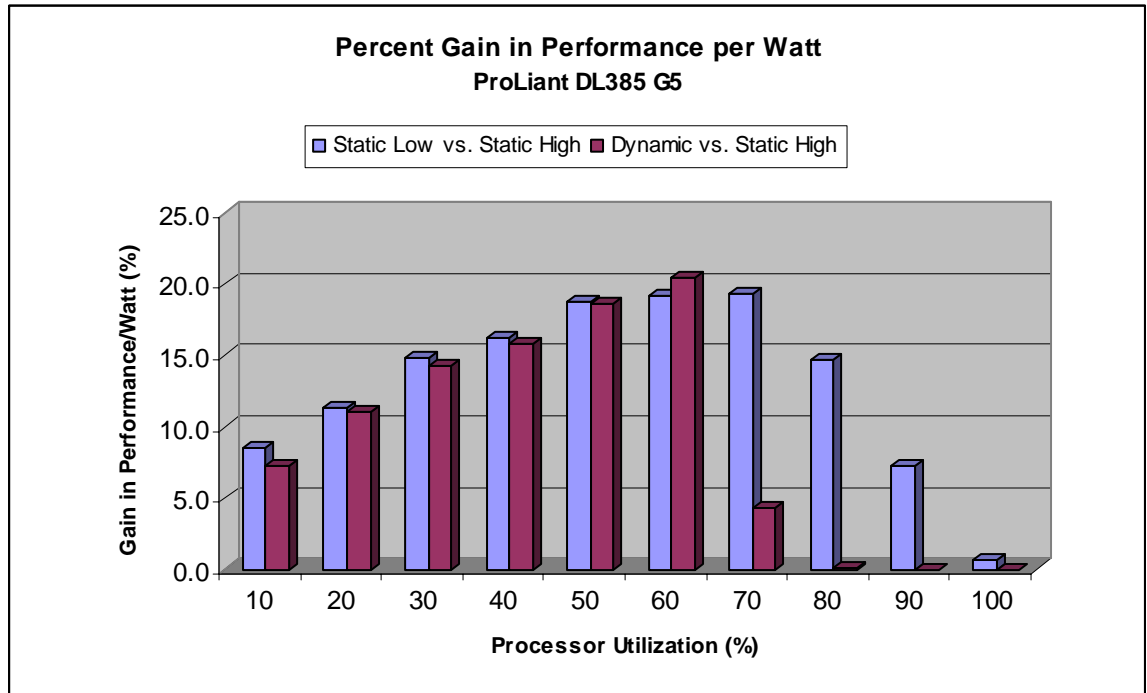


Figure 7 illustrates the effective gains in performance-per-watt for both Static Low Power and Dynamic Power Savings modes when compared to Static High Performance mode running on a single server. In both Dynamic Power Savings and Static Low Power modes the gains in performance per watt

continue through 60 percent processor utilization, indicating that the server is continuing to increase its throughput while the processor remains in Pmin. Beginning at 70 percent processor utilization, the picture begins to change. In Dynamic Power Savings mode, Power Regulator switches the processor to Pmax, losing the gains in performance per watt but allowing the server to meet the increased workload and throughput requirements. In Static Low Power mode, the server remains in Pmin, maintaining higher performance per watt efficiency at the expense of constraining the increase in system throughput.

**Figure 7.** Comparison of the gain in performance/watt of Static Low Power and Dynamic Power Savings modes compared to Static High Performance mode on a DL385 G5 server



## Power management and reporting

HP Insight Power Manager (IPM) and Integrated Lights-Out (iLO and iLO 2) management processors represent two important tools for reporting power management information on HP ProLiant servers. In addition to reporting historical and analytical server power management information, both Insight Power Manager and iLO processors can be used to manage HP Power Regulator settings.

### Insight Power Manager

Insight Power Manager (IPM) is a ProLiant Essentials product that is part of HP Systems Insight Manager. The IPM product provides graphing and historical analysis of key power and thermal data for supported ProLiant servers. IPM can store up to three years of power and thermal data and it supports graphing and analysis for single or multiple server views.

IPM provides three types of graphs and analysis for single-server views:

**Power Consumption graphs** display the power consumed by the server in watts or Btu/hr. Analysis includes an annual estimate of the power and cooling costs for the server. If desired, this information could be used for budgetary purposes or for allocating power costs.

**CPU Performance** graphs display the average utilization of the server and the average CPU frequency of the server (see Figure 8). The average CPU frequency graphs the result of the system moving in and out of different processor P-states. In HP Dynamic Power Saver mode, this graph shows the CPU moving to a higher frequency as the server becomes busier and requires additional performance. Analysis includes an estimate of the power and money saved through the use of HP Power Regulator technology.

**Inlet Air Temperature** graphs display the temperature of the air coming into the server. Most of the power consumed by the server is exhausted as heat into the data center. This graphic display shows the thermal environment in which the server is operating.

IPM also provides four types of graphs and analysis for multiple-server views:

**Group Power Consumption** graphs display in watts or Btu/hr a summary of the power consumed by the collection of servers. Analysis includes an annual estimate of the power and cooling costs for the collection of servers.

**Group Power Consumption Breakdown** graphs display the individual power consumed by each server in the collection. Analysis includes the server with the highest average and peak consumption and the theoretical peak consumption that would occur if all servers encountered peak consumption simultaneously.

**Group Inlet Air Temperature** graphs display the average temperature of the air coming into the servers in the collection. Analysis includes the temperature trend for the collection.

**Group Inlet Air Temperature Breakdown** graphs display the individual temperature of the air coming into each server in the collection. This type of graph is useful for comparing data center temperatures in different parts of the data center or the temperature change for servers in a rack.

Insight Power Manager also supports changing the Power Regulator mode for one to many ProLiant servers. System Administrators can make Power Regulator changes interactively through the web user interface or schedule the changes to occur at specific and recurring times.

For additional information and supported servers go to <http://www.hp.com/go/ipm>

Figure 8. Example Insight Power Manager “CPU Performance” graph



## Integrated Lights-Out (iLO) and Integrated Lights-Out 2 (iLO 2)

Both the iLO and iLO2 management processors can be used to control Power Regulator. The iLO reporting and management tools are limited to individual server management.

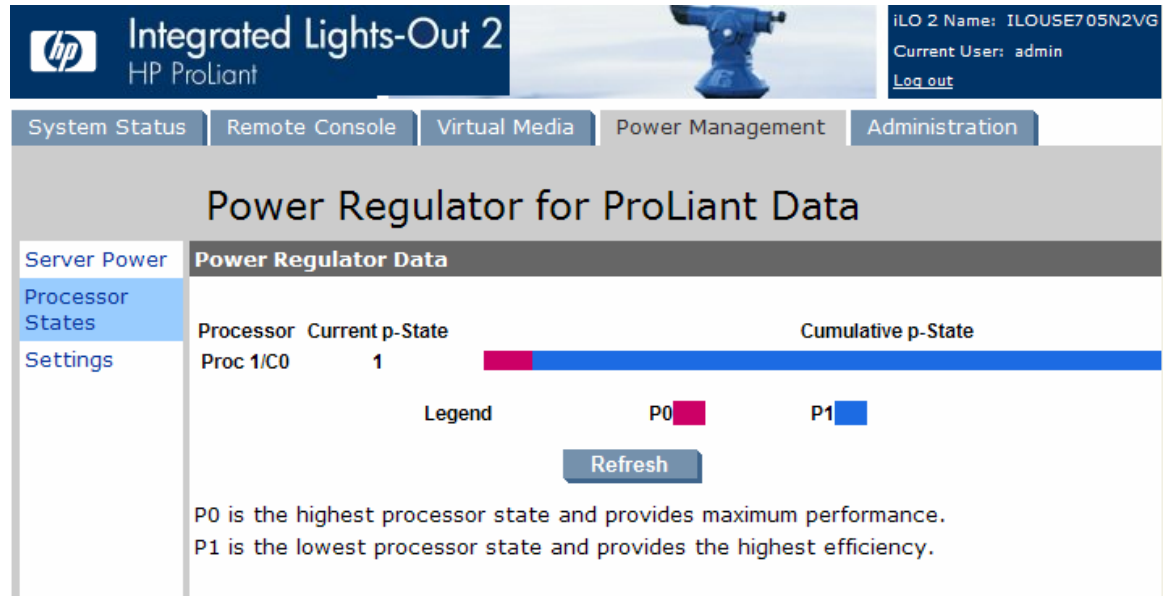
### Management

Power Regulator can be managed using the iLO user interface. Power Regulator features are accessible by accessing the **Power Management** tab of the iLO user interface.

### Reporting

In the Dynamic Power Savings mode, iLO management processors monitor the CPU operation at high and low performance levels, as shown in Figure 9. The iLO processor reports the percentage of time at each level over a 24-hour period. This provides an indication of CPU power efficiency. Results are accessible from the iLO browser, command line, and scripted reports.

**Figure 9.** Lights Out display of Power Regulator processor states



## Typical uses for Power Regulator

Power Regulator is adaptable to different environments. This section identifies three typical reasons for using Power Regulator and the expected quantified results.

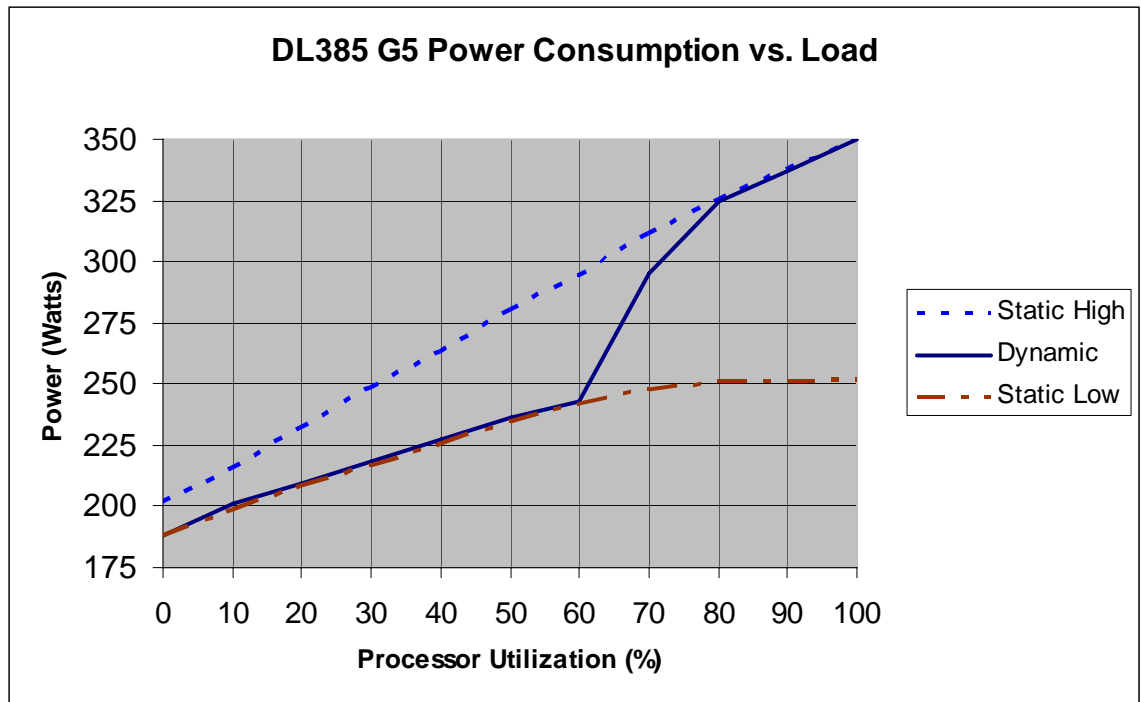
### Reducing power and cooling cost

Power Regulator can be used to improve the server power use for typical client server applications such as database, Exchange, file and print, web server, and OLTP.

Consider the effect of Power Regulator on an HP ProLiant DL385 G5 server running a database with an average processor load of 50 percent. The application involved a mixture of cache, memory, arithmetic, and floating point activities. The G5 system under test included two Quad-Core AMD Opteron 2.30-GHz processors. In Static High Performance mode the processors ran at 2.3 GHz; in Static Low Power mode they ran at 1.15 GHz.

Based on the test results displayed in Figure 10, the power reduction for the DL385 G5 server configuration is 35 watts at an average 50 percent application load. Table 3 shows the estimated cost savings of \$76.65 per year on this system under 50 percent application load in Dynamic Power Savings mode. Actual savings will vary depending on system configuration, processor utilization, and electricity cost.

**Figure 10.** Effective power consumption of Static High Performance, Dynamic Power Savings, and Static Low Power modes on a DL385 G5 server at various system loads



**Table 3.** Calculated annual power cost savings for an HP ProLiant DL385 G5 server running a database with an average processor load of 50 percent

Parameter (at 50% load reduction)	Calculation	Savings
Power saved		35 W
Cooling power saved*	1.5 x heat = 1.5 x 15	52.5 W
Total power reduction	15 + 22.5	87.5 W
Annual energy saved	0.0875 kW/hr x 24 hrs/day x 365 days	766.5 kWh
Dollars saved (at 10 cents per kWh)	766.5 kWh * \$0.10)	\$76.65

\* Calculation based on the Uptime Institute averaging cooling factor of 1.5. For more information, visit the Uptime Institute website at [www.uptime.com/TUIpages/tuihome.html](http://www.uptime.com/TUIpages/tuihome.html).

Since the results above are based on average application and processor load, the \$76.65 cost savings would be realized in either Static Low Power or Dynamic Power Savings mode for an average utilization of 50 percent. The savings should scale linearly with the number of servers, resulting in significant savings for large data center operations. As average utilization decreases, the potential savings will increase.

## Increasing server density and performance per watt

Another way to take advantage of Power Regulator is to reduce power consumption of individual systems by using the Static Low Power mode to release power capacity for additional servers in a rack. This makes it possible to deploy more servers in a rack and thereby increase density and performance in cluster environments. Figure 10 indicates the difference in system power consumption for Static High Performance, Dynamic Power Savings, and Static Low Power modes of Power Regulator over varying loads. The same server configuration used in the previous test was used to gather this data.

At 100 percent load, the system power delta is approximately 100 W less for the [HP Static Low Power Mode](#) than for the other two modes. Running a system in this mode can reduce the maximum system power consumption by 100 W. This reduction in maximum power consumption makes it possible to deploy more servers in a rack and thereby increase overall capacity or performance of clusters.

In a 100 percent load scenario, a rack limited to 5 KW of power would support 14 servers running in Static High Performance mode, as shown in Table 4. By switching to Static Low Power mode and effectively limiting 350 watt servers to 250 watts peak load, rack density can be increased. According to this calculation, an additional six servers could be supported within the 5 KW limit.

**Table 4.** Rack density calculations

<b>Peak rating with maximum no. of servers</b>	<b>No. of servers supported in an 8-KW rack</b>
350 watts per server	14 servers
250 watts per server	20 servers

If the average server application load were less than 60 percent, then adding six more servers could potentially increase the rack performance by almost 40 percent without exceeding the maximum power rating of the rack. This assumes that the application scales linearly when more servers are added. If a server does exceed 70 percent application load, the performance of that server may be limited to less than that of a fully powered server. Based upon the example given with the DL385 G5 at 100 percent utilization, the system would use 71 percent of the power while delivering 73 percent of the performance. The ROM-based Power Regulator solution allows for the power limiting feature while still delivering higher performance per watt within a rack.

## Ensuring uptime

Some customers over-populate a rack relative to its power rating because they assume that all the systems in a rack will not be simultaneously utilized at 100 percent of capacity. However, events can over-tax such systems and cause circuit breakers to trip and bring down all the servers in the rack. By setting some or all systems to Static Low Power mode, administrators can increase system availability by effectively limiting maximum system power consumption with minimal or no effect on non-peak system performance.

## Summary

HP Power Regulator allows IT administrators to directly reduce the frequency and core voltage of the processors. This, in turn, can significantly reduce costs for power and cooling. Administrators can monitor power consumption, keeping track of its use in relation to data center limitations. This capability allows policy management and permits chargeback to business units.

HP Power Regulator for ProLiant is an important technology for any data center facing power and cooling challenges. Flexible power saving modes and OS-independent ROM operation give Power Regulator a broader solution set than OS-based power management. Compatibility with other HP management products such as HP Integrated Lights-Out (iLO), Integrated Lights-Out 2 (iLO 2) and Insight Power Manager (IPM) makes the HP Power Regulator an integral part of a complete enterprise management solution.

## For more information

For more information visit the Power Regulator website at <http://www.hp.com/servers/power-regulator>.

## Call to action

Send comments about this paper to [TechCom@HP.com](mailto:TechCom@HP.com).

© 2007, 2008 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.

AMD and PowerNow are trademarks of Advanced Micro Devices, Inc.

Intel, SpeedStep and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

TC080602B, June 2008

