OpenVMS Technical Journal V14

OpenVMS Power Management

Prithvi Srihari, Burns Fisher and K Veena

Abstract

Power utilization of data centers is becoming a critical factor, often making a crucial impact to the total cost of ownership. It also can make a serious impact on the environment of our increasingly resource-conscious world. In this context, this article focuses on the Power Management facilities provided by HP OpenVMS for Integrity servers version 8.4. While Power Management features are not new to OpenVMS, the strategies employed by various versions differ. This article elucidates these and helps the administrator make a decision as to when and how to use each power setting.

Introduction

Power usage is increasingly becoming a critical issue in modern data centers. This paper provides information on how the Power Management features on the OpenVMS Operating System can be leveraged to reduce power consumption and informs the reader on changes to the Power Management features for HP OpenVMS for Integrity servers Version 8.4.

Rationale for Power Management

While Power Management for computers is not a new problem, it was for a long time confined to computers with very limited power resources, especially those that depend on batteries. Laptops, personal digital assistants, and networked sensors are obvious examples. However, in today's world, the power consumed by data centers is also becoming a critical issue. The motivating factors for these are:

The Increasing Costs of Power Consumption of Servers: The costs of cooling servers—which includes the power costs, the cooling costs as well as the infrastructure costs—has gone up much higher than the costs of the servers themselves as shown in Figure 1 [1]. Server cost has remained constant, though, if anything, they are going down. Infrastructure costs alone already exceeded the cost of the server in 2004. The combined cost of the Infrastructure and Energy (I&E), which includes the cooling needs, exceeded the cost of the server in 2001.

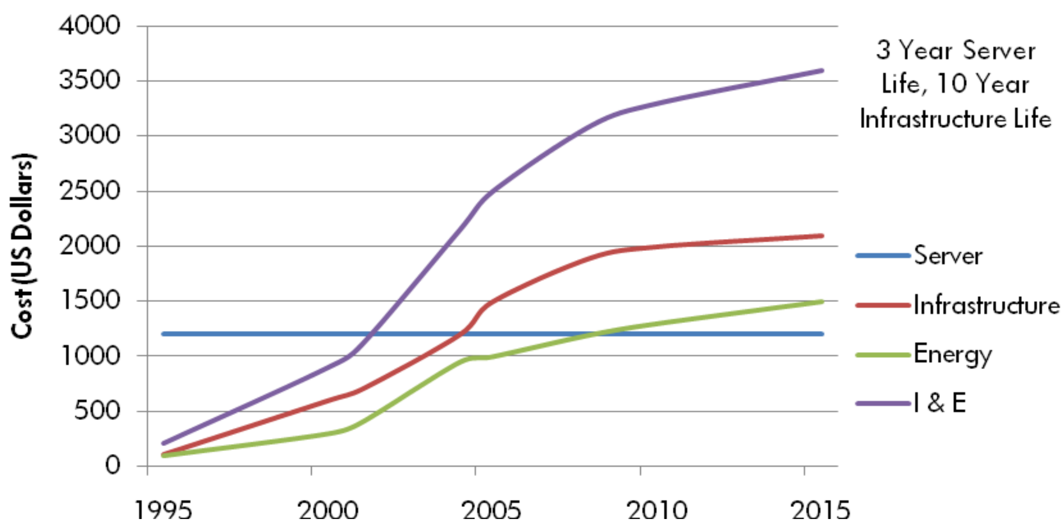

Figure 1: Approximate Annual Amortized Costs in the Data Center for a 1U Server.

The Regulatory Emphasis on Power Efficiency: Governments worldwide are increasingly focusing on encouraging companies to produce energy-efficient products and servers are no exception. Efforts are now on to bring certain classes of servers within the ambit of the US Energy Star program [2]. In future, we may expect customers to actively choose products with better energy efficiency over those that are less energy efficient.

HP's Approach for Power Management—How OpenVMS fits In

HP's Green Business Technology initiative provides leading and innovative capabilities for the data center, across the workplace, and for green IT practices and policies. A part of this initiative is HP's strategy to achieve energy efficiency for the enterprise, which is called 'HP Thermal Logic.' HP provides an extremely broad portfolio of energy-efficient systems, software, and services to help customers build new next-generation data centers or extend the capacity and life of existing ones [3]. HP Thermal Logic's goals are to:

a) **Reduce** total energy use
b) **Reclaim** 'trapped' energy capacity
c) **Extend** the life of the data center

Software	Insight control environment with dynamic power capping: 3x capacity increase
	OS/Application power management
	Insight dynamics – VSE: Forecast future power savings
	Storage thin provisioning/dynamic capacity mgt
Hardware	HP Performance Optimized Datacenter (POD)
	HP BladeSystem
	Power optimized HP servers
	Low power options: Processors, memory, SSD drives: up to half the power consumption
Energy savings from the component to the data center	

Figure 2: HP Thermal Logic

As illustrated in Figure 2, HP Thermal Logic encompasses both software and hardware, since power efficiency and management span a wide spectrum. This paper focuses on the 'OS/Application power management' layer, specific to OpenVMS Power Management. The power usage in a typical data center is composed primarily of the cooling load, followed by the IT (server) load, with UPS power contributing much less, as shown in Figure 3 [4]. Most of the power used by servers turns into heat that must be removed by cooling. So reducing server power consumption serves to reduce cooling power. Consequently, the first step in minimizing the power bill is to use the power management capabilities provided by servers.

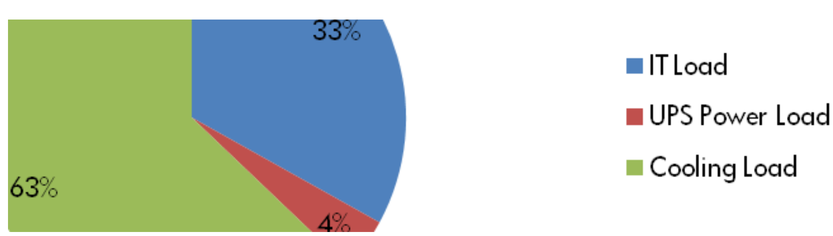

Figure 3: This Chart shows that the Drivers of Power Usage in the Data Center are the IT Load and the Cooling Load.

In the case of OpenVMS, these power management features primarily make use of the power savings features provided at the processor level.

Itanium Processor Power Features

Various Intel Itanium processors implement subsets of the Advanced Configuration and Power Interface (ACPI) Processor Power and Performance States. These are implemented as controllable power states to facilitate power management. The most common of these are the 'LIGHT_HALT' state and the 'Power/Performance States' [5].

The LIGHT_HALT State

This state corresponds to the ACPI C_1 Processor Power State [6]. It reduces power by stopping instruction execution, while maintaining cache and translation lookaside buffer coherence. Effectively, the processor is 'suspended,' not taking part in any scheduled activity. The processor enters this state when the firmware instruction PAL_HALT_LIGHT is called. The processor transitions from this state to the normal state in response to any unmasked interrupt.

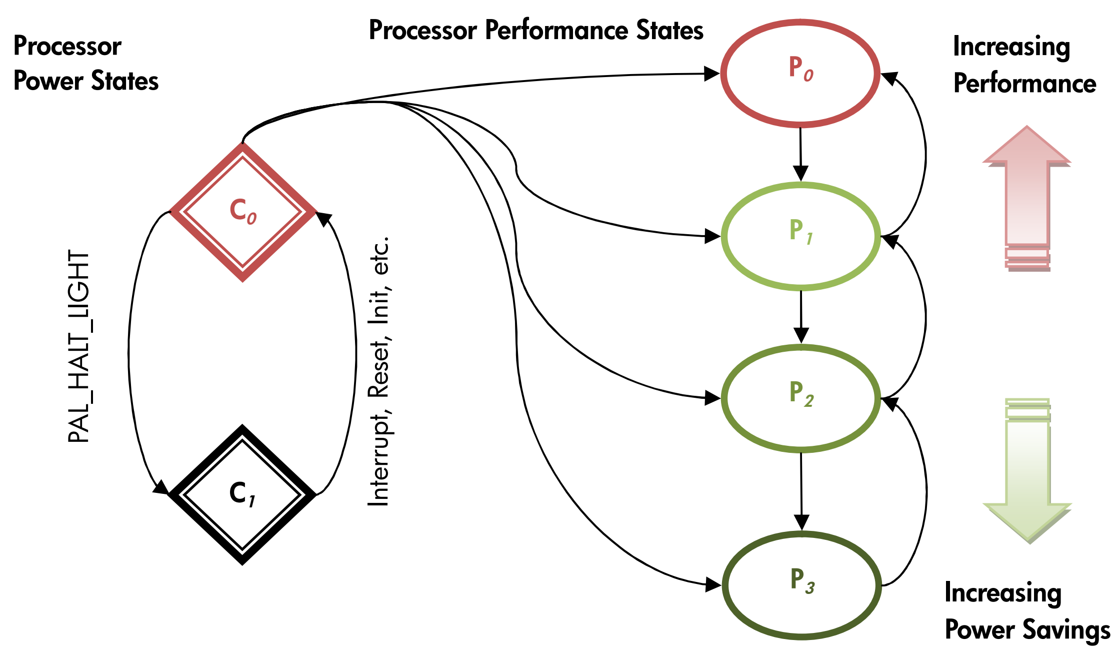

Figure 4: A Simplified View of the Power States of an Itanium Processor with two Power and four Performance States

The Power/Performance States

These states correspond to the ACPI Processor Performance States. The states range from P_0 (maximum performance, least power savings) to P_N (maximum power savings, least performance). The number of performance states available, *N*, varies with the model of the processor. For example, Figure 4 assumes *N* to be 3. As the performance state ranges from P_0, P_1, P_2,..., P_N, the power savings increase and the performance reduces. In all these states, however, the processor is still 'active,' albeit at a slower speed.

HP Integrity Platform Power Options

Processor power states enable power savings at the processor level. At the system level, the administrator can configure any of four modes of operation, described below. These modes are configured from the system firmware and internally control the processor power states.

HP Static High Performance Mode

The system's processors always operate at the highest power and performance state. In this mode, neither the firmware nor the operating system (in our case, OpenVMS) will ever program the processors to run in a lower power or performance state. This mode is useful for business operations when high system computing performance is critical and power savings is not a constraint. Effectively, running the system in this mode is as good as running a system that does not support power management features. For this reason, this mode may also be used as a baseline of power consumption data with power management.

HP Static Low Power Mode

The system's processors operate continuously at the lowest possible power state. In this mode, the processors run continuously in the low-power state to save the maximum power possible, whilst maintaining system usability. This mode is useful for business environments where power availability is constrained. It can also be used in emergency situations when power usage has to be reduced urgently, while maintaining a reduced level of computing operations. However, it might affect the system performance if the workload has high processor utilization.

HP Dynamic Power Savings Mode

Processor power use is adjusted on-the-fly to match performance levels to the application load. There is a slight performance loss when using this state, which is generally small but may be significant with some workloads. HP OpenVMS for Integrity servers version 8.4 has a new algorithm to manage this mode. The new algorithm switches the processor into LIGHT_HALT state when it is idle. Because there is some delay in restarting a processor that is in the LIGHT_HALT state, the algorithm tries to reduce interrupt latency by noting how often the processor is required to leave idle. If that frequency exceeds a threshold, the processor will not be allowed to enter the LIGHT_HALT state until the processor again leaves idle less often. This is to ensure that the response time of the system is quick to a burst of interrupts. Each processor in a system is monitored and adjusted independently. This mode allows the processors to operate responsively when high processor performance is needed and in a low power state otherwise.

OS Control Mode

Dynamic power management for the system is managed by the Operating System through a policy mechanism chosen at the OS level. The power management scheme may vary depending on what the OS that is running on the system implements as its policy. For example, currently HP-UX 11i v3 treats this mode in the same way as HP Dynamic Power Savings Mode [7]. In our specific case of OpenVMS, OS Control Mode allows the OS Administrator methods for configuring the three modes mentioned earlier (Static High Performance, Static Low Power, and Dynamic Power Savings) from OpenVMS. The three sub-modes available under OS Control with OpenVMS are:

OpenVMS Sub-Mode	Corresponding Firmware-Level Mode
OpenVMS Static High Performance	HP Static High Performance
OpenVMS Static Low Power	HP Static Low Power
OpenVMS Dynamic Power Savings	HP Dynamic Power Savings

If no mode is chosen, by default, OpenVMS Dynamic Power Savings Mode is selected. These give the OS Administrator the flexibility to select, at run-time, a mode which suits the particular load being run without having to seek recourse to the firmware interface. Importantly, it allows the Administrator to easily automate mode changes using the standard OpenVMS job automation tools such as application programs, DCL scripts, job queues, etc. For example, it could be useful to keep a system on Dynamic Power Savings Mode during the day to provide reasonable system performance, but switch to Static Low Power Mode at night, when transaction traffic is less, in order to save power. A DCL script could easily be written to automate such a schedule.

OpenVMS Power Management Features—Power Controlling Methods

Configuring the HP Static High Performance, HP Static Low Power, and HP Dynamic Power Savings Modes is done through the system firmware. This can be accomplished using either the Management Processor (MP) Console or the HP Integrated Lights-Out (iLO) Web Interface. Configuring the OpenVMS Static High Performance, OpenVMS Static Low Power, and OpenVMS Dynamic Power Savings Modes is done using an OpenVMS system service or through OpenVMS `sysgen` parameters. In either case, viewing the resulting power consumption is done using the system firmware interfaces (MP Console or iLO Web Interface).

MP Console

At the MP console, the `PM` command (Power Management) is used to switch the power state, as shown in Figure 5. This command requires MP login access.

```
MP: CM> PM
Current System Power Mode: OS Control Mode
Power Regulator Menu:
     D – Dynamic Power Savings Mode
     L – Static Low Power Mode
     H – Static High Performance Mode
     O – OS Control Mode
Enter menu item or [Q] to Quit: L
Power mode will be set to Low Power.
   Confirm? (Y/[N]): Y
   Please wait…
      ➢  Power mode has been successfully changed.
```

Figure 5: Power Setting from the MP Console.

The new mode takes effect immediately, without the need to reboot the system.

iLO Web Interface

The iLO web interface provides the same functionality as the MP console. To change the power state, the following steps have to be followed (in order):

1. From the main menu that appears at the top of the iLO web-page, 'Virtual Devices' has to be selected.
2. The sub-option 'Power Regulator' has to be selected.
3. The four available power modes will appear, as shown in Figure 6.

Power Regulator Mode:
m Enable Dynamic Power Savings Mode
m Enable Static Low Power Mode
m Enable Static High Performance Mode
l Enable OS Control Mode

Submit   Cancel

Figure 6: Power Setting Menu in the iLO Web Interface.

4. On choosing one of the options and then clicking 'Submit,' the change takes effect.

Alternately, the power setting at the firmware level can also be controlled by using HP Intelligent Power Manager (HP IPM).

OpenVMS System Service

After selecting OS Control at the firmware level, the OpenVMS system service, `sys$power_control` can be used to change the power state. This allows the Administrator to change the power state at the OpenVMS level programmatically. The syntax of the system service is:

```
int sys$power_control(
        unsigned __int64 request, unsigned __int64 *previous
);
```

…where `request` is used to choose the mode in the following manner:

Request	Mode Chosen
POWER$C_HIGH_PERF	OpenVMS Static High Performance
POWER$C_LOW_POWER	OpenVMS Static Low Power
POWER$C_EFFICIENCY	OpenVMS Dynamic Power Savings

…and `previous` is a return parameter, providing the power state that was existing before service was called.

This system service returns `SS$_NORMAL` if successful and an error otherwise. It requires `WORLD` privilege.

OpenVMS Sysgen Parameters

OpenVMS also provides dynamic sysgen parameters to modify the power state. These parameters are CPU_POWER_MGMT and CPU_POWER_THRSH. CPU_POWER_MGMT is used primarily to change the power state. Its range of values and their meaning are:

CPU_POWER_MGMT Value	Mode Chosen
0	OpenVMS Static High Performance
1	OpenVMS Static Low Power
2	OpenVMS Dynamic Power Savings

CPU_POWER_THRSH is used only in dynamic mode. It specifies the number of interrupts per 10-millisecond interval beyond which idle power savings will be turned off. The default value is 50. The higher this number, the more power is saved and the higher average interrupt latency the system will experience while processors are idle.

An Example Configuration

The system used is a HP BL870c with 4 cores, each of which is a 1.59 GHz Intel Montvale. The blade was lightly loaded to start with. The system's power was measured in the Static Low Power, Static High Performance, and Dynamic Power Savings Modes with processors being progressively loaded, the non-active cores being idle. The resulting power consumption graph is shown in Figure 7.

One can see the change in power consumption based on the mode chosen and the number of cores loaded. The difference between the power consumed at Low Power and High Performance Modes when all 4 cores are idle is about 11%. This tells us that if a system is found to be idle for long periods of time, it would be a good candidate to be switched to Low Power Mode. On the other hand, a system in which all 4 cores are active, the power gains between the two modes is about 4.5%. Hence, the power saving in this case is much less compelling. Of course, if a system is heavily loaded, its performance requirements may preclude switching it to Low Power Mode.

An interesting comparison is the Dynamic Mode. When system is lightly loaded, this mode provides power savings comparable to Low Power Mode. When the system is in heavy use, it provides very little power savings while giving the performance of High Performance Mode. Dynamic Mode may well be a very good mode to use when load characteristics vary between heavy and light usage (such a day-night scenario).

Finally, it is anticipated that future processors will be capable of more power savings.

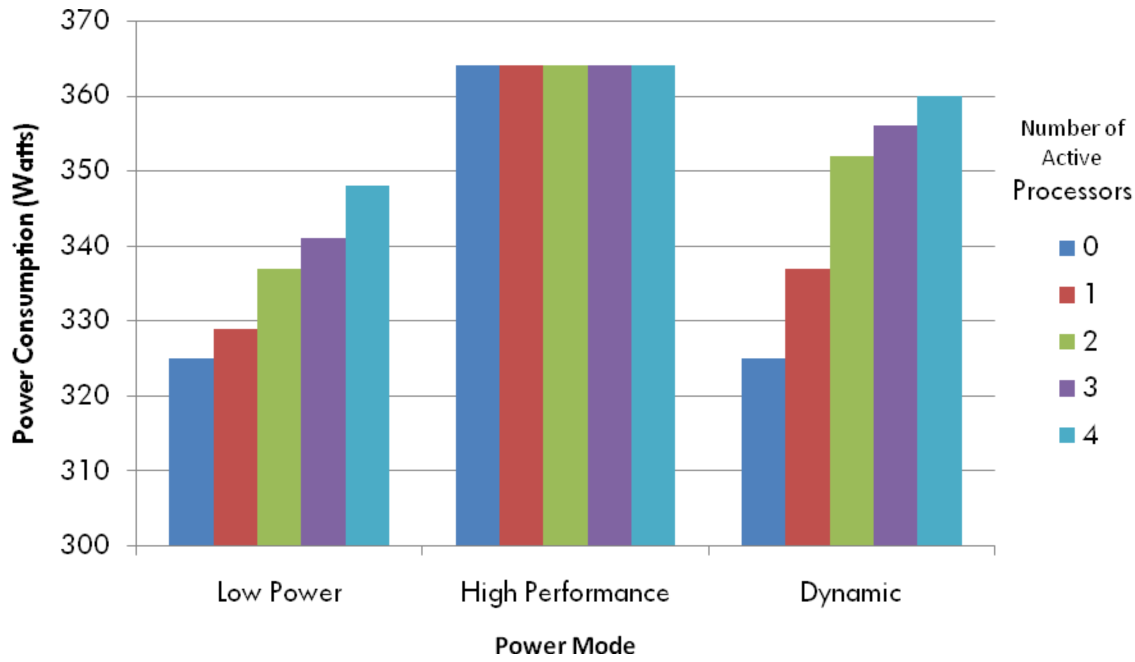

Figure 7: An Example of the Variation of the Power Consumed for each Power Mode

Conclusion

In this article, the need for Power Management was discussed. The various power states provided by HP Integrity servers and HP OpenVMS for Integrity servers version 8.4 to aid in Power Management were explored. Power savings can be controlled at the OpenVMS level and power control can be dynamic. Administrators can configure power management strategies in response to changing power constraints and load profiles. Power control can be achieved using a variety of convenient interfaces. OpenVMS power management features align with HP's overall green initiatives and help customers save money without sacrificing performance while helping to ensure a cleaner and greener world.

References

[1] Belady, Christian L., 'In the data center, power and cooling costs more than the IT equipment it supports,' Electronics Cooling Magazine, February 2007.
[2] United States Environmental Protection Agency, 'ENERGY STAR Program Requirements for Computer Servers Preliminary Draft Version 1.0 Tier 2,' September 24th 2009.
[3] Hewlett-Packard, 'HP Thermal Logic,' Internet: http://www.hp.com/go/integritythermallogic/ accessed November 16th 2009.
[4] Belady, Christian L., Wade Vinson, 'Power-Hungry Data Centers Must Develop a Plan for Curbing Costs, Increasing Efficiency,' Enterprise Networks & Servers, February 2007.
[5] Intel Corporation, 'Intel Itanium Architecture Software Developer's Manual Volume 2: System Architecture Revision 2.2,' January 2006.
[6] Hewlett-Packard, Intel Corporation, Microsoft Corporation, Phoenix Technologies Ltd, Toshiba Corporation, 'Advanced Configuration and Power Interface Specification Revision 3.0b,' October 10th 2006.
[7] Hewlett-Packard, 'Going Green with HP-UX 11i v3–Power Management Techniques to Reduce Costs and Environmental Impacts,' March 2009.