Power and Cooling Innovations in Dell PowerEdge Servers

This technical white paper describes the Dell PowerEdge Energy Smart Architecture and the new and enhanced features designed into Dell PowerEdge 13th generation servers

Dell Enterprise Systems Engineering
March 2016

John Jenne
Craig Klein

A Dell Technical White Paper

Revisions

Date	Description
March 2016	Initial release

Table of contents

Tables

Figures

Executive summary

Dell PowerEdge servers continue to evolve to deliver industry-leading value by reducing power consumption and providing intelligent power management features at the server and data center level. In this white paper, you will learn about enhancements to Dell's Energy Smart Architecture that are integrated into the latest Dell PowerEdge 13th generation servers.

Introduction

Dell's latest generation of PowerEdge servers build upon the Dell Energy Smart Architecture (DESA) introduced in previous generations of PowerEdge servers. DESA took the approach of including high power efficiency and intelligent power management as part of the platform's base features, enabling Dell customers to compute more while consuming less.

DESA features offer high efficiency, intelligent power management capabilities, increased ROI, and new or improved data center usage models. Examples include:

- Power supply right sizing
- IDLE power efficiency
- Circuit breaker power capping
- High accuracy power monitoring

Dell PowerEdge servers continue to evolve by significantly reducing power consumption with each new generation as illustrated in Figure 1.
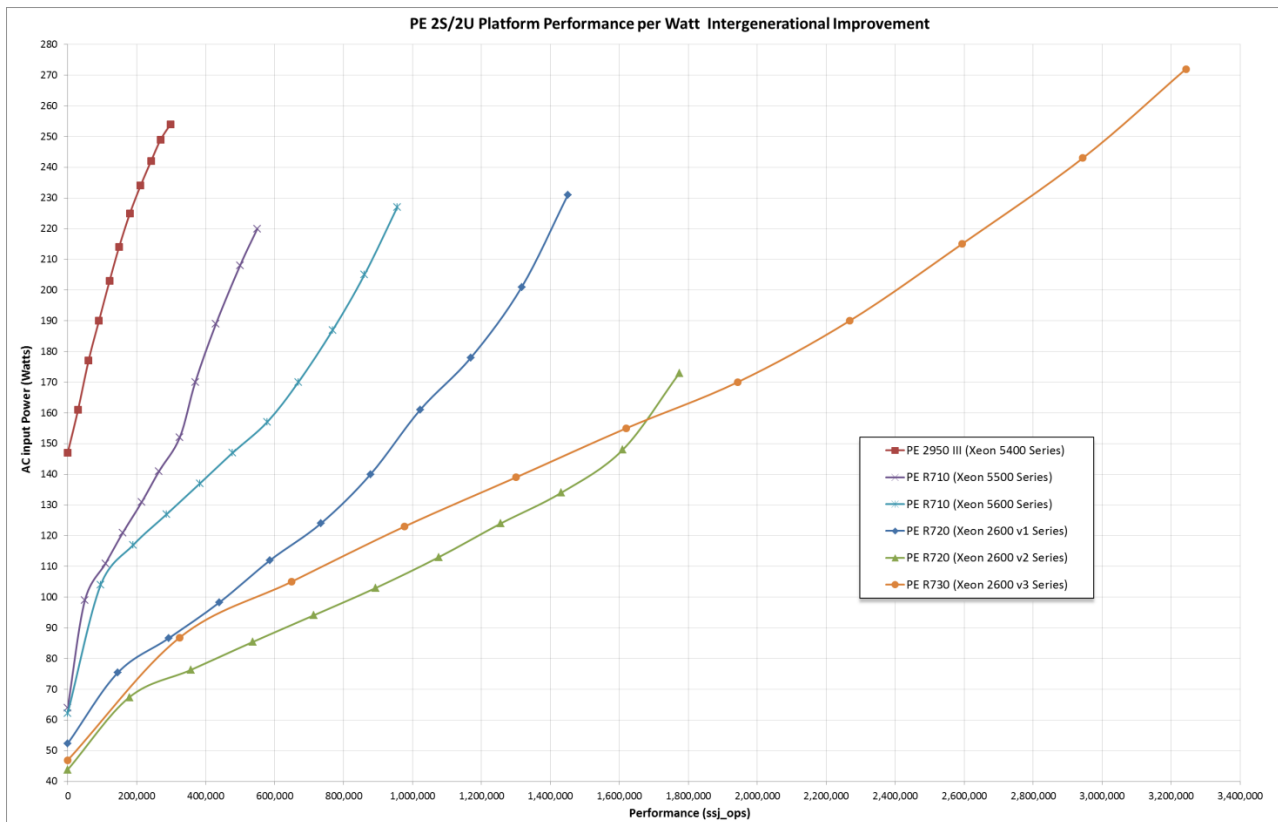

Figure 1   Dell PowerEdge SPECPower evolution

While this white paper focuses on highlighting the DESA features in Dell PowerEdge servers, visit the Dell Power and Cooling Technologies page[1] for additional information on values, strategy, implementation, components, and best practices.

Dell's Energy Smart Architecture

Dell's Energy Smart Architecture is comprised of four core tenets: design, measure, control, and report. New and enhanced Dell Energy Smart technologies have been developed for each of the four core tenets. The majority of these technologies are leveraged across the Dell PowerEdge 13th generation server portfolio (including rack, tower, and blade servers) to provide consistency and compatibility. Table 1 highlights some of the notable enhanced technologies found in Dell PowerEdge 13th generation servers.

Table 1   Design core tenets - Dell Energy Smart technologies

Design		
Power supply • Common form factor • Extended Power Range for CPU transients • High efficiency main and auxiliary power rails • Right-sized for the system • Hot Spare • Smart Power Factor Correction	Voltage regulator (VR) • High efficiency • Switching • Phase shedding • Configuration-based tuning	Components • Low voltage processors • Low voltage DDR4 DIMMs • SSDs • Software RAID option • HDDs with standby
Thermal • High-airflow chassis • Optimized heat sinks • Independent fan control • Pulse width modulation fans • Low power fans	BIOS • Processor P, C, T-states • DDR4 frequency selection • DDR4 CKE and self refresh • Turbo Boost disable	Board design • Low-loss layout • Low-loss materials • Low-loss connectors
Measure		
Temperature • Ambient, exhaust • Components: processor, chipset, DIMMs, hard drives, • PCIe adapters, power supplies	Power • Amperage, BTUs, voltage, and watts • High accuracy • Components: power supplies, processor, memory, storage, I/O, fans	Performance • Processor utilization • Memory throughput • I/O throughput

---

Control		

BIOS
• Enhanced Active Power Controller
• (OS independent process
• power management)
• OS enabled processor power management
• System Profiles
  o Performance
  o Performance per Watt
  o Dense configuration
• Processor core disable
• Host LOM port disable
• External and internal USB port disable
• Unused component disable
	Firmware
	• Data center electrical and thermal power capping
	• Non-linear fan curve
	• Adaptive thermal algorithm
	• Closed loop thermal monitoring and throttling
	• Dynamic DIMM 2x refresh
	• System thermal power capping
	• Management LOM port disable
	• Memory VR phase shedding
		• Power inventory/budget
		• Staggered power on
		• Remote power On/Off
		• Right-size PSU protection
		• Extended Power Range
		• Physical Disk Power Management
		• IDLE memory control

	Dell OpenManage™ • Remote power On/Off	Dell OpenManage Power Center • Group level (rack, aisle, data center) power capping

Report	
Dell iDRAC8 • Node level power reporting: • averages, peaks, real-time • Node level reporting mechanisms include tables and graphs • Node level alarms/alerts, including power cap not maintained	Dell OpenManage Power Center • Provides group level (rack, aisle, data center) reporting • Group level power reporting: averages, peaks, real-time • Group level reporting mechanisms include tables and graphs • Group level alarms/alerts, including power cap not-maintained

Design

With the 12th generation of Dell PowerEdge servers, Dell introduced a power supply (PSU) portfolio with a common form factor. This strategy has been carried over into the next generation, with new and enhanced designs for Dell PowerEdge 13th generation servers. The common PSU form factor lets you helps you select from a rich set of PSU options for capacity, input (AC or DC), or efficiency.

Figure 2 highlights the PSU options that are supported by the majority of the new 13th generation PowerEdge server portfolio. Some value platforms also support a low-cost cabled PSU option.

Platforms

Platforms
R930
R730
R730-XD
R630
T630
R530
T430
R430

Hot-Plug PSUs

Hot-Plug PSUs
495W WRAC Platinum
750W WRAC Platinum
750W HLAC Titanium
750W MixMode Platinum
1100W WRAC Platinum
1100W DC Gold
1600W WRAC Platinum
550W WRAC Platinum

Cabled PSUs

Cabled PSUs
450W WRAC
450W WRAC

Legend:
- ■ Rack
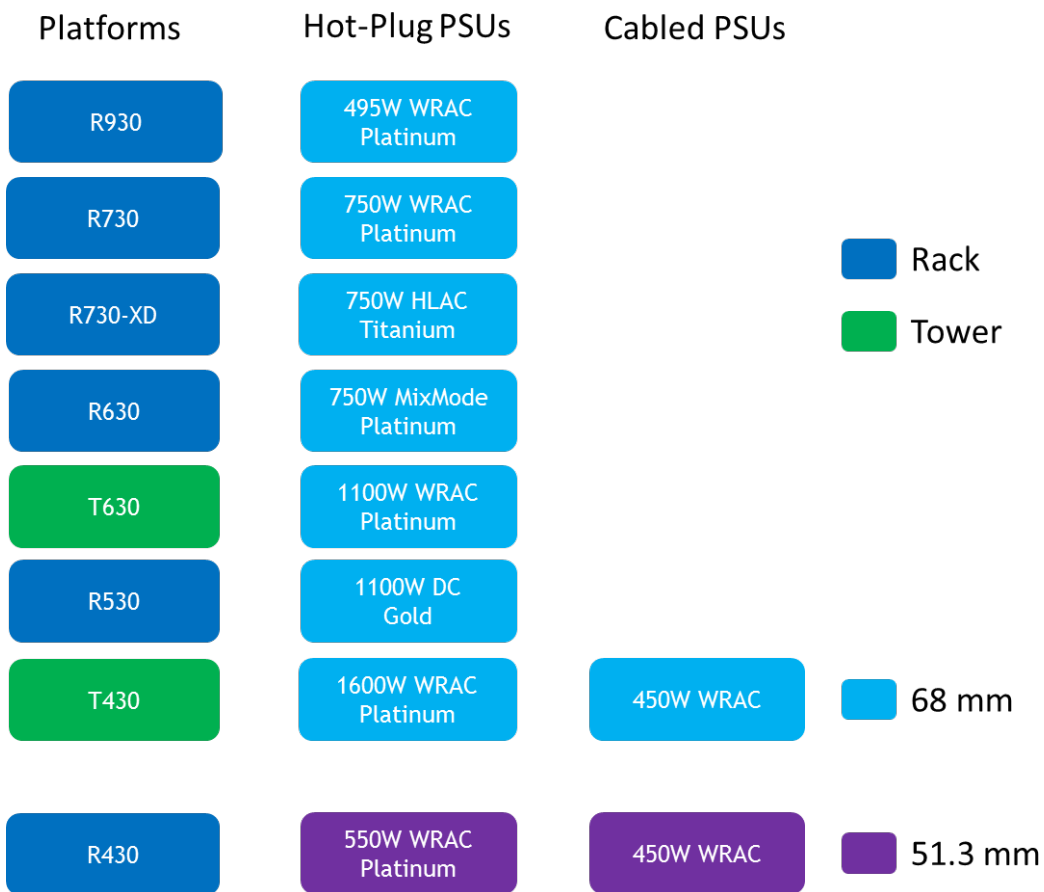- ■ Tower
- ■ 68 mm
- ■ 51.3 mm

Figure 2   PowerEdge PSU portfolio

The PowerEdge PSUs meet 80 PLUS® efficiency requirements for Platinum level with an Energy Smart option to upgrade to Titanium level. Dell PowerEdge servers are equipped with industry-leading PSU efficiencies that meet 80 PLUS's most stringent Titanium efficiency requirements.

Figure 3 highlights the reduced power dissipation benefits of right sizing and high efficiency of PowerEdge 86mm PSUs.
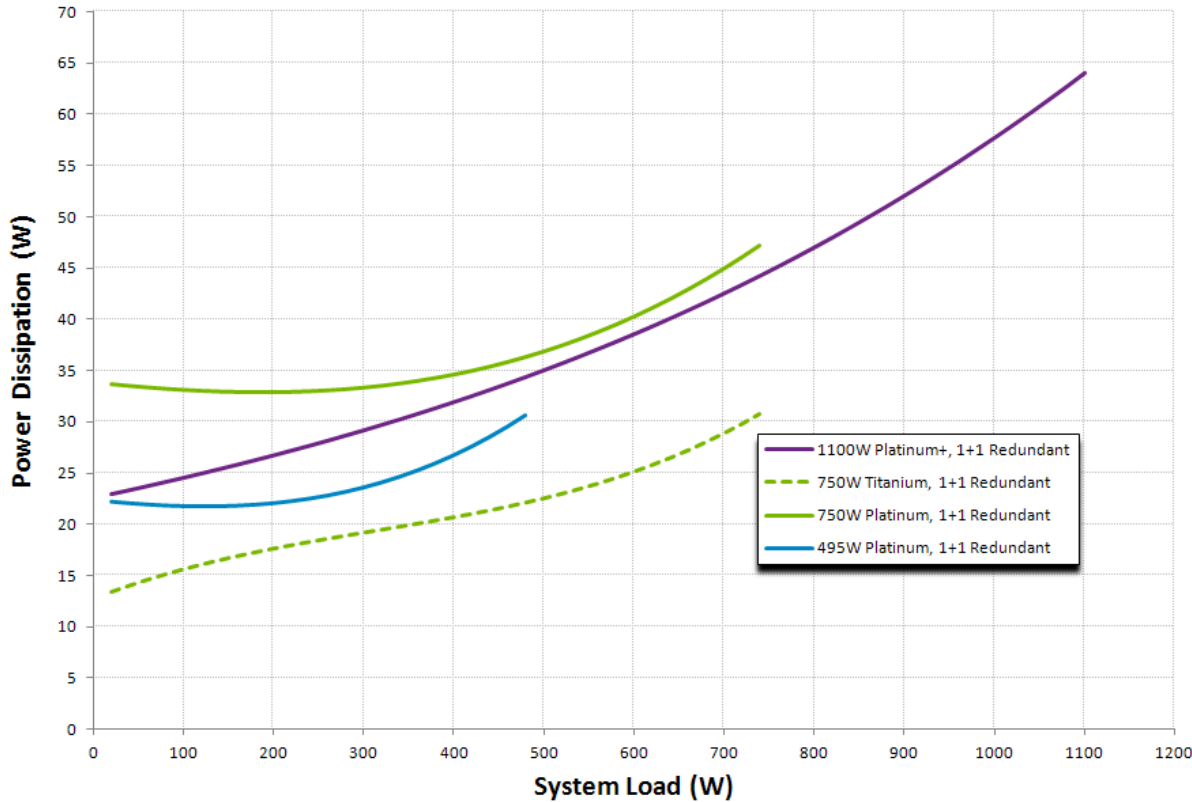

Figure 3   Power supply power loss across system loads

Power delivery efficiency is also improved through voltage regulator (VR) enhancements such as high efficiency design, intelligent phase shedding, and ability to tune operation for a given server configuration.

PowerEdge servers support a software RAID option for storage configurations with four SATA HDDs or less. While hardware RAID support improves performance and robustness, there is a power penalty. Depending on your requirements, software RAID provides an excellent RAID alternative to reduce platform power consumption.

Measure

PowerEdge servers integrate a custom version of Intel® Node Manager firmware, the product of a joint Dell and Intel development effort. Node Manager provides rich power monitoring capabilities, including system-level (PSU) and subsystem-level (processor, memory, I/O, storage, and fan) monitoring. Node Manager also provides sampling rates up to ten samples per second to improve accuracy of averaging and to enable faster response to extreme power transients.
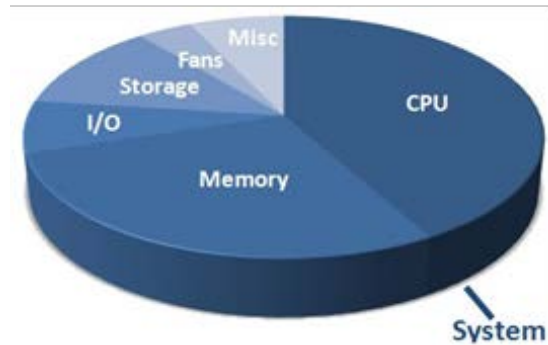

Figure 4   PowerEdge power monitoring capabilities

Power monitoring accuracy on PowerEdge servers exceeds the EPA's ENERGY STAR® certification requirements by 87%, and rivals many external high-cost power meters.
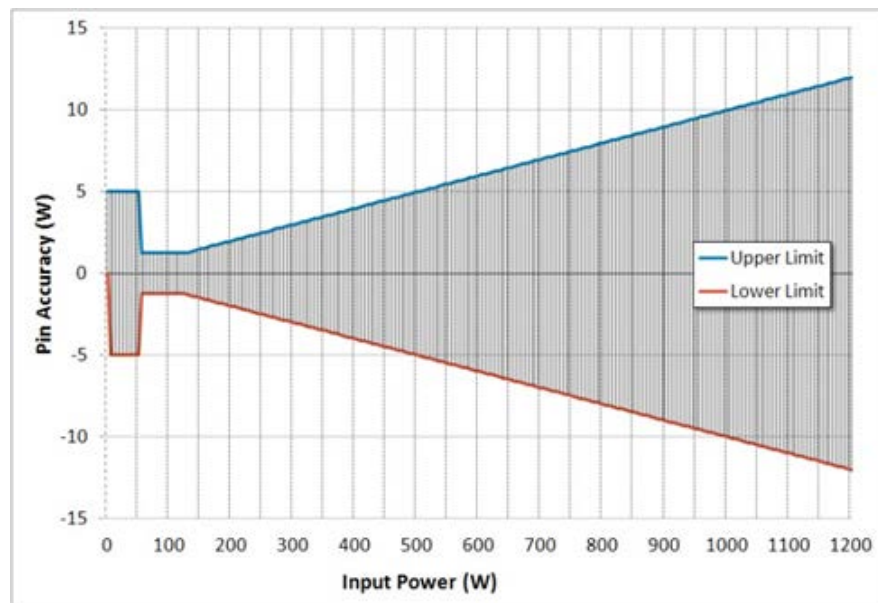

Figure 5   System-level power monitoring accuracy

Thermal monitoring capabilities include exhaust, power control unit (PCU), memory and additional component temperatures such as hard drives and PCIe adapters. The additional thermal monitoring provides pinpoint temperature information to the thermal algorithm for finely tuned control of the system thermals.

Control

Enhanced power control features designed into PowerEdge servers target power-capping, power efficiency, IDLE power, and deployment simplification.

In addition to rich power monitoring capabilities, Dell's custom version of Intel's Node Manager provides sophisticated power capping capabilities to support various data center and platform usage models. Fast and intelligent, the Node Manager power-capping solution limits power utilization below the user-defined limit within one second, while optimizing performance within the power constraints. With PowerEdge servers, you can use power capping to limit server power consumption due to data center or rack electrical or thermal limitations. Examples of other usage models include:

- Control peak power usage during a defined time period to control billing rates
- Provisioning power to workloads based on an SLA that does not require fully configured capability
- Power cap due to power utilities mandating brownout energy consumption reduction
- Limiting power due to extreme outside temperature that limits the data center's cooling capability
- Power cap to protect the data center against power and thermal excursions (such as HVAC failure)

Dell's OpenManage Power Center provides a one-to-many power management console that can be used to set group-level (rack, aisle, data center) power cap policies. This console provides a simple interface for managing the power consumption of PowerEdge servers in the data center.

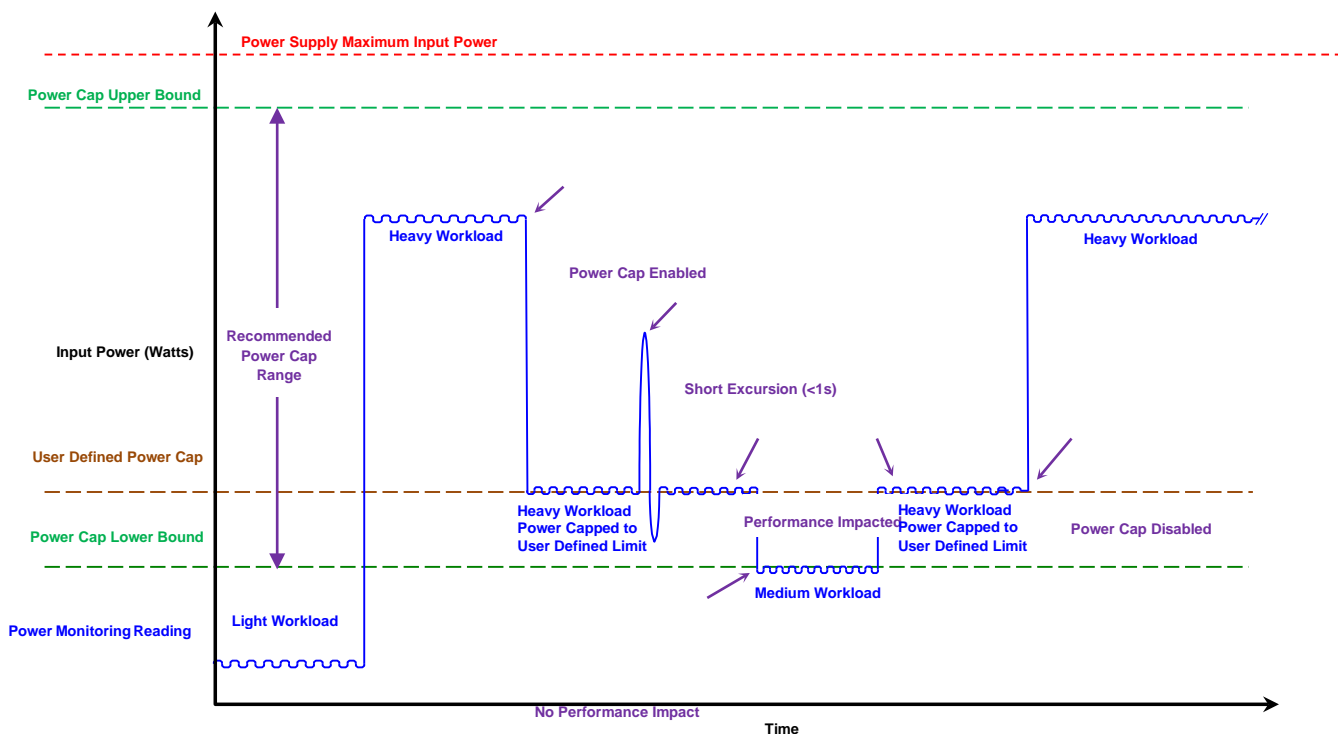The iDRAC8 Enterprise license is required for power capping through the iDRAC8 GUI or OpenManage Power Center.

Figure 6   Power capping example

Dell's PSU right-sizing strategy to improve power efficiency and reduce costs greatly benefits from a PowerEdge server feature called Extended Power Range (EPR) which temporarily extends the output capability of the power supply. Under typical workloads, power consumption is well below the output capability of the PSU, but if the workload spikes power beyond the PSU capability, ultra-fast hardware protection power cap policies are triggered to reduce power consumption. Extended Power Range allows richer system configurations to be supported within the power constraints of a smaller PSU.
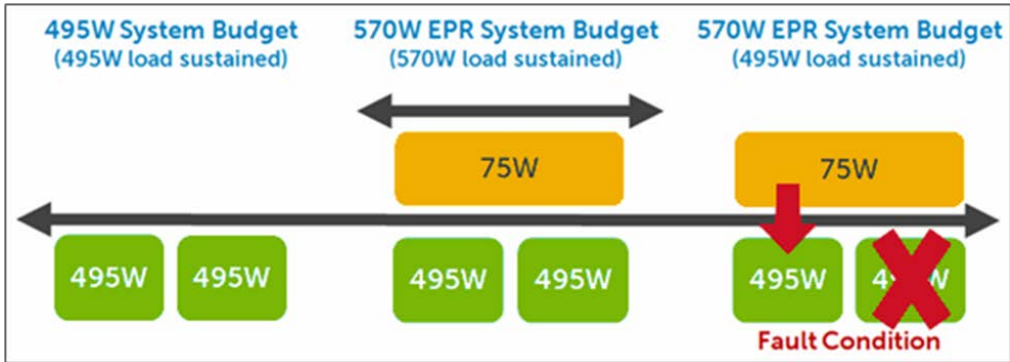

Figure 7   EPR example with redundant (1+1) PSU configuration

The PSU feature Hot Spare improves the operating efficiency of redundant PSU configurations for light to typical workloads by putting one of the PSUs in sleep state. At higher loads, the sleeping PSU autonomously wakes to optimize operating efficiency. Visit the Dell Power and Cooling Technologies page[2] to view a video demo of this feature. Figure 8 highlights the hot spare reduction in PSU power dissipation.

---

[2] Dell Power and Cooling Technologies: http://www.dell.com/en-us/work/learn/power-and-cooling-technologies
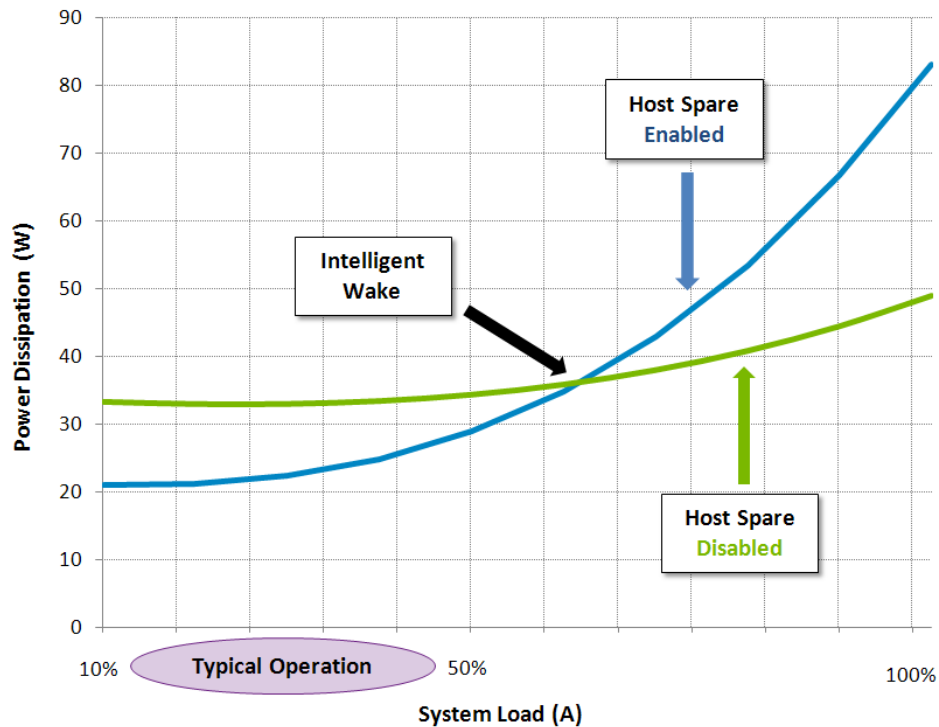
Figure 8   PSU hot spare—redundant PSU power dissipation

Dell's Active Power Controller (DAPC), which provides OS-independent processor power management, has been fine-tuned for next-generation Intel processors. DAPC intelligently manages Intel's Turbo Boost feature for a near-linear performance-per-watt response, compared to non-Turbo Boost levels. DAPC allows workloads to utilize Turbo Boost when needed, providing an optimal balance between performance and power consumption.

Other control features target IDLE power. DIMM IDLE power is minimized by supporting CKE power down and self-refresh features. Storage IDLE power is minimized by the Physical Disk Power Management feature; Dell hardware RAID controllers will spin down IDLE disks. These IDLE power controls complement other power delivery efficiencies at light loads and existing IDLE power controls (for example, P-states or C-states).

System Profiles provide "easy buttons" for configuring a server for performance, performance-per-watt, or dense (conservative) configurations that adjust various BIOS performance, power, and RAS settings. The Dense configuration focuses on prioritizing RAS over performance and power for dense memory configurations.

Available Thermal Profiles complement the System Profiles by providing options for Performance or Performance-per-Watt thermal control algorithms. These thermal controls let you set a maximum exhaust temperature or increase system fan speeds by a fixed offset for additional cooling to I/O slots.

Report

iDRAC8's power reporting in the iDRAC webserver GUI includes a power cap alert function that alerts you when a user-defined power cap cannot be maintained. iDRAC8 also allows you to specify a power cap limit below the recommended range, and lets you know when the platform cannot maintain the specified limit.
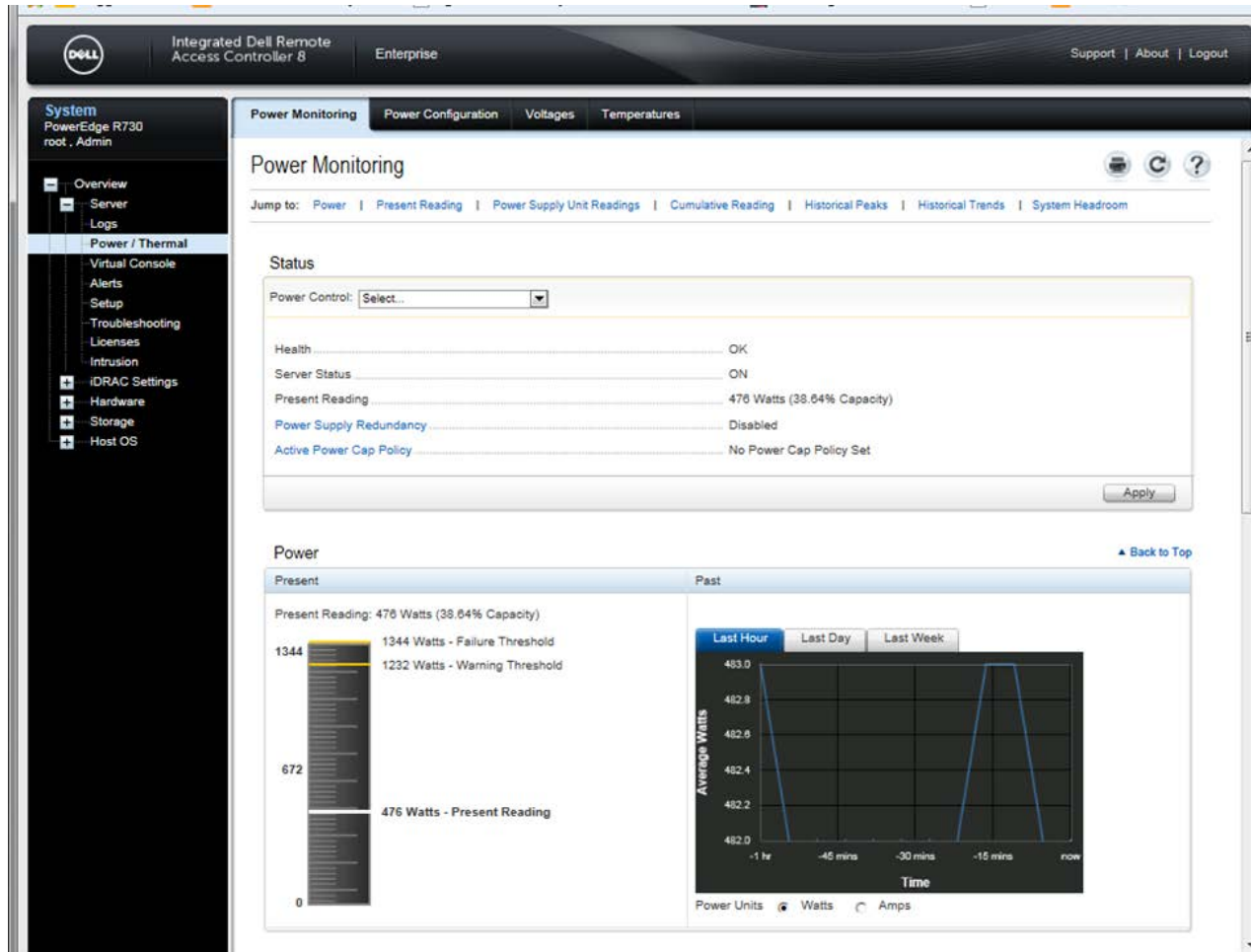
Figure 9   iDRAC8 Power Monitoring GUI

OpenManage Power Center provides powerful group-level power reporting capabilities at the rack, aisle and data center levels. This one-to-many power management console provides a clean, easy-to-use interface for users to track power consumption throughout the data center. Like iDRAC8, Power Center provides an alert function when a group-level power cap cannot be maintained. iDRAC8 Express provides power monitoring capabilities for the iDRAC8 GUI and OpenManage Power Center. An iDRAC8 Enterprise license is required for power capping in the iDRAC8 GUI or in OpenManage Power Center.

Figure 10   OpenManage Power Center power monitoring GUI

Additional Information

For additional details and information, see the following documents and resources available on Dell.com:

- "Power Consumption Reduction: Hot Spare"
- "Power Consumption Reduction: High Efficiency Power Supplies"
- "Increasing Energy Efficiency through Modular Infrastructure"
- Dell.com/PowerandCooling
- Dell.com/PowerCenter
- Dell.com/FreshAir

About the Authors

John Jenne is a Distinguished Engineer in the Dell Enterprise Product Group focused on Node Power across the Dell Server portfolio. John has previously held positions with Compaq, MaXXan Systems, and Newisys. He has a B.S. in Computer Engineering from Clemson University, an M.S.E.E. from the University of Houston, and currently holds 16 patents.

Craig Klein is a Principle Engineer in the Dell Enterprise Product Group working on Node Power across the Dell Server portfolio. Craig has previously held positions with IBM, and Jabil Circuit. He has a B.S. in Electrical Engineering from Texas A&M University, and currently holds 148 patents.