Power and Cooling Innovations in Dell PowerEdge Servers

This technical white paper describes the Dell Energy Smart Architecture and the new and enhanced features designed into Dell PowerEdge 12th generation servers.

Dell
Enterprise Systems Engineering
John Jenne
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 the Dell Energy Smart Architecture that are integrated in the Dell PowerEdge 12th generation servers, and contribute to a 44% improvement over the previous generation of PowerEdge servers in published SPECpower scores.

Contents

Introduction .................................................................................................................................................. 3
Dell Energy Smart Architecture .................................................................................................................. 4
Design ......................................................................................................................................................... 5
Measure ....................................................................................................................................................... 7
Control ......................................................................................................................................................... 8
Report ......................................................................................................................................................... 11
Additional information ............................................................................................................................... 12

This document is for informational purposes only and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

© 2012 Dell Inc. All rights reserved. Dell and its affiliates cannot be responsible for errors or omissions in typography or photography. Dell, the Dell logo, the Dell badge, OpenManage, and PowerEdge are trademarks of Dell Inc. Intel and Xeon are registered trademarks of Intel Corporation in the U.S. and other countries. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell disclaims proprietary interest in the marks and names of others.

March 2012 | Rev 1.0
Introduction

The latest generation of Dell™ PowerEdge™ servers are built upon the Dell Energy Smart Architecture (DESA) introduced in the previous generation of Dell PowerEdge servers. DESA took a new approach by including high-power efficiency and intelligent-power management as part of the platform’s base features, thus enabling Dell customers to compute more while consuming less.¹

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 at the same performance level as illustrated in Figure 1.

![Figure 1. Dell PowerEdge SPECpower evolution](image)

While this white paper focuses on highlighting the DESA features in Dell PowerEdge servers, visit the Dell Power and Cooling Technologies page\(^2\) for additional information on values, strategy, implementation, components, and best practices.

**Dell Energy Smart Architecture**

DESA is comprised of four core tenets: design, measure, control, and report. New and enhanced Dell Energy Smart technologies were developed for each of the four core tenets. The majority of these technologies were leveraged across the PowerEdge12\(^{th}\) generation server portfolio (including rack, tower, and blade servers) to provide consistency and compatibility. Table 1 highlights the new or enhanced technologies (shown in blue text).

### Table 1. Design core tenets—new and enhanced Dell Energy Smart technologies

<table>
<thead>
<tr>
<th>Design</th>
<th>Components</th>
</tr>
</thead>
</table>
| **Power supply** | - Low voltage processors  
- Low voltage DIMMs  
- DIMMs with 4Gb DRAMs  
- DIMMs with 95C capability  
- UDIMMs  
- SSDs  
- Software RAID option  
- HDDs with standby |
| **Voltage regulator (VR)** | - High efficiency  
- Switching  
- Phase shedding  
- Configuration-based tuning |
| **Thermal** | - Low-loss layout  
- Low-loss materials  
- Low-loss connectors |
| **BIOS** | - Processor P, C, T-states  
- DDR3 frequency selection  
- DDR3 CKE and self-refresh  
- FSB power management  
- Turbo Boost disable |
| **Board design** | - Low voltage form factor  
- AC and DC input  
- High efficiency main and auxiliary  
- Right-sized  
- Hot Spare  
- Smart Power Factor Correction  
- Vinyl power supplies  
- Power supplies, processor, memory, storage, I/O, fans |
| **Measure** | - Measurements including amperage, BTUs, voltage, and watts  
- High accuracy |
| **Power** | - 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
  - Performance
  - Performance per Watt
  - Dense configuration
- Processor core disable
- Host LOM port disable
- External and internal USB port disable
- Unused component disables

**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

**Dell OpenManage™**
- Remote power On/Off

**Dell OpenManage Power Center**
- Provides group level (rack, aisle, datacenter) power capping

### Report

**Dell iDRAC 7**
- 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

The most notable design improvement is the introduction of a power supply unit (PSU) portfolio with a common form factor. The common PSU form factor enables you to select from a rich set of PSU options for capacity, input (AC or DC), or efficiency to support your data center needs. Spare inventory can also be streamlined as PSU spares can be shared across server models.

Figure 2 highlights the 86 mm PSU options that are supported by the majority of the new PowerEdge server portfolio. Some value platforms also support a low-cost cabled PSU option.
The new PowerEdge PSUs meet 80 PLUS efficiency requirements for Platinum level with an Energy Smart option to upgrade to Titanium level. Dell PowerEdge servers are the first to market with industry-leading PSU efficiencies that meet 80 PLUS’s most stringent Titanium efficiency requirements.

The PowerEdge M1000e blade chassis enclosure also added a new high-capacity 2700W PSU that supports Platinum level efficiencies. The 2700W PSU refreshes the M1000e chassis to match the high-efficiency PSUs of new PowerEdge rack and tower servers, thus demonstrating the commitment Dell has to delivering high-efficiency power delivery across the PowerEdge server portfolio.

Figure 3 highlights the reduced power dissipation benefits of right sizing and high efficiency of PowerEdge 86mm PSUs.
Power delivery efficiency was also improved through voltage regulator (VR) enhancements such as a high-efficiency design, intelligent phase shedding, and the ability to tune operation for a given server configuration.

The new PowerEdge servers support a software RAID option for storage configurations with four SATA hard drives 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**

New PowerEdge servers integrate a custom version of Intel® Node Manager firmware, which is the product of a joint Dell and Intel development effort. Node Manager provides rich power monitoring capabilities that include monitoring at the system (PSU) and subsystem (processor, memory, I/O, storage, and fan) levels. Node Manager also provides sampling rates up to 10x per second to improve accuracy of averaging and response to power thresholds. Power monitoring accuracy on new PowerEdge servers exceeds the EPA's ENERGY STAR® certification requirements by 87%, and in fact, rivals external high-cost power meters.
Thermal monitoring capabilities were also expanded to include exhaust temperatures and additional component temperatures such as hard drives and PCIe adapters. The additional thermal monitoring provides pinpoint temperature information to the thermal algorithm for fine tune control.

**Control**

New or enhanced control features designed into new PowerEdge servers target power capping, power efficiency, IDLE power, and deployment simplification.

In addition to rich power monitoring capabilities, the Dell custom version of 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 within one second while optimizing performance under the limit. With new PowerEdge servers, you can use power capping to limit server power consumption due to data center/rack electrical (for example, circuit breaker protection) 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 SLA that does not require full 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 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.

iDRAC7 Enterprise is required for power capping through the iDRAC7 GUI or OpenManage Power Center.
The Dell PSU right-sizing strategy to improve power efficiency and reduce costs greatly benefits from a new 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 with a smaller PSU.

Figure 6. Extended power range example with redundant (1+1) PSU configuration
A new PSU feature called 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 PSU intelligently wakes to optimize operating efficiency. Visit the Dell Power and Cooling Technologies page to view a video demo of this feature. Figure 7 highlights the hot spare reduction in PSU power dissipation.

Figure 7. PSU hot spare—redundant PSU power dissipation

The Dell Active Power Controller (DAPC), which provides OS-independent processor power management, has been fine-tuned for next-generation Intel processors. One of the biggest enhancements is to intelligently manage the Intel Turbo Boost feature for a near-linear performance-per-watt response compared to non-Turbo Boost levels. DAPC allows workloads to use Turbo Boost when needed, but keeps a focus on power consumption.

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

---

Power and Cooling Innovations in Dell PowerEdge Servers

System Profiles (formerly known as Power Profiles in the previous generation of PowerEdge servers) provide “easy buttons” for configuring a server for performance, performance-per-watt, or dense configurations that adjust various BIOS performance, power, and RAS settings. Dense configuration is a new profile that focuses on prioritizing RAS over performance and power for dense memory configurations. System Profiles further fine-tune the system by including 3x more settings than the previous server generation.

Thermal Profiles complement System Profiles by providing options for performance or performance-per-watt thermal control algorithms. New thermal controls enable you to select a maximum exhaust temperature or to increase system fan speeds by a fixed offset for additional cooling to I/O slots.

Report

Power reporting in the iDRAC7 GUI has no notable change over the previous generation. iDRAC7 does, however, include a new power cap alert function that alerts you when a user-defined power cap cannot be maintained. iDRAC7 allows you to specify a power cap limit below the recommended range, and then notifies you when the power cap cannot be maintained.

OpenManage Power Center provides powerful group-level (rack, aisle, data center) power reporting capabilities. 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 iDRAC7, Power Center provides an alert function when a group-level power cap cannot be maintained.

iDRAC7 Express provides power monitoring capabilities for iDRAC7 GUI and OpenManage Power Center. iDRAC7 Enterprise is required for power capping in the iDRAC7 GUI or in OpenManage Power Center.
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 author

John Jenne is a Chief Engineer in the Dell Enterprise Product Group focused on Node Power across the 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.