White Paper

Prefailure alerts provided by Dell EMC PowerEdge server systems management

Abstract

Discover the various methods by which OpenManage tools can help provide better server uptime with prefailure alerts.

June 2020

ID 426

Revisions

Date	Description
September 2015	Initial release
June 2019	First Revision
June 2020	Second Revision

Acknowledgments

Authors: Aparna Giri, Damon Earley, Doug Iler, Jeff Krebs, Lori Matthews, Vish Balakrishnan

Contributors: John Abrams

Table of contents

Executive summary

The ability to receive and react to alerts for possible component issues is a critical task for any IT admin. Dell EMC PowerEdge servers provide a wide range of alerts using the integrated Dell Remote Access Controller (iDRAC) and other elements of the OpenManage portfolio. The iDRAC monitors the status of critical subsystems and notifies system administrators about any warning and critical threshold events. With this information, IT administrators can investigate and take corrective action before a component failure occurs, assuring higher server uptime and greater application availability. Often, Dell EMC PowerEdge systems use their embedded intelligence to act automatically. The OpenManage portfolio offers several methods for various IT environments. All the above is part of Dell EMCs comprehensive efforts to use agent-free management technologies to provide intelligent automation.

1 Introduction

The acronym PFA stands for prefailure alert or predictive failure analysis. Originally, PFAs focused on hard drives. The goal was, and still is, to avoid unplanned downtime. Over the years, PFAs have grown beyond hard drives, and now include many other components in the server. This expanded coverage has become increasingly important with the rise of virtualization. Today, there can be multiple virtual servers depending on the underlying physical hardware. This dependency makes taking care of the server more important than ever.

Not all failures are predictable, but the characteristics that are related to gradual mechanical wear and tear can be tracked and monitored. At a certain point, it becomes likely for a failure to occur and iDRAC triggers a warning alert. These alerts and parameters do not address every possible failure mode, but many customers find them useful.

By the end of the 1990s, the various analysis and alerting technologies began to come together under the Self-Monitoring Analysis and Reporting Technology (SMART) standard. Predictive Failure Analysis covers an entire category of predicting impending failures of various components such as drives, memory, and processors. PFA can also mean prefailure alert and predictive failure alert. This paper describes how IT administrators can use information iDRAC provides to best meet the needs of an organization.

1.1 The integrated Dell Remote Access Controller (iDRAC)

The Integrated Dell Remote Access Controller (iDRAC) is designed to make server administrators more productive and improve the overall availability of Dell EMC PowerEdge servers. The iDRAC sends alerts, helps perform remote server management, and reduces the need for physical access to a server.

The iDRAC is part of a larger data center solution that helps keep business-critical applications and workloads available. The technology allows IT administrators to deploy, monitor, manage, configure, update, troubleshoot, and remediate Dell servers from any location, and without the use of software agents. Because the iDRAC is embedded in the server, it can accomplish the above tasks regardless of operating system or hypervisor presence or state.

The iDRAC processor polls each subsystem approximately every five seconds using advanced heuristic algorithms. The iDRAC determines component performance and internal fault conditions that might lead to unscheduled downtime and provides local and remote warnings to IT staff and consoles. By monitoring alerts from the iDRAC, IT administrators can benefit from higher server availability and reduced total cost of ownership.

1.2 iDRAC monitoring and alerting

The iDRAC monitors and alerts the following PowerEdge server subsystems:

- Hard drives and SSDs
- CPU
- System memory
- System temperature
- Fans
- Power Supplies

Monitoring and alerting topics include:

- Health status
- Warning and Failure alerts
- Redundancy Warning and Failure alerts
- Predictive Failure alerts

Actions in response to these alerts include:

- SNMP traps
- Email alerts
- Redfish eventing
- IPMI events
- All events are logged and can be exported manually or remotely using Remote Syslog.
    - From the iDRAC GUI, IT administrators can export the full Lifecycle Controller log and review.
    - iDRAC Enterprise and Datacenter can connect to a Remote Syslog server for consolidation and additional auditing measures.
    - iDRAC9 Datacenter provides an option for telemetry streaming for deeper data analysis. This service sends data to ingress collectors such as Splunk or ELK stack for review.
- Integration into consoles such as Microsoft System Center Operations Manager (SCOM) and VMware vCenter to allow for seamless virtual machine migration.

Actions are user-definable, and a detailed log tracks all alerts that are received. Information from the log includes

- The date and time the alert was received.
- The type and severity associated with the alert.
- The identification of the component that generated the alert.
- Any text that was generated.
- Any actions taken.

As Dell continues to improve its agent-free monitoring and alerting capabilities, a broader range of alerts have become available, reducing the need for software agents. This out-of-band management has the principal benefit of providing critical information, regardless of what operating system or hypervisor, if any, is installed or operational.

The table below provides a high-level overview of the alerts iDRAC provides.

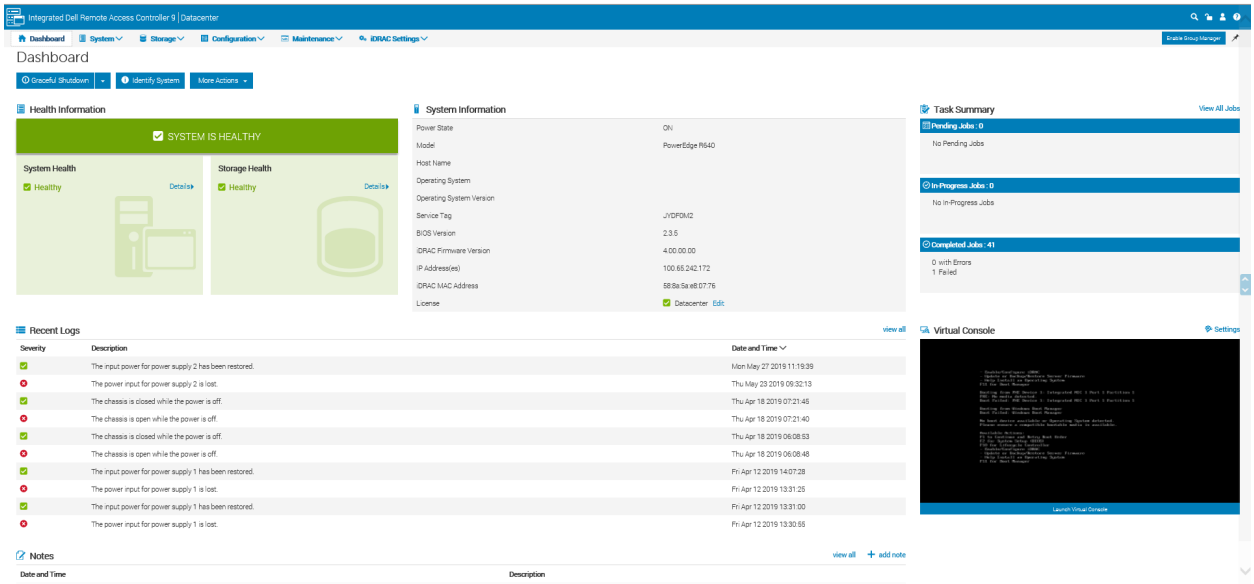
Device	Health status	Warning and Failure alerts	Redundancy Warning and Failure alerts	Predictive Failure Alerts
Drives	✓	✓	✓	✓
CPU	✓	✓	N/A	✓
Memory	✓	✓	✓	✓
Temperature	✓	✓	N/A	N/A
Fans	✓	✓	N/A	N/A
Power supplies	✓	✓	✓	N/A

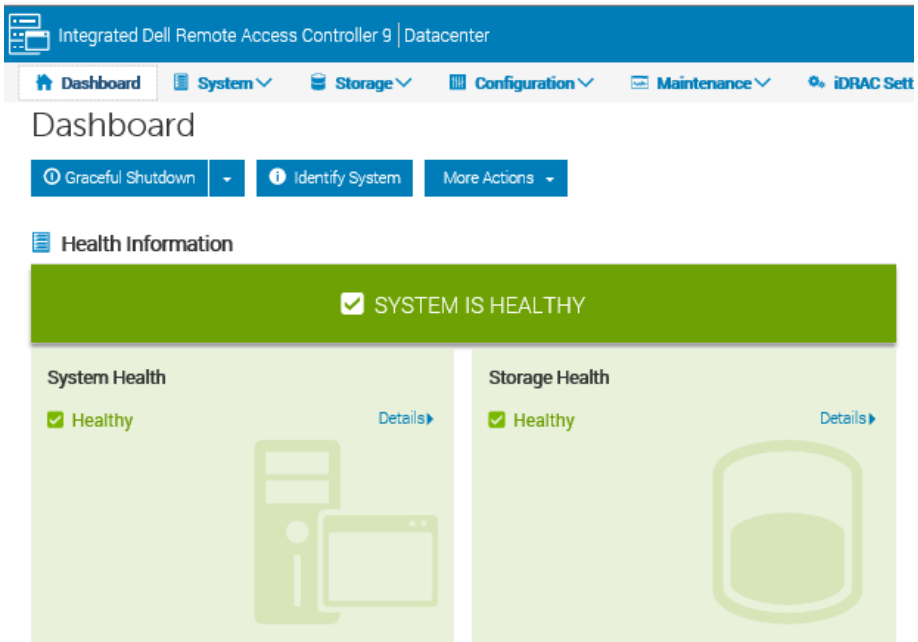
2 Alerts

This section reviews the various devises and alerts in greater detail.

- Drives – Hard Disk Drive and Solid-State Drive
- CPU
- Memory
- Temperature
- Fans
- Power supplies
- GPUs (requires iDRAC9 firmware 4.00 or higher)
- SFP I/O (requires iDRAC9 firmware 4.00 or higher)

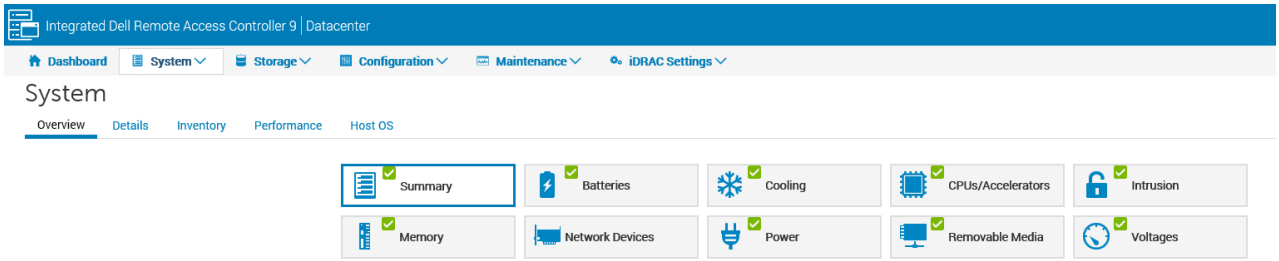
The iDRAC home page, or dashboard, provides a quick view of the health status of the server and storage.





If there were an issue with the server, the ‘details’ link would show the issue as listed in the Lifecycle Controller log.

On the “System” page, an extended view of the various components and status can be seen.



This visual provides an IT admin with quick access to key components. For example, if a warning or critical error happens in “cooling,” the icon would change color. The admin can choose that icon to directly access details to pinpoint and correct a warning or critical alert.

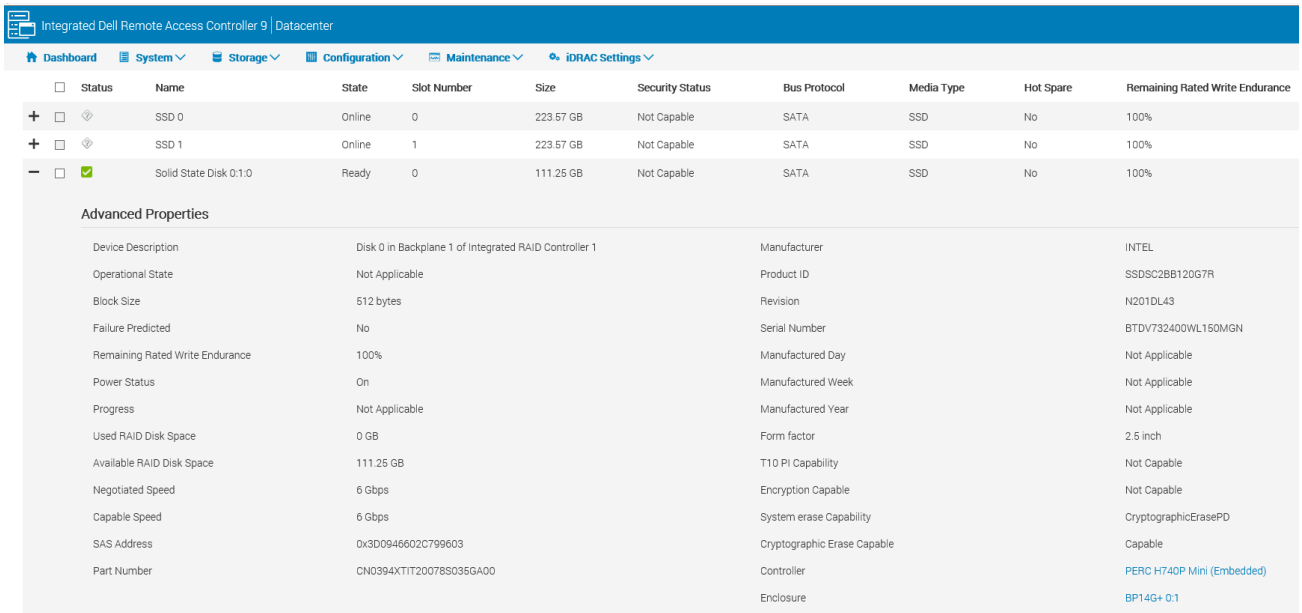
Alerts are not limited to the components mentioned previously. The iDRAC also monitors the following, with no need for a software agent in the operating system or hypervisor:

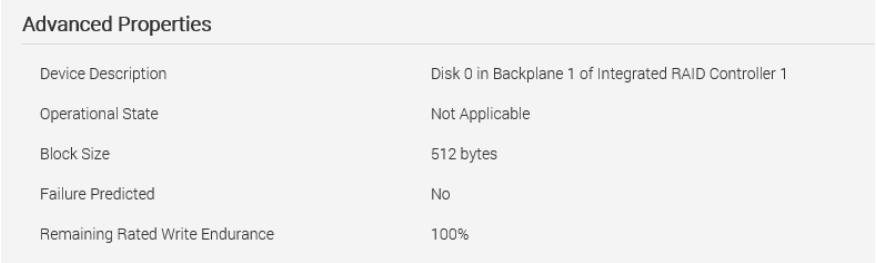
- Network interface cards (NICs)
- Converged network adapters (CNAs)
- PowerEdge RAID controllers (PERCs)
- Chassis intrusion

2.1 Drive alerts

Drive alerts are based on the SMART industry-standard specification for system drives. SMART drives are engineered to provide early warning of certain drive failure indicators. These indicators are meant to give advanced warning of certain types of failures. These warnings do not include defective components, improper handling, or static electricity discharge. However, roughly 60% of drive failures are due to gradual wear and tear. These kinds of failures usually have warning signs that trigger a SMART alert, thus enabling IT administrators to take preventive action.

The iDRAC heuristic algorithms use SMART indicators to monitor drive reliability and data availability. If certain reliability thresholds are exceeded, a predictive failure alert is generated. The iDRAC notes the event in a system log; and, if configured, sends an SNMP trap to a monitoring server. And, if configured with Dell SupportAssist, a trouble ticket is opened automatically, and can dispatch a replacement drive without any human intervention. The iDRAC interface displays a great deal of drive information, including prefailure information and remaining rated write endurance for SSDs.





2.2 System Processor (CPU) alerts

Servers have multiple CPUs, each with multiple cores, and are typically used for virtualization and high-performance applications. As system uptime service level requirements have become increasingly stringent, CPU manufacturing and testing processes have become correspondingly sophisticated. CPU faults are typically unrecoverable errors. If CPU errors occur frequently, certain problems such as L2 cache error corrections can lead to server failure. The iDRAC monitors the number of corrected errors that a CPU reports. If the number of frequencies crosses the heuristic threshold, iDRAC writes an alert to the system log. If configured, iDRAC sends an SNMP trap to a monitoring server. The failing CPU can then be replaced proactively, during a scheduled maintenance window. As with SMART alerts for hard drives, a trouble ticket can be generated, and a replacement part issued proactively.

Special actions can be put in place for servers running a hypervisor from Microsoft or VMware. If configured, the server can go into “maintenance mode” and migrate virtual machines when a CPU alert is received. Detailed information about how alerts work in tandem with virtual machine consoles is covered later in this paper.

Beginning with iDRAC8, PowerEdge servers offer Compute Usage per Second (CUPS) functionality which allows an IT administrator to monitor real-time performance the CPU, memory, and I/O. This data collection operation is independent of operating system and does not consume CPU resources. This out-of-band, real-time monitoring is available by RACADM, Redfish, and the iDRAC web Interface.

An example of CUPS is shown below.

2.3 Memory alerts

With the growing importance of memory in today's compute environment, Dell is taking steps beyond the standard monitoring and alerting on memory errors. In addition to the stand alerts, Dell has pioneered the following solutions: Memory Page Retire and Fault Resilient Memory.

2.3.1 Memory Page Retire

All Dell EMC servers ship standard with Error-Correcting Code (ECC), a first line of defense on errors in system memory. ECC looks for single-bit errors in memory and automatically corrects them, keeping the system running smoothly. Most ECC corrected errors are not isolated. Addresses that experience error corrections once tend to experience them again. If these errors cascade into nearby bits, it can expand beyond what ECC can deal with. In turn, the operating system processes the uncorrectable error and fails the system. With the introduction of iDRAC7 based systems, Dell EMC worked with hypervisor partners Microsoft and VMware to introduce Memory Page Retire (MPR).

The basic flow of MPR is:

1. The hypervisor monitors baseline ECC memory faults.
2. Should certain regions produce recoverable errors beyond a certain threshold, the section, or page, is retired.
3. After 64 Kb of page retires have occurred, the event is logged in the system event log.
4. The address and adjoining space is mapped off and unavailable to the hypervisor.
5. The defective memory can be replaced during scheduled service time.

Memory Page Retire is supported on Microsoft Windows 2012 R2 and VMware ESXi 5.1 U1 and beyond.

2.3.2 Fault Resilient Memory

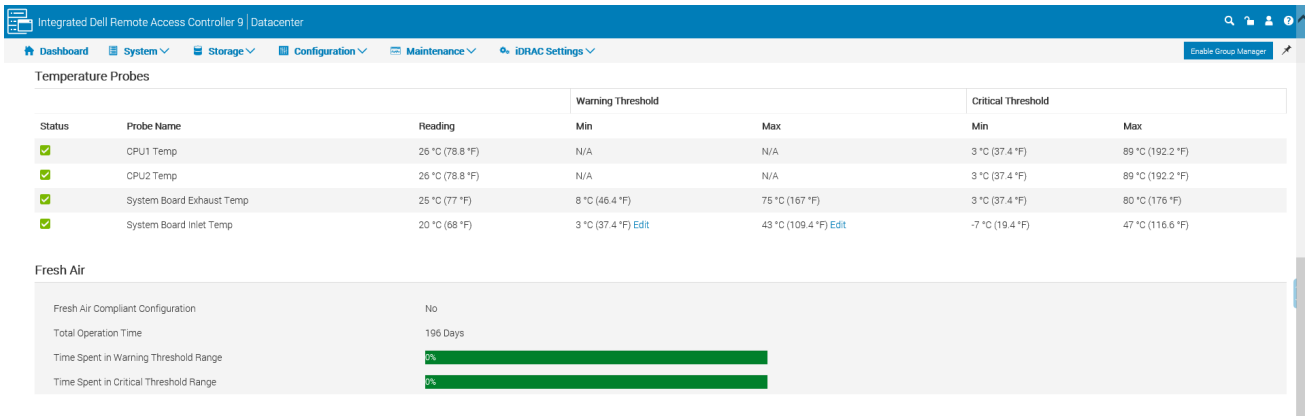
Within a virtualization environment, the hypervisor is the brain that sits below the virtual machines, controlling the server resources and distributing them as needed. Hypervisors are exposed to uncorrectable memory errors like any other operating system. However, if a hypervisor fails, they generally bring down more than one application. Fault Resilient Memory (FRM) is a patented technology Dell EMC has introduced aimed at creating more resilient memory protection for the hypervisor.

FRM works with VMware vSphere v5.5 and higher, which uses its Reliable Memory feature to work with FRM.

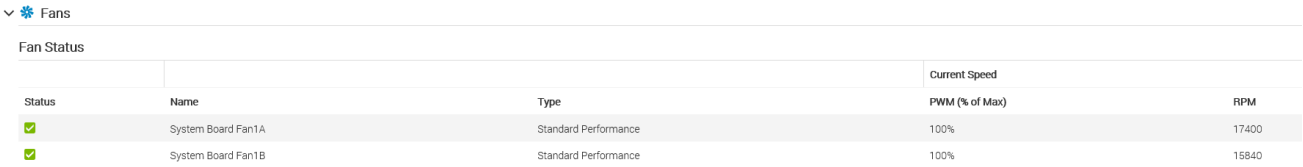
FRM creates a fault-resilient memory zone for the hypervisor within socket 0 and communicates that address up for the hypervisor to place itself into. The ESXi hypervisor in vSphere v5.5 or higher looks for this address communication, and if found, places itself in the protected zone. The protection FRM provides is as robust as Memory Mirroring. An uncorrectable error that occurs in socket 0 is logged as a System Event, without requiring a full 50% of system memory. This log event gives administrators time to become aware of the issue. They can place the system in Maintenance Mode to clear off running VMs, and then deal with the memory module showing errors.

2.4 Temperature and fan alerts

Temperature alerts provide advanced warning that either the ambient temperature is at or exceeding preset temperature ranges. Dell offers a wide array of alerts and other technologies that help monitor and manage temperature alerts. Temperatures are monitored at server CPU and at the system board inlet. Fans and blowers are well placed within the chassis to provide maximum cooling.



Fan alerts



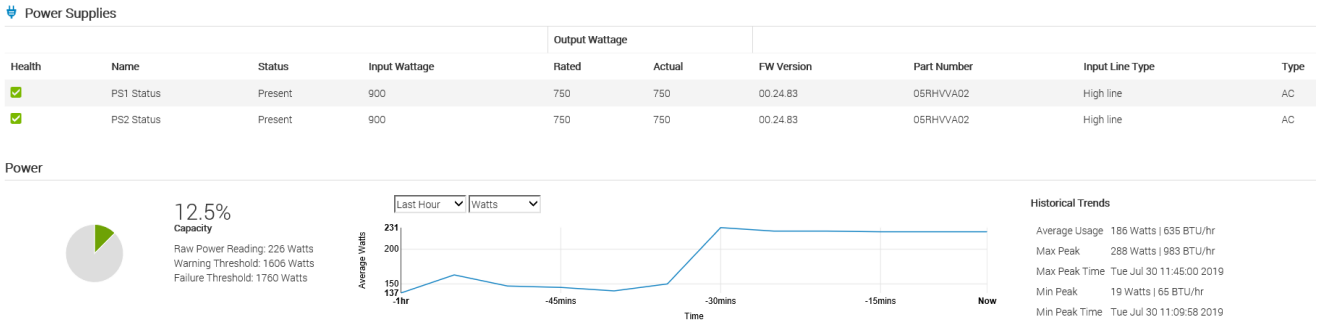
Fan status, warnings, and critical thresholds are simple and easy to read. System fan alerts provide warning that a fan may soon fail. As shown in the figure above, iDRAC monitors the system fans for circumstances when RPM speeds fall below, or exceed, certain thresholds. If a single fan has failed, PowerEdge servers are designed with redundant cooling fans, so they can continue operating in a safe temperature range.

The combination of redundant cooling fans and proactive alerts results in a well-designed solution that provides protection against failures due to overheating.

2.5 Power Supply alerts

Power-conditioning uninterruptible power supplies are a highly cost-effective step in defending servers from electrical dangers to their sophisticated and delicate electronic circuits. Once power going to servers has been conditioned properly, the next critical server subsystems to protect are their power supplies. Dell EMC PowerEdge servers are designed to offer redundant power supplies.

As shown in the figure below, iDRAC monitors power supply fan functionality and for voltage variances from preset upper and lower thresholds. If either fan performance or voltage readings are out of tolerance, iDRAC generates an entry in the log and send the alert. As with CPUs, this type of alert is often visible on properly configured hypervisor consoles and can trigger an automatic live migration of VMs, if necessary.



3 Beyond Alerts – policy-based actions

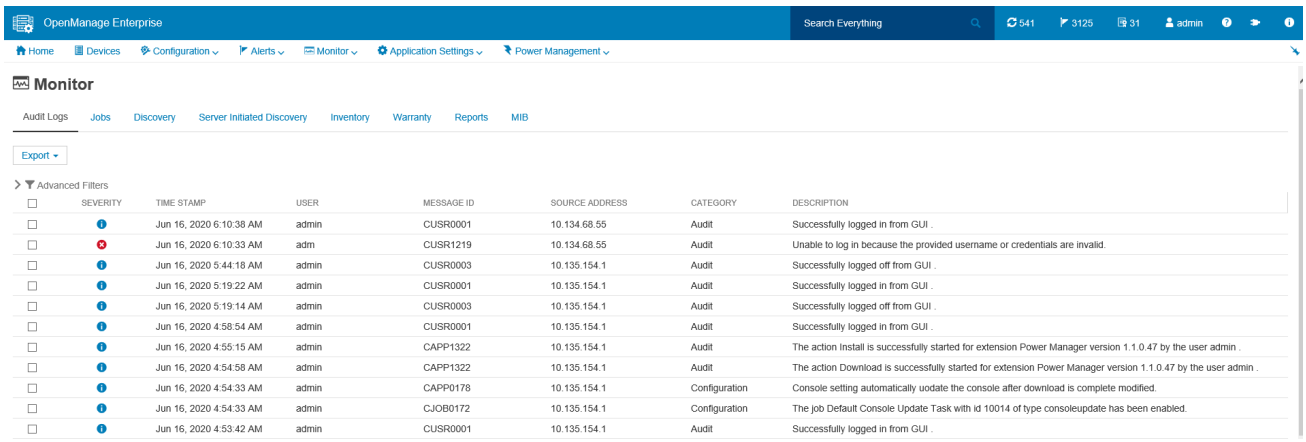
Getting an alert is one thing; acting on it is another. Some companies still have IT staff patrol the aisles of their data center, taking notes of flashing or amber lights. This process is a time-consuming task with the potential for overlooked or missed information.

For customers with a smaller IT shop with a few servers, the emailed alert option from iDRAC can be an effective solution. An iDRAC email alert is both informative and actionable. It provides the server name, service tag, an exact, easy-to-understand description of the error and the error message ID.

But for larger data centers, a more comprehensive solution is required to ensure proper uptime. This section discusses how tools such as OpenManage Enterprise, VMware vCenter, and Microsoft System Center use and consume alerts.

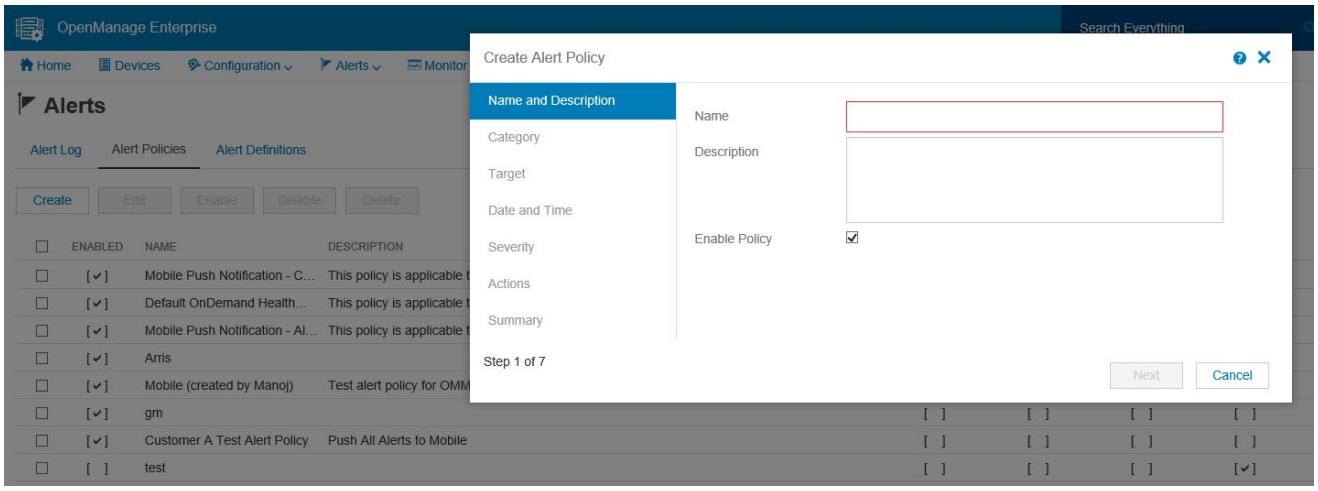
3.1 Alerts and OpenManage Enterprise

For IT shops of all sizes, Dell EMC offers the OpenManage Enterprise console. OpenManage Enterprise is a “one to many” console that can monitor and manage 8000+ PowerEdge servers. Also, OpenManage Enterprise can be set up to receive SNMP traps from almost any servers, storage, or networking device using the MIB importer. OpenManage Enterprise also allows users to set up a policy to trigger an action when the alert is received. IT administrators can also write a script to automate the actions they want to take after receiving certain alerts.



SEVERITY	TIME STAMP	USER	MESSAGE ID	SOURCE ADDRESS	CATEGORY	DESCRIPTION
	Jun 16, 2020 6:10:38 AM	admin	CUSR0001	10.134.68.55	Audit	Successfully logged in from GUI .
	Jun 16, 2020 6:10:33 AM	adm	CUSR1219	10.134.68.55	Audit	Unable to log in because the provided username or credentials are invalid.
	Jun 16, 2020 5:44:18 AM	admin	CUSR0003	10.135.154.1	Audit	Successfully logged off from GUI .
	Jun 16, 2020 5:19:22 AM	admin	CUSR0001	10.135.154.1	Audit	Successfully logged in from GUI .
	Jun 16, 2020 5:19:14 AM	admin	CUSR0003	10.135.154.1	Audit	Successfully logged off from GUI .
	Jun 16, 2020 4:58:54 AM	admin	CUSR0001	10.135.154.1	Audit	Successfully logged in from GUI .
	Jun 16, 2020 4:55:15 AM	admin	CAPP1322	10.135.154.1	Audit	The action Install is successfully started for extension Power Manager version 1.1.0.47 by the user admin .
	Jun 16, 2020 4:54:58 AM	admin	CAPP1322	10.135.154.1	Audit	The action Download is successfully started for extension Power Manager version 1.1.0.47 by the user admin .
	Jun 16, 2020 4:54:33 AM	admin	CAPP0178	10.135.154.1	Configuration	Console setting automatically uodate the console after download is complete modified.
	Jun 16, 2020 4:54:33 AM	admin	CJOB0172	10.135.154.1	Configuration	The job Default Console Update Task with id 10014 of type consoleupdate has been enabled.
	Jun 16, 2020 4:53:42 AM	admin	CUSR0001	10.135.154.1	Audit	Successfully logged in from GUI .

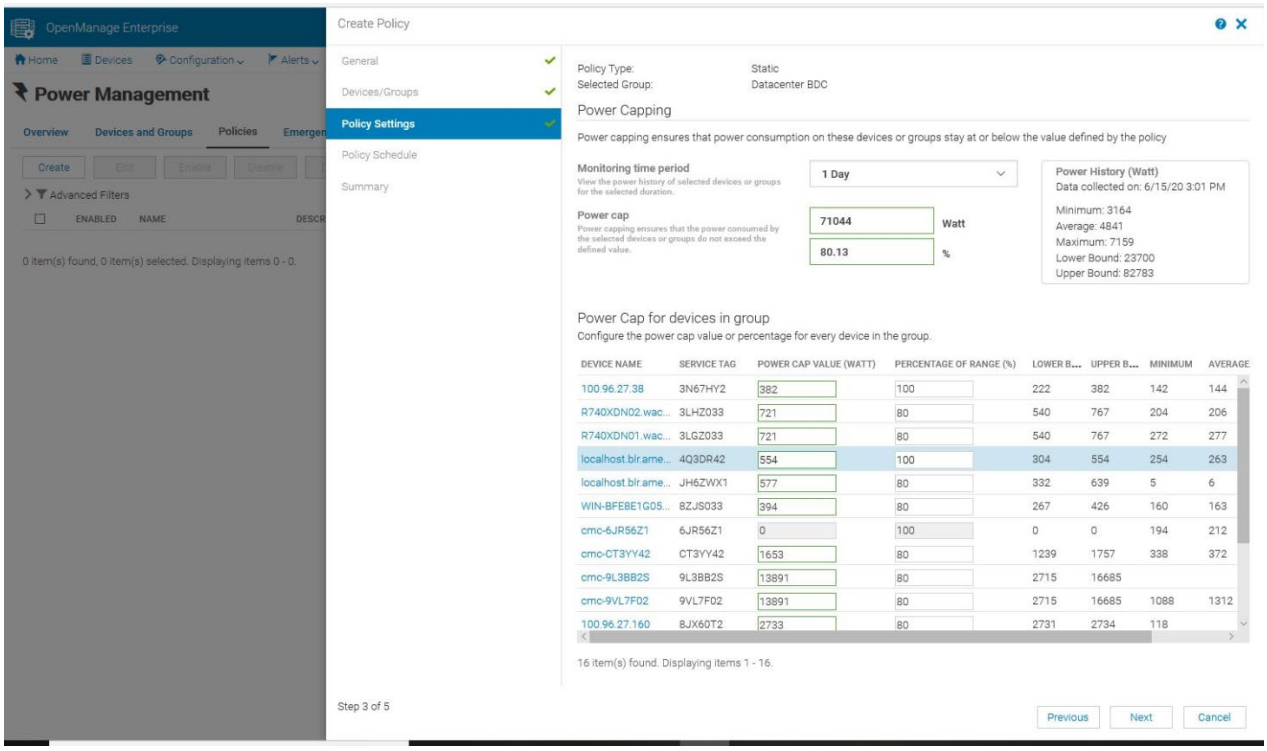
The following image shows the ‘wizard’ to help create a policy action based on event.



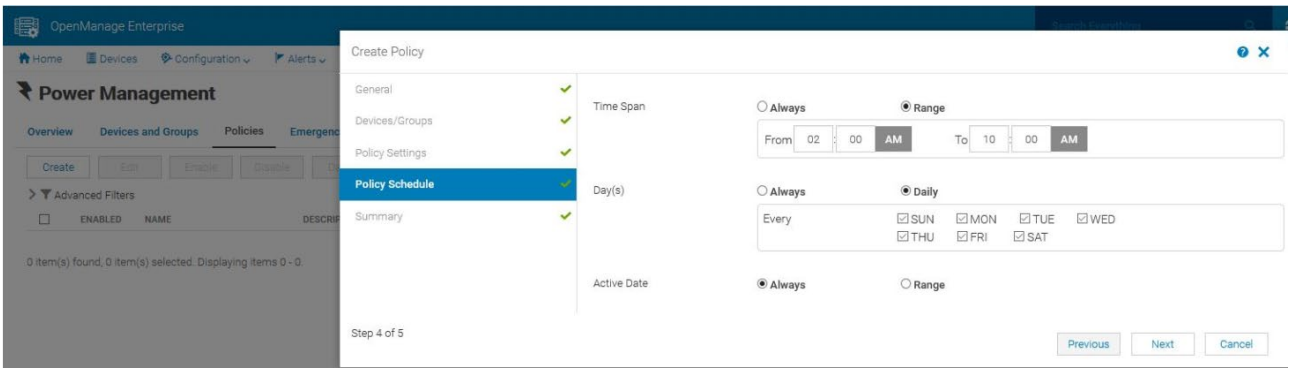
OpenManage Enterprise Power Manager is a plug-in to OpenManage Enterprise. Power Manager uses power capping to ensure power for a group of servers remains within the envelope the customer sets for the server group. An admin defines a group of servers by rack, row, or room of a data center. This policy ensures that a rack does not spike power over the allotted WATTs in order to prevent breaker tripping or grid failure. In addition, Power Manager uses power capping to immediately lower power on a group of servers that exceeds the customer set temperature. Power capping can prevent an outage of servers due to high temperatures.

OpenManage Enterprise Power Manager uses alerts to warn the customer using the OpenManage Enterprise console, email, or mapped to another vendor console. This integration allows the customer to react to the power or thermal alert for a group of servers before experiencing an outage.

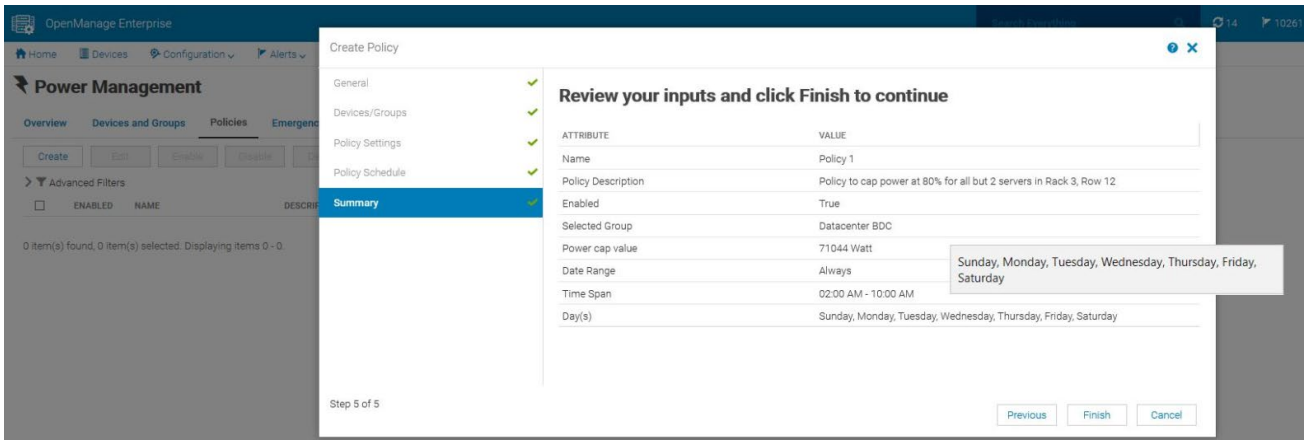
The following image shows the granular choices on the “Policy Settings” page of the wizard.



Step 4 shows the details of the “Policy Schedule.”



Step 5 is the summary slide.



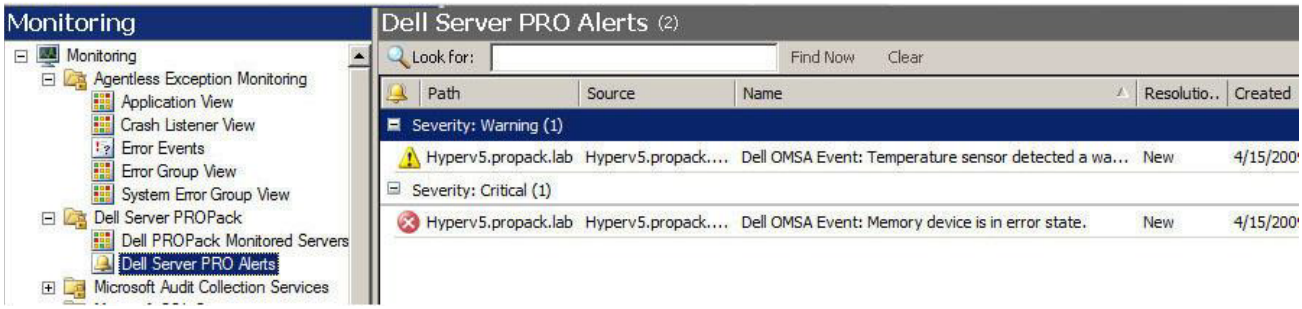
3.2 Alerts and Partner Consoles

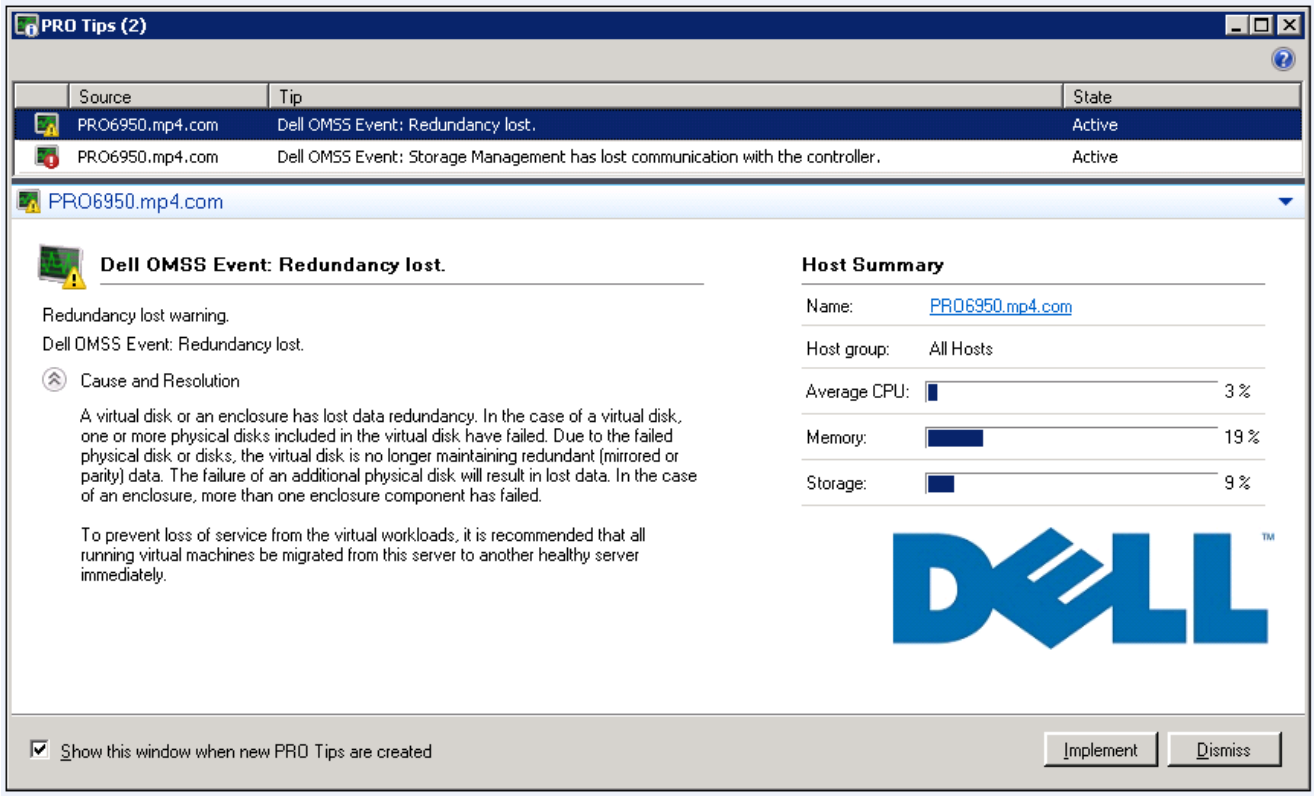
For virtualized environments running Microsoft Hyper-V or VMware ESXi, Dell EMC offers OpenManage integrations with both System Center Virtual Machine Manager (SCVMM) and VMware vCenter. This integration allows customers to set different actions that are based on alert type and severity.

Dell EMC and Microsoft have worked together since 2009 developing PRO packs which can take standard server alerts and translating them into end-user actionable events.

IT administrators can set a certain policy that is based on a specific action. For example, it is possible to place a server in maintenance mode and vacate virtual machines if a server loses a redundant power supply. Alternately, such alerts can inform an administrator without triggering a specific action.

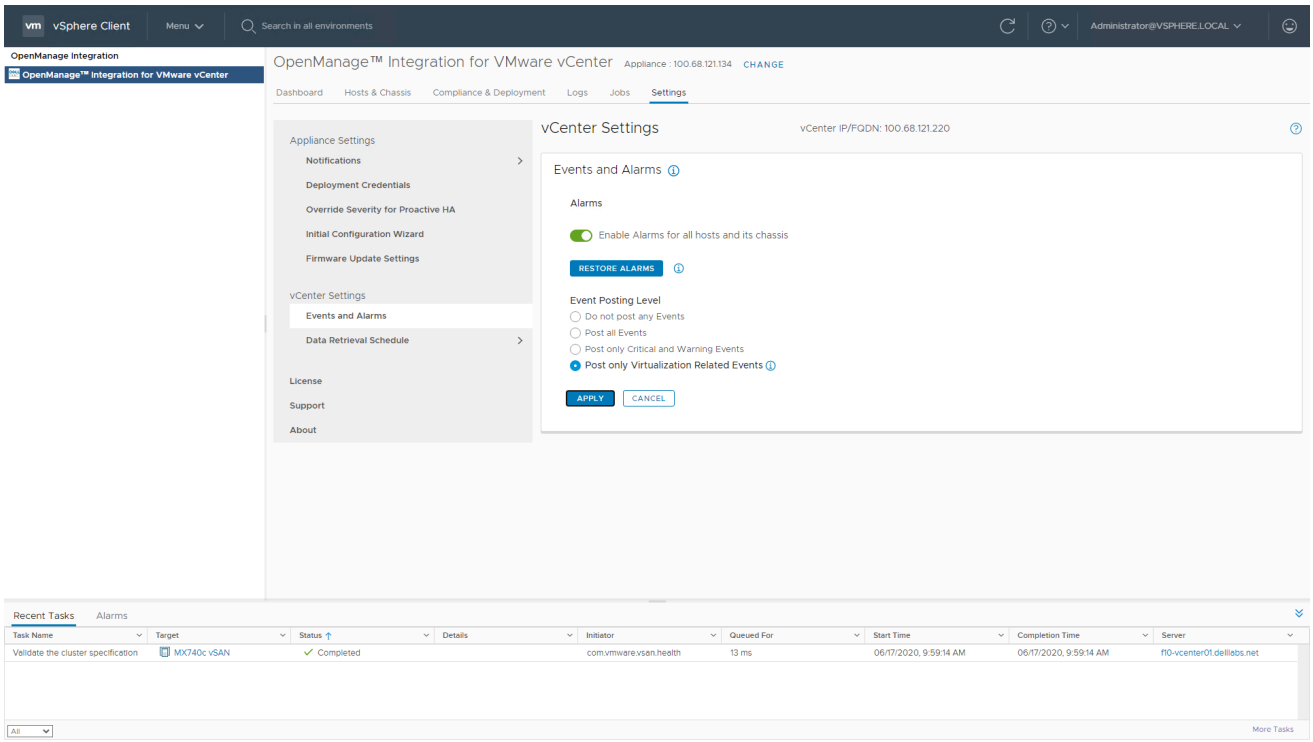
In the examples below, an IT administrator receives a PRO alert, such as a memory device or redundancy lost alert. Then, the administrator can decide to implement the action, dismiss it, or take other action.



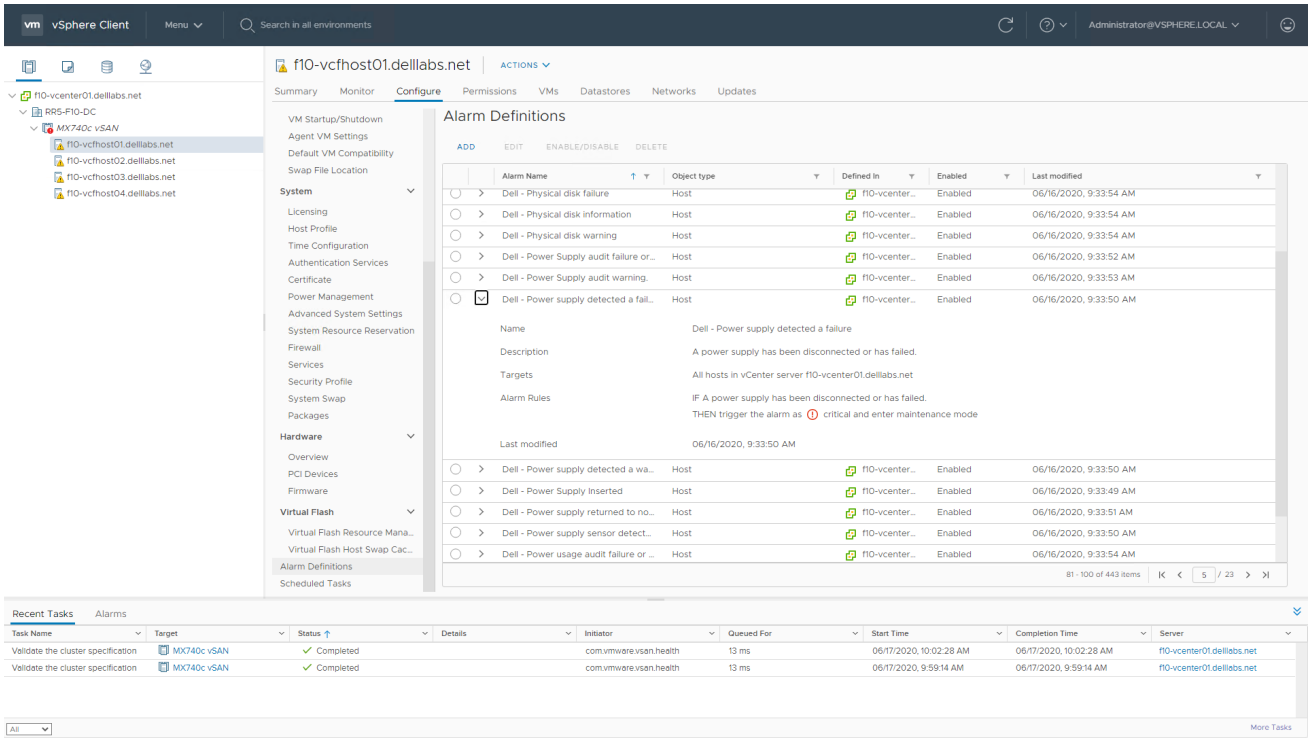


Alert integration is also available for VMware customers. The Dell EMC OpenManage Integration for VMware vCenter inserts custom alarm definitions that enable administrators to remediate failures in an automated fashion, as shown below. By integrating into VMware “events, alarms, and actions mechanism,” administrators can see and react to hardware errors. Options include placing the server in maintenance mode, running a batch file, or sending an email. This integration reduces workload downtime by ensuring virtual machines can be moved to another host in the cluster before a catastrophic physical error occurs.

The following image shows the page for enabling alarms and events in OpenManage Integration for VMware vCenter console.



The next image shows the alarm definitions at the host level in vCenter.



4 Conclusion

Dell EMC delivers an extensive PFA and performance monitoring technology by the iDRAC embedded in every PowerEdge server. The iDRAC and the comprehensive OpenManage portfolio provide effective, proactive management that is designed to make IT administrators more effective and efficient. PFA monitoring and alerts from iDRAC are the first steps in this process. The solution becomes more powerful when combined with “one to many” consoles such as OpenManage Enterprise, Microsoft System Center, or VMware vCenter.

Dell EMC systems management technology is dedicated to monitoring and managing your servers so you can manage your business.

A Technical support and resources

The iDRAC support home page provides access to product documents, technical white papers, how-to videos, and more.
www.dell.com/support/idrac

iDRAC User Guides and other manuals
www.dell.com/idracmanuals

Dell Technical Support
www.Dell.com/support