

Proactive HA with Dell EMC OpenManage Integration for VMware vCenter (OMIVV)

Pandiyan Varadharajan, Sr. Manager Software Engineering

Jonas Werner, Sr. Solutions Architect Dell EMC Customer Solutions Center

June 2017

Revisions

Date	Description
June 2017	Initial release

The information in this publication is provided "as is." Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

Copyright © 2017 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA [6/30/2017]

Dell believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

Table of Contents

1	Executive summary					
2	2 Introduction					
	2.1	Dell EMC OpenManage Integration for VMware vCenter (OMIVV)	5			
	2.2	VMware vSphere 6.5 with Proactive HA	5			
3	Solu	tion detail	6			
	3.1	Proactive HA compared to normal HA functionality	6			
	3.2	Introducing the Proactive HA provider	7			
	3.3	Monitored hardware components	7			
	3.4	Proactive HA Health Computation Engine in the Dell Provider	9			
	3.5	Dell EMC Proactive HA provider alert workflows	11			
	3.6	Customizing severity levels	11			
	3.7	Actions and responses to health status changes	12			
4	Proa	ctive HA configuration	13			
	4.1	OMIVV deployment and configuration	13			
	4.2	Proactive HA enablement and configuration	13			
5	Licer	nsing	16			
6	Tips	and troubleshooting	17			
	6.1	Enabling SNMP in the iDRAC of the ESXi host servers	17			
	6.2	Monitoring real-time alert flow	18			
	6.3	Discovery and Inventory	19			
7	Sum	mary	21			
8	Dell EMC Customer Solution Centers					
9	9 Links for further study					
10	Арре	endix	24			
	10.1	iDRAC traps and event IDs with explanations	24			
		VRTX and FX2 traps. Also valid for M1000E when "Enable Enhanced Chassis Logging and Events" is bled	25			

1 Executive summary

Virtualization has come to permeate all layers of the modern IT infrastructure, from virtual machines to virtualized storage and networking. Along with benefits in the shape of cost savings, increased efficiency and speed of service delivery also comes the risk of a hardware failure bringing down or delaying a significant number of a company's services. The solution to this is closer interaction between the underlying hardware and the virtualization platform along with a way to communicate potential hardware issues before they become fatal. With the OpenManage Integration for VMware vCenter (OMIVV) version 4.0.1 and the new VMware Proactive HA provider (vCenter version 6.5), Dell EMC have provided another layer of reliability for this challenge.

While VMware's High Availability (HA) functionality has been available on most virtualized platforms for years, it has the drawback that it reacts only once a host has completely failed and the damage is already done. HA will ensure that virtual machines will be rebooted but the impact of a non-graceful shutdown and the ensuing wait time for them to boot after a failure can have serious ramifications.

With the newly introduced Proactive HA functionality, in combination with the OMIVV virtual appliance, potential hardware issues can be detected and acted upon proactively – before they result in actual downtime.

This document provides an overview of VMware Proactive HA and OMIVV, as well as showcasing the settings and use of this new functionality to drive higher levels of uptime on VMware virtualized environments running on PowerEdge servers.

2 Introduction

The widespread virtualization of workloads have benefited organizations in countless ways from ease of management to ensuring a higher level of availability. As increasing number of applications, services and business-critical systems are virtualized, they come to rely upon the underlying hardware layer to ensure redundancy in case of failure. The more powerful the server, the larger the number of virtual machines supported and the larger the potential impact to the loss of a given system. Dell EMC PowerEdge servers have several innovations in this field to ensure uptime of the hypervisor and the workloads it supports, such as the redundant SD card system for hosting the ESXi hypervisor and Fault-Resilient memory to provide an extra layer of memory protection for the hypervisor itself.

Despite all these efforts to maintain a high level of uptime, failures do occur occasionally with increasing chances over multiple years of active usage of the hardware. Proactive HA allows for policy-driven preventive action in case an issue in the underlying hardware layer is detected. Via the policy settings, automated actions can be taken to move VMs off an affected host or prevent any new VMs to be started there. The ability to do this without human intervention also help administrators be more efficient and proactive, as well as helping improve uptime for services deployed in the virtual environment.

2.1 Dell EMC OpenManage Integration for VMware vCenter (OMIVV)

OMIVV is a plugin to VMware vCenter to bridge the gap between VMware virtualization platform and the underlying Dell EMC servers. It is easily deployed as a virtual appliance on an existing VMware cluster and hooks into the server iDRAC as well as the ESXi hypervisor deployed on each host.

OMIVV provides significant value to a PowerEdge based virtualized infrastructure solution by enabling:

- Bare-metal provisioning of new hypervisor hosts
- Cluster-aware firmware updates
- · Hardware health monitoring
- Deep system hardware inventory and health information
- Maintenance and support information

With the introduction of VMware vSphere 6.5 and OMIVV version 4.0 it also became possible to leverage Proactive HA functionality, communicating moderately degraded (warning) and severely degraded (critical) hardware states to help ensure increased uptime.

2.2 VMware vSphere 6.5 with Proactive HA

With the release of vSphere 6.5, one of the new features VMware introduced was the Proactive HA functionality, allowing hardware health information to be communicated to vCenter. Proactive HA allows a customer to set policies on what to do when a server state has degraded, including automated actions such as putting the degraded host into a specific state as to move running workloads to other hosts in the cluster or simply keep new virtual machines from being moved to the impacted host. This is actually a feature of the Distributed Resource Scheduler (DRS) feature in vCenter, which springs into action prior to the need to activate the standard HA functionality.

3 Solution detail

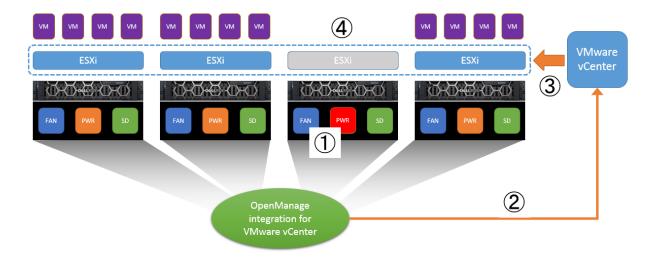
3.1 Proactive HA compared to normal HA functionality

In a normal clustered VMware environment the HA functionality will only spring to action once a failure of a host has occurred. Any virtual machines (VMs) that were running on the affected host will fail simultaneously as the host becomes unavailable but will be automatically restarted on the remaining hosts in the cluster. Since this results in a non-graceful shutdown of the VMs as well as incurring a waiting time while they boot, this is a disruptive event which in worst case can result in data loss.

As the name suggests, Proactive HA takes action prior to a host stops working due to a hardware failure. Downtime and potential data loss can potentially be prevented thanks to Proactive HA's ability to move VMs off the affected host and onto healthy servers without them having to shut down and restart.

Example: Proactive-HA event flow in case of PSU failure

- 1. A server experiences a PSU failure which is considered Severely degraded (Critical)
- 2. The OpenManage Integration for VMware vCenter plugin detects the change in server health status and informs VMware vCenter, flagging the server as Moderately Degraded (Warning) or Severely degraded (Critical) based on the alert settings.
- 3. VMware vCenter performs the pre-set Proactive HA action associated with a Severely degraded (Critical) health problem. In this case the action is to enable Maintenance mode
- 4. Maintenance mode is enabled on the affected server and if the automation is enabled DRS will migrate any VMs to healthy servers in the same cluster automatically



3.2 Introducing the Proactive HA provider

To gain insight into the hardware layer, the vCenter server in a cluster leverages a "Proactive HA provider" – a vendor provided software package with connections into the server and the ability to monitor health status in real-time. The provider has the ability to talk to the vCenter server directly in order to report any issues as they are detected. In a PowerEdge based ESXi cluster the Proactive HA provider is part of the OMIVV plug-in starting in version 4.0.

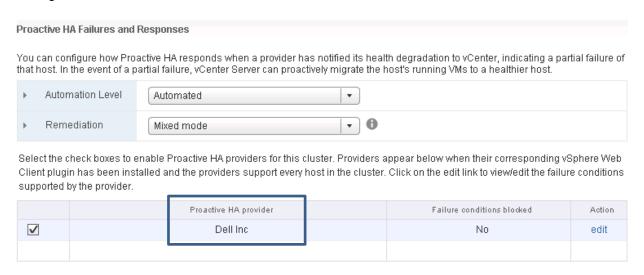


Figure 1 — Dell Inc. Proactive HA provider

3.3 Monitored hardware components

Specific subsets of hardware in a server are supported for Proactive HA. In the 4.0.x release, OMIVV support monitoring of the following three component types:

- Power supplies
- Fans
- Storage (Dell EMC IDSDM redundant SD card only)

In addition to the three component types listed above, VMware also support monitoring of the following types of hardware:

- Memory
- Network

In the initial Proactive HA support, redundancy loss was the primary focus which is why the focus on power, thermals, and the dual SD card for ESXi installs. Support for Memory and Network and extended Storage support may be introduced in a later OMIVV revision.

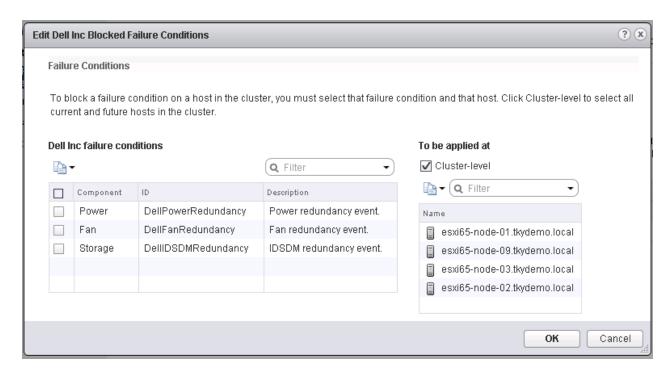


Figure 2 - Supported hardware component types and failure conditions

These hardware component types are monitored and their respective health will be reported back to vCenter as being one of the following four types: OK, Moderately Degraded (Warning), Severely degraded (Critical) or Unknown. The vCenter server will then take action to ensure that VMs won't suffer downtime. Different actions can be configured depending on the alert severity level.

Status	Color
Unknown	Gray
ОК	Green
Moderately Degraded (Warning)	Yellow
Severely degraded (Critical)	Red

3.4 Proactive HA Health Computation Engine in the Dell Provider

The provider computes these statues based on the redundancy status of the subsystem and the component status. Since PowerEdge servers come with Power Supply redundancy capabilities, the provider ensures that the host moves into Quarantine or Maintenance mode only when it is absolutely necessary and not based on any isolated failure in a PSU component.

For instance, on a standard 2U R730 rackmount server with dual power supplies, the redundancy can be configured as follows:

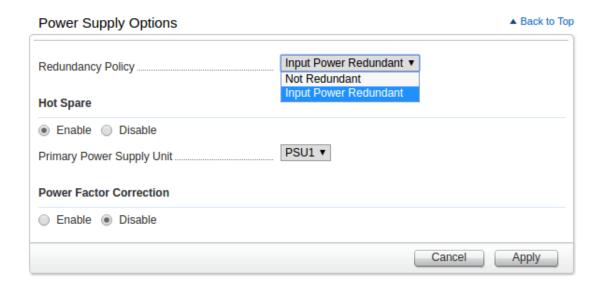


Figure 3 — Configuring PSU redundancy and priority in the server iDRAC



Figure 4 - Rear view of an R730 rackmount server with dual Power Supply Units (PSU)

To use a Power Supply subsystem as an example, the Provider would follow an intelligent inference such as the following based on how the server has the PSU redundancy policy configured,

- When the Redundancy Policy is set to Not-redundant, the provider would compute the state of Power Supply Subsystem to be Severely degraded (Critical) independent of the PSU component status. This is because for a host to be configured as part of a Proactive HA cluster, Dell EMC recommends that this policy is configured with full redundancy
- However, if the policy is set to Input Power Redundant and if the subsystem level redundancy is
 degraded or lost due to some of the PSU components failure, the provider would accordingly send
 either a Moderately Degraded (Warning) or a Severely degraded (Critical) state update to vCenter.

Note that the system has been designed to send notifications only when deemed absolutely necessary, as well as providing the ability for the vCenter administrator to override default notification levels from the OMIVV health provider to vCenter. As such, the Dell EMC provider can be mindful of the existence of administrator priorities and will notify of events in a manner which will help maximize uptime and ensure that required business service levels are met.

How host status is evaluated at a vCenter level:

Whenever a cluster is configured with a Proactive HA policy, it determines the overall status of each host by scanning through the registered events to map the current status of each of the redundancy component – it follows a principle of aggregating the rollup health by taking worst state of all components. In this case the server have both a Moderately Degraded (Warning) state and a Severely degraded (Critical) state. The result is that the overall state of the server is Severely degraded (Critical), as shown in the below graphic.



3.5 Dell EMC Proactive HA provider alert workflows

To begin with, vCenter would assume that these server components are in an Unknown state until the health provider updates it with a known state. The health provider has three major workflow cycles, namely: Init, Poll and Notify. Each is covered briefly below:

Init

Provider would check the current status of these components and initialize the state of these components in vCenter with corresponding values (OK, Moderately Degraded (Warning) or Severely degraded (Critical)).

Poll

At a periodic interval, the Provider will connect to the server and check for the current state of these components. Only upon detecting a change in the previous state, the provider will send an update to vCenter. The trigger for OMIVV for such initialization will be the following:

- Configuration of a cluster for PHA.
- vCenter restart
- OMIVV restart

Thus when OMIVV identifies a change in the cluster configuration for PHA, it will scan through all the host in that cluster and send health updates for all the PHA events on all the supported components (currently PSU, Fan and IDSDM).

Notify

When a trap is received from iDRAC, OMIVV will send updates via the health provider interface; sometimes it will also check via polling. Hence if a server is perfectly healthy the health updates would be sent as normal on the first poll. Thereafter, subsequent health update for a component will be sent to vCenter only when there is a difference in the state received from iDRAC (through poll or trap) to the value cached in the OMIVV database.

3.6 Customizing severity levels

Predefined severities are available when Proactive HA is enabled, but these can be overridden depending on how OMIVV is configured. For example, a fan event which by default would be marked as being of a Moderately Degraded (Warning) severity may be changed to be Severely degraded (Critical).

The reasoning behind the ability to change these settings is that while the defaults are set by best practices in PowerEdge servers as to what constitutes a Severely degraded (Critical) vs. a Moderately Degraded (Warning) event on the platform, the actual environment may call for something different. With test or development beds, for example, the loss of a PSU may not be a critical event because it was removed for another configuration. The intent is to provide flexibility to the administrator in circumstances where the defaults do not serve their needs.

3.7 Actions and responses to health status changes

Actions will be taken depending on the type and severity of the status reported to VMware vCenter by OMIVV. Two basic types of actions are available and they can be implemented in three different way. The first and most basic action is to enter Maintenance mode and automatically have all VM's on an affected host be live-migrated to other hosts within the cluster. If a fully-automatic DRS is configured on the cluster, this will happen without administrator intervention.

The other option is called "Quarantine" and was introduced as a new status type in vSphere version 6.5 together with Proactive HA. In quarantine status, vSphere will prevent additional VMs from being scheduled to boot on or be migrated to a particular server. VMs currently running on the server will be unaffected and remain where they are.

Maintenance mode and Quarantine mode can be combined in three modes:

Maintenance only	Maintenance mode for all failures		
Quarantine only	Quarantine for all failures		
Mixed mode	Quarantine mode for moderate failures and Maintenance mode for severe failures		

The OMIVV Proactive HA health provider works on the fundamental assumption that any change in the component level (such as PSU, Fan and IDSDM in OMIVV 4.0) redundancy degradation or lost will trigger an action on the corresponding host and move them to Quarantine or Maintenance or Mixed mode as per the user configuration. This is to ensure that the health provider alerts the admin to take any proactive measure as soon as the redundancy state is compromised for any one of the monitored parts in the server, thereby avoiding an imminent failure in the near future.

However, change in redundancy status is generally triggered by a primary event, component state change (such as PSU or Fan failure). Hence a redundancy event is almost always preceded by a component state change event, and if a Failure and Response is configured on such an event would always be performed before any action is taken on the Proactive HA Failures and Responses setting. Therefore the health provider will take ownership of hosts that are part of Proactive HA cluster and Proactive HA Failure Response is triggered by suppressing the General Failure and Response construct. A full list of the traps can be found in the <u>Appendix</u>.

The actions to take can also be either manual, where vCenter will suggest actions to take based on severity, or automated where VMs will automatically be migrated to healthy hosts from hosts experiencing problems.

Detail on how to configure these settings are included in section 4.2 - Proactive HA enablement and configuration.

4 Proactive HA configuration

This section contains a quick setup guide for getting started with OMIVV and Proactive HA on a PowerEdge server using vSphere 6.5

4.1 OMIVV deployment and configuration

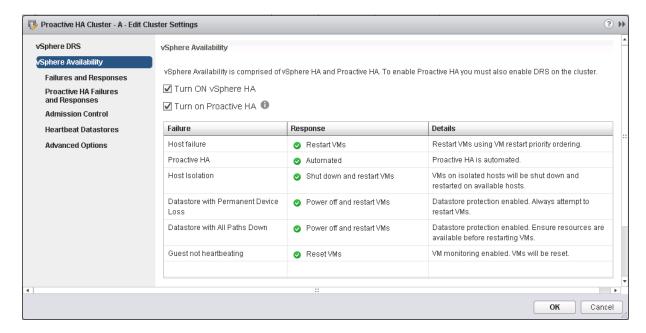
The OMIVV 4.0 virtual appliance can be downloaded directly from the Dell EMC Support webpage: <a href="http://www.dell.com/support/home/sg/en/sgbsdt1/Drivers/

Additional information pertaining to OMIVV, including videos and whitepapers can be found here: http://en.community.dell.com/techcenter/systems-management/w/wiki/1961.openmanage-integration-for-vmware-vcenter

4.2 Proactive HA enablement and configuration

Once OMIVV have been deployed and is receiving alerts from the servers it is possible to start the configuration of Proactive HA. In order to enable Proactive HA the normal HA functionality in vSphere must also be enabled for the cluster.

From the main vSphere 6.5 Web Client screen, navigate to "Hosts and Clusters", Select the cluster on which to enable Proactive HA, select "Configure", "vSphere Availability" and click the "Edit" button on the top right to open the Cluster DRS and HA settings screen. From here, first check "Turn on Proactive HA":



The next step is to highlight "Proactive HA Failures and Responses" and check the box next to the Dell Inc. Proactive HA provider.

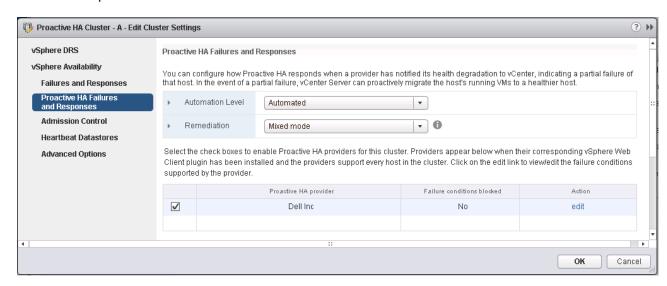


Figure 6 - Enabling the Dell Inc. Proactive HA provider

On the same screen the Automation Level as well as the Remediation settings can be found. As long as the VMware vSphere license allows it, it is recommended to enable DRS to be fully automated as well as setting the Automation Level to "Automated" for Proactive HA. A manual option is also available but will require human intervention to move VMs off a host experiencing a degraded health state.

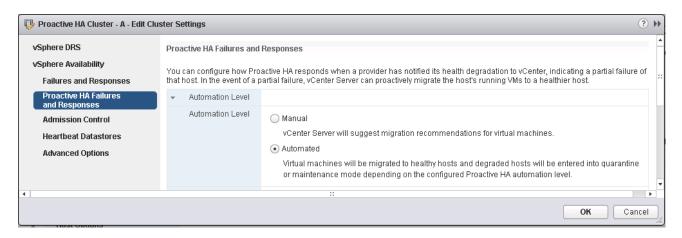


Figure 7 - Configuring the Proactive HA Automation Level

The Remediation settings decide the action to take upon detecting a degraded health state. The three options available are well explained by the VMware vSphere Web Client as per the below:

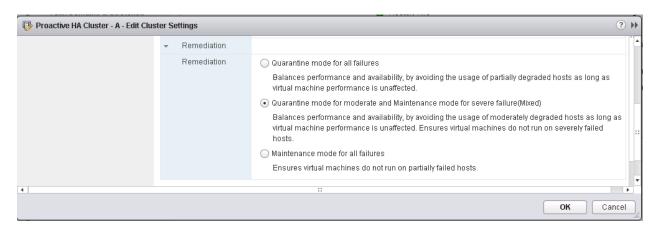


Figure 8 — Configuring the Proactive HA Remediation settings

The Proactive HA severity levels can be overridden in the OMIVV configuration screen. To do this, navigate to the main VMware vSphere Web Client dashboard, select the Dell icon, click "Manage", "Proactive HA Configuration" and "Proactive HA Events". From here it is possible to change the default severity levels for individual components.

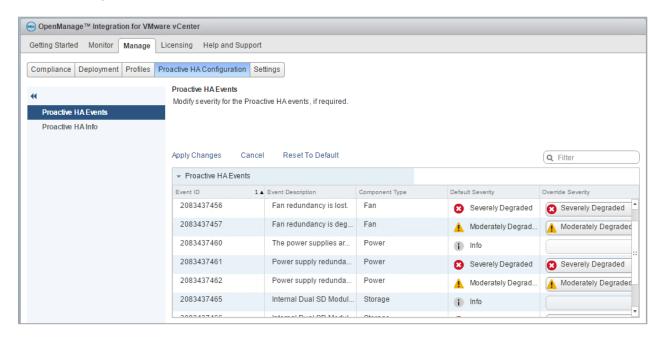


Figure 9 - Overriding the default Proactive HA health severity states

5 Licensing

Both VMware vSphere and OMIVV are licensed products. The VMware vSphere licensing requirement to use Proactive HA starts with vSphere Enterprise Plus. OMIVV is licensed per PowerEdge ESXi host on a three or five year subscription model. Please contact your Dell EMC sales representative for a quotation.

A 90 day evaluation license of OMIVV covering five ESXi hosts is available for customers who register here: https://marketing.dell.com/software-download

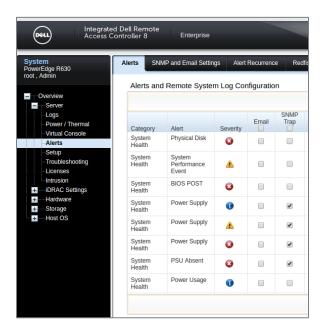
6 Tips and troubleshooting

Configuring OMIVV and Proactive HA should be a straight-forward experience but in case behavior doesn't match expectations, consider reviewing the below points to ensure configuration is accurate.

6.1 Enabling SNMP in the iDRAC of the ESXi host servers

If the timing between an event and the action taken by VMware vSphere is longer than expected it can be due to SNMP not being enabled for the correct events in the iDRAC of the ESXi host servers. If correctly configured, action will be taken within seconds. If a misconfiguration is present it could be several minutes until Quarantine or Maintenance mode is invoked. Note that this process is also dependent on the OMIVV health polling cycle.

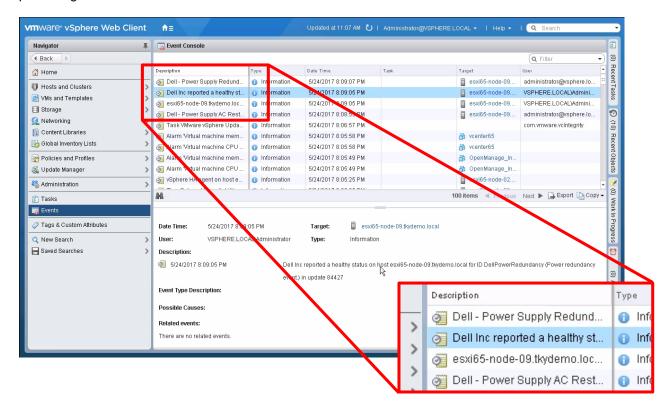
Firstly enable SNMP traps for Power Supply, Fan and IDSDM events and then set the OMIVV server as a trap destination in the SNMP Traps and Email settings. If a large number of servers require the same settings please consider using RACADM, WS-MAN or Redfish to export and import the alert settings to save time with the configuration.





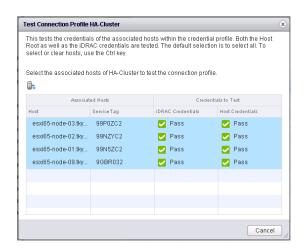
6.2 Monitoring real-time alert flow

To troubleshoot an installation that is not performing as expected it can be useful to monitor the flow of requests between the iDRAC, OMIVV and vCenter in real-time. This can be done from the "Events" panel in the VMware vSphere Web Client as per the below screenshot. Filtering on "Dell Inc" will bring up events pertaining to OMIVV and Proactive HA.

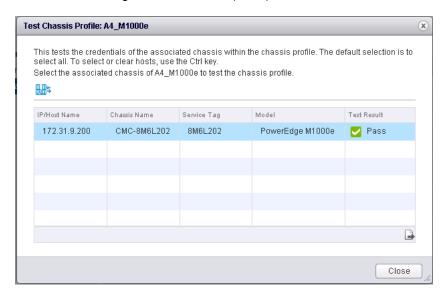


6.3 Discovery and Inventory

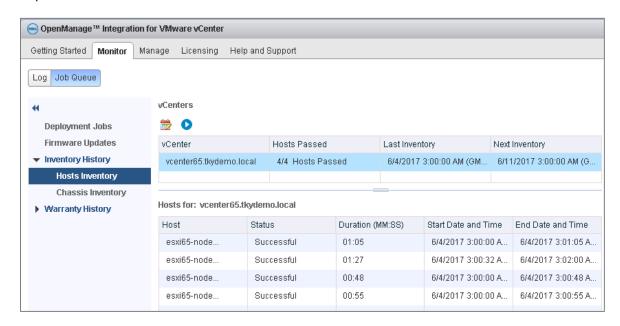
Ensuring systems have been discovered and inventoried is important for the correct functionality of Proactive HA. Firstly ensure that the servers have been properly discovered in OMIVV. This is done by entering via the main OMIVV shown as a Dell icon from the vSphere Web Client dashboard and entering "Manage" and "Profiles".



If all or a some of the servers are part of a chassis, for example an M1000e or FX2 chassis, also ensure that the Chassis Management Controller (CMC) have been added and that communication is working.



If the systems have been discovered properly, also make sure they have been inventoried at least once. This can be done from the OMIVV main menu by selecting "Monitor", "Job Queue" and "Hosts/Chassis Inventory". From here it is possible to view the last time the systems were inventoried and to kick off the inventory job if required.



7 Summary

Using the OpenManage Integration for VMware vCenter, administrators can benefit from having VMware vCenter obtain the hardware health status of PowerEdge servers and take automated action based on hardware status changes.

Dell EMC is a committed leader in the development and implementation of solutions to help build and support modern IT infrastructure. Supporting Proactive HA within the OpenManage Integration for VMware vCenter further enhances the functionality of VMware vSphere, the manageability of PowerEdge servers and provides another powerful tool to help IT administrators maintain uptime and help save time and money.

8 Dell EMC Customer Solution Centers

This white paper was produced with the assistance of the Dell EMC Customer Solution Center, Tokyo. Dell EMC Customer Solution Centers are a global network of connected labs that enable Dell EMC customers to strategize, architect, validate and build solutions, from the data center to the edge of the network. With centers located around the globe, they can help customers whether through an informal 30-60 minute briefing, a half-day workshop, or a proof-of-concept that enables "kicking the tires" of a solution prior to a purchase decision. Customers can contact their Dell EMC account team to initiate an engagement with a Customer Solution Center.

To learn more visit http://www.dell.com/customersolutioncenter



Figure 10 - Dell EMC Customer Solution Center locations

9 Links for further study

The OMIVV virtual appliance can be downloaded directly from the Dell EMC Support webpage: http://www.dell.com/support/home/sg/en/sgbsdt1/Drivers/DriversDetails?driverid=DYKN6.

Additional information pertaining to OMIVV, including videos and whitepapers can be found on the OMIVV section of the Dell EMC TechCenter webpage: http://en.community.dell.com/techcenter/systems-management/w/wiki/1961.openmanage-integration-for-vmware-vcenter

For more information on iDRAC with Lifecycle Controller, visit the **Dell TechCenter**.

10 Appendix

10.1 iDRAC traps and event IDs with explanations

Name	TrapID	Description	Category	SubCategory	Severity	Device
Fan Information	2155	Fan information.	System Health	Fan Event	Informational	Fan
Fan Warning	2154	Fan warning.	System Health	Fan Event	Minor	Fan
Fan Failure	2153	Fan failure.	System Health	Fan Event	Critical	Fan
Power Supply Normal	2187	Power supply has returned to normal.	System Health	Power Supply	Informational	PSU
Power Supply Warning	2186	Power supply has detected a warning.	System Health	Power Supply	Minor	PSU
Power Supply Failure	2185	Power supply has detected a failure.	System Health	Power Supply	Critical	PSU
Power Supply Absent	2465	Power supply is absent.	System Health	PSU Absent	Critical	PSU
Redundancy Information	2475	Redundancy information.	System Health	Redundancy	Informational	
Redundancy Degraded	2474	Redundancy is degraded.	System Health	Redundancy	Minor	
Redundancy Lost	2473	Redundancy is lost.	System Health	Redundancy	Critical	
Integrated Dual SD ModuleInformation	2211	Integrated Dual SD Module information.	System Health	IDSDM Media	Informational	IDSDM
Integrated Dual SD ModuleWarning	2210	Integrated Dual SD Module warning.	System Health	IDSDM Media	Minor	IDSDM
Integrated Dual SD ModuleFailure	2297	Integrated Dual SD Module failure.	System Health	IDSDM Media	Critical	IDSDM
Integrated Dual SD ModuleAbsent	2481	Integrated Dual SD Module is absent.	System Health	IDSDM Absent	Critical	IDSDM
Integrated Dual SD Module Redundancy Information	2491	Integrated Dual SD Module redundancy information.	System Health	IDSDM Redundancy	Informational	IDSDM
Integrated Dual SD Module Redundancy Degraded	2490	Integrated Dual SD Module redundancy is degraded.	System Health	IDSDM Redundancy	Minor	IDSDM

Integrated Dual SD
Module
Redundancy Lost

Integrated Dual
SD Module
redundancy is
lost.

Integrated Dual
SD Module
System Health
IDSDM Redundancy
Critical
IDSDM

10.2 VRTX and FX2 traps. Also valid for M1000E when "Enable Enhanced Chassis Logging and Events" is enabled.

Name	TrapID	Description	Category	SubCategory	Severity	Device
alert2FanInformation	2155	Fan information.	Status Events	Fan	Informational	Fan
alert2FanWarning	2154	Fan warning.	Status Events	Fan	Minor	Fan
alert2FanFailure	2153	Fan failure.	Error Events	Fan	Critical	Fan
alert2PowerSupplyNormal	2187	Power supply has returned to normal.	Status Events	Power Supply	Informational	PSU
alert2PowerSupplyWarning	2186	Power supply has detected a warning.	Status Events	Power Supply	Minor	PSU
alert2PowerSupplyFailure	2185	Power supply has detected a failure.	Error Events	Power Supply	Critical	PSU
alert2RedundancyInformation	2475	Redundancy information.	Status Events	Redundancy	Informational	
alert2RedundancyDegraded	2474	Redundancy is degraded.	Status Events	Redundancy	Minor	
alert2RedundancyLost	2473	Redundancy is lost.	Error Events	Redundancy	Critical	