## Chapter 7: Network Function Virtualization (NFV) support

- NUMA
- Huge Pages
- OVS-DPDK
- SR-IOV
- SR-IOV with Open vSwitch (OVS) Offload (Technology Preview)
- Distributed Virtual Router (DVR)

## Chapter 8: Additional OpenStack Features

- Barbican
- Octavia
- Satellite

## Chapter 9: Operational notes

- High availability
- Service layout
- Deployment overview

## Chapter 10: Solution architecture

- Solution common settings
  - Solution Admin Host (SAH) networking
  - Node type 802.1q tagging information
  - Solution Red Hat Ceph Storage configuration
  - Solution with 25GbE/100GbE networking overview
  - Solution 25GbE/100GbE rack layout
  - Solution 25GbE/100GbE network configuration

## Chapter 11: Professional Services

- Consulting services

## Appendix A: Bill of Materials

- Bill of Materials for Dell EMC PowerEdge R-Series Solution - Mellanox NIC's
- Bill of Materials for Dell EMC PowerEdge R-Series Solution - Intel® NIC's
- Nodes Overview
- Subscriptions and Network Switches in the Solution

## Appendix B: References

- To learn more

## Glossary
List of Figures

Figure 1: OVS Architecture ........................................................................................................ 31
Figure 2: Mellanox configuration .............................................................................................. 32
Figure 3: OVS Offload .............................................................................................................. 33
Figure 4: Solution with 25GbE/100GbE, Red Hat Ceph Storage Cluster, optional Dell EMC Unity Storage and optional SC Series Storage .................................................. 45
Figure 5: 25GbE/100GbE cluster network logical architecture for CSP profile .................... 46
Figure 6: 25GbE/100GbE cluster network logical architecture for CSP OVS Offload profile ................................................................................................................................. 47
Figure 7: Dell EMC Services for RHOSP and NFVi Ready Architecture ......................... 49
List of Tables

Table 1: Dell EMC PowerEdge R640/Dell EMC PowerEdge R740xd Tested BIOS and Firmware Versions....................................................................................................... 13

Table 2: Dell EMC SC and Unity tested software and firmware versions...................... 13

Table 3: Dell EMC tested firmware versions..................................................................... 14

Table 4: iDRAC Specification for SAH Nodes................................................................. 18

Table 5: SAH BIOS Specification.....................................................................................19

Table 6: Bonding nodes supported..................................................................................26

Table 7: Overcloud: Node type to services....................................................................37

Table 8: OpenStack Node Type to Network 802.1q Tagging.........................................41

Table 9: OpenStack Compute for xSP and CSP Profile to Network 802.1q Tagging...... 42

Table 10: Storage Node Type to Network 802.1q Tagging.............................................42

Table 11: Solution Admin Host (SAH) Dell EMC PowerEdge R640...............................52

Table 12: Controller node Dell EMC PowerEdge R640..................................................52

Table 13: Compute node Dell EMC PowerEdge R640...................................................53

Table 14: Storage node Dell EMC PowerEdge R740xd..................................................53

Table 15: Solution Admin Host (SAH) Dell EMC PowerEdge R640...............................54

Table 16: Controller node Dell EMC PowerEdge R640..................................................54

Table 17: Compute node Dell EMC PowerEdge R740....................................................55

Table 18: Storage node Dell EMC PowerEdge R740xd..................................................55

Table 19: S3048-ON switch.............................................................................................56

Table 20: S4048-ON switch.............................................................................................57

Table 21: S5232F-ON......................................................................................................57
Trademarks

Copyright © 2014-2019 Dell EMC or its subsidiaries. All rights reserved.

Intel® and Xeon® are registered trademarks of Intel Corporation. Oracle® and Java® are registered trademarks of Oracle Corporation and/or its affiliates.

Mellanox, the Mellanox logo, ConnectX, CORE-Direct, and GPUDirect are registered trademarks of Mellanox Technologies, Ltd. Mellanox Multi-Host is a trademark of Mellanox Technologies, Ltd. All rights reserved.

DISCLAIMER: The OpenStack® Word Mark and OpenStack Logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries, and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation or the OpenStack community.
Notes, cautions, and warnings

A **Note** indicates important information that helps you make better use of your system.

A **Caution** indicates potential damage to hardware or loss of data if instructions are not followed.

A **Warning** indicates a potential for property damage, personal injury, or death.

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.
Overview

Topics:

• Dell EMC Ready Architecture Guide 13.2
• Hardware options
• Networking and network services
• JetPack automation toolkit
• OpenStack Architecture

An OpenStack® based cloud is now a common need by many organizations and Dell EMC with Red Hat have worked together to build a jointly engineered and validated architecture that details software, hardware, and integration points of all solution components. The architecture provides prescriptive guidance and recommendations for:

• Hardware design
  • Compute nodes
  • Infrastructure nodes
  • Storage nodes
• Network design
• Software layout
• Offers suggestion for other system configurations
Dell EMC Ready Architecture Guide 13.2

This Architecture Guide - Version 13.2 is built on Red Hat OpenStack Platform 13 which is based on the 17th OpenStack release codename **Queens**. Red Hat OpenStack Platform 13 is a containerized version offering greater scalability, resiliency and user experience.


The Red Hat OpenStack Platform provides the foundation to build a private or public Infrastructure-as-a-Service (IaaS) cloud on Red Hat Enterprise Linux. It offers a massively scalable, fault-tolerant platform for the development of cloud-enabled workloads or Telco Cloud for Service Providers.

It is packaged so that available physical hardware can be turned into a private, public, or hybrid cloud platform including:

- Fully distributed object storage
- Persistent block-level storage
- Virtual-machine provisioning engine and image storage
- Authentication and authorization mechanism
- Integrated networking

The Red Hat OpenStack Platform is implemented by a collection of interacting services that control its computing, storage, and networking resources.

**New features**

- Support for Dell EMC PowerEdge R640 and R740/740xd servers with 2nd Generation Intel® Xeon® Scalable Processors and DDR4 2933MHz memory DIMMs
- Support for the latest release of Red Hat OpenStack Platform 13 including the latest updates
- Support for RHEL 7.7 including the latest updates
- Support for Red Hat Ceph Storage version 3.2 with BlueStore, a new back end object store for the OSD daemons replacing FileStore
- Added support for Volume Multi-Attach using Unity storage
- Added Red Hat OpenStack Platform Tech Preview for OVS off-loading using Mellanox ConnectX-5 network adapters
- Added support for Red Hat Satellite 6.5 using the JetPack automation toolkit
- Added support for Red Hat OpenStack Platform Barbican for managing security keys
- Added support for Dell EMC H740P mini raid controller for use in storage nodes

**Key benefits**

The Dell EMC Ready Architecture for Red Hat OpenStack Platform offers several benefits to help service providers and high-end enterprises rapidly implement Dell EMC hardware and Red Hat OpenStack Platform software:

- Ready-to-use solution: The reference architecture has been fully engineered, validated, tested in Dell EMC laboratories and documented by Dell EMC. This decreases your investment and deployment risk, and it enables faster deployment time.
- Long lifecycle deployment: PowerEdge R-Series servers, recommended in the architecture, include long-life Intel® Xeon® processors which reduces your investment risk and protects your investment for the long-term.
- World-class professional services: The reference architecture includes Dell EMC professional services that spans consulting, deployment, and design support to guide your deployment needs.
- Customizable solution: The architecture is prescriptive, but it can be customized to address each customer’s unique virtual network function (vNF) or other workload requirements.
Co-engineered and Integrated: OpenStack depends upon Linux for performance, security, hardware enablement, networking, storage, and other primary services. The Red Hat OpenStack Platform delivers an OpenStack distribution with the proven performance, stability, and scalability of RHEL 7.7 enabling you to focus on delivering the services your customers want, instead of focusing on the underlying operating platform.

Deploy with confidence, as the Red Hat OpenStack Platform provides hardened and stable branch releases of OpenStack and Linux. The Red Hat OpenStack Platform is a long life release product supported by Red Hat for a three (3) year “production phase” life cycle, well beyond the six-month release cycle of unsupported, community OpenStack. Red Hat OpenStack Platform life cycle support policies can be found at https://access.redhat.com/support/policy/updates/openstack/platform

Take advantage of broad application support. Red Hat Enterprise Linux, running as guest virtual machines, provides a stable application development platform with a broad set of ISV certifications. You can therefore rapidly build and deploy your cloud applications.

Avoid vendor lock-in by moving to open technologies, while maintaining your existing infrastructure investments.

Benefit from the world’s largest partner ecosystem: Red Hat has assembled the world’s largest ecosystem of certified partners for OpenStack compute, storage, networking, ISV software, and services for Red Hat OpenStack Platform deployments. This ensures the same level of broad support and compatibility that customers enjoy today in the Red Hat Enterprise Linux ecosystem.

Upgrade of Red Hat OpenStack Director-based installations.

Bring security to the cloud. Rely upon the SELinux military-grade security and container technologies of Red Hat Enterprise Linux to prevent intrusions and protect your data, when running in public or private clouds.

Hardware options

To reduce time spent on specifying hardware for an initial system, this Architecture Guide offers a full solution using validated Dell EMC PowerEdge server hardware designed to allow a wide range of configuration options, including optimized configurations for:

- Infrastructure nodes
- Compute nodes
- Storage nodes

Dell EMC recommends starting with OpenStack software using components from this Architecture Guide - Version 13.2 because the hardware and operations processes comprise a flexible foundation upon which to expand as your cloud deployment grows, so your investment is protected.

As noted throughout this Architecture Guide - Version 13.2, Dell EMC constantly adds capabilities to expand this offering, and other hardware may be available.

Networking and network services

Network configuration is based upon using the Neutron-based options supported by the Red Hat OpenStack Platform code base, and does not rely upon third-party drivers. This reference configuration is based upon the Neutron networking services using the ML2 drivers for Open vSwitch with the VLAN option.

Networking includes:

- Core and layered networking capabilities
- Network Function Virtualization (NFV)
- 25GbE and 100GbE networking
- NIC bonding
- Redundant trunking top-of-rack (ToR) switches into core routers
This enables the Dell EMC Ready Architecture for Red Hat OpenStack Platform to operate in a full production environment. See Network Architecture on page 24 for guidelines. Detailed designs are available through Dell EMC consulting services.

JetPack automation toolkit

Dell EMC has opensourced the JetPack automation toolkit\(^1\), which is an innovative automation package that is used to configure the infrastructure hardware and OpenStack software in a fully automated fashion. The toolkit includes programs from Dell EMC that work with Red Hat OpenStack Director to automate the deployment process, not only saving time, but also ensuring the process is reliable and repeatable.

The JetPack automation toolkit was used to deploy Red Hat OpenStack Platform 13.2 on the documented hardware in this Architecture Guide. The JetPack automation toolkit is used by Dell EMC, Red Hat professional service teams, and system integrator partners, and is available on GitHub to customers who prefer to do self-deployment.

This release uses the concept of profiles. There are two different validated profiles, CSP or xSP, that can be used for a deployment. The CSP profile is designed for Telecommunications Service Providers, Cable TV Operators, Satellite TV, Internet Service Providers, etc. whereas the xSP profile is designed for Business & IT Services Providers such as Hosting Service Providers, Cloud Service Providers, Software-as-a-Service/Platform-as-a-Service Providers, Application Hosting Service Providers and Private Managed Cloud Service Providers.

OpenStack Architecture

While OpenStack has many configurations and capabilities, the primary components for the Dell EMC Ready Architecture for Red Hat OpenStack Platform 13 is defined as a containerized version.

\[
\text{Note: For a complete overview of the OpenStack software, visit Red Hat OpenStack Platform and the OpenStack Project.}
\]

\(^1\) The toolkit can be found at [https://github.com/dsp-jetpack/JetPack/tree/JS-13.2](https://github.com/dsp-jetpack/JetPack/tree/JS-13.2)
Chapter
2

BIOS and firmware compatibility

Topics:

• Tested BIOS and Firmware

This chapter documents the versions of BIOS and firmware that are used to create this Dell EMC Ready Architecture for Red Hat OpenStack Platform.
Tested BIOS and Firmware

*Table 1: Dell EMC PowerEdge R640/Dell EMC PowerEdge R740xd Tested BIOS and Firmware Versions* on page 13 lists the server BIOS and firmware versions that were tested for the Dell EMC Ready Architecture for Red Hat OpenStack Platform.

**Note:** Mellanox network adapters were used in all nodes and for all features in this Reference Architecture. Additional testing was done using Intel® network adapters for all features except the OVS Offload feature.

*Table 2: Dell EMC SC and Unity tested software and firmware versions* on page 13 lists the Dell EMC Storage Center, SC Series Storage software and Dell EMC Unity firmware versions that were tested for the Dell EMC Ready Architecture for Red Hat OpenStack Platform.

*Table 3: Dell EMC tested firmware versions* on page 14 lists the default S3048-ON, and optional S4048-ON switch firmware versions that were tested for the Dell EMC Ready Architecture for Red Hat OpenStack Platform.

**CAUTION:** The versions listed below are the versions that were available at the time this Reference Architecture was developed. Ensure that the firmware on all servers, storage devices, and switches are up to date. Otherwise, unexpected results may occur.

<table>
<thead>
<tr>
<th>Product</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOS</td>
<td>2.2.11</td>
</tr>
<tr>
<td>iDRAC with Lifecycle controller</td>
<td>3.34.34.34</td>
</tr>
<tr>
<td>Intel® XXV710 NIC</td>
<td>18.8.9</td>
</tr>
<tr>
<td>Mellanox ConnectX-4</td>
<td>14.24.80.00</td>
</tr>
<tr>
<td>Mellanox ConnectX-5</td>
<td>16.24.10.00</td>
</tr>
<tr>
<td>PERC H730P Mini controller (Dell EMC PowerEdge R640)</td>
<td>25.5.5.0005</td>
</tr>
<tr>
<td>PERC H740P Mini controller (Dell EMC PowerEdge R640)</td>
<td>50.5.1-2633</td>
</tr>
<tr>
<td>HBA330 Mini (Dell EMC PowerEdge R740xd)</td>
<td>16.17.00.3</td>
</tr>
<tr>
<td>BOSS-S1 (Dell EMC PowerEdge R740xd)</td>
<td>2.5.13.3020</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Product</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dell EMC SC storage center software</td>
<td>2016 R2 Build 16.2.1.228</td>
</tr>
<tr>
<td>SC Series Storage firmware</td>
<td>6.6.11.9</td>
</tr>
<tr>
<td>Dell EMC Unity firmware</td>
<td>4.5.1.0.5.001</td>
</tr>
</tbody>
</table>
Table 3: Dell EMC tested firmware versions

<table>
<thead>
<tr>
<th>Product</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3048-ON firmware</td>
<td>Dell EMC Networking OS10 Enterprise OS version: 10.4.3.4 Build version: 10.4.3.4.213</td>
</tr>
<tr>
<td>S4048-ON firmware (optional)</td>
<td>Dell EMC Networking OS10 Enterprise OS version: 10.4.3.4 Build version: 10.4.3.4.213</td>
</tr>
<tr>
<td>S5232F-ON firmware</td>
<td>Dell EMC Networking OS10 Enterprise OS version: 10.4.3.4 Build version: 10.4.3.4.213</td>
</tr>
</tbody>
</table>
Chapter 3

Server options

Topics:

- Dell EMC PowerEdge R640 server
- Dell EMC PowerEdge R740xd servers

The base validated Solution supports the Dell EMC PowerEdge R640 and Dell EMC PowerEdge R740xd Server lines.

Note: Please contact your Dell EMC sales representative for detailed parts lists.
**Dell EMC PowerEdge R640 server**

The Dell EMC PowerEdge R640 is the ideal dual-socket, 1U platform for dense scale-out cloud computing. The scalable business architecture of the Dell EMC PowerEdge R640 is designed to maximize application performance and provide the flexibility to optimize configurations based on the application and use case.

With the Dell EMC PowerEdge R640 you can create an NVMe cache pool and use either 2.5” or 3.5” drives for data storage. Combined with up to 24 DIMM's, 12 of which can be NVDIMM's, you have the resources to create the optimum configuration to maximize application performance in only a 1U chassis. This can simplify and speed up deployments of the Red Hat OpenStack Platform.

**Dell EMC PowerEdge R740xd servers**

The Dell EMC PowerEdge R740xd delivers a perfect balance between storage scalability and performance. The 2U two-socket platform is ideal for software defined storage. The R740xd versatility is highlighted with the ability to mix any drive type to create the optimum configuration of SSD and HDD for either performance, capacity or both.

The Dell EMC PowerEdge R740xd is the platform of choice for software defined storage and is the foundation of Red Hat Ceph Storage for this Architecture Guide - Version 13.2
Chapter 4

Configuring PowerEdge R-Series hardware

Topics:

- Configuring the SAH Node
- Configuring Overcloud Nodes

This section describes manually configuring PowerEdge R-Series server hardware for the Dell EMC Ready Architecture for Red Hat OpenStack Platform:

- IPMI configuration
- BIOS configuration
- RAID configuration
Configuring the SAH Node

The Solution Admin Host (SAH) is a physical server that hosts two VMs, the Red Hat Ceph Storage dashboard and the Red Hat OpenStack Director as the Undercloud.

The SAH BIOS configuration must be configured prior to OpenStack deployment. Refer to SAH BIOS Specification on page 18

The SAH is configured using the Open Source Hardware Configuration Toolkit (OS-HCTK) tool.
Please contact your Dell EMC Sales or Professional Services representative for access to the toolkit.

Open Source Hardware Configuration Toolkit

The Open Source Hardware Configuration Toolkit (OS-HCTK) is a configuration utility with sample scripts and configuration files that is used to automate the setup and configuration of BIOS and RAID settings for Dell EMC servers used for OpenStack open source software solutions.

Note: The OS-HCTK ISO is run only on the system that will be configured as the SAH node. All other systems will be configured by OpenStack Ironic.

The OS-HCTK allows a user to create a USB key from which a Dell EMC PowerEdge R640 can be booted and the BIOS and RAID settings can be applied.

Topics discussed include:

- iDRAC settings on page 18
- SAH BIOS Specification on page 18

iDRAC settings

Note: The following settings need to be set on the SAH nodes BIOS configuration prior to a deployment.

Table 4: iDRAC Specification for SAH Nodes on page 18 lists and describes iDRAC default racadm settings that will be set by the OS-HCTK.

<table>
<thead>
<tr>
<th>Menu Choice</th>
<th>iDRAC Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>iDRAC.IPMILan.Enable</td>
<td>Enabled</td>
</tr>
<tr>
<td>iDRAC.IPMILan.PrivLimit</td>
<td>4</td>
</tr>
<tr>
<td>iDRAC.IPv4.Enable</td>
<td>Enabled</td>
</tr>
<tr>
<td>iDRAC.Users.2.Enable</td>
<td>Enabled</td>
</tr>
<tr>
<td>iDRAC.Users.2.IpmiLanPrivilege</td>
<td>4</td>
</tr>
<tr>
<td>iDRAC.Users.2.Privilege</td>
<td>0x1ff</td>
</tr>
<tr>
<td>iDRAC.WebServer.Enable</td>
<td>Enabled</td>
</tr>
</tbody>
</table>

SAH BIOS Specification

Table 5: SAH BIOS Specification on page 19 lists and describes the default BIOS settings for the SAH that will be set by the OS-HCTK.
Table 5: SAH BIOS Specification

<table>
<thead>
<tr>
<th>Display Name</th>
<th>Attribute</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boot Mode</td>
<td>BootMode</td>
<td>BIOS</td>
</tr>
<tr>
<td>Boot Sequence Retry</td>
<td>BootSeqRetry</td>
<td>Enabled</td>
</tr>
<tr>
<td>DCU IP Prefetcher</td>
<td>DculpPrefetcher</td>
<td>Enabled</td>
</tr>
<tr>
<td>DCU Streamer Prefetcher</td>
<td>DcuStreamerPrefetcher</td>
<td>Enable</td>
</tr>
<tr>
<td>Logical Processor Idling</td>
<td>DynamicCoreAllocation</td>
<td>Disabled</td>
</tr>
<tr>
<td>Integrated RAID Controller</td>
<td>IntegratedRaid</td>
<td>Enabled</td>
</tr>
<tr>
<td>Internal SD Card</td>
<td>InternalSdCard</td>
<td>Off</td>
</tr>
<tr>
<td>I/OAT DMA Engine</td>
<td>IoatEngine</td>
<td>Enabled</td>
</tr>
<tr>
<td>Logical Processor</td>
<td>LogicalProc</td>
<td>Enabled</td>
</tr>
<tr>
<td>Memory Operating Mode</td>
<td>MemOpMode</td>
<td>OptimizerMode</td>
</tr>
<tr>
<td>System Memory Testing</td>
<td>MemTest</td>
<td>Disabled</td>
</tr>
<tr>
<td>Node Interleaving</td>
<td>NodeInterleave</td>
<td>Disabled</td>
</tr>
<tr>
<td>OS Watchdog Timer</td>
<td>OsWatchdogTimer</td>
<td>Disabled</td>
</tr>
<tr>
<td>Adjacent Cache Line Prefetch</td>
<td>ProcAdjCacheLine</td>
<td>Enabled</td>
</tr>
<tr>
<td>Number of Cores per Processor</td>
<td>ProcCores</td>
<td>all</td>
</tr>
<tr>
<td>Hardware Prefetcher</td>
<td>ProcHwPrefetcher</td>
<td>Enabled</td>
</tr>
<tr>
<td>CPU Power Management</td>
<td>ProcPwrPerf</td>
<td>MaxPerf</td>
</tr>
<tr>
<td>Turbo Mode</td>
<td>ProcTurboMode</td>
<td>Enabled</td>
</tr>
<tr>
<td>Virtualization Technology</td>
<td>ProcVirtualization</td>
<td>Enabled</td>
</tr>
<tr>
<td>CPU Interconnect Bus Speed</td>
<td>CpuInterconnectBusSpeed</td>
<td>MaxDataRate</td>
</tr>
<tr>
<td>SR-IOV Global Enable</td>
<td>SrvovGlobalEnable</td>
<td>Enabled</td>
</tr>
<tr>
<td>System Profile</td>
<td>SysProfile</td>
<td>PerfOptimized</td>
</tr>
</tbody>
</table>

Configuring Overcloud Nodes

Configuring server network settings
The Overcloud hardware configurations and the settings for the Intelligent Platform Management Interface, (IPMI) are done through Ironic which is included in the JetPack automation toolkit on page 11.

Configuring Server Network Settings

1. Set the iDRAC IP address source:
   a. If the Overcloud nodes were ordered with the iDRACs configured for DHCP, or are currently configured for DHCP, then no further configuration is necessary.
   b. If you wish to use static IP addresses, then configure the Overcloud nodes’ DRAC IP address, subnet mask, default gateway IP, and default VLAN (ID = 110, if required) using the iDRAC GUI.
Chapter 5

Storage Options

Topics:

- Storage Options Overview
- Local Storage
- Red Hat Ceph Storage
- Dell EMC Unity
- Dell EMC SC Series Storage

OpenStack has several storage services, including:

- Cinder
- Glance
- Manila
- Ephemeral
- Swift

Together these services provide virtual machines (VMs) with block, image, file share, and object storage. In turn, the services employ block, file share, and object storage subsystems. Since the service design has a mechanism to replace some or all of the implementation of these services, this solution can provide alternate implementations of these services that better serve your needs.

---

2 Available through a Ceph cluster with the Swift API enabled or through a custom Professional Services engagement.
Storage Options Overview

**Cinder** virtualizes storage enabling VMs to use persistent block storage through Nova. OpenStack consumers should write data that must exist beyond the lifecycle of the guest to Cinder volumes. The volume can be accessed afterwards by a different guest.

**Glance** provides images to VMs. Generally, the images are block devices containing DVDs or virtual machines. VMs can be booted from these images or have the images attached to them. Glance storage now can use Red Hat Ceph Storage, Dell EMC Unity or Dell EMC SC Series Storage.

Note: Dell EMC Unity or SC Series Storage provides the Glance support through Cinder.

**Dell EMC Manila** driver framework (EMCShareDriver) delivers a shared filesystem in OpenStack. The plugin-based Dell EMC Manila driver design is compatible with various plugins to control Dell EMC storage products.

The Dell EMC Unity plugin manages the Dell EMC Unity Storage System for shared filesystems.

The Dell EMC Unity driver is a REST API. Unity Storage System Manila backends are a one-to-one managed storage system. Each Manila backend configures a Unity Storage System. See [https://docs.openstack.org/manila/queens/admin/emc_unity_driver.html](https://docs.openstack.org/manila/queens/admin/emc_unity_driver.html)

**Swift** provides an object storage interface to VMs and other OpenStack consumers. Unlike block storage where the guest is provided a block device of a given format and is accessible within the cluster, object storage is not provided through the guest. Object storage is generally implemented as a HTTP/HTTPS-based service through a web server. Client implementations within the guest or external OpenStack clients would interact with Swift without any configuration required of the guest other than providing the requisite network access. For example, a VM within OpenStack can put data into Swift, and later external clients could pull that data for additional processing.

Note: "Swift" in this document refers to the Swift interfaces, not the Swift implementation, of the protocol.

As with other OpenStack services, there are client and server components for each storage service. The server component can be modified to use a particular type of storage rather than the default. For example, Cinder uses local disks as the storage back-end by default. The Dell EMC Ready Architecture for Red Hat OpenStack Platform modifies the default configuration for these services.

All virtual machines will need a virtual drive that is used for the OS. Two options are available:

- Ephemeral disks
- Boot from volume or snapshot, hosted on Red Hat Ceph Storage, Dell EMC Unity Storage or SC Series Storage arrays.

Ephemeral disks are virtual drives that are created when a VM is created, and destroyed when the VM is removed. The virtual drives can be stored on the local drives of the Nova host or on a shared file system, such as Ceph Rados Block Device (RBD). During the planning process, decisions can be made to place ephemeral on local or shared storage, it is recommended that a shared backend is used which will allow live migration.

Boot from volume/snapshot will use one of the Cinder backends.

The Dell EMC Ready Architecture for Red Hat OpenStack Platform includes alternate implementations of Cinder that enable the cluster to fit many needs. Cinder has been validated using each of the back-ends independently, and multi back-ends utilizing multiple Storage back-ends consisting of two or all of:

- **Local Storage** on page 22
- **Red Hat Ceph Storage** on page 22
- **Optional Dell EMC Unity Storage**
Local Storage

When using local storage, each Compute node will host the ephemeral volumes associated with each virtual machine. Cinder will utilize LVMs that are shared by NFS for independent volumes, utilizing the local storage subsystem. With the hardware configuration for the Compute nodes see Table 13: Compute node Dell EMC PowerEdge R640 on page 53 using the eight (8) 600GB disks in a RAID 10, there will be approximately 2 TB of storage available.

Red Hat Ceph Storage

The Dell EMC Ready Architecture for Red Hat OpenStack Platform includes Red Hat Ceph Storage, which is a scale-out, distributed, software-defined storage system. Red Hat Ceph Storage is used as backend storage for Nova, Cinder and Glance. Storage nodes run the Red Hat Ceph Storage software, and Compute and Controller nodes run the Red Hat Ceph Storage block client.

Red Hat Ceph Storage also provides object storage for OpenStack VMs and other clients external to OpenStack. The object storage interface is an implementation of:

- The OpenStack Swift RESTful API (basic data access model)
- The Amazon S3 RESTful API

The object storage interface is provided by Red Hat Ceph Storage RADOS Gateway software running on the Controller nodes. Client access to the object storage is distributed across all Controller nodes in order to provide HA and IO load balancing.

Red Hat Ceph Storage is containerized in OSP13 (Mon, Mgr, Object Gateway, and Object Storage Daemon). Each OSD has an associated physical drive where the data is stored, and a journal where write operations are staged prior to being committed.

- When a client reads data from the Red Hat Ceph Storage cluster the OSDs fetch the data directly from the drives.
- When a client writes data to the storage cluster the OSDs write the data to their journals prior to committing the data.

OSD journals can be located in a separate partition on the same physical drive where the data is stored, or they can be located on a separate high-performance drive, such as an SSD optimized for write-intensive workloads. For the Architecture Guide, a recommended ratio of one (1) SSD to four (4) hard disks is used to achieve optimal performance. It should be noted that as of this writing, using a greater ratio will result in server performance degradation.

In a cost-optimized solution, the greatest storage density is achieved by forgoing separate SSD journal drives, and populating every available physical drive bay with a high-capacity HDD.

In a throughput-optimized solution, a few drive bays can be populated with high performance SSDs that will host the journals for the OSD HDDs. For example, in a Dell EMC PowerEdge R740xd system with 20 drive bays available for Red Hat Ceph Storage, 4 bays are used for SSD journal drives and 16 bays for HDD data drives. This is based upon the current industry guideline of a 4:1 HDD to SSD journal ratio.

Dell EMC Unity

One of the supported and validated Dell EMC storage options is:

- Dell EMC Unity Storage which can provide Glance (Image), Cinder (Block) and Manila (Shared FileSystems) Storage support.

Note: Other storage options are available upon request provided by Professional Services.
The Dell EMC Unity platforms accelerate application infrastructure with All-Flash unified storage platforms that simplify operations while reducing cost and datacenter footprint.

- Dell EMC Unity 350F, 450F, 550F, 650F (This RA has validated the Unity 350F and the 550F platforms)
- Unified storage with All-Flash performance
- Data reduction services for greater efficiency
- Install in minutes, manage from the cloud
- Scales up to 16 PB

Please see https://access.redhat.com/ecosystem/software/3172821 for more information about Red Hat OpenStack Platform certifications for Unity.

Please see https://docs.openstack.org/cinder/queens/configuration/block-storage/drivers/dell-emc-unity-driver.html for more information about Dell EMC Unity Cinder driver.

Please see https://docs.openstack.org/manila/queens/admin/emc_unity_driver.html for more information about the Dell EMC Unity Manila driver.

**Dell EMC SC Series Storage**

Another one of the supported and validated Dell EMC storage options is:

- Dell EMC SC Series Storage which now supports Glance (Image) storage via cinder in addition to Cinder (Block) Storage

  Note: Other storage options are available upon request provided by Professional Services.


Please see https://access.redhat.com/ecosystem/software/1610473 for more information about Red Hat OpenStack Platform certification for Dell EMC SC Series Storage.
Chapter 6

Network Architecture

Topics:

- Network architecture overview
- Infrastructure layouts
- Network components

This Architecture Guide supports consistency in rapid deployments through minimal network configuration.
Network architecture overview

The Dell EMC Ready Architecture for Red Hat OpenStack Platform uses the S5232F-ON as the Top-of-Rack (ToR) switches and the S3048-ON (S4048-ON optional) as the management switch.

Infrastructure layouts

The network consists of the following major network infrastructure layouts:

- **Core Network Infrastructure** - The connectivity of aggregation switches to the core for external connectivity.
- **Data Network Infrastructure** - The server NICs, ToR switches, and the aggregation switches.
- **Management Network Infrastructure** - The BMC management network, consisting of iDRAC ports and the out-of-band management ports of the switches, is aggregated into a 1-rack unit (RU) S3048-ON switch in one of the three racks in the cluster. This 1-RU switch in turn can connect to one of the aggregation or core switches to create a separate network with a separate VLAN.

Network components

The data network is primarily composed of the ToR and the aggregation switches. The following component blocks make up this network:

- **Server nodes** on page 25
- **Access switch or ToR** on page 26
- **Aggregation switches** on page 26
- **Core**
- **Layer-2 and Layer-3 Switching** on page 26
- **VLANs** on page 27
- **Management network services** on page 27
- **Dell EMC OpenSwitch solution** on page 27

Server nodes

In order to create a highly-available solution, the network must be resilient to loss of a single network switch, network interface card (NIC) or bad cable. To achieve this, the network configuration uses bonding across the servers and switches.

There are several types (or modes) of bonding, but only one is recommended for the Solution. The OpenStack Controller, Compute nodes, Red Hat Ceph Storage nodes, and Solution Admin Host can use:

- 802.3ad or LACP (mode = 4)

  **Note:** Other modes, such as balance-rr (mode=0), balance-xor (mode=2), broadcast (mode=3), balance-1lb (mode=5), and balance-alb (mode=6), are not supported. Please check with your technician for current support status of active-backup (mode = 1).

All nodes’ endpoints are terminated to switch ports that have been configured for LACP bonding mode across two S5232F-ON ToR switches for 25GbE/100GbE configured with a Virtual Link Trunking interconnect (VLTi) across them.

Please contact your Dell EMC sales representative for other viable options.
### Table 6: Bonding nodes supported

<table>
<thead>
<tr>
<th>Node Type</th>
<th>Bonding type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution Admin Host</td>
<td>802.3ad (LACP mode 4)</td>
</tr>
<tr>
<td>OpenStack Controller Nodes</td>
<td>Yes (solution default)</td>
</tr>
<tr>
<td>OpenStack Compute Nodes</td>
<td>Yes (solution default)</td>
</tr>
<tr>
<td>Red Hat Ceph Storage Nodes</td>
<td>Yes (solution default)</td>
</tr>
</tbody>
</table>

A single port is an option when bonding is not required. However, it is neither used nor validated in the Dell EMC Ready Architecture for Red Hat OpenStack Platform. The need to eliminate single points of failure is taken into consideration as part of the design, and this option has been eliminated wherever possible.

Please contact your Dell EMC sales representative for other configurations.

### Access switch or ToR

Dell EMC's recommended architecture uses VLT for HA between the two ToR switches, which enables the servers to terminate their Link Aggregation Group (LAG) interfaces (or bonds) into two different switches instead of one. This configuration enables active-active bandwidth utilization and provides redundancy within the rack if one ToR switch fails or requires maintenance. Dell EMC recommended ToR switch is:

- 10/25/100GbE connectivity – S5232F-ON

The ToR switches are responsible for providing the different network connections such as tenant networks, storage networks etc, between the different nodes of the OpenStack deployment such as compute, controller and storage nodes.

ToR switches are also configured to enable different network function virtualization (NFV) features such as OVS-DPDK, SR-IOV, DVR, VLAN-Aware VMs etc.

### Aggregation switches

Note: Please contact your Dell EMC sales representative for aggregation switch recommendation.

### Layer-2 and Layer-3 Switching

The layer-2 and layer-3 boundaries are separated at the aggregation layer.

The Architecture Guide uses layer-2 as the reference up to the aggregation layer, which is why VLT is used on the aggregation switches. The Red Hat OpenStack Director requires a layer-2 domain in order to provision servers.

The three network links - Provisioning, Storage, and Management - can have uplinks to a gateway device. The Provisioning network can use the Red Hat OpenStack Director as a proxy for pulling packages from a subscription server, or a gateway can be added. The Red Hat Ceph Storage, Dell EMC Unity or the SC Series Storage arrays on the Storage network may need access:

- From metrics and monitoring tools
- To enable management and updates

Note: For Dell EMC Unity Shared File Systems Storage (Manila), the Dell EMC Unity storage array requires additional access to networks. Options are External Network VLAN (Floating Network) or Internal Network VLAN for Tenants (Tenant Network) based on the use case. To support this functionality, administrators need to configure the Dell EMC Networking S5232F-ON ethernet ports to allow access for the Storage (untagged) and Floating/Tenant (tagged) VLANs.
There are many tools for Out of Band management, (OOB) for the iDRAC, which you can use after first adding the gateway to the network, and then updating the iDRAC.

The OpenStack Controllers are connected to a gateway device, usually a router or firewall. This device will handle routing for all networks external to the cluster. The required networks are:

- The floating IP range used by virtual machines
- A network for all external Public API and Graphical User Interface access

**VLANs**

This Architecture Guide implements at a minimum eight (8) separate Layer 2 VLANs:

- **External Network VLAN for Tenants** - Sets up a network that will support the floating IPs and default external gateway for tenants and virtual machines. This connection is through a router external to the cluster.
- **Internal Networks VLAN for Tenants** - Sets up the backend networks for Nova and the VMs to use.
- **Management/Out of Band (OOB) Network** - iDRAC connections can be routed to an external network. All OpenStack HA Controllers need direct access to this network for IPMI operations.
- **Private API Network VLAN** - Used for communication between OpenStack Controllers, the Red Hat OpenStack Director, and Compute nodes for Private API and cluster communications.
- **Provisioning Network VLAN** - Connects a NIC from all nodes into the fabric, used for setup and provisioning of the OpenStack servers.
- **Public API Network VLAN** - Sets up the network connection to a router that is external to the cluster. The network is used by the front-end network for routable traffic to individual VMs, access to the OpenStack API, RADOS Gateway, and the Horizon GUI. Depending upon the network configuration these networks may be either shared or routed, as needed. The Red Hat OpenStack Director requires access to the Public API Network.
- **Storage Clustering Network VLAN** - Used by all Storage nodes for replication and data checks (Red Hat Ceph Storage clustering).
- **Storage Network VLAN** - Used by all nodes for the data plane reads/writes to communicate to OpenStack Storage; setup, and provisioning of the Red Hat Ceph Storage cluster; and when included, the Dell EMC Unity Storage or SC Series Storage arrays.
- **Tenant Tunnel Network VLAN** - Used by Tenants for encapsulated networks such as, GRE or VXLAN tunnels, in place of the **Internal Networks VLAN for Tenants**.

**Management network services**

The management network and the provisioning network for all the servers and switches aggregate into a Dell EMC Networking S3048-ON switch.


The Management network services is used for several functions:

- The highly available software uses it to reboot and partition servers.
- An uplink to a router and an iDRAC configure a gateway monitoring the servers and gathering metrics.

Note: Discussion of this topic is beyond the scope of this document.

**Dell EMC OpenSwitch solution**

In addition to the Dell EMC switch-based Architecture Guide, Dell EMC provides an open standard that enables you to choose other brands and configurations of switches for your OpenStack environment.

The following list of requirements will enable other brands of switches to properly operate with Dell EMC's required tools and configurations:
Note: You are expected to ensure that the switches conform to these requirements, and that they are configured according to this Architecture Guide’s guidelines. These requirements are the minimums needed for a successful deployment and should be treated as such. Performance of the solution is directly affected by the performance of the Network.

- Support for IEEE 802.1Q VLAN traffic and port tagging
- Support for using one untagged, and multiple tagged VLANs, on the same port
- Support for using bonded interfaces as a single interface for TFTP/DHCP booting.
- Ability to provide a minimum of 96 x 25GbEs Ethernet ports in a non-blocking configuration within the Provisioning VLAN
  - Configuration can be a single switch or a combination of stacked switches to meet the additional requirements
- The ability to create LAGs with a minimum of two physical links in each LAG
- If multiple switches are stacked:
  - The ability to create a LAG across stacked switches
  - Full-bisection bandwidth
  - Support for VLANs to be available across all switches in the stack
- 250,000 packets-per-second capability per switch
- A managed switch that supports SSH and serial line configuration
- SNMP v3 support
Chapter 7

Network Function Virtualization (NFV) support

Topics:
- NUMA
- Huge Pages
- OVS-DPDK
- SR-IOV
- SR-IOV with Open vSwitch (OVS) Offload (Technology Preview)
- Distributed Virtual Router (DVR)

This Architecture Guide supports the following Network Function Virtualization (NFV) features:
- NUMA
- Huge Pages
- OVS-DPDK
- SR-IOV
- DVR

Note: All of the above NFV features can be enabled by using the JetPack 13.2 automation toolkit.
NUMA

The Dell EMC Ready Architecture for Red Hat OpenStack Platform version 13.2 provides the ability to enable NUMA optimization and CPU pinning support on all Nova compute nodes in the solution.

Non-Uniform Memory Access or NUMA allows multiple CPUs to share L1, L2, L3 caches, and main memory. The NUMA feature allows for precise placement and pinning down of programs and VM instances into specific CPUs. With that, the locality of memory accesses becomes highly predictable and does not have to contend for computational resources. These aspects help the programmer meet access-latency guarantees.

Note: NUMA can be enabled by the JetPack 13.2 automation toolkit.

Huge Pages

The Dell EMC Ready Architecture for Red Hat OpenStack Platform version 13.2 provides the ability to enable hugepage support on all Nova compute nodes in the solution.

If a program or a VM instance uses large data structures, then the use of small (i.e. 4KB) pages leads to the fragmentation of those data structures and loss of data locality. With Huge Pages of sizes 2MB and 1GB, programs and VM instances can improve data locality, thus leading to a higher performing and lower latency architecture.

Note: Hugepages can be enabled by the JetPack 13.2 automation toolkit.

OVS-DPDK

Open vSwitch (OVS) is a multilayer software/virtual switch used to interconnect virtual machines in the same host and between different hosts.

OVS makes use of the kernel for packet forwarding through a data path known as “fastpath” which consists of a simple flow table with action rules for the received packets. Exception packets or packets with no corresponding forwarding rule in the flow table are sent to the user space (slowpath). Switching between two memory spaces creates a lot of overhead, thus making the user space “slowpath”. User space makes a decision and updates the flow table in the kernel space accordingly so they can be used in the future.
As can be seen in the Figure 1, the OVS kernel module acts as a cache for the user space. And just like a cache, its performance decreases as the number of rules increase in the user space.

DPDK (Data-Plane Development Kit) eliminates packet buffer copies. It does this by running a dedicated poll-mode driver, and allocating hugepages for use as a packet buffer, then passing pointers to the packets. The elimination of copies leads to higher performance. OVS, when enabled to use DPDK-controller physical NIC interfaces, experiences a tremendous boost to packet delivery performance. It is also advantageous that both OVS and DPDK can operate in userspace, thus reducing kernel switches and improving packet processing efficiencies.

Note: OVS-DPDK can be enabled by the JetPack 13.2 automation toolkit.

**SR-IOV**

Single root I/O virtualization (SR-IOV) is an extension to the PCI Express (PCIe) specification. SR-IOV enables a single PCIe device to appear as multiple, separate virtual devices. Traditionally in a virtualized environment, a packet has to go through an extra layer of the hypervisor, that results in multiple CPU interrupts per packet. These extra interrupts cause a bottleneck in high traffic environments. SR-IOV enabled devices have the ability to dedicate isolated access to its resources among various PCIe hardware functions. These functions are later assigned to the virtual machines which allow direct memory access (DMA) to the network data.

By default, SR-IOV is not enabled in the Dell EMC PowerEdge R-Series system BIOS. When SR-IOV is deployed, virtual functions are are not created on the NIC interfaces.

Use the following steps to set the virtualization mode to SR-IOV and VF count to 64 in the Dell EMC PowerEdge R-Series system BIOS device settings for Mellanox NICs:

1. Enter system BIOS during boot by selecting F2.
2. Select Device Settings.
3. Select the Mellanox ConnectX-5 100GbE adapter.
4. Select the Device Level Configuration option.

Figure 2: Mellanox configuration

5. Change the virtualization mode from None to SR-IOV and the PCI Virtual Functions Advertised from the default setting eight to the desired value.

Note: Dell EMC has tested with a value of 64.

6. Repeat steps three through five for each Mellanox ConnectX-5 100GbE adapter where SR-IOV will be enabled.

7. Save BIOS settings and reboot the system.

Note: SR-IOV can be enabled by the JetPack 13.2 automation toolkit.

Note: SR-IOV Offload is a Red Hat OpenStack Platform Tech Preview feature

### SR-IOV with Open vSwitch (OVS) Offload (Technology Preview)

Open vSwitch (OVS) hardware offload is a technology preview and not recommended for production deployments. See Red Hat Scope of Coverage Details for Technology Preview

Open vSwitch hardware offload is a RHOSP technology preview feature which takes advantage of single root input/output virtualization (SR-IOV). An OVS software based solution is CPU intensive, affecting system performance and preventing full utilization of available bandwidth. This bottleneck can be addressed by making use of the OVS hardware offload capability, where the OVS data plane is moved to the underlying offloading capable smart NIC, while keeping the OVS control plane unmodified. This enables higher OVS performance without the associated CPU load.

In this Dell EMC Ready Architecture for Red Hat OpenStack Platform13, Dell EMC has validated this Technology Preview feature enabling 2 port or 4 port SR-IOV offload capability using Mellanox 100GbE ConnectX5 network adapters for tenant networks without bonding. Currently, bonding with SR-IOV offload capability is not supported.

Below is a logical diagram showing 2 port SR-IOV OVS Offload. For Mellanox 100GbE ConnectX-5 network adapters supporting ASAP², each PF defined on the Compute nodes must be mapped to a separate OVS bridge, and each OVS bridge will be mapped to unique physnet. This requires a corresponding OVS bridge for each PF in the Controller nodes. The SR-IOV NICs in the Controllers and Computes are connected to the same Tenant networks. If 4 port SR-IOV OVS Offload is desired, then adding additional Mellanox 100GbE ConnectX-5 network adapters is required.
Network Function Virtualization (NFV) support

Figure 3: OVS Offload

Note: Currently, NIC bonding with SR-IOV offload capability is not supported. Bonded NIC support for SR-IOV offload will be supported by the end of the Tech Preview.

Distributed Virtual Router (DVR)

Distributed Virtual Routing (DVR) offers an alternative routing design to the centralized routing model. It intends to isolate the failure domain of the Controller node and optimize network traffic by deploying the L3 agent and schedule routers on every Compute node. By eliminating a centralized layer 3 agent, the routing that was performed by single node (Master controller), is now distributed across the compute nodes using the local L3 agent. DVR follows the below routing flow rules:

- East-West traffic is routed directly on the Compute nodes in a distributed fashion.
- North-South traffic with floating IP is distributed and routed on the Compute nodes. This traffic requires the external network connects to each Compute node.
- North-South traffic without floating IP allocates and still needs a dedicated Controller node.
- The L3 agent on the Controller node is configured with a new dvr_snat mode so that the node serves only SNAT traffic.
- The neutron metadata agent distributes and deploys on all Compute nodes. The metadata proxy service hosts all the distributed routers.

Note: DVR can be enabled by the JetPack 13.2 automation toolkit.
Chapter 8

Additional OpenStack Features

Topics:
- Barbican
- Octavia
- Satellite

This Architecture Guide supports the following additional OpenStack features:
- Barbican
- Octavia
- Satellite Support

Note: These features can be enabled by using the JetPack 13.2 automation toolkit.
Barbican

Barbican is the secrets manager for Red Hat OpenStack Platform 13. You can use the barbican API and command line to centrally manage the certificates, keys, and passwords used by OpenStack services.

Symmetric encryption keys include -
• Block Storage (cinder) volume encryption
• Ephemeral disk encryption
• Object Storage (swift) encryption

Asymmetric keys and certificates include -
• Glance image signing and verification

Note: Barbican can be enabled by the JetPack 13.2 automation toolkit.

Octavia

Octavia is the OpenStack Load-balancing as a Service (LBaaS) version 2 implementation for the Red Hat OpenStack platform. It accomplishes its delivery of load balancing services by managing a fleet of virtual machines, collectively known as amphorae, which it spins up on demand.

Load Balancing Methods -
• Round robin - Rotates requests evenly between multiple instances.
• Source IP - Requests from a unique source IP address are consistently directed to the same instance.
• Least connections - Allocates requests to the instance with the least number of active connections.

Note: Octavia can be enabled by the JetPack 13.2 automation toolkit.

Satellite

JetPack support for Red Hat Satellite 6.5 gives the user the ability to deploy from a satellite instance. All nodes within the Dell EMC Ready Architecture for Red Hat OpenStack Platform can register and pull required packages and container images from the satellite instance, rather than the Red Hat Content Delivery Network (CDN).

Note: Satellite support can be enabled by the JetPack 13.2 automation toolkit.
Chapter 9

Operational notes

Topics:

- High availability
- Service layout
- Deployment overview

This section provides a basic overview of several important system aspects.
High availability

In order for the solution to be ready for production, different systems need to be fault-tolerant. The Architecture Guide design utilizes both hardware-based and software-based redundancy. This includes, but is not limited to:

- Operating Systems are hosted on either a RAID 1 or RAID 10 hard drive set.
- Critical network connections from server to switch utilize network bonding.
- Multiple Controllers host the control plane services.
- Control plane services are made highly available utilizing `ha-proxy`, `Corosync`, `Pacemaker`, and/or native resiliency.
- Red Hat Ceph Storage utilizes a minimum of three (3) servers.
- Red Hat Ceph Storage is used with either replication or erasure coding.
- Optional: Instance High Availability

This validated option utilizes remote `pacemaker` to monitor the Compute nodes. If preset criteria are met, the process of migrating instances off of the failing Compute nodes to others begins. If a Compute node completely fails, `pacemaker` can be configured to start the failed instances on different Compute nodes.

- Optional: Dell EMC Unity Storage or SC Series Storage arrays are highly available.

**Note:** The Solution Admin Host, and the servers hosted on it (Red Hat Ceph Storage Dashboard VM and Red Hat OpenStack Director), are not fault tolerant, but are not required for continued functionality of the OpenStack cluster.

Service layout

During the deployment each service configured by the Dell EMC Ready Architecture for Red Hat OpenStack Platform needs to reside upon a particular hardware type. For each server platform, two types of nodes have been designed:

- Dell EMC PowerEdge R640 or Dell EMC PowerEdge R740xd for Computes, Controllers, Solution Admin Hosts, or Infrastructure hardware type
- Dell EMC PowerEdge R740xd for Storage nodes.

Red Hat OpenStack Director is designed for flexibility, enabling you to try different configurations in order to find the optimal service placement for your workload. *Table 7: Overcloud: Node type to services* on page 37 presents the recommended layout of each service.

The Red Hat OpenStack Director and the Red Hat Ceph Storage Admin are deployed to the Solution Admin Host as individual VMs. This enables each tool to control its respective resources.

<table>
<thead>
<tr>
<th>Node to deploy</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution Admin Host (KVM)</td>
<td>Red Hat OpenStack Director</td>
</tr>
<tr>
<td>Solution Admin Host (KVM)</td>
<td>Red Hat Ceph Storage Dashboard</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Cinder-scheduler</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Cinder-volume</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Database-server</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Glance-Image</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>HAproxy (Load Balancer)</td>
</tr>
</tbody>
</table>
### Deployment overview

This is an overview of the deployment process that can be utilized for planning purposes:

1. **Hardware Setup:**
   - Rack and stack
   - Cabling
   - iDRAC setup
   - PXE NIC configuration
   - Server BIOS and RAID configuration
   - Switch configuration

2. **Software Setup:**
   - Deploy Solution Admin Host for provisioning services:
     - Deploy Red Hat Ceph Storage Dashboard VM to the Solution Admin Host
     - Deploy Red Hat OpenStack Director Virtual Server VM to the Solution Admin Host
   - Discover nodes
   - Import discovered nodes into Red Hat OpenStack Director
   - Configure Overcloud files

---

3. The number of OSDs that can be supported with 3 Controller nodes is listed at 1,000 ([https://access.redhat.com/articles/1548993](https://access.redhat.com/articles/1548993))

4. Available through a custom Services engagement.

---

<table>
<thead>
<tr>
<th>Node to deploy</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenStack Controllers</td>
<td>Heat</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Keystone-server</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Neutron-server</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Nova-Controller</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Nova dashboard-server</td>
</tr>
<tr>
<td>Three or more Compute Nodes</td>
<td>Nova-multi-compute</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Pacemaker</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>RabbitMQ-server (Messaging)</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Barbican (Secure management of secrets)</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Octavia (LBaaS)</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Red Hat Ceph Storage RADOS Gateway</td>
</tr>
<tr>
<td>OpenStack Controllers</td>
<td>Red Hat Ceph Storage Monitor</td>
</tr>
<tr>
<td>Three or more Storage Servers</td>
<td>Red Hat Ceph Storage (Block)</td>
</tr>
</tbody>
</table>

**Optional Services**

- Dell EMC Unity Storage
- Dell EMC SC Series Storage
- Dell EMC SC Series Storage Enterprise Manager Server
• Provision Overcloud
• Validate all nodes' networking
• Post-deployment, including but not limited to:
  • Enabling fencing
  • Enabling local storage for ephemeral

3. Environment Tests
• Tempest can be used to validate the deployment. At minimum the following tests should be performed:
  • Project Creation
  • User Creation
  • Network Creation
  • Image upload and launch
  • Floating IP Assignment
  • Basic network testing
  • Volume creation and attachment to VM
  • Object storage upload, retrieval and deletion
  • Deletion of all artifacts created during validation.
Chapter 10

Solution architecture

Topics:
- Solution common settings
- Solution with 25GbE/100GbE networking overview

This core architecture provides prescriptive guidance and recommendations, jointly engineered by Dell EMC and Red Hat, for deploying Dell EMC Ready Architecture for Red Hat OpenStack Platform version 13 with Dell EMC infrastructure.

The goals are to:
- Provide practical system design guidance and recommended configurations
- Develop tools to use with OpenStack for day-to-day usage and management
- Develop networking configurations capable of supporting your production system

The development of this architecture builds upon the experience and engineering skills of Dell EMC and Red Hat, and encapsulates best practices developed in numerous real-world deployments. The designs and configurations in this architecture have been tested in Dell EMC and Red Hat labs to verify system functionality and operational robustness.

The solution consists of the components shown in Figure 4: Solution with 25GbE/100GbE, Red Hat Ceph Storage Cluster, optional Dell EMC Unity Storage and optional SC Series Storage on page 45, and represents the base upon which all optional components and expansion of the Dell EMC Ready Architecture for Red Hat OpenStack Platform are built.
Solution common settings

Many settings are common through the Solution. The configurations that are tested are outlined in this section.

Solution Admin Host (SAH) networking

The Solution Admin Host is configured for 25GbE with the server internal bridged networks for the Virtual Machines. It is physically connected to the following networks:

- **Management Network** — used by the Red Hat OpenStack Director for iDRAC control of all Overcloud nodes.
- **Private API Network** — Used by the Red Hat OpenStack Director to run Tempest tests against the OpenStack private API
- **Provisioning Network** — Used by the Red Hat OpenStack Director to service DHCP to all hosts, provision each host, and act as a proxy for external network access
- **Public API Network** — used for:
  - **Inbound Access**
    - HTTP/HTTPS access to the Red Hat OpenStack Director
    - HTTP/HTTPS access to the Red Hat Ceph Storage Dashboard VM
    - Optional - SSH Access to the Red Hat OpenStack Director and Red Hat Ceph Storage Dashboard VM
  - **Outbound Access**
    - HTTP/HTTPS access for Red Hat Ceph Storage, RHEL, and RHOSP subscriptions.
    - Used by the Red Hat OpenStack Director to run Tempest tests using the OpenStack public API.
- **Storage Network** — Used by the Red Hat Ceph Storage Dashboard VM to monitor the Red Hat Ceph Storage Cluster.

Node type 802.1q tagging information

The solution is designed with the idea that different network traffic should be segregated from other traffic. This is accomplished by utilizing 802.1q VLAN Tagging for the different segments. The tables Table 8: OpenStack Node Type to Network 802.1q Tagging on page 41, Table 9: OpenStack Compute for xSP and CSP Profile to Network 802.1q Tagging on page 42, and Table 10: Storage Node Type to Network 802.1q Tagging on page 42 summarize this. This segregation is independent of network speed and used by 25GbE configuration.

<table>
<thead>
<tr>
<th>Network</th>
<th>Solution Admin Host</th>
<th>OpenStack controller</th>
<th>Red Hat Ceph Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Network VLAN for Tenants (Floating IP Network)</td>
<td>Not Connected</td>
<td>Connected, tagged</td>
<td>Not Connected</td>
</tr>
<tr>
<td>iDRAC physical connection to the Management/OOB VLAN</td>
<td>Connected, Untagged</td>
<td>Connected, Untagged</td>
<td>Connected, Untagged</td>
</tr>
<tr>
<td>Internal Networks VLAN for Tenants</td>
<td>Not Connected</td>
<td>Connected, Tagged</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Management/OOB Network VLAN</td>
<td>Connected, Tagged</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Private API Network VLAN</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Network</td>
<td>Solution Admin Host</td>
<td>OpenStack controller</td>
<td>Red Hat Ceph Storage</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Provisioning VLAN</td>
<td>Connected, Tagged</td>
<td>Connected, Untagged</td>
<td>Connected, Tagged</td>
</tr>
<tr>
<td>Public API Network VLAN</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Storage Clustering VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
<td>Connected, Tagged</td>
</tr>
<tr>
<td>Storage Network VLAN</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
</tr>
<tr>
<td>Tenant Tunnel Network</td>
<td>Not Connected</td>
<td>Connected, Tagged</td>
<td>Not Connected</td>
</tr>
</tbody>
</table>

**Table 9: OpenStack Compute for xSP and CSP Profile to Network 802.1q Tagging**

<table>
<thead>
<tr>
<th>Network</th>
<th>xSP OpenStack compute</th>
<th>CSP - OpenStack compute NFV</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Network VLAN for Tenants (Floating IP Network)</td>
<td>Not Connected</td>
<td>Connected, Tagged</td>
</tr>
<tr>
<td>iDRAC physical connection to the Management/OOB VLAN</td>
<td>Connected, Untagged</td>
<td>Connected, Untagged</td>
</tr>
<tr>
<td>Internal Networks VLAN for Tenants</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
</tr>
<tr>
<td>Management/OOB Network VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Private API Network VLAN</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
</tr>
<tr>
<td>Provisioning VLAN</td>
<td>Connected, Untagged</td>
<td>Connected, Untagged</td>
</tr>
<tr>
<td>Public API Network VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Storage Clustering VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Storage Network VLAN</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
</tr>
<tr>
<td>Tenant Tunnel Network</td>
<td>Connected, Tagged</td>
<td>Connected, Tagged</td>
</tr>
</tbody>
</table>

**Table 10: Storage Node Type to Network 802.1q Tagging**

<table>
<thead>
<tr>
<th>Network</th>
<th>Dell EMC Unity</th>
<th>Dell EMC SC Series Storage Enterprise Manager</th>
<th>Dell EMC SC Series Storage array</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Network for Tenants VLAN (Floating IP Network)</td>
<td>Connected, Tagged</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>iDRAC physical connection to the Management/OOB VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Internal Networks VLAN for Tenants</td>
<td>Connected, Tagged</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Management/OOB Network VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
</tbody>
</table>
## Solution Red Hat Ceph Storage configuration

The Red Hat Ceph Storage cluster provides data protection through replication, block device cloning, and snapshots. By default the data is striped across the entire cluster, with three replicas of each data entity. The number of storage nodes in a single cluster can scale to hundreds of nodes and many petabytes in size.

Red Hat Ceph Storage considers the physical placement (position) of storage nodes within defined fault domains (i.e., rack, row, and data center) when deciding how data is replicated. This reduces the probability that a given failure may result in the loss of more than one data replica.

The Red Hat Ceph Storage cluster services include:

- **Ceph Dashboard** — Ceph web based monitoring tool hosted on a VM.
- **RADOS Gateway** — Object storage gateway.
- **Object Storage Daemon (OSD)** — Running on storage nodes, the OSD serves data to the Red Hat Ceph Storage clients from disks on the storage nodes. Generally, there is one OSD process per disk drive.
- **Monitor (MON)** — Running on Controller nodes, the MON process is used by the Red Hat Ceph Storage clients and internal Red Hat Ceph Storage processes, to determine the composition of the cluster and where data is located. There should be a minimum of three MON processes for the Red Hat Ceph Storage cluster. The total number of MON processes should be odd.
- **Ceph Manager Daemon (ceph-mgr)** — Running on Controller nodes alongside the MON processes, it provides additional monitoring and interfaces to external monitoring and management systems.

**Note:** If MON processes on Controller nodes become a bottleneck, then additional MON processes can be added to the cluster by using dedicated machines, or by starting MON processes on storage Nodes. A custom Services engagement can be arranged; please contact your Dell EMC sales representative for assistance.

The Storage Network VLAN is described in the Red Hat Ceph Storage documentation as the public network. The Storage Cluster Network VLAN is described in the Red Hat Ceph Storage documentation as the cluster network.

A supported distribution by Red Hat with production level support of Ceph is used in this solution: Red Hat Ceph Storage 3.2, which also includes the Red Hat Ceph Storage Dashboard VM. The Red Hat Ceph Storage Dashboard also includes Red Hat Ceph Storage troubleshooting and servicing tools and utilities. Red Hat Ceph Storage Dashboard is installed on a virtual machine that runs on the Solution Admin Host (SAH). Note that:

- The SAH must have access to the Controller and Storage nodes through the Private API Access VLAN in order to manage Red Hat Ceph Storage; and for the monitoring process on all Storage nodes to return status and performance telemetry.

### Network Configuration Table

<table>
<thead>
<tr>
<th>Network</th>
<th>Dell EMC Unity</th>
<th>Dell EMC SC Series Storage Enterprise Manager</th>
<th>Dell EMC SC Series Storage array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provisioning VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Private API Network VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Public API Network VLAN</td>
<td>Not Connected</td>
<td>Connected, Untagged</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Storage Network VLAN</td>
<td>Connected, Untagged</td>
<td>Connected, Untagged</td>
<td>Connected, Untagged</td>
</tr>
<tr>
<td>Storage Clustering VLAN</td>
<td>Not Connected</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
<tr>
<td>Tenant Tunnel Network</td>
<td>Not Connected</td>
<td>Not Connected</td>
<td>Not Connected</td>
</tr>
</tbody>
</table>
The Controller nodes must have access to the Storage nodes through the Storage Network VLAN in order for the MON processes on the Controller nodes to be able to query the Red Hat Ceph Storage MON processes, for the cluster state and configuration.

The Compute nodes must have access to the Storage nodes through the Storage Network VLAN in order for the Red Hat Ceph Storage client on that node to interact with the storage nodes, OSDs, and the Red Hat Ceph Storage MON processes.

The Storage nodes must have access to the Storage Network VLAN, as previously stated, and to the Storage Cluster Network VLAN.

Solution with 25GbE/100GbE networking overview

Since the Solution is designed for a production environment, key OpenStack services are made highly available (HA) by clustering the OpenStack controller nodes. The networking is based upon 25GbE bonds for data networks, and the network switches are configured for HA. The Out of Band Management (iDRAC's) network is not HA, and is 1GbE. The 100GbE networking is used in the solution for the user/tenant traffic.

For basic hardware configuration refer to Bill of Materials on page 51.

Solution 25GbE/100GbE rack layout

The Solution includes three (3) storage nodes, configured in a Red Hat Ceph Storage cluster, which is tied into Cinder, Glance, and Nova.

See Table 11: Solution Admin Host (SAH) Dell EMC PowerEdge R640 on page 52, Table 12: Controller node Dell EMC PowerEdge R640 on page 52, Table 13: Compute node Dell EMC PowerEdge R640 on page 53, and Table 14: Storage node Dell EMC PowerEdge R740xd on page 53 for hardware configurations. The Solution includes:

- Node 1: Solution Admin Host with the Red Hat OpenStack Director and the Red Hat Ceph Storage Dashboard installed.
- Nodes 2 - 4: Dell EMC PowerEdge R640 OpenStack Controllers
- Nodes 5 - 7 Dell EMC PowerEdge R640 Nova Compute Nodes
- Nodes 8 - 10: Dell EMC PowerEdge R740xd Storage Nodes
- Network Switches: Two (2) Dell EMC Networking S5232F-ON, and one (1) Dell EMC Networking S3048-ON

Note: The following rack is not to scale but shows the node types and usage.
Figure 4: Solution with 25GbE/100GbE, Red Hat Ceph Storage Cluster, optional Dell EMC Unity Storage and optional SC Series Storage
**Solution 25GbE/100GbE network configuration**

The network for the Dell EMC Ready Architecture for Red Hat OpenStack Platform has been designed to support production-ready servers with a highly available network configuration.

The node type will determine how the switches are configured for delivering the different networks. *Table 8: OpenStack Node Type to Network 802.1q Tagging* on page 41 and *Table 10: Storage Node Type to Network 802.1q Tagging* on page 42 outline the networks to the node types.

For the CSP Profile with non OVS Offload, this is what the logical network will look like:

![Figure 5: 25GbE/100GbE cluster network logical architecture for CSP profile](image)

For the CSP Profile with OVS Offload, this is what the logical network will look like:
Figure 6: 25GbE/100GbE cluster network logical architecture for CSP OVS Offload profile
Chapter 11

Professional Services

Topics:
• Consulting services

Dell EMC Professional Services understands the complexity and challenges in an accelerated deployment, integration and operations of Red Hat OpenStack Platform. At the core, the challenges are not solely dealing with new technology – but it also includes people and process changes needed to be successful. Over the years Dell EMC Consulting has led hundreds of successful cloud implementations.

Dell EMC’s Consulting is well equipped in tackling the challenges in implementing and integrating Red Hat OpenStack Platform and Network functions virtualization infrastructure, (NFvi) technologies. Furthermore, Dell EMC has a robust, comprehensive portfolio focused on Red Hat OpenStack Platform based solutions that have evolved from our heritage in enterprise cloud solutions – and is specifically tailored to the unique and demanding Telco/Cloud Service Provider requirements for NFVi implementations. The figure below details the core service from the Dell EMC Services portfolio that our customers can leverage to support Red Hat OpenStack Platform and NFVi initiatives.
Consulting services

Consulting Services Portfolio
Dell EMC Consulting Services

Dell EMC Consulting Portfolio for Dell EMC Ready Architecture for Red Hat OpenStack Platform

**Strategy**
- Analyze business objectives, prioritize use cases & develop business case (TCO/ROI) for transformation
- Assess infrastructure and operational readiness for the insertion of NFVI, solution

**Design**
- Develop an optimized end-to-end design based on Dell EMC Ready Solutions and Industry best practice
- Provide guidance on technology options and trade-offs

**Implement**
- Implement NFV without impacting service reliability
- Plans for scaling the NFV infrastructure across network
- Testing and validating NFV implementations

**Operate**
- Onboarding VNF workload
- Aligning Organization and Processes
- Automate Lifecycle management, Netops for Automated Configuration, Placement, and Operations (CI/CT/CD)
- Operational services including Residency and out tasking services.

Figure 7: Dell EMC Services for RHOSP and NFVi Ready Architecture

- Dell EMC Red Hat OpenStack Platform, NFVi Strategy Workshop – Dell EMC NFV/SDN/Openstack workshop is intended to help customers who are starting their cloud transformation journey to develop a comprehensive NFV, RHOSP adoption strategy.

  It helps customers prioritize the use cases based on business drivers, understand the technology and platform choices such as VMware VIO, Red Hat, etc., identify infrastructure virtualization opportunities, develop strategic roadmap for implementation phases.

  This includes understanding current state, understand requirements, GAP analysis and collaborative services to define technology, process implementation paths to ensure successful transformation.

- Dell EMC Red Hat OpenStack Platform, NFVi Business Advisory – This service is intended to help customers develop a business case to adopt NFV based business drivers. It helps to demonstrate to senior management the benefits of moving towards a shared NFV, RHOSP cloud infrastructure.

  This services includes an independent and collaborative review of the business objectives to develop a comprehensive business case for the transformation including the Total cost of ownership/ Return on Investment financial model.

- Dell EMC Red Hat OpenStack Platform, NFVi Design Services – Dell EMC will help the customers develop a Low Level Design (LLD) and High Level Design (HLD) based on customer requirements and use cases. Dell EMC consultants will work with the customer to gather data, use cases and plans for the OpenStack, NFVI deployment.

  Based on that, the team will help develop the end-to-end design that can be used for implementation of the RHOSP ready architecture leveraging JetPack automation.

- Dell EMC Red Hat OpenStack Platform, NFVi Implementation Services – Dell EMC Ready Architecture for Red Hat OpenStack Platform and solutions offer a risk free option for our customers.

  To ensure faster time to value Dell EMC Implementation services provide on or off-site infrastructure planning and implementation of the RHOSP ready architecture leveraging JetPack automation.
• Dell EMC Red Hat OpenStack Platform, NFVi Residency/Operate Services – Dell EMC will help customers operate the new cloud infrastructure by providing resources as on-site/off-site residents. These residents will manage the lifecycle of the cloud and also assist with automation and onboarding of workload to the OpenStack cloud.

To engage Dell EMC consulting services, please contact your Dell EMC sales representative or send an email.
Appendix A

Bill of Materials

Topics:

• Bill of Materials for Dell EMC PowerEdge R-Series Solution - Mellanox NIC’s
• Bill of Materials for Dell EMC PowerEdge R-Series Solution - Intel NIC’s
• Nodes Overview
• Subscriptions and Network Switches in the Solution

This guide provides Bill of Material information necessary to purchase the proper hardware to deploy the Dell EMC Ready Architecture for Red Hat OpenStack Platform.

Note: For cable, racks, power please contact your Dell EMC support representative.
Bill of Materials for Dell EMC PowerEdge R-Series Solution - Mellanox NIC's

The base Dell EMC PowerEdge R-Series solution is comprised of:

- 1 Dell EMC PowerEdge R640 Solution Admin Host (SAH) node
- 3 Dell EMC PowerEdge R640 controller nodes
- 3 Dell EMC PowerEdge R640 compute nodes
- 3 Dell EMC PowerEdge R740xd storage nodes

Base configuration - Mellanox NIC's

Table 11: Solution Admin Host (SAH) Dell EMC PowerEdge R640

<table>
<thead>
<tr>
<th>Machine function</th>
<th>SAH node</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R640 (1 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel® Xeon® Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>1 x Mellanox 25GbE 2P ConnectX4LX SFP RNDC</td>
</tr>
<tr>
<td>Add-in Network</td>
<td>1 x Mellanox 25GbE 2P ConnectX4LX SFP Adapter</td>
</tr>
<tr>
<td>Disk</td>
<td>8 x 600GB 10k SAS 12Gbps</td>
</tr>
<tr>
<td>Storage Controller</td>
<td>PERC H740 Mini controller</td>
</tr>
<tr>
<td>RAID</td>
<td>RAID 10</td>
</tr>
</tbody>
</table>

Table 12: Controller node Dell EMC PowerEdge R640

<table>
<thead>
<tr>
<th>Machine function</th>
<th>Controller nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R640 (3 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel® Xeon® Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>1 x Mellanox 25GbE 2P ConnectX4LX SFP RNDC</td>
</tr>
<tr>
<td>Add-in Network</td>
<td>1 x Mellanox 25GbE 2P ConnectX4LX SFP Adapter</td>
</tr>
<tr>
<td></td>
<td>2 x Mellanox 100GbE 2P ConnectX5 SFP Adapter (Used for CSP Profile)</td>
</tr>
<tr>
<td>Disk</td>
<td>8 x 600GB 10k SAS 12Gbps</td>
</tr>
<tr>
<td>Storage Controller</td>
<td>PERC H740 Mini controller</td>
</tr>
<tr>
<td>RAID</td>
<td>RAID 10</td>
</tr>
</tbody>
</table>
### Table 13: Compute node Dell EMC PowerEdge R640

<table>
<thead>
<tr>
<th>Machine function</th>
<th>Compute nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R640 (3 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel® Xeon® Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>1 x Mellanox 25GbE 2P ConnectX4LX SFP RNDC</td>
</tr>
</tbody>
</table>
| Add-in Network   | Option 1: for xSP  
1 x Mellanox 25GbE 2P ConnectX4LX SFP Adapter  
Option 2: for CSP  
1 x Mellanox 25GbE 2P ConnectX4LX SFP  
2 x Mellanox 100GbE 2P ConnectX5 SFP Adapter |
| Disk             | 8 x 600GB 10k SAS 12Gbps |
| Storage Controller | PERC H740 Mini controller |
| RAID             | RAID 10 |

### Table 14: Storage node Dell EMC PowerEdge R740xd

<table>
<thead>
<tr>
<th>Machine function</th>
<th>Storage nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R740xd (3 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel(R) Xeon(R) Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>1 x Mellanox 25GbE 2P ConnectX4LX SFP RNDC</td>
</tr>
<tr>
<td>Add-in Network</td>
<td>1 x Mellanox 25GbE 2P ConnectX4LX SFP Adapter</td>
</tr>
</tbody>
</table>
| Disk             | Option 1:  
OSD and Journal Drives separate  
Front Drives: 12 x 2.4TB 10K SAS  
Mid Bay drive: 4 x 2.4TB 10K SAS  
Flex Bay drive: 4 x 400GB SAS SSD  
Option 2:  
OSD and Journal colocated  
Front drives: 24 x 1.8TB 10K SAS  
Mid Bay drives: 4 x 1.8TB 10K SAS  
Flex Bay drives: 4 x 1.8TB 10K SAS  
Option 3:  
NVMe Front Drives:  
12 x 1.2TB Intel NVMe  
2 x 1.8 TB 10K SAS |
| Storage Controller | HBA 330  
BOSS Cntrl + 2 M.2  
240G,R1,FH  
HBA 330  
BOSS Cntrl + 2 M.2  
240G,R1,FH  
PERC H740 Mini controller |
### Machine function | Storage nodes
--- | ---
RAID | Operating System on BOSS in RAID 1 Configuration Pass through each data disk

**Note:** Be sure to consult your Dell EMC account representative before changing the recommended hardware configurations.

## Bill of Materials for Dell EMC PowerEdge R-Series Solution - Intel® NIC’s

The base Dell EMC PowerEdge R-Series solution is comprised of:
- 1 Dell EMC PowerEdge R640 Solution Admin Host (SAH) node
- 3 Dell EMC PowerEdge R640 controller nodes
- 3 Dell EMC PowerEdge R740xd compute nodes
- 3 Dell EMC PowerEdge R740xd storage nodes

### Base configuration - Intel® NIC's

#### Table 15: Solution Admin Host (SAH) Dell EMC PowerEdge R640

<table>
<thead>
<tr>
<th>Machine function</th>
<th>SAH node</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R640 (1 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel® Xeon® Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>N/A</td>
</tr>
<tr>
<td>Add-in Network</td>
<td>2 x Intel® XXV710 DP 25GbE DA/SFP+ Adapter</td>
</tr>
<tr>
<td>Disk</td>
<td>8 x 600GB 10k SAS 12Gbps</td>
</tr>
<tr>
<td>Storage Controller</td>
<td>PERC H740 Mini controller</td>
</tr>
<tr>
<td>RAID</td>
<td>RAID 10</td>
</tr>
</tbody>
</table>

#### Table 16: Controller node Dell EMC PowerEdge R640

<table>
<thead>
<tr>
<th>Machine function</th>
<th>Controller nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R640 (3 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel(R) Xeon(R) Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>N/A</td>
</tr>
<tr>
<td>Add-in Network</td>
<td>2 x Intel® XXV710 DP 25GbE DA/SFP+ Adapter</td>
</tr>
<tr>
<td>Disk</td>
<td>8 x 600GB 10k SAS 12Gbps</td>
</tr>
<tr>
<td>Storage Controller</td>
<td>PERC H740 Mini controller</td>
</tr>
<tr>
<td>RAID</td>
<td>RAID 10</td>
</tr>
</tbody>
</table>
### Table 17: Compute node Dell EMC PowerEdge R740

<table>
<thead>
<tr>
<th>Machine function</th>
<th>Compute nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R740 (3 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel® Xeon® Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>N/A</td>
</tr>
</tbody>
</table>
| Add-in Network   | Option 1: for xSP  
  2 x Intel® XXV710 DP 25GbE DA/SFP+ Adapter  
  Option 2: for CSP  
  4 x Intel® XXV710 DP 25GbE DA/SFP+ Adapter |
| Disk             | 8 x 600GB 10k SAS 12Gbps |
| Storage Controller | PERC H740 Mini controller |
| RAID             | RAID 10 |

### Table 18: Storage node Dell EMC PowerEdge R740xd

<table>
<thead>
<tr>
<th>Machine function</th>
<th>Storage nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Dell EMC PowerEdge R740xd (3 qty)</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel® Xeon® Gold 6230 CPU @ 2.10GHz</td>
</tr>
<tr>
<td>RAM (Minimum)</td>
<td>192GB DDR-4 2933 MHz</td>
</tr>
<tr>
<td>LOM</td>
<td>N/A</td>
</tr>
<tr>
<td>Add-in Network</td>
<td>2 x Intel® XXV710 DP 25GbE DA/SFP+ Adapter</td>
</tr>
</tbody>
</table>
| Disk             | Option 1:  
  OSD and Journal Drives separate  
  Front Drives: 12 x 2.4TB 10K SAS  
  Mid Bay drive: 4 x 2.4TB 10K SAS  
  Flex Bay drive: 4 x 400GB SAS SSD  
  Option 2:  
  OSD and Journal colocated  
  Front drives: 24 x 1.8TB 10K SAS  
  Mid Bay drives: 4 x 1.8TB 10K SAS  
  Flex Bay drives: 4 x 1.8TB 10K SAS  
  Option 3:  
  NVMe Front Drives:  
  12 x 1.2TB Intel NVMe  
  2 x 1.8 TB 10K SAS  |
| Storage Controller | HBA 330  
  BOSS Ctrl + 2 M.2 240G,R1,FH  
  HBA 330  
  BOSS Ctrl + 2 M.2 240G,R1,FH  
  PERC H740 Mini controller |
### Machine function | Storage nodes
--- | ---
RAID | Operating System on BOSS in RAID 1 Configuration Pass through each data disk

**Note:** When using Intel® NIC’s, OpenvSwitch (OVS) hardware offloading is not supported. All other NFV features documented in this Architecture Guide are supported. Be sure to consult your Dell EMC account representative before changing the recommended hardware configurations.

### Nodes Overview

The minimum hardware needed is:
- 1 Solution Admin Host (SAH)
- 3 Controller nodes
- 3 Compute nodes
- 3 Storage servers

Please consult with your Dell EMC sales representative to ensure proper preparation and submission of your hardware and software orders.

### Subscriptions and Network Switches in the Solution

A Dell EMC sales representative will determine the correct software subscriptions needed for the Dell EMC Ready Architecture for Red Hat OpenStack Platform and Dell EMC Networking OS10 subscriptions.

Required subscriptions:
- Red Hat OpenStack Platform
- Red Hat Ceph Storage
- Dell EMC Networking OS10
- Red Hat Satellite - Optional

**Note:** Please contact your Dell EMC sales representative.

### Default network switch - Dell EMC Networking S3048-ON switch

**Table 19: S3048-ON switch**

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3048-ON</td>
<td>48 line-rate 1000BASE-T ports, 4 line-rate 10GbE SFP+ ports (1 qty)</td>
</tr>
</tbody>
</table>
| Redundant power supplies | AC Power Supply  
\ or 
DC Power Supply |
| Fans                | Fan Module I/O Panel to PSU Airflow  
\ or 
Fan Module PSU to I/O Panel Airflow |
<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validated operating systems</td>
<td>Dell EMC Networking OS10</td>
</tr>
</tbody>
</table>

**Dell EMC Networking S4048-ON optional switch**

Table 20: S4048-ON switch

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4048-ON</td>
<td>48x 10GbE SFP+, 6x QSFP+ (1 qty) - optional</td>
</tr>
<tr>
<td>Redundant power supplies</td>
<td>AC Power Supply</td>
</tr>
<tr>
<td></td>
<td>or</td>
</tr>
<tr>
<td></td>
<td>DC Power Supply</td>
</tr>
<tr>
<td>Fans</td>
<td>Fan Module I/O Panel to PSU Airflow</td>
</tr>
<tr>
<td></td>
<td>or</td>
</tr>
<tr>
<td></td>
<td>Fan Module PSU to I/O Panel Airflow</td>
</tr>
<tr>
<td>Validated operating systems</td>
<td>Dell EMC Networking OS10</td>
</tr>
</tbody>
</table>

**Dell EMC Networking S5232F-ON switch**

Table 21: S5232F-ON

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S5232F-ON</td>
<td>100GbE, 40GbE, and 25GbE (2 qty)</td>
</tr>
<tr>
<td>Redundant power supplies</td>
<td>AC Power Supply</td>
</tr>
<tr>
<td></td>
<td>or</td>
</tr>
<tr>
<td></td>
<td>DC Power Supply</td>
</tr>
<tr>
<td>Fans</td>
<td>Fan Module I/O Panel to PSU Airflow</td>
</tr>
<tr>
<td></td>
<td>or</td>
</tr>
<tr>
<td></td>
<td>Fan Module PSU to I/O Panel Airflow</td>
</tr>
<tr>
<td>Validated operating systems</td>
<td>Dell EMC Networking OS10</td>
</tr>
</tbody>
</table>
Appendix B

References

Topics:

- To learn more


Note: If you need additional services or implementation help, please contact your Dell EMC sales representative.
To learn more


For more information on Dell EMC Service Provider Solutions, visit https://www.dell EMC.com/en-us/service-providers/index.htm

Copyright © 2014-2019 Dell EMC or its subsidiaries. All rights reserved. Trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Specifications are correct at date of publication but are subject to availability or change without notice at any time. Dell EMC and its affiliates cannot be responsible for errors or omissions in typography or photography. Dell EMC’s Terms and Conditions of Sales and Service apply and are available on request. Dell EMC service offerings do not affect consumer’s statutory rights.

Dell EMC, the DELL EMC logo, the DELL EMC badge, and PowerEdge are trademarks of Dell EMC
Glossary

API
Application Programing Interface is a specification that defines how software components can interact.

BMC/iDRAC Enterprise
Baseboard management controller. An on-board microcontroller that monitors the system for critical events by communicating with various sensors on the system board, and sends alerts and log events when certain parameters exceed their preset thresholds.

BOSS
The Boot Optimized Storage Solution (BOSS) enables customers to segregate operating system and data on server-internal storage. This is helpful in the Hyper-Converged Infrastructure (HCI) and Software Defined Storage (SDS) arenas, to separate operating system drives from data drives, and implement hardware RAID mirroring (RAID1) for OS drives.

CDH
Cloudera Distribution for Apache Hadoop

Cloud computing
Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Cluster
A set of servers dedicated to OpenStack that can be attached to multiple distribution switches.

Compute node
The hardware configuration that best supports the hypervisor server or Nova compute roles.

DevOps
Development Operations (DevOps) is an operational model for managing data centers using improved automated deployments, shortened lead times between fixes, and faster mean time to recovery. See https://en.wikipedia.org/wiki/DevOps.
**DIMM**
Dual In-line Memory Module

**DNS**
The domain name system (DNS) defines how Internet domain names are located, and translated into Internet Protocol (IP) addresses.

**FQDD**
A fully qualified device descriptor (FQDD) is a method used to describe a particular component within a system or subsystem, and is used for system management and other purposes.

**FQDN**
A fully qualified domain name (FQDN) is the portion of an Internet Uniform Resource Locator (URL) that fully identifies the server to which an Internet request is addressed. The FQDN includes the second-level domain name, such as "dell.com", and any other levels as required.

**GUI**
Graphical User Interface - A visual interface for human interaction with the software, taking inputs and generating easy to understand visual outputs.

**Hypervisor**
Software that runs virtual machines (VMs).

**IaaS**
Infrastructure as a Service.

**Infrastructure node**
Systems that handle the control plane and deployment functions.

**ISV**
Independent Software Vendor.

**JBOD**
Just a Bunch of Disks
LAG
Link Aggregation Group.

LOM
LAN on motherboard.

LVM
Logical Volume Management

ML2
The Modular Layer 2 plug-in is a framework that allows OpenStack to utilize different layer 2 networking technologies.

NFS
The Network File System (NFS) is a distributed filesystem that allows a computer user to access, manipulate, and store files on a remote computer, as though they resided on a local file directory.

NIC
Network Interface Card

Node
One of the servers in the cluster.

NUMA
Non-Uniform Memory Access

Overcloud
The functional cloud that is available to run guest VMs and workloads.

Pod
An installation comprised of three racks, and consisting of servers, storage, and networking.

REST
REST - Representational State Transfer (also ReST). Relies upon stateless, client-server, cacheable communications protocol to access the API.
**RHOSP**

Red Hat OpenStack Platform

**RPC**

Remote Procedure Call

**SAH**

The Solution Admin Host (SAH) is a physical server that supports VMs for the Undercloud machines needed for the cluster to be deployed and operated.

**SDS**

Software-defined storage (SDS) is an approach to computer data storage in which software is used to manage policy-based provisioning and management of data storage, independent of the underlying hardware.

**SDN**

Software-defined Network (SDN) is where the software will define, create, use and destroy different networks as needed.

**Stamp**

A stamp is the collection of all servers and network switches in the solution.

**Storage Node**

The hardware configuration that best supports SDS functions such as Red Hat Ceph Storage.

**ToR**

Top-of-rack switch/router.

**U**

U used in the definition of the size of server, example 1U or 2U. A "U" is a unit of measure equal to 1.75 inches in height.

**Undercloud**

The Undercloud is the system used to control, deploy, and monitor the Overcloud - it is a single node OpenStack deployment completely under the administrators control. The Undercloud is not HA configured.
**VLT**

A Virtual Link Trunk (VLT) is the combined port channel between an attached device (ToR switch) and the VLT peer switches.

**VLTi**

A Virtual Link Trunk Interconnect (VLTi) is an interconnect used to synchronize states between the VLT peer switches. Both endpoints must be the same speed, i.e. 40Gb → 40Gb; 1G interfaces are not supported.

**VM**

Virtual Machine - a simulation of a computer system.