Dell Red Hat Cloud Solutions Reference Architecture Guide

A Dell Reference Architecture Guide

April 2, 2014

Contents

Tables

Figures

Trademarks

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.

Terminology and Abbreviations

<div align="center">Table 1: Terminology and Abbreviations</div>

Term	Meaning	Substitute term
BMC/IDRAC	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.	IPMI
Compute Node	The hardware configuration that bests supports the hypervisor server or Nova compute roles	Host, Box Unit
DevOps	An operational model for managing data centers using automated deployments.	Puppet™
Hypervisor	Software that runs virtual machines (VMs).	KVM, Xen, VMware, HyperV, QEMU, RHEV
iDRAC	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.	IPMI
Infrastructure Node	The hardware configuration that bests supports controller, hypervisor server, database and other non-storage type roles	Host, Box, Unit
LAG	Link Aggregation Group	
LOM	LAN on motherboard.	
Node	One of the servers in the system.	Host, Box, Unit
Storage Node	The hardware configuration that bests supports storage functions such as Red Storage Server, Swift, CEPH.	Host, Box, Unit

Overview

Approach

This reference architecture focuses on helping organizations begin evaluating OpenStack® software and/or create proofs of concept. Dell and Red Hat can provide guidance for more sophisticated deployments; however, they are beyond the scope of this document. The expected focus for the Dell Red Hat Cloud Solutions with Red Hat Enterprise Linux™ OpenStack Platform encompasses software, hardware, operations, and integration in unique packages known as bundles. See Solutions Sizing Terms on page 16 for more information.

This reference architecture advocates an operational approach based upon highly automated solution deployments, using the components of the Dell Red Hat cloud solutions with RHEL OpenStack Platform. We believe that this operational model is the best practice for both initial cloud evaluations, and long-term maintenance of both moderate-scale and hyperscale data centers.

OpenStack Maturity

The code base for OpenStack is evolving at a very rapid pace. The November 2013 OpenStack release is known as Havana. Please see the https://access.redhat.com/site/support/policy/updates/openstack/platform to find out about the solution.

The eighth release of OpenStack, Havana, builds upon previous releases, with more than 1000 contributors and nearly 400 new contributors, represented by the 20,000 commits. OpenStack is being deployed and tested 700 times per day. Please see http://www.slideshare.net/openstack/openstack-havana-release for verification.

We designed this reference architecture to make it easy for Dell – Red Hat customers to build their own operational readiness cluster and design their initial offerings, using the current releases. Dell and Red Hat will provide the enterprise support and services customers need to stand up production-ready, enterprise-grade OpenStack clusters.

Taxonomy

In the Havana design, this solution contains the core components of a typical OpenStack solution:

- Compute (Nova)
- Object Storage (Swift)
- Block Storage (Cinder)
- Image Service (Glance)
- Identity(Keystone)
- Dashboard (Horizon)
- Networking (Nova networking default; Neutron networking available)[1]
- Telemetry (Ceilometer)
- Orchestration (Heat)

---

[1] Neutron networking is available via custom Services engagements, for customer evaluation only. We currently recommend using Nova networking for production deployments of this solution.

The taxonomy presented in Figure 1 reflects both included infrastructure components (shown in light green) and OpenStack-specific components (shown in red) that are under active development by the community, Dell, and Red Hat. The taxonomy reflects a DevOps perspective that there are two sides for cloud users: standards-based API (shown in pink) interactions and site-specific infrastructure. The standards-based APIs are the same between all OpenStack deployments, and let customers and vendor ecosystems operate across multiple clouds. The site-specific infrastructure combines open and proprietary software, Dell hardware, and operational processes to deliver cloud resources as a service.

The implementation choices for each cloud infrastructure are highly specific to the needs and requirements of each site. Many of these choices can be standardized and automated using the tools in this reference architecture, and by following DevOps processes. Conforming to best practices helps reduce operational risk by leveraging the accumulated experience of Dell, Red Hat and the broader OpenStack community.
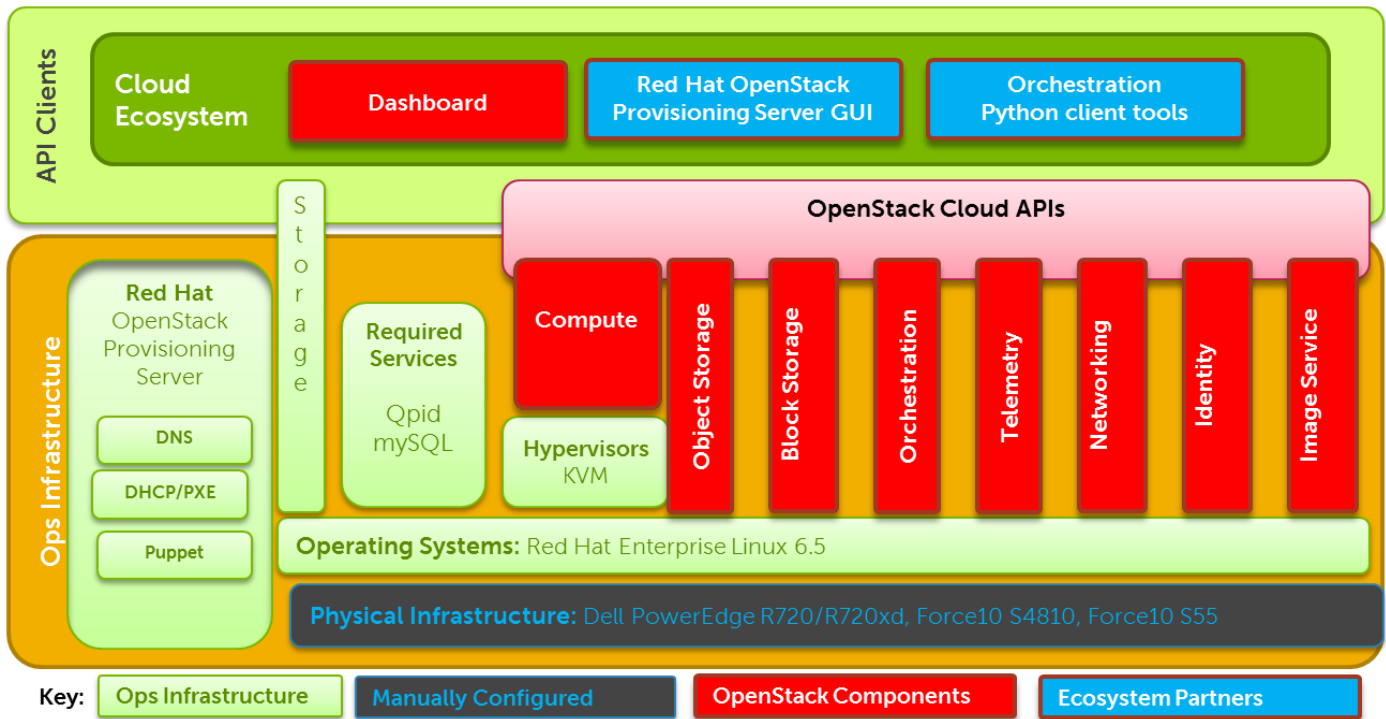
Figure 1: OpenStack Taxonomy

Hardware Options

To reduce time spent on hardware specification for an initial system this reference architecture offers specific choices for hardware and networking. For evaluations, the recommended hardware is general-purpose, and allows a wide range of configuration options. For pilots, the recommended hardware has been optimized for infrastructure, compute, and storage roles. As noted throughout this reference architecture, we are constantly adding capabilities to expand this offering.

Each of the Dell PowerEdge R720 and R720xd server configurations in this reference architecture is designed as a getting-started setup for OpenStack compute, OpenStack storage, or both simultaneously. We recommend starting with OpenStack software using components from this configuration because the hardware and operations processes are a flexible foundation to expand upon. By design, you can repurpose the reference architecture configuration as your cloud deployment grows, so your investment is protected.

Networking and Network Services

This reference architecture uses a starter configuration for networking, which provides access to OpenStack capabilities, while reducing the initial complexity during evaluation. After a review of the OpenStack networking options, Nova-Network and Neutron, we have determined that Nova-Network is the best fit for initial deployments. Nova-Network spreads the network load across all nova servers, keeps all routing and higher functions on the nodes themselves, and eliminates the Network Controller as a single point of failure.

For a production system, additional networking and redundancy configurations are required. This reference architecture supports the addition of these components as the system grows. This includes core and layered networking capabilities, 10GbE networking, NIC teaming, and redundantly trunking top-of-rack (ToR) switches into core routers. See Network Architecture on page 22 for guidelines. Detailed designs are available through Dell consulting services.

OpenStack Architecture

While OpenStack has many configurations and capabilities, we focus on the nine (9) primary components for Red Hat Enterprise Linux® OpenStack Platform 4.0 (Havana), as we have defined in the above taxonomy.

📝   For a complete overview of OpenStack software, visit www.openstack.org.

OpenStack Components

The following component descriptions are from the http://openstack.org site. Extensive documentation for the OpenStack components is available at http://docs.openstack.org/.

Table 2: OpenStack Components		
Function	**Code Name**	**Description**
Identity	Keystone	http://www.openstack.org/software/openstack-shared-services/
		Identity Service provides a central directory of users mapped to the OpenStack services they can access. It acts as a common authentication system across the cloud operating system and can integrate with existing backend directory services.
Dashboard/ Portal	Horizon	http://www.openstack.org/software/openstack-dashboard/
		The OpenStack dashboard provides administrators and users a graphical interface to access, provision and automate cloud-based resources. The extensible design makes it easy to plug in and expose third party products and services.
Object Storage	Swift	http://www.openstack.org/software/openstack-storage/
		OpenStack Object Storage (Swift) is open source software for creating redundant, scalable object storage using clusters of standardized servers to store petabytes of accessible data. It is not a file system or real-time data storage system, but rather a long-term storage system for a more permanent type of static data that can be retrieved, leveraged, and then updated if necessary. Primary examples of data that best fit this type of storage model are virtual machine images, photo storage, email storage, and backup archiving. Having no central "brain" or master point of control provides greater scalability, redundancy, and permanence.
		Objects are written to multiple hardware devices in the data center, with the OpenStack software responsible for ensuring data replication and integrity across the cluster. Storage clusters can scale horizontally by adding new nodes. Should a node fail, OpenStack works to replicate its content from other active nodes. Because OpenStack uses software logic to ensure data replication and distribution across different devices, inexpensive commodity hard drives and servers can be used in lieu of more expensive equipment.
Compute/IaaS	Nova	http://www.openstack.org/software/openstack-compute/
		OpenStack Compute is open source software designed to provision and manage large networks of virtual machines, creating a redundant and scalable cloud computing platform. It gives you the software, control panels, and APIs required to orchestrate a cloud, including running instances, managing networks, and controlling access through users and projects. OpenStack Compute strives to be both hardware and hypervisor agnostic, currently supporting a variety of standard hardware configurations and seven major hypervisors.

Function	Code Name	Description
Virtual Images	Glance	http://www.openstack.org/software/openstack-shared-services/
		OpenStack Image Service (Glance) provides discovery, registration, and delivery services for virtual disk images. The Image Service API server provides a standard REST interface for querying information about virtual disk images stored in a variety of back-end stores, including OpenStack Object Storage. Clients can register new virtual disk images with the Image Service, query for information on publicly available disk images, and use the Image Service's client library for streaming virtual disk images.
Block Storage	Cinder	http://www.openstack.org/software/openstack-storage/
		OpenStack provides persistent block level storage devices for use with OpenStack compute instances. The block storage system manages the creation, attaching and detaching of the block devices to servers. Block storage volumes are fully integrated into OpenStack Compute and the Dashboard allowing for cloud users to manage their own storage needs. In addition to using simple Linux ® server storage, it has unified storage support for numerous storage devices. Block storage is appropriate for performance sensitive scenarios such as database storage, expandable file systems, or providing a server with access to raw block level storage. Snapshot management provides powerful functionality for backing up data stored on block storage volumes. Snapshots can be restored or used to create a new block storage volume.
Networking	Neutron	http://www.openstack.org/software/openstack-networking/
		OpenStack Networking is a pluggable, scalable and API-driven system for managing networks and IP addresses. Like other aspects of the cloud operating system, it can be used by administrators and users to increase the value of existing datacenter assets. OpenStack Networking ensures the network will not be the bottleneck or limiting factor in a cloud deployment and gives users real self-service, even over their network configurations.
Telemetry	Ceilometer	https://www.openstack.org/software/openstack-shared-services/
		The OpenStack Telemetry service aggregates usage and performance data across the services deployed in an OpenStack cloud. This powerful capability provides visibility and insight into the usage of the cloud across dozens of data points and allows cloud operators to view metrics globally or by individual deployed resources.
Orchestration	Heat	https://www.openstack.org/software/openstack-shared-services/
		OpenStack Orchestration is a template-driven engine that allows application developers to describe and automate the deployment of infrastructure. The flexible template language can specify compute, storage and networking configurations as well as detailed post-deployment activity to automate the full provisioning of infrastructure as well as services and applications. Through integration with the Telemetry service, the Orchestration engine can also perform auto-scaling of certain infrastructure elements.

Red Hat Enterprise Linux OpenStack Platform 4.0

Red Hat Enterprise Linux OpenStack Platform delivers an integrated foundation to create, deploy, and scale a secure and reliable public or private OpenStack cloud. It delivers a cloud platform built on Red Hat Enterprise Linux, optimized for and integrated with Red Hat's OpenStack technologies, giving you the agility to scale and quickly meet customer demands without compromising on availability, security, or performance. Additionally, when combined with the Red Hat Storage Server[2] add-on, you gain an integrated and distributed highly available cloud storage platform across public and private clouds.

Red Hat Enterprise Linux OpenStack Platform 4 is built from the combination of Red Hat Enterprise Linux 6.5 and the latest Red Hat OpenStack technology, which is an enterprise-hardened version of the community "Havana" release. This version boasts all the community core features and functions, as well as some additional innovations by Red Hat.

Specifically, Red Hat Enterprise Linux OpenStack Platform 4 now includes OpenStack orchestration (project "Heat"), which provides template-based orchestration for provisioning infrastructure resources. It also includes OpenStack telemetry (project "Ceilometer"), which collects and stores metering and usage data and makes that data available via APIs to custom billing systems. Additionally, Red Hat Enterprise Linux OpenStack Platform 4 now utilizes the highly scalable Foreman provisioning tool for enterprise deployments, which includes bare-metal provisioning for new compute nodes. Finally, in an effort to provide a more seamless infrastructure to customers, it is now integrated with additional Red Hat infrastructure tools, including Red Hat Storage to provide optional object, block, and image storage services.

Key Benefits:

- Integrated and optimized: OpenStack is dependent on the underlying Linux for performance, security, hardware enablement, networking, storage, and other basic servers. Red Hat Enterprise Linux OpenStack Platform delivers an OpenStack distribution with the performance, security, and stability of Red Hat Enterprise Linux, allowing you to focus on delivering the services your customers want instead of the underlying operating platform.
- Take advantage of broad application support. Red Hat Enterprise Linux running as guest virtual machines provides a stable application development platform with many independent software vendor (ISV) certifications so you can rapidly build and deploy your cloud applications.
- Avoid vendor lock-in by moving to open technologies while maintaining your existing infrastructure investments.
- 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 Enterprise Linux OpenStack Platform deployments, ensuring the same level of broad support and compatibility customers enjoy today in the Red Hat Enterprise Linux ecosystem.
- Bring security to the cloud. Rely on 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.
- Enterprise hardened: Red Hat tests, fixes, and certifies each OpenStack release running on Red Hat Enterprise Linux OpenStack Platform for broad hardware and software compatibility and performance.
- Enterprise life cycle: Red Hat Enterprise Linux OpenStack Platform provides a stable branch release of OpenStack and Linux which can be supported by Red Hat for extended life cycles beyond the six-month release cycle of OpenStack. Security fixes, bug fixes, performance enhancements, and some features can be back-ported from future releases without disrupting production environments.

---

[2]http://www.redhat.com/products/storage-server/

Server Infrastructure Options

The solution includes Dell PowerEdge R720/R720xd series high-density rack servers. Table 3: Server Platform Attributes summarizes the high level attributed involved in a server platform choice.

Table 3: Server Platform Attributes		
Shared Infrastructure Platform	**Customer Environment Attributes**	**Workload Attributes**
R720/R720XD	Chosen for • Rack Density 10-20 Server Per Rack • Power Per Rack < 10Kw • Standard Rack/Rear cabling	Chosen for • Higher Frequency CPU • Require High Memory Density (768GB, 24 DIMMs) • Need High Spindle >12 x 2.5" Drives • Ideal for PoC or-Pilot OpenStack Cluster

The following sections describe the supported server models and configurations required. Detailed part lists and rack layouts are included in the appendices

PowerEdge R720 and R720xd Server

The PowerEdge R720 and R720xd servers are Dell's 12G PowerEdge mainstream 2-socket 2U rack servers. They are designed to deliver the most competitive feature set, best performance, and best value. In this generation, Dell offers a large storage footprint, best-in-class I/O capabilities, and more advanced management features. The PowerEdge R720 and R720xd are technically similar, except that the R720xd has a backplane that can accommodate up to 24 drives.

Figure 2: PowerEdge 720xd Server

PowerEdge R720xd feature summary:

- Intel® Romley platform and Intel Xeon® E5-2600 processors
- 1600MHz DDR3
- Network daughter cards for customer choice of LOM speed, fabric, and brand at point of sale
- PCIe SSD in a front-accessible, hot-plug format
- Internal GPGPU support
- Intel® Node Manager power management technology
- Hardware RAID
- Platinum efficiency power supplies, common across 600 and 700 series platforms

Hardware Configurations

Table 4: Hardware Configurations – PowerEdge R720 and R720xd			
Machine Function	PoC Bundle Nodes	Pilot Bundle Infrastructure Nodes	Pilot Bundle Storage Nodes
Platform	PowerEdge R720	PowerEdge R720	PowerEdge R720xd
CPU	2 x E5-2650v2 (8-core)	2 x E5-2670v2 (8-core)	2 x E5-2670v2 (8-core)
RAM (Minimum)	128 GB	128 GB	128 GB
LOM	4 x 1GbE	4 x 1GbE	4 x 1GbE
Add In Network	<none>	2 x Intel X520 DP 10Gb DA/SFP+	2 x Intel X520 DP 10Gb DA/SFP+
DISK	8 x 500GB 7.2K SATA 3.5 inch	8 x 1TB 7.2K SAS 3.5-inch	12 x- 2TB SATA 7.2K 3.5-inch
Storage Controller	PERC H710	PERC H710	PERC H710
RAID	RAID 10	RAID 10	RAID1(OS)/RAID 6(Data)

> 📝 Be sure to consult your Dell account representative before changing the recommended hardware configurations.

Configuration Notes

The *Dell Red Hat Cloud Solutions Bill of Materials Guide* contains the full bill of materials (BOM) listing for the PowerEdge R720 and R720Xd server configurations.

The R720 and R720xd configurations can be used with 10GbE networking. To use 10GbE networking support, an additional network card is required in each node. Refer to the *Dell Red Hat Cloud Solutions Bill of Materials Guide* for the details on the supported card.

Solutions Sizing Terms

In setting up deployments, we have found that there are two distinct user types. The first user type is the person or team that wants to use Openstack in an environment that starts to duplicate production, with only enough hardware to prove the concepts that have been proposed. This team is looking to develop tools to use with Openstack for day to day usage and management, develop an application that requires actual hardware as the base and can start to replicate networking issues that will be seen in a production system.

We recommend for the first user type a PoC Bundle of 5 Servers and networking.

The second user type is the team that has already identified Openstack as the platform of choice, began building tools to support OpenStack, and have applications ready for deployment. They are usually ready to design the first production implementation, with networking, Operating System and management.

We recommend for the second user type a Pilot Bundle, which is suitable for non-production, heavy workloads.

PoC Bundle

The PoC Bundle has been created for the customer who wants to start learning Openstack and developing applications in a Proof of Concept environment. This bundle provides enough infrastructures to exercise OpenStack administration tasks, network configuration, application development, and to prove how to execute on a deployment plan.

Given a virtual machine with the characteristics of 2 cores, 4GB of memory and 40GB of ephemeral hard disk, you can expect to run around 70 virtual machines with a 1.5 oversubscription of CPU cores. At 90 virtual machines, you will have 2 to 1 CPU core oversubscription and still have ephemeral storage under subscribed and memory just starting to be oversubscribed.

The bundle consists of the following:

All nodes are R720 nodes with Dual 8 Core Xeon Processors, 4TB raw (2TB usable) drive space, 128 GB of memory and 1 Gb Ethernet:

- Node 1: Red Hat Openstack Manager Server
- Node 2: Openstack Controller
- Nodes 3-5: Nova Compute Nodes
- Network Switch: Force 10 S55

Pilot Bundle

The Pilot Bundle is designed for pilot environments with storage. This bundle is for customers who want entry-level testing of an application, and the beginning of a production environment. This does not preclude the need to become familiar with Openstack administration and day to day management. These bundles consist of the components described in Figure 3: Pilot Bundle on page 18.

Given a virtual machine with the same characteristics as above, you can expect to run around 170 virtual machines with a 1.5 oversubscription of CPU cores. At 228 virtual machines, you will have 2 to 1 CPU core oversubscription, still have ephemeral storage under subscribed, and memory just starting to be oversubscribed.
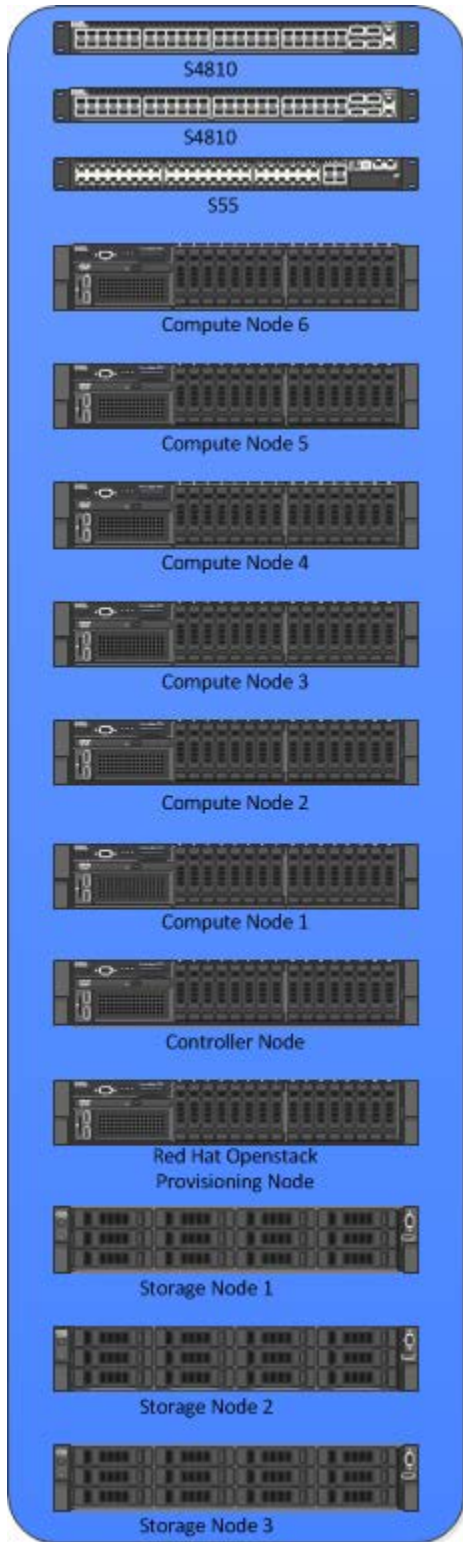
Pilot Bundle

R720 with Dual 10 Core Xeon processors, 8TB (4TB usable) HDD, 128 GB of memory, four (4) 1Gb Ethernet, and four (4) 10 Gb Ethernet

- Node 1 Red Hat Openstack Manager

- Node 2: Openstack Controller

- Nodes 3-8 Nova Compute Nodes

- Nodes 9 – 11 Storage Nodes- These are the R720xd nodes added to this bundle to bring in 3 Storage nodes. These nodes have the same specifications as the above, except for the hard drives. The R720xd is configured with 12 2TB hard drives, for a RAW space of 24 TB. Please see your Account Representative for hard drive options.

- Network Switches: Two (2) Force 10 S4810 and one (1) Force 10 S55

Figure 3: Pilot Bundle

Pilot Bundle Cluster Building Blocks

The Pilot Bundle can be increased in capability by adding additional building blocks. As each building block changes the solution, the requirement for additional time to architect, deploy and maintain the solution needs to be taken into consideration.

The base building blocks are:

- Compute Building Blocks: Consists of three R720's with Dual 10 Core Xeon processors, 8TB (4TB usable) HDD, 128 GB of memory four (4) 1Gb Ethernet and four (4) 10 Gb Ethernet

- Storage Building Block: Consists of three(3) R720xd with Dual 10 Core Xeon processors, 12TB (11TB usable) HDD, 128 GB of memory four (4) 1Gb Ethernet and four (4) 10 Gb Ethernet

- Infrastructure Building Block: Consist of single R720's with Dual 10 Core Xeon processors, 8TB (4TB usable) HDD, 128 GB of memory, four (4) 1Gb Ethernet and four (4) 10 Gb Ethernet

Production Bundle

Clusters larger than the Pilot bundle (i.e., Production bundles) must be designed, sized and configured based on your requirements. Please work with your Sales Representative to properly architect Production cluster deployments.

Operational Notes

Backup/Recovery

Since the system is designed for exploration that could later be extended to a production stage, backup and recovery have not been addressed in this configuration. The Red Hat Openstack Manager node, while not needed for normal operations of the services, is not redundant or backed-up.

Service Layout

During the deployment each service configured by the solution needs to be on a particular hardware type. For each server platform, two types of nodes have been designed: Infrastructure and Storage. The configurations in the Appendices have them separated that way. Red Hat Openstack Manager is designed for flexibility, enabling you to try different configurations in order to find the optimal service placement for your workload. Table 5 is the recommended layout of each service.

Table 5: Node Type to Services		
Hardware Type	**Service**	**Node to Deploy Upon**
Infrastructure	Red Hat Openstack Manager	Admin Node
Infrastructure	Database-server	Openstack Controller
Infrastructure	Qpid-server(Messaging)	Openstack Controller
Infrastructure	Keystone-server	Openstack Controller
Infrastructure	Cinder-scheduler	Openstack Controller
Infrastructure	Cinder-volume (LVM Block Storage)	Openstack Controller
Infrastructure	Neutron-server	Openstack Controller (Optional)
Infrastructure	Nova-Controller	Openstack Controller
Infrastructure	Nova-multi-compute	1 or more Compute Nodes
Infrastructure	Nova dashboard-server	Openstack Controller
Infrastructure	Ceilometer	Openstack Controller
Infrastructure	Heat	Openstack Controller
Infrastructure	Nova-Network	Same nodes as Nova-multi-compute
Storage	Red Hat Storage Server	3 or more Storage Servers
Storage	Swift Object Store	3 or more Storage Servers

Deployment

Deployment consists of three phases:

- Hardware Setup:
    - Rack, Stack
    - Cable
    - BIOS Config
    - RAID config
    - Switch config
- Software Setup:
    - Deploy provisioning server
    - Deploy storage services
    - Deploy controller node(s)
    - Deploy compute nodes
- Test environment

Network Architecture

The Dell Red Hat cloud solutions with RHEL OpenStack Platform uses Dell Force10 S55 1/10-Gigabit and/or Dell Force10 S4810 10-Gigabit Ethernet switches as the top-of-rack connectivity to all OpenStack-related nodes. This reference architecture is used to support consistency in rapid deployments through the minimal differences in the network configuration.

This reference architecture implements at a minimum six (6), separate vLANs:
- **Management/OOB network —** iDRAC connections can be routed
- **Nova Network Private LAN–** Sets up the backend network for nova and the VM's to use
- **Nova Network Public LAN –** Sets up the front network for routable traffic to individual VMs
- **Provisioning Network LAN**—Connects all nodes NICs into the fabric used for setup and provisioning of the servers
- **Private API Network Cluster Management LAN**— Used for communication between OS controller and nodes
- **Public API Network Access** – Sets up access to the RESTFUL API and the Horizon GUI
- **Storage Network –** Used by all the storage nodes for replication/writes/reads

The network consists of three major network infrastructure layouts:
- **Data network infrastructure** — The data network consists of the server NICs, the top-of-rack (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-RU s55 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.
- **Core network infrastructure** — The connectivity of aggregation switches to the core for external connectivity.

Network Components

The data network is primarily composed of the ToR and the aggregation switches. Configurations for 1GbE and 10GbE are included in this reference architecture. The following component blocks make up this network:

Server Nodes

Server connections to the network switches could be one of four possible configurations:
- Active-Active LAG in load-balance bond formation
- Active-Backup in failover/failback formation
- Active-Active round robin based on gratuitous ARP
- Single port

In the first case the connectivity on the switch side must be in a LAG (or port-channel). In cases 2 and 3, we recommend that you do the configuration as a LAG but the ports should still be part of the same layer-2 domain. In some cases all members of the LAG connect to a single ToR switch. In others the LAG splits into two ToR switches. This is an optional setup as OpenStack has redundancy built into the application.

The teaming configuration that Dell recommends is transmit-lb (mode = 5). This configuration setting is explained in greater deal in the *Dell Red Hat Cloud Solutions Deployment Guide*. Please contact your sales representative for a copy of the deployment guide.

Access Switch or Top of Rack (ToR)

The servers connect to ToR switches. Typically there are two in each rack. The switches recommended by Dell are the Force10 S55 for 1GbE and S4810 for 10G connectivity.

The 10GbE configuration utilizes Force10 S4810 switches as the ToR switches. We recommend this pair of switches run Virtual Link Trunking (VLT). This feature allows the servers to terminate their LAG interfaces into two different switches instead of one. This allows active-active bandwidth utilization. This feature provides redundancy within the rack if one switch fails, or needs maintenance. The uplink to the aggregation pair is 80 Gb, using a LAG from each ToR switch. This is achieved by using two 40G interfaces in a LAG connecting to the aggregation pair. Therefore, a collective bandwidth of 160G is available from each rack.

Each rack is managed as a separate entity from a switching perspective, and ToR switches connect only to the aggregation switches.

Aggregation Switches

For a deployment of three to nine racks of 10G servers (12 racks max), we recommend the Dell Force10 S4810 as the aggregation switch. It is both 10GbE and 40GbE capable. The 40GbE interfaces on the S4810 could be converted into four 10GbE interfaces, thereby converting this switch into 64 10GbE-capable ports. ToR switches connect to aggregate switches via uplinks of 10GbE interfaces from the ToR Force10 S4810 to the Force10 S4810.

Our recommended architecture uses Virtual Link Trunking (VLT) between the two Force10 S4810 switches in aggregation. This feature enables a multi-chassis LAG from the ToR switches in each rack. The stacks in each rack can divide their links between this pair for switches to achieve powerful active-active forwarding, while using full bandwidth capability, with no requirement for spanning tree. Running 40GbE Ethernet switches like the Dell Force10 Z9000 in aggregation can achieve a scale of up to hundreds of 1G deployed nodes.

For the 10G server deployment, our recommendation depends upon:

- The scale at which the rack layouts are planned

- Required future scaling

When designing a large deployment, we recommend the Force10 S4810 for aggregation for smaller scale and the Force10 Z9000 for larger deployments. The Force10 Z9000 is a 32-port, 40G high-capacity switch. It can aggregate up to 15 racks of high-density PowerEdge R720 and R720xd servers. The rack-to-rack bandwidth needed in OpenStack would be most suitably handled by a 40G-capable, non-blocking switch. The Force10 Z9000 can provide a cumulative bandwidth of 1.5TB of throughput at line-rate traffic from every port.

Core

The aggregation layer could itself be the network core in many cases, but otherwise it would connect to a larger core. Details on this topic are beyond the scope of this document.

Layer-2 and Layer-3

The layer-2 and layer-3 boundaries are separated at the aggregation layer. The Reference Architecture uses layer-2 as the reference up to the aggregation layer. That is why VLT is used on the aggregation switches. Red Hat Foreman Openstack Manger requires a layer-2 domain in order to provision servers.

PoC Bundle Configuration

The PoC bundle configuration logical networking will look like Figure 4: PoC Bundle Rack and Networking Layout. The bundle includes a Dell Force10 S55 switch to which all nodes will be connected. This bundle is configured with NO network redundancy. The rack layout will be similar to Figure 4. Please note that all nodes are stacked together, with a single switch above them.
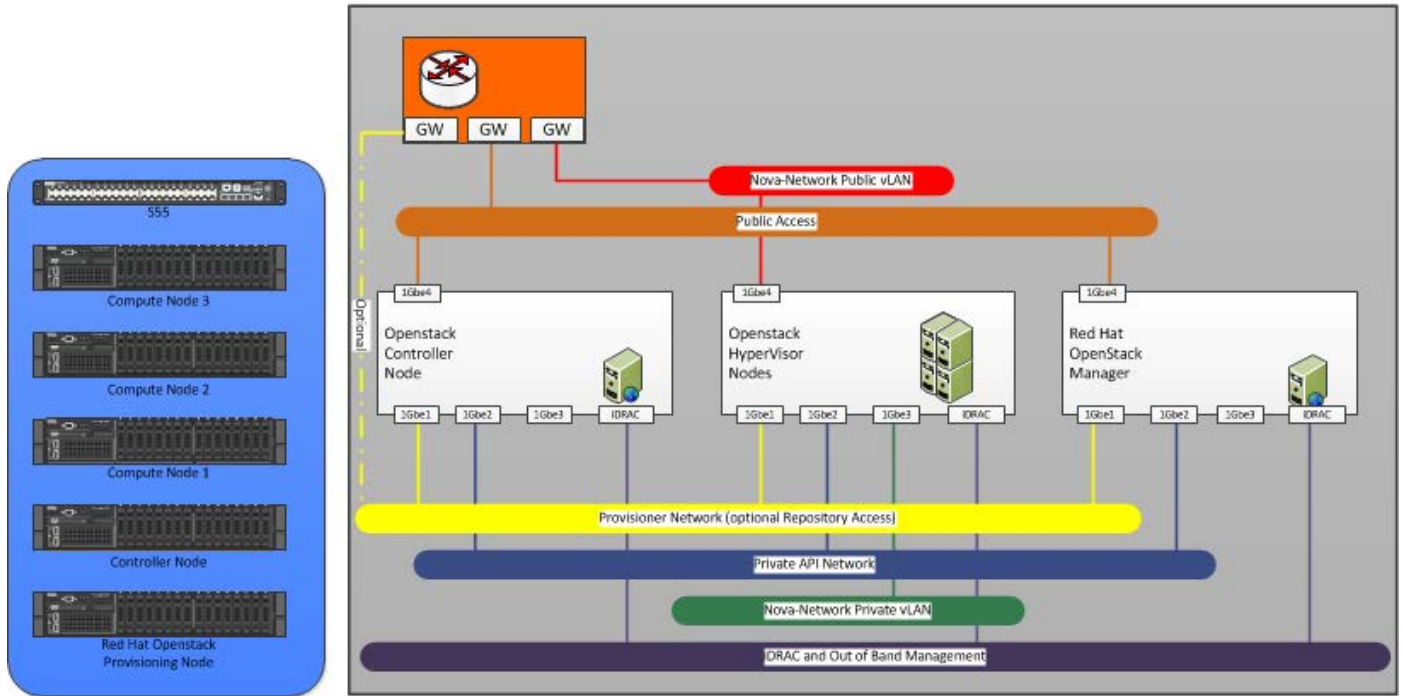
Figure 4: PoC Bundle Rack and Networking Layout

Pilot Bundle Configuration

The networks for the Pilot and Larger bundles have been designed to support production-ready servers with a highly available network configuration. The rack layout consists of two Dell Force 10 S4810 switches as the ToR, and one Dell Force 10 S55 as management (a second Force 10 S55 should be considered for the Heart Beat to be truly redundant). The switches take advantage of Virtual Link Trunking (VLT). Using additional Dell Force 10 features, the ports can be configured to support one or many virtual LANs; allowing the nodes to communicate between themselves, and allow network segregation of tenants utilizing vLANs.
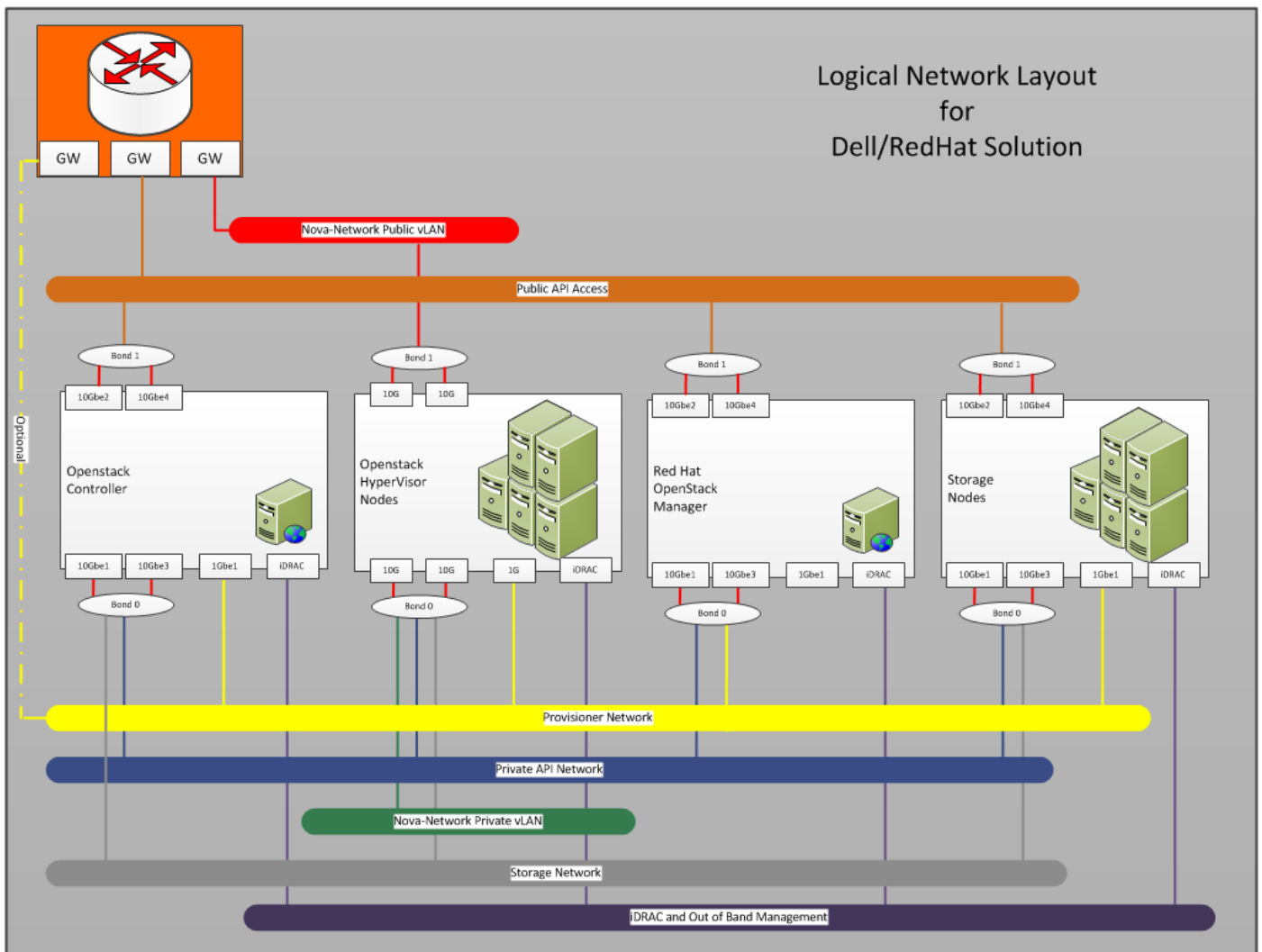
Figure 5: Pilot Logical Network

Figure 5: Pilot Logical Network shows how each node is connected to the different networks used in the solution.

Out of Band Management Network

The management network of all the servers and switches is aggregated into a Dell Force10 S55 switch that is located in each rack of the POD. It uplinks on a 10G link to the S4810 switches.

Dell Open Switch Solution

In addition to the Dell switch-based reference architecture, Dell provides an open standard that allows you to choose other brands and configurations of switches for your OpenStack environment. The customer is expected to ensure that the switches conform to these requirements, which are configured according this RA's guidelines. The following list of requirements will enable other brands of switches to properly operate with the tools and configurations in the *Dell Red Hat Cloud Solutions Reference Architecture*:

- Support for IEEE 802.1Q vLAN traffic and port tagging
- Support using one untagged and multiple tagged vLANs on the same port
- Ability to provide a minimum of 170 Gigabit Ethernet ports in a non-blocking configuration within Provisioning vLAN
  - Configuration can be a single switch or a combination of stacked switches to meet the additional requirements

- The ability to create link aggregation groups (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

Update History

Version 1

First Reference Architecture for the Dell Red Hat Cloud Solutions with Red Hat Enterprise Linux™ OpenStack Platform

Getting Help

Contacting Dell

For customers in the United States, call 800-WWW-DELL (800-999-3355).

> ✎ If you do not have an active Internet connection, you can find contact information on your purchase invoice, packing slip, bill, or Dell product catalog.

Dell provides several online and telephone-based support and service options. Availability varies by country and product, and some services may not be available in your area. To contact Dell for sales, technical support, or customer service issues:

- Visit support.dell.com.
- Click your country/region at the bottom of the page. For a full listing of country/region click **All**.
- Click **All Support** from the **Support** menu.
- Select the appropriate service or support link based on your need.

Choose the method of contacting Dell that is convenient for you.

References

For more information about deploying OpenStack in your solution portfolio, see:

- Dell OpenStack®-Powered Cloud Solution Deployment Guide
- Red Hat® Enterprise Linux® OpenStack® Platform Deployment Guide
- Red Hat ® Enterprise Storage Server

To Learn More

For more information on the ell Red Hat Cloud Solutions with Red Hat Enterprise Linux™ OpenStack Platform, visit:

www.dell.com/openstack