

SUSE® OpenStack Cloud Production Deployment Architecture

Guide

www.suse.com

Solution Guide

Cloud Computing

STACK[™]
VELOCITY



Table of Contents	page
Introduction	2
High Availability Configuration	6
Network Topography	8
Services and Support	10
Learn More	10

Introduction

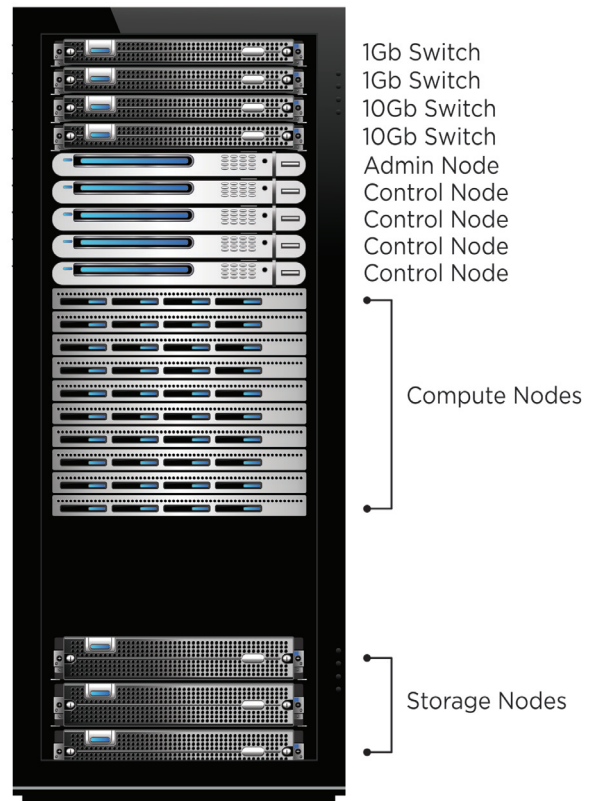
Delivering a fully orchestrated OpenStack cloud can be a complex and time-consuming task. StackVelocity and SUSE have teamed up to deliver solutions that allow you to accelerate your time to value while maintaining the economics required for success in the marketplace.

The following document provides a reference solution for an OpenStack implementation intended for a production cloud environment supporting approximately 768 guest instances and includes a control plane configured for highly available operations. The capacity can be expanded to 1600 guests when the rack is fully populated.

The solution utilizes white box hardware in combination with SUSE® OpenStack Cloud to provide a pre-integrated, optimized and supported configuration that is ready for production deployment.

The entire configuration is integrated, tested and configured with software in the factory prior to shipment.

Pilot Production Rack Configuration



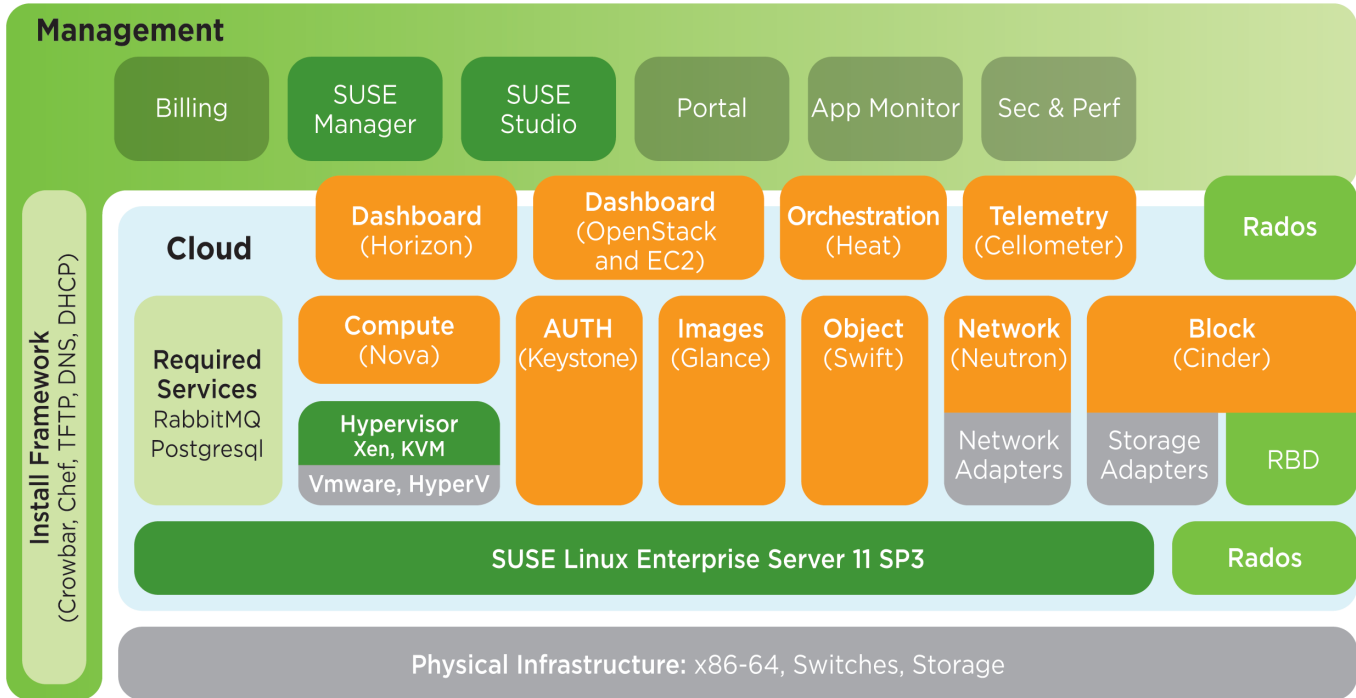
Hardware Specifications

ADMIN NODE (1)	1U server enclosure Single 8 core CPU processor 16 GB DDR-3 memory, 1600 MT/s 3 HDDs, 3 TBs each, SATA
STORAGE NODES (3)— CINDER / CEPH (Expandable to 4)	(3) OSD NODES 2U server enclosure Dual 12 core CPU processors 64 GB DDR-3 memory, 1600 MT/s Internal SFF HDDs (internal or rear-mounted) (2) 128GB SSD (OS—RAID1 mirror) External SFF HDDs (24) 1TB SATA SSD 2 x 10G and 2 x 1G Ethernet ports
CONTROLLER NODE (4)— Controller / RGW	1U server enclosure Single 8 core CPU processor 24 GB DDR-3 memory, 1600 MT/s 4 HDDs, 2 TBs each, SATA 2 x 10G and 2 x 1G Ethernet ports
COMPUTE NODES (12) (Expandable to 25)	1U server enclosure Dual 8 core CPU processors 256 GB DDR-3 memory, 1600 MT/s 6 1TB SATA SSD (data) 2 128GB SSD RAID1 (OS) 2 x 10G and 2 x 1G Ethernet ports
NETWORK	2 1Gb Managed Ethernet Switch (48P) IPMI network for server management (public) Management network for deployment and configuration (private) 2 10Ge Ethernet switch (48P) Cluster, application and data traffic (public)
RACK AND POWER	42U—19" rack Dual PDUs 208VAC single-phase inputs—L6-30 plugs

Sizing and Workload Assumptions

Workload Profile: Guest specification: 2C processor // 4GB memory // 18GB storage Guests per node: -64 Total guest capacity (base config): -768 Total guest capacity (full rack with 25 computes): -1600
Performance: Base config IOPS per guest: -93+ IOPS per storage node: -240,000 Total base IOPS capacity: 720,000
Design Criteria and Parameters: Compute: 18GB persistent storage per guest (full rack capacity) 2CPU, 4GB RAM per guest 175% core hyper-threading, 4 to 1 guest to core ratio IOPS: 1000 IOPS per SSD deliverable (extreme conservative value)
Storage Sizing: 2.0–2.6GB RAM per OSD, conservative over commit 1 OSD per spindle 1 CPU core per OSD
Storage Performance: Low-to-mid density 100% availability with 1 node failure 100% uptime with 1 OSD failure Total capacity = 96TB Total capacity with 3x replication = 32TB Storage capacity per RU = 12TB IOPS per compute node = 38,400

SUSE OpenStack Cloud Solution Architecture



The software stack employs an optimized OpenStack distribution from SUSE called “SUSE OpenStack Cloud” in conjunction with SUSE Linux Enterprise Server as the operating system. Deployment and management are provided through the SUSE OpenStack Cloud Administration Server.

OpenStack Component Distribution

Cluster Management (Administration Server)	Bootstrap images: Chef/Crowbar Environment management (GUI) Add/remove nodes/roles
OpenStack Management (Horizon)	High Availability architecture (4 nodes) 2 nodes dedicated to database service (PostgreSQL) 2 nodes dedicated to remaining services Cluster collocation constraints to separate networking (Neutron) services on 1 node and remaining services on the other, except in the case of a failover Remaining services: RabbitMQ (messaging daemon) Glance (image repository) Cinder (volume storage—backed by Ceph) Nova (compute) controller Heat (orchestration) Ceilometer (metrics)
Compute (Nova)	Nova compute instances Local storage L2 networking
Storage (Ceph)	Gateway RADOS API Proxy services Storage (object + block) Object storage (replacing Swift) Block storage layered on top (Cinder)

Out-of-Box Functionality

OpenStack Capabilities:

Fully-configured and functional OpenStack environment
Full control over Ceph's storage platform
Add/remove volumes
Add/remove objects as needed
Full control to create/destroy instances
Full control over network IP space and bridging
Full control over images

Service Scalability:

Ability to add/remove nodes as required to scale out infrastructure at will
Ability to add/remove OpenStack services at any time
Production deployment enables storage replication and hyper-expansion to/from geo-dispersed regions and zones

Cluster Management:

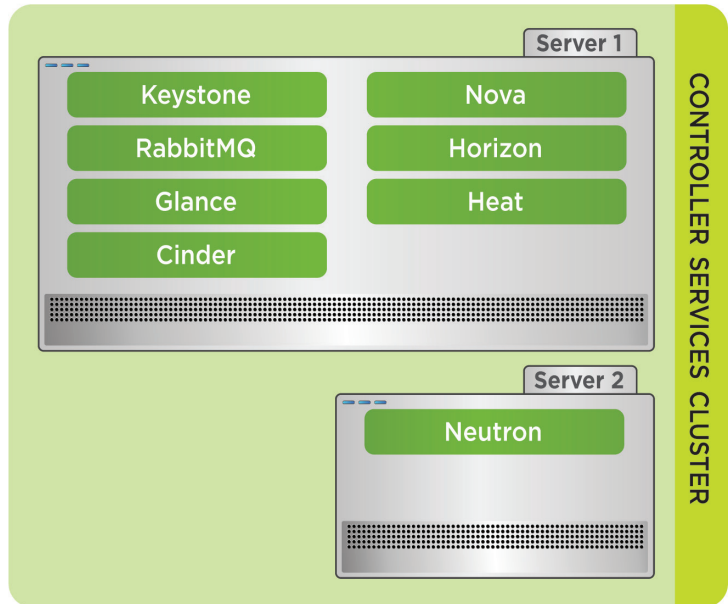
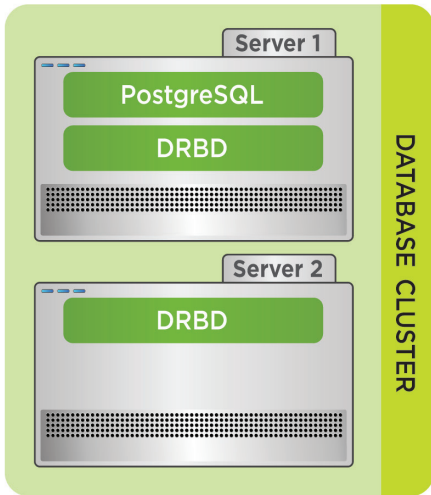
Provisioning (SUSE Linux Enterprise Server, Windows Hyper-V, Windows 2012)
Automated high availability
Dynamic role assignment
Integrated "closed circuit" DHCP PXE cluster management network

High Availability Configuration

One of the key differences between proof of concept (PoC) and production deployments is future-proofed designs with High Availability (HA). Although HA is relatively atypical in smaller deployments, this solution gets companies immediately headed in the right direction by mitigating single points of failure. This translates into fewer growth-related service interruptions and industry best practices that facilitate the redundancy of stateless API service planes, guaranteed message delivery and improved load balancing.

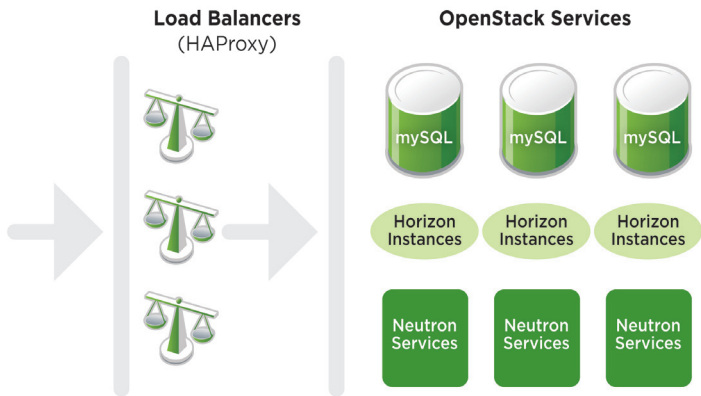
HA OpenStack Management

The first step in establishing a future-proofed OpenStack deployment is with the redundancy of OpenStack and other services running on control nodes. OpenStack controllers house the OpenStack services, PostgreSQL, load balancing and L4-L7 network service management if desired. To optimize performance while maintaining a reasonable hardware footprint, High Availability via four controller nodes with redundant configurations is being used in this rack configuration.



HA Service Availability

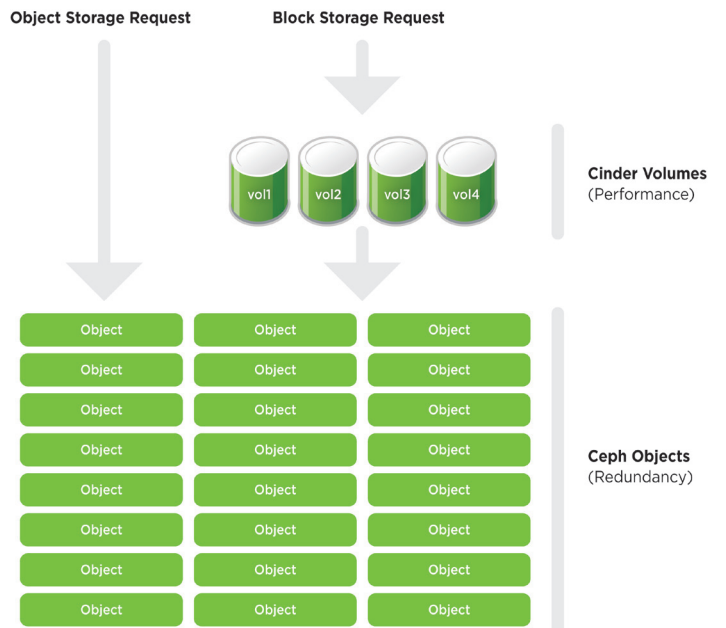
Incoming requests are accepted by HAProxy, where the load is balanced across all three control nodes. The request is then distributed to the first available compute node. This process is generally handled in a "round robin" fashion. The first request is issued to the first compute node, the second request to the second compute node, and so on.



HA Storage

Ceph is a scalable software defined storage solution which provides object, block and file system storage all within the same platform. Redundancy of object stores is achieved by distributing them over five zones. In addition, OpenStack users are able to access Cinder volumes on the same platform.

Through the use of this architecture we are also able to decentralize volume storage. In the event of a host failure, this allows an instance to be migrated to another host without affecting network and storage resources.



Network Topography

IP Address Scheme

NETWORK	SCHEME
Management	IPMI: 10.1.0.0/16 Admin: 10.1.0.0/16 Public: 10.0.0.0/16 Neutron: 10.2.0.0/16
Storage	Ceph Object Storage: 10.3.0.0/16
Virtual Instances	Private: 172.16.0.0/16 Public (Floating IP): 192.168.0.0/16 (for this doc, use appropriate public IP range)

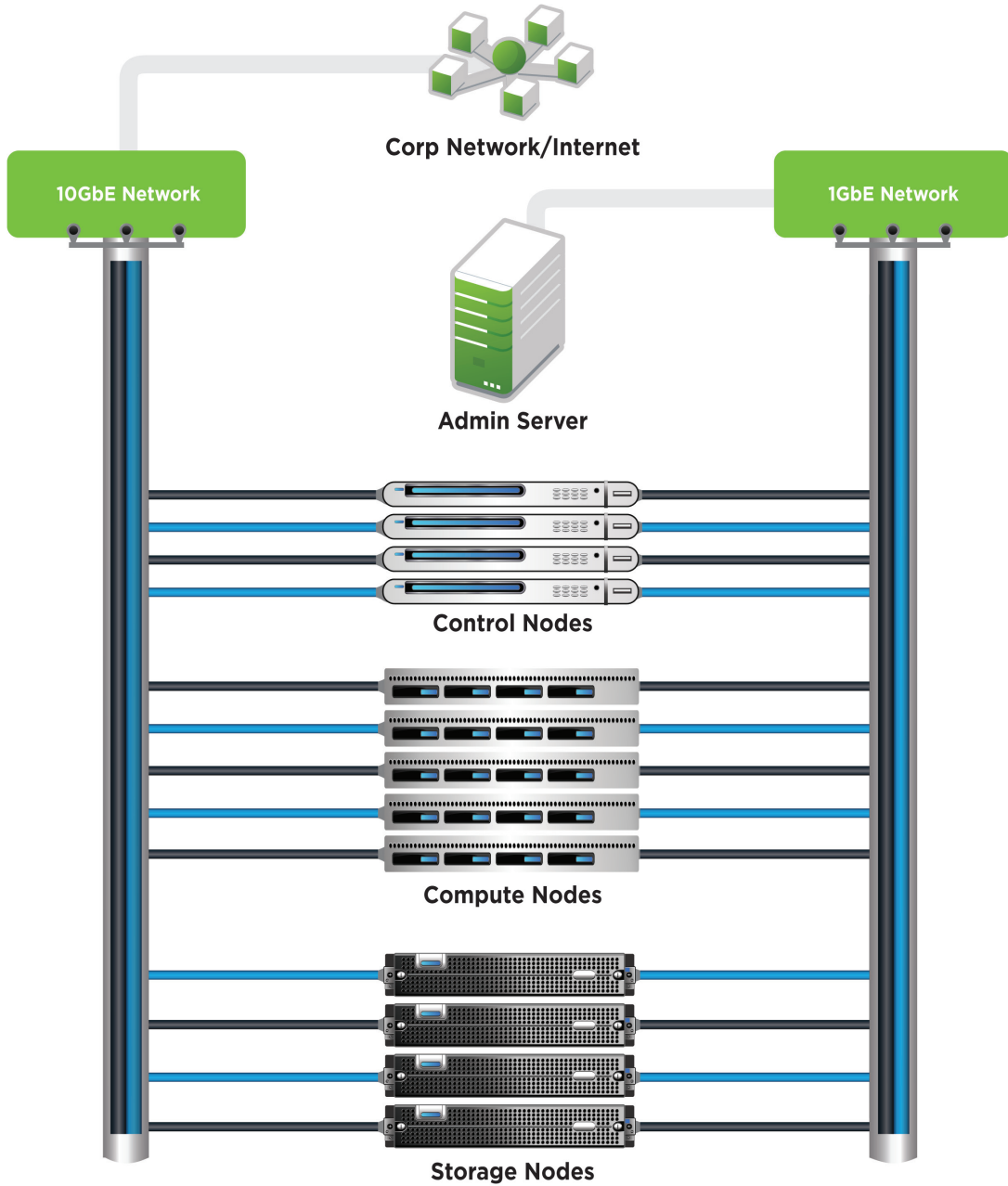
VLAN Definitions

NETWORK	VLAN
IPMI Network	Untagged—1GB
Admin Network	Untagged—1GB
Neutron	700
Storage Network	200
Public and Floating IPs	500
Private Network	300

Physical Interface to VLAN Mapping

SYSTEM	NIC	NETWORK SPEED	VLAN
SUSE OpenStack Cloud Admin Node	NIC #0	1GbE	Bond1—untagged
	NIC #1	1GbE	Bond1—untagged
Controller Nodes / Storage Gateways (1-3)	NIC #0	1GbE	Bond1—untagged
	NIC #1	1GbE	Bond1—untagged
	NIC #2	10GbE	Bond2—200,300,500,700
	NIC #3	10GbE	Bond2—200,300,500,700
Storage Nodes (1-5)	NIC #0	1GbE	Bond1—untagged
	NIC #1	1GbE	Bond1—untagged
	NIC #2	10GbE	Bond2—200,300,500,700
	NIC #3	10GbE	Bond2—200,300,500,700
Compute Nodes (1-11)	NIC #0	1GbE	Bond1—untagged
	NIC #1	1GbE	Bond1—untagged
	NIC #2	10GbE	Bond2—200,300,500,700
	NIC #3	10GbE	Bond2—200,300,500,700

Network Diagram



www.suse.com

SUSE OpenStack Cloud

DEPLOYMENT AND MANAGEMENT	The SUSE OpenStack Cloud Administration Node is an open source deployment and management tool for OpenStack. It simplifies and automates the process of deploying and maintaining OpenStack across the rack configuration through a GUI-driven interface.
OPENSTACK DISTRIBUTION	SUSE OpenStack Cloud supports components and packages from multiple vendors. It supports multiple hypervisors (Xen, KVM, VMware vSphere and Microsoft Hyper-V) and multiple plug-ins for storage back-ends and networking.
SOFTWARE SUPPORT	SLA options including 8x5 and 24x7 remote support options. Phone and web support portal. Global training programs.

Services and Support

Hardware

- *One year support included*
 - Hardware warranty (parts and labor)
 - Advanced replacement
 - 8x5 remote support
- *Additional available services and support options*
 - Onsite deployment and installation
 - Two or three year warranty
 - On-site support / replacement
 - Advanced spares
 - 24x7 remote support

SUSE OpenStack Cloud

- *One year support included*
 - 24x7 priority support
- *Intermediate training*
- *Additional available services and support options*
 - Professional services
 - Advanced training

Learn More

**STACK
VELOCITY**

Stack Velocity

30 Great Oaks Blvd
San Jose, CA 95119

www.stackvelocity.com

contact: info@stackvelocity.com



SUSE LINUX Products GmbH

Maxfeldstrasse 5
90409 Nuremberg
Germany

Tel: +49 (0)911-740 53-0

www.suse.com



**Contact your local SUSE Solutions Provider,
or call SUSE at:**

1 800 796 3700 U.S./Canada
1 801 861 4500 Worldwide

SUSE
Maxfeldstrasse 5
90409 Nuremberg
Germany