# Planning an OpenStack Deployment

SUSE OpenStack Cloud 7

# Planning an OpenStack Deployment: generated by DAPS

SUSE OpenStack Cloud 7

by Rick Ashford and Cameron Seader

Publication Date:  February 08, 2017 , Version: Version 0.1

SUSE LLC
10 Canal Park Drive
Suite 200
Cambridge MA 02141
USA
https://www.suse.com/documentation

# Contents

# 1 Introduction

## 1.1 Why do I need this?

The actual physical implementation and deployment of your cloud will be the shortest part of the whole project to get it up and running. The vast majority of your project time will be spent in conference rooms with whiteboards and take-out food hashing out the details of what this project is going to look like. The cloud project will require the cooperation and input from a wide variety of sources, such as the storage team, network team, physical infrastructure management, end-users, legal, procurement and more. Without this coordinated effort, users will get frustrated, your project will not be successful, and the whole thing is doomed from the start.

## 1.2 What is the intent of this document?

This document is intended as a starting point for a SUSE OpenStack Cloud implementation. Use this guide as a guide to help you prepare for deployment and as a worksheet as you discuss how to configure SUSE OpenStack Cloud for your environment.

The items below are discussed in greater detail in the SUSE OpenStack Cloud deployment guide. Reference the deployment guide for an understanding of the terms and clarification. Throughout the document you will see references to the section(s) in the deployment guide.

As of the time of this writing, the current version is SUSE OpenStack Cloud 7 (Newton-based release). Deployment guide links are based off of this version and may need to be adjusted for subsequent product versions.

Community documentation may be found on the OpenStack site, http://docs.openstack.org/ ↗. All SUSE-provided documentation, including the deployment guide, is available at: https://www.suse.com/documentation/suse-cloud-7/ ↗

# 2 Preliminary Considerations

## 2.1 Know Your Philosophy

Before you begin planning the technical details of your implementation, there are several overarching questions you will want to consider that will directly impact the overall architecture

- What problem are you trying to solve by implementing a private Infrastructure-as-a-service (IaaS) cloud?
- Who are the users of your cloud? What do they want out of it?
- What are your business and technical requirements?
- What are the constraints for this project?
- What additional resources will you require to have a successful deployment?

### 2.1.1 What problem are you trying to solve?

This question is probably the most vital one to know before you begin planning your cloud. If you don't know what the goal is that you are trying to accomplish, your likelihood of achieving it is dramatically reduced.

If, for example, your overall goal is to provide a playground for your developers so they will stop annoying IT with constant requests for additional resources, you will make dramatically different decisions than if your overall goal is to streamline your production environment processes.

In the first scenario, you would probably have a single, non-commercial hypervisor such as KVM or Xen. You would likely be looking to implement a relatively cheap backend for volume storage, and high-availability of your control plane may not be significant. You would not likely need to have a large address space reserved for floating IP addresses to expose your cloud workloads to the outside world.

In the second scenario, a production environment, you will be much more stringent in your requirements. You will likely need multiple hypervisors to accommodate varying virtual environments (ie – production will be on VMware, developers will have the cheaper KVM environment, and Windows workloads will be deployed on Hyper-V to maximise the efficiency of licensing costs). You will likely be looking for a more reliable storage infrastructure, leveraging a SAN instead of using local disk storage. You may even have several storage backends that you need to accommodate. You will likely need to reserve a significant number of IP addresses for exposing these production workloads to the outside world.

Without understanding exactly what you are trying to accomplish with your cloud implementation, you are much more likely to make decisions that will be sub-optimal, which may result in significant costs (in time and money) to rectify. Establishing a correct course here will inform every other decision you make for the better.

### 2.1.2 Who are the users of your cloud?

This question is similar to the first, but it is still worthy of considering individually. If the intended users of the cloud are not in-line with the overall goal of the cloud, then significant frustration will likely emerge. Knowing who all of the intended users are and what their needs and expectations are can help you head that frustration off before it festers and impacts productivity.

If, for example, your stated goal is to provide a playground for developers, but there is also going to be a significant number of less-technical users, you will want to make sure you cater to both user groups. The less-technical users will require significantly more documentation of your processes for using the cloud, and formal training may be appropriate to ensure they don't get frustrated and refuse to use it.

### 2.1.3 What are your business and technical requirements?

Most enterprise IT environments have specific expectations for uptime, typically in the form of a Service-Level Agreement, or SLA. Stringent SLA's require higher budgets to accommodate higher quality and higher quantities of hardware, networking, and physical infrastructure (power, cooling, disaster-recovery processes, etc). In addition, if you work in government, retail, or healthcare, you will likely have specific compliance requirements for things like PCI, HIPAA, or Common Criteria and these need to be taken into account as well.

### 2.1.4 What are the constraints for this project?

Every project is short on something, whether it's manpower, money, or time. This falls squarely into the realm of the old adage, “You can pick any two of fast, cheap, and high quality. You can't have all three.” Often-times ill-informed management will attempt to defy the laws of physics and human nature and require all three, but realistically speaking, that's not going to happen. Understanding the priority of your constraints can help you set management expectations appropriately so that you have achievable success criteria.

### 2.1.5 What additional resources will you need?

Typically it is a fairly small team that is given the charge to build out your private IaaS cloud. You will need to plan on getting input and assistance from a wide variety of other teams, such as storage, networking, and physical infrastructure. Identifying whose help you will need and when you will need it allows them to plan and give your needs the full attention they require.

# 3 Performing a Proof-of-Concept

If you are using this guide to deploy a proof-of-concept (POC), here are some additional steps to consider.

## 3.1 Pre-POC planning meeting

**BEFORE THE POC BEGINS, SCHEDULE A MEETING TO DISCUSS THE FOLLOWING TOPICS:**

- Is the hardware you are planning to use for the proof-of-concept (POC) certified for SUSE Linux Enterprise Server? You may want to compare your planned hardware setup with the published reference architectures available at https://www.suse.com/products/suse-cloud/resource-library/#RA ↗
- Verify your intended method for installing the Admin node. Do you have physical access to the node, and if so can you use a DVD or USB drive to install? If you do not have physical access are you planning on attaching the ISO as a virtual DVD, or will you be using a network-based PXE setup?
- Confirm that the BIOS, fiber-channel cards, network cards, IPMI devices, and BMC' controllers are all updated to the latest firmware from the manufacturer.
- Discuss the planned network, including IP address ranges, VLAN tags, Neutron plugins, and, if using the Open vSwitch plugin, the encapsulation method. Make sure there is an established schedule to have that configuration done before you begin the POC.
- Are you planning on having a highly-available control plane? If so, will there be a difference in the hardware between the POC and production? For example, in the POC you may choose to use a two-node cluster hosting all services, but in production use three-node clusters and split out services onto their own individual clusters.
- If you choose to have a highly available control plane we strongly recommend you choose teaming mode and have at least 2 NIC’s available on all nodes. We recommend that you have at least 4 for redundancy on the network stack.
- Identify your STONITH method and verify that the appropriate resources are available to implement.
- Source for installation media and SMT repositories (ex. USB drive brought by SE)

- Use Admin Appliance? (Could remove the next requirement for backup/restore)
- Backup/restore plan for the admin server if the installation must be repeated
- What will the patching/upgrade strategy be? (update barclamp, SUSE Manager, none)
- Build a SLE12 SP1 JeOS Live USB or ISO image using SUSE Studio with supportutils package and setup network configuration with yast at firstboot. This will come in handy for troubleshoting network issues on any hardware so that we can setup the network configuration properly for crowbar.
- On the first day of the POC:
  - Bring JeOS image
  - Collect a supportconfig from all systems that will be part of the cloud infrastructure
  - Use the supportconfigs to review interface map/network conduit/network info in network.json DG Appendix D (https://www.suse.com/documentation/suse-cloud-5/singlehtml/book_cloud_deploy/book_cloud_deploy.html#app.deploy.network_json) ↗

# 4 Worksheet

## 4.1 High-Availability

In order to scale your cloud effectively and still provide the levels of service that your users expect, you will likely need to implement high-availability services for the control plane. The level of clustering will depend greatly on the mixture of your level of paranoia, combined with the practical limits of budget, physical resources, and ability/willingness to introduce additional complexity to your implementation.

If you decide to cluster, you can take two different approaches:

- Build a single cluster and put all the services on it. If you need to minimize resource contention, you may introduce affinity rules to the cluster that will have it prefer to keep the active/passice services (PostgreSQL and RabbitMQ) on their own server, and allow haproxy to load-balance the rest of the services.
- Build multiple clusters that can each be dedicated to specific OpenStack services. This can be as simple and efficient as a single 2-node cluster for PostgreSQL/RabbitMQ, and a separate cluster for the remaining load-balanced services, or even wholly-dedicated clusters for any number of services.

The decision for your cluster configuration will directly impact the number of servers you will need to buy, although usually budget is the limiting factor for how many systems you can have, and you will need to optimize your resources and configuration accordingly. DG 2.6.2 (https://www.suse.com/documentation/suse-cloud-5/singlehtml/book_cloud_deploy/book_cloud_deploy.html#sec.depl.reg.ha.control)↗

Regardless of your eventual cluster configuration, you will need to prepare shared storage for housing your PostgreSQL data, and again for RabbitMQ. What will you use for your shared storage for the database-server and rabbitmq-server? (NFS, DRBD (requires an extra disk for each node in the cluster) external NAS, SAN)

DG section 2.7.5

DG section 2.7.2

The next step is to determine what mechanism you want to use for cluster fencing, or STONITH (Shoot The Other Node In The Head). The most popular fencing method is SBD, which requires a small shared block device between all the nodes in the cluster. You may even want to consider multiple SBD devices via multipathing to reduce false alarms and create a more robust solution. Additional STONITH methods include integrating with the IPMI device in your server, or leveraging a network-connected APC. Please note that if you decide not to use SBD for fencing that you will want to have an odd number of servers in your cluster so that quorum can be attained.

DG section 2.7.5

## 4.2 Networking

Networking is probably the single most important area for planning, as it is the most complex and is extremely difficult to change after deployment without having to rebuild your cloud.

### 4.2.1 Understanding Network Choices

Therefore, prior to commencing any cloud deployment, please define the following network subnets (defaults in parentheses):

- Admin (`192.168.124.0/24`) – Administrative network that will be used to deploy and maintain all the cloud infrastructure nodes. The size of this subnet is the upper limit to the number of compute/control/storage nodes that can be provisioned. This network will have a PXE boot server on it, and therefore must be isolated from other networks. The TFTP service required for PXE does not understand VLAN tags, so it must appear as an untagged network (ie - trunk the ports).

- BMC (`192.168.124.0/24`) – Administrative network for accessing IPMI devices for the cloud infrastructure. It must be on the same subnet as the admin network unless you use a VLAN.

- Fixed (`192.168.123.0/24 – tag 500`) – Cloud-wide network that most cloud instances will connect to in order to interact with each other and the outside world. While some number of systems may exist in isolated networks and are never connected to the fixed network, generally speaking the range of this subnet defines the maximum number of concurrent cloud workloads that can be run.

- SDN (`192.168.130.0/24 – tag 700`) – Cloud-wide network that will be used for the software-defined networking stack. The subnet for this needs to be at least as large as the admin network.

- Storage (`192.168.125.0/24 — tag 200`) – Network used for communicating with Swift storage nodes. Even if there are no plans to use Swift, the Admin server still requires that this network be defined. If you opt to use SUSE Enterprise Storage and deploy it via the cloud admin node, it may live on this network as well. It is generally, however, a better idea to deploy SUSE Enterprise Storage as its own stand-alone service and integrate it with your cloud, allowing it to provide storage to other areas in your datacenter besides just your cloud.
- Public (`192.168.126.0/24 — tag 300`) – Used for exposing OpenStack services and APIs to external users. Must be routable on your existing LAN. Only the control node servers will be on this network, so a small address space may be used when defining the DHCP range.
- Floating (`192.168.126.0/24 — tag 300`) – Used for exposing cloud instances externally. Must be the same subnet as the public network. The number of IP's designated for this network will be the hard upper limit for the number of cloud instances that can be concurrently exposed outside of your cloud.

## 4.2.2 Network Tables

TABLE 4.1: ADMIN/BMC NETWORK ADDRESS ALLOCATION

| Function | Default Address(es) | VLAN | IP Address(es) | Notes |
|---|---|---|---|---|
| Netmask | 255.255.255.0 | N/A | | |
| router | 192.168.124.1 | None | | |
| admin | 192.168.124.10-11 | None | | |
| DHCP | 192.168.124.21-80 | None | | |
| Host | 192.168.124.81-160 | None | | |
| BMC VLAN host | 192.168.124.161 | 100 | | Optional - If the IP- |

| Function | Default Address(es) | VLAN | IP Address(es) | Notes |
|---|---|---|---|---|
| | | | | MI devices are already configured on a VLAN and you do not wish to change their configuration, use this network. |
| BMC host | 192.168.124.162-240 | None | | |
| Switch | 192.168.124.241-250 | None | | |

TABLE 4.2: BASTION NETWORK ADDRESS ALLOCATION

| Subnet | Gateway | DNS | IP Address | Proxy |
|---|---|---|---|---|
| | | | | |

TABLE 4.3: STORAGE NETWORK ADDRESS ALLOCATION

| Function | Default Address(es) | VLAN | IP Address(es) | Notes |
|---|---|---|---|---|
| host | 192.168.125.10-239 | 200 | | |

TABLE 4.4: FIXED NETWORK ADDRESS ALLOCATION

| Function | Default Address(es) | VLAN | IP Address(es) | Notes |
|---|---|---|---|---|
| Netmask | 255.255.255.0 | N/A | | |
| router | 192.168.123.1-49 | 500 | | |
| dhcp | 192.168.123.50-254 | 400 | | |

TABLE 4.5: PUBLIC/FLOATING NETWORK ADDRESS ALLOCATION

| Function | Default Address(es) | VLAN | IP Address(es) | Notes |
|---|---|---|---|---|
| Netmask | 255.255.255.0 | N/A | | |
| router | 192.168.126.1 | 500 | | |
| public host | 192.168.126.2-49 | 300 | | |
| public dhcp | 192.168.126.50-127 | 300 | | |
| floating host | 192.168.126.129-191 | 300 | | |

TABLE 4.6: SOFTWARE-DEFINED NETWORKING (SDN) ADDRESS ALLOCATION

| Function | Default Address(es) | VLAN | IP Address(es) | Notes |
|---|---|---|---|---|
| Netmask | 255.255.255.0 | N/A | | |

| Function | Default Address(es) | VLAN | IP Address(es) | Notes |
|---|---|---|---|---|
| host | 192.168.130.10-254 | 700 | | |

### 4.2.3 Compute Infrastructure Information

**Hypervisor(s)**

With which hypervisor or hypervisors do you intend to deploy and integrate?

Options include:

- KVM
- Xen
- VMware vSphere
- Hyper-V
- zVM
- Bare metal

**Caution**

OpenStack security groups are currently implemented via iptables for non-proprietary software-defined networking stacks (Open vSwitch, Linuxbridge, Flat). As such, non-Linux hosts (VMware, Hyper-V, zVM) will not take advantage of these security groups and will have to handle firewall rules inside the guest workload using their native toolsets.

**Containers**

If you intend to manage containers with this implementation, which deployment/management model(s) do you intend to use?

Options include:

- Bare metal (dedicated Docker host)
- Virtualized (via the Magnum service)

**Caution**

The nova-docker plugin for doing Docker deployments to bare metal, outside of an orchestration service like Heat or Magnum, has been deprecated and will eventually be removed from the product.

______________________________

### 4.2.4 Additional Network Information

**Interface Mode**

Which network mode will you will be using? ((Single, Dual, Teaming) Remember to take into account whether you are using a highly available control plane or not, if so then you should choose Teaming) DG 2.1.2

______________________________

**FQDN**

Fully Qualified Domain Name (FQDN) of the Admin server (cannot be changed after deployment).

______________________________

**External DNS**

Are there additional DNS servers you want SUSE Cloud to forward to for non-local DNS records?

______________________________

**Neutron L2 Plugin**

Which plugin do you wish to use for your software-defined networking stack?

Options include:

- Open vSwitch (Default)
- Linuxbridge
- VMware NSX
- Cisco Nexus
- Cisco ACI
- Midokura

### Caution

VMware only supports vSphere as part of an OpenStack implementation if used with their NSX SDN.

---

**Neutron Encapsulation Plugin**

Which packet encapsulation protocol do you wish to use for your software-defined networking stack?

Options include:

- GRE(default)
- VXLAN
- VLAN
- Flat/None

### Caution

The GRE plugin is incompatible with VMware and Hyper-V

---

### 4.2.5 Storage Information

**Image Repository**

Which storage plugin do you wish to use for your image repository (Glance)

Options include:

- Local
- Ceph RBD
- Swift

______________________________

**Block Storage**

What storage service do you wish to use as the backend for your block storage (Cinder)

Options include:

- Local
- Ceph RBD
- NetApp
- Fujitsu
- EMC
- EqualLogic
- VMware

______________________________

### 4.2.6 Package Source Information

**Subscription-Management Tool (SMT)**

Will you utilize an existing SMT server or install one on the admin node? If existing, then indicate the FQDN below.

________________________________

**SUSE Manager**

As an alternative to SMT, do you wish to use SUSE Manager as the source for your packages? If so, then indicate the FQDN below.

________________________________

# 5 Going from Proof-of-Concept to Production

## 5.1 Patching Strategy

Identify a strategy for how you plan to patch your cloud infrastructure

**Updater barclamp**

The admin node has an "Updater" barclamp available which can be set to automatically deliver patches to the underlying cloud infrastructure servers (*NOT* the guest workloads).

These patches will automaatically be distributed when the following criteria have been met:

1. They are available in the repositories the admin server is using for provisioning systems
2. They fit the criteria identified in the barclamp
3. The infrastructure server initiates its chef-client run to verify its configuration

## 5.2 Operational Planning

What are the environmental and logistical differences between your proof-of-concept environment for which you will need to adjust your configuration?

**Networking**

Subnets and VLANs

**Physical Infrastructure**

Availability zones

**Control Plane**

Clustering

**Authentication**

Active Directory

**Third-Party applications**

SUSE Studio