CANONICAL ubuntu

iVolve Prerequisites Checklist

Private Cloud Build

Prepared by: Ivan Dobos <ivan.dobos@canonical.com>

Prepared on: 15 March, 2021

Version: 1.0

Revisions

Date	Author	Version	Notes
2021/03/15	Ivan Doboš	1.0	Initial version

1. Introduction

This document holds the requirements and guidelines laid down by Canonical to ensure a smooth Private Cloud Build (PCB) deployment process. The following list should be reviewed and verified by the Customer prior to Canonical commencing the PCB deployment. The goal of this document is to reduce the number of potential pitfalls and unknowns encountered when Canonical employees start the deployment, to ensure a positive experience for both the Customer and Canonical. Any discrepancies may result in delay of the PCB deployment.

If for any reason a requirement on this checklist cannot be met prior to Canonical being present to commence deployment we ask that it is brought to the attention of the stakeholders as early as possible.

2. Checklist

2.1. Infrastructure Nodes

2.1.1. 3x infrastructure nodes shall be provided and console access made available so that Ubuntu 18.04 can be installed remotely by Canonical;

2.1.2. Infrastructure nodes must be connected to the out of band management network (with or without native VLAN configuration) and to the network used for PXE booting by other machines (native VLAN, no vlan tags);

2.1.3. Infrastructure nodes have to have either public internet access or access via HTTP/HTTPS proxy servers to endpoints mentioned in the “Security and Firewalls” section;

2.1.4. Infrastructure nodes must be dedicated to the target deployment and not used for other purposes.

2.2. Node Inventory

2.2.1. A detailed inventory sheet listing each node being deployed in the cloud, including BMC address, out of band management information (credentials, including username and password, other parameters) and hostname;

2.2.2. Rack design: machine and switch placement information in racks and pods;

2.2.3. A unique hostname for each node used for the Cloud Build. Hostnames should belong to a subdomain that would be delegated to the MAAS DNS service or used independently but consistently with a corporate DNS service;

2.2.4. Datacenter network design related to node connectivity: routing and switching configuration, DNS servers, proxy servers, NTP servers, egress routing setup.

2.3. Machine hardware configuration

2.3.1. UEFI needs to be prioritized over legacy boot. Alternatively legacy boot needs to be completely disabled;

2.3.2. UEFI/BIOS should be configured such that CPU Frequency is possible to manage via the operating system. The relevant settings are vendor-specific and UEFI/BIOS-version specific.

- **HP**:
  - HP Power Profile: “Custom”;Huawei FusionServer Pro (ex: 2288H v5)
  - HP Power Regulator: “OS Control Mode”;
- **Dell** PowerEdge 11th Generation
  - Power Management: “OS Control”;

- **Dell** PowerEdge 12th Generation:
  - System Setup > System BIOS > System Profile > "OS Control";
  - CPU power management: "OS DBPM";
    - Note that this can be set with the following command:
      `racadm set BIOS.SysProfileSettings.SysProfile PerfPerWattOptimizedOs`
- **SuperMicro**:
  - Power Technology > Custom;
  - EIST > Enable.
- **Huawei FusionServer Pro** (ex: 2288H v5)
  - Socket Configuration -> Advanced Power Mgmt. Configuration
    - Select "Performance Profile" with "Customer"
    - Select "Power Policy" with "Customer"

    Note: The "Customer" profile will auto set EIST(P-State) as "Enabled"

2.3.3. OOB management card (IPMI) needs to be enabled and configured (note that for IPMI 2.0 maximum password length is 20 characters, and 16 for IPMI 1.5)

2.3.4. PXE boot needs to be enabled;

- Preference needs to be given to on-motherboard networks cards (unless agreed upon otherwise);

2.3.5. Boot order should be as follows to ensure proper boot process if MAAS is unavailable:

- PXE;
- internal drives.

2.3.6. Intel VT-x virtualization features enabled;

2.3.7. Hyperthreading enabled (unless the contrary is agreed upon);

2.3.8. Intel VT-d enabled if supported;

2.3.9. SR-IOV enabled (if agreed during the design session and supported).

2.3.10. 1xRAID-1 created (to hold the O/S) on all the nodes (Infra + Cloud)

2.3.11. Serial console redirection should be enabled. In order for the entire boot process to be visible via IPMI, both "serial redirection" and "serial over lan" need to be enabled. Settings should be consistent across all nodes;

2.3.12. Power policy in case of loss of AC power should be always-on (not 'last-state' or 'always-off').

2.3.13. IPMI over LAN needs to be enabled;

2.4. Firmware

2.4.1. All hardware (motherboard UEFI/BIOS, NIC, NVMe, RAID controllers) is updated to have the latest firmware version supplied by the hardware vendor unless a legacy version is explicitly required by Canonical for the purpose of security or deploying its tooling;

2.4.2. Hardware should not have mixed firmware versions on different servers unless there are different generations or models used which prevents this.

2.5. Certificates

Any certificates required to authenticate with Customer's internal services are provided to Canonical, e.g.

2.5.1. An Internal CA certificate (or a certificate chain) for a CA that issued LDAP/AD TLS certificates;

2.5.2. Artifacts necessary for a PKI setup (I or II):

I. Certificates and keys for OpenStack TLS termination:
   A. Internal or public CA certificates used for issuing OpenStack API endpoint certificates;
   B. Certificates and keys to be used for TLS termination of individual OpenStack services (Keystone, Glance, Cinder, Nova API, Neutron, Gnocchi, OpenStack Dashboard, Heat, Aodh, Rados Gateway or Swift, Vault, Landscape);

II. An Intermediate CA certificate generated based on a CSR provided by Canonical and used for the purposes of creating endpoint certificates via Vault service deployed by Canonical during the Build process (the intermediate CA should have the capability of issuing certificates with CN and subjAltName fields for names that belong to a region-specific zone *.region-one.openstack.example);

2.5.3. If TLS termination is used on corporate HTTP proxy servers, issuer CA certificates used for validation of proxy server certificates must be provided;

2.5.4. If SSLBump configuration is used at a corporate proxy server to dynamically decrypt HTTPS traffic and generate destination server certificates to impersonate them, a CA certificate or chain used to validate these dynamic certificates needs to be provided by the Customer.

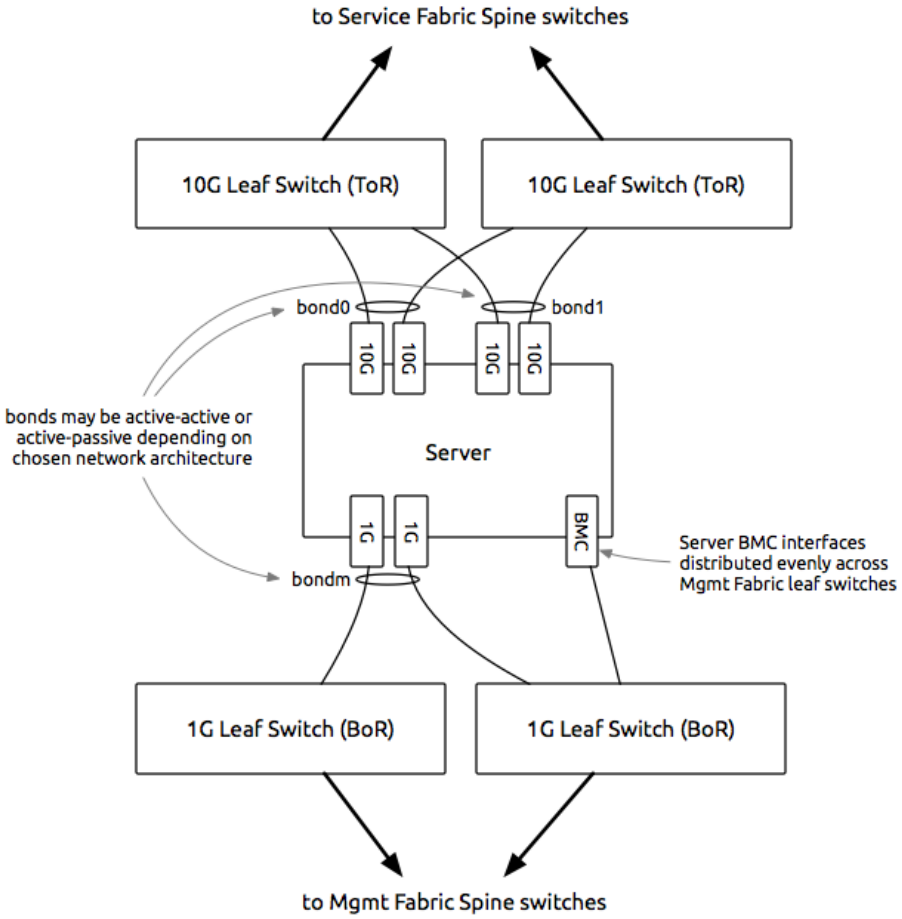
2.6. Physical Connectivity

2.6.1. Ensure that cabling is aligned with the physical network layout.

2.6.2. Depending on the design, two or three bonds connected to the Service Fabric switch pair ("bond0" and "bond1" below) where ports of an individual bond are connected to different switches running in the Multi-Chassis Link Aggregation (MLAG, MC-LAG, Virtual PortChannel or other vendor-specific names.) configuration;

All bonds are to be configured on the switch with 802.3ad Link Aggregation in "Fast" Mode.

2.6.3. One bond ("bondM") connected to the Management Fabric (BoR) switch pair (with MLAG); A single 1G ethernet connection to the Management Fabric switch pair for the BMC (IPMI, DRAC, iLO etc) interface;



Interface	Member	Switch
bond0	eth3+eth5 x 10Gbps	10G Leaf Switch ToR

bond1	eth2+eth4 x 10Gbps	10G Leaf Switch ToR
bondM	eth0+eth1 x 1Gbps	1G Leaf Switch BoR
BMC (DRAC/IPMI/iLO)	BMC	1G Leaf Switch BoR (No Port Channel/Bonding)

2.7. Networking Switching

2.7.1. Read-only access to network switches needs to be provided in order to facilitate debugging of possible routing and switching issues. Alternatively, a copy of switch configs needs to be provided (without any sensitive data like credentials) and subsequent copies if changes are made.

2.7.2. A chosen network layout needs to provide resilient, high bandwidth network connectivity for a defined set of separate Layer 2 broadcast domains delivered to every host.

- ➢ Leaf-spine with paired leaf switches is the recommended L3 fabric topology with EVPN provided to end hosts. While L3 setups without shared L2 provided to end hosts across leaves are supported, we need to account for a few low-level design considerations related to HA and Neutron provider networks.
- ➢ Other topologies like multi-tier design (with core, aggregation and access switches) are acceptable as well.

2.7.3. Switch ports must be configured to support bonded host interfaces;

2.7.4. Switch ports must support IEEE 802.1q VLAN tagged and untagged traffic;

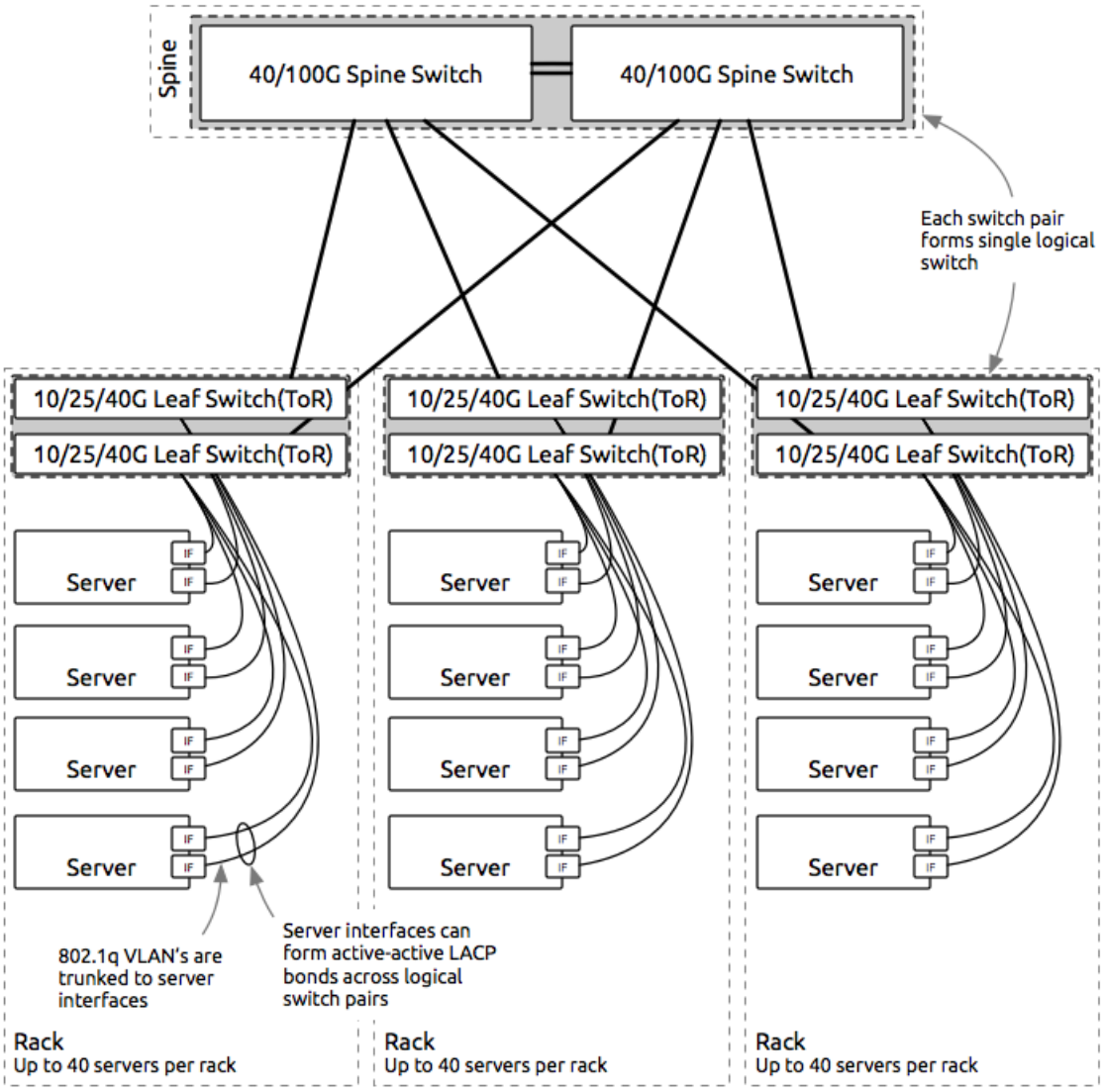
2.7.5. A single layer 2 network and associated layer 3 subnet must be designated for OAM (Operation, Administration, and Maintenance);

- ➢ Multi-segment OAM setups are also supported, however, require more careful planning.

2.7.6. All BMC and operating system provisioning interfaces must be connected or routed to the OAM network. We recommend to implement separate layer 2 broadcast domains (such as VLANs) for BMC and OAM boot networks, however, the requirement is for infrastructure nodes to have unrestricted network access to both of those networks;

2.7.7. If host OAM interfaces are to be bonded, the switch ports must be configured to support IEEE 802.3ad LACP with LACP rate set to “Fast timeout” and LACP fallback (sometimes called “LACP bypass”, “LACP suspend individual”, “LACP force-up” or “LACP fallback” depending on the hardware vendor) options enabled;

2.7.8. In order to set the LACP fallback option mentioned above it is important to set it with the lowest possible priority value in order to bring the LAG up and keeping a single port active until the expiry of a timeout period, otherwise the server will not be able to boot from PXE;



2.7.9. If STP is used, ensure the PortFast (or equivalent) option is enabled on server facing ports;

2.7.10. Port security should not be enabled on OAM switch ports as traffic from both infrastructure and non-infrastructure nodes will contain MAC addresses of virtual machines and containers which may lead to switchport interface downtime due to policy enforcement;

2.7.11. Different types of traffic generated by the deployed infrastructure will require Individual broadcast domains (such as VLANs):

- L3-oriented setups such as L3 leaf-spine where no L2 stretching is set up are also supported, however, should be discussed separately as every leaf will terminate a L2 domain introducing a requirement of per-leaf VLANs and subnets related to a single logical network. Alternatively, EVPN can be used to provide a stretched L2 setup.

VLAN	MTU[1]	Bond	Name	Subnets
~~TBD~~	~~1500~~	~~bondM~~	~~Public package mirror/corporate proxy server network - connected with the Infra nodes only (used for package downloads only)~~	~~A public or private range: TBD~~
TBD (untagged)	1500	bondM	Provisioning Network - OAM (could be merged with the 'Public package mirror/corporate proxy server network')	At least a private /24 range: TBD
TBD	1500	bondM	Management network connected with a VPN terminator from which server management ports (OOB and PXE) are L3-reachable (could be merged with 'Public package mirror/corporate proxy server network' and 'Provisioning Network')	At least a private /24 range: TBD
TBD	1500	bond0	Internal OpenStack communication	At least a private /24 range: TBD
TBD	9216 OR 9000	bond0	Ceph Replication Network	At least a private /24 range: TBD
TBD	9216 OR 9000	bond1	Ceph Access Network	At least a private /24 range: TBD
TBD	9216 OR 9000	bond1	Underlay network for Overlay network traffic (VXLAN or GRE). IP addresses are assigned to VTEPs.	At least a private /24 range TBD
TBD	1500	bond1	Instance (VM or container) Floating IPs	A public range to cover the amount of instances planned in the environment by the Customer
TBD	1500	bond1	Externally consumable	A public or private range.

[1] *The MTU values mentioned will be configured at the host side. If there is additional encapsulation at the switching fabric side (e.g. fabric-side VXLAN), please communicate the MTU available for use by hosts, not the one configured on switch ports.*

			OpenStack API	TBD
TBD	1500	bond1	Cloud-provided DNS access VLAN. Could be the same as external OpenStack API.	A /28 subnet (6 usable addresses minimum; 3 for designate-bind units and 3 for linux bridges on the hosts). Will be used for “glue records” in subdomain delegation from corporate servers.

2.7.12. Ensure MTU configuration is consistent throughout the Customer network and configured to the agreed upon setting laid out in the design phase;

2.7.13. Jumbo frames must be enabled (if supported) on Ceph access and Replication VLANs in addition to the underlay network for the overlay traffic.

2.8. Network layer and Reachability

2.8.1. Verify that a logical network configuration matches the one agreed upon during the design phase;

2.8.2. A subnetting design should be provided - every VLAN needs to have a subnet allocation enough to cover hosts and infrastructure containers and virtual machines used for the deployment;

2.8.3. If a layer 3 topology like Leaf-Spine is used make sure that each VLAN configured on each switch terminating the L2 domain has an associated subnet allocation and routing is implemented between the different segments by using a combination of a routing protocol and possibly VRF-lite (or similar) functionality. In this case, gateway addresses for each subnet must be provided by the Customer. Alternatively, an EVPN setup needs to be provided;

2.8.4. Ensure there are no DHCP agents or relay agents in broadcast domains (VLANs) that MAAS will be serving DHCP on. The expectation is that MAAS will provide DHCP services for each of the networks defined in the Cloud configuration or will otherwise manage static IP address allocations for provisioned servers;

2.8.5. If any IPv6 usage is expected (control plane or data plane), this has to be discussed in advance;

2.8.6. Verify NTP server connectivity from the target environment - correctly configured NTP is critical for the operation of any cloud platform. Upstream NTP should be configured to provide resilient, high quality NTP sources;

2.8.7. Verify corporate or public DNS server connectivity from the target environment; Corporate DNS should resolve public domains. Verify HTTP proxy server or direct public internet connectivity by performing basic deb and snap package installation on an Ubuntu system from an installed infrastructure node:

```
sudo apt update && sudo apt upgrade
sudo snap install --classic juju
```

2.9. DNS

2.9.1. Delegation of subdomains to name servers deployed as a part of the cloud needs to be done. For that a subdomain for API endpoints and a separate subdomain for instances deployed in OpenStack should be dedicated per region (if there are multiple cloud regions), such as:

- Region one: region-one.openstack.example.;
  - keystone.region-one.openstack.example;
  - cinder.region-one.openstack.example;
  - glance.region-one.openstack.example;
  - nova.region-one.openstack.example;
  - neutron.region-one.openstack.example;
  - designate.region-one.openstack.example;
  - radosgw.region-one.openstack.example;
  - aodh.region-one.openstack.example;
  - gnocchi.region-one.openstack.example;
  - heat.region-one.openstack.example;
  - openstack-dashboard.region-one.example;
- compute.region-one.openstack.example.;
  - ws-vm1.<tenant-zone>.compute.region-one.openstack.example;
  - ...
  - vm-name.<tenant-zone>.compute.region-one.openstack.example;
- (DNS server IP addresses to be specified in Glue records will be provided after the deployment.).

2.9.2. Use of .local top-level domain is highly discouraged as it is reserved for multicast DNS (RFC6762) and affects systemd-resolved behavior such that it does not forward DNS queries to upstream DNS servers;

- .local is reserved by IANA as a special-use domain[2]
- Note that by default lookups for domains with the ".local" suffix are not routed to DNS servers[3]

2.9.3. In order to access cloud-provided DNS for auto-generated instance name-based DNS records, corporate DNS servers must be configured to use subdomain delegation to a set of name servers in a dedicated address range. Appropriate NS records and “glue” records must be created for delegated subdomains;

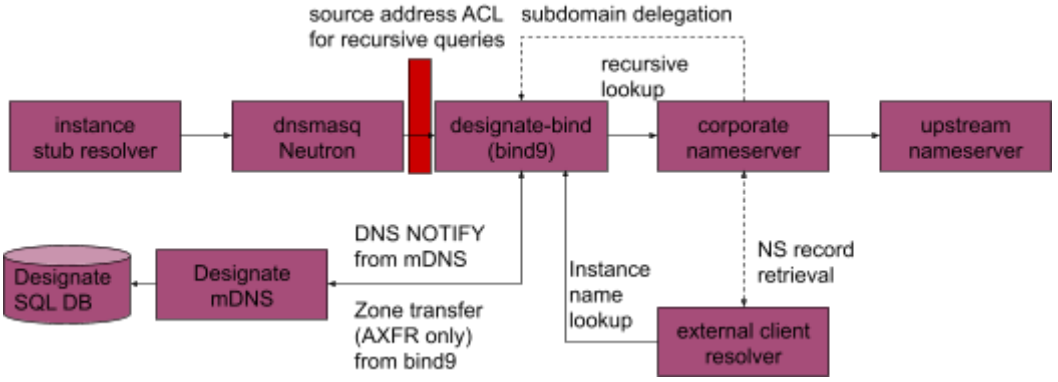
- Usage of Designate requires support in a Neutron core plugin such as the reference ML2 core plugin. While ML2 supports Designate integration, SDN solutions implementing their own core plugin may not have that support.

---

[2] *https://www.iana.org/assignments/special-use-domain-names/special-use-domain-names.xhtml*

[3] *https://www.freedesktop.org/software/systemd/man/systemd-resolved.service.html#Protocols%20and%20Routing*

- A logical overview of instance DNS resolution and external client DNS resolution is shown below:



2.9.4. In order to access public OpenStack API endpoints a set of virtual IPs for OpenStack services (keystone, nova etc.) in a given region needs to have names assigned with address records added to the corporate name servers.

- For L3-oriented setups virtual IPs are not used and the procedure to provide a single endpoint per API service is different and requires subdomain delegation to MAAS DNS servers.

2.10. Security and Firewalls

2.10.1. All firewall and iptables rules communicated in the design phase are consistent with deployed hardware and a copy of those rules are provided to Canonical;

2.10.2. IPMI/iLO/AMT or any other baseboard management traffic must be allowed between all infrastructure nodes and server out of band interfaces

- Secure Shell (SSH) TCP 22;
- Remote Console/Telnet TCP 23;
- Web Server Non-TLS TCP 80;
- Web Server TLS TCP 443;
- Terminal Services TCP 3389;
- Virtual Media TCP 17988;
- Shared Remote Console TCP 9300;
- Console Replay TCP 17990;
- Raw Serial Data TCP 3002;
- IPMI: 623 (UDP);
- Remote console: 5900 (TCP);
- Virtual media: 623 (TCP);
- WS-MAN: 8889 (TCP);
- 16992 Intel AMT HTTP;
- 16993 Intel AMT HTTPS;
- 16994 Intel AMT Redirection/TCP;
- 16995 Intel AMT Redirection/TLS;

- 623 ASF Remote Management and Control Protocol (ASF-RMCP);
- 664 DMTF out-of-band secure web services management protocol;
- 5900 VNC (Virtual Network Computing).

2.10.3. Access to Canonical package repositories and any third party resources required by the infrastructure deployment must be provided unless otherwise specified during the design phase:

- ubuntu-cloud.archive.canonical.com, nova.cloud.archive.ubuntu.com, cloud.archive.ubuntu.com, nova.clouds.archive.ubuntu.com, clouds.archive.ubuntu.com, TCP/80, TCP/443;
- cloud-images.ubuntu.com - TCP/80, TCP/443;
- keyserver.ubuntu.com - TCP/80, TCP/443;
- archive.ubuntu.com - TCP/80, TCP/443;
- security.ubuntu.com - TCP/80, TCP/443;
- usn.ubuntu.com - TCP/80, TCP/443;
- launchpad.net TCP/80, TCP/443;
- git.launchpad.net TCP/22
- ppa.launchpad.net - TCP/80, TCP/443;
- jujucharms.com - TCP/80, TCP/443;
- entropy.ubuntu.com - TCP/443;
- streams.canonical.com - TCP/80, TCP/443;
- public.apps.ubuntu.com - TCP/80, TCP/443;
- https://login.ubuntu.com[4] - TCP/443;
- images.maas.io - TCP/80, TCP/443;
- api.jujucharms.com - TCP/443;
- api.snapcraft.io - TCP/443;
- landscape.canonical.com - TCP/443 (Landscape SaaS only);
- livepatch.canonical.com - TCP/443;
- dashboard.snapcraft.io - TCP/443;
- access to internal NTP server or access to ntp.ubuntu.com - UDP/123, TCP/123;
- access to internal DNS server or access to root DNS servers - UDP/53;
- [Elastic]
  - packages.elastic.co TCP/80, TCP/443;
  - artifacts.elastic.co TCP/80 TCP/443;
  - packages.elasticsearch.org: TCP 80/443;
- [Landscape, Nagios, etc]
  - Outgoing SMTP Relay on customer network: TCP 25;
- [CDN]
  - *.cdn.snapcraftcontent.com/443
  - *.cdn.snapcraft.io/443
- [Kubernetes only]
  - image-registry.canonical.com, TCP/443, TCP/5000;
  - rocks.canonical.com, TCP/443, TCP/5000;
  - quay.io TCP/443 (Kubernetes/Calico/Canal only);

[4] *See* *https://docs.ubuntu.com/snap-store-proxy/en/install*

- *.cloudfront.net TCP/443 (Kubernetes/Calico/Canal only);
- gcr.io TCP/443 (Kubernetes only);
- k8s.gcr.io TCP/443 (Kubernetes only)
- storage.googleapis.com TCP/443 (Kubernetes only);
- auth.docker.io, TCP/443;
- registry-1.docker.io TCP/443 (Kubernetes only);
- Production.cloudflare.docker.com

➢ [NVIDIA only]
- Nvidia.github.io
  - libnvidia-container, nvidia-container-runtime, nvidia-docker2 packages;
- Developer.download.nvidia.com
  - Various NVIDIA driver packages;
  - https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86_64/

➢ [Contrail only]
- https://hub.juniper.net/contrail TCP/443

➢ [Trilio]
- apt.fury.io TCP/443

➢ [BootStack only]
- archive.admin.canonical.com TCP/80 (does not go via VPN)
- portal.admin.canonical.com TCP/443 (does not go via VPN)
- events.pagerduty.com TCP/443 (does not go via VPN) (PagerDuty)
- bootstack-infra.canonical.com TCP/22 for ud-ldap (does not go via VPN).
  - Customer must provide source public IP addresses for BootStack to allow incoming traffic

2.11. Public Network Prefixes

The following are public network prefixes corresponding to Canonical-managed Hostnames:

➢ 91.189.88.0/21 (AS41231);
➢ 162.213.32.0/22 (AS11210).

2.12. Remote Access requirements

2.12.1. Remote access to the environment is mandatory for deployment;

2.12.2. Continuous and stable post-deployment remote access is required for managed services with BootStack;

2.12.3. Remote access can be done via VPN (OpenVPN or IPsec). Once connected, a Canonical representative must be able to reach all the VLANs/subnets present in the design;

2.12.4. The customer must provide all information necessary to open the VPN connection (protocol, hostname/IP of the VPN terminator, credentials, keys).

2.13. Other Requirements

2.13.1. In case HTTP/HTTPS proxy is in place, customer needs to make sure that proxy will allow access to any of the mentioned destinations and their respective ports;

- The proxy server must be able to resolve destination domain names as DNS resolution does not happen at the client side when proxy servers are used;
- The proxy server must allow both HTTP and HTTPS traffic to be sent by clients (if security is a concern for HTTPS traffic tunneling, decryption schemes such as SSLBump need to be set up with relevant certificates communicated to Canonical);
- The proxy server must support at least one of the following ciphersuites:
    - TLS_ECDHE_ECDSA_WITH_AES_256_GCM_SHA384;
    - TLS_ECDHE_ECDSA_WITH_AES_256_CBC_SHA;
    - TLS_ECDHE_ECDSA_WITH_AES_128_GCM_SHA256;
    - TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA;
    - TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384;
    - TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA;
    - TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256;
    - TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA.

2.13.2. A Canonical representative (field engineer) needs to have unrestricted access (TCP/22, TCP/80, TCP/443) to all machines that will be parts of the deployment as well as unrestricted access to out of band management interfaces;

2.13.3. When on-site, a Canonical representative also needs to have unrestricted internet access, at least on TCP/22, TCP/80 and TCP/443 as well as access to uk.sesame.canonical.com (UDP/9062) and us.sesame.canonical.com (UDP/9062). This is not required for a remote deployment.

2.13.4 For Nagios and Landscape to send alerts by email, an SMTP relay should be pre configured by the Customer. The port 25 on the Landscape Server and the Nagios unit should be whitelisted to communicate with this relay. The relay address must be made available to Canonical.

2.14. Internet Network Bandwidth

2.14.1. To ensure a timely deployment, a minimum of 100Mbps of Internet download bandwidth is required for the deployment.

2.15. LDAP Integration

The following prerequisites apply for the Keystone LDAP integration:

2.15.1 An existing LDAP or LDAP enabled AD server deployed and maintained by the customer. This server *must* be reachable by every node within the Private Cloud Build including the infrastructure hosts using the OAM or provisioning network.

2.15.2 An LDAP service account which can perform lookups on the customer AD tree, the credentials of which are provided before the end of the first week of deployment. The account should be provided using the full CN. For example:

'CN=openstack-bind,OU=Service Accounts,OU=Users,DC=example,DC=com'

2.15.3 The base DN for both users and groups as well as the user and group filters with which the Canonical team can form the appropriate query scope to lookup users and groups using the above account. This DN should reflect only the subset of users within the customer AD tree which will have access to the OpenStack cloud. For example:

User DN:
'ou=Users,dc=example,dc=com'
User Filter:
(memberof=cn=openstack-users,ou=Users,dc=example,dc=com)
Group DN:
‘ou=Groups,dc=example,dc=com’
Group Filter:
(&(ObjectClass=group)(memberof=cn=openstack-groups,ou=Groups,dc=example,dc=com))

2.15.4 Prior to and during the deployment access to an Identity Management expert within the customer organization to provide support for the integration and testing process.

Appendix A - Prerequisites checklist table

ID	Description	Verification Comment
2.1	**Infrastructure Nodes**	
2.1.1	3x infrastructure nodes shall be provided and console access made available so that Ubuntu 18.04 can be installed remotely by Canonical;	
2.1.2	Infrastructure nodes must be connected to the out of band management network (with or without native VLAN configuration) and to the network used for PXE booting by other machines (native VLAN, no vlan tags);	
2.1.3	Infrastructure nodes have to have either public internet access or access via HTTP/HTTPS proxy servers to endpoints mentioned in the “Security and Firewalls” section;	
2.1.4	Infrastructure nodes must be dedicated to the target deployment and not used for other purposes.	
2.2	**Node Inventory**	
2.2.1	A detailed inventory sheet listing each node being deployed in the cloud, including BMC address, out of band management information (credentials, including username and password, other parameters) and hostname;	
2.2.2	Rack design: machine and switch placement information in racks and pods;	
2.2.3	A unique hostname for each node used for the Cloud Build Hostnames should belong to a subdomain that would be delegated to the MAAS DNS service or used independently but consistently with a corporate DNS service;	
2.2.4	Datacenter network design related to node connectivity: routing and switching configuration, DNS servers, proxy servers, NTP servers, egress routing setup.	
2.3	**Machine hardware configuration**	
2.3.1	UEFI needs to be prioritized over legacy boot Alternatively legacy boot needs to be completely disabled;	
2.3.2	UEFI/BIOS should be configured such that CPU Frequency is possible to manage via the operating system The relevant settings are vendor-specific and UEFI/BIOS-version specific.	
2.3.3	OOB management card (IPMI) needs to be enabled and configured (note that for IPMI 2.0 maximum password length is 20 characters, and 16 for IPMI 1.5)	
2.3.4	PXE boot needs to be enabled;	

2.3.5	Boot order should be as follows to ensure proper boot process if MAAS is unavailable:	
2.3.6	Intel VT-x virtualization features enabled;	
2.3.7	Hyperthreading enabled (unless the contrary is agreed upon);	
2.3.8	Intel VT-d enabled if supported;	
2.3.9	SR-IOV enabled (if agreed during the design session and supported).	
2.3.10	1xRAID-1 created (to hold the O/S) on all the nodes (Infra + Cloud)	
2.3.11	Serial console redirection should be enabled In order for the entire boot process to be visible via IPMI, both “serial redirection” and “serial over lan” need to be enabled Settings should be consistent across all nodes;	
2.3.12	Power policy in case of loss of AC power should be always-on (not ‘last-state’ or ‘always-off’).	
2.4	**Firmware**	
2.4.1	All hardware (motherboard UEFI/BIOS, NIC, NVMe, RAID controllers) is updated to have the latest firmware version supplied by the hardware vendor unless a legacy version is explicitly required by Canonical for the purpose of security or deploying its tooling;	
2.4.2	Hardware should not have mixed firmware versions on different servers unless there are different generations or models used which prevents this.	
2.5	**Certificates**	
2.5.1	An Internal CA certificate (or a certificate chain) for a CA that issued LDAP/AD TLS certificates;	
2.5.2	Artifacts necessary for a PKI setup (I or II):	
2.5.3	If TLS termination is used on corporate HTTP proxy servers, issuer CA certificates used for validation of proxy server certificates must be provided;	
2.5.4	If SSLBump configuration is used at a corporate proxy server to dynamically decrypt HTTPS traffic and generate destination server certificates to impersonate them, a CA certificate or chain used to validate these dynamic certificates needs to be provided by the Customer.	
2.6	**Physical Connectivity**	
2.6.1	Ensure that cabling is aligned with the physical network layout.	
2.6.2	Depending on the design, two or three bonds connected to the Service Fabric switch pair (“bond0” and “bond1” below) where ports of an individual bond are connected to different switches running in the Multi-Chassis Link Aggregation (MLAG, MC-LAG, Virtual PortChannel or other vendor-specific names.) configuration;	

2.6.3	One bond (“bondM”) connected to the Management Fabric (BoR) switch pair (with MLAG); A single 1G ethernet connection to the Management Fabric switch pair for the BMC (IPMI, DRAC, iLO etc) interface;	
2.7	**Networking Switching**	
2.7.1	Read-only access to network switches needs to be provided in order to facilitate debugging of possible routing and switching issues Alternatively, a copy of switch configs needs to be provided (without any sensitive data like credentials) and subsequent copies if changes are made.	
2.7.2	A chosen network layout needs to provide resilient, high bandwidth network connectivity for a defined set of separate Layer 2 broadcast domains delivered to every host.	
2.7.3	Switch ports must be configured to support bonded host interfaces;	
2.7.4	Switch ports must support IEEE 802.1q VLAN tagged and untagged traffic;	
2.7.5	A single layer 2 network and associated layer 3 subnet must be designated for OAM (Operation, Administration, and Maintenance);	
2.7.6	All BMC and operating system provisioning interfaces must be connected or routed to the OAM network We recommend to implement separate layer 2 broadcast domains (such as VLANs) for BMC and OAM boot networks, however, the requirement is for infrastructure nodes to have unrestricted network access to both of those networks;	
2.7.7	If host OAM interfaces are to be bonded, the switch ports must be configured to support IEEE 802.3ad LACP with LACP rate set to “Fast timeout” and LACP fallback (sometimes called “LACP bypass”, “LACP suspend individual”, “LACP force-up” or “LACP fallback” depending on the hardware vendor) options enabled;	
2.7.8	In order to set the LACP fallback option mentioned above it is important to set it with the lowest possible priority value in order to bring the LAG up and keeping a single port active until the expiry of a timeout period, otherwise the server will not be able to boot from PXE;	
2.7.9	If STP is used, ensure the PortFast (or equivalent) option is enabled on server facing ports;	
2.7.10	Port security should not be enabled on OAM switch ports as traffic from both infrastructure and non-infrastructure nodes will contain MAC addresses of virtual machines and containers which may lead to switchport interface downtime due to policy enforcement;	
2.7.11	Different types of traffic generated by the deployed infrastructure will require Individual broadcast domains (such as VLANs):	
2.7.12	Ensure MTU configuration is consistent throughout the Customer network and configured to the agreed upon setting laid out in the design phase;	

2.7.13	Jumbo frames must be enabled (if supported) on Ceph access and Replication VLANs in addition to the underlay network for the overlay traffic.	
2.8	**Network layer and Reachability**	
2.8.1	Verify that a logical network configuration matches the one agreed upon during the design phase;	
2.8.2	A subnetting design should be provided - every VLAN needs to have a subnet allocation enough to cover hosts and infrastructure containers and virtual machines used for the deployment;	
2.8.3	If a layer 3 topology like Leaf-Spine is used make sure that each VLAN configured on each switch terminating the L2 domain has an associated subnet allocation and routing is implemented between the different segments by using a combination of a routing protocol and possibly VRF-lite (or similar) functionality In this case, gateway addresses for each subnet must be provided by the Customer Alternatively, an EVPN setup needs to be provided;	
2.8.4	Ensure there are no DHCP agents or relay agents in broadcast domains (VLANs) that MAAS will be serving DHCP on The expectation is that MAAS will provide DHCP services for each of the networks defined in the Cloud configuration or will otherwise manage static IP address allocations for provisioned servers;	
2.8.5	If any IPv6 usage is expected (control plane or data plane), this has to be discussed in advance;	
2.8.6	Verify NTP server connectivity from the target environment - correctly configured NTP is critical for the operation of any cloud platform Upstream NTP should be configured to provide resilient, high quality NTP sources;	
2.8.7	Verify corporate or public DNS server connectivity from the target environment; Corporate DNS should resolve public domains Verify HTTP proxy server or direct public internet connectivity by performing basic deb and snap package installation on an Ubuntu system from an installed infrastructure node:	
2.9	**DNS**	
2.9.1	Delegation of subdomains to name servers deployed as a part of the cloud needs to be done For that a subdomain for API endpoints and a separate subdomain for instances deployed in OpenStack should be dedicated per region (if there are multiple cloud regions), such as:	
2.9.2	Use of .local top-level domain is highly discouraged as it is reserved for multicast DNS (RFC6762) and affects systemd-resolved behavior such that it does not forward DNS queries to upstream DNS servers;	
2.9.3	In order to access cloud-provided DNS for auto-generated instance name-based DNS records, corporate DNS servers must be configured to use subdomain delegation to a set of name servers in a dedicated address range Appropriate NS records and “glue” records must be created for delegated subdomains;	

2.9.4	In order to access public OpenStack API endpoints a set of virtual IPs for OpenStack services (keystone, nova etc.) in a given region needs to have names assigned with address records added to the corporate name servers.	
2.1	**Security and Firewalls**	
2.10.1	All firewall and iptables rules communicated in the design phase are consistent with deployed hardware and a copy of those rules are provided to Canonical;	
2.10.2	IPMI/iLO/AMT or any other baseboard management traffic must be allowed between all infrastructure nodes and server out of band interfaces	
2.10.3	Access to Canonical package repositories and any third party resources required by the infrastructure deployment must be provided unless otherwise specified during the design phase:ubuntu-cloud.archive.canonical.com, nova.cloud.archive.ubuntu.com, cloud.archive.ubuntu.com, nova.clouds.archive.ubuntu.com, clouds.archive.ubuntu.com, TCP/80, TCP/443;	
2.11	**Public Network Prefixes**	
2.12	**Remote Access requirements**	
2.12.1	Remote access to the environment is mandatory for deployment;	
2.12.2	Continuous and stable post-deployment remote access is required for managed services with BootStack;	
2.12.3	Remote access can be done via VPN (OpenVPN or IPsec) Once connected, a Canonical representative must be able to reach all the VLANs/subnets present in the design;	
2.12.4	The customer must provide all information necessary to open the VPN connection (protocol, hostname/IP of the VPN terminator, credentials, keys).	
2.13	**Other Requirements**	
2.13.1	In case HTTP/HTTPS proxy is in place, customer needs to make sure that proxy will allow access to any of the mentioned destinations and their respective ports;	
2.13.2	A Canonical representative (field engineer) needs to have unrestricted access (TCP/22, TCP/80, TCP/443) to all machines that will be parts of the deployment as well as unrestricted access to out of band management interfaces;	
2.13.3	When on-site, a Canonical representative also needs to have unrestricted internet access, at least on TCP/22, TCP/80 and TCP/443 as well as access to uk.sesame.canonical.com (UDP/9062) and us.sesame.canonical.com (UDP/9062) This is not required for a remote deployment.	
2.15	**LDAP Integration**	