Shared storage (with ceph)

Configuration

The inventory may define these hosts to run the step completely:

- `clients`: Set of hosts that must be *ceph-client*. These hosts will access to the storage cluster
- `mons`: Set of hosts to that must be *ceph-mon*. These hosts will maintain a map of the state of the cluster
- `osds`: Set of hosts to that must be *ceph-osd*. These hosts will interact with the logical disk to stock the data

More details in the documentation here.

The inventory must define these variables to run the step:

- `configure_firewall`: Boolean to configure the firewall (`true` by default)
- `ceph_origin`: Origin of the ceph installed files (must be set to `distro`). SEAPATH installs ceph with an installer (see the installation section)
- `ceph_osd_disk`: Device to stock datas (only for *ceph-osd* hosts), it's on this disk ceph will build a rbd. To success the CI, the path should be in `/dev/disk/by-path`
- `cluster_network`: Address block to access to cluster network
- `dashboard_enabled`: Boolean to enable a dashboard (must be set to `false`)
- `devices`: List of devices to use for the shared storage. All specified devices will be used entirely
- `lvm_volumes`: List of volumes to use for the shared storage. To use to replace `devices` to take a part of the device
- `monitor_address`: Address where the host will bind
- `ntp_service_enabled`: Boolean to enable the NTP service (must be set to `false`). SEAPATH installs a specific NTP client and configure it
- `osd_pool_default_min_size`: Minimal number of available OSD to ensure cluster success (best: `ceil(osd_pool_default_size / 2.0)`)
- `osd_pool_default_size`: Number of OSD in the cluster
- `public_network`: Address block to access to public network

Volumes specifications

If `lvm_volumes` is defined, the `devices` variables is ignored.

When a volume is defined for the shared storage, some fields should be set for *seapath-ansible* and *ceph-ansible*.

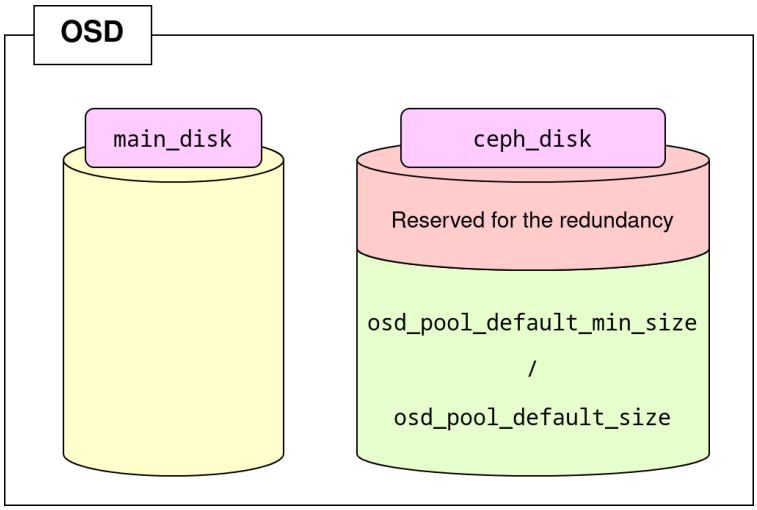
- `data`: Logical volume to use for the shared storage (*ceph-ansible* variable)
- `data_vg`: Volume group where the logical volume is (*ceph-ansible* variable)
- `data_size`: Size of the logical volume (in megabytes by default). Change the unit with the appropriate suffix
- `device`: Device to use to create the logical volume
- `device_number`: Number of the partition
- `device_size`: Size of the partition to stock the logical volume

Override configuration

ceph offers to override the configuration with `ceph_conf_override`:

Override global configuration

- `mon_osd_min_down_reporters`: (must be set to `1`)
- `osd_crush_chooseleaf_type`: (must be set to `1`)
- `osd_pool_default_min_size`: Minimal number of available OSD to ensure cluster success (best: `ceil(osd_pool_default_size / 2.0)`)
- `osd_pool_default_pg_num`: (must be set to `128`)
- `osd_pool_default_pgp_num`: (must be set to `128`)
- `osd_pool_default_size`: Number of OSD in the cluster



Override mon configuration

- `auth_allow_insecure_global_id_reclaim`: Boolean (must be set to `false`)

Override osd configuration

- `osd_max_pg_log_entries`: (must be set to `500`)
- `osd_min_pg_log_entries`: (must be set to `500`)
- `osd_memory_target`: Size of the memory (in bytes)

Ceph provides ansible rules to configure the software, you can read the documentation here.

If this step is failed, you must restart it at the previous step. Use snapshot LVM to do this.

At end of this step, make sure that:

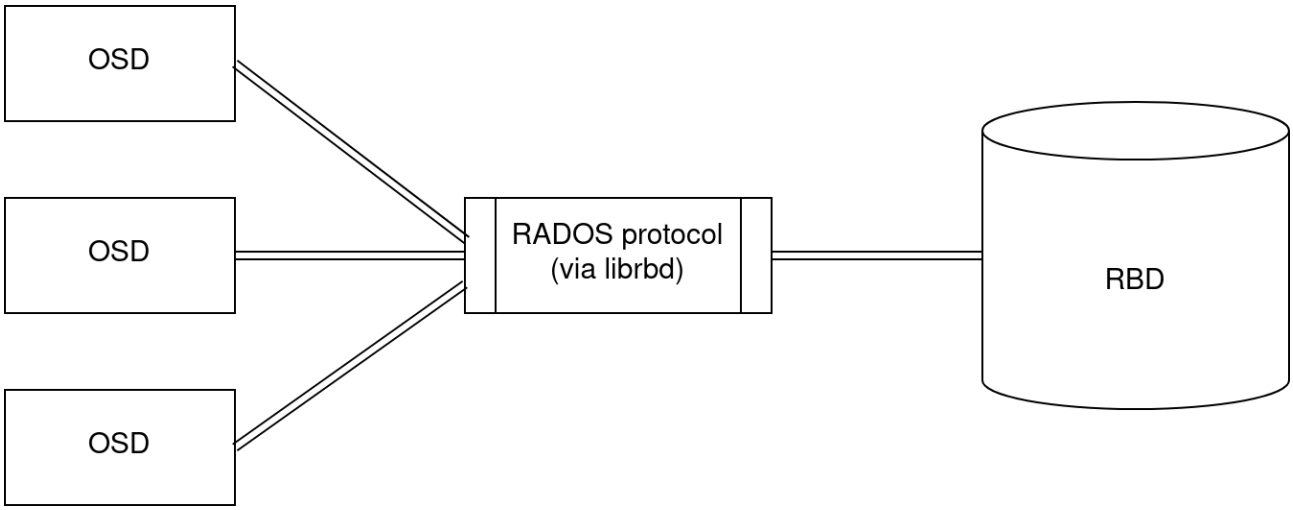
- there is a `rbd` pool with this command: `ceph osd pool ls`.
- there is a `ceph` pool with this command: `virsh pool-list`.

RADOS Block Devices

During this step, ceph will build a RADOS block device (RBD) from `ceph_osd_disk`. A storage entry (a pool) will be automatically generated for libvirt. When the service will be started, the hypervisor should be used to launch VMs.

On the `ceph_osd_disk` of all machines, there are the same data.

This disk will be used with the librbd library provided by ceph.



More details here.