VSM Introduction & Major Features

Wang, Yaguang

Ferber, Dan

March 2015

Agenda

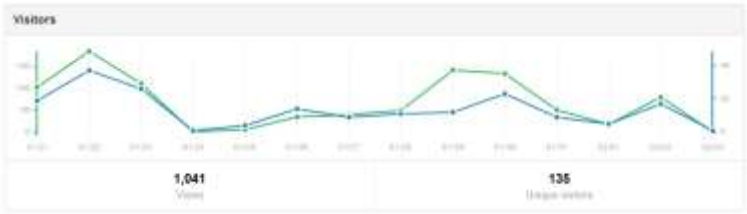
- Overview
- Architecture & Key Components
- Major Features

Overview

VSM Overview

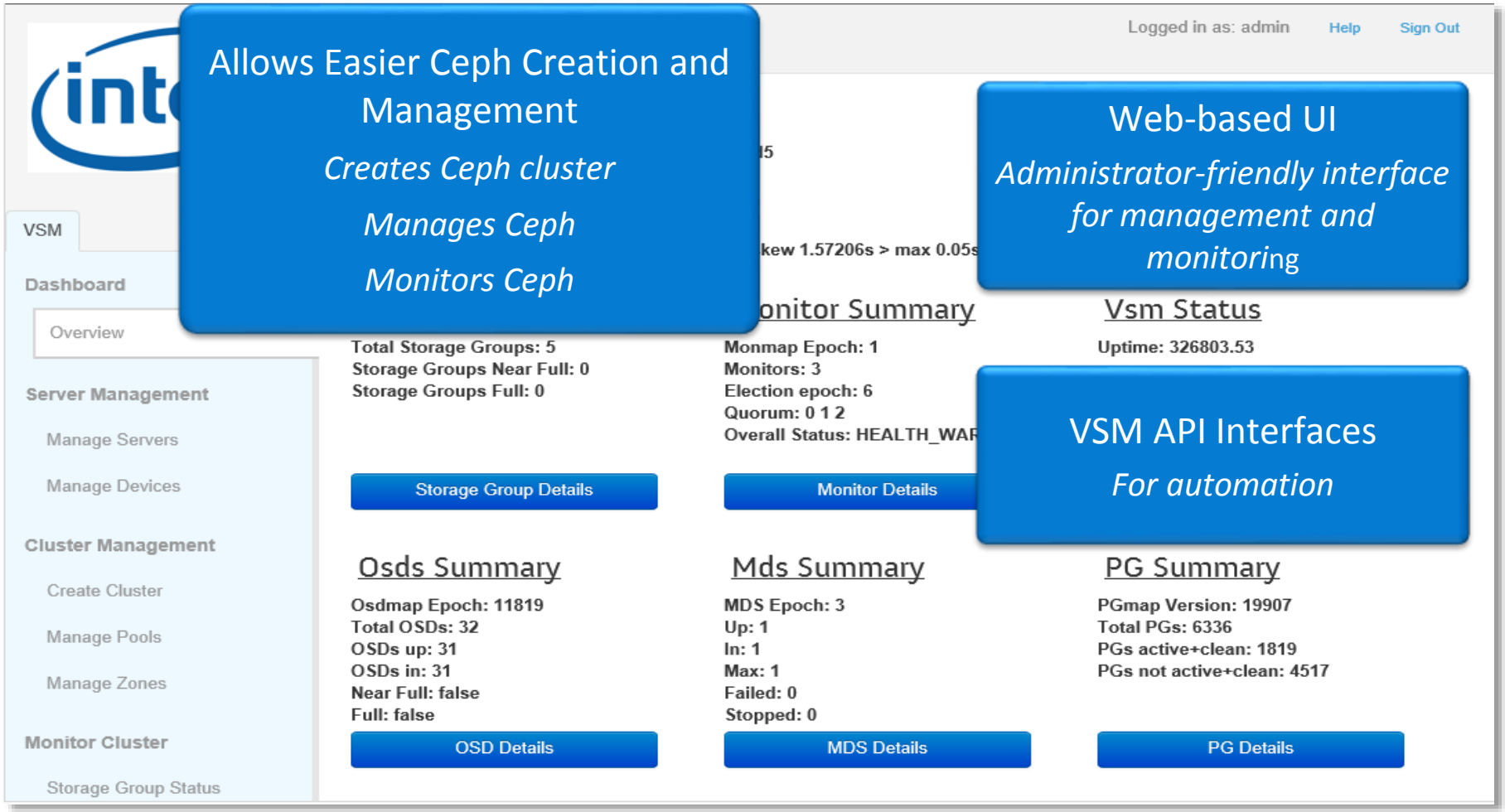
- VSM (Virtual Storage Manager) is an open source ceph management tool developed by Intel, and announced on 2014 Nov's openstack Paris summit. It is designed to help make the day to day management of Ceph easier for storage administrators.

- Home page:
  - https://01.org/virtual-storage-manager
- Code Repository:
  - https://github.com/01org/virtual-storage-manager
- Issue Tracking:
  - https://01.org/jira/browse/VSM
- Mailing list:
  - http://vsm-discuss.33411.n7.nabble.com/



VSM Overview

- VSM is designed for bundled ceph storage appliance, it creates ceph cluster for management and monitoring. So far it runs on Centos 6.5 with Ceph firefly and openstack havana icehouse.



Typical VSM-Managed Cluster

VSM Controller – Dedicated server or server instance

Server Nodes

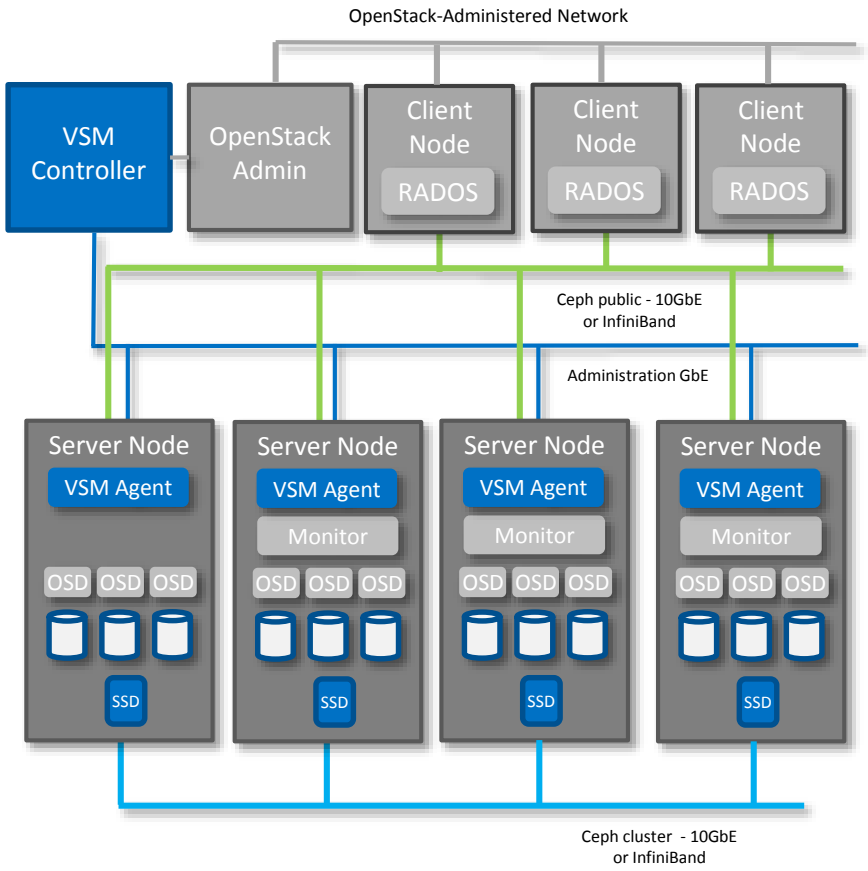
- Are members of VSM-managed Ceph cluster
- VSM agent runs on every server in VSM-managed cluster

Network Configuration

- Ceph public subnet – Carries data traffic between clients and Ceph cluster servers
- Administration subnet – Carries administrative communications between VSM controller and agents
  - Also administrative comms between Ceph daemons
- Ceph cluster subnet – Carries data traffic between Ceph storage nodes – replication and rebalancing

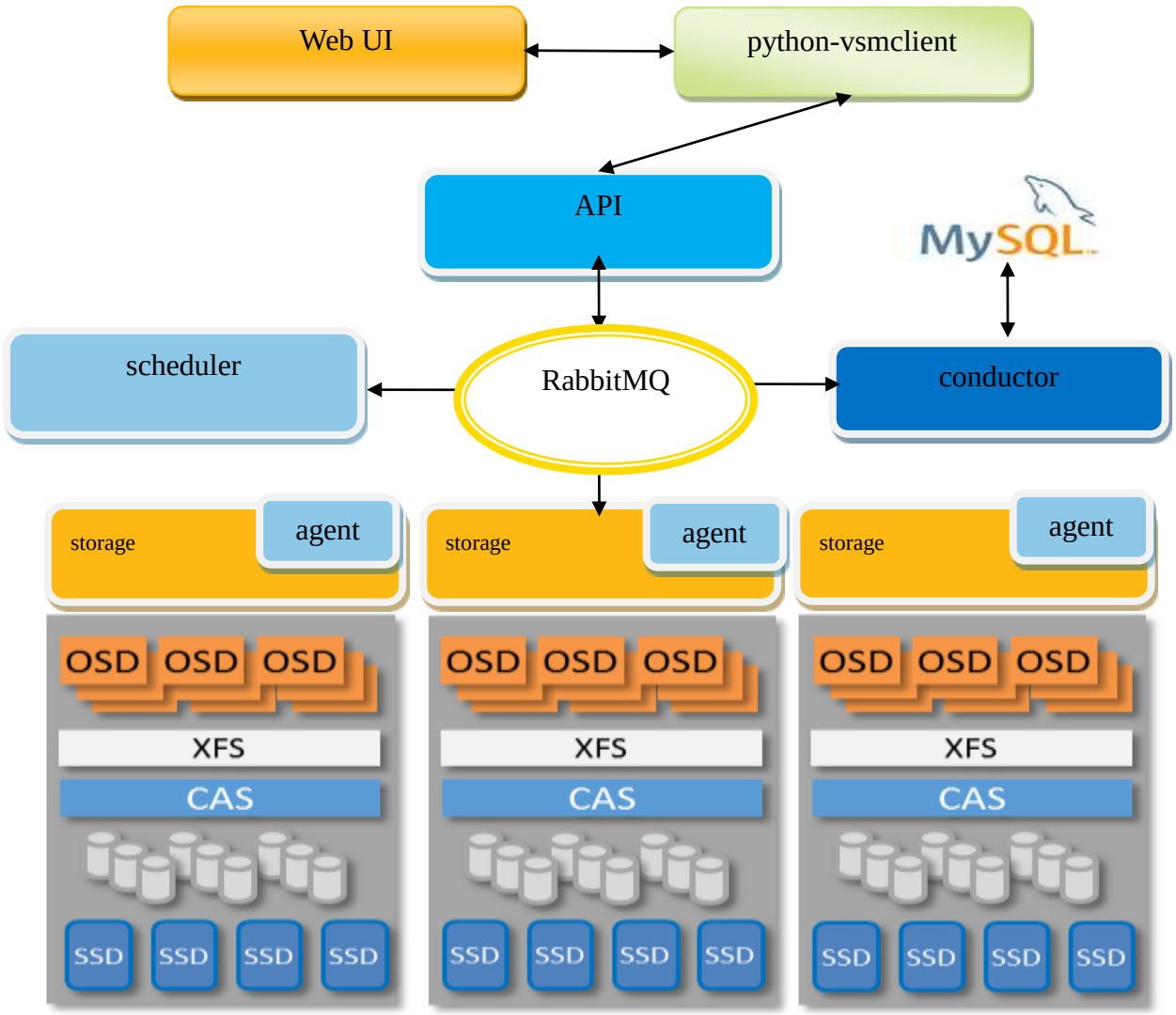
OpenStack admin (optional)

- Optionally connected to VSM via SSH connection
- Allows VSM to “tell” OpenStack about Ceph storage pools



Architecture & Key components

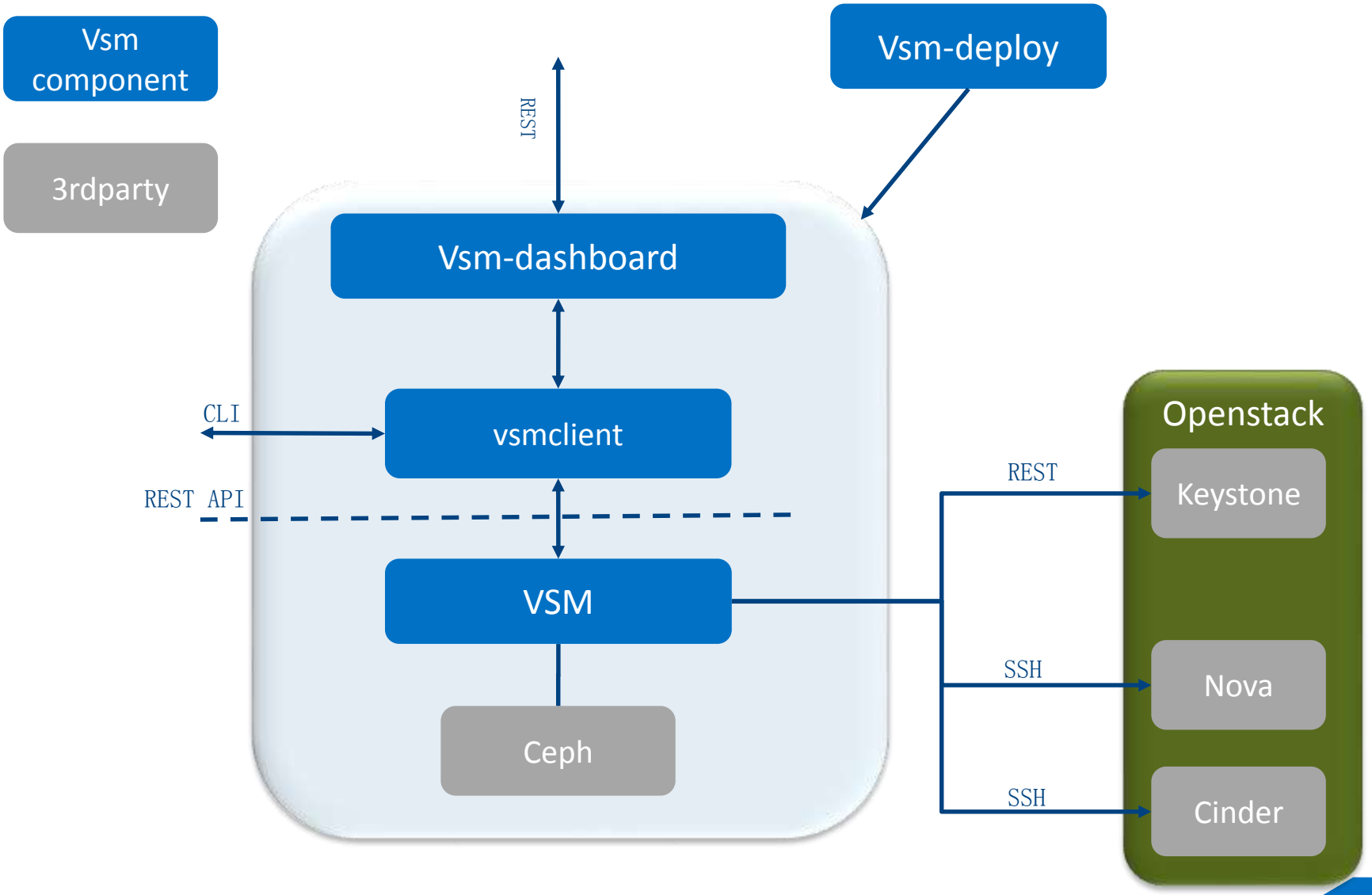
VSM Architecture



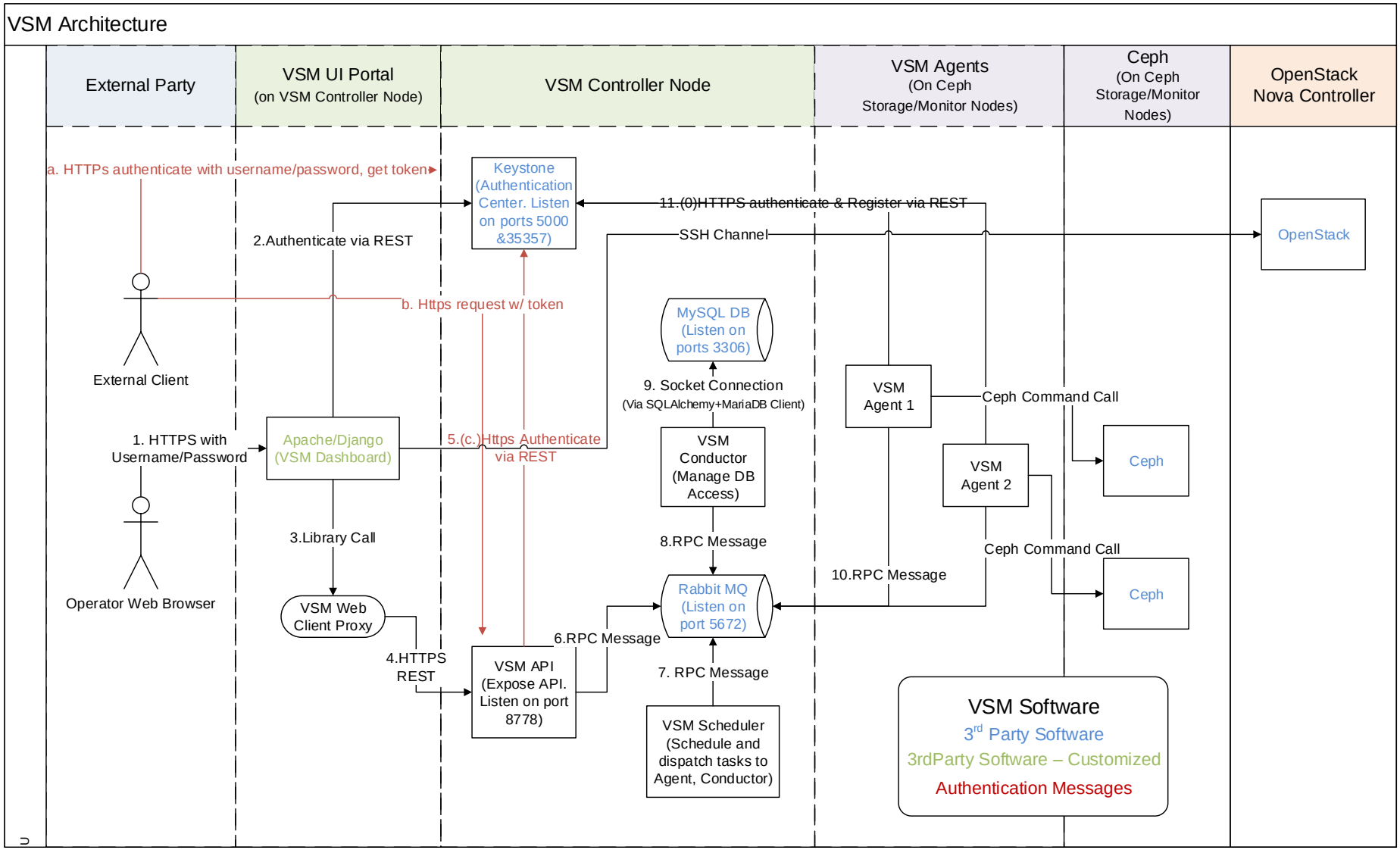
Key Components: Overview

- *Python-vsmclient*
  - This is a client for the vsm API, it consists of
    - a Python API (the vsmclient module),
    - a command-line script (vsm). Each implements 100% of the vsm API.
- *Vsm*
  - A major module for ceph management
- *Vsm-dashboard*
  - web based management interface for VSM.
- *Vsm-deploy*
  - The ceph deployment tool kit provided by VSM .

Key Components: relationships

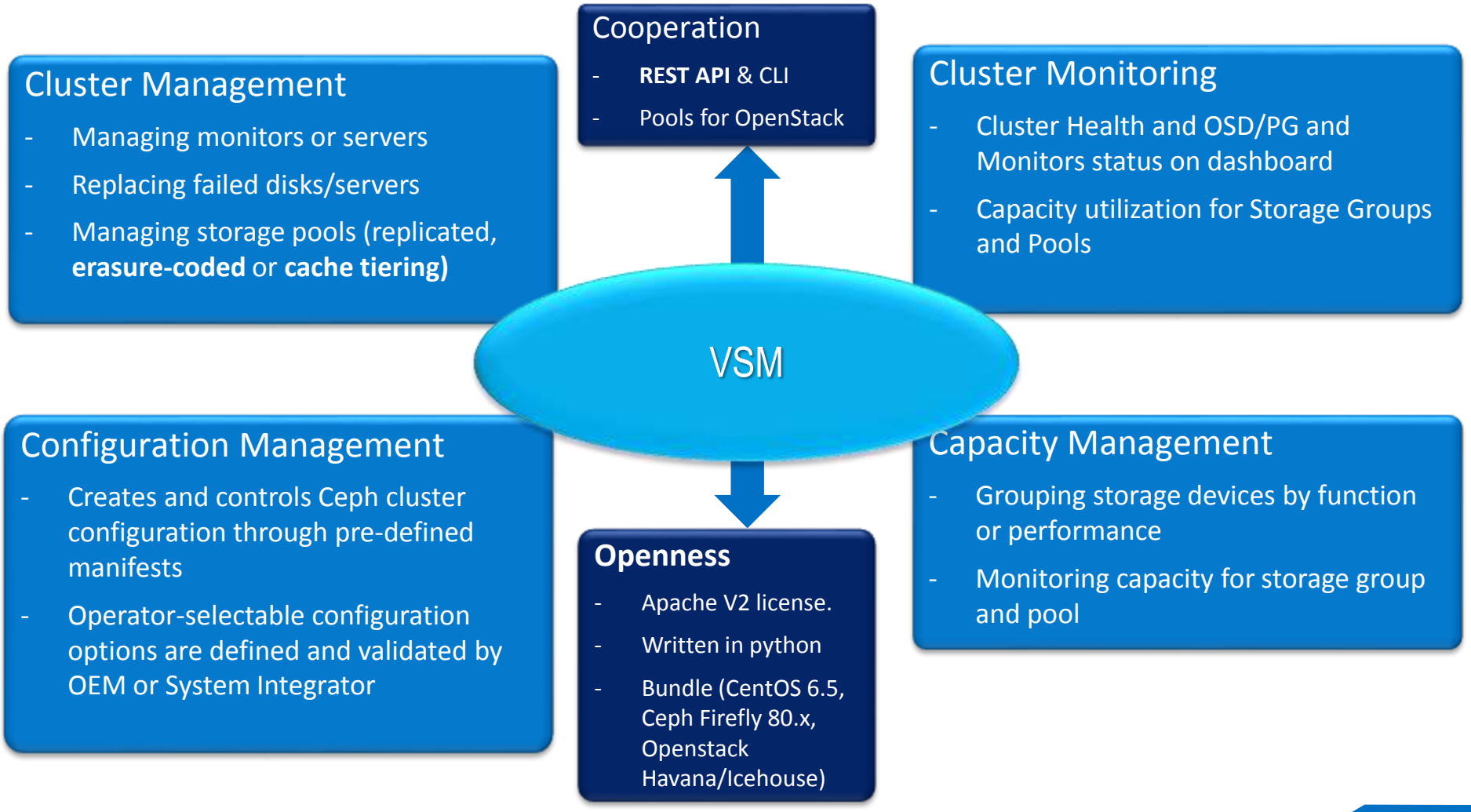


Key Components: interaction flow



Major Features

Major Features



What it does…

Web-based UI

- Administrator-friendly interface for management and monitoring

Configuration management

- Storage Groups aggregate similar drives
- Zones aggregate drives within failure domains

Capacity Management

- Pools segregate allocation by application
- Capacity utilization by storage group and pool

Cluster Management

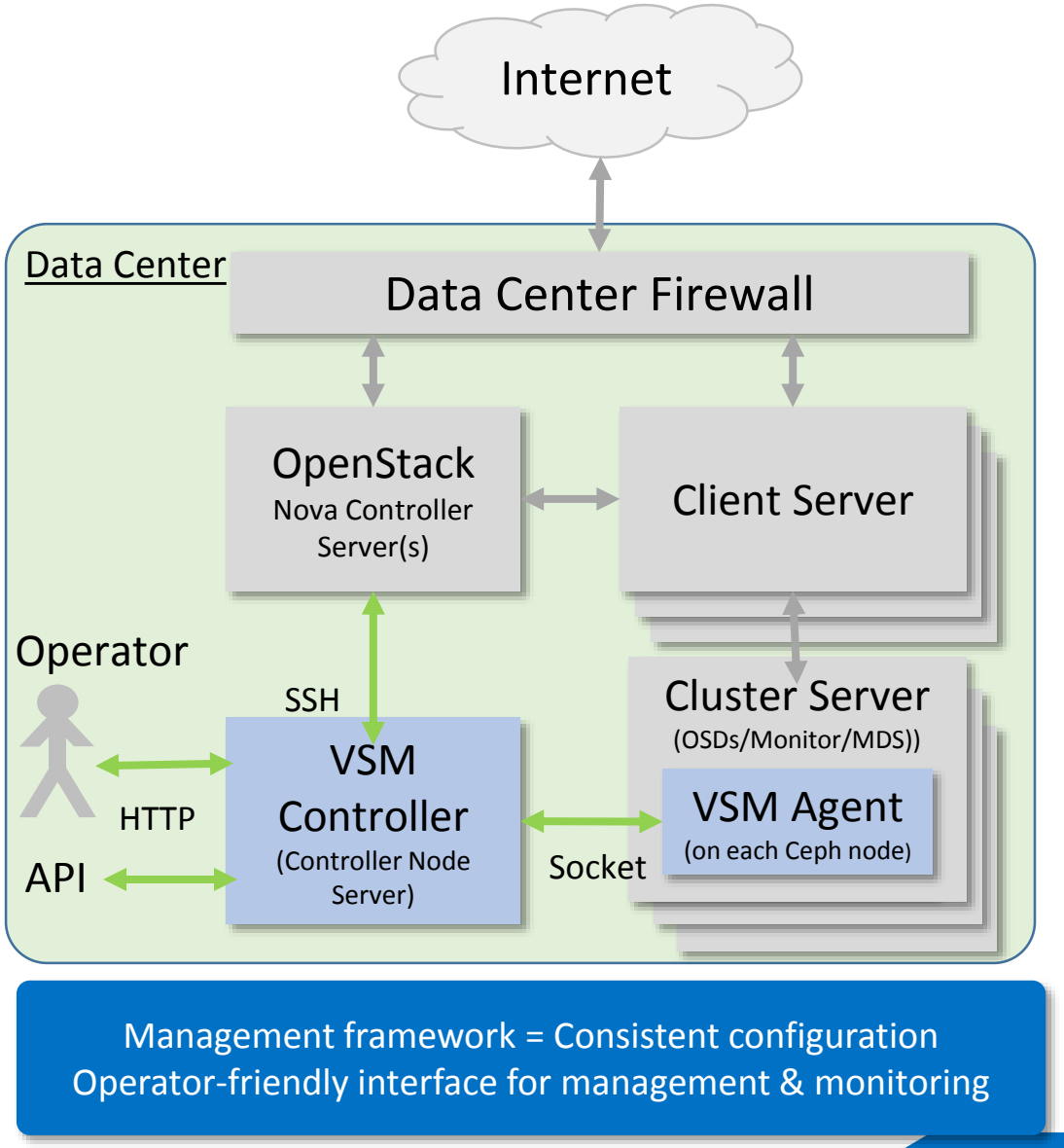
- Manage capacity growth
- Manage Server & disk maintenance

Cluster Monitoring

- Capacity and performance
- OSD, PG, and monitor state

VSM APIs

- Software interface supports automation



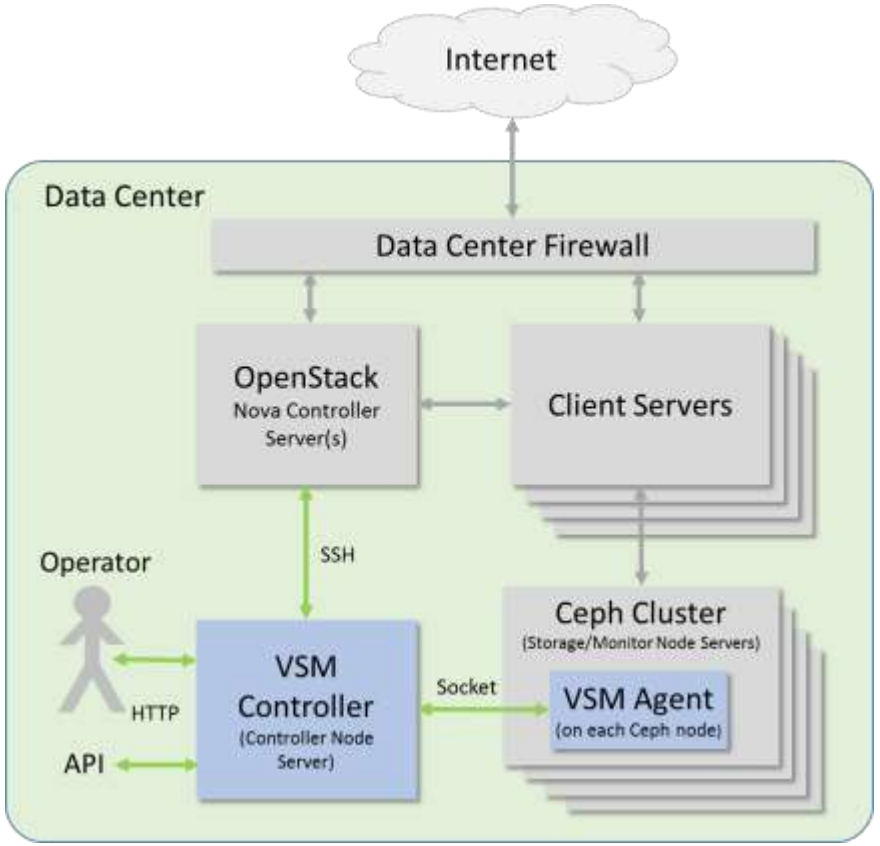
What it is…

VSM Controller Software

- Runs on dedicated server (or server instance)
- Connects to Ceph cluster through VSM agent
- Connects to OpenStack Nova controller (optional) via SSH
- Never touches clients or client data

VSM Agent Software

- Runs on every server in the Ceph cluster
- Relays server configuration & status information to VSM controller



Managing Servers and Disks

Servers can host more than one type of drive, drives with similar performance characteristics are identified by **Storage Class**.

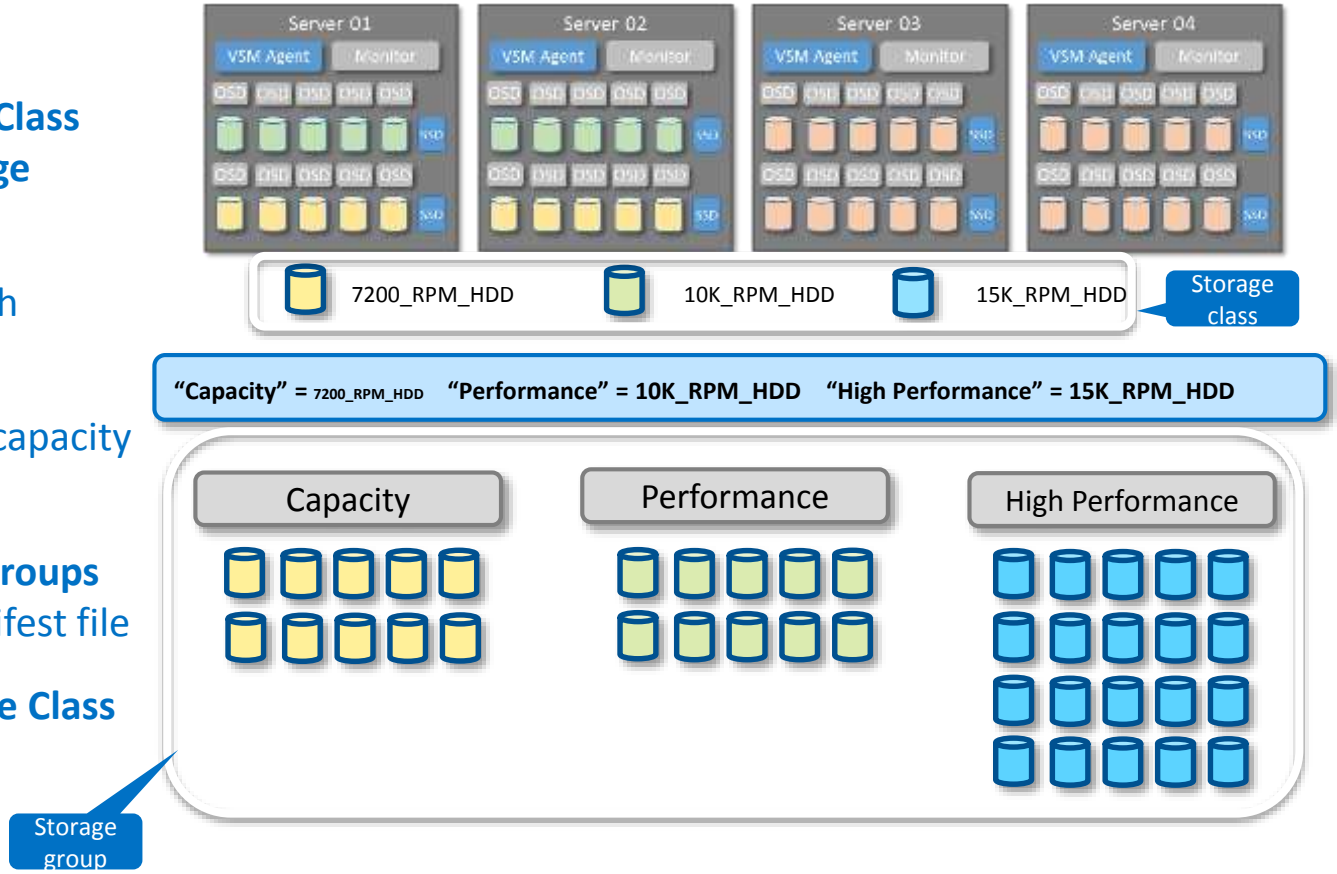
Drives with the same **Storage Class** are grouped together in **Storage Groups**

Storage Groups are paired with specific **Storage Classes.**

VSM monitors Storage Group capacity utilization, warns at threshold.

Storage Classes and **Storage Groups** are defined in the cluster manifest file

Drives are identified by **Storage Class** in the server manifest file



Managing Failure Domains

Servers can be grouped into failure domains. In VSM, failure domains are indented by **zones**.

Zones are placed under each **Storage Group**

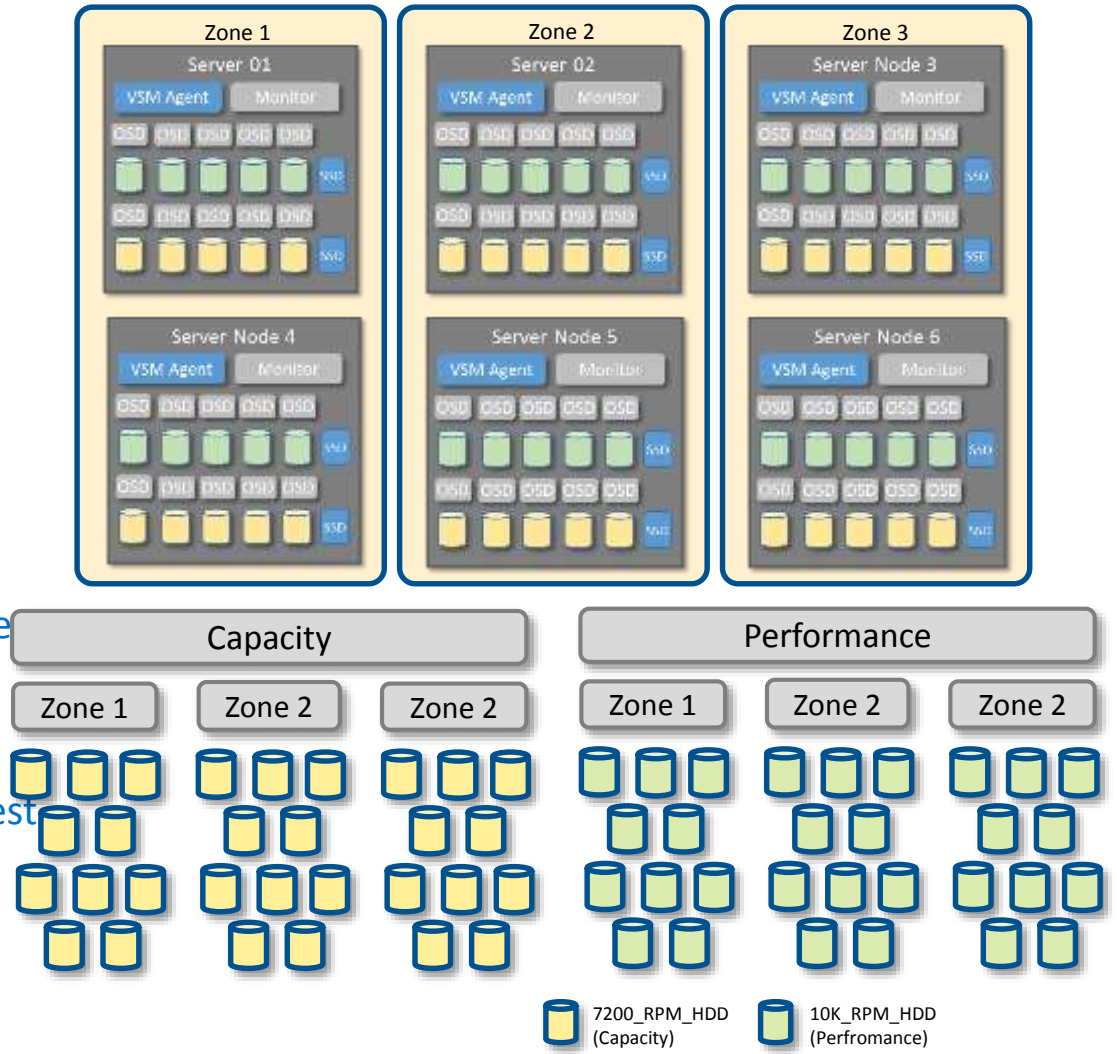
Drives in each **zone** are placed in their respective **storage group**

In the example at right, six servers are placed in three different zones. VSM creates three zones under each storage group, and places the drives in their respective storage groups and zones.

Zones are defined in the cluster manifest file

Zone membership is defined in the server manifest file

One Zone with server-level replication



VSM Controller: Cluster Manifest

VSM Controller runs on dedicated server or server instance

- VSM controller uses the cluster manifest file to define the cluster meta information, like storage classes, storage groups, zones, subnet, and also cluster settings, like capacity threshold for warning, ec and cache tier settings.

Meta information

- The cluster manifest defines the storage classes, storage groups and zones to be used in the ceph cluster, also it defines the subnets for management, public or cluster.

Cluster Setting

- The cluster manifest provides a few settings to help tune cluster behaviors, like capacity thresholds, heartbeat intervals.
- Also the default settings for ec pool and cache tiering are also covered.

VSM Controller: Cluster Manifest File

Cluster Manifest File

- Resides on the VSM controller server.
- Tells VSM how to organize storage devices, how the network is configured, and other management details

```
[storage_class]
7200_rpm_sata
10krpm_sas
ssd
ssd_cached_7200rpm_sata
ssd_cached_10krpm_sas

[storage_group]
#format: [storage group name]  ["user friendly storage group name"] [storage
class]
high_performance  "High_Performance_SSD"    ssd
capacity          "Economy_Disk"            7200_rpm_
performance       "High_Performance_Disk"   10krpm_sas
value_performance "High_Performance_Disk_with_ssd_cached_Acceleration"
ssd_cached_10krpm_sas
value_capacity    "Capacity_Disk_with_ssd_cached_Acceleration"
ssd_cached_7200rpm_sata

[cluster]
cluster_a

[file_system]
xfs

[management_addr]
192.168.123.0/24

[ceph_public_addr]
192.168.124.0/24

[ceph_cluster_addr]
192.168.125.0/24

[zone]
Zone0
zone1
```

Storage classes defined

Storage groups

Cluster name

Data disk file system

Network configuration

zone

```
[settings]
storage_group_near_full_threshold  65
storage_group_full_threshold  85
ceph_near_full_threshold  75
ceph_full_threshold  90
pg_count_factor 100
heartbeat_interval 6
osd_heartbeat_interval 7
osd_heartbeat_grace 21

[ec_profiles]
#format: [profile-name] [path-to-plugin] [plugin-name]
[pg_num value] [json format key/value]
#the key/value strings
spacesdefault_profile                                    code
jerasure  3
{"k":2,"m":1,"technique":"reed_sol_van"}#profile_name2
/usr/lib64/ceph/erasure-code  jerasure  6
{"k":4,"m":2,"technique":"reed_sol_van"}

[cache_tier_defaults]
ct_hit_set_count 1
ct_hit_set_period_s 3600
ct_target_max_mem_mb 1000000
ct_target_dirty_ratio 0.4
ct_target_full_ratio 0.8
ct_target_max_objects 1000000
ct_target_min_flush_age_m 10
ct_target_min_evict_age_m 20
```

settings

Ec pool profiles

Cache tier default parameters

VSM Agent: Server Manifest

VSM Agent runs on every server managed by VSM

- VSM Agent uses the server manifest file to identify and authenticate with the VSM controller, and describe server configuration

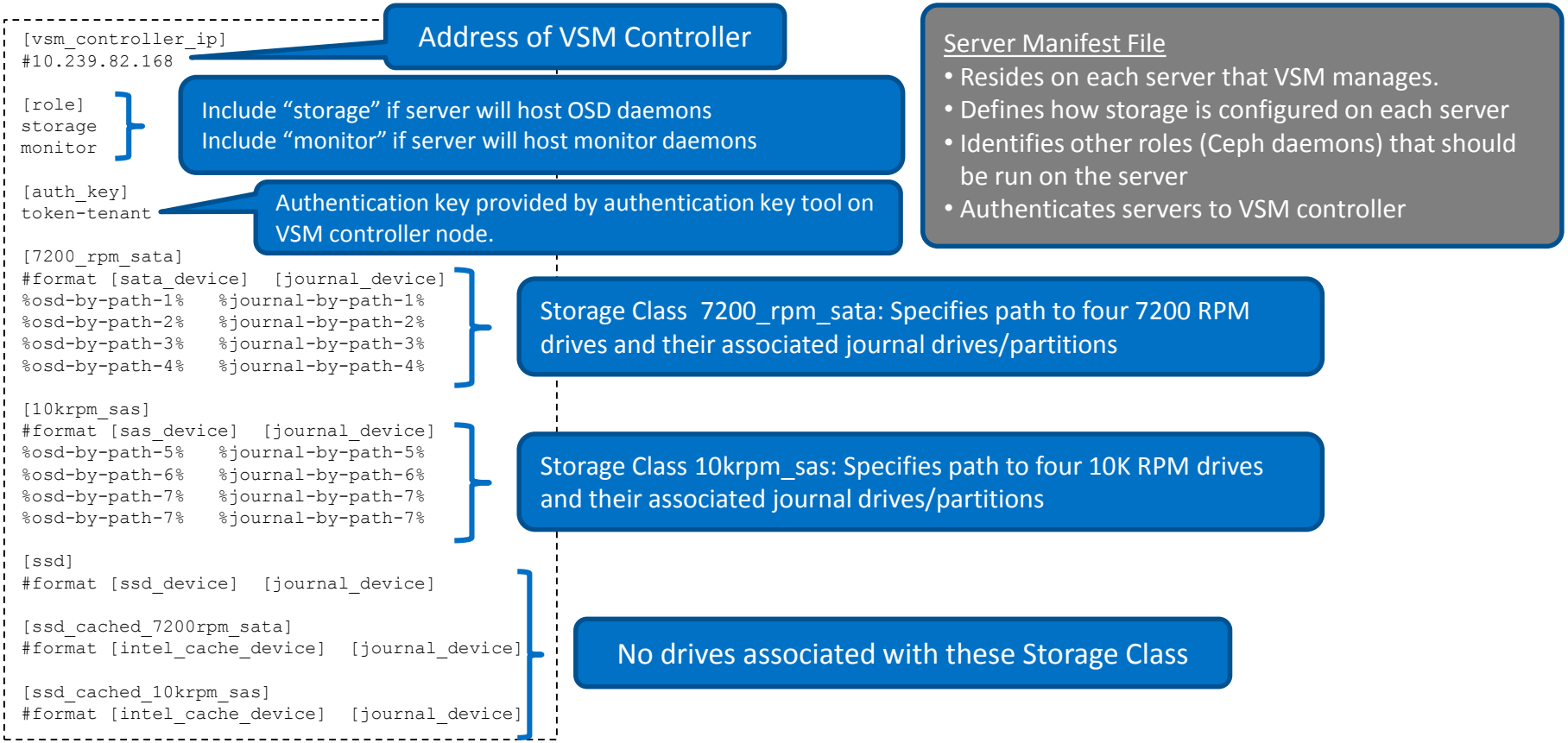
Discovery and authentication

- To be added to a cluster, the server manifest file must contain the IP address of the VSM controller, and a valid authentication key
  - The authentication key is generated by vsm controller in advance, to be applied into server manifest before launching agent daemon.

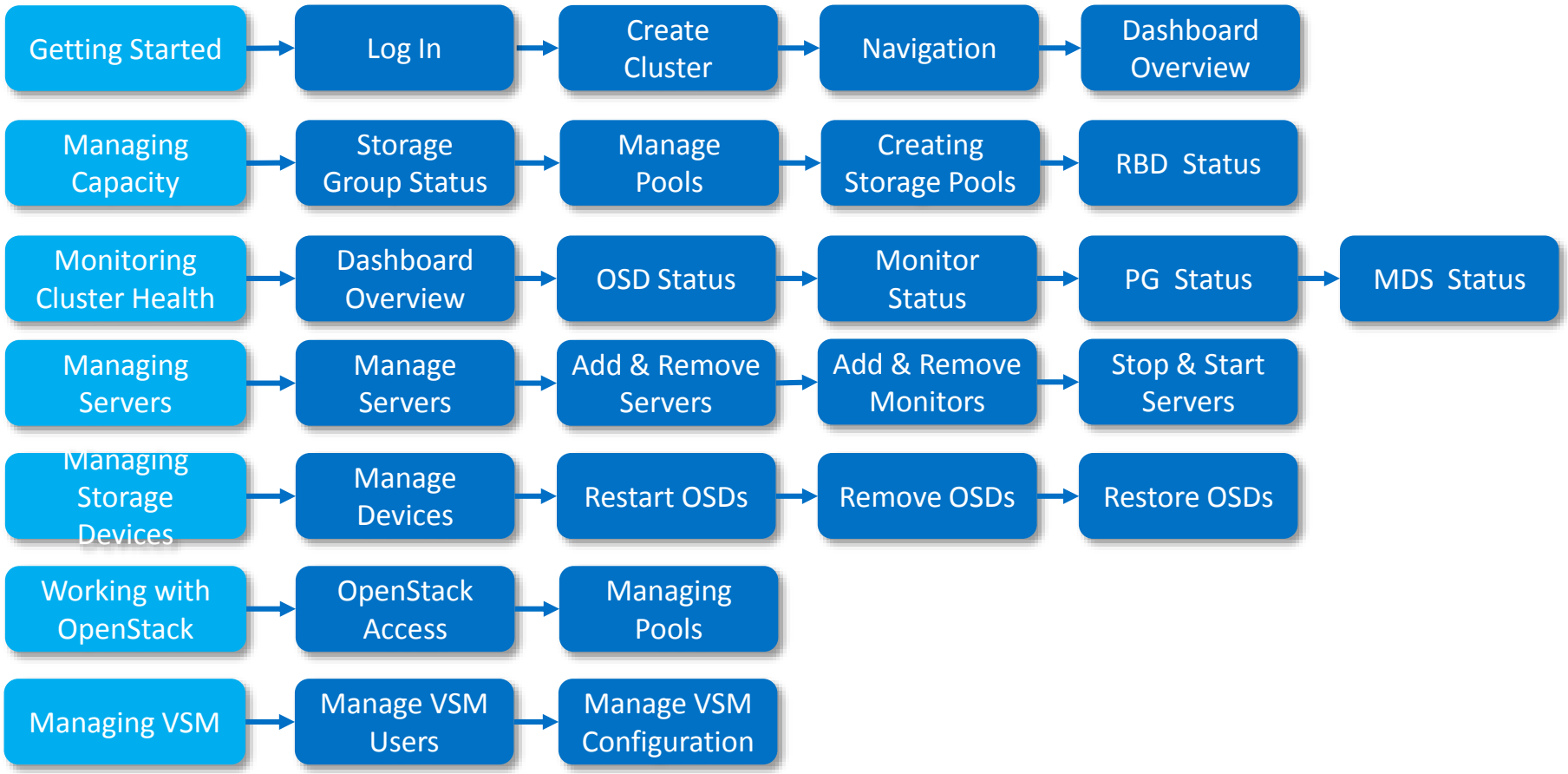
Server Configuration

- VSM relies on the server manifest to identify and classify data devices and associated journals. VSM does not have knowledge of how storage devices have been partitioned.
- Devices and partitions are specified “by path” to ensure that paths remain constant in the event of a device removal or failure.

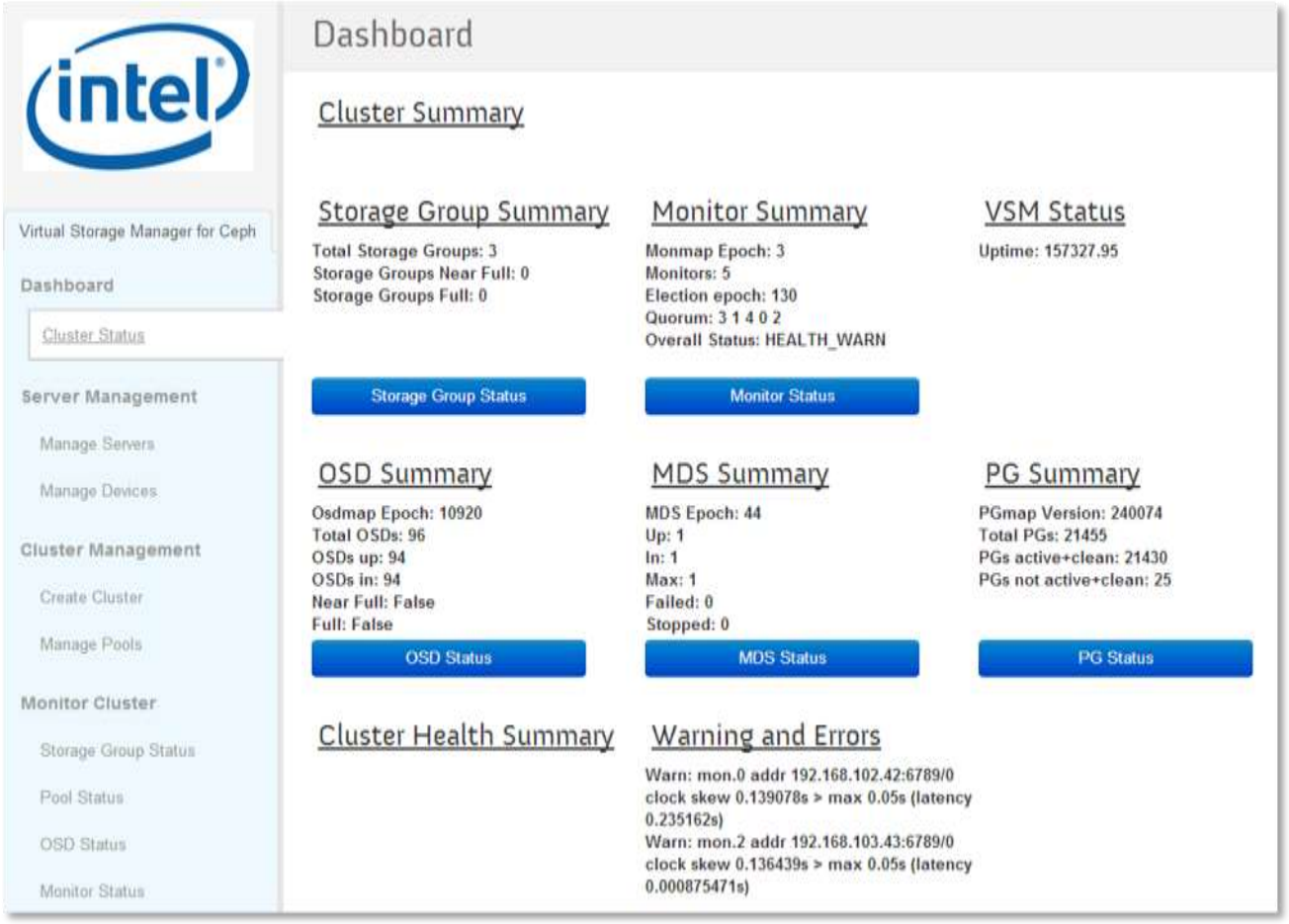
VSM Agent: Server Manifest



VSM Operation Map



Live Demo


Dashboard
Cluster Summary
Virtual Storage Manager for Ceph
Dashboard
Cluster Status
Server Management
Manage Servers
Manage Devices
Cluster Management
Create Cluster
Manage Pools
Monitor Cluster
Storage Group Status
Pool Status
OSD Status
Monitor Status
Storage Group Summary
Total Storage Groups: 3
Storage Groups Near Full: 0
Storage Groups Full: 0
Storage Group Status
Monitor Summary
Monmap Epoch: 3
Monitors: 5
Election epoch: 130
Quorum: 3 1 4 0 2
Overall Status: HEALTH_WARN
Monitor Status
VSM Status
Uptime: 157327.95
OSD Summary
Osdmap Epoch: 10920
Total OSDs: 96
OSDs up: 94
OSDs in: 94
Near Full: False
Full: False
OSD Status
MDS Summary
MDS Epoch: 44
Up: 1
In: 1
Max: 1
Failed: 0
Stopped: 0
MDS Status
PG Summary
PGmap Version: 240074
Total PGs: 21455
PGs active+clean: 21430
PGs not active+clean: 25
PG Status
Cluster Health Summary
Warning and Errors
Warn: mon.0 addr 192.168.102.42:6789/0 clock skew 0.139078s > max 0.05s (latency 0.235162s)
Warn: mon.2 addr 192.168.103.43:6789/0 clock skew 0.136439s > max 0.05s (latency 0.000875471s)

Thank You!