

# Ceph Outage Impact Analysis for StarOS VNF

## Contents

[Introduction](#)

[Prerequisite](#)

[Requirements](#)

[Components used](#)

[Abbreviations](#)

[Ceph in Cisco VIM](#)

[Basics of Monitoring Mechanism in Ceph](#)

[Impact of Blocking I/O on StarOS VNF](#)

[Long Blocking I/O Scenarios](#)

[Laggy Timer Mechanism](#)

[RAID Card Hardware Failure](#)

[How to Mitigate the Impact?](#)

[Move to Local Disk from Ceph Storage](#)

[Ceph Configuration Tuning](#)

[Monitor RAID Card Hardware Issue](#)

[CEPH OSD RESERVE PCORES Tuning](#)

## Introduction

This document describes how StarOS VNF, which runs on the Cisco Virtualized Infrastructure Manager (VIM) is impacted when Ceph storage service is impaired, and what can be done to mitigate the impact. It is explained on the assumption that Cisco VIM is used as an infrastructure but the same theory can be applied to any Openstack environment.

## Prerequisite

### Requirements

Cisco recommends that you have knowledge of these topics:

- Cisco StarOS
- Cisco VIM
- Openstack
- Ceph

### Components used

The information in this document is based on these software and hardware versions:

- StarOS: 21.16.c9
- Cisco VIM: 3.2.2 (Openstack Queens)

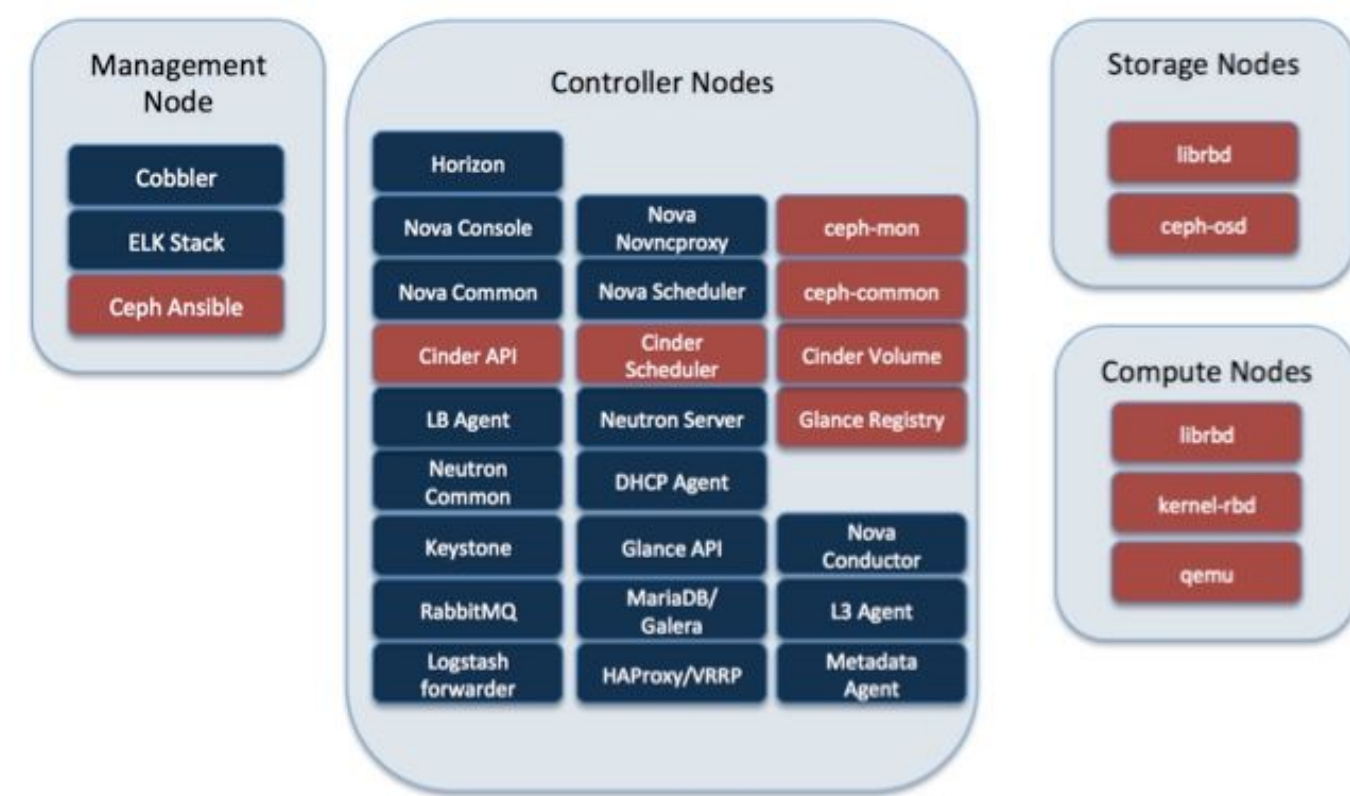
The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

## Abbreviations

Cisco VIM	Cisco Virtualized Infrastructure Manager
VNF	Virtual Network Function
Ceph OSD	Ceph Object Storage Daemon
StarOS	Operating System for Cisco Mobile Packet Core solution

## Ceph in Cisco VIM

This image here is taken from the Cisco VIM Administrator Guide. Cisco VIM uses Ceph as a storage back end.



Ceph supports both block and object storage and is therefore used to store VM images and volumes that can be attached to VMs. Multiple OpenStack services that depend on the storage backend include:

- Glance (OpenStack image service) — Uses Ceph to store images.
- Cinder (OpenStack storage service) — Uses Ceph to create volumes that can be attached to VMs.
- Nova (OpenStack compute service) — Uses Ceph to connect to the volumes created by Cinder.

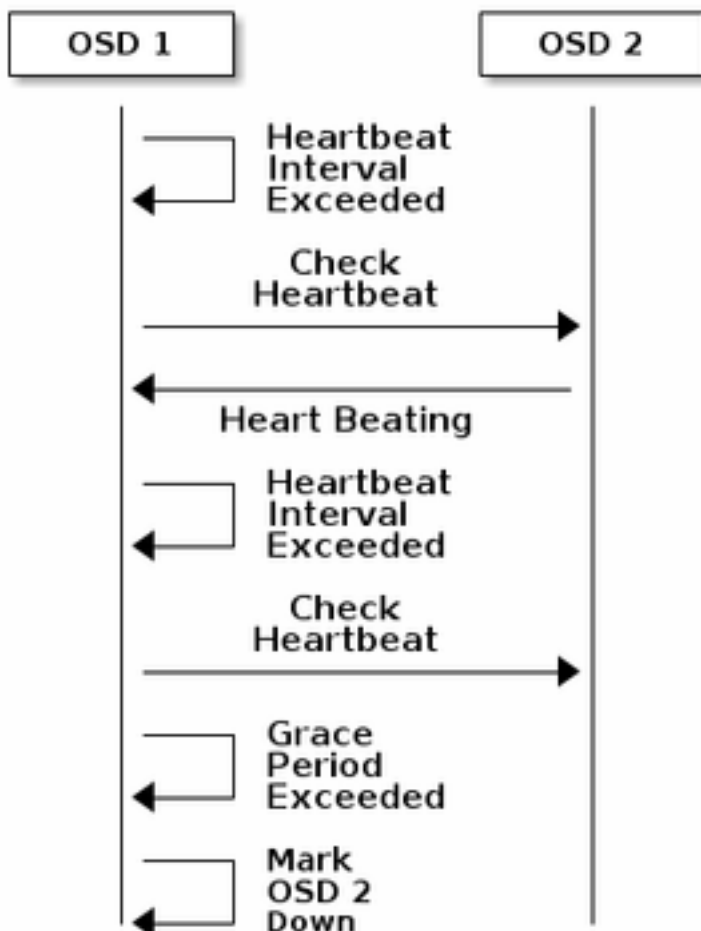
In many cases, a volume is created in Ceph for **/flash** and **/hd-raid** for StarOS VNF like the example here.

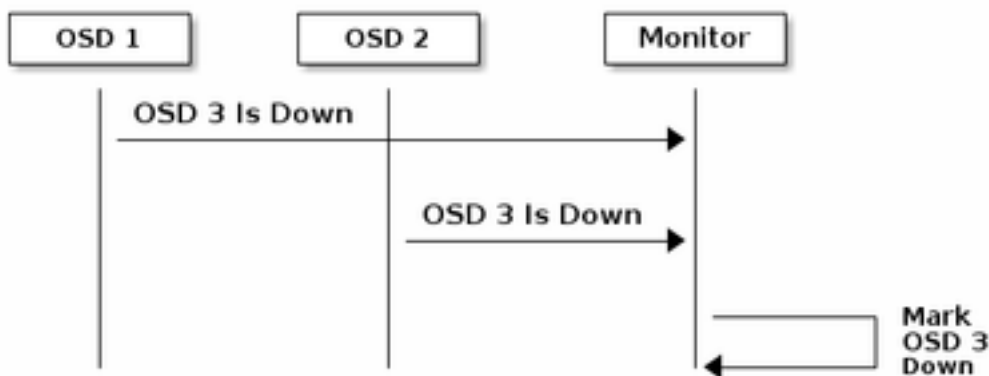
```
openstack volume create --image `glance image-list | grep up-image | awk '{print $2}` --size 16
--type LUKS upl-flash-boot
openstack volume create --size 20 --type LUKS upl-hd-raid
```

## Basics of Monitoring Mechanism in Ceph

Here is the explanation from Ceph document regarding monitoring:

Each Ceph OSD Daemon checks the heartbeat of other Ceph OSD Daemons at random intervals less than every 6 seconds. If a neighboring Ceph OSD Daemon doesn't show a heartbeat within a 20 second grace period, the Ceph OSD Daemon may consider the neighboring Ceph OSD Daemon down and report it back to a Ceph Monitor, which updates the Ceph Cluster Map. By default, two Ceph OSD Daemons from different hosts must report to the Ceph Monitors that another Ceph OSD Daemon is down before the Ceph Monitors acknowledge that the reported Ceph OSD Daemon is down.





So, in general, it takes around 20 seconds to detect OSD down and the Ceph cluster Map is updated, only after this VNF can use a new OSD. During this time disk, I/O is blocked.

## Impact of Blocking I/O on StarOS VNF

If disk I/O is blocked for more than 120 seconds, StarOS VNF reboots. There is a specific check for `xfssyncd/md0` and `xfs_db` processes which are related to disk I/O and StarOS intentionally reboots when it detects a stuck on these processes more than 120 seconds.

StarOS debug console log:

```
[ 1080.859817] INFO: task xfssyncd/md0:25787 blocked for more than 120 seconds.
[ 1080.862844] "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message.
[ 1080.866184] xfssyncd/md0 D ffff880c036a8290 0 25787 2 0x00000000
[ 1080.869321] ffff880aacf87d30 0000000000000046 0000000100000a9a ffff880a00000000
[ 1080.872665] ffff880aacf87fd8 ffff880c036a8000 ffff880aacf87fd8 ffff880aacf87fd8
[ 1080.876100] ffff880c036a8298 ffff880aacf87fd8 ffff880c0f2f3980 ffff880c036a8000
[ 1080.879443] Call Trace:
[ 1080.880526] [<ffffffff8123d62e>] ? xfs_trans_commit_iclog+0x28e/0x380
[ 1080.883288] [<ffffffff810297c9>] ? default_spin_lock_flags+0x9/0x10
[ 1080.886050] [<ffffffff8157fd7d>] ? _raw_spin_lock_irqsave+0x4d/0x60
[ 1080.888748] [<ffffffff812301b3>] _xfs_log_force_lsn+0x173/0x2f0
[ 1080.891375] [<ffffffff8104bae0>] ? default_wake_function+0x0/0x20
[ 1080.894010] [<ffffffff8123dc15>] _xfs_trans_commit+0x2a5/0x2b0
[ 1080.896588] [<ffffffff8121ff64>] xfs_fs_log_dummy+0x64/0x90
[ 1080.899079] [<ffffffff81253cf1>] xfs_sync_worker+0x81/0x90
[ 1080.901446] [<ffffffff81252871>] xfssyncd+0x141/0x1e0
[ 1080.903670] [<ffffffff81252730>] ? xfssyncd+0x0/0x1e0
[ 1080.905871] [<ffffffff81071d5c>] kthread+0x8c/0xa0
[ 1080.908815] [<ffffffff81003364>] kernel_thread_helper+0x4/0x10
[ 1080.911343] [<ffffffff81580805>] ? restore_args+0x0/0x30
[ 1080.913668] [<ffffffff81071cd0>] ? kthread+0x0/0xa0
[ 1080.915808] [<ffffffff81003360>] ? kernel_thread_helper+0x0/0x10
[ 1080.918411] *** xfssyncd/md0 stuck, resetting card
```

But it is not limited to the 120 seconds timer, if disk I/O is blocked for a while, even less than 120 secs, VNF may reboot for a variety of reasons. The output here is one example that shows a reboot due to the disk I/O issue, sometimes continuous StarOS task crash, and so on. It depends on the timing of active disk I/O versus storage issue.

```
[ 2153.370758] Hangcheck: hangcheck value past margin!
[ 2153.396850] ata1.01: exception Emask 0x0 SAct 0x0 SErr 0x0 action 0x6 frozen
```

```
[ 2153.396853] ata1.01: failed command: WRITE DMA EXT
--- skip ---
```

```
SYSLINUX 3.53 0x5d037742 EBIOS Copyright (C) 1994-2007 H. Peter Anvin
```

Basically a long blocking I/O can be considered as a critical issue for StarOS VNF and should be minimized as much as possible.

## Long Blocking I/O Scenarios

Based on the research from multiple customer deployments and lab testing, there are 2 main scenarios identified that can cause a long blocking I/O in Ceph.

### Laggy Timer Mechanism

There is a heartbeat mechanism between OSDs, to detect OSD down. Based on **osd\_heartbeat\_grace** value(20 secs by default), OSD is detected as failed. And there is a laggy timer mechanism, when there is a fluctuation or flap in OSD status the grace timer is automatically adjusted(become longer). This may make **osd\_heartbeat\_grace** value larger.

In the normal situation, heartbeat grace is 20 secs

```
2019-01-09 16:58:01.715155 mon.ceph-XXXXXX [INF] osd.2 failed (root=default,host=XXXXXX) (2
reporters from different host after 20.000047 >= grace 20.000000)
```

But after multiple network flaps of a storage node, it becomes a bigger value.

```
2019-01-10 16:44:15.140433 mon.ceph-XXXXXX [INF] osd.2 failed (root=default,host=XXXXXX) (2
reporters from different host after 256.588099 >= grace 255.682576)
```

So in the example above, it takes 256 seconds to detect OSD as down.

### RAID Card Hardware Failure

Ceph may not be able to detect RAID card hardware failure in a timely manner. RAID card failure ends up with a sort of OSD hang situation. In this case, OSD down is detected after a few minutes which is enough to make StarOS VNF reboots.

When RAID card hangs, some CPU cores take 100% on wa status.

```
%Cpu20 : 2.6 us, 7.9 sy, 0.0 ni, 0.0 id, 89.4 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu21 : 0.0 us, 0.3 sy, 0.0 ni, 99.7 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu22 : 31.3 us, 5.1 sy, 0.0 ni, 63.6 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu23 : 0.0 us, 0.0 sy, 0.0 ni, 28.1 id, 71.9 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu24 : 0.0 us, 0.0 sy, 0.0 ni, 0.0 id,100.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu25 : 0.0 us, 0.0 sy, 0.0 ni, 0.0 id,100.0 wa, 0.0 hi, 0.0 si, 0.0 st
```

And it eats up all CPU cores gradually and OSD is also down gradually with some time gap.

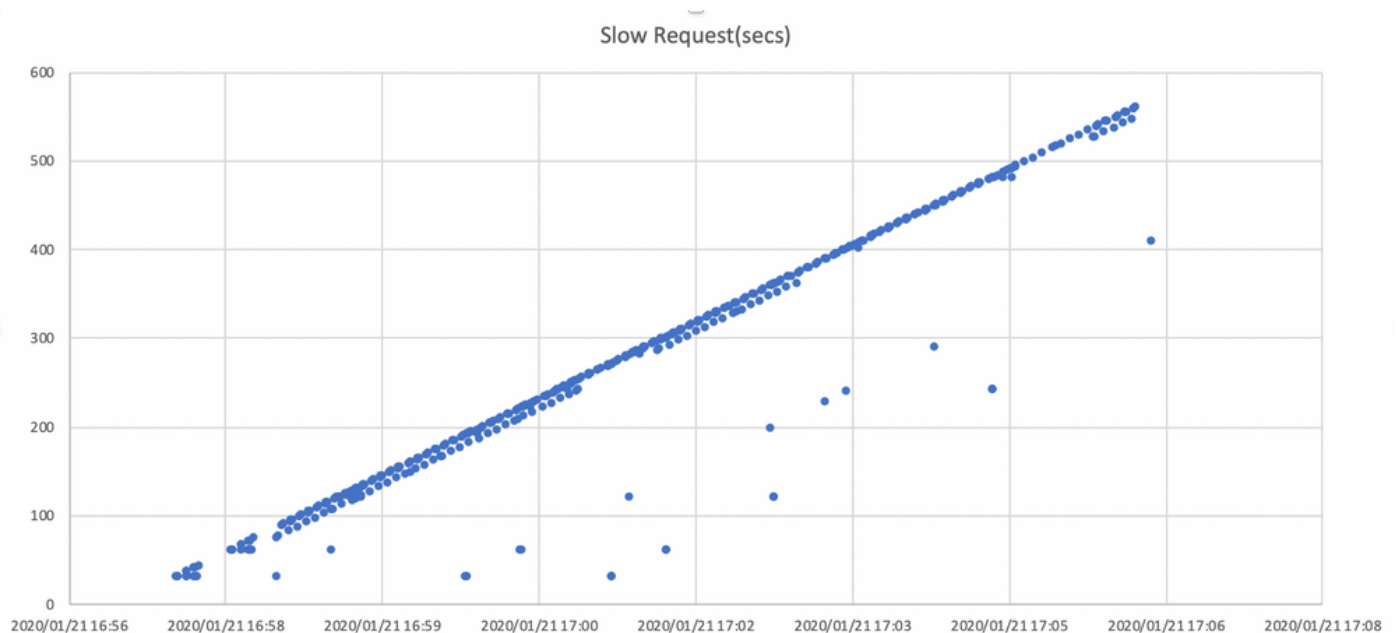
```
2019-01-01 17:08:05.267629 mon.ceph-XXXXXX [INF] Marking osd.2 out (has been down for 602
seconds)
```

```
2019-01-01 17:09:25.296955 mon.ceph-XXXXXX [INF] Marking osd.4 out (has been down for 603 seconds)
2019-01-01 17:11:10.351131 mon.ceph-XXXXXX [INF] Marking osd.7 out (has been down for 604 seconds)
2019-01-01 17:16:40.426927 mon.ceph-XXXXXX [INF] Marking osd.10 out (has been down for 603 seconds)
```

In parallel, slow requests are detected in **ceph.log**.

```
2019-01-01 16:57:26.743372 mon.XXXXXX [WRN] Health check failed: 1 slow requests are blocked > 32 sec. Implicated osds 2 (REQUEST_SLOW)
2019-01-01 16:57:35.129229 mon.XXXXXX [WRN] Health check update: 3 slow requests are blocked > 32 sec. Implicated osds 2,7,10 (REQUEST_SLOW)
2019-01-01 16:57:38.055976 osd.7 osd.7 [WRN] 1 slow requests, 1 included below; oldest blocked for > 30.216236 secs
2019-01-01 16:57:39.048591 osd.2 osd.2 [WRN] 1 slow requests, 1 included below; oldest blocked for > 30.635122 secs
-----skip-----
2019-01-01 17:06:22.124978 osd.7 osd.7 [WRN] 78 slow requests, 1 included below; oldest blocked for > 554.285311 secs
2019-01-01 17:06:25.114453 osd.4 osd.4 [WRN] 19 slow requests, 1 included below; oldest blocked for > 546.221508 secs
2019-01-01 17:06:26.125459 osd.7 osd.7 [WRN] 78 slow requests, 1 included below; oldest blocked for > 558.285789 secs
2019-01-01 17:06:27.125582 osd.7 osd.7 [WRN] 78 slow requests, 1 included below; oldest blocked for > 559.285915 secs
```

The graph here shows how long I/O requests are blocked with a timeline. The graph is created by plotting the slow request logs in ceph.log. It shows that the blocking time is getting longer over time.



## How to Mitigate the Impact?

### Move to Local Disk from Ceph Storage

The simplest way to mitigate the impact is to move to a local disk from Ceph storage. StarOS uses 2 disks, /flash and /hd-raid, it is possible to move only /flash to local disk which makes StarOS VNF more robust for the Ceph issues. The negative side of using shared storage such as Ceph is, all VNF using it are impacted at the same time when an issue happens. By using local disk the

impact of storage issue can be minimized to VNF running on the affected node only. And the scenarios mentioned in the previous section are applicable to Ceph only so not applicable to local disk. But the flip side of the local disk is, the content of the disk, such as StarOS image, configuration, core file, billing record, can not be retained when VM is redeployed. It may impact to VNF auto-healing mechanism too.

## Ceph Configuration Tuning

From StarOS VNF point of view, the following new Ceph parameters are recommended to minimize the blocking I/O time mentioned above.

<default settings>

```
"mon_osd_adjust_heartbeat_grace": "true",  
"osd_client_watch_timeout": "30",  
"osd_max_markdown_count": "5",  
"osd_heartbeat_grace": "20",
```

<new settings>

```
"mon_osd_adjust_heartbeat_grace": "false",  
"osd_client_watch_timeout": "10",  
"osd_max_markdown_count": "1",  
"osd_heartbeat_grace": "10",
```

It consists of:

- The laggy timer mechanism is disabled, no automatic adjustment
- The heartbeat grace time is shortened
- OSD is marked as down immediately (by default 5 times in last 600 seconds)

The new parameters are tested in a lab, the detection time for OSD down is reduced to approximately less than 10 seconds, it was originally around 30 seconds with the default configuration of Ceph.

## Monitor RAID Card Hardware Issue

For the RAID card hardware scenario, it may be still difficult to detect in a timely manner as the nature of the issue, as it creates a situation where OSD is intermittently working while I/O is blocked. There is no single solution for this, but it is recommended to monitor server hardware log for RAID card failure, or the slow request log in ceph.log by some script and take some action such as to make affected OSD down proactively.

## CEPH\_OSD\_RESERVE\_PCORES Tuning

This is not related to the scenarios mentioned, but if there is an issue with Ceph performance due to heavy I/O operation, increasing CEPH\_OSD\_RESERVE\_PCORES value can make Ceph I/O performance better. By default CEPH\_OSD\_RESERVE\_PCORES on Cisco VIM is configured as 2 and can be increased.