ACI Switch Node SSD Lifetime Explained

Contents

Introduction

This document describes the Solid State Drive (SSD) lifespan exhaustion behavior on Application Centric Infrastructure (ACI) switch platforms.

If you have landed on this page as a result of an SSD fault raised against Application Policy Infrastructure Controller (APIC) (not a switch node), refer to FN64329.

APIC SSD Replacement documents the APIC SSD replacement procedure.

Background Information

ACI Leaf and Spine nodes leverage SSDs for storage and event logging. All SSDs, irrespective of the vendor type, have a fixed lifespan based on specific attributes defined by the manufacturer. Some examples include the number of writes and erases as well as the volume of data transferred to the drives over a period of time. The SSD lifespan on Leaf and Spine nodes can become exhausted due to regular wear and tear. If the user does not take proper action against a switch that has an SSD approaching exhaustion, the switch could crash due to the SSD going into read-only mode. For this reason, it is imperative to prioritize SSD replacement in the event the F3073 fault is raised against a given node. SSD are field replaceable by Cisco certified Field Engineer.

F3073 was implemented in releases 2.1(4), 2.2(4), 2.3(1o), and 3.1(2m). Any subsequent software versions have the functionality to raise F3073.

Note: If you run an ACI software version that does not have SSD monitoring functionality, it is strongly advised to upgrade to a version where such monitoring is available.

Typical Symptoms if the SSD is Not Replaced

1. When the switch fails to boot, on the console you can see these errors:

```
/dev/hd-cfg0: ********** WARNING: Filesystem still has errors **********
e2fsck 1.42.1 (17-Feb-2012)
/dev/hd-cfg1: recovering journal
/sbin/e2fsck: unable to set superblock flags on /dev/hd-cfg1
```

2. The system is able to proactively raise faults related to the SSD (F3073 or F3074 based on the SSD lifespan usage) on later code releases (covered in a later section).

- <#root>

  F3074:

  ```
  fltEqptFlashFlash-minor-alarm (80% lifetime)
  ```

- <#root>

  F3073:

  ```
  fltEqptFlashFlash-worn-out     (90% lifetime)
  ```

3. You might also see fault F3525 raised for high SSD usage. This fault is often confused with F3073/F3074.

<#root>

```
# fault.Inst
code : F3525
ack : no
annotation :
cause : equipment-flash-warning
changeSet : deltape (New: 21), peCycles (New: 1678), tbw (New: 32.465179), warning (New: yes)
childAction :
created : 2019-08-05T18:22:01.455-07:00
delegated : no
descr :

High SSD usage observed. Please check switch activity and contact Cisco Technical Support about high SSD

dn : topology/pod-1/node-206/sys/ch/supslot-1/sup/flash/fault-F3525
domain : infra
extMngdBy : undefined
highestSeverity : warning
lastTransition : 2019-08-05T18:24:02.029-07:00
lc : raised
modTs : never
occur : 1
origSeverity : warning
prevSeverity : warning
rn : fault-F3525
rule : eqpt-flash-flash-warning-alarm
severity : warning
status :
subject : flash-warning-alarm
type : operational
```

Fault F3525 is raised if P/E (program-erase) cycles increment by more than 21 in seven days. It does not mean that the SSD is worn out, simply that there is a lot of churn which might eventually cause the SSDs to be worn out.

You need to work with the Technical Assistance Center (TAC) to understand what is causing this churn and address it. There is also a policy to change the thresholds that cause F3525 to be raised for a specific type of SSD (See **Access Policies > Policies > Switch > Equipment Flash Config Policies**).

How to Verify the Current Health of an SSD

There are two ways to check the SSD lifespan status:

- Proactive System Alerting
- Manual Data Analysis

Proactive System Alerting

SSD lifetime monitoring was added as part of Cisco bug ID [CSCve88634](#) in these versions of code:

- 2.1(4) and later for 2.1 train
- 2.2(4) and later for 2.2 train
- 2.3(1o) and later for 2.3 train
- 3.1(2m) and all future releases

A minor fault is raised when the drive approaches 80% of its lifetime and a major fault is raised when it exceeds 90% of its lifetime.

This allows network operators the capability to monitor and proactively replace any switch before it fails due to SSD lifetime exceeded.

- **F3074:** fltEqptFlashFlash-minor-alarm (80% lifetime)
- **F3073:** fltEqptFlashFlash-worn-out (90% lifetime)

Earlier versions of code did not have an SSD lifetime check for switch SSDs. Therefore, even if the switch was approaching the SSD expiration, there would not be a fault raised.

Manual Data Analysis

SSD usage logging was introduced starting in these releases: 3.2(5d), 3.2(6i), 3.2(7f), and 4.1(1i).

```
<#root>

a-leaf101#

cat /mnt/pss/ssd_log_amp.log

   (Tested on 4.2.3l)
Model                    SerialNo         Software-Version         Date
Micron_M600_MTFDDAT064MBF MSA195207V1                              Mon Dec 16 19:00:01 EST 2019
Model                    SerialNo         Software-Version         Date
Micron_M600_MTFDDAT064MBF MSA195207V1                              Tue Dec 17 19:00:01 EST 2019
Model                    SerialNo         Software-Version         Date
Micron_M600_MTFDDAT064MBF MSA195207V1                              Wed Dec 18 19:00:01 EST 2019
Model                    SerialNo         Software-Version         Date
```

```
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Thu Dec 19 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Fri Dec 20 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Sat Dec 21 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Sun Dec 22 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Mon Dec 23 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Tue Dec 24 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Wed Dec 25 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Thu Dec 26 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Fri Dec 27 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Sat Dec 28 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Sun Dec 29 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Mon Dec 30 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Tue Dec 31 19:00:01 EST 2019
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Wed Jan  1 19:00:01 EST 2020
Model                       SerialNo        Software-Version           Date
Micron_M600_MTFDDAT064MBF MSA195207V1                                   Thu Jan  2 19:00:01 EST 2020

Model                            SerialNo       SW-Version   Date         R.Error(1) GBB(5)    UpTim
Micron_M600_MTFDDAT064MBF        MSA195207V1    14.2(3l)     2020-01-03   16557977   0         29832
Micron_M600_MTFDDAT064MBF        MSA195207V1    14.2(3l)     2020-01-04   16558440   0         29856
Micron_M600_MTFDDAT064MBF        MSA195207V1    14.2(3l)     2020-01-05   16559031   0         29880
Micron_M600_MTFDDAT064MBF        MSA195207V1    14.2(3l)     2020-01-06   16561567   0         29904
Micron_M600_MTFDDAT064MBF        MSA195207V1    14.2(3l)     2020-01-07   16564701   0         29928
Micron_M600_MTFDDAT064MBF        MSA195207V1    14.2(3l)     2020-01-08   16565955   0         29952
```

ssd_log_amp.log file is in the 1of3 switch tech supports.

For any additional manufacturer-related commands outputs would need privileges of a root user which would require a TAC service request.

Triggers Leading to the Switch SSD Lifespan Exhaustion

- In an effort to ensure serviceability and the ability to provide root cause for events long after that event has occurred, ACI APICs and switches have verbose logging for each component that is persistent to SSD.

- Those logging capabilities require a high amount of disk writes and overwrite. The capability of the disks to perform this is not infinite, hence, there is a lifetime associated with the drive. Instability in the network (huge number of moves or policy updates, and so on), for a long-continued period can lead to premature exhaustion of a switch SSD.

- There have been many enhancements to optimize the writes and maintain serviceability while decreasing the amount of data written to disk. These changes were introduced across different versions. Refer to the Code Optimization section.

- For a version with the code optimization enhancements integrated, if you still experience fast SSD consumption like 2/3 P/E cycle each day, it is very likely caused by DB write instead of logging as most of them are written to RAMFS. Logs are only moved to the SSD after ZIP when rotation is required.

Code Optimization

There have been many enhancements to optimize the writes to maintain serviceability, while decreasing the amount of data written to disk. These changes were introduced across different versions.

- Cisco bug ID CSCve88634 - Raise a fault when the SSD exceeds its lifetime.
  Added to releases 2.1(4), 2.2(4), 2.3(1o), and 3.1(2m)
- Cisco bug ID CSCvi32353- F3073 Revision to SSD lifetime exceeded detection logic
- Cisco bug ID CSCvh73803 - create ramfs partition for tmp_logs to decrease logging impact on SSD
- Cisco bug ID CSCvm97108 - Bug to track SSD overprovisioning changes
- Cisco bug ID CSCvt36458 - Removed SQL DB persistent Layer for switches.
  Added to releases 4.2(6d) and 5.1(1h)

Releases 3.2.4d and later and 4.x and later have all the fault monitoring and optimization enhancements required to prolong the SSD lifespan. This of course does not mean that upgrading to these releases absolutely prevents the SSD lifespan issue.

Reload Type Behavior Change

Cisco bug ID CSCvt36458 changed the expected behavior of an ACI Switch node ungraceful reload:

Reload Type	Before Cisco bug ID CSCvt36458	After Cisco bug ID CSCvt36458 (4.2(6d) and later and 5.1(1h) and later)
Upgrades	Stateless	Stateless
Graceful Reloads (Manual reloads, hap-resets)	Stateful	Stateful
Ungraceful Reloads (Kernel panic, power cycle)	Stateful	**Stateless**

In summary, graceful switch reloads continue to use their DB on bringup. Ungraceful switch reloads now require that all policies are pulled from the APIC, similar to an upgrade.

Miscellaneous FAQ

Can the SSD be replaced in the field?

Yes. Cisco is now replacing SSDs in the field with the help of a certified Field Engineer. On modular Spines, the entire supervisor is replaced.

Should you proactively replace switch hardware?

TAC engineers can check smartctl values and evaluate which switches have more lifetime, so that replacements can be made for the most urgent switches immediately, and then gradually move to replace those switches that are not as urgent.

Are there specific hardware PIDs that are more susceptible to SSD failures?

No. The SSD lifespan usage can vary across different devices in the same environment. Usage of the SSD disks depends on how often the device needs to write the logs for event keeping. A device that registers a large number of events contributed by network instability has the potential of reaching its SSD lifespan sooner than others. Hence Cisco advocates the need to use later releases to ensure that code optimization helps with controlling the SSD writes and prolongs the SSD lifespan.