WHITE PAPER

Hot-Plug Guidance

Intel® Solid State Drive Data Center Family for PCIe®

Hot-Plug Capability of NVMe™ SSDs in Server Platforms

Intel® SSD Data Center Family for PCIe®

Expanding platform capabilities by enabling Hot-Plug support of Intel NVMe SSDs.

Overview

This White Paper is provided for OEM partners to support Intel® Solid State Drive (SSD) Data Center Family for PCIe® surprise hot-add/remove. It outlines a simple best known method (BKM) required for PCIe-based SSDs to work in the referenced platform, including operating systems (OS) and settings that Intel has validated. Some of the steps are workarounds, and are intended to be temporary until a permanent solution is in place.

"With the introduction of the U.2 form factor, OS and NVMe driver support, as well as Intel PCIe SSD support, Hot-Plug capability is now a widely spread standard feature for servers with NVMe SSD support."

– Zhdan Bybin
Sr. Application Engineer,
Intel Corporation

TERMS AND DEFINITIONS

Hot-Plug	A general term for adding or removing a device while the system is running.
Hot-Swap	A combined operation of a hot-removal followed by a hot-add of a different drive of the *same type/model.*
Surprise Hot-Add	Inserting an Enterprise NVMe SSD into a powered system while the OS is running, without notification; typically to add capacity or replace a failed drive. This is also known as **Hot-Insertion**.
Surprise Hot-Remove	Removal of an Enterprise NVMe SSD drive without any notification while the system is actively using the device.

Metrics of success:

- Surprise hot-add or hot-remove of an Intel SSD does not disrupt the functionality of the SSD; any acknowledged writes are guaranteed committed.
- Surprise hot-remove of an Intel SSD does not crash the system.
- Surprise hot-add of an Intel SSD is detected by the OS without requiring a reboot.

Current State of Capabilities

For surprise hot-remove/add of an Intel NVMe-based SSD to function, the following components of a system must be surprise hot-plug capable:

- CPU / Platform / OS
- NVM Express™ SSD & Driver

Intel SSD DC P3700/P3600/P3500 Series devices are *surprise hot-plug ready.*

All Intel SSD DC P3700/P3600/P3500 Series U.2 2.5" SFF SKUs are tested and validated by Intel to verify:

- Drive is electrically protected against power surge.
- Integrity of committed drive data remains intact.
- Acknowledged writes to drive will be committed in the event of an unplanned power loss.

Linux* NVMe driver support of surprise hot-plug starts with mainline kernel 3.10 with additional corner-case hot-plug fixes in 3.14.

Intel NVMe drivers for Windows* continue to support hot-plug as they have since the initial release. Current version of the NVMe

Zhdan Bybin
Intel Corporation

Nudrat Owens
Intel Corporation

Sneha Sreekantaiah
Intel Corporation

Annie Foong
Intel Corporation

driver is available on the Intel support page: https://downloadcenter.intel.com/download/23929/Intel-Solid-State-Drive-Data-Center-Family-for-PCIe-Drivers

At the time of this publication, the latest NVMe driver version is 1.3.0.1007.

Red Hat* support can be found at: http://access.redhat.com/products/red-hat-enterprise-linux

In the past two years, Red Hat has matured and stabilized the functionality of their driver. Red Hat Enterprise Linux 7.0 is fully featured, including hot-plug support and hot-boot capability.

Intel continuously validates platforms and operating systems. Please contact your Intel representative to determine any additional validated platforms not lisited in the table below.

BKM (Best Known Method)

1. Use Native OS hot-plug support

- In BIOS, set ACPI hot-plug = ***N***

  If ACPI hot-plug is enabled, hot-plug is handled by the BIOS, not the OS. This option is not covered by this BKM.

2. Ensure surprise hot-plug registers are set correctly

- Set slot capabilities register (offset 14h) ***bit 5*** (hot-plug surprise)

  When set, this bit indicates that an adapter present in this slot might be removed from the system without any prior notification. This bit is an indication to the operating system to allow for such removal without impacting continued software operation.

- Set slot control register (offset 18h) to ***bit 3*** set (Presence Detect Changed Enable).

  When set, this bit enables software notification on a presence detect changed event.

NOTE: Slot Status Register – 0x1A bit 6 (Presence Detect State) should present on all downstream ports that implement slots. This bit indicates the presence of an adapter in the slot.

- 0b = Slot empty
- 1b = Card present in slot

If a PCIe slot is surprise hot-plug capable, the BIOS will have configured it as so. On some platforms there is a BIOS setting in the **PCIe configuration** tab for each slot, usually named **hot-plug enable**. Otherwise this is a write-once setting, accessible only by the BIOS and set at the factory. If you believe that it is not configured correctly, use OS PCI tools to determine status[1]. (For example: Linux: lspci –vvv)

3. Workaround L1 limitation

Intel® Xeon® E3/E5 v3 Family of server CPUs currently do not trigger an interupt on surprise link down when link is in L1 state. (This limit will be removed in future CPUs.) The workaround is to turn off link power management, using one of the following options:

- Use PCI tool to set LinkControlRegister -> ActiveStatePowerManagement (ASPM) Control = ***0***
- In BIOS set PCIE ASPM to ***disable***

4. Workaround Linux PCIe Max Payload Size mismatch

On hot-add after boot, Linux does not re-tune the device's Maximum_Payload_Size to match the Downstream port (DSP) that the device is connected to. As a result, the device defaults to 0x0 (128B). This will result in a mismatch if the associated DSP, previously set by BIOS, is set to a MPS instead of 128B.

- Set kernel boot time parameter to ***pci=pcie_bus_perf*** as a workaround.

The MPS issues are intended to be addressed by Linux 4.3.

Please contact Intel Customer Support (ICS) or your Intel representative for more information or to obtain a copy of **Hot-plug Test Plan** for Intel PCIe SSDs.

Limitations and considerations

- The focus of this document is NVMe SSDs directly connected to CPU PCIe root ports.
- In this document, only 2.5" drives with U.2 connections (formerly SFF-8639) are applicable. Standard PCIe add-in cards may also be hot-plug capable, although it's not required by the industry or applicable in this case.
- This document applies only to Intel SSD Data Center Family for PCIe P3500/P3600/P3700 Series.

PLATFORMS AND OS DISTRIBUTIONS VALIDATED

OS	Kernel/Build	CPU/Platforms
Windows 2012*	Build 9200	**Romley** Supermicro* X9DRD-CNT+ Quanta* S210-X52S* (S2SP)
Windows 2012 R2*	Build 9600	
RHEL 7.0	3.10.0-123.el7.x86_64	
RHEL 7.0	3.10.0-131.el7.x86_64	**Grantley** Wildcat Pass Supermicro X10QBi Supermicro X10DDW Supermicro X10DRU-i+
RHEL 7.0	2.6.32-500.el6.x86_64	
RHEL 7.0	2.6.32-558.el6.x86_64	
SLES 12	3.12.28-2 rc 3	

[1] Linux tool – lspci/setpci; Windows tool - ReadWriteEverything or (Mindshare's) Arbor to access PCI register settings. Settings done this way will not be persistent across boot .

Printed in USA 160104/zb/ra ♻ Please Recycle 333596-001US