CXL ENGINEERING CHANGE NOTICE

TITLE:	PPR Enhancement
DATE:	First showing: 01/17/2024 Introduced:
AFFECTED DOCUMENT:	CXL 3.1
SPONSOR:	Daniele Balluchi – Micron Jordan Chin – Dell

Part I

1. **Summary of Functional Changes**

   CXL 3.x defines PPR maintenance operations which may be supported by CXL devices with DRAM components that have PPR capabilities. A PPR maintenance operation requests the CXL device to perform a repair operation on its media.

   DRAM components may support two types of PPR: Hard PPR (hPPR), for a permanent row repair, and Soft PPR (sPPR), for a temporary row repair. sPPR is much faster than hPPR, but the repair is lost with a power cycle.

   PPR operation may be initiated by host (issuing Perform Maintenance command) or by the device. Host may enable device initiated PPR setting **Device Initiated** bit to 1b in sPPR Feature or hPPR Features. hPPR can be performed only at boot time.

   Since **Device Initiated** bit is cleared upon a Cold Reset, it's not possible to perform device initiated PPR at boot.

   To permit device initiated PPR at boot time (after Warm or Cold Reset), this ECN defines a new configuration bit that is preserved across any type of reset: **Device Initiated at Device Boot**.

2. **Benefits as a Result of the Changes**

   The ECN enables device initiated hPPR. Since it is not possible to perform DRAM refresh during Hard PPR, the only time a device may perform hPPR is at boot.

3. **Assessment of the Impact**

   This ECN defines a capability bit to manage backward compatibility. Old software with new device should not check this capability bit, new software with old device will discover that device initiated at boot is not supported.

4. **Analysis of the Hardware Implications**

   Hardware implications should be limited because CXL devices already have the capability to store persistent information in non-volatile memory.

5. **Analysis of the Software Implications**
   Software should be modified to enable device initiated at boot.

6. **Analysis of the Compliance and Test Implications**

   - There is no compliance test for PPR in CXL 3.1.

Part II

Detailed Description of the change

- In sPPR Feature and hPPR Feature a new capability bit and a new configuration bit were added to enable PPR at device boot:
  - sPPR Flags and hPPR Flags:
    - Bit[3]: **Device Initiated at Device Boot Capability**
  - sPPR Operation Mode and hPPR Operation Mode:
    - Bit[1]: **Device Initiated at Device Boot**
- Since **Device Initiated at Device Boot** mode bit shall persist across power cycle, support for **Saved Selection Supported** = 1 was added.
- In CXL 3.x, Saved Selection Supported = 0, therefore **Device Initiated** bit is cleared to 0 by any reset event (Deepest Reset Persistence = 000b: None. Any reset will restore the default value). With this ECN, a device with **Saved Selection Supported** = 1 will restore the saved **Device Initiated** bit value at conventional reset.
- This ECN permits Saved Selection Supported = 1 and Deepest Reset Persistence = 010b (Conventional reset will restore the saved value), therefore
  - **Device Initiated** bit may be set just one time and preserved.
  - If **Device Initiated** = 0 and **Device Initiated at Device Boot Capability** = 1, the CXL Device may initiate PPR only at device boot and not at runtime.

8.2.9.7.1.1 PPR Maintenance Operations

Post Package Repair (PPR) maintenance operations may be supported by CXL devices that implement CXL.mem protocol. A PPR maintenance operation requests the CXL device to perform a repair operation on its media.
For example, a CXL device with DRAM components that support PPR features may implement PPR Maintenance operations. DRAM components may support two types of PPR: Hard PPR (hPPR), for a permanent row repair, and Soft PPR (sPPR), for a temporary row repair. sPPR is much faster than hPPR, but the repair is lost with a power cycle.
Based on DRAM PPR features, two maintenance operations are defined: sPPR and hPPR. Note that PPR maintenance operations may also apply to other types of media component.

During the execution of a PPR Maintenance operation, a CXL memory device:

- May or may not retain data
- May or may not be able to process CXL.mem requests correctly, including the ones that target the DPA involved in the repair If the device is not capable of correctly processing a CXL.mem request during a PPR

Maintenance operation, then:

- A read shall return poison
- A write shall be dropped, and an NDR shall be sent
- Any subsequent reads of DPA for which writes may have been incorrectly processed shall return poison

These CXL Memory Device capabilities are specified by Restriction Flags in the sPPR Feature and hPPR Feature (see Section 8.2.9.7.2.1 and Section 8.2.9.7.2.2, respectively).

sPPR maintenance operation may be executed at runtime, if data is retained and CXL.mem requests are correctly processed. For CXL devices with DRAM components, hPPR maintenance operation may be executed only at boot because data would not be retained.

When a CXL device identifies a failure on a memory component, the device may inform the host about the need for a PPR maintenance operation by using an Event Record, where the Maintenance Needed flag is set. The Event Record specifies the DPA that should be repaired. A CXL device may not keep track of the requests that have already been sent and the information on which DPA should be repaired may be lost upon power cycle.

The Host or the FM may or may not initiate a PPR Maintenance operation in response to a device request. It is possible to check whether resources are available by issuing a Perform Maintenance command for the PPR maintenance operation with the Query Resources flag set to 1. If the controller does not support reporting whether a resource is available, and a Perform Maintenance operation for PPR is issued with Query Resources set to 1, the controller shall return Invalid Input.

If resources are available, then the command shall be completed with the Success Return Code; otherwise, the command shall be completed with the Resources exhausted Return Code.

The host or the FM may initiate a repair operation by issuing a Perform Maintenance command, setting the Maintenance Operation Class to 01h (PPR), the Maintenance Operation Subclass to either 00h (sPPR) or 01h (hPPR), and indicating the DPA (if supported).

During the execution of a PPR maintenance operation, the device operation may be restricted as indicated by the Restriction Flags in the sPPR Feature and hPPR Feature (see Section 8.2.9.7.2.1 and Section 8.2.9.7.2.2, respectively).

Upon completion of a PPR maintenance operation, the device shall produce a Memory Sparing Event Record with updated resource availability, if the Memory Sparing Event Record Enable bit is set (see Table 8-113 or Table 8-116).

8.2.9.7.2 Features Associated with Maintenance Operations

Maintenance operations leverage the Features command set (see Section 8.2.9.6).

A Feature that provides capabilities and configurations may be defined for a maintenance operation. The list of maintenance operations that the device supports can be discovered by analyzing the device's supported Features. This can be accomplished by issuing the Get Supported Features command.

…
These Features represent maintenance operations capabilities and settings. Some fields of the Features are writable to configure the desired device behavior. Table 8-111 shows the Maintenance Operation Feature format. The first 16 bytes are common to all Maintenance Operation Features.

Table 8-111. Common Maintenance Operation Feature Format

Byte Offset	Length	Attribute	Description
00h	1	RO	**Maximum Maintenance operation latency**: This field indicates the maximum latency of the maintenance operation. • Bits[3:0] These bits specify the time scale. – 0h = 1 us, – 1h = 10 us – 2h = 100 us – 3h = 1 ms – 4h = 10 ms – 5h = 100 ms – 6h = 1 s – 7h = 10 s – All other encodings are reserved. • Bits[7:4] These bits specify the maximum operation latency with the time scale indicated in bits[3:0].
01h	2	RO	**Operation Capabilities** • Bit[0] **Device Initiated capability**: A value of 1 indicates that the device has the capability to initiate a maintenance operation without host involvement. Individual Feature (e.g. sPPR or hPPR) may provide additional description. • Bits[15:1]: Reserved.
03h	2	RW	**Operation Mode** • Bit[0] **Device Initiated**: value of 1 indicates that the device may initiate the maintenance operation without host involvement. ~~If cleared to 0, the device shall initiate the maintenance operation only when receiving an explicit maintenance command.~~ Individual Feature may provide additional description. If bit[0] of the Operation Capabilities field returns 0, this bit is considered reserved. • Bit[15:1] Reserved. Operation Mode default value is 0000h.
05h	1	RO	**Maintenance Operation Class**: This field specifies the Maintenance Operation Class. This value is set in Event Records to request this operation and in the Perform Maintenance input payload to initiate this maintenance operation.
06h	1	RO	**Maintenance Operation Subclass**: This field specifies the Maintenance Operation Subclass. This value is set in the Perform Maintenance input payload to initiate this maintenance operation.
07h	9	RsvdZ	Reserved
10h	Varies	-	Operation specific fields

8.2.9.7.2.1 sPPR Feature Discovery and Configuration Feature

The UUID of this feature is defined in Table 8-110.
Table 8-112 shows the information returned in the Get Supported Features output payload for the sPPR Feature. Some Feature attributes are changeable. ~~Feature attributes cannot be saved~~.

Table 8-112. Supported Feature Entry for the sPPR Feature.

Byte Offset	Length	Attribute	Value
0	10h	Feature Identifier	892ba475-fad8-474e-9d3e-692c917568bb
10h	2	Feature Index	Device specific
12h	2	Get Feature Size	14h
14h	2	Set Feature Size	03h
16h	4	Attribute Flags	• Bit[0]: Vendor specific value (Changeable). • Bits[3:1] ~~000b~~ 010b if saved selection is supported (Bit[6] = 1), otherwise 000b (Deepest Reset Persistence) ○ ~~None. Any reset will restore the default value.~~ • Bit[4]: ~~0~~ Vendor specific value (Persist Across Firmware Update). • Bit[5]: 1 (Default Selection Supported). • Bit[6]: ~~0~~ Vendor specific value. It shall be set to 1 if Device Initiated at Device Boot Capability is 1, see Table 8-113. (Saved Selection Supported). • Bit[31:7] Reserved
1Ah	1	Get Feature Version	0~~3~~2h
1Bh	1	Set Feature Version	0~~3~~2h
1Ch	2	Set Feature Effects	• Bit[0]: 0b (Configuration Change after Cold Reset). • Bit[1]: 1b (Immediate Configuration Change). • Bit[2]: 0b (Immediate Data Change). • Bit[3]: 0b (Immediate Policy Change). • Bit[4]: Vendor specific value (Immediate Log Change). • Bit[5]: 0b (Security State Change). • Bit[6]: 0b (Background Operation). • Bit[7]: Vendor specific value (Secondary Mailbox Supported). • Bit[8]: 0 (Request Abort Background Operation Supported) • Bit[9]: 1 is recommended, 0 is permitted (CEL[11:10] Valid) • Bit[10]: 0 (Configuration Change after Conventional Reset) • Bit[11]: 0 (Configuration Change after CXL Reset) • Bits[15:12]: 0h
1Dh	18	Reserved	-

Table 8-113 shows the output payload returned by a Get Feature command with Selection cleared to 0h (Current value) or set to 1h (Default value).

Table 8-113. sPPR Feature Readable Attributes

Byte Offset	Length	Description
00h	1	**Maximum Maintenance operation latency**: This field is defined in Table 8-111.
01h	2	**Operation Capabilities**: This field is defined in Table 8-111. If Device Initiated capability bit is set to 1, the device has the capability to initiate sPPR maintenance without host involvement at runtime. Device Initiated capability bit shall be cleared to 0 if Restriction Flags bit[0] or bit[2] is set to 1.

Byte Offset	Length	Description
03h	2	**Operation Mode**[1]: This field is defined in Table 8-111. If Device Initiated bit is set to 1, the device may initiate sPPR maintenance without host involvement at runtime.
05h	1	**Maintenance Operation Class**: It shall be set to 01h (PPR).
06h	1	**Maintenance Operation Subclass**: It shall be set to 00h (Soft PPR).
07h	9	Reserved
10h	1	~~**Soft s**~~**PPR Flags** Bit[0]: **DPA support Flag** If set, the device supports DPA argument in the Perform Maintenance command input payload. Bit[1]: **Nibble support Flag** If set, the device supports Nibble Mask argument in the Perform Maintenance command input payload. Bit[2]: **Memory Sparing Event Record Capability Flag**[2]: If set, the device has the capability to produce a Memory Sparing Event Record upon completion of sPPR maintenance operation. This bit shall be set if Get Feature Version is 02h or greater. Bit[3] **Device Initiated at Device Boot Capability**[3]**:** A value of 1 indicates that the device has the capability to initiate the sPPR maintenance operation without host involvement when Memory_Active = 0.This bit may be set to 1 even if Restriction Flags bit[0] or bit[2] is set to 1. This bit may be set to 1 even if the Device Initiated capability bit in the Operation Capabilities is cleared to 0. Bit[7:~~3~~4] Reserved.
11h	2	**Restriction Flags** Bit[0]: This bit indicates the ability of the device to process CXL.mem requests during Soft PPR maintenance operation execution. 0b: CXL.mem requests are properly processed. 1b: Media is not accessible. The device shall return poison on read access and drop write access. Bit[1]: Reserved. Bit[2]: This bit indicates the ability of the device to retain data across a Soft PPR maintenance operation execution. 0b: Data is retained. 1b: Data may or may not be retained. Bit[15:3] Reserved.
13h	1	**sPPR Operation Mode**[2] • Bit[0]: **Memory Sparing Event Record Enable**: If set, the Device shall produce a Memory Sparing Event Record upon completion of sPPR maintenance operation. This bit is reserved if ~~Soft s~~PPR Flags bit[2] is 0. • ~~Bit[1]:~~ **Device Initiated at Device Boot**[3,4]: value of 1 indicates that the device may initiate the sPPR maintenance operation without host involvement when Memory_Active = 0. This bit may be set to 1 even if the Device Initiated bit in the Operation Mode is cleared to 0. This bit is reserved if sPPR Flags bit[3] is 0. • Bits[7:~~1~~2]: Reserved. sPPR Operation Mode default value is 00h.

1. If Saved Selection Supported is 0, any reset will clear Device Initiated bit to 0 (default value). If Saved Selection Supported is 1, Deepest Reset Persistence will be 010b therefore the saved value is restored upon Conventional Reset.
2. Introduced as part of Get Feature Version = 02h.
3. Introduced as part of Get Feature Version = 03h.
~~2~~4. If Device Initiated and Device Initiated at Device Boot are cleared to 0, the device shall initiate sPPR maintenance operation only when receiving an explicit maintenance command.

Table 8-114 shows the Feature Data for the Set Feature command.

Table 8-114. sPPR Feature Writable Attributes

Byte Offset	Length	Description
00h	2	**Operation Mode** • Bit[0]: **Device Initiated**: A value of 1 indicates that the device may initiate the sPPR maintenance operation without host involvement at runtime. ~~If cleared to 0, the device shall initiate the maintenance operation only when receiving an explicit maintenance command. This bit is cleared upon a Cold Reset.~~ • Bits[15:1]: Reserved.
02h	1	**sPPR Operation Mode**[1] • Bit[0]: **Memory Sparing Event Record Enable**: If set, the device shall produce a Memory Sparing Event Record upon completion of sPPR maintenance operation. This bit is reserved if ~~Soft PPR~~sPPR Flags bit[2] is 0. • Bit[1] **Device Initiated at Device Boot**[2]: value of 1 indicates that the device may initiate the sPPR maintenance operation without host involvement when Memory Active = 0. This bit is reserved if sPPR Flags bit[3] is 0. • Bits[7:~~1~~2]: Reserved.

1. Introduced as part of Get Feature Version = 02h. This field is not included the Minimum Feature Data Size.
2. Introduced as part of Get Feature Version = 03h.

8.2.9.7.2.2 hPPR Feature Discovery and Configuration

The UUID of this feature is defined in Table 8-110.
Table 8-115 shows the information returned in the Get Supported Features output payload for the hPPR Feature. Some Feature attributes are changeable. ~~Feature attributes cannot be saved.~~

Table 8-115. Supported Feature Entry for the hPPR Feature

Byte Offset	Length	Attribute	Value
00h	10h	Feature Identifier	80ea4521-786f-4127-afb1-ec7459fb0e24
10h	2	Feature Index	Device specific
12h	2	Get Feature Size	14h
14h	2	Set Feature Size	03h
16h	4	Attribute Flags	• Bit[0]: Vendor specific value (Changeable). • Bit[3:1]: ~~000b~~ 010b if saved selection is supported (Bit[6] = 1), otherwise 000b (Deepest Reset Persistence) ~~○ None. Any reset will restore the default value.~~ • Bit[4]: ~~0~~ Vendor specific value (Persist Across Firmware Update). • Bit[5]: 1 (Default Selection Supported). • Bit[6]: ~~0~~ Vendor specific value. It shall be set to 1 if Device Initiated at Device Boot Capability is 1, see Table 8-116. (Saved Selection Supported). • Bit[31:7] Reserved
1Ah	1	Get Feature Version	0~~3~~2h
1Bh	1	Set Feature Version	0~~3~~2h
1Ch	2	Set Feature Effects	• Bit[0]: 0b (Configuration Change after Cold Reset). • Bit[1]: 1b (Immediate Configuration Change). • Bit[2]: 0b (Immediate Data Change). • Bit[3]: 0b (Immediate Policy Change).

			• Bit[4]: Vendor specific value (Immediate Log Change). • Bit[5]: 0b (Security State Change). • Bit[6]: 0b (Background Operation). • Bit[7]: Vendor specific value (Secondary Mailbox Supported). • Bit[8]: 0 (Request Abort Background Operation Supported) • Bit[9]: 1 is recommended, 0 is permitted (CEL[11:10] Valid) • Bit[10]: 0 (Configuration Change after Conventional Reset) • Bit[11]: 0 (Configuration Change after CXL Reset) • Bits[15:12]: 0h
1Eh	2	Reserved	All 0s

Table 8-116 shows the output payload returned by a Get Feature command with Selection set to 0h (Current value) or 1h (Default value).

Table 8-116. hPPR Feature Readable Attributes

Byte Offset	Length	Description
00h	1	**Maximum Maintenance Operation Latency**: This field is defined in Table 8-111.
01h	2	**Operation Capabilities**: This field is defined in Table 8-111. If Device Initiated capability bit is set to 1, the device has the capability to initiate hPPR maintenance without host involvement at runtime. Device Initiated capability bit shall be cleared to 0 if Restriction Flags bit[0] or bit[2] is set to 1.
03h	2	**Operation Mode**[1]: This field is defined in Table 8-111. If Device Initiated bit is set to 1, the device may initiate hPPR maintenance without host involvement at runtime.
05h	1	**Maintenance Operation Class**: It shall be set to 01h (PPR).
06h	1	**Maintenance Operation Subclass**: It shall be set to 01h (Hard PPR).
07h	9	**Reserved**
10h	1	**hPPR Flags** • Bit[0] **DPA support flag** If set, the device supports the DPA argument in the Perform Maintenance command input payload. • Bit[1] **Nibble support flag** If set, the device supports the Nibble Mask argument in the Perform Maintenance command input payload. • Bit[2]: **Memory Sparing Event Record Capability Flag**[2]: If set, the device has the capability to produce a Memory Sparing Event Record upon completion of hPPR maintenance operation. This bit shall be set if Get Feature Version is 02h or greater. • Bit[3] **Device Initiated at Device Boot capability**[3]: A value of 1 indicates that the device has the capability to initiate the hPPR maintenance operation without host involvement when Memory_Active = 0. This bit may be set to 1 even if Restriction Flags bit[0] or bit[2] is set to 1. This bit may be set to 1 even if the Device Initiated capability bit in the Operation Capabilities is cleared to 0. • Bit[7:3~~4~~] Reserved.

11h	2	**Restriction Flags** Bit[0]: This bit indicates the ability of the device to process CXL.mem requests during Hard PPR maintenance operation execution. 0b: CXL.mem requests are properly processed. 1b: Media is not accessible. The device shall return poison on read access and drop write access. Bit[1]: Reserved. Bit[2]: This bit indicates the ability of the device to retain data across a Hard PPR maintenance operation execution. 0b: Data is retained. 1b: Data may or may not be retained. Bit[15:3] Reserved.
13h	1	**hPPR Operation Mode**~~[1]~~[2] • Bit[0]: **Memory Sparing Event Record Enable**: If set, the device shall produce a Memory Sparing Event Record upon completion of hPPR maintenance operation. This bit is reserved if hPPR Flags bit[2] is 0. • Bit[1] **Device Initiated at Device Boot**[3,4]: value of 1 indicates that the device may initiate the hPPR maintenance operation without host involvement when Memory_Active = 0. This bit may be set to 1 even if the Device Initiated bit in the Operation Mode is cleared to 0. This bit is reserved if hPPR Flags bit[3] is 0. • Bits[7:~~1~~2]: **Reserved**. hPPR Operation Mode default value is 00h.

1. If Saved Selection Supported is 0, any reset will clear Device Initiated bit to 0 (default value). If Saved Selection Supported is 1, Deepest Reset Persistence will be 010b therefore the saved value is restored upon Conventional Reset.
2. Introduced as part of Get Feature Version = 02h.
3. Introduced as part of Get Feature Version = 03h.
~~2~~4. If Device Initiated and Device Initiated at Device Boot are cleared to 0, the device shall initiate hPPR maintenance operation only when receiving an explicit maintenance command.

Table 8-117 shows the input payload for Set Feature command.

Table 8-117 hPPR Feature Writeable Attributes

Byte Offset	Length in Bytes	Description
00h	2	**Operation Mode** • Bit[0] **Device Initiated**: A value of 1 indicates that the device may initiate the hPPR maintenance operation without host involvement at runtime. ~~If clear to 0, the device shall initiate the maintenance operation only when receiving an explicit maintenance command. This bit is cleared upon a Cold Reset.~~ • Bit[15:1] Reserved.
02h	1	**hPPR Operation Mode**[1] • Bit[0]: **Memory Sparing Event Record Enable**: If set, the device shall produce a Memory Sparing Event Record upon completion of hPPR maintenance operation. This bit is reserved if hPPR Flags bit[2] is 0. • Bit[1] **Device Initiated at Device Boot**[2]: value of 1 indicates that the device may initiate the hPPR maintenance operation without host involvement when Memory_Active = 0. This bit is reserved if hPPR Flags bit[3] is 0. • Bits[7:~~1~~2]: Reserved.

1. Introduced as part of Get Feature Version = 02h. This field is not included the Minimum Feature Data Size.
2. Introduced as part of Get Feature Version = 03h.