PCI-SIG ENGINEERING CHANGE NOTICE

TITLE:	Multicast
DATE:	December 14, 2007; approved by PWG May 8, 2008
AFFECTED DOCUMENT:	PCI Express Base Specification version 2.0
SPONSORS:	Hewlett-Packard, Integrated Device Technology, PLX Technology, NextIO, Tundra Semiconductor

Part I

1. Summary of the Functional Changes

This optional normative ECN adds Multicast functionality to PCI Express by means of an Extended Capability structure for applicable Functions in Root Complexes, Switches, and components with Endpoints. The Capability structure defines how Multicast TLPs are identified and routed. It also provides means for checking and enforcing send permission with Function-level granularity. The ECN identifies Multicast errors and adds an MC Blocked TLP error to AER for reporting those errors.

2. Benefits as a Result of the Changes

Multicast allows a single Posted Request TLP sent from a source to be distributed to multiple recipients, resulting in a very high performance gain when applicable.

3. Assessment of the Impact

This is an optional capability that is applicable to all components that either route TLPs or that might participate in Multicast as either as a source or a sink of a Multicast TLP. With the exception of multi-headed graphics, the usages for Multicast are primarily in the embedded spaces. Furthermore, the methodology accommodates components with Endpoints that don't implement the Multicast Capability structure as sources and sinks of Multicast TLPs. Thus, the ecosystem impact will be limited to Switches and to Root Complexes and Endpoints specifically designed to participate in Multicast.

4. Analysis of the Hardware Implications

This is an optional capability that is useful for Functions whose applications benefit from Multicast. Components that this does not apply to or that do not wish to add this capability are not required to. There are no hardware implications for a component that does not implement this capability.

5. Analysis of the Software Implications

There is no impact to current software. The registers associated with this capability default to a benign state so that current resource allocations algorithms continue to operate as they currently do.

The Multicast Capability structure, when enabled, creates a Multicast address range that exists outside the base and limit registers of the Switches that interconnect the virtual hierarchy of which it is a part. Resource allocation software that supports Multicast must recognize the Multicast Capability structure and take into account the existence of the Multicast address range.

Part II

Detailed Description of the change

Add to Terms and Acronyms:

Multicast, MC	A feature and associated mechanisms that enables a single Posted Request TLP sent by a source to be distributed to multiple targets.
Multicast Group, MCG	A set of Endpoints that are the target of Multicast TLPs in a particular address range.
Multicast Hit	The determination by a Receiver that a TLP will be handled as a Multicast TLP.
Multicast TLP	A TLP that is potentially distributed to multiple targets, as controlled by Multicast Capability structures in the components through which the TLP travels.
Multicast Window	A region of Memory Space where Posted Request TLPs that target it will be handled as Multicast TLPs.

Change 1.3.3. Switch

 A Switch forwards transactions using PCI Bridge mechanisms; e.g., address based routing except when engaged in a Multicast, as defined in section 6.xx. Multicast Operations

Change 2.2.3 TLP Digest Rules

 If an intermediate or ultimate PCI Express Receiver of the TLP does not support ECRC checking, the Receiver must ignore the TLP Digest[1]

Change 2.7 Data Integrity

Such a Switch can optionally check ECRC on TLPs that it forwards. On TLPs that the Switch forwards, the Switch must preserve the ECRC (forward it untouched) as an integral part of the TLP, regardless of whether the Switch checks the ECRC or if the ECRC check fails[2].

Change 2.7.1 ECRC Rules

 Switches must pass TLPs with ECRC unchanged from the Ingress Port to the Egress Port[2]

Change 2.7.1 ECRC Rules

Note that a Switch may optionally perform ECRC checking on TLPs passing through the Switch. ECRC Errors detected by the Switch are reported as described in Table 6-4, but do not alter the TLPs' passage through the Switch.[2]

---

[1] An exception is an Intermediate Receiver forwarding a Multicast TLP out an Egress Port with MC_Overlay enabled. See Section 6.xx.5.

[2] An exception is a Multicast TLP that an Egress Port is modifying due to the MC_Overlay mechanism. See Section 6.xx.5.

Change 6.2.3.2.3 Error Pollution. Note: this section already incorporates Base Specification errata that adds ACS Violation to the precedence list, as well as the AtomicOps ECN, which adds AtomicOp Egress Blocked to the list.

For errors detected in the Transaction layer, it is permitted and recommended that no more than one error be reported for a single received TLP, and that the following precedence (from highest to lowest) be used:

- ❑ Receiver Overflow
- ❑ Flow Control Protocol Error
- ❑ ECRC Check Failed
- ❑ Malformed TLP
- ❑ AtomicOp Egress Blocked
- ❑ ACS Violation
- ❑ MC Blocked TLP
- ❑ Unsupported Request (UR), Completer Abort (CA), or Unexpected Completion[3]
- ❑ Poisoned TLP Received

Change 6.12. Access Control Services, last sentence

ACS hardware functionality is disabled by default, and is enabled only by ACS-aware software. With the exception of ACS Source Validation, ACS access controls are not applicable to Multicast TLPs (see Section 6.xx), and have no effect on them.

Change Table 6-4 Transaction Layer Error List (add entry for MC Blocked TLP error)

Error Name	Error Type (Default Severity)	Detecting Agent Action	References
MC Blocked TLP	Uncorrectable (non-fatal)	*Receiver (if checking):* Send ERR_NONFATAL to Root Complex. Log the header of the Request TLP that encountered the error.	6.xx.4

[3] These are mutually exclusive errors, so their relative order does not matter.

Add new Section 6.xx

6.xx. Multicast Operations

The Multicast Capability structure defines a Multicast address range, the segmentation of that range into a number, N, of equal sized Multicast Windows, and the association of each Multicast Window with a Multicast Group, MCG. Each Function that supports Multicast within a component implements a Multicast Capability structure that provides routing directions and permission checking for each MCG for TLPs passing through or to the Function. The Multicast Group is a field of up to 6 bits in width embedded in the address beginning at the MC_Index_Position, as defined in section 7.xx.4.

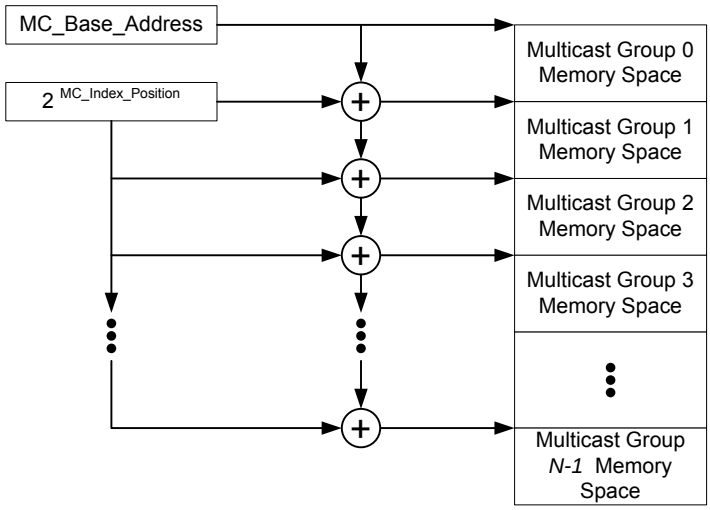

Figure 6.xx: Segmentation of the Multicast Address Range

6.xx.1. Multicast TLP Processing

A Multicast Hit occurs if all of the following are true:

- MC_Enable is Set
- TLP is a Memory Write or an Address Routed Message, both of which are Posted Requests
- $Address_{TLP}$ >= MC_Base_Address
- $Address_{TLP}$ < (MC_Base_Address+ ($2^{MC\_Index\_Position}$ * (MC_Num_Group + 1)))

In this step, each Switch Ingress Port and other components use values of MC_Enable, MC_Base_Address, MC_Index_Position, and MC_Num_Group from any one of their Functions. Software is required to configure all of these fields identically and results are indeterminate if this is not the case.

If the address in a Non-Posted Memory Request hits in a Multicast Window, no Multicast Hit occurs and the TLP is processed normal per the base specification – i.e. as a unicast.

If a Multicast Hit occurs, the only ACS access control that can still apply is ACS Source Validation. In particular, neither ACS redirection nor the ACS Egress Control vector affects operations during a Multicast Hit.

If a Multicast Hit occurs, normal address routing rules do not apply. Instead, the TLP is processed as follows:

The Multicast Group is extracted from the address in the TLP using any Function's values for MC_Base_Address and MC_Index_Position. Specifically:

$$MCG = ((Address_{TLP} - MC\_Base\_Address) >> MC\_Index\_Position) \& 3Fh$$

In this process, the component may use any Function's values for MC_Base_Address and MC_Index_Position. Which Function's values are used is device-specific.

Components next check the MC_Block_All and the MC_Block_Untranslated bits corresponding to the extracted MCG. Switches and Root Ports check Multicast TLPs in their Ingress Ports using the MC_Block_All and MC_Block_Untranslated registers associated with the Ingress Port. Endpoint Functions check Multicast TLPs they are preparing to send, using their MC_Block_All and MC_Block_Untranslated registers. If the MC_Block_All bit corresponding to the extracted MCG is set, the TLP is handled as an MC Blocked TLP. If the MC_Block_Untranslated bit corresponding to the extracted MCG is set and the TLP contains an Untranslated Address, the TLP, is also handled as an MC Blocked TLP.

IMPLEMENTATION NOTE

Multicast Window Size

Each ultimate Receiver of a Multicast TLP may have a different Multicast Window size requirement. At one extreme, a Multicast Window may be required to cover a range of memory implemented within the device. At the other, it may only need to cover a particular offset at which a FIFO register is located. The MC_Window_Size_Requested field within the Multicast Capability register is used by an Endpoint to advertise the size of Multicast Window that it requires.

Unless available address space is limited, resource allocation software may be able to treat each request as a minimum and set the Multicast Window size via MC_Index_Position to accommodate the largest request. In some cases, a request for a larger window size can be satisfied by configuring a smaller window size and assigning the same membership to multiple contiguous MCGs.

IMPLEMENTATION NOTE

Multicast, ATS, and Redirection

The ACS P2P Request Redirection and ACS Direct Translated P2P mechanisms provide a means where P2P Requests with Untranslated Addresses can be redirected to the Root Complex (RC) for access control checking, whereas P2P Requests with Translated Addresses can be routed "directly" to their P2P targets for improved performance. No corresponding redirection mechanism exists for Multicast TLPs.

To achieve similar functionality, an RC might be configured to provide one or more target Memory Space ranges that are not in the Multicast address range, but the RC maps to "protected" Multicast Windows. Multicast TLP senders either with or without ATS capability then target these RC Memory Space ranges in order to access the protected Multicast Windows indirectly. When either type of sender targets these ranges with Memory Writes, each TLP that satisfies the access control checks will be reflected back down by the RC with a Translated Address targeting a protected Multicast Window. ATS-capable senders can request and cache Translated Addresses using the RC Memory Space range, and then later use those Translated Addresses for Memory Writes that target protected Multicast Windows directly and can be Multicast without a taking a trip through the RC.

For hardware enforcement that only Translated Addresses can be used to target the protected Multicast Windows directly, software Sets appropriate MCG bits in the MC_Block_Untranslated register in all applicable Functions throughout the platform. Each MCG whose bit is set will cause its associated Multicast Window to be protected from direct access using Untranslated Addresses.

If the TLP is not blocked in a Switch or Root Complex it is forwarded out all of the Ports, except its Ingress Port, whose MC_Receive bit corresponding to the extracted MCG is set. In an Endpoint, it is consumed by all Functions whose MC_Receive bit corresponding to the extracted MCG is set. If no Ports forward the TLP or no Functions consume it, the TLP is silently dropped. Note that to prevent loops it is prohibited for a Root Port or a Switch Port to forward a TLP back out its Ingress Port, even if so specified by the MC_Receive register associated with the Port.

A Multicast Hit suspends normal address routing, including default Upstream routing in Switches. When a Multicast Hit occurs, the TLP will be forwarded out only those Egress Ports whose MC_Receive bit associated with the MCG extracted from the address in the TLP is set. If the address in the TLP does not decode to any Downstream Port using normal address decode, the TLP will be copied to the Upstream Port only if so specified by the Upstream Port's MC_Receive Register.

6.xx.2. Multicast Ordering

No new ordering rules are defined for processing Multicast TLPs. All Multicast TLPs are Posted Requests and follow Posted Request ordering rules. Multicast TLPs are ordered per normal ordering rules relative to other TLPs in a component's ingress stream through the point of replication. Once copied into an egress stream, a Multicast TLP follows the same ordering as other Posted Requests in the stream.

6.xx.3. Multicast Capability Structure Field Updates

Some fields of the Multicast Capability structure may be changed at any time. Others can't be changed with predictable results unless the MC_Enable bit is Clear in every Function of the component. The latter group includes MC_Base_Address and MC_Index_Position.

Fields which software may change at any time include MC_Enable, MC_Num_Group, MC_Receive, MC_Block_All, and MC_Block_Untranslated. Updates to these fields must themselves be ordered. Consider, for example, TLPs A and B arriving in that order at the same Ingress Port and in the same TC. If A uses value X for one of these fields, then B must use the same value or a newer value.

6.xx.4. MC Blocked TLP Processing

When a TLP is blocked by the MC_Block_All or the MC_Block_Untranslated mechanisms, the TLP is dropped. The Function blocking the TLP serves as the Completer. The Completer must log and signal this MC Blocked TLP error as indicated in Figure 6-2. In addition, the Completer must set the Signaled Target Abort bit in either its Status register or Secondary Status register as appropriate. To assist in fault isolation and root cause analysis, it is highly recommended that AER be implemented in Functions with Multicast capability.

In Root Complexes and Switches, if the error occurs with a TLP received by an Ingress Port, the error is reported by that Ingress Port. If the error occurs in an Endpoint Function preparing to send the TLP, the error is reported by that Endpoint Function.

6.xx.5 MC_Overlay Mechanism

The MC Overlay mechanism is provided to allow a single BAR in an Endpoint that doesn't contain a Multicast Capability structure to be used for both Multicast and unicast TLP reception. Software can configure the MC_Overlay_BAR in a Downstream Port so that the Multicast address range, or a portion of it, is remapped (overlaid) onto the Memory Space range accepted by the Endpoint's BAR. At the Upstream Port of a Switch, the mechanism can be used to overlay a portion of the Multicast address range onto a Memory Space range associated with host memory.

When enabled, the overlay operation specifies that bits in the address in the Multicast TLP whose bit numbers are equal to or higher than the MC_Overlay_Size field be replaced by the corresponding bits in the MC_Overlay_BAR. In other words:

```
If (MC_Overlay_Size == 0)
Then Egress_TLP_Addr = Ingress_TLP_Addr;
Else Egress_TLP_Addr = {MC_Overlay_BAR[63:MC_Overlay_Size],
                                    Ingress_TLP_Addr[MC_Overlay_Size-1:0]};
```

If the TLP with modified address contains the optional ECRC, the unmodified ECRC will almost certainly indicate an error. The action to be taken if a TLP containing an ECRC is Multicast copied to an Egress Port that has MC_Overlay enabled depends upon whether or not optional support for ECRC regeneration is implemented. All of the contingent actions are outlined in Table 6-xx below. If MC_Overlay isn't enabled, the TLP is forwarded unmodified. If MC_Overlay is enabled and the TLP has no ECRC, the modified TLP, with its address replaced as specified in the previous paragraph is forwarded. If the TLP has an ECRC but ECRC regeneration isn't supported, then the

modified TLP is forwarded with its ECRC dropped and the TD bit in the header cleared to indicate no ECRC attached. If the TLP has an ECRC and ECRC regeneration is supported, then an ECRC check is performed before the TLP is forwarded. If the ECRC check passes, the TLP is forwarded with regenerated ECRC. If the ECRC check fails, the TLP is forwarded with inverted regenerated ECRC.

Table 6-xx ECRC Rules for MC_Overlay

MC_Overlay Enabled	TLP has ECRC	ECRC Regeneration Supported	Action if ECRC Check Passes	Action if ECRC Check Fails
No	x	x	Forward TLP unmodified	
Yes	No	x	Forward modified TLP	
Yes	Yes	No	Forward modified TLP with ECRC dropped and TD bit clear	
Yes	Yes	Yes	Forward modified TLP with regenerated ECRC	Forward modified TLP with inverted regenerated ECRC

IMPLEMENTATION NOTE

MC_Overlay and ECRC Regeneration

Switch and Root Complex Ports have the option to support ECRC regeneration. If ECRC regeneration is supported, then it is highly advised to do so robustly by minimizing the time between checking the ECRC of the original TLP and replacing it with an ECRC computed on the modified TLP. The TLP is unprotected during this time, leaving a data integrity hole if the pre-check and regeneration aren't accomplished in the same pipeline stage.

Stripping the ECRC from Multicast TLPs passing through a Port that has MC_Overlay enabled but doesn't support ECRC regeneration allows the receiving Endpoint to enable ECRC checking. In such a case, the Endpoint will enjoy the benefits of ECRC on non-Multicast TLPs without detecting ECRC on Multicast TLPs modified by the MC_Overlay mechanism.

When Multicast ECRC regeneration is supported, and an ECRC error is detected prior to TLP modification, then inverting the regenerated ECRC ensures that the ECRC error isn't masked by the regeneration process.

IMPLEMENTATION NOTE

Multicast to Endpoints That Don't Have Multicast Capability

An Endpoint Function that doesn't contain a Multicast Capability structure can't distinguish Multicast TLPs from unicast TLPs. It is possible for a system designer to take advantage of this fact to employ such Endpoints as Multicast targets. The primary requirement for doing so is that the base and limit registers of the virtual PCI to PCI Bridge in the Switch Port above the device be configured to overlap at least part of the Multicast address range or that the MC_Overlay mechanism be employed. Extending this reasoning, it is even possible that a single Multicast target Function could be located on the PCI/PCI-X side of a PCI Express to PCI/PCI-X Bridge.

If an Endpoint without a Multicast Capability structure is being used as a Multicast target and the MC_Overlay mechanism isn't used, then it may be necessary to read from the Endpoint's Memory Space using the same addresses used for Multicast TLPs. Therefore, Memory Reads that hit in a Multicast Window aren't necessarily errors. Memory Reads that hit in a Multicast Window and that don't also hit in the aperture of a Root Complex Integrated Endpoint or the Downstream Port of a Switch will be routed Upstream, per standard address routing rules, and be handled as a UR there.

IMPLEMENTATION NOTE

Multicast in a Root Complex

A Root Complex with multiple Root Ports that supports Multicast may implement as many Multicast Capability structures as its implementation requires. If it implements more than one, software should ensure that certain fields, as specified in Section 6.xx.3, are configured identically. To support Multicast to Root Complex Integrated Endpoints, the implementation needs to expose all TLPs identified as Multicast via the MC_Base_Address Register to all potential Multicast target Endpoints integrated within it. Each such Integrated Endpoint then uses the MC_Receive register in its Multicast Capability structure to determine if it should receive the TLP.

IMPLEMENTATION NOTE

Multicast and Multi-Function Devices

All Port Functions and Endpoint Functions that are potential Multicast targets need to implement a Multicast Capability structure so that each has its own MC_Receive vector. Within a single component, software should configure the MC_Enable, MC_Base_Address, MC_Index_Position, and MC_Num_Group fields of these Capability structures identically. That being the case, it is sufficient to implement address decoding logic on only one instance of the Multicast BAR in the component.

IMPLEMENTATION NOTE

Congestion Avoidance

The use of Multicast increases the output link utilization of Switches to a degree proportional to both the size of the Multicast groups used and the fraction of Multicast traffic to total traffic. This results in an increased risk of congestion and congestion spreading when Multicast is used.

To mitigate this risk, components that are intended to serve as Multicast targets should be designed to consume Multicast TLPs at wire speed. Components that are intended to serve as Multicast sources should consider adding a rate limiting mechanism.

In many applications, the application's Multicast data flow will have an inherent rate limit and can be accommodated without causing congestion. Others will require an explicit mechanism to limit the injection rate, selection of a Switch with buffers adequate to hold the requisite bursts of Multicast traffic without asserting flow control, or selection of Multicast target components capable of sinking the Multicast traffic at the required rate. It is the responsibility of the system designer to choose the appropriate mechanisms and components to serve the application.

IMPLEMENTATION NOTE

The Host as a Multicast Recipient

For general-purpose systems, it is anticipated that the Multicast address range will usually not be configured to overlap with Memory Space that's directly mapped to host memory. If host memory is to be included as a Multicast recipient, the Root Complex may need to have some sort of I/O Memory Management Unit (IOMMU) that is capable of remapping portions of Multicast Windows to host memory, perhaps with page-level granularity. Alternatively, the MC Overlay mechanism in the Upstream Port of a Switch can be used to overlay a portion of the Multicast address range onto host memory.

For embedded systems that lack an IOMMU, it may be feasible to configure Multicast Windows overlapping with Memory Space that's directly mapped to host memory, thus avoiding the need for an IOMMU. Specific details of this approach are beyond the scope of this specification.

Modify Figure 7-32 and Table 7-29

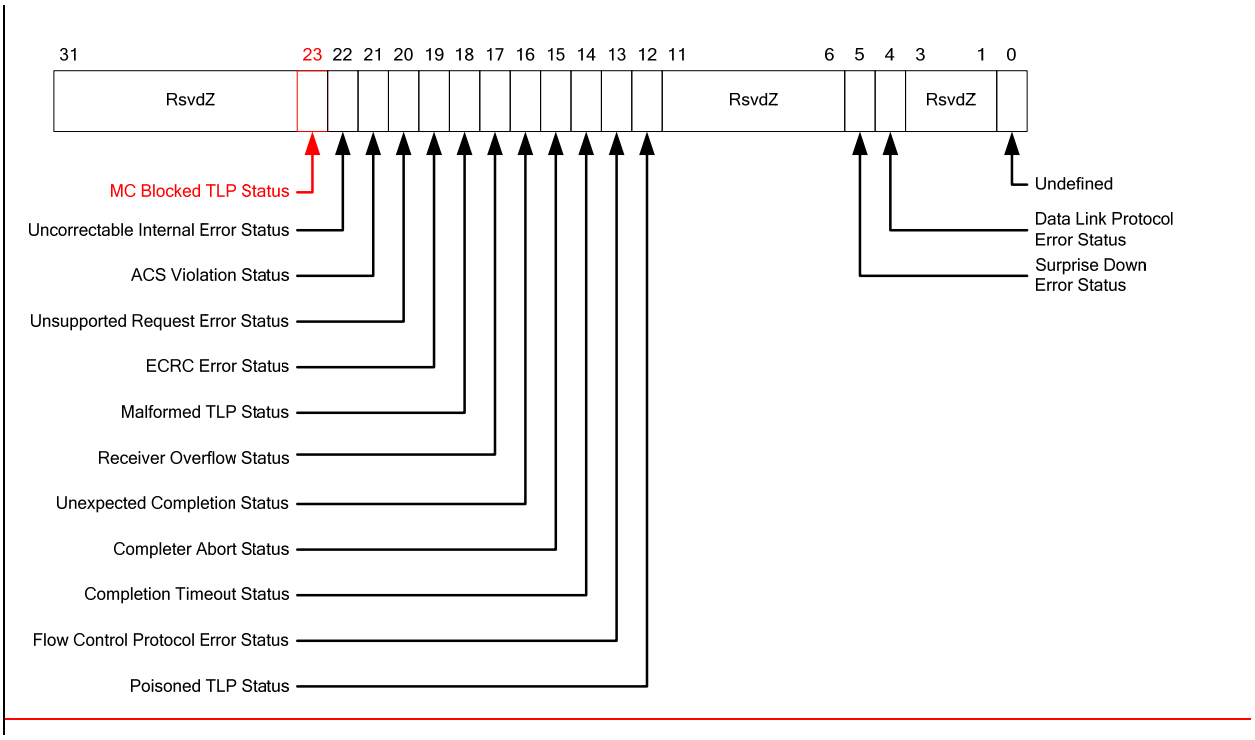

Figure 7-32 Uncorrectable Error Status Register

Table 7-29: Uncorrectable Error Status Register

Bit Location	Register Description	Attributes	Default
…	…	…	…
23	**MC Blocked TLP Status (Optional)**	RW1CS	0b

Modify Figure 7-33 and Table 7-30

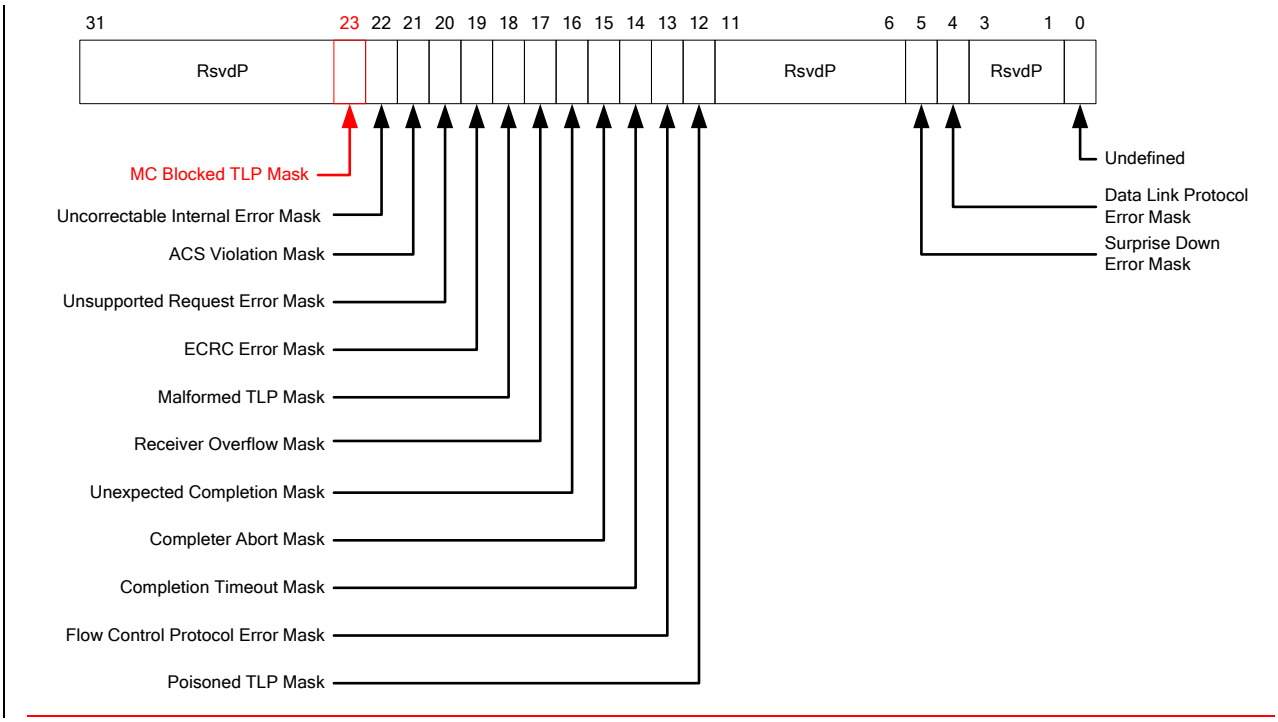

Figure 7-33: Uncorrectable Error Mask Register

Table 7-30: Uncorrectable Error Mask Register

Bit Location	Register Description	Attributes	Default
…	…	…	…
23	**MC Blocked TLP Mask (Optional)**	RWS	0b

Modify Figure 7-34 and Table 7-31

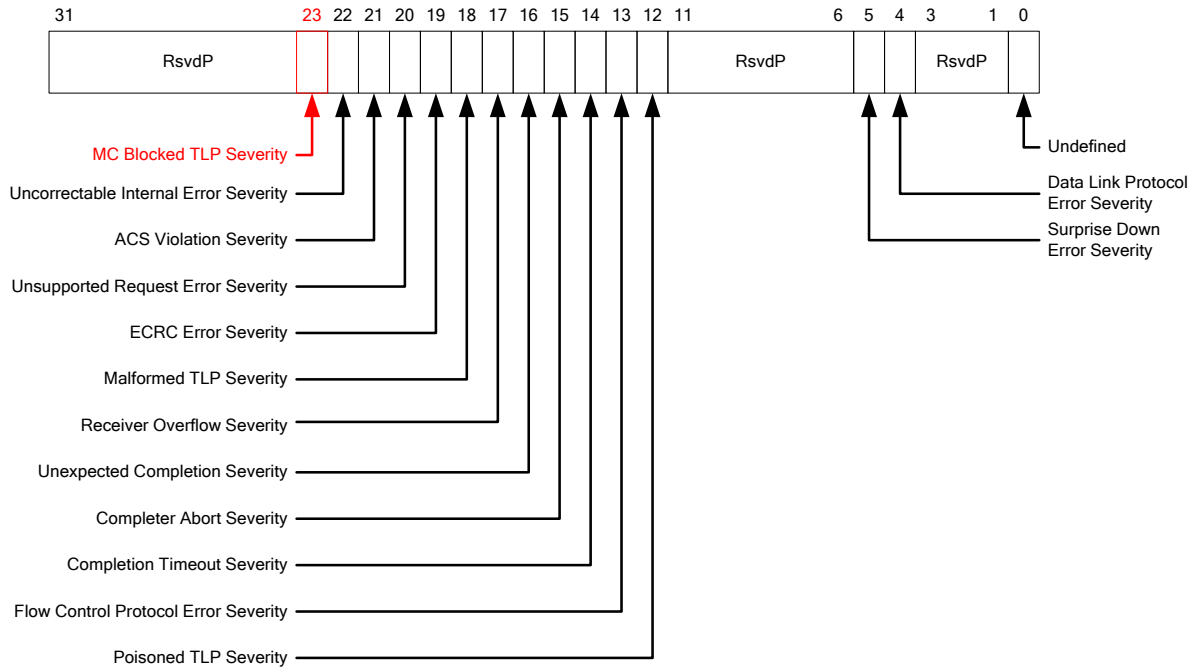

Figure 7-34: Uncorrectable Error Severity Register

Table 7-31: Uncorrectable Error Severity Register

Bit Location	Register Description	Attributes	Default
…	…	…	…
23	**MC Blocked TLP Severity (Optional)**	RWS	0b

Add new Section 7.xx

7.xx. Multicast Capability

Multicast functionality is controlled by the Multicast Capability structure. The Multicast capability is applicable to Root Ports, RCRBs, Switch Ports, Endpoint Functions, and Root Complex Integrated Endpoints. It is not applicable to PCI Express to PCI/PCI-X Bridges.

In the cases of a Switch or Root Complex or a component that contains multiple Functions, multiple copies of this Capability structure are required – one for each Endpoint Function, Switch Port, or Root Port that supports Multicast. To provide implementation efficiencies, certain fields within each of the Multicast Capability structures within a component must be programmed the same and results are indeterminate if this is not the case. The fields and registers that must be configured with the same values include MC_Enable, MC_Num_Group, MC_Base_Address and MC_Index_Position. These same fields in an Endpoint's Multicast Capability structure must match those configured into a Multicast Capability structure of the Switch or Root Complex above the Endpoint or in which the Root Complex Integrated Endpoint is integrated.

	31 0	Byte Offset
	PCI Express Extended Capability Header	00h
	Multicast Control Register / Multicast Capability Register	04h
	MC_Base_Address Register	08h
		0Ch
	MC_Receive Register	10h
		14h
	MC_Block_All Register	18h
		1Ch
	MC_Block_Untranslated Register	20h
		24h
Root Ports and Switch Ports	MC_Overlay_BAR	28h
		2Ch

Figure 7-xx Multicast Extended Capability Structure

7.xx.1. Multicast Extended Capability Header (Offset 00h)

Figure 7-xx details allocation of the fields in the Multicast Extended Capability Header and Table 7-xx provides the respective bit definitions.

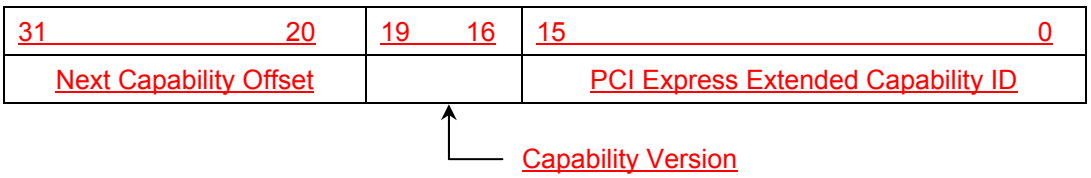

Figure 7-xx Multicast Extended Capability Header

Table 7-xx Multicast Extended Capability Header

Bit Location	Register Description	Attributes
15:0	**PCI Express Extended Capability ID** – This field is a PCI-SIG defined ID number that indicates the nature and format of the Extended Capability. PCI Express Extended Capability ID for the Multicast Capability is **0012h**	RO
19:16	**Capability Version** – This field is a PCI-SIG defined version number that indicates the version of the Capability structure present.	RO

	Must be 1h for this version of the specification.	
31:20	**Next Capability Offset** – This field contains the offset to the next PCI Express Extended Capability structure or 000h if no other items exist in the linked list of capabilities.	RO

7.xx.2. Multicast Capability Register (Offset 04h)

Figure 7-xx details allocation of the fields in the Multicast Capability Register and Table 7-xx provides the respective bit definitions.

15	14	13 8	7 6	5 0
	RsvdP	MC_Window_Size_Requested	RsvdP	MC_Max_Group

MC_ECRC_Regeneration_Supported

Figure 7-xx Multicast Capability Register

Table 7-xx Multicast Capability Register

Bit Location	Register Description	Attributes
5:0	MC_Max_Group –Value indicates the maximum number of Multicast Groups that the component supports, encoded as M-1. A value of 00h indicates that one Multicast Group is supported.	RO
13:8	MC_Window_Size_Requested – In Endpoints, the $\log_2$ of the Multicast Window size requested. RsvdP in Switch and Root Ports	RO
15	MC_ECRC_Regeneration_Supported – If set, indicates that ECRC regeneration is supported	RO

7.xx.3. Multicast Control Register (Offset 06h)

Figure 7-xx details allocation of the fields in the Multicast Control Register and Table 7-xx provides the respective bit definitions.

15	14 6	5 0
	RsvdP	MC_Num_Group

MC_Enable

Figure 7-xx Multicast Control Register

Table 7-xx Multicast Control Register

Bit Location	Register Description	Attributes
5:0	MC_Num_Group –Value indicates the number of Multicast Groups configured for use, encoded as N-1. The default value of 00h indicates that one Multicast Group is configured for use. Behavior is undefined if value exceeds MC_Max_Group. This parameter indirectly defines the upper limit of the Multicast address range. This field is ignored if MC_Enable is Clear. Default is 0.	RW
15	MC_Enable – When set, Multicast Capability is enabled for the component. Default is 0.	RW

7.xx.4. Multicast Base Address Register (Offset 08h)

The MC_Base_Address Register contains the MC_Base_Address and the MC_Index_Position. Figure 7-xx details allocation of the fields in the MC_Base_Address Register and Table 7-xx provides the respective bit definitions.

31 12	11 6	5 0
MC_Base_Address [31:12]	RsvdP	MC_Index_Position
MC_Base_Address [63:32]		

Figure 7-xx MC_Base_Address Register

Table 7-xx Multicast Base Address Register

Bit Location	Register Description	Attributes
5:0	MC_Index_Position – The location of the LSB of the Multicast Group number within the address. Behavior is undefined if this value is less than 12 and MC_Enable is Set. Default is 0.	RW
63:12	MC_Base_Address – The base address of the Multicast address range. The behavior is undefined if MC_Enable is Set and bits in this field corresponding to address bits that contain the Multicast Group number or address bits less than MC_Index_Position are non-zero. Default is 0.	RW

7.xx.5. MC_Receive Register (Offset 10h)

The MC_Receive register provides a bit vector denoting which Multicast groups the Function should accept, or in the case of Switch and Root Complex Ports, forward Multicast TLPs. This register is required in all Functions that implement the MC Capability structure.

Figure 7-xx details allocation of the fields in the MC_Receive Register and Table 7-xx provides the respective bit definitions.

31 0
MC_Receive [31:0]
MC_Receive [63:32]

Figure 7-xx MC_Receive Register

Table 7-xx MC_Receive Register

Bit Location	Register Description	Attributes
MC_Max_Group:0	MC_Receive– For each bit that's Set, this Function gets a copy of any Multicast TLPs for the associated Multicast Group. Bits above MC_Num_Group are ignored by hardware. Default is 0.	RW
All other bits	Reserved	RsvdP

7.xx.6. MC_Block_All Register (Offset 18h)

The MC_Block_All register provides a bit vector denoting which Multicast groups the Function should block. This register is required in all Functions that implement the MC Capability structure. Figure 7-xx details allocation of the fields in the MC_Block_All Register and Table 7-xx provides the respective bit definitions.

31 0

MC_Block_All [31:0]
MC_Block_All [63:32]

Figure 7-xx MC_Block_All Register

Table 7-xx MC_Block_All Register

Bit Location	Register Description	Attributes
MC_Max_Group:0	MC_Block_All– For each bit that is Set, this Function is blocked from sending TLPs to the associated Multicast Group. Bits above MC_Num_Group are ignored by hardware. Default is 0.	RW
All other bits	Reserved	RsvdP

7.xx.7. MC_Block_Untranslated Register (Offset 20h)

The MC_Block_Untranslated Register is used to determine whether or not a TLP that includes an Untranslated Address should be blocked. This register is required in all Functions that implement the MC Capability structure. However, an Endpoint Function that doesn't implement the ATS capability may implement this register as RsvdP.

Figure 7-xx details allocation of the fields in the MC_Block_Untranslated Register and Table 7-xx provides the respective bit definitions.

31 0
MC_Block_Untranslated [31:0]
MC_Block_Untranslated [63:32]

Figure 7-xx MC_Block_Untranslated Register

Table 7-xx MC_Block_Untranslated Register

Bit Location	Register Description	Attributes
MC_Max_Group:0	MC_Block_Untranslated – For each bit that is Set, this Function is blocked from sending TLPs containing Untranslated Addresses to the associated MCG. Bits above MC_Num_Group are ignored by hardware. Default is 0.	RW
All other bits	Reserved	RsvdP

7.xx.8. MC_Overlay_BAR (Offset 28h)

The MC_Overlay_BAR is required in Switch and Root Complex Ports that support the Multicast Capability and not implemented in Endpoints. Software must interpret the Device/Port Type Field in the PCI Express Capabilities Register to determine if the MC_Overlay_BAR is present in a Function.

The MC_Overlay_BAR specifies the base address of a window in unicast space onto which Multicast TLPs going out an Egress Port are overlaid by a process of address replacement. This allows a single BAR in an Endpoint attached to the Switch or Root Port to be used for both unicast and Multicast traffic. At a Switch Upstream Port, it allows the Multicast address range, or a portion of it, to be overlayed onto host memory.

Figure 7-xx details allocation of the fields in the MC_Overlay BAR and Table 7-xx provides the respective bit definitions.

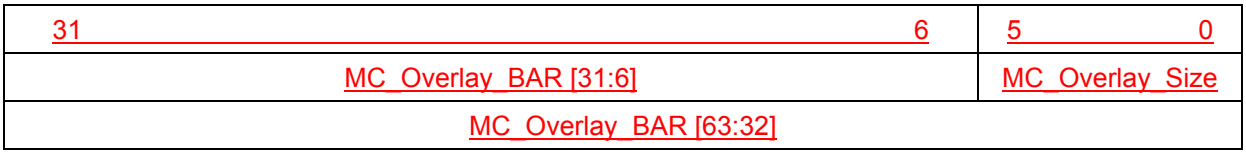

Figure 7-xx MC_Overlay_BAR

Table 7-xx MC Overlay BAR

Bit Location	Register Description	Attributes
5:0	MC_Overlay_Size – If 6 or greater, specifies the size in bytes of the overlay aperture as a power of 2. If less than 6, disables the overlay mechanism. Default is 0.	RW
63:6	MC_Overlay_BAR – Specifies the base address of the window onto which MC TLPs passing through this Function will be overlaid. Default is 0.	RW