

1. Background

Current ACPI HMAT (Heterogeneous Memory Attribute Table) mechanisms provide a matrix-based description of bandwidth and latency between memory initiators and targets, but do not adequately describe non-coherent peer-to-peer (P2P) communication capabilities between PCIe endpoint.

2. Problem Statement

Platform software (OS, drivers) requires:

1. Discovery of P2P connectivity between endpoints
2. Quantitative characterization (latency/bandwidth) of P2P paths e.g. P2P BW across sockets vs. within a socket
3. Differentiation of capability levels such as:
 - No P2P support
 - Conditional P2P (e.g., UIO-only ordering)
 - high-performance P2P

However, current ACPI structures do not provide:

- A standardized topology-level representation of P2P
 - A way to indicate partial or conditional support modes
 - A mechanism to encode non-existent paths vs non-optimal paths
-

3. Proposed Approach

Extend HMAT to Model P2P Between Host Bridges

Leverage the existing matrix abstraction of HMAT to describe P2P interactions:

- PCIe Host Bridges (HBs) can be represented via existing Generic Port construct in ACPI.
- Define HB-to-HB latency/bandwidth entries representing P2P data movement

Marked up ACPI ECN content

Add new structure Type to Table 5.178

Table 5.178 **HMAT Structure Types**

Value	Description
0	Memory Proximity Domain Attributes Structure
1	System Locality Latency and Bandwidth Information Structure
2	Memory Side Cache Information Structure
3	PCIe P2P Latency and Bandwidth Information Structure
4-0xFFFF	Reserved

Add a new table Table 5.18x **PCIe P2P Latency and Bandwidth Information Structure**

Field	Byte Length	Byte Offset	Description
Type	2	0	3 – PCIe P2P Latency and Bandwidth Information Structure
Reserved	2	2	Reserved
Length	4	4	Length in bytes for entire structure.
Flags	1	8	Bit [0] – Applies to non-UIO traffic Bit [1] - Applies to UIO traffic Bits [7:2] Reserved
Data Type	1	9	Type of data represented by this structure instance: - 0 - Access Latency (if read and write latencies are same)

Field	Byte Length	Byte Offset	Description
			- 1 - Read Latency - 2 - Write Latency - 3 - Access Bandwidth (if read and write bandwidth are same) - 4 - Read Bandwidth - 5 - Write Bandwidth Other values are reserved.
Reserved	2	10	Reserved
Number of Initiator Generic Port (s)	4	12	Indicates total number of Generic Ports that can initiate PCIe P2P requests to other Generic Ports. This is typically a PCIe Host Bridge.
Number of Target Generic Port (s)	4	16	Indicates total number of Generic Ports that can act as target. This is typically a PCIe Host Bridge.
Reserved	4	20	Reserved
Entry Base Unit	8	24	Base unit for Matrix Entry Values (latency or bandwidth). Base unit for latency in picoseconds. Base unit for bandwidth in megabytes per second (MB/s). This field shall be non-zero.
Initiator Proximity Domain List[0]	4	32	
Initiator Proximity Domain List[1]	4		
...			
Initiator Proximity Domain List[s-1]	4		
Target Proximity Domain List[0]	4	32 + 4 x s	
Target Proximity Domain List[1]	4		
...			
Target Proximity Domain List[t-1]	4		
Latency / bandwidth values			Total Number Entry shall be equal to Number of Initiator Proximity Domains * Number of Target Proximity Domains
Entry[0][0]	2	32 + 4 x s + 4 x t	Matrix entry (Initiator Proximity Domain List[0], Target Proximity Domain List[0])
Entry[0][1]	2		Matrix entry (Initiator Proximity Domain List[0], Target Proximity Domain List[1])
...			
Entry[0][Number of Target Proximity Domains -1]	2		Matrix entry (Initiator Proximity Domain List[0], Target Proximity Domain List[t-1])
Entry[1][0]	2		Matrix entry (Initiator Proximity Domain List[1], Target Proximity Domain List[0])

Field	Byte Length	Byte Offset	Description
Entry[1][1]	2		Matrix entry (Initiator Proximity Domain List[1], Target Proximity Domain List[1])
...			
Entry[1][Number of Target Proximity Domains -1]			Matrix entry (Initiator Proximity Domain List[1], Target Proximity Domain List[t-1])
...			
Entry[Number of Initiator Proximity Domains - 1][Number of Target Proximity Domains -1]	2		Matrix entry (Initiator Proximity Domain List[s-1], Target Proximity Domain List[t-1])

This structure provides a matrix that describes the normalized P2P read/write latency, the read/write bandwidth between Initiator Proximity Domains (I/O) and Target Proximity Domains (I/O). See section 5.2.29.4 for encoding of the Latency and Bandwidth fields.