PCI-SIG ENGINEERING CHANGE NOTICE

TITLE:	Hierarchy ID Message
DATE:	Introduced: Feb 23, 2016 Updated: Sept 8, 2016 14-day Cross WG Review: Sept 15, 2016 Member Review: October 20, 2016 Final Approval: February 22, 2017
AFFECTED DOCUMENT:	PCI Express Base Specification, Version 3.1a Single Root I/O Virtualization and Sharing Specification, Revision 1.1 PCI Code and ID Assignment Specification, Version 1.7
SPONSOR:	Steve Glaser, Nvidia

Part I

1. Summary of the Functional Changes

Defines a new, optional PCI-SIG Defined Type 1 Vendor Defined Message.

This message provides software and/or firmware, running on a Function, additional information to uniquely identify that Function, within a large system or a collection of systems.

When a single system contains multiple PCI Express Hierarchies, this message tells a Function which Hierarchy it resides in. This value, in conjunction with the Routing ID number uniquely identifies a Function within that system.

In clustered system, this message can include a System Globally Unique Identifier (System GUID) for each system. This value, in conjunction with the Hierarchy ID and Routing ID uniquely identifies a Function within that cluster.

2. Benefits as a Result of the Changes

Currently, a Function knows its Routing ID, but it doesn't need to know which System or which Hierarchy within that System it belongs to. With the advent of clustered systems, and the potential for consolidating Sideband signals into a distinct bus (with its own addressing structure), it is helpful, in some systems, to provide additional information to help software determine the topology.

3. Assessment of the Impact

No impact to existing systems. No impact to future systems that don't need this functionality.

4. Analysis of the Hardware Implications

New optional message. New extended capability structure in configuration space.

5. Analysis of the Software Implications

No required changes. Software for clustered systems can benefit by using this mechanism. Future software that uses the consolidated Sideband signaling mechanism currently under discussion could also benefit.

6. Analysis of the C&I Test Implications

New extended capability structure requires the creation of a new test for register field attributes.

Tests should be able to generate this message and see that supporting Endpoints record it. Optional Downstream Port tests should attempt to generate the message from a supporting Downstream Port and verify that the message was transmitted.

Part II

Detailed Description of the change

In the PCI Express Base Specification, add new Section 2.2.8.6.5 as follows:

2.2.8.6.5 Hierarchy ID Message

Hierarchy ID uses the PCI-SIG-Defined VDM mechanism (see Section 2.2.8.6.1). The Hierarchy ID Message is a PCI-SIG-Defined VDM (Vendor-Defined Type 1 Message) with payload (MsgD).

Beyond the rules for other PCI-SIG-Defined VDMs, the following rules apply to the formation of Hierarchy ID Messages:

- Table 2-xx and Figure 2-xx illustrate and define the Hierarchy ID Message.
- The TLP Type must be MsgD.
- Each Message must include a 4-DWORD data payload.
- The Length field must be 4.
- The TC[2:0] field must be 000b.
- The Attr[2:0] field is Reserved.
- The Tag field is Reserved.
- The Subtype field is 01h.
- The Message Routing field must be 011b – Broadcast from Root Complex.

Receivers may optionally check for violations of these rules (but must not check reserved bits). These checks are independently optional (see Section 6.2.3.4). If a Receiver implementing these checks determines that a TLP violates these rules, the TLP is a Malformed TLP.

1. If checked, this is a reported error associated with the Receiving Port (see Section 6.2).

The payload of each Hierarchy ID Message contains the lower 128-bits of the System GUID.

For details of the Hierarchy ID, GUID Authority ID, and System GUID fields see Section 6.x.

Table 2-xx: Hierarchy ID Message

Name	Code[7:0] (b)	Routing r[2:0] (b)	Support				Description/Comments
			RC	Ep	Sw	Br	
Hierarchy ID Message	0111 1111	011	t	r	tr		Hierarchy ID

The format of the Hierarchy ID Message is shown in Figure 2-xx below.

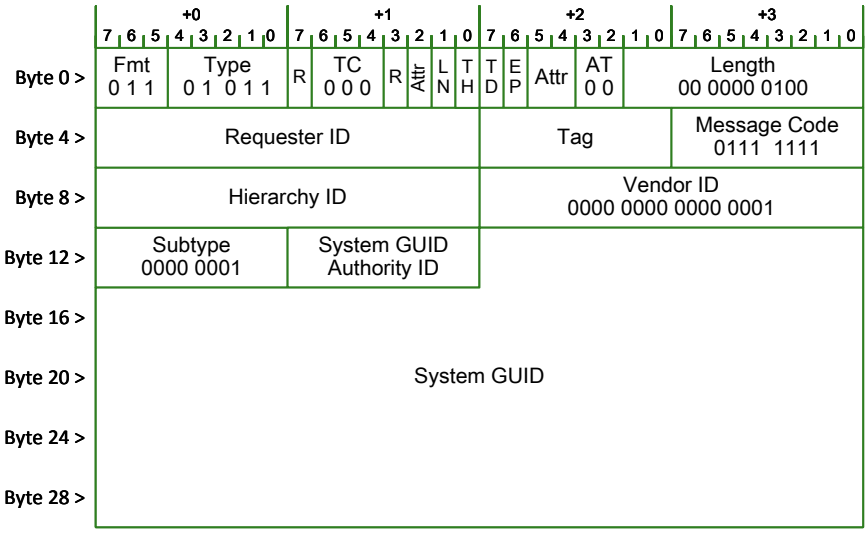

Figure 2-xx: Hierarchy ID Message

In the PCI Express Base Specification, add new Section 6.x at the end of Chapter 6 as follows:

6.x Hierarchy ID Message

When software initializes a PCI Hierarchy, it assigns unique Bus and Device numbers to each component so that every Function in the Hierarchy has a unique Routing ID within that Hierarchy. To ensure that Routing IDs are unique in large systems that contain more than one Hierarchy and in clustered systems that contain multiple Hierarchies, additional information is required to augment the Routing ID to produce a unique number. Functions can be uniquely identified by the combination of:

- Unique Identifier for the System (or Root Complex)
- Unique Identifier for the Hierarchy within that Root Complex
- Routing ID within that Hierarchy

The Hierarchy ID message (see Section 2.2.8.6.5) is used to provide the additional information needed for a Function to uniquely identify itself in a multi-hierarchy platform.

Hierarchy ID Messages are generated by a Downstream Port upon software request. Received messages at an Upstream Port are reported in the Hierarchy ID Extended Capability (see Section 7.x).

Hierarchy ID Messages are a PCI-SIG-Defined Type 1 VDM. Hierarchy ID Messages can safely be sent at any time and components that do not comprehend them will silently ignore them.

Hierarchy ID Messages typically are sent from a Downstream Port at the top of the Hierarchy (e.g., a Root Port). In systems where the Root Port does not support Hierarchy ID Messages, Hierarchy ID Messages can be sent from Switch Downstream Ports.

The Hierarchy ID Message is intended for use by software, firmware, and/or hardware. When using the Hierarchy ID Message, all bits of the Hierarchy ID, System GUID, System GUID Authority ID fields must be compared, without regard to any internal structure. How this information is used is outside the scope of this specification.

Layout of the Hierarchy ID Message is shown in Figure 2-xx. Fields in the Hierarchy ID Message are as follows:

Hierarchy ID contains the Segment Group Number associated with this Hierarchy (as defined by the *PCI Firmware Specification*). This field can be used in conjunction with the Routing ID to uniquely identify a Function within a System. The value 0000h indicates the default (or only) Hierarchy of the Root Complex. Non-zero values indicate additional Hierarchies.

System GUID[143:0], in conjunction with System GUID Authority ID, provides a globally unique identification for a System.

System GUID[143:136] is byte 14 in the Hierarchy ID Message.
System GUID[135:128] is byte 15 in the Hierarchy ID Message.
System GUID[127:120] is byte 16 in the Hierarchy ID Message.
System GUID[119:112] is byte 17 in the Hierarchy ID Message.
System GUID[111:104] is byte 18 in the Hierarchy ID Message.
System GUID[103: 96] is byte 19 in the Hierarchy ID Message.
System GUID[95: 88] is byte 20 in the Hierarchy ID Message.
System GUID[87: 80] is byte 21 in the Hierarchy ID Message.
System GUID[79: 72] is byte 22 in the Hierarchy ID Message.
System GUID[71: 64] is byte 23 in the Hierarchy ID Message.
System GUID[63: 56] is byte 24 in the Hierarchy ID Message.
System GUID[55: 48] is byte 25 in the Hierarchy ID Message.
System GUID[47: 40] is byte 26 in the Hierarchy ID Message.
System GUID[39: 32] is byte 27 in the Hierarchy ID Message.
System GUID[31: 24] is byte 28 in the Hierarchy ID Message.
System GUID[23: 16] is byte 29 in the Hierarchy ID Message.
System GUID[15: 8] is byte 30 in the Hierarchy ID Message.
System GUID[7: 0] is byte 31 in the Hierarchy ID Message.

System GUID Authority ID identifies the mechanism used to ensure that the System GUID is globally unique. The mechanism for choosing which Authority ID to use for a given system is implementation specific. The defined values are shown in Table 6-x.

Table 6-x: System GUID Authority ID Encoding

Authority ID	Description
00h	**None** – System GUID[143:0] is not meaningful. System GUID[143:0] must be 0.
01h	**Timestamp** – System GUID[63:0] contains a timestamp associated with the particular system. Encoding is a Unix 64 bit time (number of seconds since midnight UTC January 1, 1970).

	The mechanism of choosing the timestamp to represent a system is implementation specific. System GUID[143:64] must be 0.
02h	**IEEE EUI-48** – System GUID[47:0] contains a 48 bit Extended Unique Identifier (EUI-48) associated with the particular system. Encoding is defined by the IEEE. See EUI-48[1] for details. EUI-48 values are frequently used as network interface MAC addresses. The mechanism of choosing the EUI-48 value to represent a system is implementation specific. System GUID[143:48] must be 0.
03h	**IEEE EUI-64** – System GUID[63:0] contains a 64 bit Extended Unique Identifier (EUI-64) associated with the particular system. Encoding is defined by the IEEE. See EUI-64[2] for details. The mechanism of choosing the EUI-64 value to represent a system is implementation specific. System GUID[143:64] must be 0.
04h	**RFC-4122 UUID** – System GUID[127:0] contain a UUID as defined by the IETF in RFC-4122[3]. This definition is technically equivalent to ITU-T Rec. X.667[4] | ISO/IEC 9834-8. The mechanism of choosing the UUID value to represent a system is implementation specific. System GUID[143:128] must be 0
05h	**IPv6 Address** – System GUID[127:0] contains the unique IPv6 address of one of the network interfaces of the system. The mechanism of choosing the IPv6 value to represent a system is implementation specific. System GUID[143:128] must be 0.
06h to 7Fh	**Reserved** – System GUID[143:0] contains a unique value. The mechanism used to ensure uniqueness is outside the scope of this specification.
80h to FFh	**PCI-SIG Vendor Specific** – System GUID Authority ID values 80h to FFh are reserved for PCI-SIG vendor-specific usage. System GUID[143:128] contains a PCI-SIG assigned Vendor ID.

[1] https://standards.ieee.org/develop/regauth/tut/eui48.pdf

[2] https://standards.ieee.org/develop/regauth/tut/eui64.pdf

[3] https://tools.ietf.org/html/rfc4122

[4] http://www.itu.int/rec/T-REC-X.667-201210-I/en

	System GUID[127:0] contain a unique number assigned by that vendor. The mechanism used for assigning numbers is implementation specific. One possible mechanism would be to use the serial number assigned to the system. The mechanism used to choose between these System GUID Authority IDs is implementation specific. One usage would be to allow a vendor to define up to 128 distinct 128-bit System GUID schemes.

IMPLEMENTATION NOTE

System GUID Consistency and Stability

To support the purpose of System GUID, software should ensure that a single system uses identical System GUID and System GUID Authority ID values everywhere.

Implementers should carefully consider their stability requirements for the System GUID value. For example, some use cases may require that the value not change when the system is rebooted. In those cases, a mechanism that picks the EUI-48 value associated with the first Ethernet MAC address discovered might be problematic if the result changes due to hardware failure, system reconfiguration, or variations/parallelism in the discovery algorithm.

IMPLEMENTATION NOTE

Hierarchy ID vs. Device Serial Number

The Device Serial Number mechanism can also be used to uniquely identify a component (see Section 7.17). Device Serial Number may be a more expensive solution to this problem if it involves a ROM associated with each component.

IMPLEMENTATION NOTE

Virtual Functions and Hierarchy ID

The Hierarchy ID capability can be emulated by the Virtualization Intermediary (VI). Doing so provides VF software access to this Hierarchy ID information.

When VF hardware needs access to this information, the VF should implement the Hierarchy ID capability. This provides access to both VF software and hardware.

In some situations, the VF should get the same information as the PF. In other situations, particularly those involving migration of Virtual Machines, it may be appropriate to present the VF with Hierarchy ID information that differs from the associated PF and from other VFs associated with that PF.

The following mechanisms are supported:

	VF Hierarchy ID Capability	Hierarchy ID VF Configurable	Hierarchy ID Writeable	VF Software has access	VF Hardware has access	VF Hierarchy Data / GUID
1	Not Present	n/a	n/a	No	No	Not Emulated
2				Yes	No	Emulated
3	Present	0b	0b	Yes	Yes	Same as PF
4		1b	0b	Yes	Yes	Same as PF
5		1b	1b	Yes	Yes	Configured by VI

In mechanism 1, the the Virtualization Intermediary does not emulate the capability. VF software and hardware have no access.

In mechanism 2, the Virtualization Intermediary emulates the capability and returns whatever Hierarchy ID information is desired. VF software has access. VF hardware does not have access.

In mechanisms 3 and 4, VF information is the same as the PF and is automatically filled in from received Hierarchy ID messages. Both VF hardware and software have access.

In mechanism 5, VF information is configured by software (probably the VI). Both VF hardware and software have access.

In the PCI Express Base Specification, add new Section 7.x at the end of Chapter 7 as follows

7.x Hierarchy ID Extended Capability

The Hierarchy ID Extended Capability provides an optional mechanism for passing a unique identifier to Functions within a Hierarchy. At most one instance of this capability is permitted in a Function. This capability is not applicable to Bridges, Root Complex Event Collectors, and RCRBs.

This capability takes three forms:

In Upstream Ports:

- This capability is permitted any Function associated with an Upstream Port.
- This capability is optional in Switch Upstream Ports. Support in Switch Upstream and Downstream Ports is independently optional.
- This capability is mandatory in Functions that use the Hierarchy ID Message. This includes use by the Function's driver.
- Functions, other than VFs, that have Hierarchy ID Writeable Clear, must report the Message Routing ID, Hierarchy ID, System GUID Authority ID, and System GUID fields from the most recently received Hierarchy ID Message.
- All VFs that have Hierarchy ID Writeable Clear, must report the same Hierarchy ID Valid, Message Routing ID, Hierarchy ID, System GUID Authority ID, and System GUID values as their associated PF.
- PFs must implement this capability if any of their VFs implement this capability.
- Functions that have Hierarchy ID Writeable Set must report the Hierarchy ID Valid, Message Routing ID, Hierarchy ID, System GUID Authority ID, and System GUID values programmed by software.

In Downstream Ports:

- This capability is permitted in any Downstream Port. It is recommended that it be implemented in Root Ports.
- When present in a Switch Downstream Port, this capability must be implemented in all Downstream Ports of the Switch. Support in Switch Upstream and Downstream Ports is independently optional.
- In Downstream Ports, the Hierarchy ID, System GUID Authority ID, and System GUID fields are Read / Write and contain the values to send in the Hierarchy ID Message.
- A Hierarchy ID capability is not affected by Hierarchy ID Messages forwarded through the associated Downstream Port.

In Root Complex Integrated Endpoints:

- VFs that have Hierarchy ID Writeable Clear must report the same Message Routing ID, Hierarchy ID, System GUID Authority ID, and System GUID values as their associated PF.
- PFs must implement this capability if any of their VFs implement this capability.
- Functions, other than VFs, that have Hierarchy ID Writeable Clear, must report the same Hierarchy ID Valid, Message Routing ID, Hierarchy ID, System GUID Authority ID, and System GUID values. The source of this information is outside the scope of this specification.
- Functions that have Hierarchy ID Writeable Set must report the Hierarchy ID Valid, Message Routing ID, Hierarchy ID, System GUID Authority ID, and System GUID values programmed by software.

Figure 7-x1 details the layout of the Hierarchy ID Extended Capability.

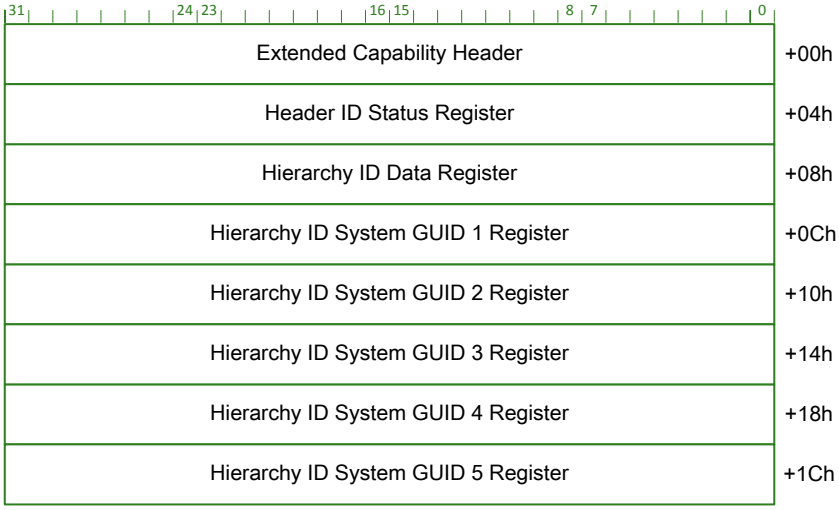

Figure 7-x1: Hierarchy ID Extended Capability

7.x.1 Hierarchy ID Extended Capability Header

Figure 7-x2 and Table 7-x2 detail allocation of fields in the Hierarchy ID Extended Capability Header.

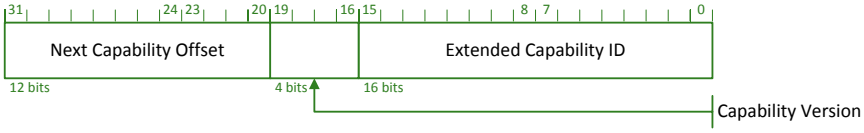

Figure 7-x2: Hierarchy ID Extended Capability Header

Table 7-x1: Hierarchy ID Extended Capability Header

Bit Location	Description	Attributes
15:0	**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 Hierarchy ID Capability is 0028h.	RO
19:16	**Capability Version** – This field is a PCI-SIG defined version number that indicates the version of the Capability structure present. Must be 1h for this version of the specification.	RO

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. For Extended Capabilities in configuration space, this offset is relative to the beginning of PCI-compatible Configuration Space and thus must always be either 000h (for terminating the list of Capabilities) or greater than 0FFh.	RO

7.x.2 Hierarchy ID Status Register

Figure 7-x3 and Table 7-x3 detail allocation of fields in the Hierarchy ID Status Register.

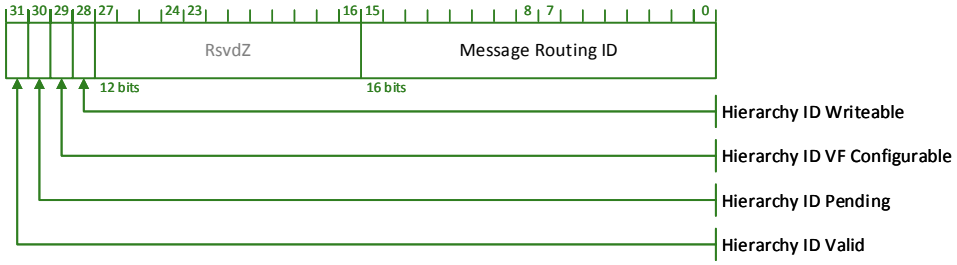

Figure 7-x3: Hierarchy ID Status Register

Table 7-x2: Hierarchy ID Status Register

Bit Location	Description	Attributes
15:0	**Message Routing ID** – In an Upstream Port, this field contains the Routing ID from the most recently received Hierarchy ID Message. This field is meaningful only if Hierarchy ID Valid is 1b. In a Downstream Port, this field is RsvdZ. For Root Complex Integrated Endpoints, this field is RsvdZ. This field defaults to 0000h.	RO/RsvdZ
29:16	RsvdZ	RsvdZ
28	**Hierarchy ID Writeable** – This bit is Set to indicate that the Hierarchy ID Data and GUID registers are read/write. This bit is Clear to indicate that the Hierarchy ID and GUID registers are read only. In Downstream Ports this bit is hardwired to 1b. In Upstream Ports, Functions that are not VFs must hardwire this bit to 0b. Root Complex Integrated Endpoints that are not VFs, must hardwire this bit to either 0b or 1b.	RW/RO

	VFs in an Upstream Port and Root Complex Integrated VFs are permitted to either: • hardwire this bit to 0b or • implement this bit as read / write with a default value of 0b.	
29	**Hierarchy ID VF Configrable** – This bit indicates that Hierarchy ID Writeable can be configured. If Hierarchy ID Writeable is implemented as read / write, this bit is 1b. Otherwise this bit is 0b.	RO
30	**Hierarchy ID Pending** – In Downstream Ports this requests the transmittion of a Hierarchy ID Message. Setting it requests transmission of a message based on the Hierarchy Data and GUID registers in this capability. This bit is cleared when either the transmit request is satisfied or the Link enters DL_Down. Behavior is undefined if the Hierarchy Data or GUID registers in this capability are written while this bit is Set. In Downstream Ports, this bit is Read / Write defaulting to 0b. In all other Functions, this bit is RsvdZ.	RW/RsvdZ
31	**Hierarchy ID Valid** – This bit indicates that the remaining fields in this capability are meaningful. In Downstream Ports, this bit is hardwired to 1b. In all other Functions, the following rules apply: ❑ If Hierarchy ID Writeable is Set, this bit is read/write, default 0b. ❑ If Hierarchy ID Writeable is Clear, this bit is read only, default 0b. • In VFs, this bit contains the same value as the associated PF. • In Functions other than VFs that are associated with an Upstream Port, this bit is Set when a Hierarchy ID Message is received, and Cleared when the Link is DL_Down. • In Root Complex Integrated Endpoints other than VFs, this bit contains a system provided value. The mechanism for determining this value is outside the scope of this specification.	RW/RO

7.x.3 Hierarchy ID Data Register

Figure 7-x4 and Table 7-x4 detail allocation of fields in the Hierarchy ID Data Register.

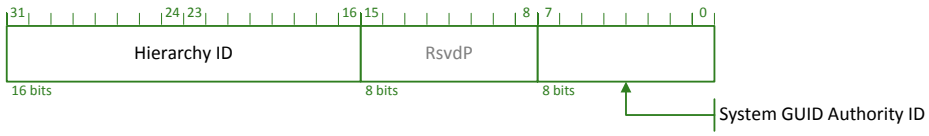

Figure 7-x4: Hierarchy ID Data Register

Table 7-x3: Hierarchy ID Data Register

Bit Location	Description	Attributes
7:0	**System GUID Authority ID** – This field corresponds to the System GUID Authority ID field in the Hierarchy ID Message. See Section 6.x for details. This field is meaningful only if Hierarchy ID Valid is 1b. If Hierarchy ID Writeable is Set, this field is read-write and contains the value programmed by software. If Hierarchy ID Writeable is Clear, this field is read only. The value is determined using the rules defined in Section 7.x. This field defaults to 00h.	RO/RW
15:8	RsvdP	RsvdP
31:16	**Hierarchy ID** – This field corresponds to the Hierarchy ID field in the Hierarchy ID Message. See Section 6.x for details. This field is meaningful only if Hierarchy ID Valid is 1b. If Hierarchy ID Writeable is Set, this field is read-write and contains the value programmed by software. If Hierarchy ID Writeable is Clear, this field is read only. The value is determined using the rules defined in Section 7.x. This field defaults to 0000h.	RO/RW

7.x.4 Hierarchy ID GUID 1 Register

Figure 7-x5 and Table 7-x5 detail allocation of fields in the Hierarchy ID GUID 1 Register.

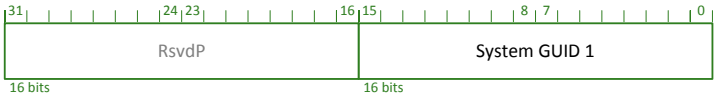

Figure 7-x5: Hierarchy ID GUID 1 Register

Table 7-x4: Hierarchy ID GUID 1 Register

Bit Location	Description	Attributes
15:0	**System GUID 1** – This field corresponds to bits [143:128] of the System GUID in the Hierarchy ID Message. See Section 6.x for details. This field is meaningful only if Hierarchy ID Valid is 1b. If Hierarchy ID Writeable is Set, this field is read-write and contains the value programmed by software. If Hierarchy ID Writeable is Clear, this field is read only. The value is determined using the rules defined in Section 7.x. This field defaults to 0000h.	RO/RW
31:0	RsvdP	RsvdP

7.x.5 Hierarchy ID GUID 2 Register

Figure 7-x6 and Table 7-x6 detail allocation of fields in the Hierarchy ID GUID 2 Register.

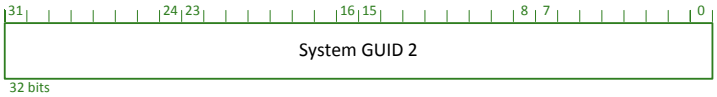

Figure 7-x6: Hierarchy ID GUID 2 Register

Table 7-x5: Hierarchy ID GUID 2 Register

Bit Location	Description	Attributes
31:0	**System GUID 2** – This field corresponds to bits [127:96] of the System GUID field in the Hierarchy ID Message. See Section 6.x for details. This field is meaningful only if Hierarchy ID Valid is 1b. If Hierarchy ID Writeable is Set, this field is read-write and contains the value programmed by software.	RO/RW

	If Hierarchy ID Writeable is Clear, this field is read only. The value is determined using the rules defined in Section 7.x. This field defaults to 0000 0000h.	

7.x.6 Hierarchy ID GUID 3 Register

Figure 7-x7 and Table 7-x7 detail allocation of fields in the Hierarchy ID GUID 3 Register.

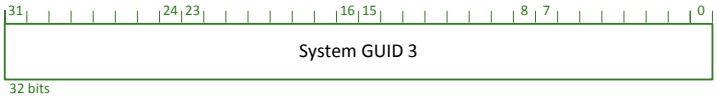

Figure 7-x7: Hierarchy ID GUID 3 Register

Table 6-x7: Hierarchy ID GUID 3 Register

Bit Location	Description	Attributes
31:0	**System GUID 3** – This field corresponds to bits [95:64] of the System GUID field in the Hierarchy ID Message. See Section 6.x for details. This field is meaningful only if Hierarchy ID Valid is 1b. If Hierarchy ID Writeable is Set, this field is read-write and contains the value programmed by software. If Hierarchy ID Writeable is Clear, this field is read only. The value is determined using the rules defined in Section 7.x. This field defaults to 0000 0000h.	RO/RW

7.x.7 Hierarchy ID GUID 4 Register

Figure 7-x8 and Table 7-x8 detail allocation of fields in the Hierarchy ID GUID 4 Register.

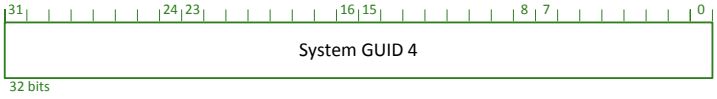

Figure 7-x8: Hierarchy ID GUID 4 Register

Table 7-x7: Hierarchy ID GUID 4 Register

Bit Location	Description	Attributes
31:0	**System GUID 4** – This field corresponds to bits [63:32] of the System GUID field in the Hierarchy ID Message. See Section 6.x for details. This field is meaningful only if Hierarchy ID Valid is 1b. If Hierarchy ID Writeable is Set, this field is read-write and contains the value programmed by software. If Hierarchy ID Writeable is Clear, this field is read only. The value is determined using the rules defined in Section 7.x. This field defaults to 0000 0000h.	RO/RW

7.x.8 Hierarchy ID GUID 5 Register

Figure 7-x9 and Table 7-x9 detail allocation of fields in the Hierarchy ID GUID 5 Register.

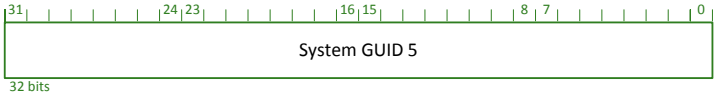

Figure 7-x9: Hierarchy ID GUID 5 Register

Table 7-x8: Hierarchy ID GUID 5 Register

Bit Location	Description	Attributes
31:0	**System GUID 5** – This field corresponds to bits [31:0] of the System GUID field in the Hierarchy ID Message. See Section 6.x for details. This field is meaningful only if Hierarchy ID Valid is 1b. If Hierarchy ID Writeable is Set, this field is read-write and contains the value programmed by software. If Hierarchy ID Writeable is Clear, this field is read only. The value is determined using the rules defined in Section 7.x. This field defaults to 0000 0000h.	RO/RW

In Appendix F, change Table F-2 as follows:

Table F-2: PCI-SIG-Defined VDM Subtype Usage

Subtype	Routing r[2:0]	Type	Description
…	…	…	…
0000 0001	011	MsgD	Hierarchy ID Message, See Section 2.2.8.6.5 and Section 6.x

For SR-IOV Specification Modify Section 3.7 as shown:

Table 3-22: SR-IOV Usage of PCI Express Extended Capabilities

Extended Capability ID	Description	PF Attributes	VF Attributes
…	…	…	…
0028h	Hierarchy ID	Base	Base

For PCI Code and ID Assignment Specification Modify Section 3 as shown:

Table 3-1 Extended Capability IDs

ID	Extended Capability
…	…
0028h	Hierarchy ID Message