Enterprise
SSD Form Factor

Version 1.0a

SSD Form Factor Version1_a.docx

December 12, 2012

Please send comments to Mark Myers

specification@ssdformfactor.org

INTELLECTUAL PROPERTY DISCLAIMER

THIS SPECIFICATION IS PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER INCLUDING ANY WARRANTY OF MERCHANTABILITY, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, SPECIFICATION, OR SAMPLE.

NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED OR INTENDED HEREBY.

THE SSD FORM FACTOR WORK GROUP DISCLAIMS ALL LIABILITY, INCLUDING LIABILITY FOR INFRINGEMENT OF PROPRIETARY RIGHTS, RELATING TO IMPLEMENTATION OF INFORMATION IN THIS SPECIFICATION. THE SSD FORM FACTOR WORK GROUP DOES NOT WARRANT OR REPRESENT THAT SUCH IMPLEMENTATION(S) WILL NOT INFRINGE SUCH RIGHTS.

The document revision is changed to 1.0a.

Changes are:

1. Changed 12V Tolerance to unspecified, and added an Implementation Note regarding power tolerance. Leaves Max continuous current to 2.45A to match wider voltage range.
2. Changed text to show SFF-8639 as the approved normative document for mechanical specification. Section 3 is now informative, not normative. There is still an overview and primary drawings, but removed more detailed sections and reference SFF-8639. Specifically removed sections on intermateability, layouts (footprints), and blind mate & wipe,
3. Changed pin numbers to match approved SFF-8639 (SFF-8639, 2012). Specifically E17-E39 are relabeled S15-S28, E17-E25. No change in pin placement, aligning pin label with SFF-8630 (SFF-8630, 2012).
4. Changes to align with SFF-8630 (SAS Multilink). Removed references to SAS Multilink using SMBus, 3.3Vaux, or Act2. These changes are only in pin list tables.
5. Added multiple Extended Capability enumeration fields to Section 9.2. This allows out of band discovery of Dual Port support, RefClk extensions, SM-Bus Temp sensor, Active/Passive mode.
6. Added description to improve clarity, Section 2.6 and 2.7, added P-Init to Figure 29.
7. Editorial changes, including updating references and bibliography.

Table of Contents

Table of Figures

1 Overview

1.1 Overview

This specification defines the electrical and mechanical requirements for a PCI Express connection to the existing standard 2.5" and 3.5" disk drive form factors. This is intended for PCIe connections to SSDs (Solid State Drives) for the enterprise market of servers and storage systems. This provides a new PCIe form factor that is storage friendly, leveraging both the existing PCIe specification and the existing 2.5" and 3.5" drive mechanical specifications. This standard allows system designs that can support a flexible mix of new enterprise PCI Express drives, and existing SAS and SATA drives.

Note: The ***References and Bibliography*** are listed at the end, in section 8 on page 39. The ***Specification Conventions*** are listed at the end, in section 11 on page 54. This includes the standard material such as the definition of "Shall", "Should", "May", etc. and bibliographic material to reference specifications. Also included are definitions of terms such as "PCIe", "PCIe Gen 3", "PCIe x4", "SFF", and "SSD". The bibliographic reference use the original author name, for example (SFF-8482, 2006) while the section 8.3 give the current designator, for example EIA-966.

1.2 Goals

The following are the characteristics of the product envisioned using this specification:

- PCIe[1] connection to Enterprise Solid State Drives (SSDs).
- Standardize connector and form factor – but enabling innovation using the PCIe model.
- Fits in existing drive mechanical enclosures using a disk backplane. Supports both 2.5" drive mechanical enclosures and 3.5" drive mechanical enclosures.
- Support customer expectation for storage device:
    - Externally accessible & Hot Swappable (with surprise removal)
    - Support existing OEM's existing drive infrastructure (device detection and indicators)
    - Support existing OEM thermal architectures, and support for future enhancements.
- Support for both single port (typical servers), or dual port (typically storage systems).
- Meet projected system storage device performance requirements for systems introduced in the 2H 2011 to 2016 timeframe.
- Enables flexible system designs that support Enterprise PCIe Express SSDs and SAS or SATA drives using the same connector family, allow systems to support flexible mix of capacity (SAS or SATA hard drives) or Enterprise PCIe SSDs.

1.3 Technical Summary

A technical summary of the "SSD Form Factor Working Group" is:

- Focus is on connector which is supported in both the existing 2.5" and 3.5" form factors.
- Supports the 2.5" drive (SFF-8223, 2006) specification[2] unchanged, or supports the 3.5" drive (SFF-8301, 2010) specification unchanged.
- Compatible with existing SAS connector (SFF-8680, 2012). Meaning the connectors allow interoperation, see Section 10. For example a new backplane receptacle would accept new or existing drives. The connector is doubled sided extension to existing SAS connector. New pins are 0.8mm pitch across the complete B-side and on top of SAS key.
- Support for PCIe x4 (Gen2 and Gen3) signaling. References PCIe specifications for high speed electrical specifications – unchanged.
- The electrical channel is assumed to be similar to SAS-3 channel – short board trace, longer cable, trace on storage backplane, SSD.

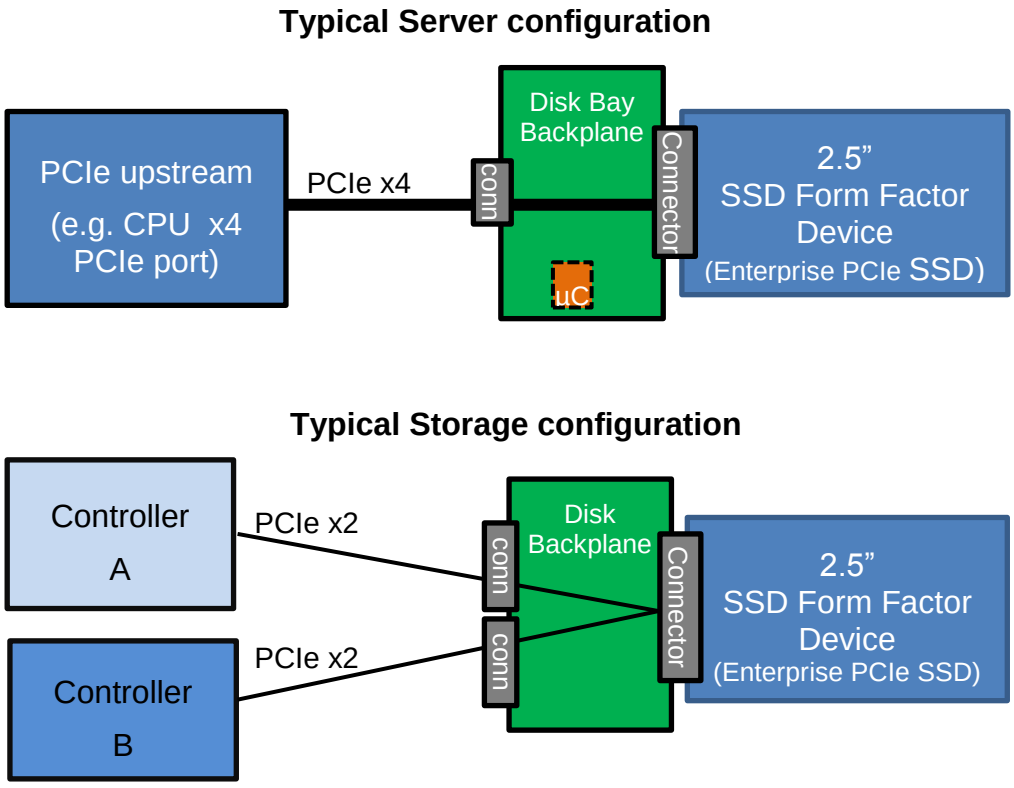
[1] Definition of PCIe and reference specification are included in section 8 on page 40

[2] Definition of SAS, SATA, SFF and reference specifications including official designations (EIA, ANSI, INCITIS numbers) are included in section 8 on page 40

- 25W connector limit, with 12V only delivery. Expecting Enterprise PCIe SSDs at range of power levels, and expecting Enterprise PCIe SSDs to support software settable power limit, and temperature monitoring.
- There is optional support for 3.3V aux to allow probing device information before system power on (Vital Product Data – VPD – over SM bus.)
- Pins list is the merging of SATA/SAS signals and the PCIe x4 plus PCIe side band. The actual high speed lanes for SAS and SATA are separate pins from the Enterprise PCIe pins. This enables direct connection to SAS or SATA controller and separate pins to high performance Enterprise PCIe upstream device without an intervening repeater, mux, or controller.
- Support Single x4 or Dual x2 port as (2@x2) with dual PCIe side band
- Supports Hot swap support details, basics same as SAS
- Supports SM bus for out of band discovery
- Supports PCIe power management features

Figure 1 shows example configurations.

Figure 1: Example configurations



1.3.1 Flexible Drive Backplane option

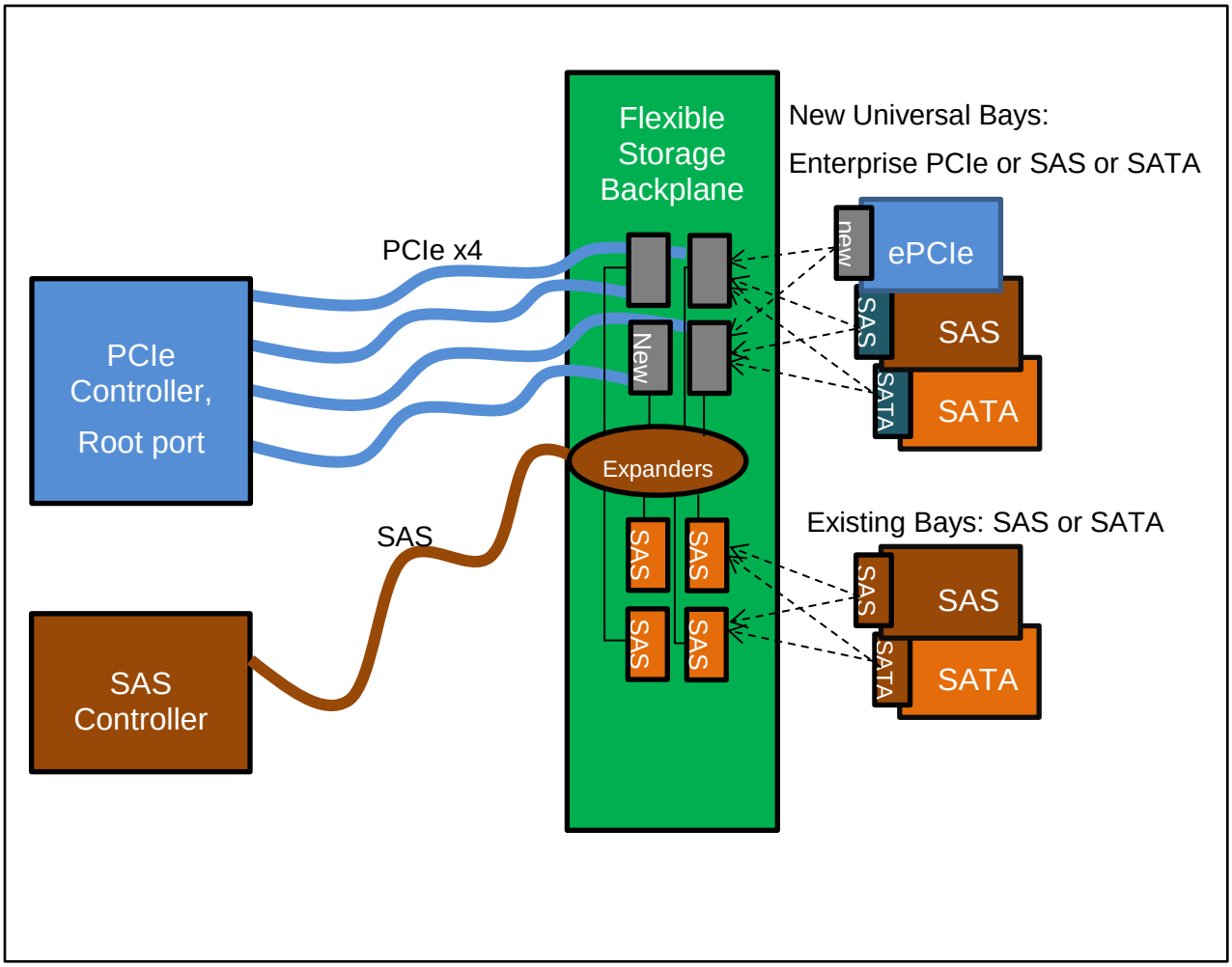
The current technical direction allows for system implementations that support flexible storage backplane bays supporting a mix of Enterprise PCIe or SAS or SATA or SATA Express devices. The system would provide both a PCIe connection along with a separate connection to a SAS or SATA controller. The system support infrastructure would also have to be extended to support the super-set capabilities (power amperage, management signaling).

The concept is that a single system or backplane design would support either Enterprise PCIe SSD or SAS/SATA drives allowing an optimal balance of performance and/or capacity to be achieved. The assumption is that SAS expanders allow adding SAS ports at low additional cost. This is shown in Figure 2.

The drawing of Flexible Backplane (Figure 2) is not the only implementation since a Hard Disk Drive (HDD) optimized using just SAS/SATA and other implementations for just PCIe are also expected. There are multiple variations of flexible backplane based on the capabilities of the supporting controllers. The informative appendix, Section 10.2 on page 50, shows multiple other options.

The capability to build a flexible backplane is based on extending the existing SAS connector (SFF-8680, 2012) sharing functions like power and giving it the ability to meet SAS MultiLink (SFF-8630, 2012) functionality as well as giving it the capability to deliver four ports of high speed PCIe. SFF-8630 does not accommodate the additional 2 x PCIe ports nor does it provide the PCIe sidebands.

Figure 2: Flexible Backplane Implementation



1.4 Scope

This is a form factor specification that focuses on extending the existing connector for PCIe use. The new connector is a backward compatible extension of the existing SAS connector (SFF-8482, 2006). The

overall mechanical form factor is compliant with the 2.5" form factor (SFF-8223, 2006) or 3.5" form factor (SFF-8301, 2010).

This specification defines the following:

- Pin list and pin out of the new connector.
- Mechanical definition of the new connector, including latching, keying, and retention.
- Limited electrical specification (largely reference PCIe CEM and SAS Specifications).
- Hardware hooks for Hot Swap
- Signal definitions for Dual Port
- Definitions of a basic product data accessible from SM-Bus

1.5 Outside of Scope

This specification is only normative for connector and related definitions (just listed)

The actual usage is not defined. While we envisioned having an SSD using NVM Express interface (NVM Express, 2011) or SOP (SCSI over PCIe) (T10, SOP, Draft 2012), these are not required and other PCIe device interface models would work.

The specification does give possible usage (aka. implementation notes) for Hot Swap, Dual Port, and SM-bus but the details are system implementation specific. These are just meant to illustrate a possible usage.

Specifically this specification is informative (not normative) in the following areas:

- The drive outline is defined by EIA/ECA-720 (SFF-8223, 2006) for 2.5" drives or by EIA-740 (SFF-8301, 2010) for 3.5" drives. The drive carrier and storage enclosure are implementation specific.
- Mechanical drawings normative document is SFF-8639 (SFF-8639, 2012). Section 3 is informative, for ease of reference.
- SAS signal definitions and specifications – normative are the SAS Specifications (T10, Various).
  Note: *SAS-x4 signal definitions in this specification are a draft proposal to T10. This includes SOP (SCSI over PCIe).*
- PCIe signal definitions and specifications – normative are the PCIe Specifications (PCI SIG, various).
- SATA/SATA Express signal definitions and specification – normative are the SATA Specifications (SATA-IO, various).

  Note: *SATA Express signal definitions in this specification are a proposal to the SATA-IO Cabcon*

- SM Bus signal definitions and specifications – normative are the SM Bus specifications (SMBus, 2000)
- PCIe device functionality and register details and the SM-bus registers are defined by the specific device specifications. Likely referencing the PCIe Specification for standard PCIe device discovery and setup.
- Specific system implementation of presence pins (hot plug)
- Specific system implementation of dual port functionality
- References to other related connectors - specifically keying and overlaying of pins.

This is envisioned as a PCIe SSD specification and there is no standardization for additional connections beyond the drive PCIe connector. There are both backplane and cable versions of the connector receptacle.

2 Signal List

The signal list is a combination of SAS/SATA and PCIe signals. Figure 3: Signal List Summary gives an overview of the signal groups. Figure 4: Signal List Table gives the complete signal list, usage, and which document is normative for the signal definition. Most of the signals used in this specification are defined in other documents. Figure 5 gives the mapping of signal pins to connector pins.

2.1 Signal Lists

Figure 3: Signal List Summary (pin counts)

	Usage	Signals	Contacts
New Enterprise PCIe, (E1E25. S16-S28)	x4 PCIe Gen3	4 lanes (Tx,Rx) (diff+gnd) = 4*(4sig+2gnd) +1gnd(end) (S16-S28 shared with lanes 2-3 of SAS x4)	25 (16 sig, 9 Gnd)
	PCIe Sideband -	Enterprise Reset, RefClk[4](diff+gnd)	4
	PCIe optional	*SM-bus(2), 3.3V-aux, Reserved(2)*	*5*
	Dual Port support	*2nd Reset, RefClk[1] (diff+gnd), DualPortEn#*	*5*
Existing SAS/SATA (S1-S14)	*2 SATA/SATA Express/SAS Ports[2]*	*2 lanes (Tx,Rx) (diff+gnd) = 2 ports * 7 pins*	*14 (8 sig 6 Gnd),*
Existing Power (P1-P15)	Power return	Gnd	3
	Power	12V (25W max)	3
	SATA/SAS Power Rail[2]	*5V*	*3*
	Shared side band	Presence, Interface Detect[2], Activity	3
	SATA Express Sideband[2]	*Client Reset, Reserved(Wake/OBFF#)[3], Reserved(ClkReq#/DevSLP)[3]*	*3*
	Total		**68** 39 new, 29 existing, (4 redefined[2])

The following are the notes on the signal list:

1. The high speed signals (PCIe lanes) are separate from SATA Express and Enterprise PCIe SSD. This is driven by system requirements. There are separate PCIe Reset signals. See Section 10.1 for details.
2. This specification and SATA Express replace the 3.3V signals and 1 ground with PCIe sideband signals: Interface Detect, Client Reset, and new reserved signals, Reserved(WAKE#/OBFF#), Reserved(CLKReq#/DevSLP).
3. Reserved(WAKE#/OBFF#) and Reserved(ClkReq#/DevSLP) are optional, and are for system and device power management. The usage for power management of SSD devices is not yet completely defined and are new proposals to the PCI Express standard (PCI SIG ECNs, various).
4. SATA Express (client PCIe) proposal includes an optional RefClk. When the optional RefClk is not present it will be dependent on PCI-SIG (PCI SIG ECNs, various) changes to support independent spread spectrum clocks (with increased skip ordered set insertion). This is not currently available but is anticipated in future revisions of the PCI-SIG.

Figure 4: Signal List Table (defining specification)

Pin Group	Pin Name	Signal Description	Usage	Defining Specification
PCIe High Speed	PETp0, PETn0, PETp1, PETn1, PETp2, PETn2, PETp3, PETn3	PCIe x4 Transmit	Enterprise PCIe	PCIe
	PERp0, PERn0, PERp1, PERn1, PERp2, PERn2, PERp3, PERn3	PCIe x4 Receive	Enterprise PCIe	PCIe
PCIe Side Band	RefClk0+, RefClk0-	PCIe RefClk (primary port A)	Enterprise PCIe Optional SATAe	PCIe
	ePERst0#	Enterprise PCIe Reset (Port A)	Enterprise PCIe	This Spec
PCIe Optional Side Band	SMClk, SMDat	SM Bus	Enterprise PCIe, SAS x4	SM Bus
	3.3VAux	Power for SMbus accesses	Enterprise PCIe, SAS x4	This Spec
	RSVD(WAKE#/OBFF#)	Reserved for system/device power management	ePCIe, SATAe	PCIe
	RSVD(ClkReq #/DevSLP)		ePCIe, SATAe	PCIe
	sPERst#	SATA Express/SAS Reset	SAS, SATAe only	PCIe
Dual Port PCIe Side Band	RefClk1+, RefClk1-	PCIe RefClk (dual port, port B)	Dual Port ePCIe	PCIe
	ePERst1#	Enterprise PCIe Reset (Port B)	Dual Port ePCIe	This Spec
	DualPortEn#	Dual Port PCIe enable(2@x2)	Dual Port ePCIe	This Spec
SAS High Speed	S0T+ (A+), S0T- (A-)	SAS 0 Transmit (SATA/SATA Express Transmit)	SAS, SATA/SATAe	SAS, SATA
	S0R- (B-), S0R+ (B+)	SAS 0 Receive (SATA/SATA Express Receive)	SAS, SATA/SATAe	SAS, SATA
2nd Port SAS High speed	S1T+, S1T-	SAS 1 Transmit (SATA Express Transmit)	SAS, SATAe only	SAS, SATAe
	S1R+, S1R-	SAS 1 Receive (SATA Express Receive)	SAS, SATAe only	SAS, SATAe
SAS x4 High speed	S2T+, S2T-, S3T+, S3T-	Additional SAS X4 transmit	Overlapped with PET1/PER1, PET2/PER2	SAS x4
	S2R+, S2R-, S3R+, S3R-	Additional SAS x4 receive		SAS x4
SAS support	SAS Reset	Additional SAS Reset	Overlapped with SATAe sPERst#, OBFF#	SAS x4
	SAS 2nd Activity	Additional SAS 2nd Activity		SAS x4
Shared Support	Activity	Drive Active	Shared	SAS, SATA
	PRSNT#	Presence; (Drive type encoded on PRSNT# and IfDet#)	Shared	This Spec, SATAe
	IfDet#	Interface Detect (Drive type encoded on PRSNT# and IfDet#)	Shared	This Spec, SATAe
Power	12V	Only power for Enterprise PCIe	Shared	This Spec
	5 V	Power for SAS, SATA, SATAe	SAS, SATA/SATAe	SAS, SATA
	3.3 V	Reclaimed pins for SATA Express sideband	Not used	SAS, SATA
	GND		Shared	

Notes: SATAe → SATA Express ePCIe → Enterprise PCIe

2.2 Connector pin out

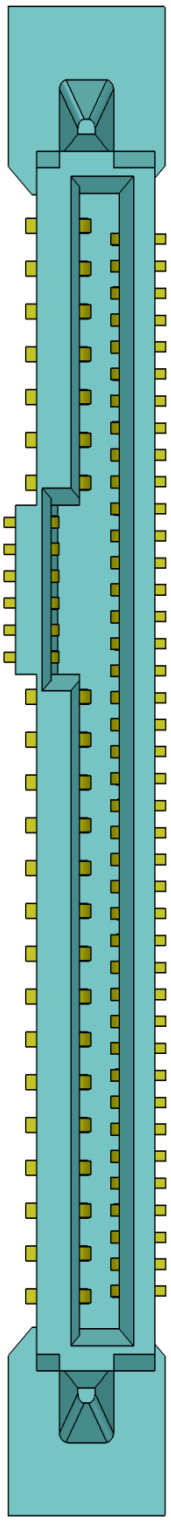
Figure 5: Pin Out Drawing (Receptacle pin naming).

Drive	Usage	Signal Description	Name	Mating	Pin #
		Ground	GND	2nd	S1
input	SAS+SATA	SAS/SATA/SATAe 0 Tx+	S0T+ (A+)	3rd	S2
input	SAS+SATA	SAS/SATA/SATAe 0 Tx -	S0T- (A-)	3rd	S3
		Ground	GND	2nd	S4
output	SAS+SATA	SAS/SATA/SATAe 0 Rcv -	S0R- (B-)	3rd	S5
output	SAS+SATA	SAS/SATA/SATAe 0 Rcv +	S0R+ (B+)	3rd	S6
		Ground	GND	2nd	S7
input	Dual Port	ePCIe RefClk + (port B)	RefClk1+	3rd	E1
input	Dual Port	ePCIe RefClk – (port B)	RefClk1-	3rd	E2
input	ePCIe opt	3.3V for SM bus	3.3Vaux	3rd	E3
input	Dual Port	ePCIe Reset (port B)	ePERst1#	3rd	E4
input	ePCIe	ePCIe Reset (port A)	ePERst0#	3rd	E5
Bi-Dir		Reserved (vendor)	RSVD(Vendor)	3rd	E6
input	SATAe	Reserved(WAKE#/OBFF)	RSVD(Wake#)	3rd	P1
input	SATAe	SATAe Client	sPCIeRst	3rd	P2
input	SATA+PCIe	Reserved(ClkReq#/DevSLP)	RSVD(DevSLP)	2nd	P3
output	SATAe + ePCIe	Interface Detect (Was GND-precharge)	IfDet#	1st	P4
	all	Ground	GND	2nd	P5
	all			2nd	P6
input	SAS+SATA	Precharge	5 V	2nd	P7
input	SAS+SATA	SATA, SATAe, SAS only		3rd	P8
input	SAS+SATA			3rd	P9
	all	Presence (Drive type)	PRSNT#	2nd	P10
Bi-Dir	all	Activity(output)/Spinup	Activity	3rd	P11
	all	Hot Plug Ground	GND	1st	P12
input	all	Precharge	12 V	2nd	P13
input	all	All – 12V		3rd	P14
input	all	Only power for ePCIe SSD		3rd	P15

ePCIe → Enterprise PCIe (separate from SATA/SAS)

SATAe → SATA Express
(Client PCIe- muxed on SATA/SAS signals)

SAS4 → SAS x4



Pin #	Mating	Name	Signal Description	Usage	Drive
E7	3rd	RefClk0+	ePCIe Primary RefClk +	ePCIe+SATAe	input
E8	3rd	RefClk0-	ePCIe Primary RefClk -	ePCIe+SATAe	input
E9	2nd	GND	Ground		
E10	3rd	PETp0	ePCIe 0 Transmit +	ePCIe	input
E11	3rd	PETn0	ePCIe 0 Transmit -	ePCIe	input
E12	2nd	GND	Ground		
E13	3rd	PERn0	ePCIe 0 Receive -	ePCIe	output
E14	3rd	PERp0	ePCIe 0 Receive +	ePCIe	output
E15	2nd	GND	Ground		
E16	3rd	RSVD	Reserved		
S8	2nd	GND	Ground		
S9	3rd	S1T+	SAS/SATAe 1 Transmit +	SAS+SATAe	input
S10	3rd	S1T-	SAS/SATAe 1 Transmit -	SAS+SATAe	input
S11	2nd	GND	Ground		
S12	3rd	S1R-	SAS/SATAe 1 Receive -	SAS+SATAe	output
S13	3rd	S1R+	SAS/SATAe 1 Receive +	SAS+SATAe	output
S14	2nd	GND	Ground		
S15	3rd	RSVD	Reserved		
S16	2nd	GND	Ground		
S17	3rd	PETp1/S2T+	ePCIe 1 /SAS 2 Transmit +	ePCIe+SAS4	input
S18	3rd	PETn1/S2T-	ePCIe 1 /SAS 2 Transmit -	ePCIe+SAS4	input
S19	2nd	GND	Ground		
S20	3rd	PERn1/S2R-	ePCIe 1 /SAS 2 Receive -	ePCIe+SAS4	output
S21	3rd	PERp1/S2R+	ePCIe 1 /SAS 2 Receive +	ePCIe+SAS4	output
S22	2nd	GND	Ground		
S23	3rd	PETp2/S3T+	ePCIe2 / SAS 3 Transmit +	ePCIe+SAS4	input
S24	3rd	PETn2/S3T-	ePCIe2 / SAS 3 Transmit -	ePCIe+SAS4	input
S25	2nd	GND	Ground		
S26	3rd	PERn2/S3R-	ePCIe 2 / SAS 3 Receive -	ePCIe+SAS4	output
S27	3rd	PERp2/S3R+	ePCIe 2 / SAS 3 Receive +	ePCIe+SAS4	output
S28	2nd	GND	Ground		
E17	3rd	PETp3	ePCIe 3 Transmit +	ePCIe	input
E18	3rd	PETn3	ePCIe 3 Transmit -	ePCIe	input
E19	2nd	GND	Ground		
E20	3rd	PERn3	ePCIe 3 Receive -	ePCIe	output
E21	3rd	PERp3	ePCIe 3 Receive +	ePCIe	output
E22	2nd	GND	Ground		
E23	3rd	SMClk	SM-Bus Clock	PCIe opt	Bi-DIr
E24	3rd	SMDat	SM-Bus Data	PCIe opt	Bi-Dir
E25	3rd	DualPortEn#	ePCIe 2x2 Select	Dual Port	input

Notes on Pin out table

- Relative to Standard SATA and SAS connector's retention detents: new Pins E9-E12 & S21-S24 & E18-E21 (yellow shaded pin numbers) displaced required SAS connector retention detents. A side latch is used for Enterprise PCIe cable retention described in Section 3.5.
- "Drive" column shows signal direction relative to the plug connector (Enterprise PCIe SSD). Input means sourced by the system. Output means sourced by the Enterprise PCIe SSD device and an input to system.
- "Name" and "Signal Description" columns are relative to the receptacle connector (backplane).
    - The SAS specification labels the pins based on if it is a plug (drive) or receptacle (backplane). The SAS receptacle connector pin-out (shown) uses naming relative to the host, while the SAS plug connector pin-out (not-shown) is relative to the Enterprise PCIe SSD.
    - This document uses PCIe convention. PCIe pin-out is always relative to the host. The following is from the PCIe Specification (PCI SIG CEM, 2007).
        - "[PCIe Transmit (i)+] PETp<i> and [PCIe Transmit (i)-] PETn<i> pins (the transmitter differential pair of the connector) shall be connected to the PCI Express transmitter differential pair on the system board, and to the PCI Express receiver differential pair on the add-in card.
            - These pins are inputs on an Enterprise PCIe SSD.
        - "[PCIe Receive (i)+] PERp<i> and [PCIe Receive (i)-] PEPn<i> pins (the receiver differential pair of the connector) shall be connected to the PCI Express receiver differential pair on the system board, and to the PCI Express transmitter differential pair on the add-in card.
            - These pins are outputs of an Enterprise PCIe SSD.
- Pins P1-P3 SATA Express sideband signals are redefined from SATA & SAS specifications which defined these pins as 3.3V. The 3.3V rail is not widely used on existing products and SATA Express is using these for sideband signals. Drive pins P1-P3 should be tolerant of 3.3V applied from the system.
- Pin P4 is now Interface-Detect (IfDet#) and is redefined from SATA & SAS specifications which defined this pin as a ground. It remains 1st mate.
- Pin P3-Reserved(ClkReq#/DevSLP) is listed as Reserved since the usage is not yet specified for Enterprise PCIe SSDs. The ClkReq# mode is used in some PCIe client devices (SATA-IO, Draft 2011). The DevSLP mode is defined for legacy SATA devices (SATA-IO, 2011).
- Pin P1-Reserved(WAKE#/OBFF#) is listed as Reserved since the usage for Enterprise PCIe SSD is not completely understood.

SATA Express pin placement is only a proposal on how to align this specification's Enterprise PCIe pin-out with proposals for pins in SATA Express. The SATA Express signals are not yet standardized, and this document is not the normative document on pin out or pin definition for SATA Express.

SAS-x4 pin placement is a proposal on how to align this specification's Enterprise PCIe pin-out with proposals for pins to support SAS-x4

2.3 PRSNT# - Presence Detect and IfDet# - Interface Detect Signal Definition

The Presence (PRSNT#) is similar to Presence on the existing SFF-8482 connector. There is a redefinition of the IfDet# (pin P4) to encode drive type. The usage of the combined signals is to detect a drive is present and the drive type. Figure 6 shows how the drive type can be decoded from PRSNT# and IfDet# pins. .

Figure 6: Drive type - Pin decoding

	P10→ PRSNT# Drive-output	P4→ IfDet# (was GND) Drive-output	E25→ DualPortEn# Drive= Input	Comment
SATA/SAS	Gnd (drv)	Gnd (drv)	Open(NC)	Existing Drive (P4 is GND, E25 is NC)
SATA Express (Client PCIe)	Gnd (drv)	Open (drv)[1]	Open(NC)	PCIe on SATA/SAS lanes (S1-S14) sPERst# is used, no RefClk used (P4 is NC, E25 is NC)
Enterprise PCIe (SFF-8639)	Open (drv)	Gnd (drv)	High (by system)	PCIe are on new Enterprise lanes (E7-E15, S16-S28, E17-E22) RefClk0 and ePERst0# are used
Dual Port Enterprise PCIe (SFF-8639)	Open (drv)	Gnd (drv)	Low (by system)	PCIe are on new Enterprise lanes (PortA = E7- E15, S16-S22 PortB = S22-S28, E17-E22) RefClk0, ePERst0#, and RefClk1, ePERst1# are used
No drive present	Open (drv)	Open (drv)	-	All open

1. SATA Express may define IfDet# as a pull up & resistance. The SATA Express documentation is normative.

Either P10 or P4 being low indicates a drive is present. Further decoding is when P4 is open, then the PCIe lanes use the SATA pins (S1-S14) as defined for SATA Express.

In-band signal discovery is used to determine SATA vs. SAS. In-band signal discovery is used for width (x1, x2, x4) determination for SAS and PCIe. This is consistent with the existing SAS and PCIe standards.

2.4 DualPortEn# - PCIe Dual Port Enable

The Enterprise PCIe SSD can be configured to train as either a single x4 controller or dual x2. The mechanism for enabling dual port operation is DualPortEn# (pin E25). DualPortEn# is pulled high internal to the Enterprise PCIe SSD. If DualPortEn# is left open then the PCIe is configured as Single x4 port. If DualPortEn is pulled low by the system (driven low or grounded by backplane), then the Enterprise PCIe SSD has dual Ports enabled. DualPortEn# is part of the drive type determination shown in Figure 6.

DualPortEn# is a static signal. DualPortEn# must be stable 1uS before either ePERst[1:0]# are de-asserted, and DualPortEn# can only change if both ePERst[1:0]# are asserted. The Enterprise PCIe SSD may sample this signal at any time, and operation is undefined if DualPortEn# changes state during operation. In practice is sampled during Enterprise PCIe SSD local reset sequence to configure the ports.

The electrical signal characteristics of DualPortEn# follows the sideband interface signals of PCIe as stated in ***PCIe CEM Specification 2.6 Auxiliary Signal Parametric Specifications*** (PCI SIG CEM, 2007).

Section 5.1 on page 33 gives more details on Dual Port operation.

2.5 Reserved(WAKE#/OBFF#), Reserved(ClkReq#/DevSLP)

The Reserved(Wake#/OBFF#), Reserved(ClkReq#/DevSLP), signals are all optional and are used in device power management. These signals are defined by the PCIe Specification (PCI SIG ECNs, various). At this time the pin definition and usage in Enterprise PCIe SSD is not completely defined, so the pins are listed as reserved.

2.6 ePERst[1:0]# - PCIe Reset

The Enterprise PCIe Reset (ePERst[1:0]#) are logically the same as defined in the PCIe Specification (PCI SIG CEM, 2007) but for Enterprise PCIe SSD there are additional input current requirements. This allows for a drive to have pull up or pull down circuitry to support hot removal. This is defined in Section 7.3. There are two resets to allow for dual port operation, as defined in Section 5.

The Enterprise PCIe Reset (ePERst0#) is required to be supplied by the system for the Enterprise PCIe SSD.

If DualPort# is asserted, then Port B Enterprise PCIe Reset (ePERst1#) is required to be supplied by the system for the Port B of the Enterprise PCIe SSD.

2.7 RefClk[1:0][+/-] – Enterprise PCIe Clocks

The Enterprise PCIe Reference Clock (RefClk0[+/-]) **should** be supplied by the system for all Enterprise PCIe SSDs.

If DualPort# is asserted, then Port B Enterprise PCIe Reference Clock (RefClk1[+/-]) **should** be supplied by the system for all Port B (second port) of the Enterprise PCIe SSDs.

Implementation Note: *RefClk is required for systems that use Spread Spectrum Clocking (SSC) to track the changes in the reference clock. In a typical server system it is expected that Enterprise PCIe SSD will need to use RefClk# for correct operation. This is the same as PCIe slots. RefClk does not need to be supplied if both the platform and drive support SRIS - Separate RefClk with Independent SSC* (PCI SIG ECNs, various) *or SRNS - Separate RefClk with No SSC (independent clocking).*

Implementation Note: *RefClk is required for Enterprise PCIe SSD but note that SATA-Express is pursuing a PCI SIG specification change in future root ports that will enable no RefClk for SATA-Express.*

2.8 Activity - Activity indication signal

The Activity signal asserts with drive activity. The blink patterns the same as for existing SAS drives, and may have vendor defined blink patterns. The system use of the Activity signal is optional.

Implementation Note: *An overview is the Activity signal is asserted when drive is active, but blinks off "occasionally" during periods of continuous activity to distinguish it from other LED states. The Activity signal is called "Ready" in the SAS documentation and the definition is in the SAS Protocol Layer - 2 (SPL-2) Specification* (T10, SPL 2, 2011) *in* Section 9.4.1 SAS Ready signal behavior.

3 Connector Mechanical

Note: this section with the connector mechanical drawing is informative. The normative specification is SFF-8639(SFF-8639, 2012).

Figure 7 gives an overview of the mechanical layout of the connector. The new signals are shown in red. Figure 8 gives an isometric view of representative examples of the connectors from multiple angles.

Figure 7: Connector Mechanical Overview

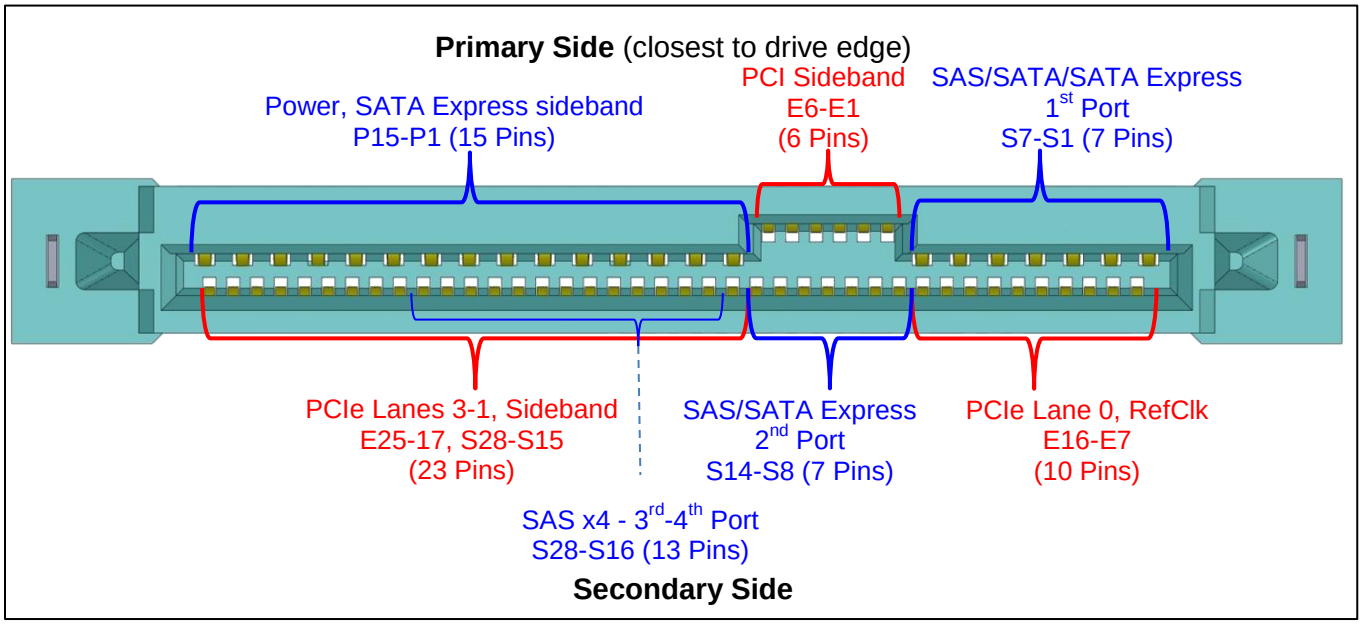
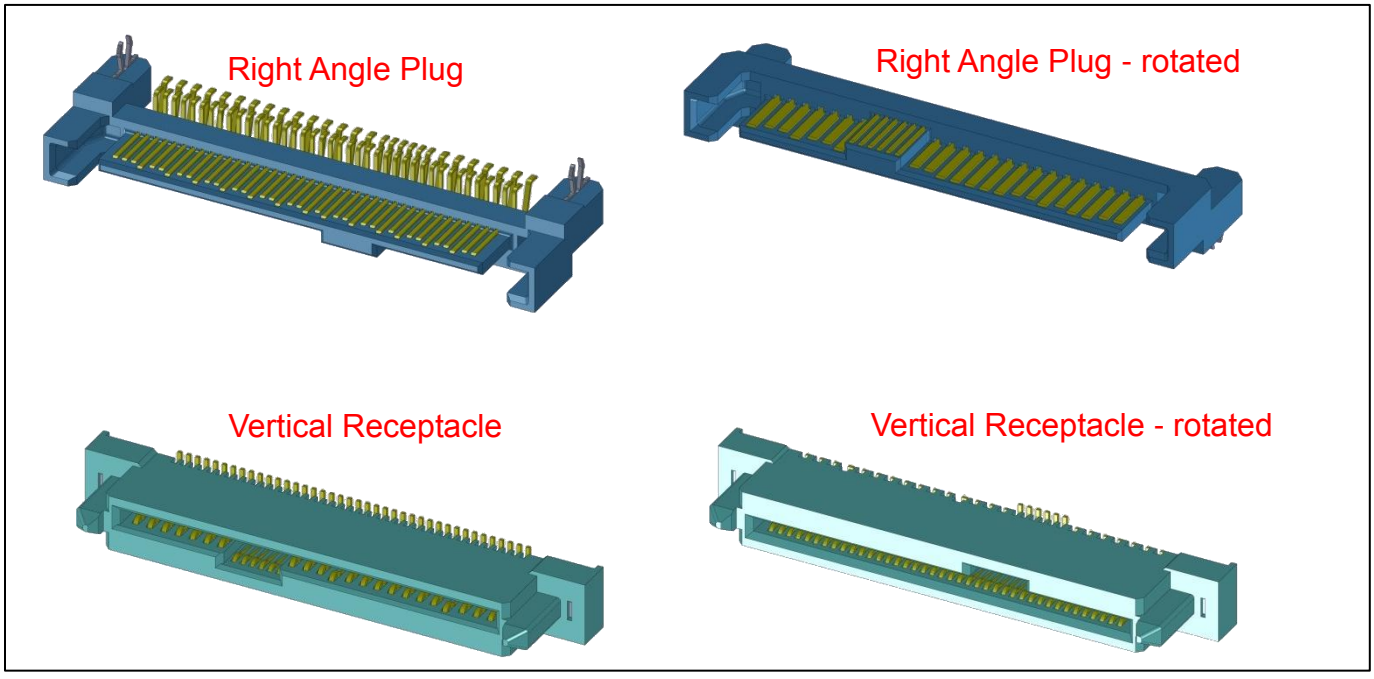


Figure 8: Connector Examples



3.1 Connector Mechanical Drawing notes

The mechanical drawing of the connector plug (drive side) are shown in Figure 9 though Figure 17. The mechanical drawings of the connector receptacle are shown in Figure 18 and Figure 19.

3.2 Connector Plug Mechanical Drawings

Figure 9: Connector Plug Mechanical Drawing – Top View

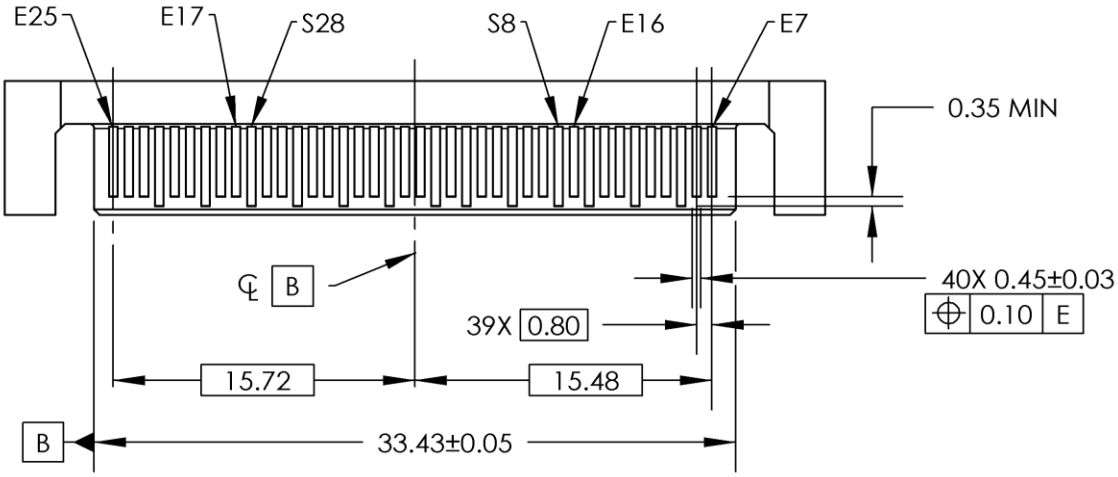


Figure 10: Connector Plug Mechanical Drawing – Drive Insertion View

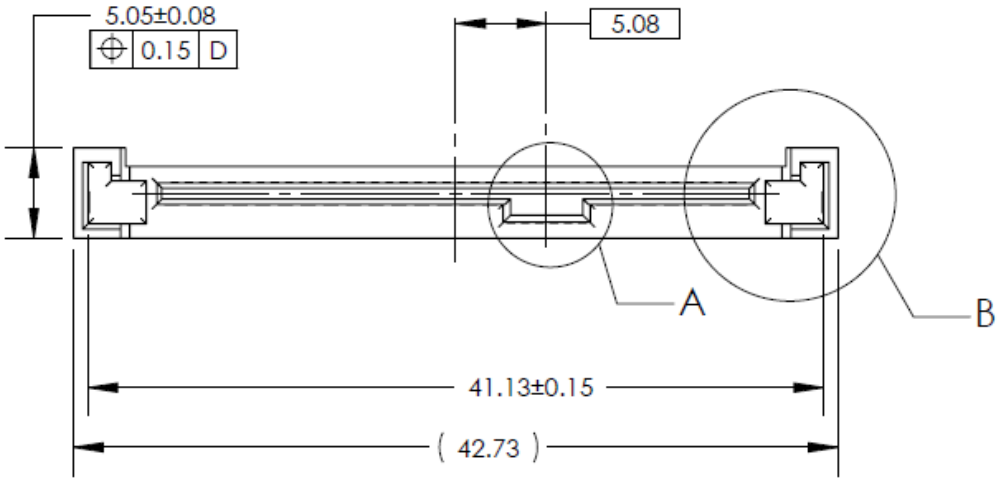


Figure 11: Connector Plug Mechanical Drawing – Bottom View (outside of enclosure)

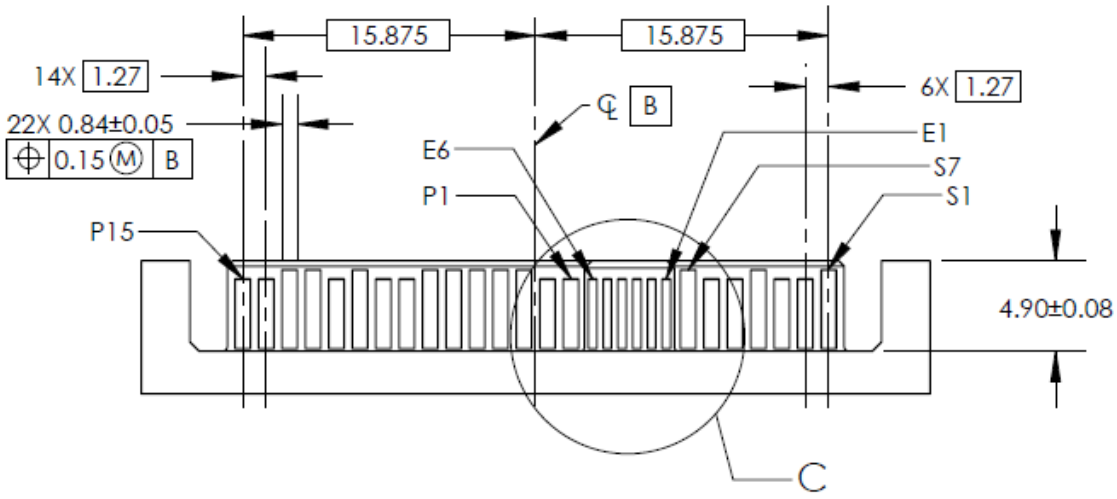
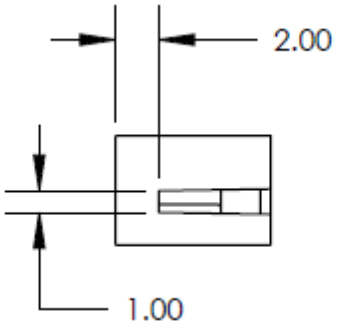


Figure 12: Connector Plug Mechanical Drawing – End View (showing latch slot)



3.3 Connector Plug Mechanical Drawings - Details

Figure 13: Connector Plug Mechanical Drawing – Detail A (key)

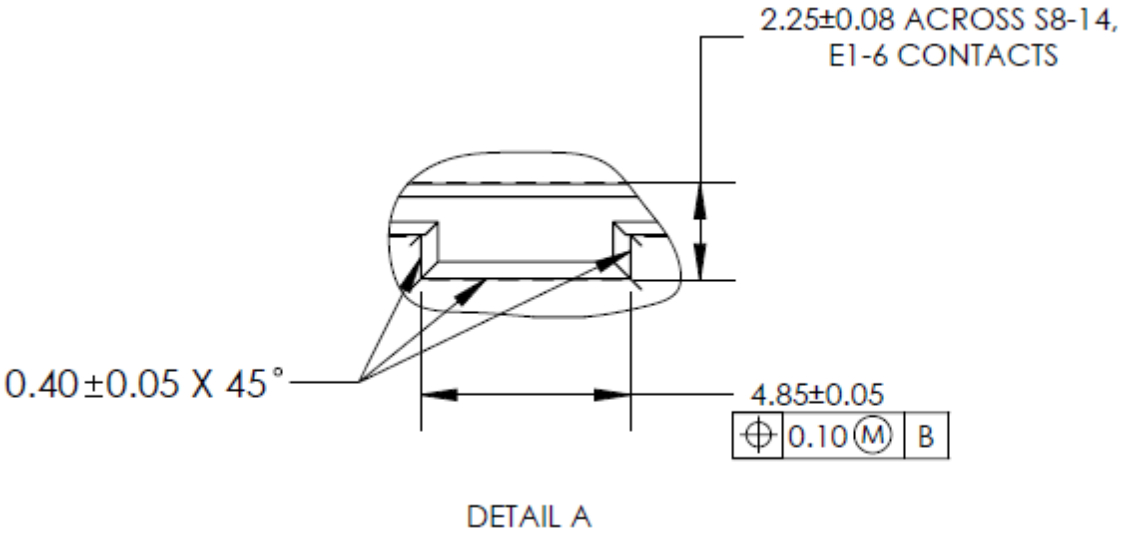


Figure 14: Connector Plug Mechanical Drawing – Detail C (key pins)

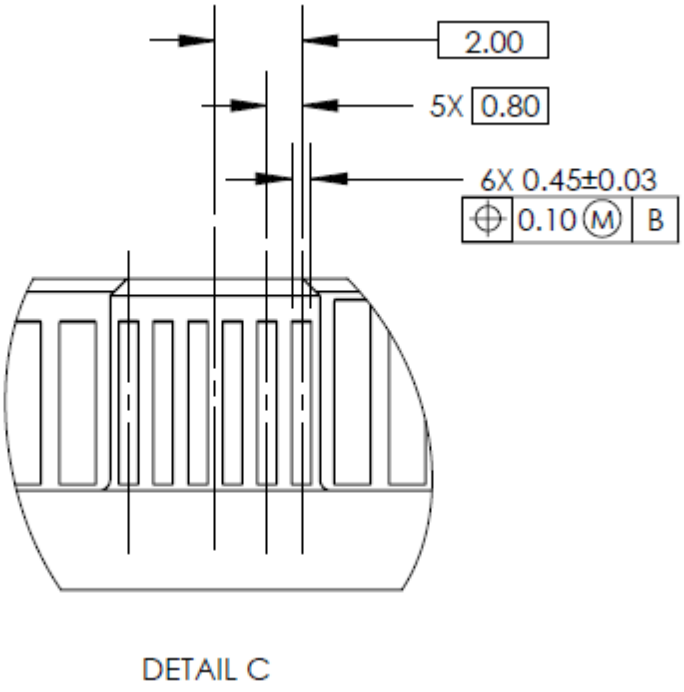


Figure 15: Connector Plug Mechanical Drawing – Detail B (Tongue)

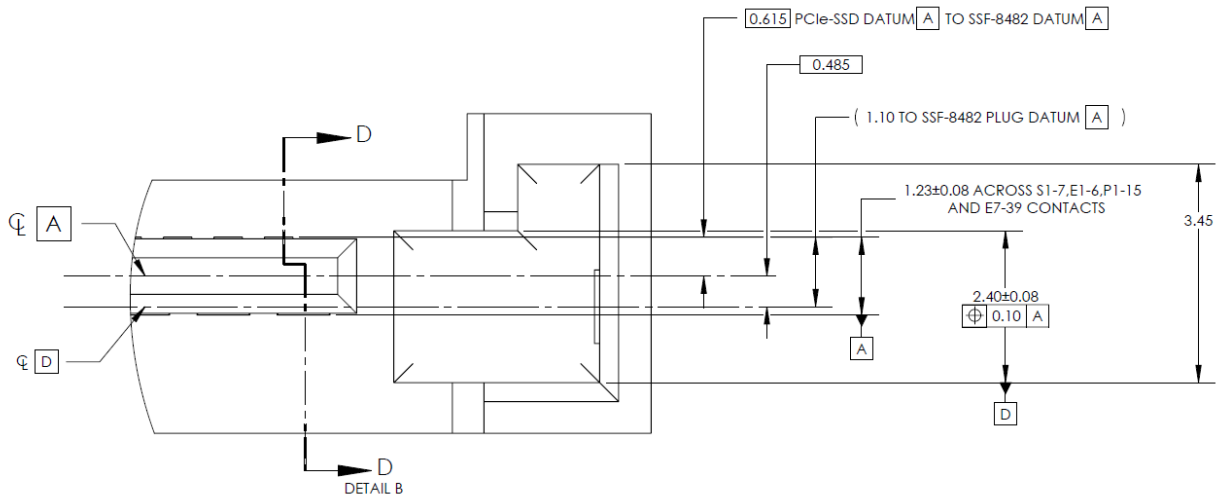


Figure 16: Connector Plug Mechanical Drawing – Section D-D (Tongue cross section)

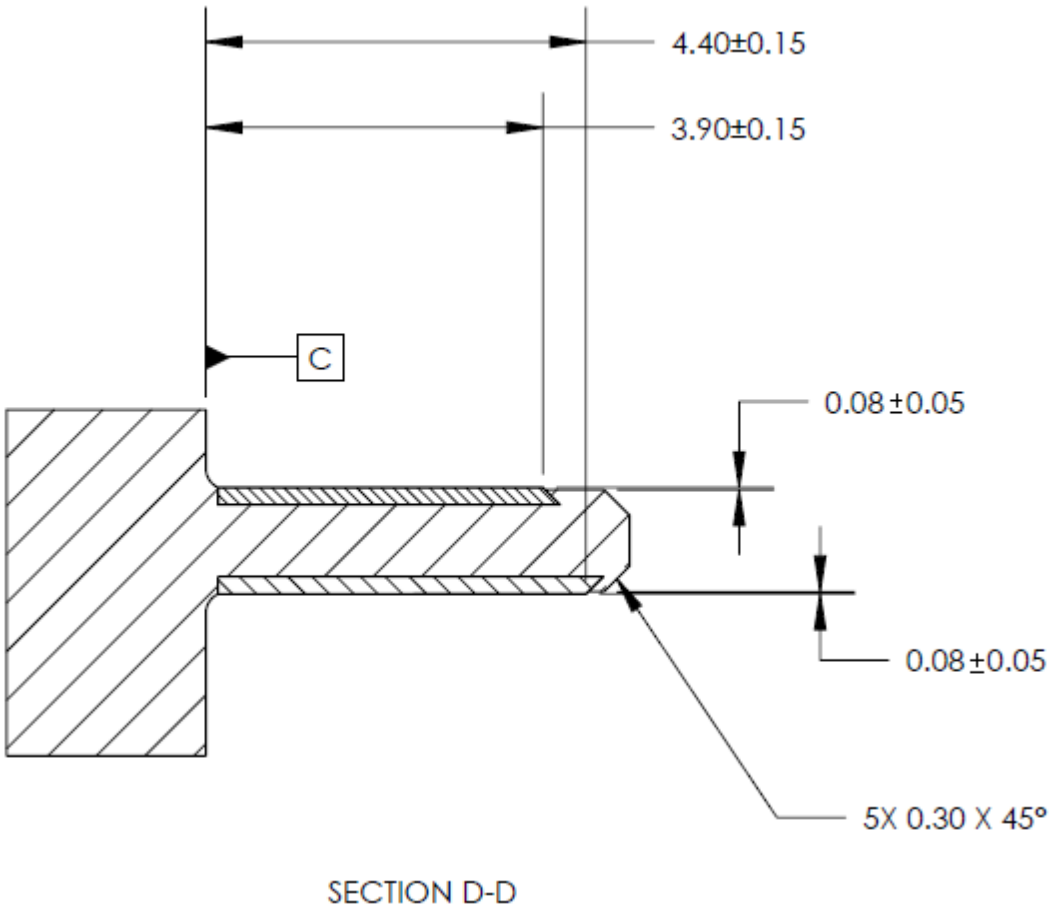
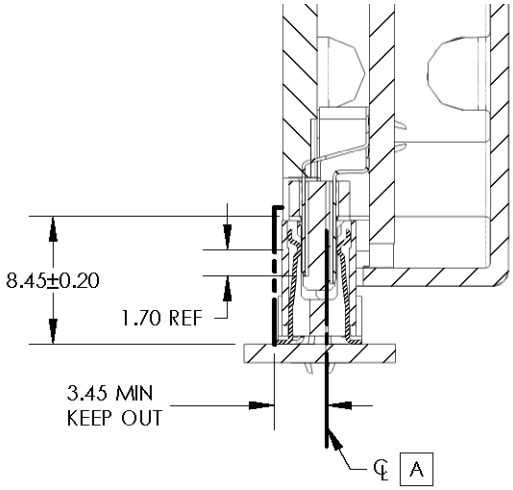


Figure 17: Connector Plug Mechanical Drawing – Drive Keep-out Area



3.4 Connector Receptacle Mechanical Drawings

Figure 18: Connector Receptacle Mechanical Drawing – Drive insertion view

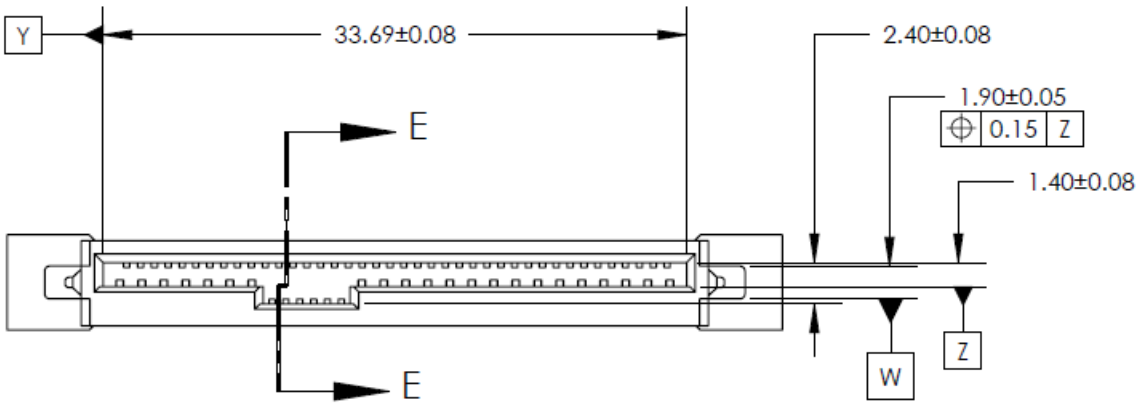
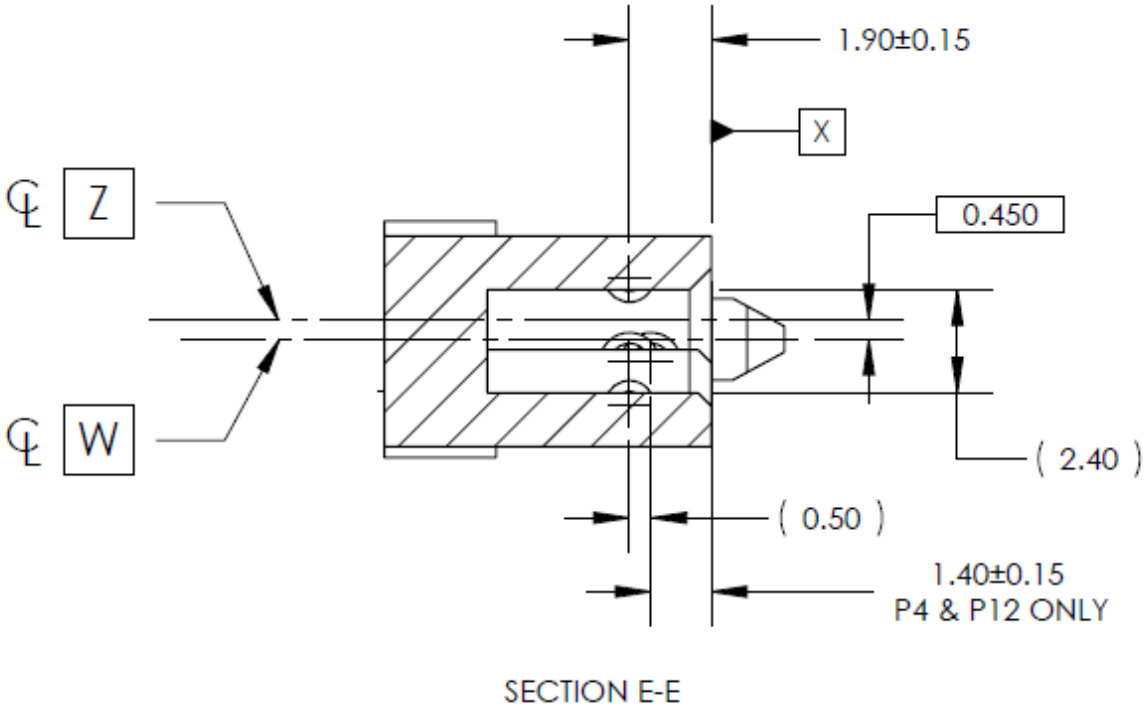


Figure 19: Connector Receptacle Mechanical Drawing – Section E-E (Receptacle)



3.5 Cable Retention Mechanical Drawings

The primary server and storage Enterprise PCIe SSD connection is a backplane with retention provided by the drive carrier and storage enclosure. This is typically a universal driver carrier with insertion guides and insertion/ejection lever and any required keying. The definition of the drive carrier is outside the scope of this specification. The backplane receptacle connector is intended to be universal accepting all the standard drive keys, relying on more mechanically robust storage enclosure keying.

A cable receptacle is defined for applications needing a cabled interface to the drive. The Enterprise PCIe SSD cable (receptacle) interface incorporates an active latching retention mechanism to prevent accidental disconnecting of the interface. The cable retention is a snapping latch when mated and a press-to-release finger actuated button on the sides of the housing. This is different from existing SATA and SAS connectors which use a friction detent on the secondary side of the tongue. An overview of the Enterprise PCIe SSD cable retention is shown in Figure 20.

The Enterprise PCIe SSD cable (receptacle) has an extra keying feature to allow the Enterprise Cable to be blocked from mating with SATA, SATA Express, and SAS drives. The “L” shaped key **shall** be implemented on the Enterprise PCIe SSD cable receptacle. The “L” shaped key should not be implemented on the Enterprise PCIe SSD backplane receptacle. A keying compatibility matrix is shown in Section 10.3.

Figure 21 shows the mechanical drawing of the cable retention features that mate with the plug feature shown in Figure 12. The latch slot **shall** be implemented on Enterprise PCIe SSDs plugs. (The drive latch slot should not be implemented in SATA, SATA Express, and SAS drive plugs.) The latch **shall** be implemented on Enterprise PCIe SSD cables. The size cable housing and location of finger release are not specified, and these drawing are just informative. The clearance for a representative cable housing is shown in Figure 22. A representative cable assembly with flat ribbon cables is shown in Figure 23.

Implementation note: An existing SAS cable would accept an Enterprise PCIe SSD drive but would not function (no signal connections) and would have not engaged the SAS cable detent retention mechanism.

Figure 20: Cable Retention: Latch and Key Overview

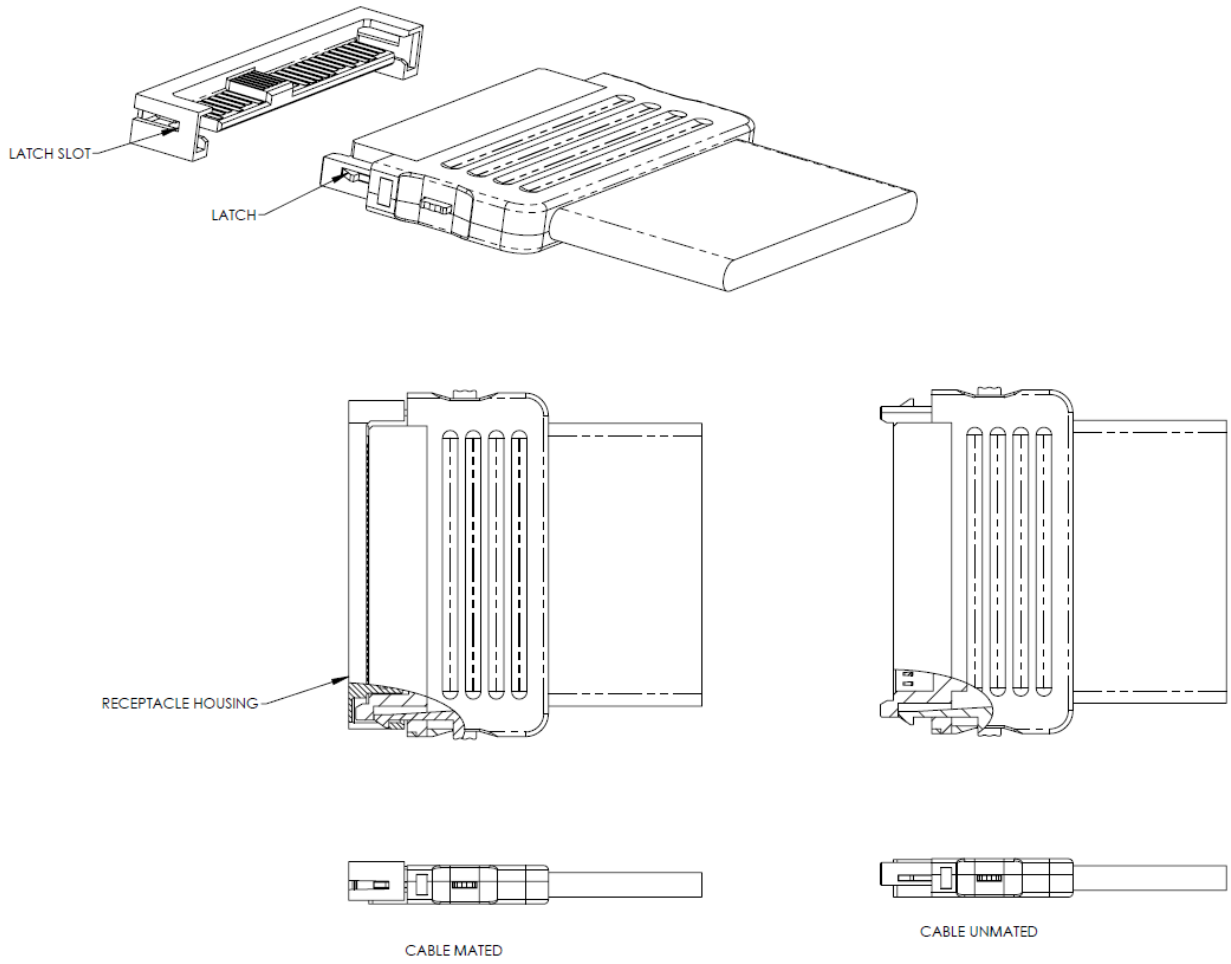


Figure 21: Cable Retention Mechanical Drawing – Representative Cable Housing

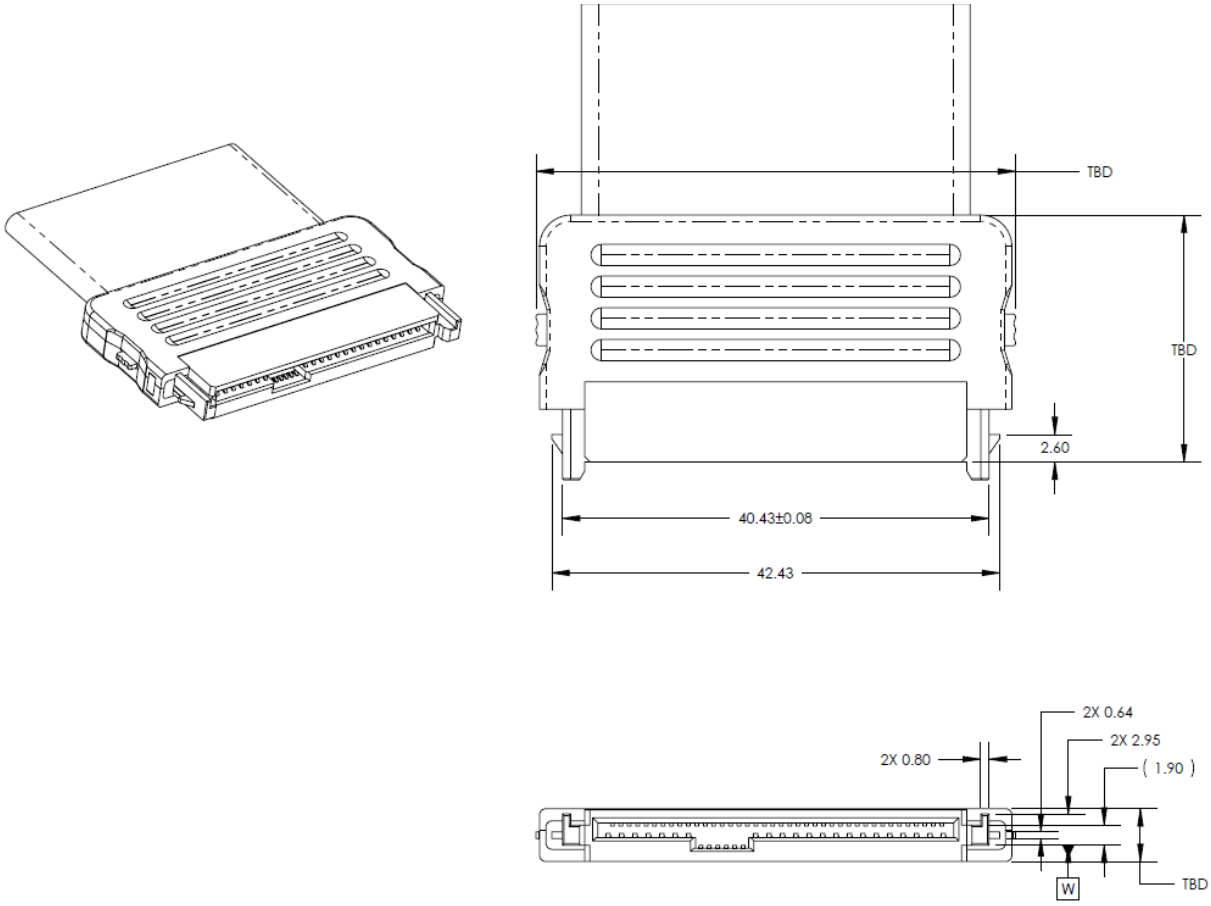


Figure 22: Cable Release Clearance - Representative Cable Housing

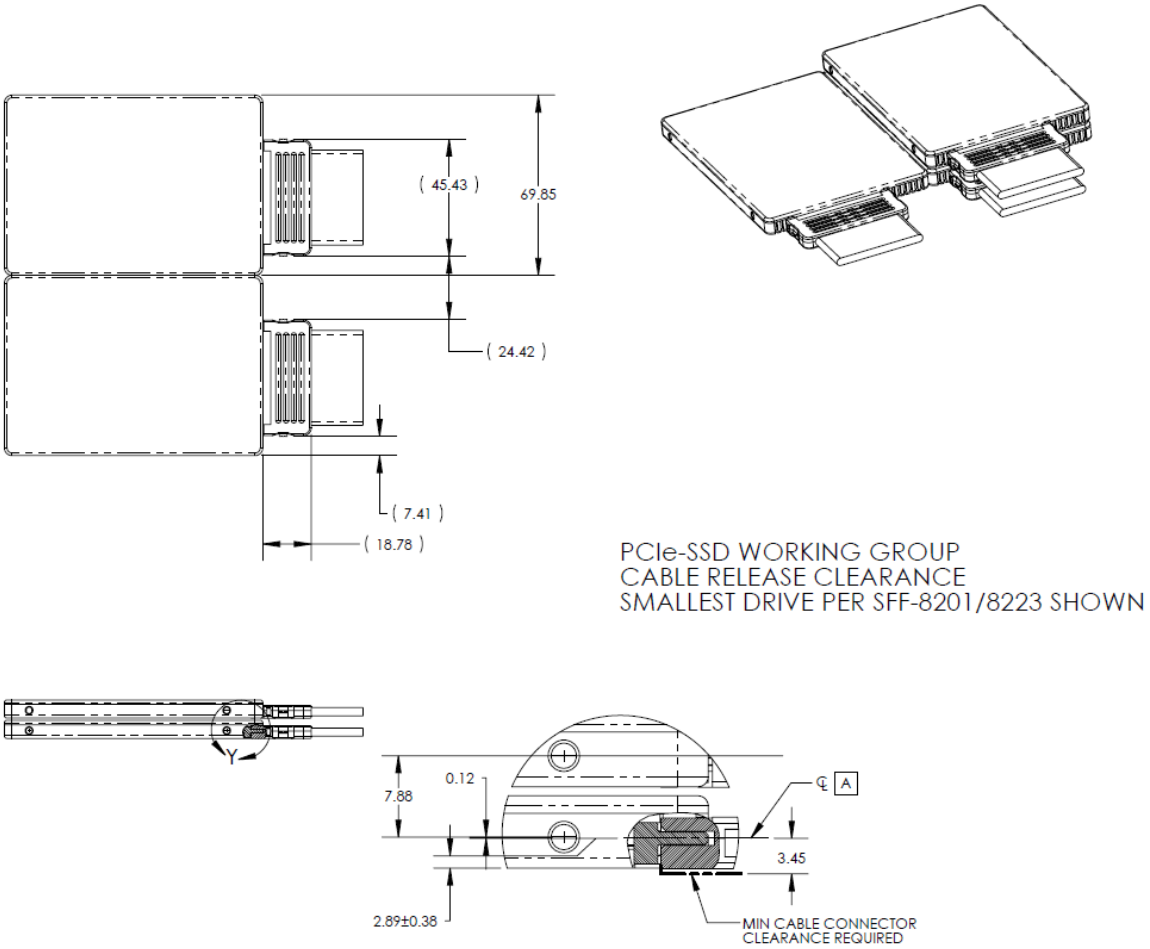
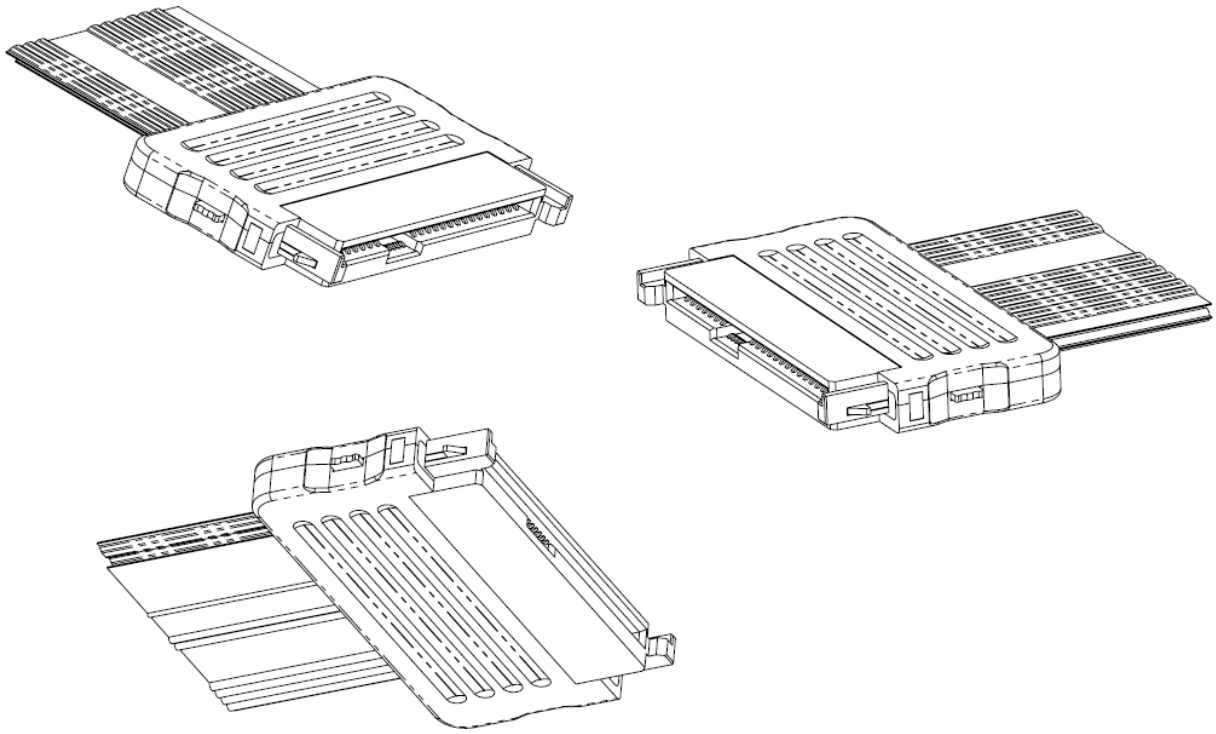


Figure 23: Representative Cable Housing with flat ribbon cables



4 Hot Plug and Removal

Enterprise PCIe SSDs **shall** support hot plug and removal, both with and without prior system notification (surprise removal).

The fact that this is a physical hard drive form factor suggests that some characteristics of a hard drive usage model will be supported. In today's servers hard drives are hot pluggable with some limitations. The Enterprise SSD drive shall be hot pluggable for both hot insertion and hot removal with and without prior system notification or preparation. Without prior system notification is typically referred to as "surprise insertion" and "surprise removal".

Whether the user provides the system with notification of the hot plug event or not, many aspects of the system are involved in processing such an event. Elements of the overall system that are impacted are dependent on platform architecture and design but will typically include components of the PCIe infrastructure such as switches, platform firmware (BIOS), the Operating System's PCIe bus driver stack and its IO stack. However, typically, a user would announce a hot plug event through one of the available system management interfaces.

The goal of this specification is not to define an overall systems architecture and design to support hot plug for the Enterprise PCIe SSDs but to define those elements of the hardware infrastructure that can be utilized to support such a solution.

The objectives of the hot plug design are:

- To provide a user experience similar to existing disk drive hot plug behavior.
- Support all forms of hot insertion/removal
  - OS aware insertion/removal
  - Surprise insertion/removal
    - Insertion
      - Insertion into a slot that has power and clocks on
    - Removal
      - Removal from a slot that has power and clocks on
      - Removal with TLPs in flight and outstanding host non-posted requests
  - SAS based hot plug signaling or PCIe native hot-plug
- No special controller requirements beyond those already defined in the PCIe base and in the PCIe Card Electromechanical specs
  - Allows standard PCIe controllers to be used in Enterprise PCIe SSD without modification

The issues to be addressed are:

1. Limiting of inrush current.
2. Detection of Enterprise PCIe SSD addition and removal and Determination of Drive Type – PCIe vs. SAS
3. Drive management – reset
4. SW implications

System support for Enterprise PCIe SSD hot plug and removal as well as the level of support (i.e., OS aware insertion/removal vs. OS and surprise insertion/removal) is implementation specific and outside the scope of this specification.

Implementation Note: *Surprise Hot Removal is a must have requirement from system vendors for PCIe SSD devices. More specifically, it is expected that system vendors will require Enterprise PCIe SSD to retain all "committed data" following a surprise hot removal. This implies that the Enterprise PCIe SSD locally detect hot removal and store/destage any cache data. For a typical implementation this means the following is implemented*

- *Having power sustaining capability, such as a battery or super cap.*
- *Detecting removal (by loss of 12V input or link loss)*

- *Copying all "committed data" from volatile storage (DRAM based write back cache) to non-volatile storage. The details of this operation are complex, and implementation specific. The definition of "committed data" is specific to the higher level interface but typically means acknowledged write data. This "committed data" includes all indirection pointers to the data.*
- *This copying of data has to be recovered upon PCIe SSD power on or hot addition. It is common for data or indirection tables to be written in to a staging location to simplify power down. This has to be reconstructed on power on/hot add to recover the full committed PCIe SSD state.*

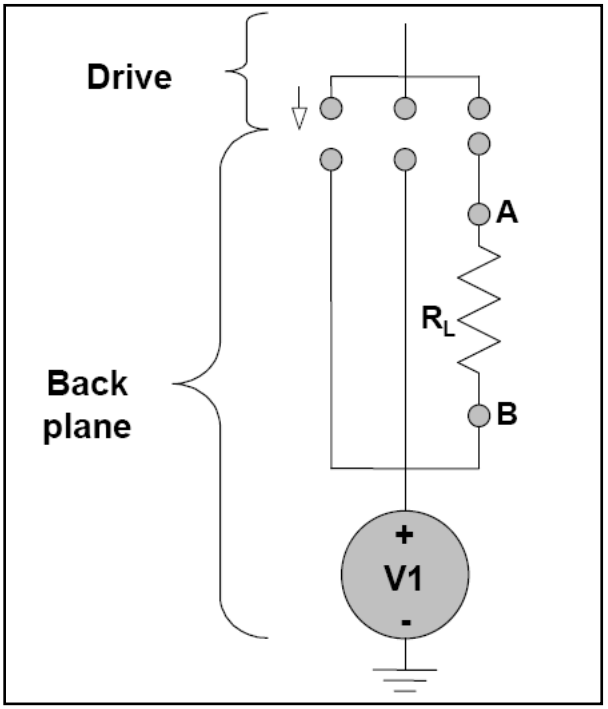
From a practical perspective for the Enterprise PCIe SSD, surprise hot removal is not that different from surprise power fail. In system power fail 12V likely has a longer power down ramp, but that is not guaranteed.

4.1 Limit current inrush

The Enterprise PCIe SSD shall limit the current inrush by leveraging the pre-charge approach from SAS/SATA and combining this with the slot power limit approach from PCIe. The Enterprise PCIe SSD shall:

a. Bus together all power and ground pins.
b. Limit the exposed capacitive load as specified in Section 7.4.
c. Adhere to the mechanical pin length as shown in Figure 28
   - 1 ground pin mate first (pin P12)
   - 12V precharge pin mates second (pin P13). In addition all the signal ground returns are second mate.
   - Remaining 12V pins mate third (Pin P14 and P15)
d. Initialize and enforce the Slot Power Limit PCIe configuration register to 10W value (Section 7.4). This eliminates the need for the system power supply to handle worse case power at start up.

Figure 24: Hot Plug Current Inrush Limiting



Implementation Note: *Limiting the inrush and hot plug support require merging information from a number of sources. SFF-8639* (SFF-8639, 2012) *has the mechanical specifications for blind mate tolerance and the 0.5mm difference in mating pin length. The SATA Specification* (SATA-IO, various) *gives a typical existing system insertion rates of 3ms delay from the 1st precharge 12V power to*

remaining power pins contacting. Given these parameters and the 12Vcap specification in Section 7.4, the backplane resistance (RL in Figure 24*) can be determined.*

4.2 Detecting Enterprise PCIe SSD Insertion/Removal

There are multiple methods to detect PCIe SSD insertion or removal.

- Use the drive type pins (PRSNT# pin P10, and IfDet#, pin P4). Note drive type can be determined as described in Section 2.3**.**
- Use of In-Band signaling. The PCIe or SAS links may be enabled to train periodically, succeeding when device is plugged in. The PCIe or SAS links are lost when device is removed.

How system software is signaled is beyond the scope of this specification.

Implementation Note: *There are three broad classes of usage of Presence (PRSNT#) and SATA Detect (IfDet#) during hot insertion or removal.*

1. *A common implementation is expected to be that PRSNT# and IfDet# are routed to the existing storage enclosure microcontroller and the microcontroller sends specific drive insertion or removal information to system software using the management stack. This is a common SAS implementation extended to two pins for 3 type encodings.*
2. *Alternatively the PRSNT# & IfDet# signals are simply turned into a general storage attention interrupt, enabling the system software to probe the system for changes. This can be augmented by examining if any PCIe link has changed status, or used as a trigger to enable Enterprise PCIe SSD to test for link training.*
3. *Optionally an existing storage enclosure microcontroller can poll and read the SM-bus to read vital product data (VPD) to learn specific information about the inserted device.*
4. *It is acceptable for the PRSNT# and IfDet# signals to be ignored completely by system hardware. Enterprise PCIe SSD insertion detection can be done in-band by detecting the training of the PCIe link. This may be complicated by PCIe links being disabled (powered down), so periodic re-arming may be required to probe for newly added Enterprise PCIe SSDs. In this case Enterprise PCIe SSD removal is detected by loss of PCIe link or PCIe link errors. From a timing perspective the PCIe link error/loss is the common 1st detection of surprise removal since it is tightly coupled with ongoing operation.*

4.3 PCIe Hardware and Reset System Requirements

Enterprise PCIe SSD Reset/power on and PCIe Hot Add follows the PCIe specification (PCI SIG CEM, 2007). Enterprise PCIe SSDs must generate their own internal power-on reset by detecting the power ramp on 12V.

In all cases the Root Complex and system software shall allow at least 1.0 second after a Conventional Reset of an Enterprise PCIe SSD, before determining that an Enterprise PCIe SSD which fails to return a Successful Completion status for a valid Configuration Request, is an unresponsive or failed Enterprise PCIe SSD. This 1.0 second is true for both PCIe and SMbus operations. This period is independent of how quickly Link training completes.

PCIe Reset Card Requirements

- Normal System Operation (not Hot Insertion/Removal)
  - An Enterprise PCIe SSD shall adhere to PCIe system reset and card reset requirements (Figure 25 and Figure 26)
- Hot Insertion
  - If reset asserted then an Enterprise PCIe SSD must adhere to PCIe system reset and slot reset requirements.
    - This also means that the system & root complex must adhere to these requirements
  - If reset negated
    - Power and clock shall be stable 10ms after assertion of IfDet# (Enterprise PCIe SSD's Presence Detect)
      - PRSNT# and IfDet# is in the last (3rd) group of pins to mate

        - This consumes 10ms of the 100ms power stable time in Figure 25 and Figure 26.
    - Enterprise PCIe SSD **shall** generate its own internal reset
    - Logically PCIe reset is considered to have occurred 100ms after PRSNT# pin mating
        - An Enterprise PCIe SSD must adhere to PCIe system reset requirements from this point on
        - This simply shifts timing parameters on Figure 26 by 100ms
            - This shift must also be comprehended by the platform if hot insertion occurs during a normal PCIe system reset sequence.
- Hot Removal
    - On a Hot Removal the Enterprise PCIe SSD shall perform any necessary clean-up and retain on the PCIe SSD any previously acknowledged write operations

Figure 25: System Reset Timing

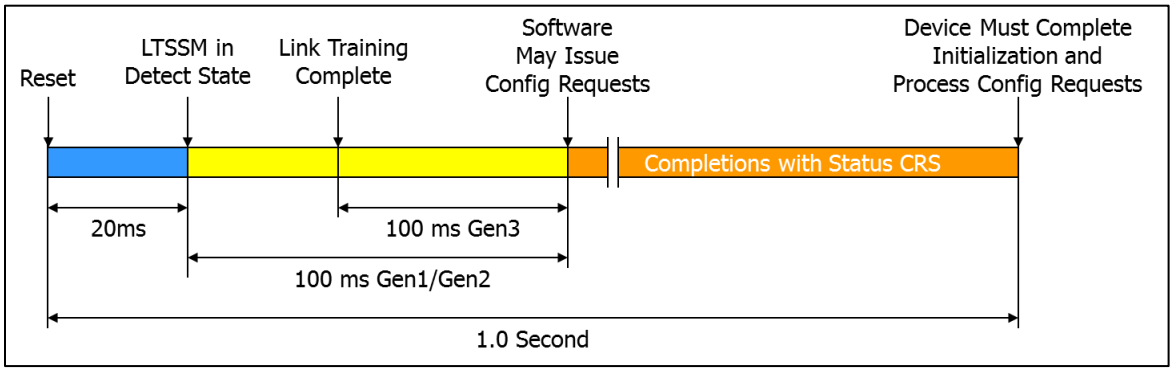
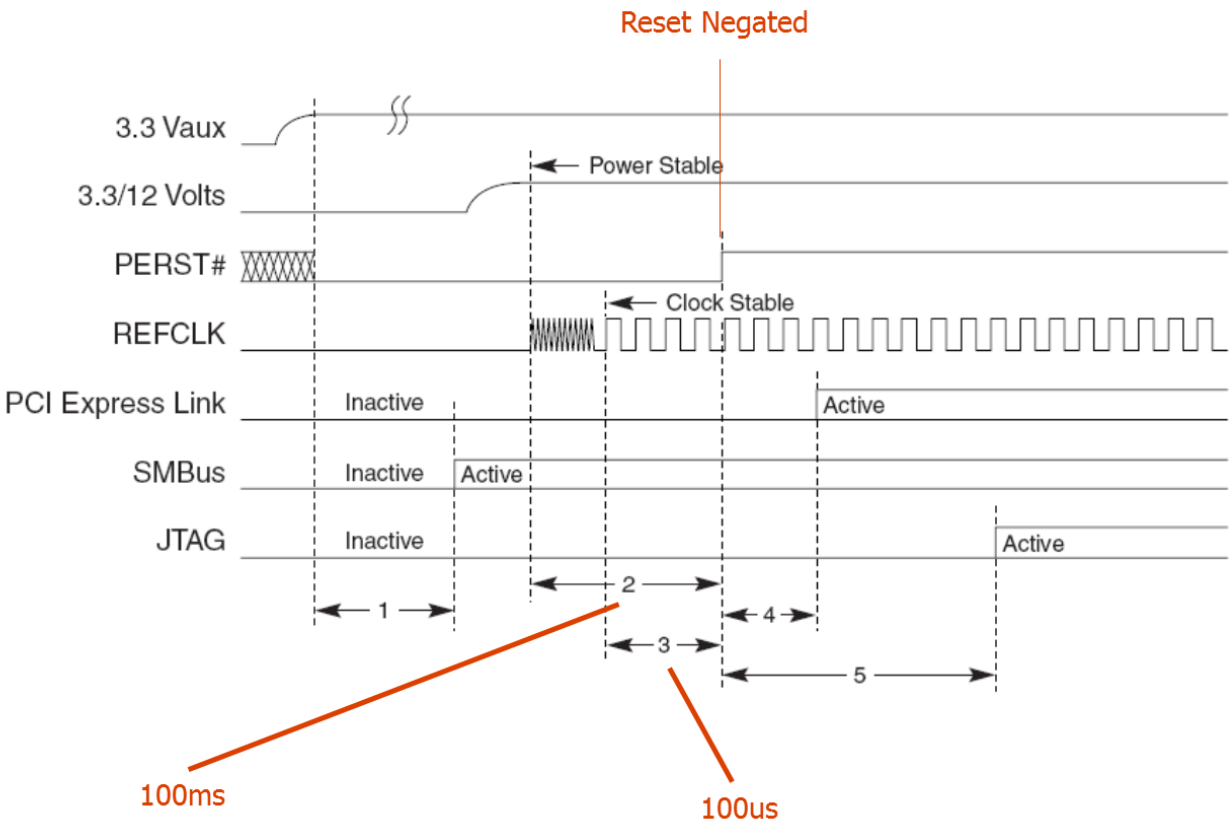


Figure 26: Reset Timing Details (PCI SIG CEM, 2007)



4.4 Software Implications of Hot Swap (Informative)

The software implementation is outside the scope of the specification; however this section provides some guidelines for OS and BIOS/Firmware implementers. Please refer to appropriate system BIOS/Firmware and system software (OS, VMM) for normative documentation.

The typical issues SW must handle on Hot Add include:

- Discover the Enterprise PCIe SSD location
  - Discovery may be assisted by management software or PRSNT# pin.
  - Software may have to enable a PCIe link (kick link out of power down mode)
- PCIe plug and play
  - PCIe configuration: map the new Enterprise PCIe SSD into the PCIe space, and do any space allocation and resource rebalancing.
  - Load the appropriate software drivers.
- Map the Enterprise PCIe SSD into the name space of the file system. This may include rebuilding any RAID arrays damaged by an earlier failure, or mapping to new on-the-fly RAID or tiering configuration.
- Finally start using the PCIe SSD

The types of issues SW must handle on Hot Remove include:

- Map the Enterprise PCIe SSD out of any RAID, tiering configuration – prepare stack for Enterprise PCIe SSD removal.
- Stop using the Enterprise PCIe SSD, send commands to “prepare for removal”
  - Typically this causes any write cache to flush, and any mapping table to be written back

    - If Enterprise PCIe SSD is not notified before removal, then it is surprise removal, see next section.
- Optionally use management system to identify the Enterprise PCIe SSD to remove. For example blink a red light for failure indication.
- User safely removes the Enterprise PCIe SSD.

If the removal is “surprise” meaning that the Enterprise PCIe SSD is active as it is being removed then the system software implications grow. These include, but are not limited to, the handling of errors during surprise removal event.

The software implications vary with the PCIe root port implementation and OS/SW versions. The software implications are beyond the scope of this specification.

***Implementation Note**: A complication of Hot Removal is discovering and testing all the boundary cases of the hardware/software interaction. In general a clean HW/SW interface and a clean software implementation are warranted – for example not a legacy interrupt model. The basic mechanism for an error detected by an instruction (ld/st) failure is to fault then handle the fault with by aborting the transfer.*

It is expected that any data in flight at the time of removal is lost or corrupted. Hardware is only expected to retain “acknowledged and committed data”.

***Implementation Note**: It is expected that older versions of the operating system and PCIe root port will not be able to support surprise Hot Removal. With older OS/HW the errors during surprise removal might generate a fatal error.*

New versions of the PCIe root port will be design to signal to the system software (BIOS and OS) to handle the hot removal event. The details are beyond the scope of this specification. The documents are not yet released. This detail is expected in documents such as the System BIOS Writers Guide for the root port component.

***Implementation Note**: Windows client OS is surprise hot remove capable because it was designed to support ExpressCard form factor. The Slot Capabilities Register contains a bit called Hot Plug Surprise which is a form factor specific bit. Express Card was expected to set this bit to indicate to Windows that card could be removed without notification and software stack should handle this without failure or data loss. For example, any writes to media behind such a card should never be cached by OS.*

Since the Server OS shares a common code base with the Client OS at the kernel level, the basic capability exists in Windows Server but is not validated.

An outline of how Enterprise PCIe SSD Hot Swap will work:

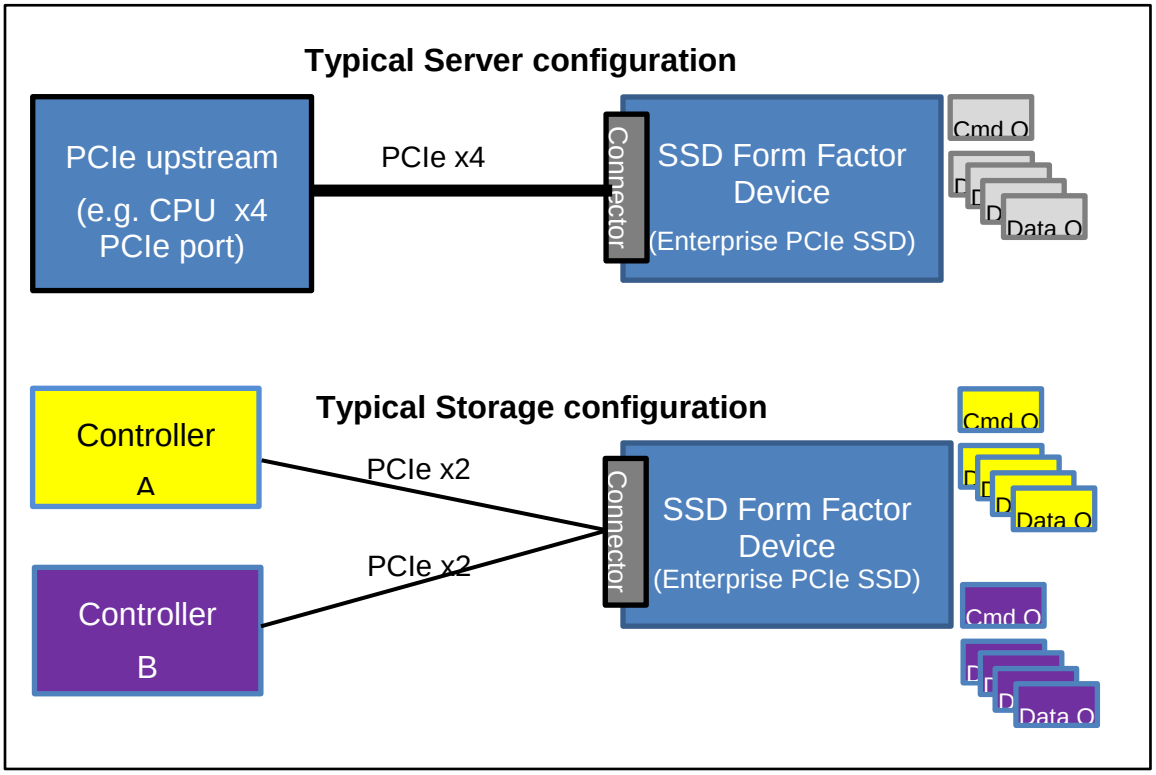
- *The OS gets signaled that an Enterprise PCIe SSD is inserted. (Using PCIe Hot Plug mechanism, a custom OEM leveraged mechanism, or detecting a change in PCIe Link Status.)*
- *OS enables read caching but not software write caching since the Enterprise PCIe SSD is surprise remove capable (if this impacts performance, then IT has to ensure they go through software UI to prep for ejection first).*
- *When the Enterprise PCIe SSD is removed cleanly, the error handling will abort any active transfers, then the presence detect toggle will indicate the hot remove event to OS*
- *OS will unload the corresponding driver stack and complete all pending IO request packets.*

5 Dual Port Operation

The Enterprise PCIe SSD **may** optionally support Dual port operation. The expectation is that servers and other single controller systems will use Enterprise PCIe SSDs in a single x4 configuration. High Reliability Storage Controllers with redundant controllers will use two x2 links. In dual port mode the data on the Enterprise PCIe SSD is accessible from either controller allowing for failure of a controller. In dual port modes the Enterprise PCIe SSD acts independently as a PCIe device on each link. This is shown in Figure 27: Single Port and Dual Port example.

Implementation Note: *Dual port operation is optional but is expected to be a required feature by storage vendors*

Figure 27: Single Port and Dual Port example



.

The two PCIe links must operate independently. It is expected that both ports are fully PCIe compliant devices and as such each port presents a fully a PCIe compliant endpoint.

It is possible that the PCIe device is a multi-root PCIe switch, but that is not expected to be the common solution. In this solution resets and failures (such as errors and clock failures) on one port must not affect operation on the other port.

Implementation Note*: In the most common dual port usage the Enterprise PCIe SSD appears as a totally independent Enterprise PCIe SSD on each bus sharing media. In Figure 27 both Controller A and Controller B see standard PCIe devices. There is not expected to be a PCIe-probe SW discoverable connection to the other port. There are two completely independent interfaces – such as two independent sets of PCIe configuration space, memory mapped control registers and queues. A higher level mechanism (such as reporting the same GUID on both links, or inherent knowledge of the topology) allow high level storage software to detect that the two Enterprise PCIe SSD controllers are sharing a common pool of storage and the higher level software coordinates the allocation of resources.*

Implementation Note: *All Enterprise PCIe Devices have a single supply and single ground return plane. For signal integrity there* ***MUST*** *be a single ground plane. Any redundant supply is handled outside the Enterprise PCIe SSD.*

This specification places no access restriction or interlocks on the usage of the two ports. Specifically writing independently to the same storage block from both ports the ordering is not guaranteed. Higher level software or system architecture is assumed to prevent any unsafe overlapping commands.

5.1 Dual Port Mode (implications of DualPortEn#)

The Enterprise PCIe SSD may be configured to train as either a single x4 port (port A) or dual x2 ports (i.e., ports A and B). When DualPortEn# (pin E25) is asserted (typically grounded by storage backplane), then the Enterprise PCIe SSD **shall** be configured for dual port operation. DualPortEn# is left unconnected or undriven by the system for single port operation. The Enterprise PCIe SSD **shall** pull DualPortEn# high (deasserted) if left unconnected by the system.

When configured for single port operation, the following rules apply:

- Enterprise PCIe SSD **shall not** train as two independent links.
- Enterprise PCIe SSD **may** train as a single x4, x2 or x1 link. Support for x2 is optional for both the system and the PCIe SSD.
- Enterprise PCIe SSD **shall** use ePERst0# to control the PCIe interface as defined by PCIe specifications (PCI SIG CEM, 2007).
- Enterprise PCIe SSD **may** use RefClk0 (if operating in common clock mode)
- Enterprise PCIe SSD **shall** ignore dual port signals, RefClk1, ePERst1#.

When configured for dual port operation, the following rules apply:

- Port A **shall** use PCIe lanes 0 and 1. Port B **shall** use PCIe lanes 2 and 3.
- Enterprise PCIe SSD **shall not** train as a single x4 link.
- Enterprise PCIe SSD **shall** train as a dual independent x2 if both upstream controllers are active and the Enterprise PCIe SSD supports dual port.
- Either port **may** train only as x2 or x1.
- Enterprise PCIe SSD **may** train one port, both ports, or neither port.
- Each PCIe port **shall** operate completely independently
    - ePERst0# assertion/negation or loss of RefClk0 or errors affecting PCIe port A **shall** not affect the operation on the PCIe Port B.
    - ePERst1# assertion/negation or loss of RefCLk1 or errors affecting PCIe port B **shall** not affect the operation on the PCIe Port A.
- When an Enterprise PCIe SSD is not capable of dual port operation but is configured for dual port operation, the following rules apply:
    - Enterprise PCIe SSD **should** train on either port A or Port B in x2 or x1. This is to allow just a single upstream controller active. This requires the appropriate PCIe Clock and PCIe Reset to be used. Specifically if Port A is trained then RefClk0 and ePERst0# are used, and if Port B is trained then RefClk1 and ePERst1# are used. For a single port PCIe SSD support for Port B is optional, and support for Port A is required.
    - Enterprise PCIe SSD **shall** not drive or terminate the signals on the port that is not trained. Specifically: If Port A is trained, then Enterprise PCIe SSD shall not drive or terminate Port B signals; or if only Port B is trained, then PCIe shall not drive or terminate Port A signals.

Implementation Note: *For a single port device that is presented with DualPortEn# asserted, it is important to disable driving and terminating the unused pins because some upstream ports may detect the electrical drive or termination as an electrical presence and try to train the port which would create training errors. This capability can be implemented using standard PCIe IP block sitting on Port A, and using the force width capability to disable the unused lanes, or by a separate HW disable derived from DualPortEn# to disable lanes 2-3.*

Implementation Note: *There is a partial solution called Active/Passive mode. This is defined in Section 9.2.4.*

6 SM-Bus Operation

The Enterprise PCIe SSD **should** support SM-Bus operation. The SM-Bus protocol and signaling are defined in the SM-Bus Specification (SMBus, 2000). SM-Bus operation is at 100khz or less.

SM-Bus is defined as independent of the PCIe links. There is no defined relationship between SM-Bus accesses and accesses on the PCIe buses.

There are two level of SM-Bus access.

1. Single reads and writes to a PROM device holding VPD (Vital Product Data) used for Enterprise PCIe SSD discovery and power allocation. Supported on 3.3Vaux power and during normal operation. The register definitions are listed in **Section 9.1 SM-Bus Vital Product Data (VPD).** Each Enterprise PCIe SSD's VPD is expected to be independently addressed on a unique SM-Bus segment with VPD at a fixed address. SM-Bus ARP is not supported for VPD access. An Enterprise PCIe SSD may require 3.3Vaux to be supplied to access this VPD data.
2. Extended management and access using MCTP-over-SMbus and this mode is only supported during normal operation (e.g., when 12V is present). MTCP Commands are defined in **Section 9.2**

Implementation Note: *While SM-Bus is optional for Enterprise PCIe SSD operation, it is expected to be included in the Enterprise PCIe SSD purchase criteria by many system vendors.*

Implementation note: *This is specified to allow implementation with an I2C PROM connected to SM-Bus and to 3.3Vaux. It is allowed to have the VPD with 3.3Vaux applied but required to apply 12V to access optional extended management capabilities.*

Implementation Note: *The SM-Bus is not replicated for dual port operation. In dual controller configuration, this is envisioned as going to a shared management agent, using the multi-master characteristics of SM-Bus, or independent SM-Buses merged external to the Enterprise PCIe SSD.*

6.1 Simple Accesses to Vital Product Data (VPD)

Enterprise PCIe SSD **should** support simple Reads and Write to Vital Product Data. This is defined in **Section 9.1 SM-Bus Vital Product Data (VPD)** in summary it contains:

- Basic inventory information such as type and size of Enterprise PCIe SSD, manufacturer, date, revision, and GUID.
- Power management data such as power level and power modes
- Vendor specific data

Implementation note: *This is likely implemented using a SM-Bus PROM connected to SM-Bus and to 3.3Vaux.*

If a system vendor wants access to this data, then 3.3V **shall** be supplied to 3.3Vaux pins. If the system vendor does not want access to this data via SM-bus, then 3.3Vaux **may** be optional.

6.2 Extended SM-Bus management

An Enterprise PCIe SSD **may** support full MCTP device access over SMbus (MCTP-over-SMbus). This is only possible during normal operation (e.g. when 12V present).

The MCTP commands are beyond the scope of this specification.

Implementation Note:

The MCTP-over-SMbus is implementation dependent but is likely to support:

- *Access to VPD using MCTP semantics*
- *Access to Enterprise PCIe SSD control registers*
- *Access to underlying media (with protection mechanisms)*

MCTP-over-SMbus is likely to be nearly symmetric with MCTP-over-PCIe.

7 Electrical (pointer to other specifications)

The electrical characteristics of signals that are defined as part of other standards, particularly PCIe and SAS, are referenced, rather than defined in this specification. This specification does define 2 signals, and new absolute maximum values for the 12V power specification. The new signal DualPortEn# complies with the PCIe auxiliary signal parametric specifications. The signals at Pins P1-P3 (Reserved(ClkReq#/DevSLP), Reserved(Wake#/OBFF#), sPERst#) **shall** be 3.3V tolerant.

Figure 4: Signal List Table (defining specification) on page 11 defines which specification is normative for each signal. This section gives electrical details on signals that are listed as “this spec” in Figure 4.

Note: For PCIe signals electrical compliance will be specified in a new PCI SIG document that is being created by a PCI SIG electrical work group. A reference to that document will be included when it is released.

7.1 Electrical Channel

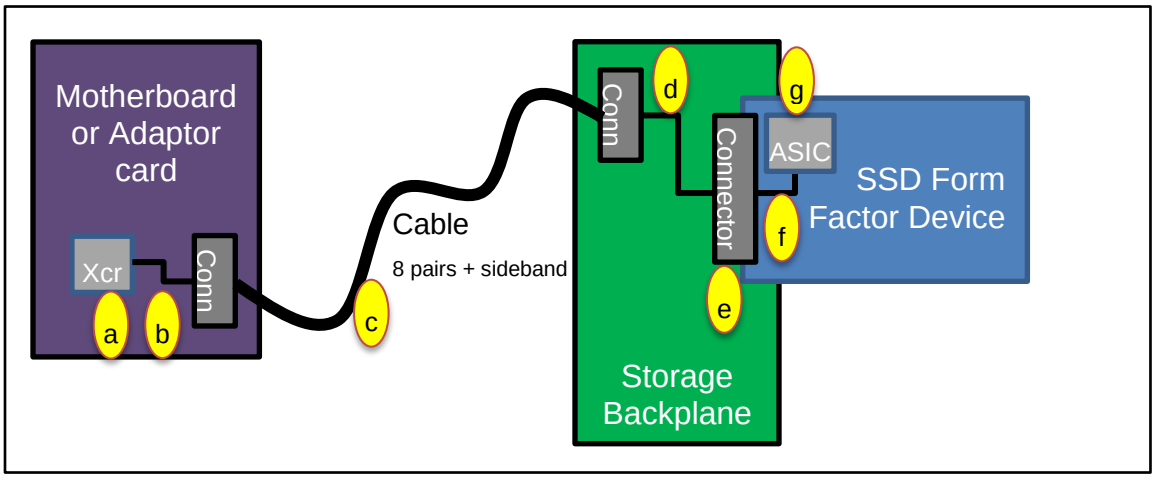
The high speed electrical characteristics of the connector are the same as the SAS Specification (T10, Various).

Implementation Note: *There is ongoing work to get to higher frequencies and that will change the channel definition, and possibly the connector requirements. That is outside the scope of this specification.*

Implementation Note: *The electrical channel used in simulation is modeled after the typical server disks connection. The typical channel is shown in Figure 28. Itemizing the channel from the upstream PCIe transceiver toward the downstream Enterprise PCIe SSD port (letters correlate to Figure 28).*

a. *Upstream PCIe port transceiver and package*
b. *Short (<3”) lead in trace between upstream PCIe device and cable connector.*
c. *Internal cable – assuming a cable similar to SFF-8087* (SFF, various). *The details are outside the scope of this specification.*
d. *Short (<4”) lead in trace between cable connector and Enterprise PCIe SSD connector*
e. *Enterprise PCIe SSD connector – specified in this specification*
f. *Short (<2”) trace inside Enterprise PCIe SSD*
g. *Downstream PCIe port package and transceiver.*

Figure 28: Typical SSD Form Factor Channel



7.2 Presence and Interface Detect

Presence and Interface Detect are used to signal the SSD has been inserted. This is the same purpose as SAS and SATA. The signal usage is given in Section 2.3 on page 14. In the drive the PRSNT# and IfDet# are connected to ground or left not-connected to signal drive type with a system supplying a pull up for signal detection.

7.3 ePERst[1:0]

The PCIe reset signal is logically identical as defined in the PCIe Specification (PCI SIG CEM, 2007). There is a difference in the input current.

Parameter	Definition	Value	Comment
I_{in}	Reset pin Input Current	1mA	10uA in PCIe speciation (PCI SIG CEM, 2007)

Implementation Note: This extended specification allows an Enterprise SSD drive to implement a circuitry, such as a pull up, to control drive reset when a drive is being inserted or removed.

7.4 Power – 12V

An Enterprise PCIe SSD(s) shall get all its power from the 12V pins (P13-P15). However, 5V (P7-P9) and 3.3V (P1-P3) may be available for other drive types. For SM-bus support see 3.3Vaux below. Note: The max current provided on 12V is greater than that specified in the SAS specification.

Figure 29: Power (12V) specifications

Parameter	Definition	Value	Comment
P_{max}	Maximum power	25W (max)	
Pinit-limit	Initial Slot Power Limit	10W (max)	Initial value until updated by system using a PCIe Slot Power Limit message[4]
$12V_{tol}$	Voltage Tolerance (at pin)		See Note 5
$12V_{amp}$	Max continuous current[1]	2.45A (max)	Higher than PCI CEM
$12V_{peak\text{-}amp}$	Max peak current[2]	4.5A (max)	
$12V_{cap}$	Max cap load[3]	5uF (max)	
$12V_{drop}$	Voltage drop across connector	80mV	
SM-Bus delay	Delay from 12V being within spec before SM-bus access to any SM-bus slave address other than VPD.	20ms (min) 1.0s (max)	The Enterprise PCIe SSD has up to 1.0s before responding to SM-bus transactions. (Section 4.3)
	Delay from PRSNT# connector mating to an SM-bus access to any SM-bus slave address other than VPD.	20ms (min) 1.0s (max)	

1. Maximum continuous current is defined as the highest averaged current value over any one second period.
2. Maximum current to limit connector damage and limit instantaneous power.
3. Maximum capacitance presented by the Enterprise PCIe SSD on the 12V power rail at the backplane connector.
4. *System Software can set a different lower limit using the PCIe Express Slot Capabilities register defined in the PCIe Specifications* (PCI SIG 3.0, 2010). *This value can be set before an*

Enterprise PCIe SSD is installed, and the value will be sent to the device just after link training and overrides this default. This is standard PCIe behavior.

5. No voltage tolerance is specified. Due diligence is need to confirm system and drive have compatible specifications.

***Implementation Note**: To accommodate all drive types possible with a flexible backplane using SFF-8639 connector, a 12V voltage tolerance of ±5% would conservatively support shipping SATA and SAS drives. The SATA and SAS standards do not specify 12V voltage tolerance, but many existing drives specify ±5% tolerance. The PCIe CEM* (PCI SIG CEM, 2007) *specifies ±9% on 12V in the standard PCIe card edge form factor. System vendors commonly specify a voltage tolerance in their purchasing specifications, commonly in the ±5-±8% range but not always symmetrical.*

Note that Rev 1.0 of this specification specified a 12V voltage tolerance of ±15%, but this changed to unspecified in Rev 1.0a. This is due to no single value being optimal for all flexible backplane configurations.

7.5 Power 3.3Vaux

Enterprise PCIe SSDs **may** support 3.3V operation. SM-bus access is supported only when 3.3Vaux is supplied.

The system **may** choose to not supply 3.3Vaux if limitation to SM-bus access is acceptable. The Enterprise PCIe SSD **shall** operate even with no 3.3Vaux supplied. Enterprise PCIe SSD operation shall not be dependent on 3.3Vaux except for SM-bus.

***Implementation Note:** Systems may connect 3.3VAux to either a standard 3.3V rail, or a special rail that is powered even when system is in a standby state with main 12V rail disabled.*

***Implementation Note:** A typical Enterprise PCIe SSD could connect a small PROM to SM-Bus and power this device with 3.3Vaux. This VPD is only accessible if 3.3Vaux is applied by the system. SM-Bus could also be routed to the controller ASIC and these SM-bus operations would be dependent on 12V power rail.*

Figure 30: Power 3.3VAux specifications

Parameter	Definition	Value	Comment
$3.3VAux_{tol}$	Voltage Tolerance (at pin)	±15% (max)	Relaxed from PCIe CEM
$3.3VAux_{amp}$	Max continuous current[1]	20µA (SM-Bus inactive) 1mA (SM-Bus active or 12V applied)	Lower than PCI CEM 12V applied so system not in S5-Soft Off (S5)
$3.3VAux_{cap}$	Max cap load[2]	5µF (max)	
SM-Bus delay	Delay from 3.3V being within spec before SM-Bus access to VPD serial EEPROM may be performed.	20ms (min) 1.0s (max)	The Enterprise PCIe SSD has up to 1.0s before responding to SM-bus transactions. (Section 4.3)
	Delay from PRSNT# & IfDet# connector mating to when SM-bus access to VPD serial EEPROM may be performed.	20ms (min) 1.0s (max)	

1. Maximum continuous current is defined as the highest averaged current value over any one second period.
2. Maximum capacitance presented by the PCIe SSD on the 12V power rail at the backplane connector.

8 References and Bibliography

8.1 References

PCI Express® Specifications are available from http://www.pcisig.com.

- *PCI Express Base Specification, revision 3.0* (PCI SIG 3.0, 2010).
- *PCI Express Base Specification, revision 2.1* (PCI SIG 2.1, 2009).
- *PCI Express® Card Electromechanical Specification, Revision 2.0.* April 11, 2007. Commonly known as the "PCIe CEM" spec (PCI SIG CEM, 2007).
- *PCI Express Engineering Change Notices,* (PCI SIG ECNs, various)

Small Form Factor (SFF) specifications are available from ftp://ftp.seagate.com/sff

- *SFF-8680 Specification for Unshielded Dual Port Serial Attachment Connector*. Commonly known as the "SAS drive connector" (SFF-8680, 2012). This is the new 12Gbps version of SFF-8482 which has been transferred to the ECIA and is now officially known as EIA-966 (SFF-8482, 2006).
- *SFF-8223 Specification for 2.5" Drive Form Factor with Serial Connector*. Commonly known as the "2.5" drive". (SFF-8223, 2006)
- *SFF-8301 Specification for 3.5" Drive Form Factor*. Commonly known as the "3.5" drive" (SFF-8301, 2010).

SSD HW/SW interface specification available from http://www.nvmexpress.org/ .

- *NVM Express (formerly eNVMHCI).* This is the HW/SW interface for a PCIe based controller Non-Volatile (Flash, PCM) memory. Commonly known as "NVMe" (NVM Express, 2011),

SOP/PQI HW/SW interface specification available from http://www.t10.org/drafts/ .

- *SCSI over PCIe (SOP)* and *PCIe Queuing Interface (PQI) .* This is an SCSI derived HW/SW interface for a PCIe based SSDs. Commonly known as "SOP/PQI" (T10, SOP, Draft 2012) (T10, PQI, 2012),

8.2 References Under Development

The following are documents that are under development, but are not current publically available.

- *STA SAS-**3** WG – SCSI Trade Association's Serial Attached SAS **3 generation** Working Group.* Defining a 12Gbps version for SAS disk standard. Commonly known as the "SAS-3" (T10, SAS-3, Draft 2012).
- *SATA Express*. Defining a next generation of the SATA standard that uses PCIe interface for high performance drives. Also referred to in this specification as "Client PCIe" (SATA-IO, Draft 2011).

8.3 Formal Bibliography

NVM Express. (2011, July 12). *NVM Express Revision 1.0b.* Retrieved December 13, 2011, from NVM Express: http://www.nvmexpress.org/

PCI SIG 2.1. (2009, March 4). *PCIe® Base Specification, Revision 2.1.* Retrieved from PCI SIG: http://www.pcisig.com/specifications/pciexpress/

PCI SIG 3.0. (2010, November 10). *PCIe® Base Specification Revision 3.0.* Retrieved from PCI SIG: http://www.pcisig.com/specifications/pciexpress/

PCI SIG CEM. (2007, April 11). *PCI Express® Card Electromechanical Specification Revision 2.0.* Retrieved from PCI SIG: http://www.pcisig.com/specifications/pciexpress/

PCI SIG ECNs. (various). *PCI Express Specifications & ECNs*. Retrieved from PCI SIG: http://www.pcisig.com/specifications/pciexpress/specifications/#ecn2

PCI SIG. (various). *PCI Express Specifications*. Retrieved from PCI-SIG: http://www.pcisig.com/specifications/pciexpress

SATA-IO. (2011, July). *Serial ATA Revision 3.1 Specification.* Retrieved from Serial ATA Internation Organization (SATA-IO): https://www.sata-io.org/developers/purchase_spec.asp

SATA-IO. (Draft 2011, Not yet published). *SATA Express.* Retrieved from Serial ATA International Organization: https://www.sata-io.org/technology/sataexpress.asp

SATA-IO. (various). *SATA and SATA-Express*. Retrieved from Serial ATA Internation Organization (SATA-IO): http://www.sata-io.org/

SFF. (various). *SFF Specifcations*. Retrieved from Small Form Factor Web Site: http://ftp.seagate.com/sff

SFF-8223. (2006, May). *2.5" Drive Form Factor with Serial Connector.* Retrieved from SFF site: ftp://ftp.seagate.com/pub/sff/SFF-8223.PDF

SFF-8301. (2010, March 16). *2.5" Drive Form Factor with Serial Connector.* Retrieved from SFF site: ftp://ftp.seagate.com/sff/SFF-8301.PDF

SFF-8482. (2006, February 1). *EIA-966, Serial Attachment 2X Unshielded Connector.* Retrieved from SFF site in lieu of EIA: ftp://ftp.seagate.com/pub/sff/SFF-8482.PDF

SFF-8630. (2012, June 15). *Serial Attachment 12 Gb/s 4X Unshielded Connector, rev 1.1.* Retrieved from SFF: ftp://ftp.seagate.com/sff/SFF-8630.PDF

SFF-8639. (2012, June). *Multifunction 12 Gbps 6X Unshielded Connector.* Retrieved from SFF: ftp://ftp.seagate.com/sff/SFF-8639.PDF

SFF-8680. (2012, July 11). *Serial Attachment 12 Gb/s 2x Unshielded Connector, rev 1.6.* Retrieved from SFF: ftp://ftp.seagate.com/sff/SFF-8680.PDF

SMBus. (2000, August 3). *System Management Bus (SMBus) Specification.* Retrieved from SM Bus: http://smbus.org/specs/smbus20.pdf

T10. (Various). *INCITIS, Technical Committee T10 - SCSI Storage Interfaces*. Retrieved from SCSI Storage Interfaces: http://www.t10.org

T10, PQI. (2012, 10 9). *2239-D, PCIe(r) architecture Queuing Interface (PQI).* Retrieved from T10 Drafts: http://www.t10.org/drafts.htm#SCSI3_PCI

T10, SAS 2.1. (2011). *ANSI/INCITS 478-2011, Serial Attached SCSI - 2.1 (SAS-2.1).* Retrieved from INCITIS: http://www.techstreet.com/standards/incits/478_2011

T10, SAS-3. (Draft 2012, 10 18). *T10/2212-D, Serial Attached SCSI - 3 (SAS-3).* Retrieved from T10: http://www.t10.org/drafts.htm#SCSI3_SAS

T10, SOP. (Draft 2012, 9 24). *2239-D, SCSI over PCIe(r) architecture (SOP).* Retrieved from T10 Drafts: http://www.t10.org/drafts.htm#SCSI3_PCI

T10, SPL 2. (2011). *ANSI/INCITS 476-2011, SAS Protocol Layer (SPL).* Retrieved from INCITIS: http://www.techstreet.com/standards/incits/476_2011

9 SM-Bus - Informative Appendix

9.1 SM-Bus Vital Product Data (VPD)

Vital Product Data (VPD is stored in an SM-Bus device with a slave address of 0xA6 (i.e., slave address bits 7-1 correspond to 1010_011 on the SM-Bus). The field definitions align with PCIe Specification field and capability definitions.

Implementation note: *This is likely implemented using a SM-Bus PROM connected to SM-Bus and to 3.3Vaux. This is analogous with DDR's SPD, using a small PROM to definition the capabilities, starting with VPD fields followed by a linked list of PROM fields that define what optional "Extended Capabilities" the Enterprise PCIe SSD supports and the define access details like addresses for extended registers.* .

Figure 31: SM Bus Vital Product Data

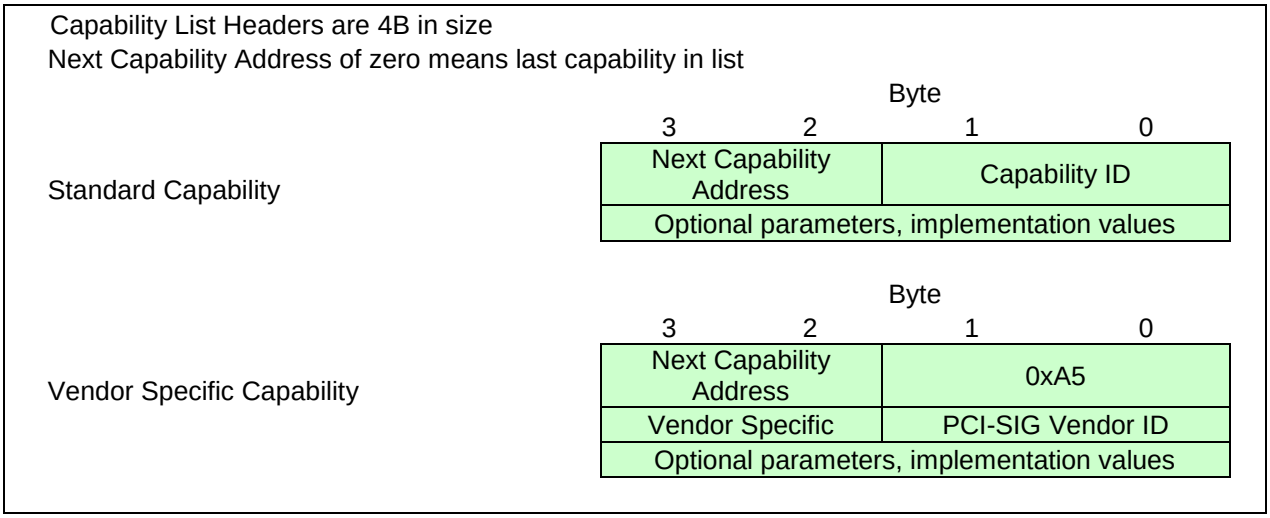
Addr	Function	Type	Size (B)	Default Value	Description	Controlling Specification
0	Class Code	RO	3	Vendor	Device type and programming interface	PCI Code and ID Assignment Specification Rev. 1.0 (PCI-SIG) - Class Code
3	ID	RO	2	Vendor	PCI-SIG Vendor ID	Enterprise NVMHCI Rev 1.0 - Identify Controller Data Structure
5		RO	20	Vendor	Serial Number (vendor unique)	
25		RO	40	Vendor	Model Number (ASCII string)	
65	PCIe Port 0 Capabilities	RO	1	Vendor	Maximum Link Speed	PCIe Base 3.0 (PCI-SIG) - Link Capabilities Register / Maximum Link Speed
66		RO	1	Vendor	Maximum Link Width	PCIe Base 3.0 (PCI-SIG) - Link Capabilities Register / Max Link Width
67	PCIe Port 1 Capabilities	RO	1	Vendor	Maximum Link Speed	PCIe Base 3.0 (PCI-SIG) - Link Capabilities Register / Maximum Link Speed
68		RO	1	Vendor	Maximum Link Width	PCIe Base 3.0 (PCI-SIG) - Link Capabilities Register / Max Link Width
69	Initial Power Requirements	RO	1	Vendor	12V power rail initial power requirement (W)	
70		RO	1	0	Reserved	
71		RO	1	0	Reserved	
72	Maximum Power Requirements	RO	1	Vendor	12V power rail maximum power requirement (W)	
73		RO	1	0	Reserved	
74		RO	1	0	Reserved	
75	Capability List Pointer	RO	2	Vendor	16-bit address pointer to start of capability list (zero means no capability list)	

9.2 SM-Bus Enterprise SSD Capability Definitions

The Drive Capability Definitions allow discovery of optional “extended” features or discovery of implementation variable parameters such as address. These structures in SMbus PROM are optional, but enumerating the features in this way simplify feature discovery and simplify driver development.

Figure 32 shows the format for Capability Definitions enumerating optional features of an Enterprise PCIe SSD. The first two bytes identify which capability or feature is implemented. The next two bytes are pointer to next capability definition (a linked list). A value of “zero” is the end of the linked list. For features that have parameters such as size or register definitions, these values follow, and the definition is specific for each capability.

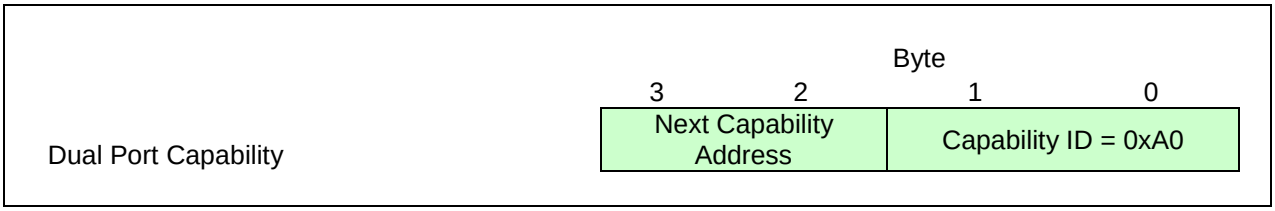
Figure 32: SM Bus Capability Definition



9.2.1 Dual Port Extended Capability.

Figure 33 shows Capability Definition for Dual Port.

Figure 33: Dual Port Extended Capability Definition



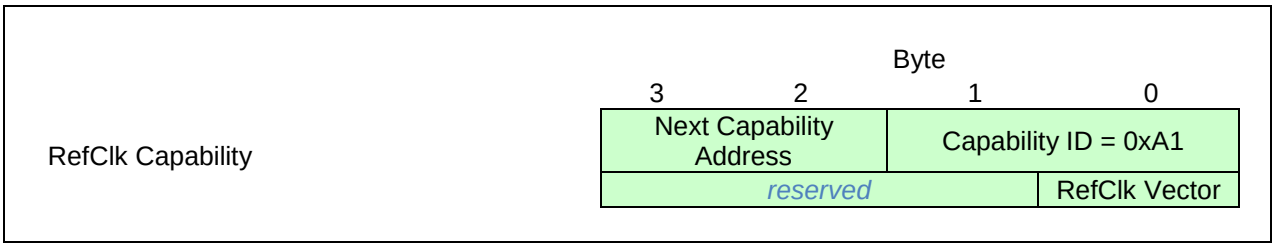
- Capability ID value of 0xA0 specifies that Enterprise PCIe SSD support Dual Port Mode as defined in **Section 5**.
- Next Capability Address is offset to next capability with 0x0 meaning end of capability list.

9.2.2 RefClk Extended Capability.

Figure 34 shows Capability Definition for RefClk. This is not required for standard RefClk implementations, but only for Enterprise PCIe SSD that implement extended RefClk capabilities.

Implementation Note: *This is providing hooks for future implementations as designs move from current standard (RefClk required) to new implementation that implement the PCIe SIG change that allow independent clocking (SRIS ECN - Separate RefClk Independent SSC Engineering Change Notice)* (PCI SIG ECNs, various).

Figure 34: RefClk Extended Capability Definition



- Capability ID value of 0xA1 specifies that SSD supports extended RefClk functionality.
- Next Capability Address is offset to next capability with 0x0 meaning end of capability list.
- RefClk Vector is a bit vector decoded as follows: (Set all that apply for a device)
  - Bit 0 – 0x01 – Common RefClk required as defined – default.
  - Bit 1 – 0x02 – Separate RefClk with no SSC (SRNS)
  - Bit 2 – 0x04 – Separate RefClk with SSC (SRIS)
  - Bit 3 – 0x08 – Can select Common RefClk pin or Separate RefClk using Vendor specific mechanism.
  - Bit 4 - 0x10 – Will automatically sense RefClk pin and use if provided, otherwise will use Separate RefClk.

9.2.3 SM-Bus Temperature Extended Capability.

Figure 35 shows Capability Definition for SM-Bus Temperature Sensor.

Implementation Note: *This is for an SM-Bus Sensor, and does not enumerate the PCIe in-band sensors such as temperature reported via SMART structures.*

Implementation Note: *If a design has multiple SM-Bus temperature sensors, then this Extended Capability can be repeated.*

Figure 35: SM Bus Temperature Extended Capability Definition

Temperature Sensor Capability

Byte 3	Byte 2	Byte 1	Byte 0
Next Capability Address		Capability ID = 0xA2	
reserved		SM-Bus Address	Sensor Type
Overtemp Threshold		Warning Threshold	

- Capability ID value of 0xA2 specifies that a SM-Bus temperature sensor is present.
- Next Capability Address is offset to next capability with 0x0 meaning end of capability list.
- Sensor Type is decoded as follows
  - 0x00 = JC-42.4 TSE2004av compliant temperature sensor on 3.3Vaux
  - 0x01 = JC-42.4 TSE2004av compliant temperature sensor on 12V
    - The sensor shall implement the features, accuracy, temperature register set, and behavior of the TSE2004av temperature sensor. Presence of this temperature sensor shall not imply that TSE2004av EEPROM functionality is available
  - 0x02 = JC-42.4 TSE2002av compliant temperature sensor on 3.3Vaux
  - 0x03 = JC-42.4 TSE2002av compliant temperature sensor on 12V
    - The sensor shall implement the features, accuracy, temperature register set, and behavior of the TSE2004av temperature sensor. Presence of this temperature sensor shall not imply that TSE2002av EEPROM functionality is available
- SM-Bus Address, bits 1 through 7, specify the SM-Bus slave address of the sensor. A typical value is 0x36.
- Warning Threshold specify the minimum temperature that if read from the temperature sensor means the PCIe SSD is warning that it is approaching overheating. Additional cooling is advisable to limit the need for thermal throttling.
- Overtemp Threshold specify the minimum temperature that if read from the temperature sensor means the PCIe SSD is overheating. Accelerated data loss, device shutdown or extreme thermal throttling may result, addition cooling is required.
- The alignment of temperatures is standard for TSE2004av. The temperatures are given in twos compliment centigrade in bits 2 through 12, with bit 2 representing 0.125ºC, but 11 representing 128ºC, and bit 12 representing twos compliment sign bit.

Implementation Note: *The temperatures reported are device specific implementation, and should be treated as a relative value and used to compare to the Warning and Overtemp, Thresholds. The temperature is allowed to be measured at some significant or sensitive location on the drive, and does not have to be the hottest location. Two drives, even from the same vendor, could measure temperature at different locations in the drive, so comparison of temperature between drives is meaningless.*

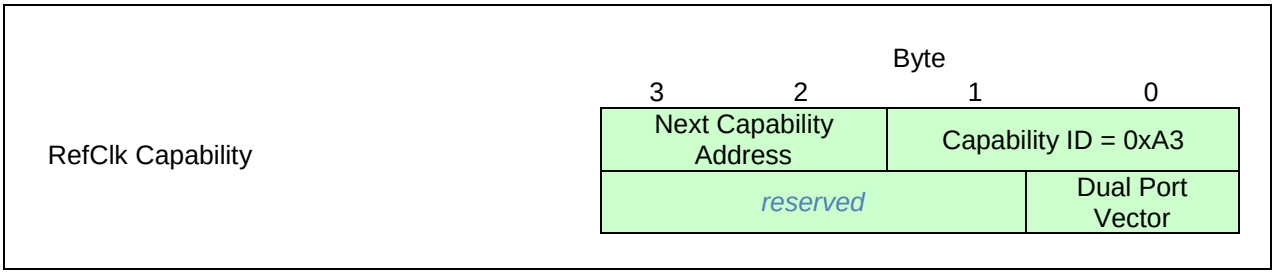
Implementation Note: *If the SM-Bus temperature sensor is connected to 3.3Vaux then to meet the 3.3Vaux power specification of 20µA SM-Bus inactive implies that this temperature sensor is disabled when 12V is off,*

9.2.4 Dual Port and Active/Passive Extended Capability.

Figure 36 shows Extended Capability definition for Active/Passive capability. This is for a partial Dual Port mode but only one PCIe port is active, but this can be Port A (lanes 0-1, RefClk0, ePERst0#), or Port B (lanes 2-3, RefClk1,ePERst1#). The default behavior is Port A is Active and if Link training is successful, then Port-A is used. If Port-A Link training fails Port B is used.

Implementation Note: *This is viewed as a transitional step to get partial Dual Port capability before full dual PCIe controller and dual clocking is supported. This allows the side band interface to force the drive to use Port B signals. This would be done to force failover to a slave controller.*

Figure 36: Dual Port and Active/Passive Extended Capability Definition



- Capability ID value of 0xA3 specifies that SSD supports extended Dual Port functionality.
- Next Capability Address is offset to next capability with 0x0 meaning end of capability list.
- Dual Port Vector is a bit vector decoded as follows: (Set all that apply for a device)
  - Bit 0 – 0x01 – Dual Port mode as specified.
  - Bit 1 – 0x02 – Active/Passive mode

9.3 SM-Bus MCTP commands

Not yet defined.

10 System Usage – Informative Appendix

There are multiple standards that use the same family of connectors – SAS and SATA standards, and emerging standards of Enterprise PCIe SSD (this specification), SATA Express, and SAS. This section gives an overview of how they interoperate.

SATA Express pin placement is only for reference on how to align this specification's Enterprise PCIe pin-out with SATA Express standard (SATA-IO, Draft 2011)**.**

SAS-x4 pin placement is only for reference on how to align this specification's Enterprise PCIe pin-out with SAS x4 specification (SFF-8630, 2012)**.**

10.1 Pin usage across standards

Figure 37 show the pin map across SATA and SAS connector to the Enterprise PCIe SSD. Figure 38 shows the pin map across connectors to the Enterprise PCIe SSD. The figures all use the Enterprise PCIe SSD signal names for consistency.

Implementation Note: *There are multiple signals that are similar across the standards that are driven by system requirements. Even for signals that are logically the same there are multiple copies on the connector since the source or destination on the system design is not the same.*

- *Existing SATA and existing SAS standards do not share high speed signals with Enterprise PCIe SSD. Only support signals are shared (12V, GND, PRSNT#, IfDet#). The SAS/SATA/SATA Express high speed signals are driven by separate SAS/SATA controllers from Enterprise PCIe ports.*
- *SATA Express and Enterprise PCIe SSD standards do not share similar signals including separate high speed signals for PCIe lanes. This signal duplication is driven by the upstream connection. For SATA Express this is the SATA controller typically in a chipset or IO controller. For Enterprise PCIe SSD this is the highest performance PCIe root ports typically off a CPU or PCIe fan-out chipset.*
    - *There are two sets of PCIe high speed signals for Lane 0 and 1. The SATA Express PCIe signals do not use the same wires as the Enterprise PCIe signals. The SATA Express needs the client PCIe to use the same signals as existing SATA (S1-S7). This allows the high speed pins to be reused in the cost constrained client systems. The Enterprise PCIe SSD uses separate signals (E7-E36) since these signals typically go direct to the PCIe root complex.*
    - *There are independent PCIe Reset signals. SATA Express use pin P2 while Enterprise PCIe SSD use pin E5. This allows either controller to independently control the link reset. A controller may drive a reset output to enable simplified device reset.*
    - *There is no PCI RefClk (+/-) for the SATA Express since this is after expected PCIe changes (a mode for greater skip order set insertion) to allow independent device and upstream clocking even with spread spectrum clock source.*
    - *The Reserved(WAKE#/OBFF#) PCIe signal is shared, as are shared support 12V, Gnd, PRSNT#, IfDet#.*
- *SAS-x4* (SFF-8630, 2012) *and Enterprise PCIe SSD high speed signals are mixed – some separate and some shared. This prevents a flexible backplane from supporting Enterprise PCIe SSDs and SAS-x4 in one board, but connectors are mechanically compatible allowing shared components.*
    - *The high speed data signals (lanes) are separate for SAS lanes 0-1 but shared for SAS lanes 2-3.*
    - *SAS-x4 may optionally use new SATA Express support signals such as Reset (P2)*
    - *SAS-x4 may optionally use new Enterprise PCIe SSD support signals such as SM-Bus, and 3.3Vaux.*

- *SAS-x4 may redefine support pins for example layering a second Active (SAS-Active2) on existing Reserved(WAKE#/OBFF#) signal. If this is done the system must be able to gate the signal usage based on drive type.*

Figure 37: Pin usage across Existing standards

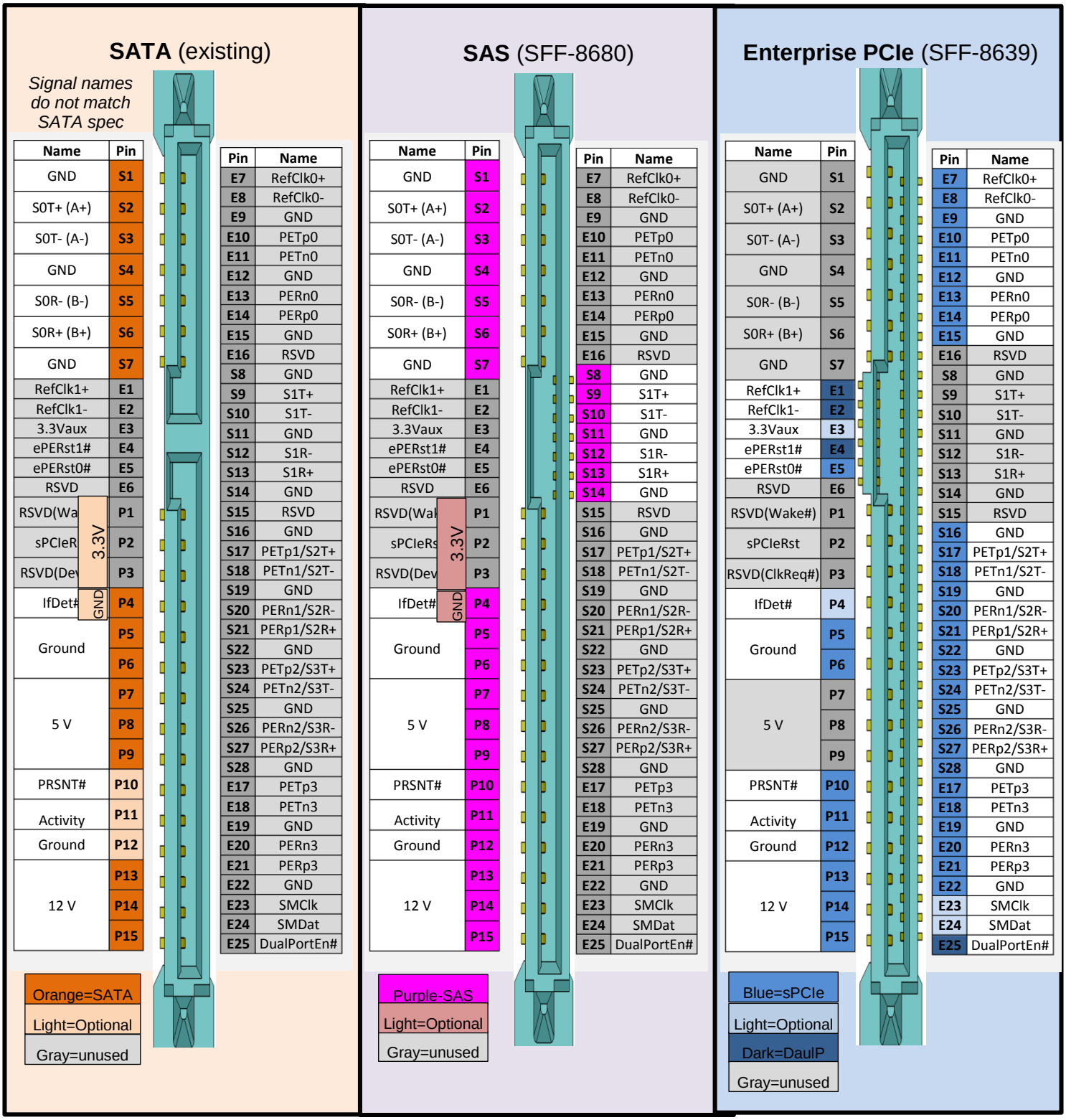
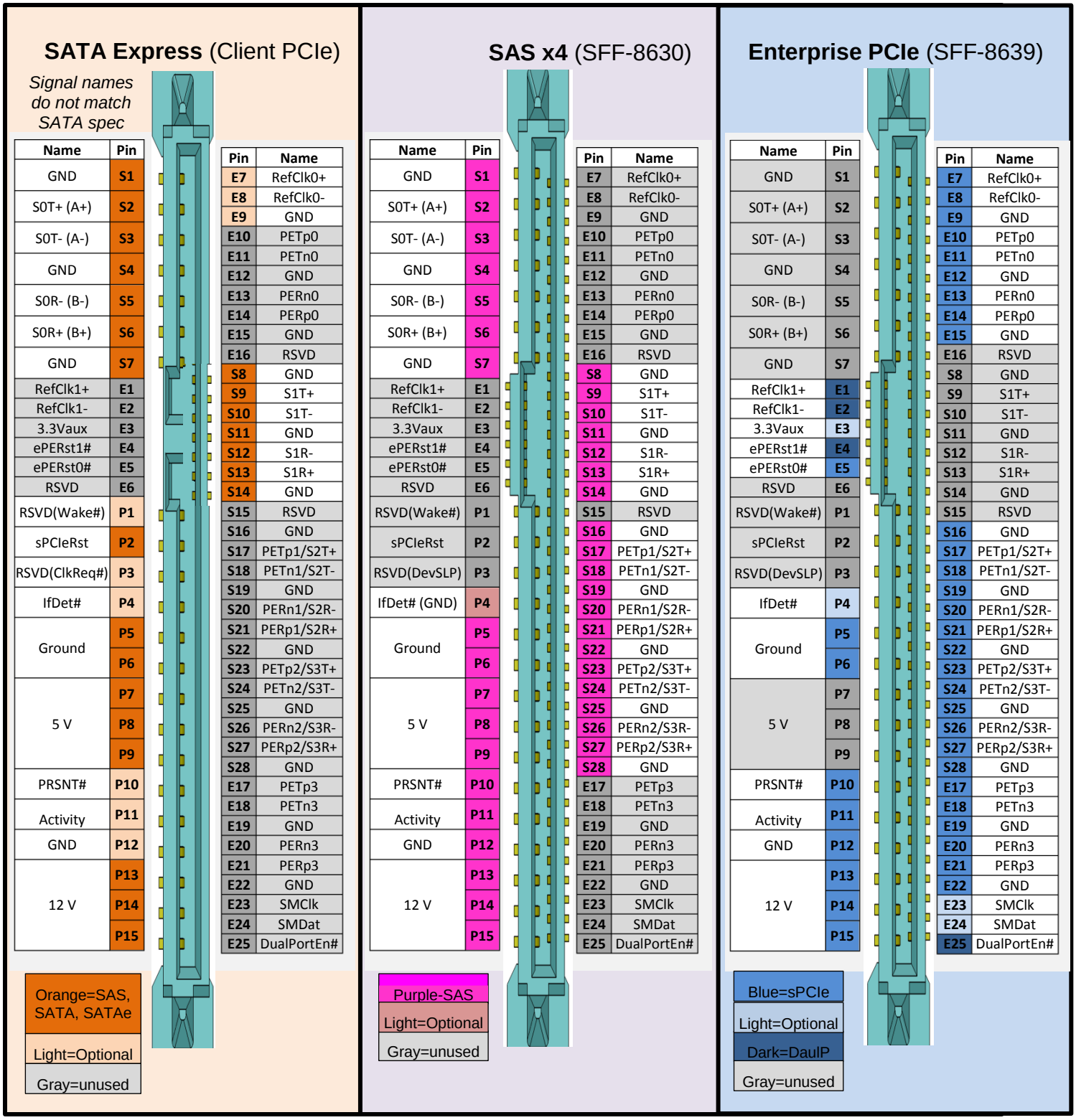


Figure 38: Pin usage across Emerging standards



10.2 System Usage Examples

The drive connector supports five interfaces: SATA, SAS (single & dual), SAS X4, SATA Express (Client PCIe x1-x2 muxed high speed lanes with SATA), Enterprise PCIe x4 (separate high speed lanes). If the system implementation uses cables to attach to a drive, then all 5 drives types could be supported by using different flavors of cables and changing the connection at the system end to either SATA/SATA Express(client PCIe), SAS controller, or Enterprise PCIe x4 root port.

A fixed backplane based configuration (common in servers), the system implementation would have to select which interfaces to support. The following diagrams show the options:

a. Figure 39 shows a backplane and connections for supporting Enterprise SSD PCIe x4, Client PCIe (x1-x2) and SATA. (SATA and Client PCIe x2 are muxed).
b. Figure 40 shows a backplane and connections for supporting Enterprise SSD PCIe x4, SAS (single & dual), and SATA. (SATA and SAS are muxed).
c. Figure 41 shows a backplane and connections for supporting SAS X4, SAS (single & dual), and SATA. (SATA and SAS are muxed).

Figure 39: Backplane for Enterprise SSD (PCIe x4), SATA & SATA Express (Client PCIe)

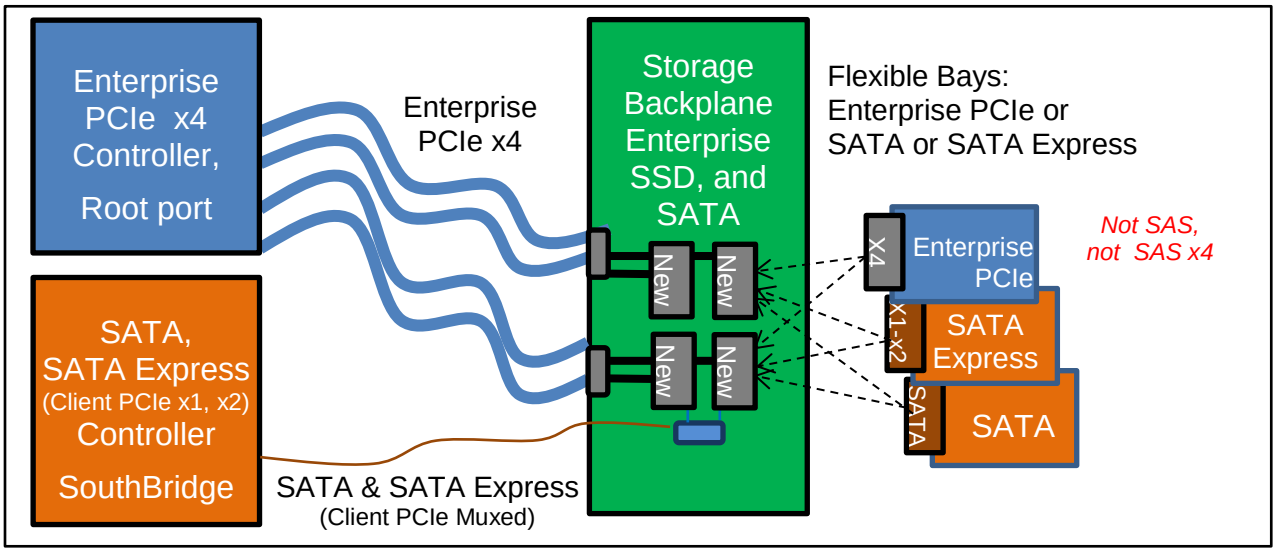


Figure 40: Backplane for Enterprise SSD (PCIe x4), SAS & SATA

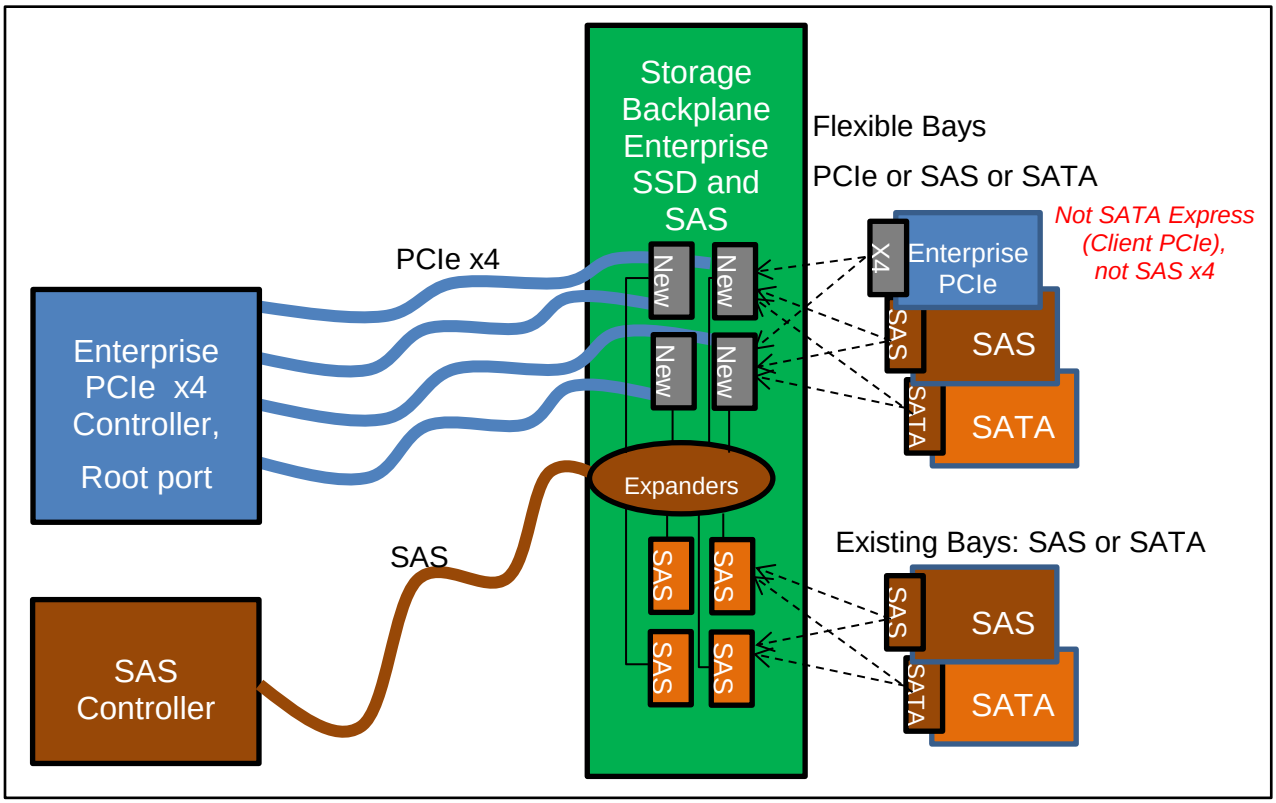
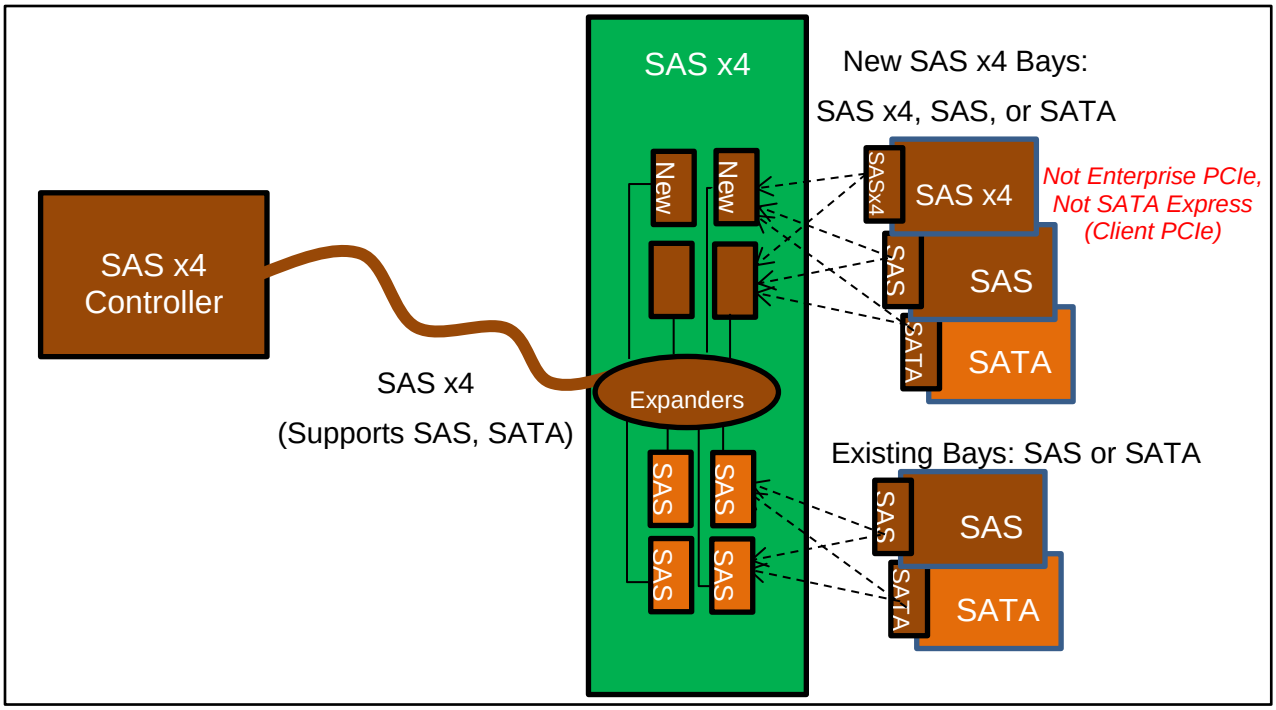


Figure 41: Backplane for SAS X4, SAS & SATA



10.3 Connector Keying (Informative)

The primary goal of connector keying is to make sure that end user pluggable devices and cables can be keyed to prevent insertion, if the device would not function.

Examples:

- Block plugging of a SAS drive into client system with SATA and SATA Express cable.
  - Achieved by mechanical keys in region of pin E3-E4.
- Block plugging of Enterprise PCIe SSD Cable into a SAS or SATA drive.
  - Achieved by cable keying in the alignment posts.
- The standard must also support a universal backplane connector that allows any drive to mate. (Any keying is at the drive carrier or enclosure level.)

SATA Express keying features are only a proposal on how to align this specification's with proposals for SATA Express. The SATA Express keying features are not yet standardized, and this document is not the normative document for SATA Express.

Figure 42 Cross Standard Connector Keying

	 SATA drive	 SATA Express drive	 SAS drive	 Enterprise PCIe drive
 Enterprise backplane	Works- system supports (carrier key)	Works- if system supports (carrier key)	Works- if system supports (carrier key)	Works
 SAS backplane	Works with STP	Mates-Nonfunctional (requires STP+) (carrier key)	Works	Mates-nonfunctional (carrier key)
 SATA Express backplane/laptop	Works	Works	Blocked-Key	Blocked-Key
 SATA backplane/laptop		Blocked-Key	Blocked-Key	Blocked-Key
 Enterprise cable	Blocked-Key	Blocked-Key	Blocked-Key	Works
 SAS cable	Works	Mates-Nonfunctional (requires STP+)	Works	Mates-nonfunctional & no detent retention
 SATA Express cable	Works	Blocked-Key	Blocked-Key	Blocked-Key
 SATA cable	Works	Blocked-Key	Blocked-Key	Blocked-Key

Key:

Works	Works as expected
Works with STP	Works as expected since SATA Tunnel Protocol widely deployed
Works if system supports	Connectors will mate, will work if system designed to support - Section 10.2
Blocked-key	Connectors do not mate, blocked by connector keying features Colored light green because this is desired behavior
Mates-nonfunctional	Connector will mate, but will not function.
Carrier key	Disk carrier is used, and system could be designed to prevent insertion

11 Specification Conventions

11.1 Definitions

11.1.1 PCI Express® (PCIe)

PCIe refer to the PCIe Express IO bus standards. Please see the ***8.1 References*** on *page 39* for a list of the most relevant PCIe documents (PCI SIG, various).

11.1.2 PCI Express® Generation 2 (PCIe Gen2)

In this context Gen 2 is used to refer to a link speed of 5Gbps using 8b/10b encoding.

11.1.3 PCI Express® Generation 3 (PCIe Gen3)

In this context Gen 3 is used to refer to a link speed 8Gbps using a 128/130 encoding with scrambling.

11.1.4 PCI Express® x4 (PCIe x4)

x4 means using 4 PCIe lanes for communication. This is 4 differential pairs in each direction. Total signals is 4 lanes times 2 (differential) times 2 (Transmit and receive) = 16 signals.

11.1.5 SFF standards

Small Form Factor Committee is an ad hoc group that defined the 2.5" & 3.5" drive and related specifications. Documents created by the SFF Committee are submitted to bodies such as EIA (Electronic Industries Association) or an ASC (Accredited Standards Committee).

Please see the ***8.1 References*** on *page 39* for a list of the most relevant (SFF, various).

11.1.6 SSD

Solid State Disk.

11.2 Keywords

Several keywords are used to differentiate between different levels of requirements.

11.2.1 mandatory

A keyword indicating items to be implemented as defined by this specification.

11.2.2 may

A keyword that indicates flexibility of choice with no implied preference.

11.2.3 optional

A keyword that describes features that are not required by this specification. However, if any optional feature defined by the specification is implemented, the feature shall be implemented in the way defined by the specification.

11.2.4 R, RSVD

"R" and RSVD are used as an abbreviation for "reserved" when the figure or table does not provide sufficient space for the full word "reserved".

11.2.5 reserved

A keyword indicating reserved signal pins, bits, bytes, words, fields, and opcode values that are set-aside for future standardization. Their use and interpretation may be specified by future extensions to this or other specifications. A reserved bit, byte, word, field, or register shall be cleared to zero, or in accordance with a future extension to this specification. The recipient shall not check reserved bits, bytes, words, or fields.

11.2.6 shall

A keyword indicating a mandatory requirement. Designers are required to implement all such mandatory requirements to ensure interoperability with other products that conform to the specification.

11.2.7 should

A keyword indicating flexibility of choice with a strongly preferred alternative. Equivalent to the phrase “it is recommended”.