



Yosemite V3: Vernal Falls E1.S 10U Flash Blade and Expansion Board Design Specification

Version 1.5

Authors:

Matt Bowman, Hardware Engineer, Facebook

Abe Garcia, Hardware Engineer, Facebook

Jun Shen, Thermal Engineer, Facebook

Michael Haken, Mechanical Engineer, Facebook

Wei Zhang, Software Engineer, Facebook

Ross Stenfort, Hardware System Engineer, Facebook

1 OCP

1.1. **OPTION B: Open Web Foundation (OWF) CLA**

Contributions to this Specification are made under the terms and conditions set forth in Open Web Foundation Contributor License Agreement (“OWF CLA 1.0”) (“Contribution License”) by:

Facebook

You can review the signed copies of the applicable Contributor License(s) for this Specification on the OCP website at <http://www.opencompute.org/products/specsanddesign> Usage of this Specification is governed by the terms and conditions set forth in **Open Web Foundation Final Specification Agreement (“OWFa 1.0”) (“Specification License”)**.

You can review the applicable Specification License(s) executed by the above referenced contributors to this Specification on the OCP website at <http://www.opencompute.org/participate/legal-documents/>

Notes:

- 1) The following clarifications, which distinguish technology licensed in the Contribution License and/or Specification License from those technologies merely referenced (but not licensed), were accepted by the Incubation Committee of the OCP:

NONE

- 2) The above license does not apply to the Appendix or Appendices. The information in the Appendix or Appendices is for reference only and non-normative in nature.

NOTWITHSTANDING THE FOREGOING LICENSES, THIS SPECIFICATION IS PROVIDED BY OCP "AS IS" AND OCP EXPRESSLY DISCLAIMS ANY WARRANTIES (EXPRESS, IMPLIED, OR OTHERWISE), INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, FITNESS FOR A PARTICULAR PURPOSE, OR TITLE, RELATED TO Date: XXXX, 2XXXX Page 5 Open Compute Project

- THE SPECIFICATION. NOTICE IS HEREBY GIVEN, THAT OTHER RIGHTS NOT GRANTED AS SET FORTH ABOVE, INCLUDING WITHOUT LIMITATION, RIGHTS OF THIRD PARTIES WHO DID NOT EXECUTE THE ABOVE LICENSES, MAY BE IMPLICATED BY THE IMPLEMENTATION OF OR COMPLIANCE WITH THIS SPECIFICATION. OCP IS NOT RESPONSIBLE FOR IDENTIFYING RIGHTS

FOR WHICH A LICENSE MAY BE REQUIRED IN ORDER TO IMPLEMENT THIS SPECIFICATION. THE ENTIRE RISK AS TO IMPLEMENTING OR OTHERWISE USING THE SPECIFICATION IS ASSUMED BY YOU. IN NO EVENT WILL OCP BE LIABLE TO YOU FOR ANY MONETARY DAMAGES WITH RESPECT TO ANY CLAIMS RELATED TO, OR ARISING OUT OF YOUR USE OF THIS SPECIFICATION, INCLUDING BUT NOT LIMITED TO ANY LIABILITY FOR LOST PROFITS OR ANY CONSEQUENTIAL, INCIDENTAL, INDIRECT, SPECIAL OR PUNITIVE DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND EVEN IF OCP HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

2 Scope

This specification describes the design of the E1.S expansion system that can be used for the next generation Yosemite V3 platform.

3 Contents

1	OCP	2
2	Scope	3
3	Contents	3
4	Overview.....	5
4.1	Reference Documents	6
5	Mechanical	7
5.1	Platform Details.....	7
5.2	Expansion Card Requirements	9
6	Electrical	10
6.1	Design Guidelines	10
6.2	Front Expansion Pinout	10
6.3	E1.S Drive Pinout	17
7	Power	18
7.1	Input Voltage Level.....	18
7.2	E1.S Expansion Slot Power Budget.....	18
7.3	Capacitive Load	19
7.4	VR Efficiency	19
7.5	Power Sequence and Power States.....	19
7.6	Hot Swap Controller	21
8	Hardware System Overview	22

8.1	PCIe.....	23
8.2	Clock Tree	25
8.3	I2C Topology.....	25
8.4	LED Definitions	25
9	Functional.....	26
9.1	Low Speed Sideband	26
9.2	Debug	26
9.3	EEPROM.....	26
9.4	Expansion System Management	26
10	Thermal	28
10.1	Data Center Environmental Conditions.....	28
10.2	Server Operational Conditions	29
11	Environmental Requirements	29
11.1	Vibration and Shock	29
12	Labels and Markings.....	30
13	Prescribed Materials	31
13.1	Disallowed Components.....	31
13.2	Capacitors and Inductors.....	31
13.3	Component De-rating.....	32
14	Revision History	33
15	Appendix A	33

4 Overview

This document describes the 10U E1.S front expansion board for the Delta Lake server blade. This board has a codename of “Vernal Falls”.

The E1.S front expansion board interfaces with the 1S server blade through the following connectors seen in Figure 4-1 - 10U Server Board with no Expansion Cards.

- 1x Straddle mount 4C+
- 1x Straddle mount 4C (mechanical only, no signaling)

The front expansion system uses the OCP NIC 3.0 interface as a starting point. The use of this interface provides flexibility for different add-in card use cases. Not all signals on this interface are applicable to all expansion system use cases.

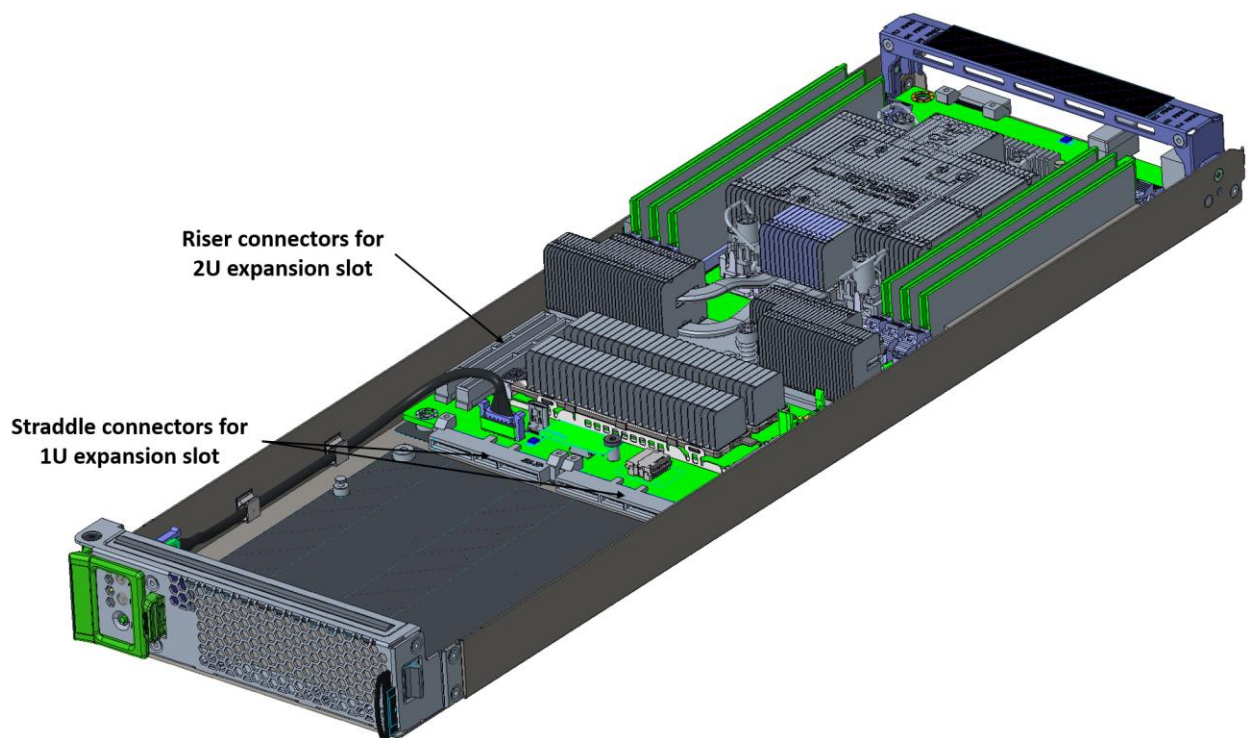


Figure 4-1 - 10U Server Board with no Expansion Cards

The E1.S expansion board will provide an interface for the Delta Lake attach to E1.S media. The expansion board is targeting a unit density of four sockets capable of supporting the E1.S-25mm form factor. This specification is focused on enabling the storage use cases on the Yosemite V3/Delta Lake platform.

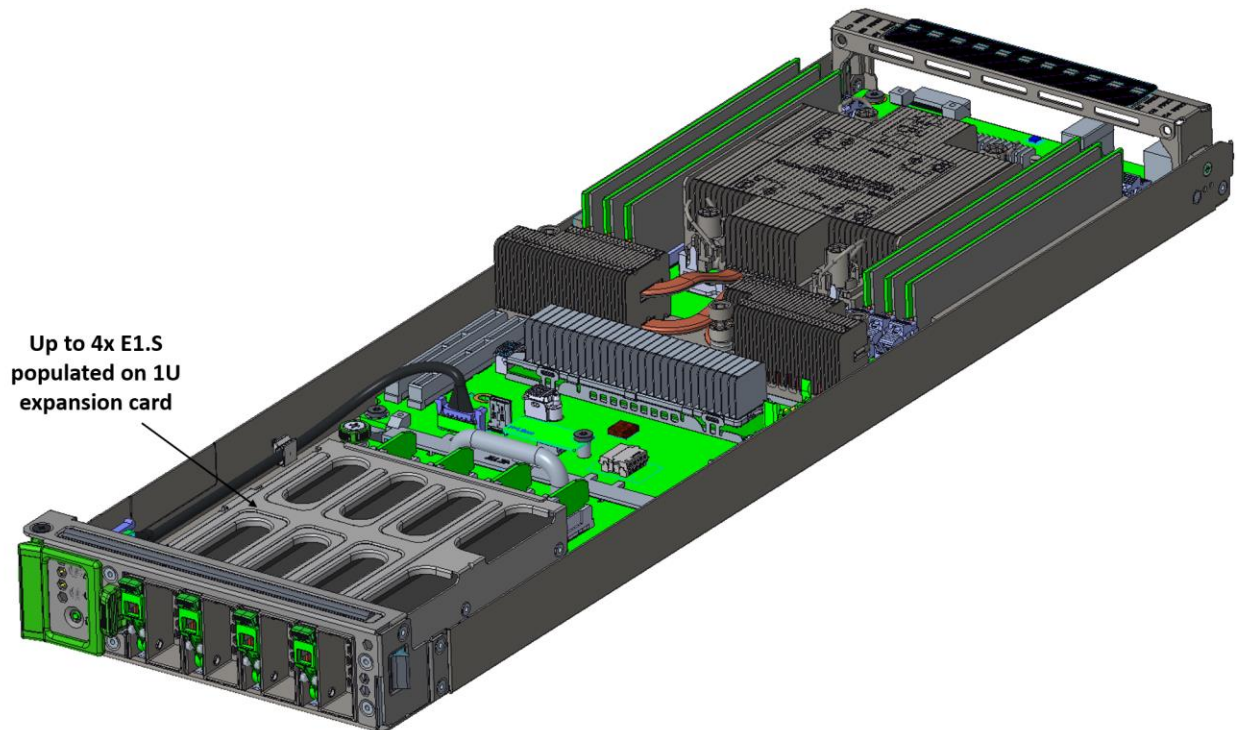


Figure 4-2- 10U Server with E1.S Expansion Card

4.1 Reference Documents

Cooper Lake Server Processor External Design Specification, Volume One: Architecture (Intel)

OCP NIC 3.0 Design Specification (OCP Server Workgroup, OCP NIC Subgroup)

PCIe CEM Specification: Revision 4 (PCI-SIG)

SFF-TA-1002: Protocol Agnostic Multi-Lane High Speed Connector (SNIA)

SFF-TA-1006: Enterprise and Datacenter 1U Short SSD Form Factor (E1.S) (SNIA)

SFF-TA-1009 – Enterprise and Datacenter SSD Pin and Signal Specification (SNIA)

Facebook Indicator Specification Version 1.0 (Facebook)

Yosemite V3: Facebook Multi-Node Server Platform Design Specification (Facebook)

Yosemite V3: Yosemite V3: Delta Lake 1S Server Design Specification (Facebook)

Yosemite V3: Generic Expansion Design Specification (Facebook)

5 Mechanical

The E1.S expansion board will be designed to fit into the 10U Delta Lake single socket server in a Yosemite V3 chassis.

5.1 Platform Details

These details are for informational use only – use the Yosemite V3: Delta Lake 1S Server Design Specification and Yosemite V3: Generic Expansion Design Specification for detailed information on the mechanical constraints of the expansion board.

The Delta Lake 1S server provides two edge connectors for connectivity to the front expansion options. They are the straddle-mount 4C+ and 4C connectors as described in the *SFF-TA-1002: Protocol Agnostic Multi-Lane High Speed Connector* specification.

Figure 5-1 defines the locations of the 4C+ and 4C connectors. The 1U E1.S expansion board must be confined in the regions defined in Figure 5-1 - Delta Lake Expansion Card Keep-in Area and Figure 5-2 - Expansion Card Keep-in Areas (1U and 2U). Note that the 2U expansion region is inclusive of the 1U expansion region.

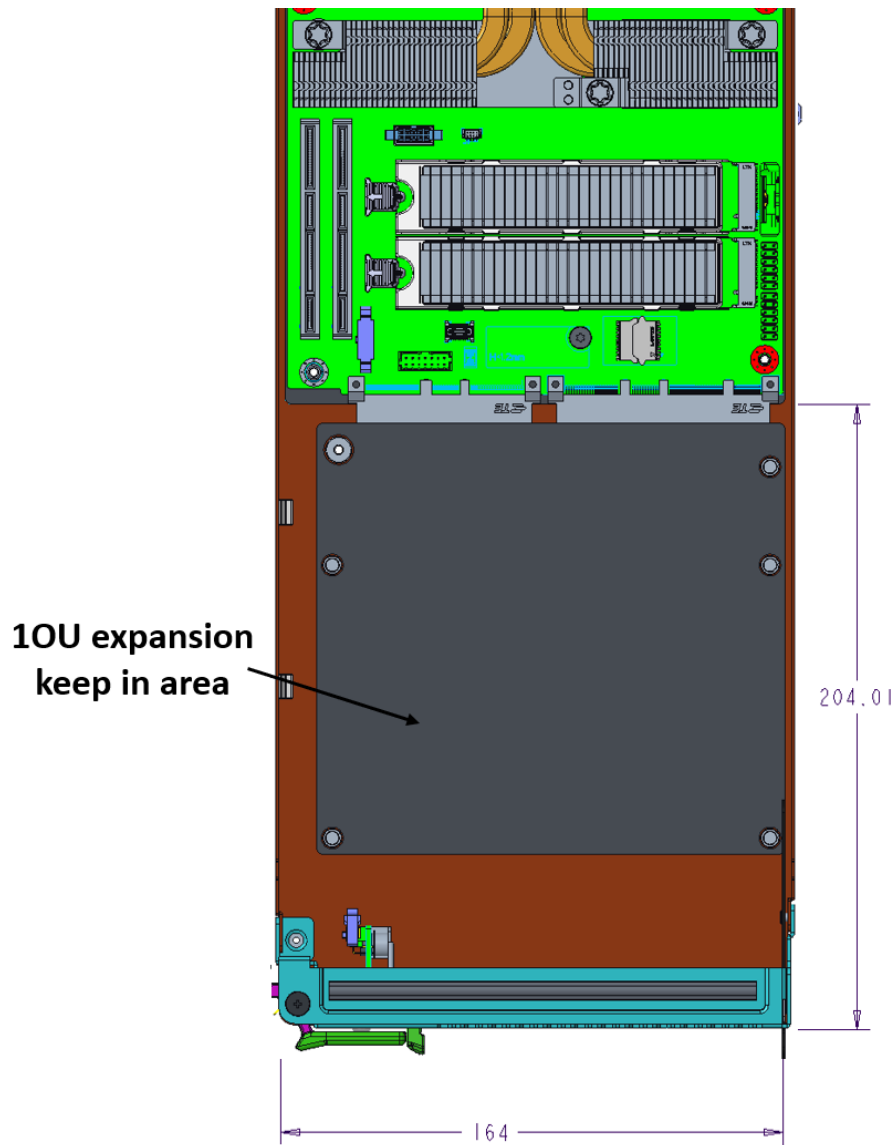


Figure 5-1 - Delta Lake Expansion Card Keep-in Area

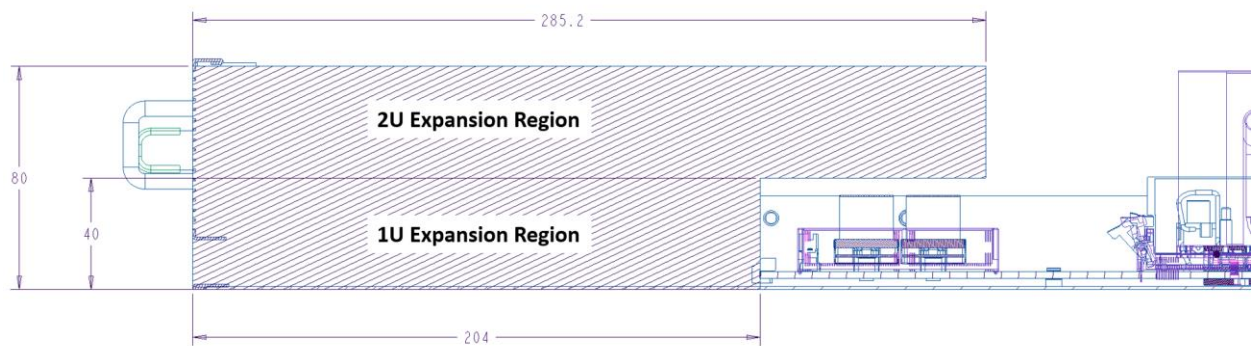


Figure 5-2 - Expansion Card Keep-in Areas (1U and 2U)

5.2 Expansion Card Requirements

A high-level overview of the mechanical requirements for the expansion card are as follows:

- The E1.S expansion card must fit in the 1U expansion region.
- The E1.S expansion card must support four front accessible E1.S SSDs
 - Each EDSFF connector must be orthogonal
 - Each EDSFF connector must be a x4 width
 - Each EDSFF connector must have clearance to support x8 E1.S components to be inserted without interference
 - Each EDSFF drive slot must support the E1.S-25mm asymmetric drive as described in *SFF-TA-1006:Enterprise and Datacenter 1U Short SSD Form Factor (E1.S)*
- The E1.S Status and Attention LEDs must be visible from the front panel
- The E1.S expansion board must be compliant to SFF-TA-1002 with respect to PCB thickness.
- The E1.S expansion board must be compliant to SFF-TA-1002 with respect to the gold finger physical dimensions and tolerances.
 - The E1.S expansion board must use the straddle mount 4C+ connector found on the Delta Lake 1 socket server.
 - The E1.S expansion board should also include a card edge to connect to the 4C connector on the Delta Lake 1S server. However, this connector may be depopulated on the Delta Lake board so no signals from this connection should be utilized in the expansion board design.

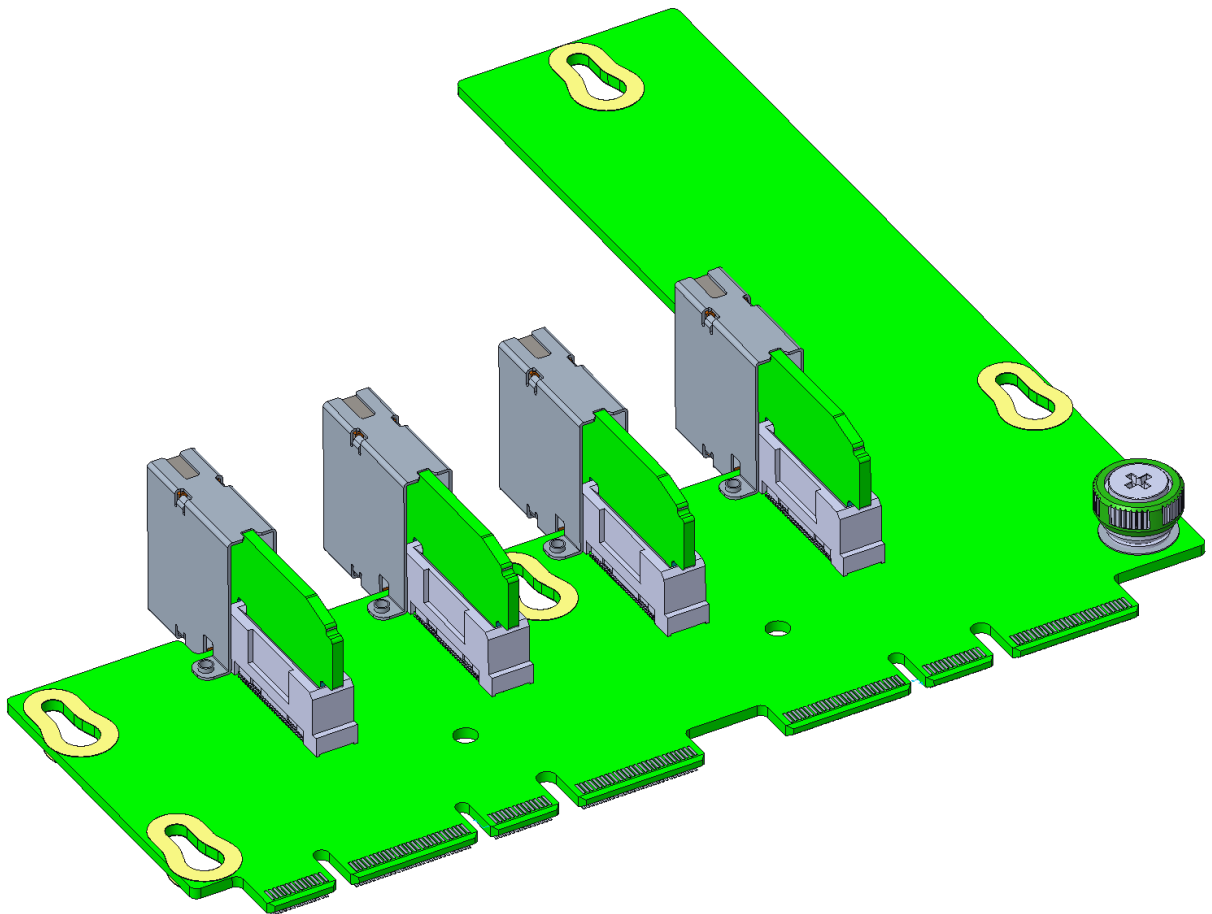


Figure 5-3 – Potential E1.S Expansion Card

6 Electrical

6.1 Design Guidelines

Most of the electrical design of the expansion modules would be referred to PDG documents of the CPU platform as well as the OCP NIC 3.0 specifications to provide consistent impedance and power connectivity.

6.2 Front Expansion Pinout

The connectivity requirements of the E1.S front expansion board are as follows:

1. One x16 PCIe Gen3 link to the host that is bifurcated into four x4 links.
2. One Gen3 compliant PCIe Reference clock
3. 4-wire serial interface for faster management/debug status
4. SMBUS or I2C interface for sideband management/thermal information
5. Control/alert signals
6. Power

The secondary 4C connector has no signals except for power signals connected on the Delta Lake Server side. It may be used for additional power delivery if needed. Vernal Falls will not populate this connector.

All unused pins adjacent to utilized signals shall be tied to ground through an RC network as defined in the PCIe CEM specification¹. The resistance and capacitance of this network shall be optimized to minimize any resonance and the resulting high-frequency crosstalk on adjacent high-speed signals.

Unless otherwise noted, the 4C+ connector is expected to follow the pin and signal specifications defined in the following specifications:

- *OCP NIC 3.0 Design Specification*

Table 1: Front Expansion Pin-out (Primary)

Notes	Side B		Side A		Notes
PWR_GOOD to host	OCP_B1	NIC_PWR_GOOD	PERST2#	OCP_A1	Unused
PWR_EN from host	OCP_B2	MAIN_PWR_EN	PERST3#	OCP_A2	Unused
SGPIO bus	OCP_B3	Unused	WAKE_N	OCP_A3	Unused
SGPIO bus	OCP_B4	Unused	RBT_ARB_IN	OCP_A4	Unused, N/C on Delta Lake
SGPIO bus	OCP_B5	Unused	RBT_ARB_OUT	OCP_A5	Unused, N/C on Delta Lake
SGPIO bus	OCP_B6	Unused	SLOT_ID1	OCP_A6	To expansion card FRU EEPROM A0 pin
To expansion card FRU EEPROM A1 pin ²	OCP_B7	SLOT_ID0	RBT_TX_EN	OCP_A7	Unused, N/C on Delta Lake

¹ PCIe CEM Specification Revision 4.0, Version 1.0

See ² OCP NIC 3.0 Specification for more detail

E1.S 10U Flash Blade and Expansion Board Design Specification – Vernal Falls

Unused	OC_P_B8	RBT_RXD1	RBT_TXD1	OC_P_A8	Unused, N/C on Delta Lake
Unused	OC_P_B9	RBT_RXD0	RBT_TXD0	OC_P_A9	Unused, N/C on Delta Lake
	OC_P_B10	GND	GND	OC_P_A10	
SMB_IPMB_3V3SB_CLK on Delta Lake	OC_P_B11	REFCLKn2	REFCLKn3	OC_P_A11	SMBUS_ALERT_N on Delta Lake
SMB_IPMB_3V3SB_DAT on Delta Lake	OC_P_B12	REFCLKp2	REFCLKp3	OC_P_A12	CARD_TYPE on Delta Lake
	OC_P_B13	GND	GND	OC_P_A13	
Unused, N/C on Delta Lake	OC_P_B14	RBT_CRS_DV	RBT_CLK_IN	OC_P_A14	Unused, N/C on Delta Lake
Mechanical Key					
P12V_STBY	B1	+12V_EDGE	GND	A1	
P12V_STBY	B2	+12V_EDGE	GND	A2	
P12V_STBY	B3	+12V_EDGE	GND	A3	
P12V_STBY	B4	+12V_EDGE	GND	A4	
P12V_STBY	B5	+12V_EDGE	GND	A5	
P12V_STBY	B6	+12V_EDGE	GND	A6	
N/C on Delta Lake	B7	BIF0#	SMCLK	A7	SMB_BMC_CPLD_CLK
N/C on Delta Lake	B8	BIF1#	SMDAT	A8	SMB_BMC_CPLD_DAT
N/C on Delta Lake	B9	BIF2#	SMRST#	A9	Unused, N/C on Delta Lake
To Expansion Board BIC	B10	PERST0#	PRSNTA#	A10	Board Presence Indicator
P3V3_AUX	B11	+3.3V_EDGE	PERST1#	A11	Unused
AUX_PWR_EN	B12	AUX_PWR_EN	PRSNTB2#	A12	Board Bifurcation Indicator (see 6.2.1)
	B13	GND	GND	A13	
To CLK Mux	B14	REFCLKn0	REFCLKn1	A14	UART_RX on Delta Lake
To CLK Mux	B15	REFCLKp0	REFCLKp1	A15	UART_TX on Delta Lake
	B16	GND	GND	A16	
To Drive 0	B17	PETn0	PERn0	A17	From Drive 0
To Drive 0	B18	PETp0	PERp0	A18	From Drive 0
	B19	GND	GND	A19	
To Drive 0	B20	PETn1	PERn1	A20	From Drive 0
To Drive 0	B21	PETp1	PERp1	A21	From Drive 0
	B22	GND	GND	A22	
To Drive 0	B23	PETn2	PERn2	A23	From Drive 0
To Drive 0	B24	PETp2	PERp2	A24	From Drive 0
	B25	GND	GND	A25	

E1.S 10U Flash Blade and Expansion Board Design Specification – Vernal Falls

To Drive 0	B26	PETn3	PERn3	A26	From Drive 0
To Drive 0	B27	PETp3	PERp3	A27	From Drive 0
	B28	GND	GND	A28	
Mechanical Key					
	B29	GND	GND	A29	
To Drive 1	B30	PETn4	PERn4	A30	From Drive 1
To Drive 1	B31	PETp4	PERp4	A31	From Drive 1
	B32	GND	GND	A32	
To Drive 1	B33	PETn5	PERn5	A33	From Drive 1
To Drive 1	B34	PETp5	PERp5	A34	From Drive 1
	B35	GND	GND	A35	
To Drive 1	B36	PETn6	PERn6	A36	From Drive 1
To Drive 1	B37	PETp6	PERp6	A37	From Drive 1
	B38	GND	GND	A38	
To Drive 1	B39	PETn7	PERn7	A39	From Drive 1
To Drive 1	B40	PETp7	PERp7	A40	From Drive 1
	B41	GND	GND	A41	
Board Bifurcation Indicator (see 6.2.1)	B42	PRSENTB0#	PRSENTB1#	A42	Board Bifurcation Indicator (see 6.2.1)
Mechanical Key					
	B43	GND	GND	A43	
To Drive 2	B44	PETn8	PERn8	A44	From Drive 2
To Drive 2	B45	PETp8	PERp8	A45	From Drive 2
	B46	GND	GND	A46	
To Drive 2	B47	PETn9	PERn9	A47	From Drive 2
To Drive 2	B48	PETp9	PERp9	A48	From Drive 2
	B49	GND	GND	A49	
To Drive 2	B50	PETn10	PERn10	A50	From Drive 2
To Drive 2	B51	PETp10	PERp10	A51	From Drive 2
	B52	GND	GND	A52	
To Drive 2	B53	PETn11	PERn11	A53	From Drive 2
To Drive 2	B54	PETp11	PERp11	A54	From Drive 2
	B55	GND	GND	A55	
To Drive 3	B56	PETn12	PERn12	A56	From Drive 3
To Drive 3	B57	PETp12	PERp12	A57	From Drive 3
	B58	GND	GND	A58	
To Drive 3	B59	PETn13	PERn13	A59	From Drive 3

E1.S 10U Flash Blade and Expansion Board Design Specification – Vernal Falls

To Drive 3	B60	PETp13	PERp13	A60	From Drive 3
	B61	GND	GND	A61	
To Drive 3	B62	PETn14	PERn14	A62	From Drive 3
To Drive 3	B63	PETp14	PERp14	A63	From Drive 3
	B64	GND	GND	A64	
To Drive 3	B65	PETn15	PERn15	A65	From Drive 3
To Drive 3	B66	PETp15	PERp15	A66	From Drive 3
	B67	GND	GND	A67	
Unused, N/C on Delta Lake	B68	RFU1, NC	USB_DATn	A68	Unused
Unused, N/C on Delta Lake	B69	RFU2, NC	USB_DATp	A69	Unused
Board Presence Indicator	B70	PRSNTB3#	PWRBRK0#	A70	Power Break, Unused

Table 2: Front Expansion System Pin-out (secondary)

Notes	Side B		Side A		Notes
P12V_STBY	B1	+12V_EDGE	GND	A1	
P12V_STBY	B2	+12V_EDGE	GND	A2	
P12V_STBY	B3	+12V_EDGE	GND	A3	
P12V_STBY	B4	+12V_EDGE	GND	A4	
P12V_STBY	B5	+12V_EDGE	GND	A5	
P12V_STBY	B6	+12V_EDGE	GND	A6	
Unused, N/C on Delta Lake	B7	BIF0#	SMCLK	A7	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B8	BIF1#	SMDAT	A8	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B9	BIF2#	SMRST#	A9	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B10	PERST4#	PRSENTA#	A10	Board Presence Indicator, Unused, N/C on Delta Lake
P3V3_AUX	B11	+3.3V_EDGE	PERST5#	A11	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B12	AUX_PWR_EN	PRSENTB2#	A12	Board Bifurcation Indicator, Unused, N/C on Delta Lake
	B13	GND	GND	A13	
Unused, N/C on Delta Lake	B14	REFCLKn4	REFCLKn5	A14	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B15	REFCLKp4	REFCLKp5	A15	Unused, N/C on Delta Lake
	B16	GND	GND	A16	
Unused, N/C on Delta Lake	B17	PETn16	PERn16	A17	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B18	PETp16	PERp16	A18	Unused, N/C on Delta Lake
	B19	GND	GND	A19	
Unused, N/C on Delta Lake	B20	PETn17	PERn17	A20	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B21	PETp17	PERp17	A21	Unused, N/C on Delta Lake
	B22	GND	GND	A22	
Unused, N/C on Delta Lake	B23	PETn18	PERn18	A23	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B24	PETp18	PERp18	A24	Unused, N/C on Delta Lake
	B25	GND	GND	A25	
Unused, N/C on Delta Lake	B26	PETn19	PERn19	A26	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B27	PETp19	PERp19	A27	Unused, N/C on Delta Lake
	B28	GND	GND	A28	

E1.S 10U Flash Blade and Expansion Board Design Specification – Vernal Falls

Mechanical Key					
	B29	GND	GND	A29	
Unused, N/C on Delta Lake	B30	PETn20	PERn20	A30	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B31	PETp20	PERp20	A31	Unused, N/C on Delta Lake
	B32	GND	GND	A32	
Unused, N/C on Delta Lake	B33	PETn21	PERn21	A33	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B34	PETp21	PERp21	A34	Unused, N/C on Delta Lake
	B35	GND	GND	A35	
Unused, N/C on Delta Lake	B36	PETn22	PERn22	A36	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B37	PETp22	PERp22	A37	Unused, N/C on Delta Lake
	B38	GND	GND	A38	
Unused, N/C on Delta Lake	B39	PETn23	PERn23	A39	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B40	PETp23	PERp23	A40	Unused, N/C on Delta Lake
	B41	GND	GND	A41	
Board Bifurcation Indicator, Unused, N/C on Delta Lake	B42	PRSENTB0#	PRSENTB1#	A42	Board Bifurcation Indicator, Unused, N/C on Delta Lake
Mechanical Key					
	B43	GND	GND	A43	
Unused, N/C on Delta Lake	B44	PETn24	PERn24	A44	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B45	PETp24	PERp24	A45	Unused, N/C on Delta Lake
	B46	GND	GND	A46	
Unused, N/C on Delta Lake	B47	PETn25	PERn25	A47	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B48	PETp25	PERp25	A48	Unused, N/C on Delta Lake
	B49	GND	GND	A49	
Unused, N/C on Delta Lake	B50	PETn26	PERn26	A50	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B51	PETp26	PERp26	A51	Unused, N/C on Delta Lake
	B52	GND	GND	A52	
Unused, N/C on Delta Lake	B53	PETn27	PERn27	A53	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B54	PETp27	PERp27	A54	Unused, N/C on Delta Lake
	B55	GND	GND	A55	
Unused, N/C on Delta Lake	B56	PETn28	PERn28	A56	Unused, N/C on Delta Lake

E1.S 10U Flash Blade and Expansion Board Design Specification – Vernal Falls

Unused, N/C on Delta Lake	B57	PETp28	PERp28	A57	Unused, N/C on Delta Lake
	B58	GND	GND	A58	
Unused, N/C on Delta Lake	B59	PETn29	PERn29	A59	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B60	PETp29	PERp29	A60	Unused, N/C on Delta Lake
	B61	GND	GND	A61	
Unused, N/C on Delta Lake	B62	PETn30	PERn30	A62	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B63	PETp30	PERp30	A63	Unused, N/C on Delta Lake
	B64	GND	GND	A64	
Unused, N/C on Delta Lake	B65	PETn31	PERn31	A65	Unused, N/C on Delta Lake
Unused, N/C on Delta Lake	B66	PETp31	PERp31	A66	Unused, N/C on Delta Lake
	B67	GND	GND	A67	
Unused, N/C	B68	RFU3	UART_RX	A68	Unused, N/C on Delta Lake
Unused, N/C	B69	RFU4	UART_TX	A69	Unused, N/C on Delta Lake
Board Presence Indicator, N/C on Delta Lake	B70	PRSENTB3#	PWRBRK1#	A70	Unused, N/C on Delta Lake

6.2.1 Front Expansion Pin Definitions

The expansion board connector was derived from the OCP NIC 3.0 connector. This section highlights any differences in implementation.

Signal name	Direction (with respect to server blade)	Description
CARD_TYPE_N	I	Found in front expansion module. Used to determine the card type. 1.65V -1.75V: Expansion with NIC 0.8-0.9V: Expansion with 4 EDSFF 0.0-0.2V: Expansion with 6 M.2
(Primary) PRSENTB[0:3]#	I	Assign to 0b1100 – SFF 4C+, single host, 1 upstream socket, 4 links per socket, 4 lanes per link
(Primary) BIF[0:2]#	O	Unused, N/C on Delta Lake
(Secondary) SMCLK, SMRST#	O	Unused, N/C on Delta Lake

(Secondary) SMDAT	I/O	Unused, N/C on Delta Lake
(Secondary) PRSNTB[0:3]#	I	Unused, N/C on Delta Lake
(Secondary) BIF[0:2]#	O	Unused, N/C on Delta Lake
PER[pn][16:31]	I	Unused, N/C on Delta Lake
PET[pn][16:31],	O	Unused, N/C on Delta Lake
NIC_PWR_GOOD	I	Expansion board power good signal – equivalent in function to NIC_PWR_GOOD from the OCP3.0 NIC specification.
PWRBRK[01]#	O	Unused, N/C on Delta Lake
PERST4#, PERST5#	O	Unused, N/C on Delta Lake
REFCLK[pn]4, REFCLK[pn]5	O	Unused, N/C on Delta Lake
WAKE#	I	Unused

6.3 E1.S Drive Pinout

The connectivity requirements of each E1.S SSD are as follows:

1. One x4 PCIe Gen3 link to the host
2. One Gen3 compliant PCIe Reference clock per drive
3. SMBUS/I2C interface for sideband management/thermal information
4. LED Control functionality
5. Presence Detect monitoring
6. Power

There is no expectation for dual port support:

- DUALPORTEN# can be pulled into a disabled state (strap to +3.3V)
- REFCLK[pn]1 does not need to be provided
- PERST1#/CLKREQ# is not supported

The E1.S drives are expected to support the EDSFF x4 pin and signal specifications defined in the following specifications:

- SFF-TA-1009 – Enterprise and Datacenter SSD Pin and Signal Specification

Table 3: E1.S Connector Pin-out

Notes	Side B		Side A		Notes
P12V_STBY	B1	+12V	GND	A1	
P12V_STBY	B2	+12V	GND	A2	
P12V_STBY	B3	+12V	GND	A3	
P12V_STBY	B4	+12V	GND	A4	
P12V_STBY	B5	+12V	GND	A5	
P12V_STBY	B6	+12V	GND	A6	

E1.S 10U Flash Blade and Expansion Board Design Specification – Vernal Falls

No Connect	B7	MFG	SMCLK	A7	
No Connect	B8	RFU	SMDAT	A8	
Pull-up to 3.3V	B9	DUALPORTEN#	SMRST#	A9	
	B10	PERST0#	LED#/ACTIVITY	A10	Service LED Indicator
P3V3_AUX	B11	+3.3V_AUX	PERST1#/CLKREQ#	A11	Unused
	B12	PWRDIS	PRSNT0#	A12	SSD Presence Indicator
	B13	GND	GND	A13	
PCIE REFCLK	B14	REFCLKn0	REFCLKn1	A14	Unused
PCIE REFCLK	B15	REFCLKp0	REFCLKp1	A15	Unused
	B16	GND	GND	A16	
Lane 0 To Drive	B17	PETn0	PERn0	A17	Lane 0 From Drive
Lane 0 To Drive	B18	PETp0	PERp0	A18	Lane 0 From Drive
	B19	GND	GND	A19	
Lane 1 To Drive	B20	PETn1	PERn1	A20	Lane 1 From Drive
Lane 1 To Drive	B21	PETp1	PERp1	A21	Lane 1 From Drive
	B22	GND	GND	A22	
Lane 2 To Drive	B23	PETn2	PERn2	A23	Lane 2 From Drive
Lane 2 To Drive	B24	PETp2	PERp2	A24	Lane 2 From Drive
	B25	GND	GND	A25	
Lane 3 To Drive	B26	PETn3	PERn3	A26	Lane 3 From Drive
Lane 3 To Drive	B27	PETp3	PERp3	A27	Lane 3 From Drive
	B28	GND	GND	A28	

7 Power

7.1 Input Voltage Level

The expected nominal input voltage delivered by the Delta Lake 1S server's power delivery subsystem is 12.5VDC; however, it has a varying range of 11.5V to 13.5V.

The E1S expansion board shall operate normally with an input voltage range of 12.5V +/- 7%.

7.1.1 Input Voltage Level – SSD

The input voltage level for the E1.S SSDs shall adhere to the requirements defined in SFF-TA-1009.

7.2 E1.S Expansion Slot Power Budget

As described in the Yosemite V3: Delta Lake 1S Server Design Specification, the Delta Lake 1S Server shall support various expansion configurations and blade chassis sizes to accommodate

various future use cases. However, the expansion subsystem must adhere to specific power design requirements:

1. The sustained current draw from the front expansion shall be at most 15A
2. Although the power budget may allow for high power commodities, feasibility is ultimately determined by whether the cooling solution implemented is adequate and does not introduce long term reliability degradation of the hardware
3. The expansion system must consume minimal power when not in Main Power Mode (see **Error! Reference source not found.** for details).

7.2.1 E1.S Expansion Board Power Budget – SSD

The maximum 12V input power for the E1.S SSDs shall be a subset of the requirements defined in SFF-TA-1009. A maximum continuous power draw of 70W is supported by the EDSFF connector. However, the recommended sustained power an E1.S device with an asymmetric enclosure is 25W. We are planning to support a sustained E1.S SSD power draw of 16W per SSD.

The 3.3V edge pin is rated to 1.1A for a maximum derated power delivery of 3.6W

7.3 Capacitive Load

The capacitance on the input 12V power rail for Delta Lake 1S Server's expansion system shall be optimized to meet the system's power supply requirements and does not cause instability.

In addition, the E1.S expansion board shall be designed such that it adheres to the maximum capacitive load allowable as defined in the Yosemite V3 Platform Design Specification document.

7.4 VR Efficiency

High efficiency Voltage Regulators (VR) shall be used on all Delta Lake 1S server expansion systems with at least 91% efficiency over the 30% to 90% load range. If higher efficiency VRs are available at additional cost and/or design complexity, then the vendor is encouraged to present the tradeoffs prior to implementation.

7.5 Power Sequence and Power States

The Delta Lake 1S Server host system must have the capability to identify the expansion systems before turning them on. To facilitate this operation the existence of different power zones is required.

Design consideration must be considered to avoid any leakage paths among different power subsystems (e.g. Yosemite V3 platform, 1S Delta Lake Server, and expansion cards).

7.5.1 Power Sequence – Expansion Board

The power sequence for the E1S expansion board is expected to follow the power sequence for the OCP NIC. Figure 7-1 and Figure 7-2 are for illustrative purposes only. For exact timing information, refer to the OCP NIC 3.0 specification.

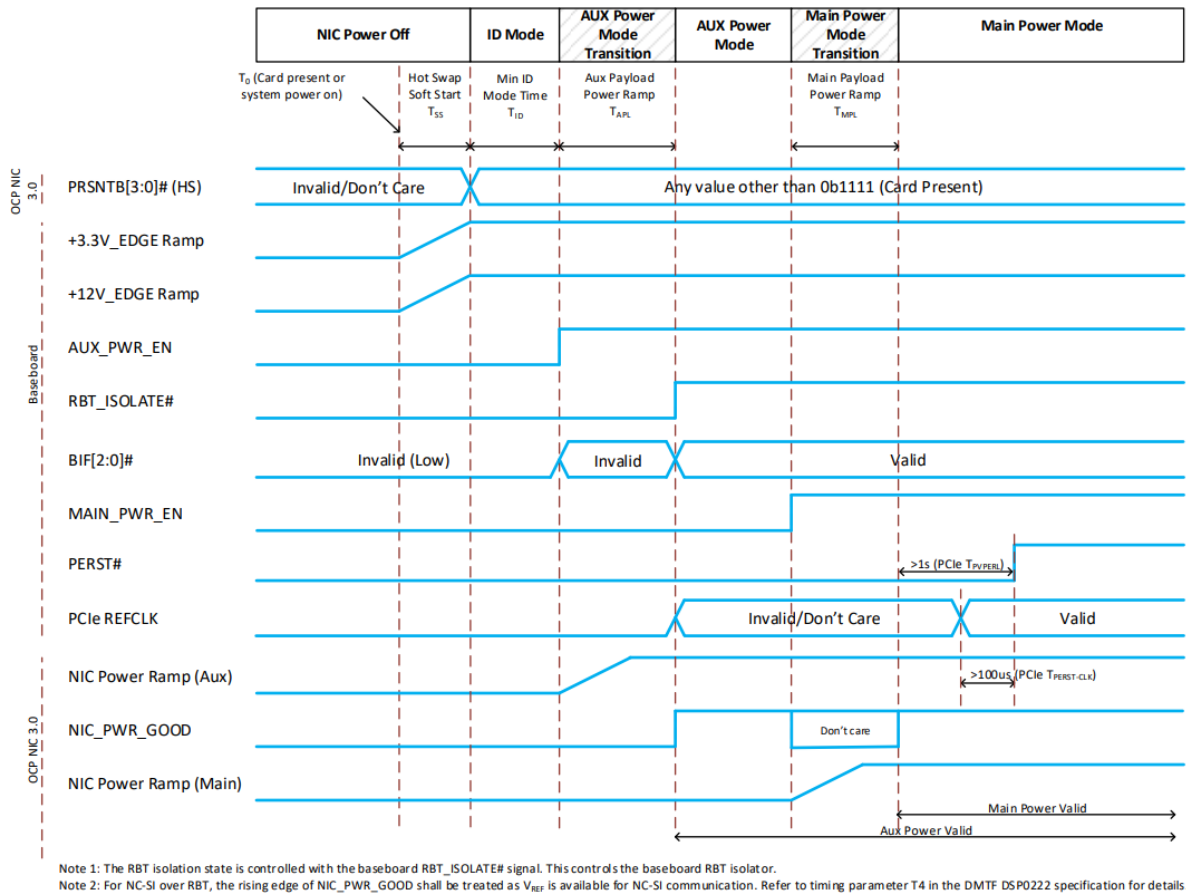


Figure 7-1 - OCP NIC 3.0 Power-on Sequence

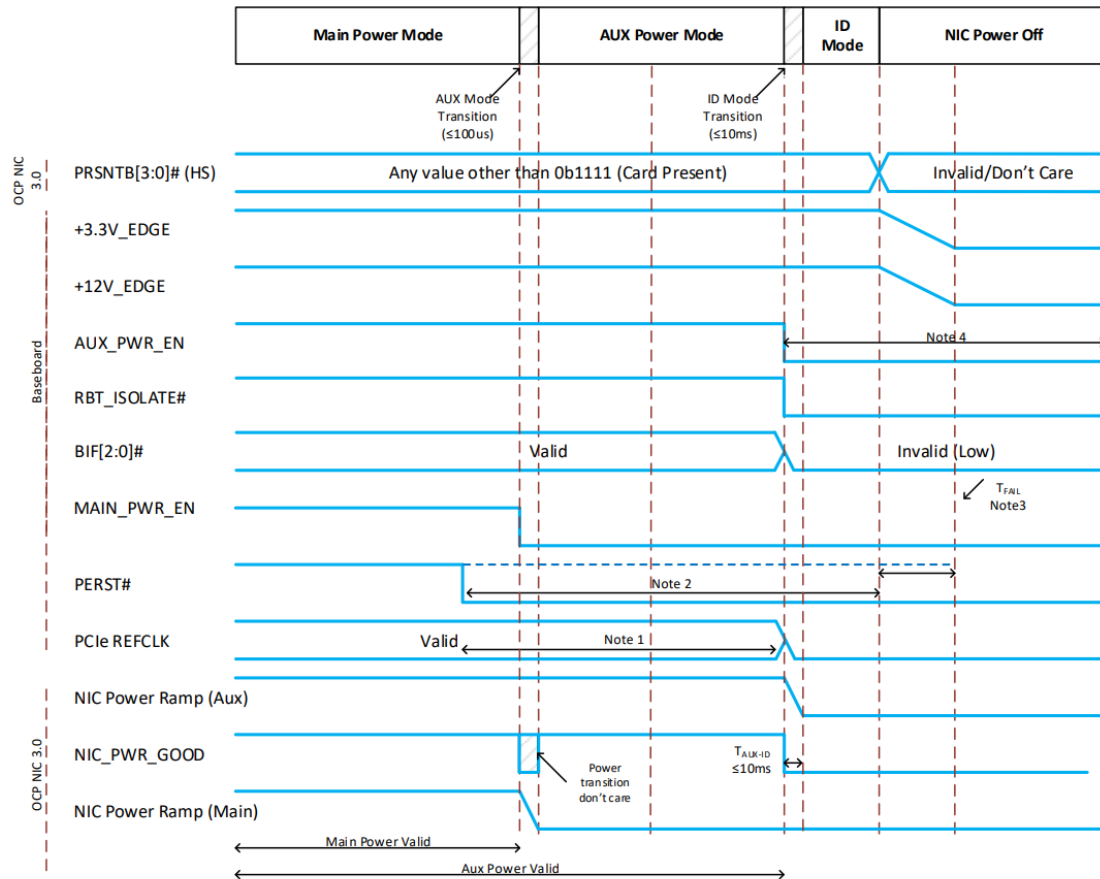


Figure 7-2 - OCP NIC 3.0 Power-down Sequence

The expansion board BIC must use the standby power provided by +3.3V_EDGE pins on the expansion board connector.

The SSDs will not sequence on until after the `NIC_PWR_GOOD` (or equivalent for expansion board) is asserted.

7.5.2 Power Sequence and Standby Power - SSD

There are no power sequencing requirements for the +3.3Vaux and +12V for the SSD.

7.6 Hot Swap Controller

Hot plug for the E1.S expansion board is not a planned feature.

In order to provide protection for the attached Delta Lake 1S Server in the event of a power failure on the expansion board, a hot swap controller is required on the +12V_EDGE input from the host.

The +3.3V_EDGE input for the expansion board must have a protection mechanism to prevent damage to the host in the event of a power failure.

All hot-swap controllers must support the following:

1. In-rush current control when the expansion board is inserted and powered up
2. MOSFETs must be kept within safe operating area during all operational conditions such as power on/off and fault conditions
3. Signals that indicate power status, alerts, interrupts must be used to allow rapid response upon impending fault conditions and/or warnings
4. Current limit protection for over current and short circuit whereby overcurrent threshold should be configured to > 2.09A to allow 25W E1.S operation.
5. Undervoltage and overvoltage protection shall be configured to support the ranges defined in **Error! Reference source not found..**
6. Default HSC response for fault conditions shall be latch off with retry as stuff option
7. PMBus interface that supports the following features:
 - a. Report voltage, current, and power (VIP) telemetry with accuracy of +/- 2.0% or better from 10% load to 100% load at room temperature.
 - b. Status registers that allow the definition of upper and lower critical thresholds for VIP of which are logged upon being triggered
8. The voltage drop across the hot-swap solution should be less than or equal to 25 mV at 100% load.

8 Hardware System Overview

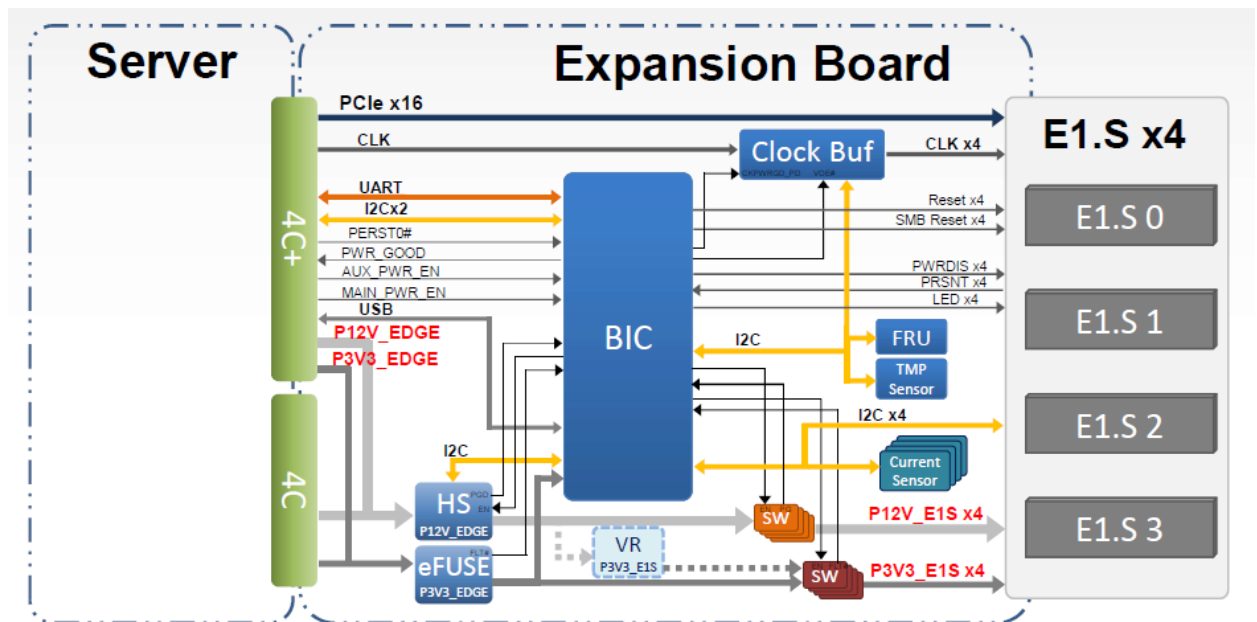


Figure 8-1 - E1.S Hardware Block Diagram

8.1 PCIe

8.1.1 PCIe Topology

The E1.S front expansion board interface has one x16 PCIe interface which will be bifurcated into four x4 PCIe interfaces. Each of these x4 interfaces will be assigned to an SSD.

8.1.2 PCIe Clock Tree

Only one instance of the PCIe reference clock is provided by Delta Lake server to the expansion board. The 100 MHz reference clock (CLKOUT_SRC_P[N]_4) is a differential reference clock provided by the PCH. It is specified to meet the PCIe 4.0 requirements for a reference clock.

The PCH clock is routed to REFCLK0 on the 4C+ interface for the expansion board. This PCIe reference must be multiplexed to the four drives on the expansion board.

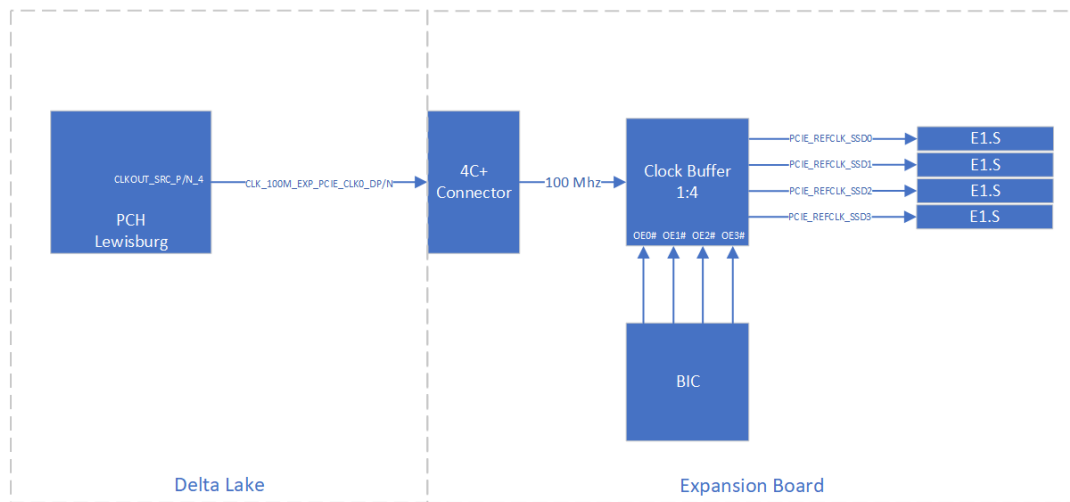


Figure 8-2

8.1.3 PCIe Reset Signals

The PCIe reset signals (PERST[0-3]_N) are driven from a CPLD on the Delta Lake system. These resets will be directly connected to the BIC which will drive the PERST pins on the SSDs as shown in Figure 9-3

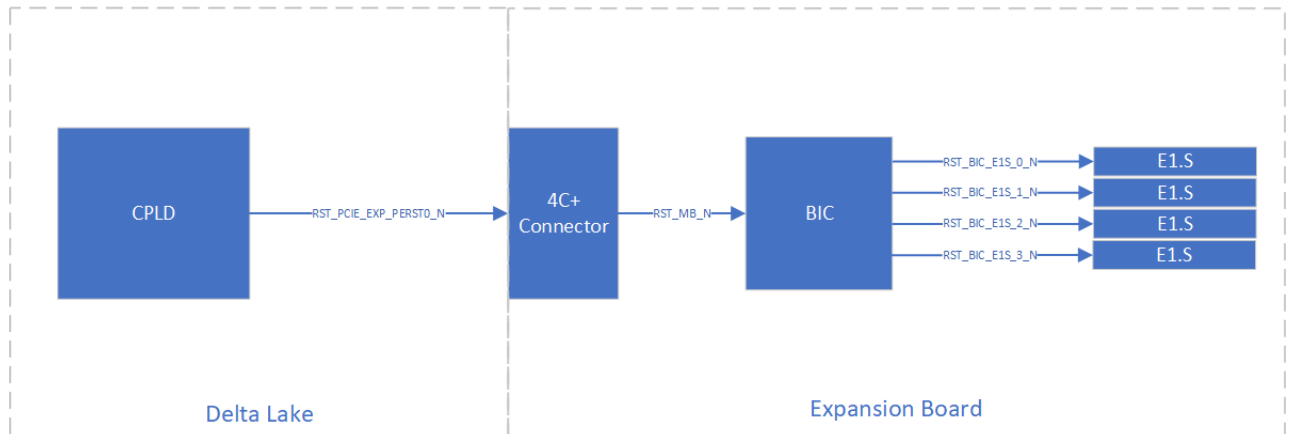


Figure 8-3

8.1.4 PCIe Lane and Polarity Reversal

The E1.S expansion card is connected to the Port 3a x16 interface of the Cooper Lake processor on the Delta Lake server.

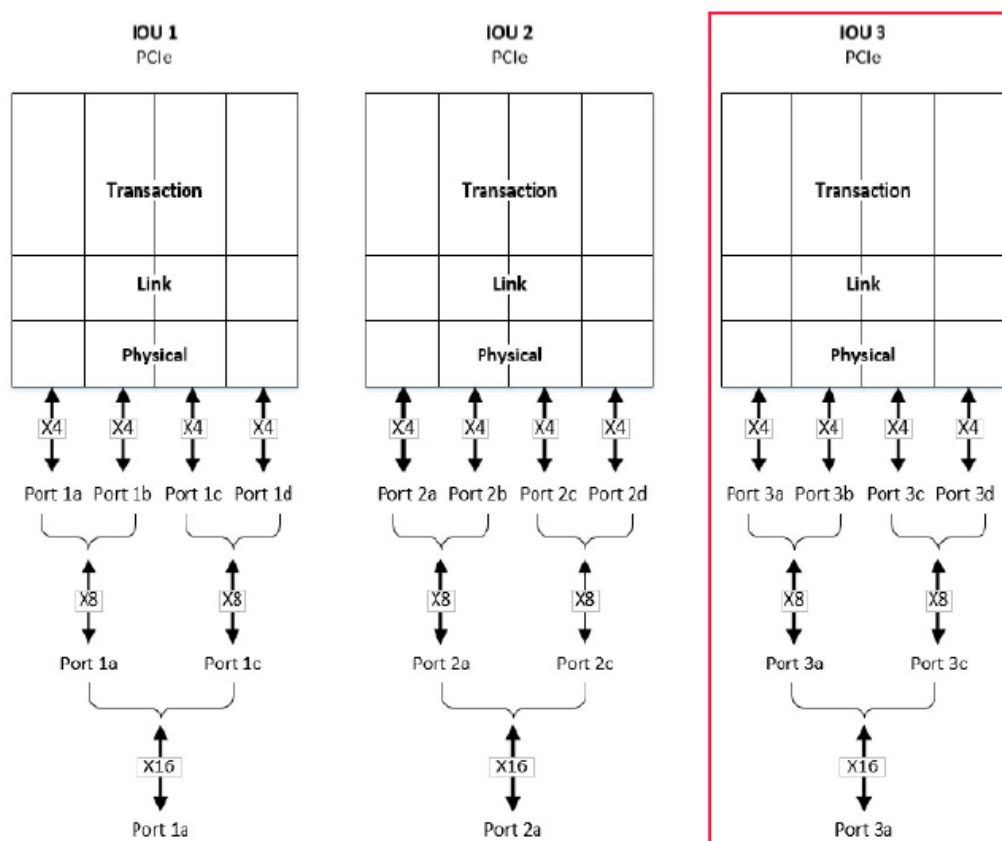


Figure 8-4 - PCI Express lane Partitioning for Cooper Lake

The port supports the following features that will be needed to support the E1.S expansion board:

- PCIe G3 (8 GT/s)
- Bifurcation of the x16 link to four x4 links
 - o Configured in BIOS
- Training of each width in lane-reversed and non-lane-reversed modes
 - o The front panel interface was lane-reversed on the Delta Lake platform to ease routing
 - o The mapping of port 3a, 3b, 3c, and 3d to front panel drive is shown in the Table 7 below while Figure 12 shows the physical location on the chassis of each drive:

Table 4

Port 3a	SSD0
Port 3b	SSD1
Port 3c	SSD2
Port 3d	SSD3



Figure 8-5

The PCIE port also supports the following features that may be useful depending on the application:

- PCIe Live Error Recovery – An advanced RAS feature that is similar in functionality to Downstream Port Containment

8.2 Clock Tree

The reference clock from the Delta Lake 1S Server should be used as an input to a clock buffer and fanned out to all four E1.S drive slots.

8.3 I2C Topology

8.3.1 Voltage Monitors

To ensure proper operation of the expansion board the voltages of all major rails should be monitored.

These monitors must have fault limits as well as a mechanism to alert the BIC or host that a voltage fault has occurred.

The monitors must be accessible over I2C.

8.3.2 Temperature Sensors

To ensure proper operation of the expansion board temperature sensors are required.

There must be an inlet sensor as well, one sensor per hotspot location.

There must be a way to configure fault limits as well as a mechanism to alert the BIC or host that a temperature fault has occurred.

The sensors must be accessible over I2C.

8.3.3 E1.S SSDs

There must be a sideband connection between the BIC and each E1.S SSD.

The controller for these interfaces must be compliant to the SMBUS protocol version 2.0 and electrically compliant to the “DC Specification for 3.3V Logic Signaling” as defined in SFF-TA-1009 revision 2.0.

8.4 LED Definitions

All front panel LEDs shall follow the guidelines defined in the Facebook Indicator Specification Version 1.0.

9 Functional

9.1 Low Speed Sideband

The expansion board must have an SMBus connection to allow host system to obtain management on the devices on board.

9.2 Debug

There should be a debug access path for the devices on the expansion board.

9.3 EEPROM

There should be an EEPROM that is accessible from the platform via the Bridge IC. The EEPROM contains the Field Replaceable Unit Identification (FRU ID) information and any additional configuration information that may be required. The FRU ID is formatted in accordance with the IPMI Platform Management FRU Information Storage Definition document.

The EEPROM must contain the following entries:

- Board Manufacturer
- Board Name
- Board Serial Number
- Board Part Number
- Product Manufacturer
- Product Name
- Product Part Number
- Product Serial Number
- Product Asset Tag
- Product Build: e.g. EVT, DVT, PVT, MP
- Product Version: e.g. C1
- Manufacturing Date and Time
- Manufacturing Lot Code: (preferred, but optional)
- Manufacturing Work Order: (preferred, but optional)
- PCB Revision
- SoC Model Name/Number
- SoC Revision
- SoC T_{jMAX} (Maximum Junction Temperature)

9.4 Expansion System Management

The system must:

- Include bridge micro-controller to manage I2C/SMBUS devices on expansion system. This includes but not limited to VR, temperature sensors, PCIE switches or clock buffers.
- Include functionality to perform complete power cycle of the expansion card and associated SSDs when system is operational.
 - The system must not crash during these events
 - Support both graceful and surprise shutdown

- Support VPP functionality for the four E1.S drives
 - ATNLED, PWREN#, PRSNT# pins are required
 - MRL#/EMILS, EMIL and BUTTON# pins are not used
- Devices with external flash must support Facebooks security requirements.

9.4.1 Expansion Bridge IC

The Bridge IC is the key component for the host to control devices on the expansion board. It will communicate with the host management device with IPMI messages. It shall run in high-speed mode with a minimum speed of 400KHz. When possible, a 1MHz or better speed is strongly recommended.

The Bridge IC on the expansion card must have at least one local I²C bus. It must have the ability to poll and/or manage devices on that bus. The bus could contain any of the following devices:

- FRU EEPROM
- Thermal sensors
- Voltage Regulators
- Clock generators
- Other I²C responder devices

The Bridge IC must have access to any devices with SMBUS on board. It must have the capability to update any program or firmware on these devices.

The generic connection diagram from the baseboard to the expansion modules on the server card is as shown below. Take note that the links shown are conceptual. Connectivity to the different devices are to have the following configurations to avoid bus lock or design errors.

1. Devices of different power state should NOT be on the same bus line. Do consider separate channels and disabling of the bus controller when it isn't in use.
2. Unique bus address for devices on a bus line
3. Bus multiplexers should be used in when a bus is to access devices of same bus address
4. Design need to cater for separate I2C bus for those connecting to connector.

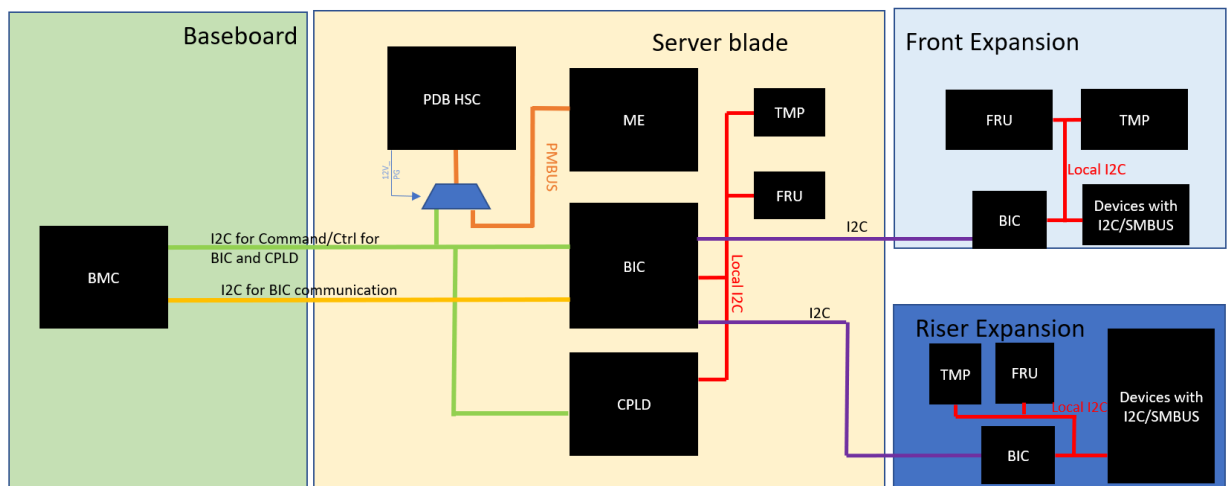


Figure 9-1: SMBus/I2C connection from baseboard to a server card with expansions.

9.4.2 GPIO Register

The Bridge IC shall provide a GPIO interface to the BMC through a GPIO register block. In this way, the BMC can control the GPIO behind the Bridge IC through accessing this register block.

The Bridge IC shall provide a way for BMC to configure GPIO pin direction, interrupt capability, and provide a way to get/set the current status of GPIO signals. It shall send an interrupt message to the BMC when the interrupt enabled GPIO signal changes its state.

9.4.2.1 Thermal Alerts

The Bridge IC must have a mechanism to provide thermal alerts and over temperature notifications to the host. The external manager must be able to receive these alerts in a timely fashion to allow it to act quickly.

9.4.2.2 Event Log

There must be an event logged whenever the devices on the expansion system fail to operate correctly. Some events include over-temperature, over-voltage, and over-current events detected.

10 Thermal

10.1 Data Center Environmental Conditions

This section outlines Facebook data center operational conditions. The expansion card shall operate in all Facebook data center conditions.

10.1.1 Location of Data Center/Altitude

Maximum altitude is 6,000 ft above sea level. Any variation of air properties or environmental difference due to the high altitude needs to be considered for the thermal design.

10.1.2 Cold-Aisle Temperature

Data centers generally maintain cold aisle temperatures between 18°C and 30°C (65°F to 85°F). The mean temperature in the cold aisle is usually 25°C with 3°C standard deviation. The cold aisle temperature in a data center may fluctuate minutely depending to the outside air temperature. Every component must be cooled and must maintain a temperature below its maximum specification temperature in the cold aisle.

10.1.3 Cold-Aisle Pressurization

Data centers generally maintain cold aisle pressure between 0 inches H₂O and 0.005 inches H₂O. The thermal solution of the system should consider the worst operational pressurization possible, which generally is 0 inches H₂O and 0.005 inches H₂O with a single fan (or rotor) failure.

10.1.4 Relative Humidity

Components are expected to handle a relative humidity between 20% and 90%.

10.2 Server Operational Conditions

10.2.1 System Volumetric Flow

The unit of airflow (or volumetric flow) used for this spec is cubic feet per minute (CFM). The CFM can be used to determine the thermal expenditure or to calculate the approximate Delta T of the system. The thermal expenditure is quantified by the metric CFM/W, which is calculated by the following formula:

$$\text{Thermal Expenditure} = \frac{\text{System airflow}}{\text{Total system power consumption, including fans}} \text{ [CFM/W]}$$

The server airflow with the E1.S expansion board installed should meet the data center's requirement. A lower CFM/watt ratio is preferred.

10.2.2 Thermal Margin

We define thermal margin as the difference between the maximum theoretical safe operating temperature and the actual temperature.

The expansion board shall be designed to operate with a server inlet temperature of 35°C (95°F) with a minimum 5% thermal margin for every component on the card.

10.2.3 Upper Critical Threshold

The upper critical threshold (UCT) setting should allow the detection of abnormal thermal behaviors in the system. The UCT values for the sensors that are not used in Fan Speed Control (FSC) should use a 15% thermal margin from the worst experiment data. The UCT values for the sensors used in FSC, except for CPU, inlet, and outlet sensors, should use a 20% thermal margin from the worst experiment data.

10.2.4 Thermal Testing

Thermal testing must be performed at a low inlet temperature of 15°C (59°F) and up to 35°C (95°F) inlet temperature to guarantee the design is free of thermal defect and has high temperature reliability.

11 Environmental Requirements

The full system with the server card installed meets the following environmental requirements:

- Gaseous contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40°C to +70°C (long-term storage) *
- Transportation temperature range: -40°C to +70°C (short-term storage) *
- Operating altitude with no de-rating to 6000 ft

11.1 Vibration and Shock

The motherboard meets shock and vibration requirements according to the following IEC specifications: IEC78-2-(*) and IEC721-3-(*) Standard & Levels.

Table 5: Vibration and Shock Requirements

	Operating	Non-Operating
Vibration	0.5G _{rms} , 5-500-5Hz, Random Vibe, 20 min, along all 3 axes (+/-) 5-20Hz – 6db/Oct 20-250Hz – 0.0007 G ² /Hz 250-500 – -6db/Oct	1.2G, 5-500-5Hz, 1 sweep at 0.5 Octave/min, along all 3 axes (+/-) 5-20Hz – 0.5G 10-350Hz – 1.2G 350-500 – 0.5G
Shock	6G, half sine, 11ms, total 6 shocks along three axes	12G, half sine, 11ms, total 6 shocks along three axes

12 Labels and Markings

The motherboard shall include the following labels on the component side of the expansion board. The labels shall not be placed in a way which may cause them to disrupt the functionality or the airflow path of the motherboard.

Description	Type	Barcode Required?
Safety Markings	Silk Screen	No
Vendor P/N, S/N, REV (Revision would increment for any approved changes)	Adhesive label	Yes
Vendor Logo, Name & Country of Origin	Silk Screen	No
PCB Vendor Logo, Name	Silk Screen	No

Date Code (Industry Standard: WEEK / YEAR)	Adhesive label	Yes
RoHS compliance	Silk Screen	No
WEEE Symbol. The expansion board will have the crossed out wheeled bin symbol to indicate that it will be taken back by the Manufacturer at the end of its useful life. This is defined in the European Union Directive 2002/96/EC of January 27, 2003 on Waste Electrical and Electronic Equipment (WEEE) and any subsequent amendments.	Silk Screen	No
CE Marking	Silkscreen	No
UL Marking	Silkscreen	No

13 Prescribed Materials

13.1 Disallowed Components

The following components are not used in the design of the motherboard:

- Components disallowed by the European Union's Restriction of Hazardous Substances Directive, **RoHS 2 Directive (2011/65/EU)**
- Trimmers and/or potentiometers
- Dip switches

13.2 Capacitors and Inductors

The following limitations apply to the use of capacitors:

- Only aluminum organic polymer capacitors made by high-quality manufacturers are used; they must be rated 105°C.
- All capacitors have a predicted life of at least 50,000 hours at 45°C inlet air temperature, under the worst conditions.
- Tantalum capacitors using manganese dioxide cathodes are forbidden.
- SMT ceramic capacitors with case size > 1206 are forbidden (size 1206 are still allowed when installed far from the PCB edge and with a correct orientation that minimizes the risk of cracking).
- X7R ceramic material for SMT capacitors should be used by default and at minimum X6S for portions of design subject to thermal hotspots such as CPU and/or DIMM cavities.
- COG or NP0 type should be used for tolerance sensitive portions of design

- Conditional usage of X5R ceramic material must be based on evaluation of worst-case thermal conditions and upon approval from Facebook

The following limitations apply to the use of inductors:

- Only SMT inductors may be used as the use of through-hole inductors is disallowed.

13.3 Component De-rating

For inductors, capacitors, and FETs, derating analysis is based on at least 20% derating.

14 Revision History

Author	Description	Revision	Date
Sean Ceballos-McGee	Initial draft – based on Delta Lake 1S Server Expansion Design Spec 0x12.	0.1	10/3/2019
Sean Ceballos-McGee	- Add requirement for BIC I2C interface - Add detail for unused pins on expansion board connectors - Updated pin notes on front panel expansion - Updated Vibration and Shock Requirements to match latest Facebook specification. - Update wording from +12V_AUXpmax to +12V_AUXpavg - Added LED visibility requirement to Section 5.2 Expansion Card Requirements - Updated Figure 8-1 - E1.S Hardware Block Diagram	0.2	11/1/2019
Matt Bowman	- Updated pin usage on page 13 to indicate SGPIO signals are unused - Added clarification that the 4C connector on Delta Lake may be depopulated, and signals from it should not be utilized - Added Vernal Falls codename to description and header	0.3	2/13/2020
Matt Bowman	Updated block diagram, stackup, and power budgets to reflect EVT design	1.0	8/12/2020
Abe Garcia	- Removed development phase specific data. - Added more contributor names - Updated drive to PCIe port mapping	1.1	1/5/2021
Abe Garcia	- Updated renderings - Added contributor names - Updated terminology to align with OCP Terminology Guidelines for Inclusion and Openness	1.2	1/25/2021
Abe Garci	- Updated formatting - Removed confidential information	1.3	3/31/2021
Ross Stenfort	Updated for publication	1.4	7/29/2021
Ross Stenfort	Additional updates for publication	1.5	8/15/2021

15 Appendix A

Requirements	Details	Link to which Section in Spec
--------------	---------	-------------------------------

Contribution License Agreement	OPTION B: Open Web Foundation (OWF) CLA	1.1OCP
Are All Contributors listed in Sec 1: License?	Yes	1.1OCP
Did All the Contributors sign the appropriate license for this spec? Final Spec Agreement/HW License?	Yes	1.1OCP
Which 3 of the 4 OCP Tenets are supported by this Spec?	Efficiency Scalability Openness Impact	▪ Efficiency Enables improved performance per Watt from previous generations ▪ Scalability Simple design enables E1.S devices to be front servicable and simplicity in servicing expansion board. ▪ Openness Yosemite V3 expansion board provides a flexible baseline which can be used with delta lake contribution or for any member of the community to mix and match to their specific needs. ▪ Impact This contribution enables 1OU flash server with E1.S flash devices. This has enables improved thermals, servicabiltiy

E1.S 10U Flash Blade and Expansion Board Design Specification – Vernal Falls

		and performance.
Is there a Supplier(s) that is building a product based on this Spec?	Yes	
Will Supplier(s) have the product available for GENERAL AVAILABILITY within 120 days?		