Advanced Accelerator Adapter

Electro-Mechanical Specification

Workgroup Specification

Revision 1.0 (November 8, 2017)

OpenPOWER™

www.openpowerfoundation.org

Advanced Accelerator Adapter: Electro-Mechanical Specification

25G IO Interoperability Mode Work Group <25giomode-chair@openpowerfoundation.org>
OpenPower Foundation

Revision 1.0 (November 8, 2017)

Abstract

This document defines an electro-mechanical specification for advanced accelerator adapters within the OpenPOWER eco-system supported by IBM® POWER9™. POWER9 modules present a 25Gbps interface organized into groups of 8 bit-lanes. Innovation within the community is encourgaged for systems that support accelerated computing and the accelerator adapters needed to make heterogeneous / accelerated computing solutions available to the market.

This document is a Standard Track, Workgroup Specification work product owned by the 25G IO Interoperability Compatability Workgroup and handled in compliance with the requirements outlined in the *OpenPOWER Foundation Work Group (WG) Process* document. It was created using the *Document Development Guide* version 1.1.0. Comments, questions, etc. can be submitted to the public mailing list for this document at <25giomode-p9_25gbps_phy@mailinglist.openpowerfoundation.org>.

Table of Contents

List of Figures

List of Tables

Preface

1. Conventions

The OpenPOWER Foundation documentation uses several typesetting conventions.

Notices

Notices take these forms:

Note

A handy tip or reminder.

Important

Something you must be aware of before proceeding.

Warning

Critical information about the risk of data loss or security issues.

Changes

At certain points in the document lifecycle, knowing what changed in a document is important. In these situations, the following conventions will used.

- *New text will appear like this.* Text marked in this way is completely new.
- ~~Deleted text will appear like this.~~ Text marked in this way was removed from the previous version and will not appear in the final, published document.
- Changed text will appear like this. Text marked in this way appeared in previous versions but has been modified.

Command prompts

In general, examples use commands from the Linux operating system. Many of these are also common with Mac OS, but may differ greatly from the Windows operating system equivalents.

For the Linux-based commands referenced, the following conventions will be followed:

$ prompt Any user, including the `root` user, can run commands that are prefixed with the $ prompt.

 # prompt The `root` user must run commands that are prefixed with the # prompt. You can also prefix these commands with the **sudo** command, if available, to run them.

Document links

Document links frequently appear throughout the documents. Generally, these links include a text for the link, followed by a page number in parenthesis. For example, this link, Preface [vii], references the Preface chapter on page vii.

2. Document change history

This version of the guide replaces and obsoletes all earlier versions.

The following table describes the most recent changes:

Revision Date	Summary of Changes
November 8, 2017	• Version 1.0 - Final WG Specification
October 19, 2017	• (pre-8) Updates from Public Review • Modified mezzanine card section 2.1,2., and 2.3 to reflect a reference design and connector PN • Added abstract to point to the OIF spec and appropriate copyright languange
August 17, 2017	• Public Review Draft - approved by 25GIO WG
August 8, 2017	• (pre-7) Added section on Conformance, additional minor edits, removed Section 6.2 Mezzanine Net Model
August 2, 2017	• (pre-6) Updates to Table 5.1 about ILD, ILDB, and "moving average smoothing"
August 1, 2017	• (pre-5) Minor updated from workgroup review of pre-4
July 20, 2017	• (pre-4) Major update from workgroup review and comments
June 27, 2017	• (pre-3) Added part on cable extender version
April 27, 2017	• (pre-2) Updates from Dan Dreps review • Editorial updates from Jeff Brown review
March 13, 2017	• (pre-1) Creation of Electo-Mechanical consolidated specification. Merging 25gio content with the mezzanine card content

1. Introduction

1.1. Scope

This document defines an electro-mechanical specification for advanced accelerator adapters within the OpenPOWER eco-system supported by IBM® POWER9™. POWER9 modules present a 25Gbps interface organized into groups of 8 bit-lanes. Innovation within the community is encouraged for systems that support accelerated computing and the accelerator adapters needed to make heterogeneous / accelerated computing solutions available to the market.

This document defines two accelerator approaches. The first approach is a mezzanine card attached to the system planar via two connectors. This approach is defined in Part I, “Mezzanine Adapter Card” [2]. The second approach is via cable and is defined in Part II, “Cabled Interface Extension” [24]. While the accelerator card form factor is not defined in this specification Part II, “Cabled Interface Extension” [24] assumes a PCIe® card for illustration.

The electrical characteristics of the 25Gbps channel are defined in Part III, “25 Gbit/sec Electrical Channel” [38].

1.2. Conformance to this Specification

Mezzanine adapter cards and systems designed to accept them must satisfy the requirements defined in Part I, “Mezzanine Adapter Card” [2].

Cabled adapter cards and systems designed to accept them must satisfy the requirements defined in Part II, “Cabled Interface Extension” [24].

Systems, adapter cards, and the chip/modules that provide the channel endpoints must satisfy the requirements defined in Part III, “25 Gbit/sec Electrical Channel” [38].

Part I. Mezzanine Adapter Card

This part describes the mezzanine adapter card for OpenPOWER systems based on the POWER9™ processor. These adapter cards attach to the 25 gbit/sec interface native to the POWER9 and plug into the mezzanine card connectors.

This specification describes the system level details required of mezzanine add-in cards. This specification is constrained to the use of the tested and verified connector technology and IO assignments. It also includes an overview of the mechanical requirements including system level airflow and pressure drop requirements.

2. System Mechanical Details

Table of Contents

This section describes the mechanical constraints and details of designing the Advanced Accelerator Adapter add in card.

2.1. Mezzanine Card Connectors

Adapter add-in cards use high-speed array connectors as the electrical interface to the system planar. The connector is a 400 pin, 4mm stack height, high-speed interconnect. The reference system planar uses the plug connector (FCI PN 84740-102LF). The adapter add-in card is required to use a compatible receptacle connector. An example/reference connector known to be compatible and demonstrated in this application is FCI PN 74221-101LF (see Figure 2.1, “Reference System Planar Connector (MEG-Array® 84740-102LF)” [4]. Each add-in card co-docks two connectors, one of which is primarily for high speed signals and the other for power and lower speed signals. Co-docking these connectors drive the specific PCB tolerances described in Section 2.2, “Adapter Mezzanine Card Outline” [4] . The low nominal pin wipe in the connectors drives specific mechanical tolerances on the bottom side stiffener, detailed in Section 2.3, “Mezzanine Card Mechanical and Attachment Requirements” [9] .

Figure 2.1. Reference System Planar Connector (MEG-Array® 84740-102LF)

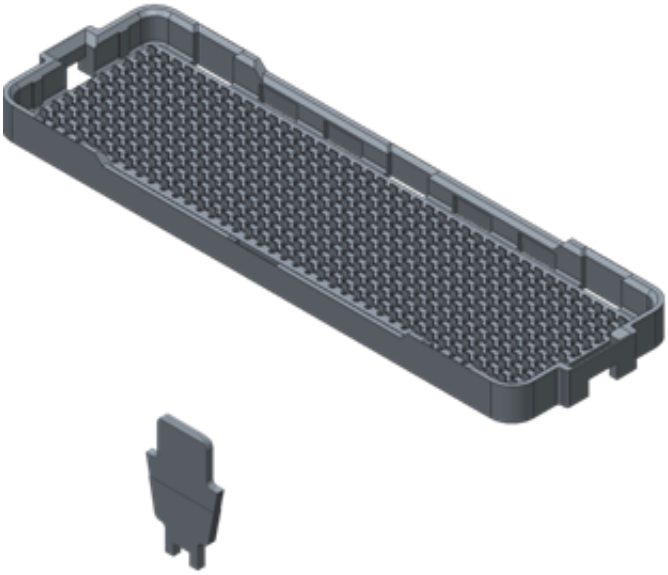
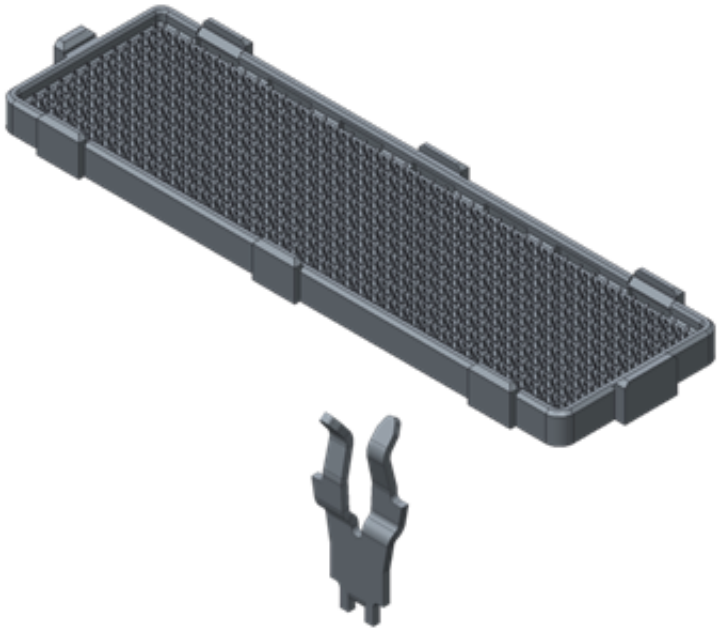


Figure 2.2. Reference Mezzanine Card Plug connector (MEG-Array® 74221-101LF)



2.2. Adapter Mezzanine Card Outline

The top view of the mezzanine card outline is shown in Figure 2.3, “Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters” [6] and a bottom view of the mezzanine card is shown in Figure 2.4, “Bottom view of the reference mezzanine

card with datum identification, connector placement and pin identification” [7] . The maximum add-in card planar dimensions are 78mm x 140mm. For orientation in the system, the north side of the card is identified on the drawing. Of particular note is the alignment hole tolerance. Two alignment pins are used to orient the card within the system. The origin pin (south side) drives the alignment and docking of the card to the planar. This feature is critical for proper alignment of the add-in card and the system planar. The north side alignment pin is used to provide angular alignment and may be slotted on the card for hole positional tolerance.

The dimensions for mounting holes are depicted in Figure 2.3, “Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters” [6]. The alignment holes and pins A1 and K40 are depicted in Figure 2.3, “Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters” [6]. The orientation and connector identification (1 and 2) is also defined in this view as well as the location of two copper grounding pads located between the connector pair. Additional details for add-in card grounding requirements are detailed in Section 2.3, “Mezzanine Card Mechanical and Attachment Requirements” [9]. The dimensions and tolerances for the connector pin holes and alignment holes are required to be replicated to add-in cards in order to ensure proper co-docking.

Figure 2.3. Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters

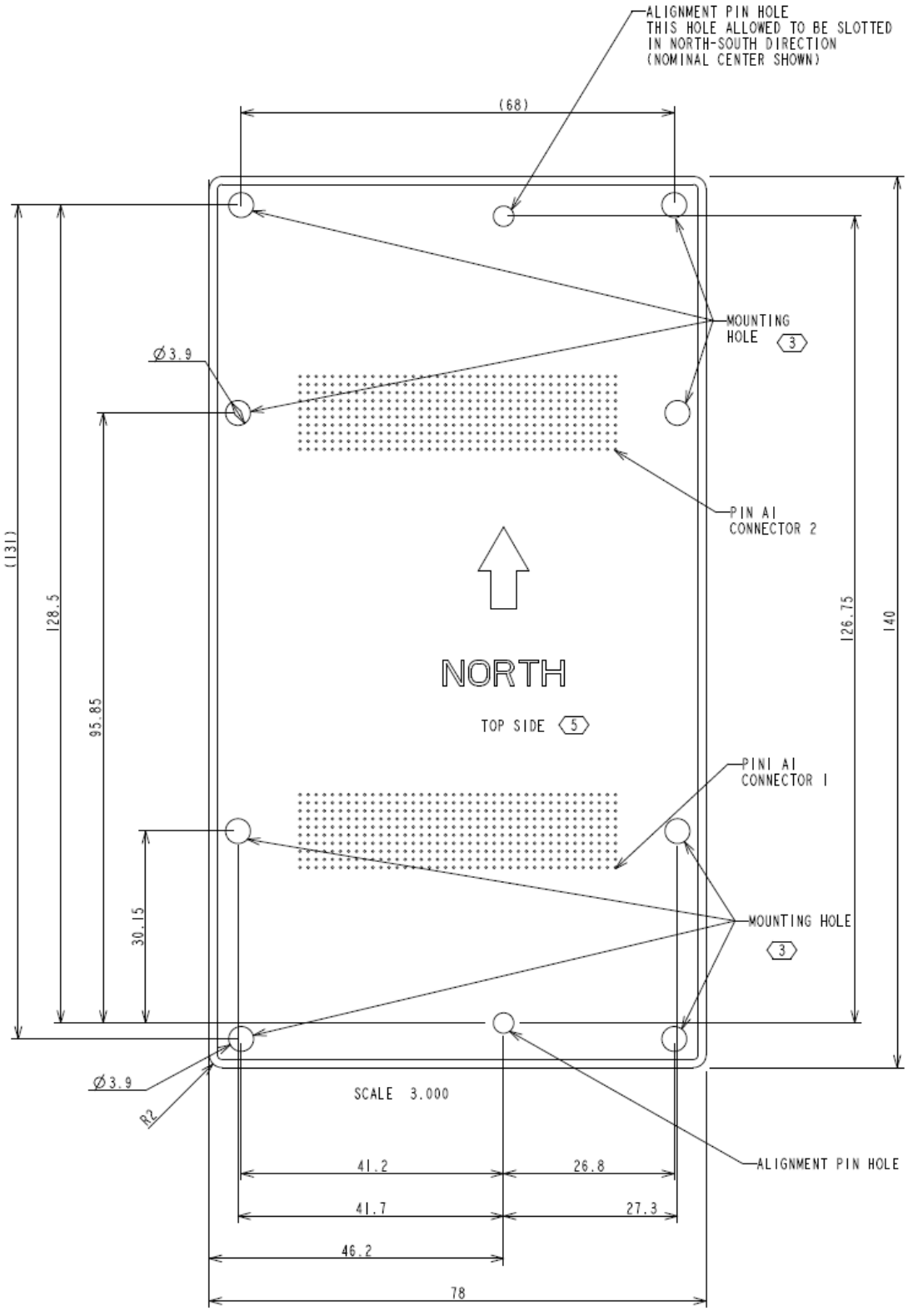
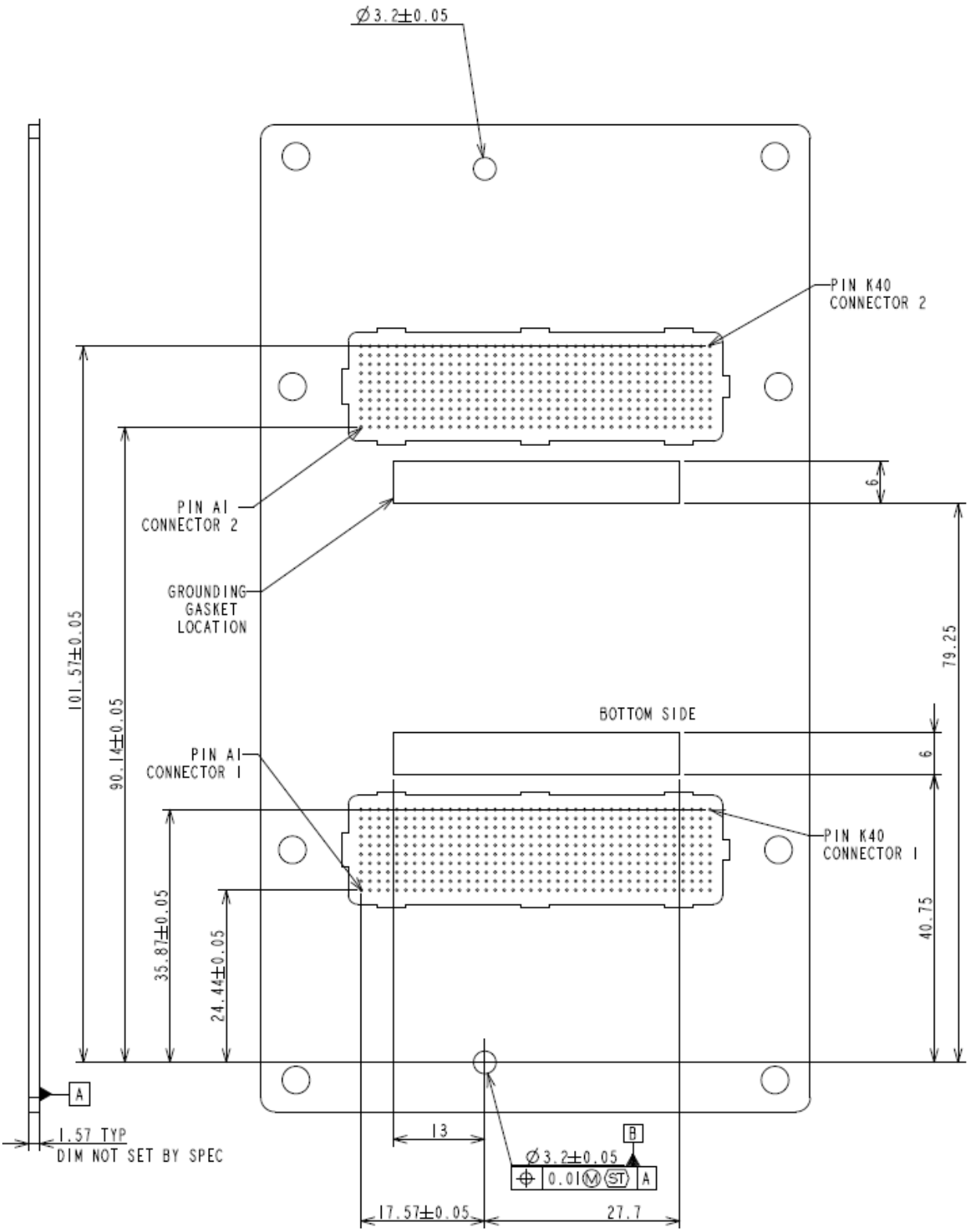


Figure 2.4. Bottom view of the reference mezzanine card with datum identification, connector placement and pin identification



Maximum component height on the top side of the card is approximately 66mm. The total height of adapter cards must be less than 71.75mm. It is preferred that the card extend the full height (including the heat sink), or provide airflow blocking features. This is covered in the Section 2.5, “Adapter Thermal Requirements” [13].

Figure 2.5. Mezzanine Card Height restriction

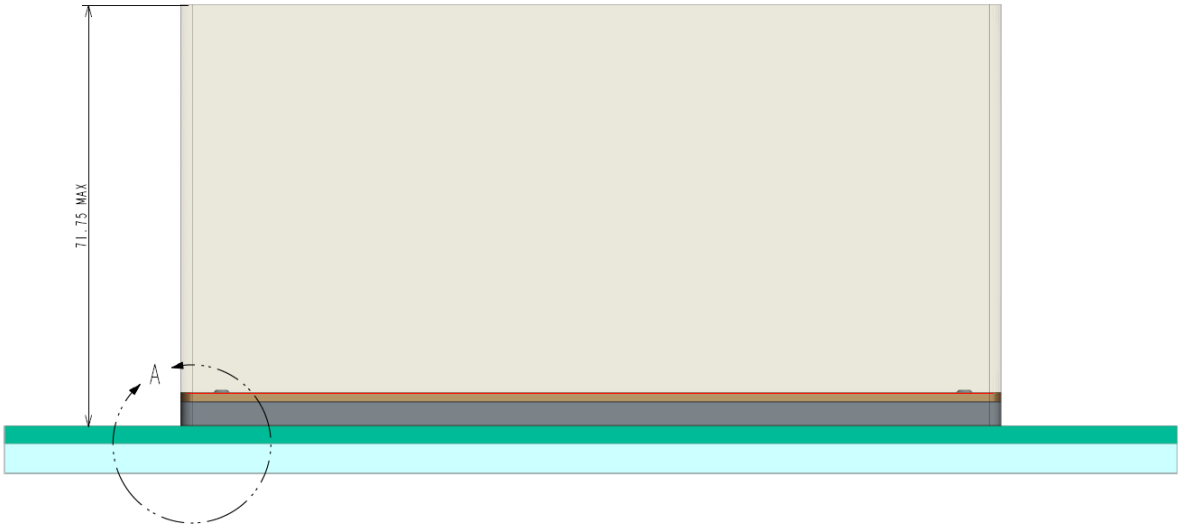
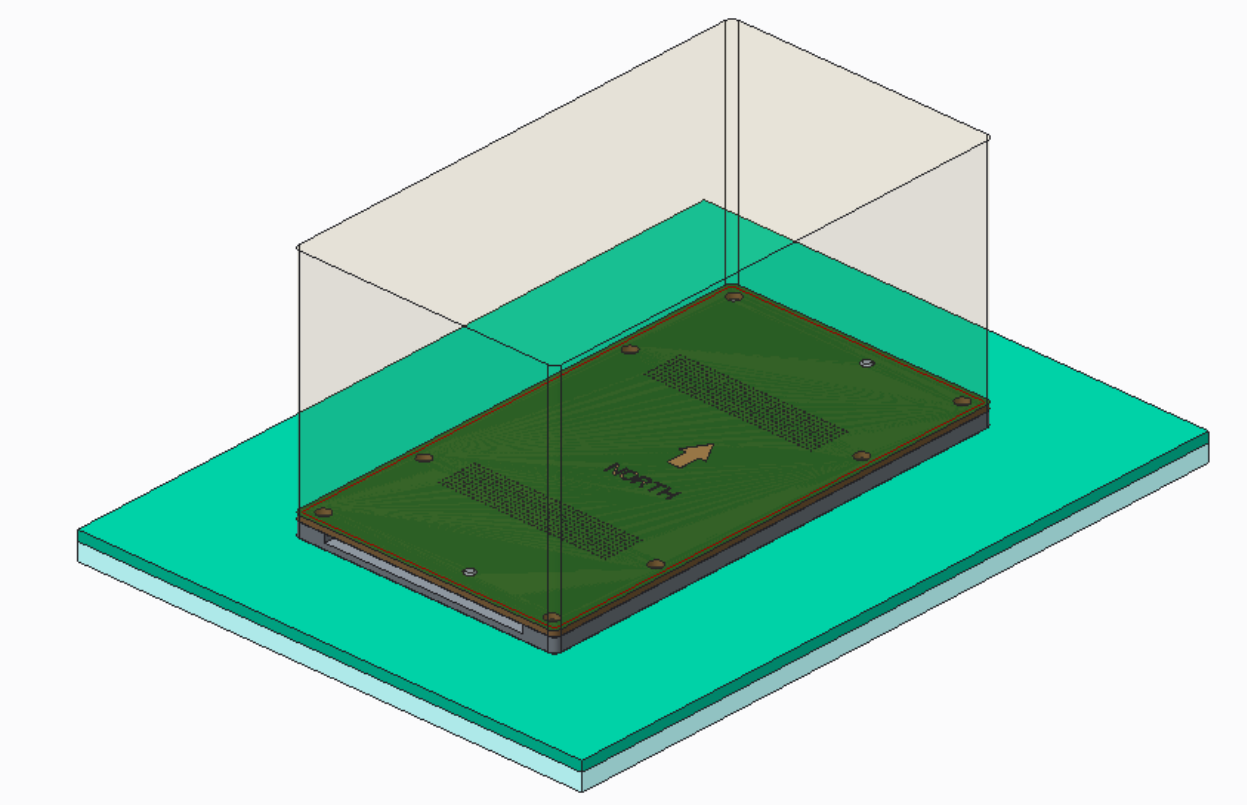


Figure 2.6. Mezzanine Card Envelope



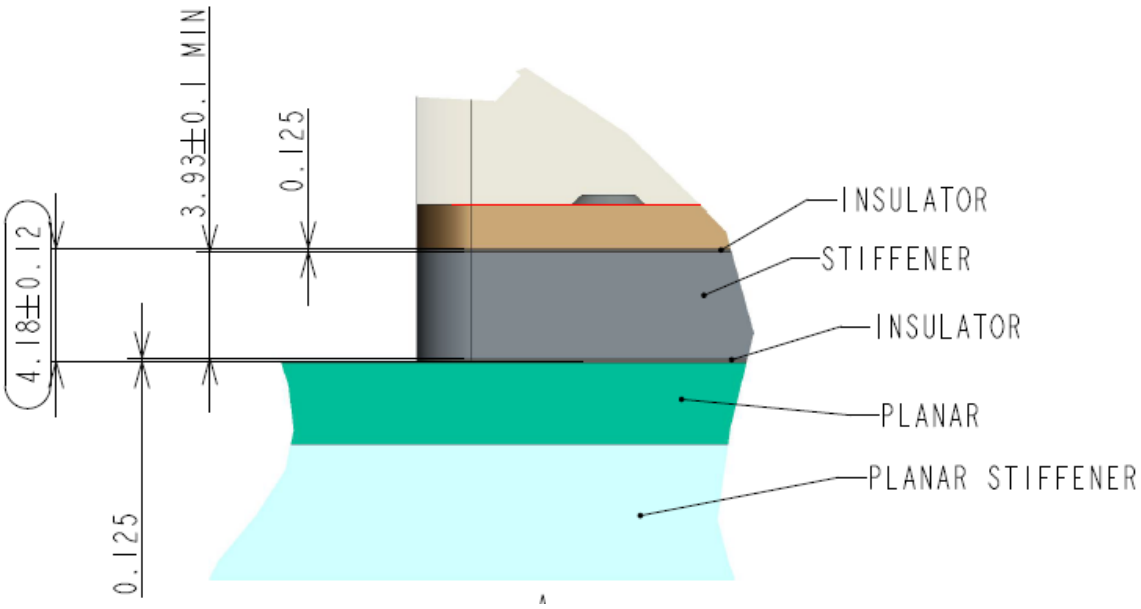
2.3. Mezzanine Card Mechanical and Attachment Requirements

A stiffener is required on the bottom side of the add-in cards to ensure that the connectors are appropriately loaded. The stiffener requires a minimum thickness of 3.93 +/- 0.1 mm with a minimum insulator thickness of 0.125 mm on the bottom of the stiffener. For the reference card design the total thickness of stiffener and insulators (bottom and top insulators with adhesive) is to be 4.18 +/- .12 mm to ensure proper mating of the MEG-Array connectors.

Note

This stack-up has been statistically analyzed to maximize pin wipe of the MEG-Array connectors but also directs clamp loads through the stiffener and not the MEG-Array connectors. See Figure 2.7, “Reference Card Stiffener and Insulator Detail” [9].

Figure 2.7. Reference Card Stiffener and Insulator Detail



The stiffener is required to be full thickness around the mounting locations, but can be less than full thickness in other locations to allow for components to be placed on the bottom side of the mezzanine card. The design of the stiffener should allow full force insertion of the connectors (for example, at ~ 31lbf per the reference MEG-Array connector) and maintain integrity of the BGA grid on the connector.

Note

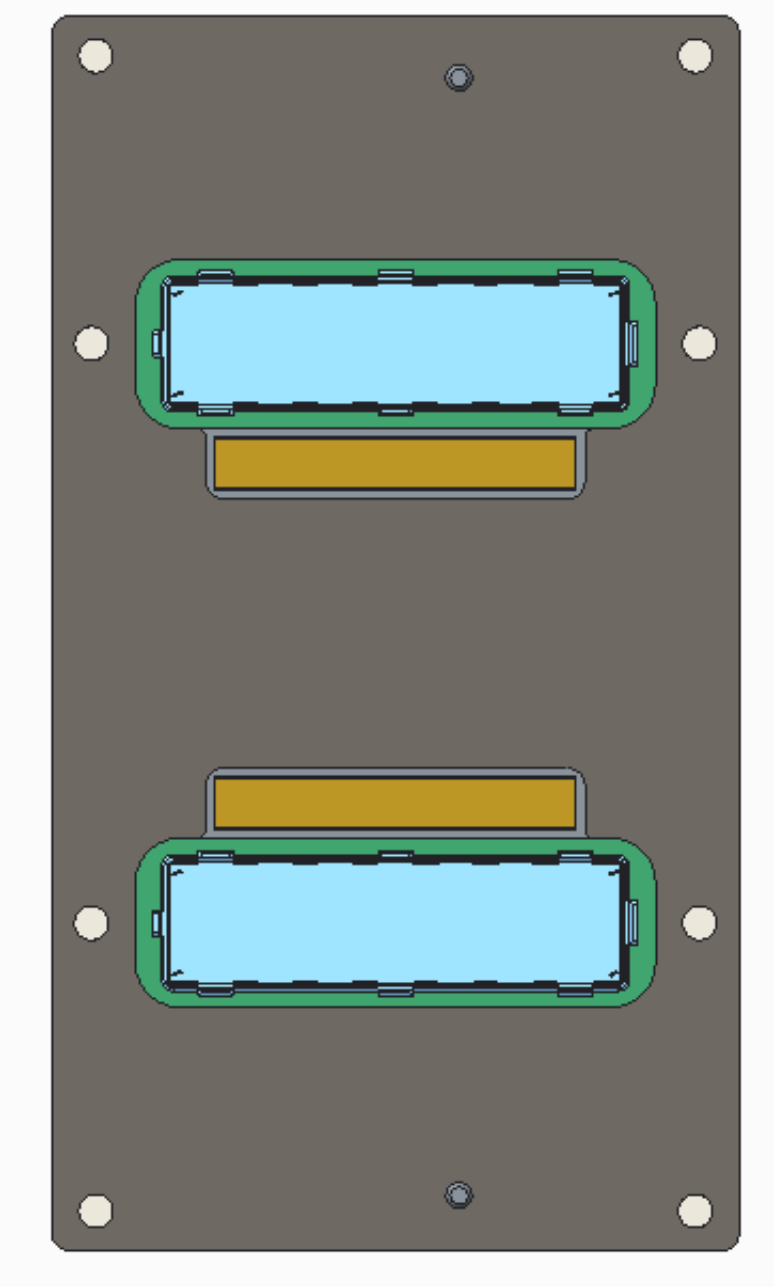
Care should be taken when designing bottom side stiffener (as well as top side stiffener if required) to minimize board strain in the area around the connector BGA to reduce damage to the connector and solder joints during card insertion and extraction.

The stiffener is required to include two conductive fabric over foam gaskets to touch the copper grounding pads dimensioned in Figure 2.3, “Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters” [6]; Schlegel profile E1B with a nominal compression of 0.265 mm is preferred (gasket attach plane would be nominally inset 0.35mm from the bottom stiffener plane to achieve this compression). This fabric over foam gaskets should also electrically connect to the adapter card bottom stiffener and provide a ground path the card. This is chassis ground.

The System planar should have a package keep-out for the full 78 mm x 140 mm on the top side on the main planar. However, top side traces are allowed in this region. Add-in cards must insure that areas of the adapter card that are in contact with the main planar are electrically isolated (with the exception of the conductive foam gaskets). It is recommended that a thin insulator (.125mm) be use for this purpose. Non-conductive coatings may be used, but may not be consider the primary method of isolation and damage mitigation to the planar card.

Attachment to the planar is by 8 threaded fasteners at the mounting locations noted in figure 2 (and included in the appendix drawing). These 8 mounting locations are defined as M3 x 0.5 and the threaded depth is 5 mm. accounting for a planar thickness of 3.01mm, the maximum threaded depth beyond the bottom surface to the adapter card is 8mm. An exposed threaded fastener with a length between 6.5 and 7.5mm from bottom of the adapter card should be used.

Figure 2.8. Bottom view ReferenceMezzanine Card with Stiffener



Removing the mezzanine card from the system planar requires an upward force to un-mate the connectors.

2.4. Operating conditions

Mechanical qualification (operational and non-operational) for adapter add-in cards should be performed at a system level. The suggested environmental product specifications are presented below for reference.

2.4.1. Runtime Specification

- Useful operational life of 5 years
- Maximum of 8,760 power-on-hours per year (typical 8,640)
- Nominal ambient input air temperature of 25C
- Ambient temperature range of 5C to 40C
- Relative humidity range of 8% to 85%
- Air Quality is that of a typical business office (Class G1 & P1)
- Altitude range of 0 to 3050 meters (see dry bulb temperature degrade table for details of max altitude at given ambient temperature)
- Nominal AC power input
- Nominal usage of 100% during the power-on-hours for electronics
- No preventive maintenance

2.4.2. Temperature, Humidity and Altitude

This section defines the storage, shipping and operational atmospheric environmental limitation for which the product should handle without physical and functional degradation. The system units should conform to these requirements in both horizontal and vertical configurations. The system units should comply with the classifications detailed in Table 2.1, “Enviromental Specifications” [12].

Table 2.1. Enviromental Specifications

ENVIROMENTAL CLASS		Temperature (Dry Bulb)	Altitude	Relative Humidity
ASHRAE Class A3	Operating	5C - 40C	Up to 950 meters Uper limit is derated 1C for ever 175 meters increase in altitude	8% - 80%
	Power - off	5C - 45C		5% - 80%
	Storage	-40C - 60C		5% - 100%

2.4.3. Shock and Vibration

Table 2.2. Shock and Vibration Specification

Item	Test Specification	Test Configuration	Test Level	Test Criteria
Packaging test	Packaged System Products, Testing for Shipment, Test Levels, and Procedures.	Full-load system (with designed package)	Sine Sweep 0.5g for 30 minutes all there axes Random Vibration 1.04grms all there axes Drop test (drop 18 inch for 6 faces) Horizontal impact 1m/s all 4 sides	No physical and functional damage to the system unit.
Operational Vibration	Vibration Levels for System Products; Product Environments, Product Classes.	Loaded in a rack		System functional

2.5. Adapter Thermal Requirements

This thermal section will describe the minimum requirements for a single add-in card site, but is applicable to multiple sites. Due to slight variance in the airflow within the system, the site with the minimum requirements will be described in this section.

The expected boundary condition for the adapter site is for airflow to enter the front (card south side) on the site and is exhausted toward the rear of the system (card north side). Airflow is not expected to exit the boundary to the left or right and the mounted heatsink solution must ensure that a does not happen as it may interfere with the operation of the system.

Table 2.3, “Boundary Conditions” [13] summarizes the inlet and outlet conditions. Depending on the system planar layout, the adapter site or sites could be directly behind the processors. The slots must accept pre-heated air under the conditions for operation detailed in Table 2.3, “Boundary Conditions” [13].

Table 2.3. Boundary Conditions

Condition	Average Temperature	Airflow Nominal	Notes
Nominal	38C	30 CFM	25C room
Low Flow	TBD	11 CFM	Minimum flow expected in system
High Temperature	50 C	50 CFM	

Air flow distribution across the system is accomplished by managing the impedance of the processor and the adapter heatsinks. It is important that the adapter card adheres to the impedance requirements described in this section. In order to maintain proper airflow within the system, the cooling solution must restrict the flow such that all of the air passes through the heatsink and does not provide an airflow bypass path above or around the heatsink. The total pressure drop across the heatsink must include the entire area as outlined in Figure 2.10, “Impedance Area” [14]. The allowable range of heatsink impedance is shown in Figure 2.9, “Allowable Heatsink Impedance” [14], which outlines the upper and lower bounds of the heatsink impedance. The bounds help ensure adequate airflow balancing for all components that may go into the adapter slots.

Figure 2.9. Allowable Heatsink Impedance

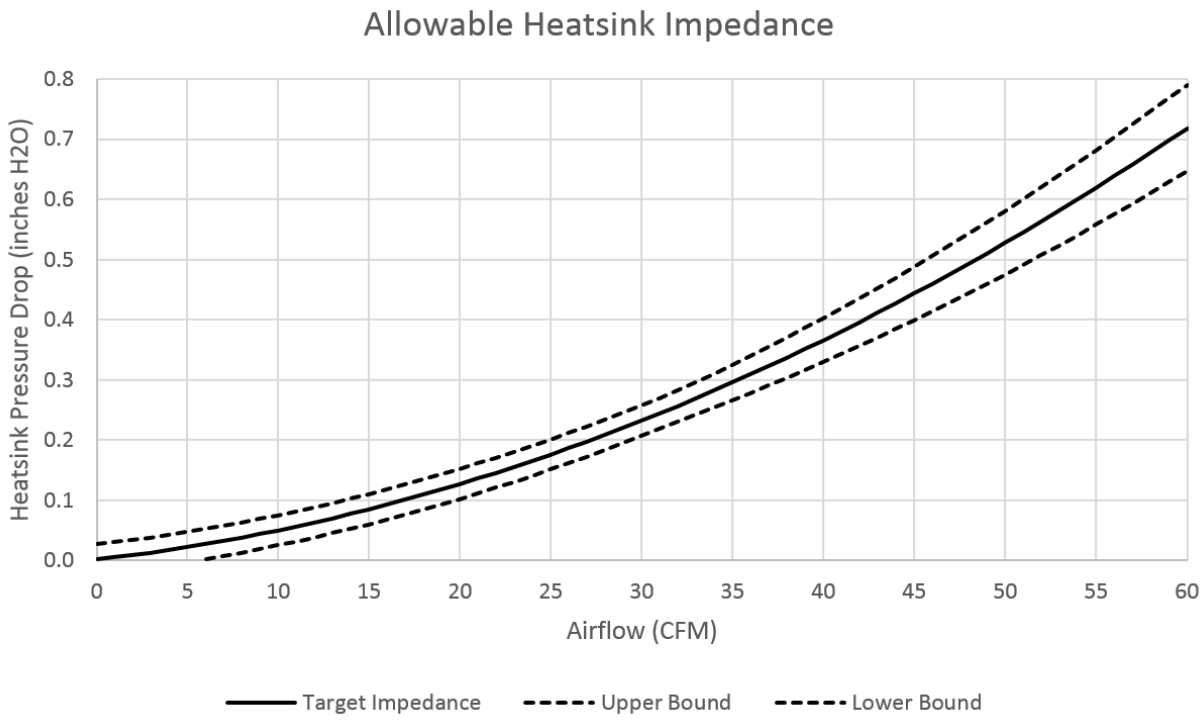
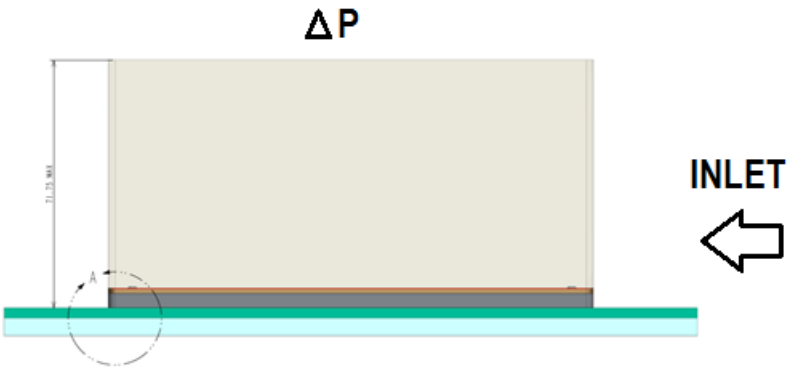


Figure 2.10. Impedance Area



3. System Electrical Details

Table of Contents

This section details the power delivery capability and pinout of the connectors labeled 1 and 2 in Figure 2.3, “Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters” [6].

3.1. Power Delivery Capability

Table 3.1, “System power delivery capability per add-in card (set of 2 connectors)” [15] details the typical power delivery capability of the system as a function of voltage domain per add-in card. All amperages are assumed to be equally distributed amongst all pins on that domain.

Table 3.1. System power delivery capability per add-in card (set of 2 connectors)

Voltage Domain	Amperage	Allowable Excursion
12 V	25 A	30% for <100ms
5 V	1 A	30% for <100ms

The system capacitance is 10uF for 5V without an inductor requirement.

The system capacitance is 10uF for 12V without an inductor requirement.

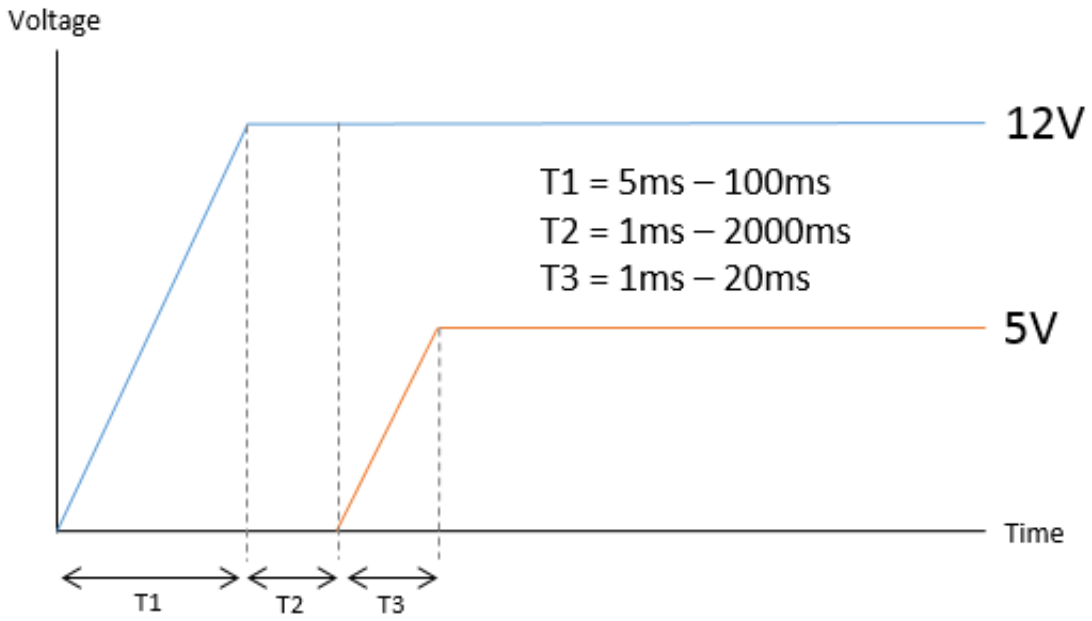
The capacitance recommendations for the adapter card are as follows:

- The maximum capacitance for 12V on the adapter card shall not exceed 1mF. (Required to meet the power initialization timing requirements.)
- The maximum capacitance for 5V on the adapter card card shall not exceed 300uF.

3.2. Power Initialization Timing

The power initialization timing from the planar to the mezzanine card is detailed in Figure 3.1, “Mezzanine OP Card Power Initialization” [16]. As can be seen, 12V ramps to max over a duration of 20 to 100 ms, followed by a timeout of 1 to 2000 ms, after which 5V ramps to max over a duration of 1 to 20 ms. PWR_EN rises after 12V and 5V are stable typically 1 millisecond after T3 shown in Figure 3.1, “Mezzanine OP Card Power Initialization” [16].

Figure 3.1. Mezzanine OP Card Power Initialization

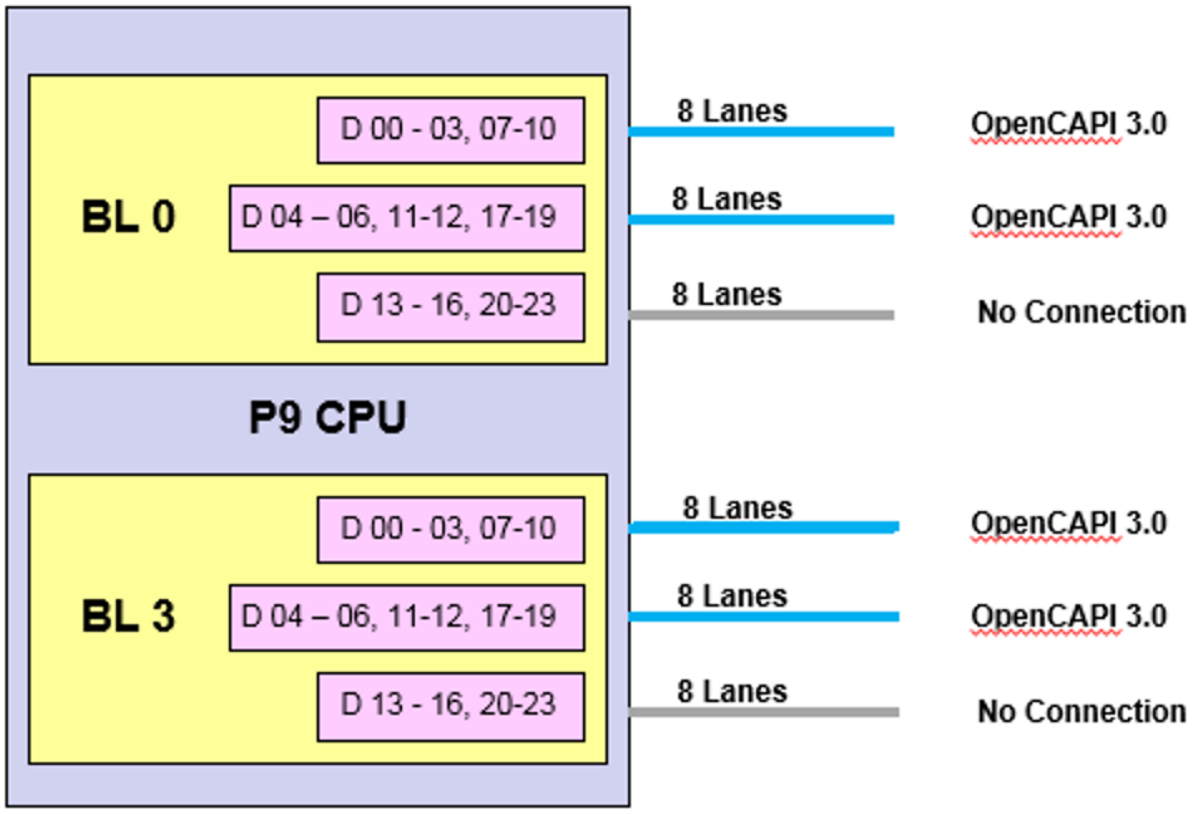


3.3. Power Down Timing

The adapter card is powered down by de-asserting the PWR_EN signal. The power rails from the system remain on during this time. If the system turns off the adapter device power, it is assumed that the devices internal power rails stays powered on for 1 millisecond after PWR_EN de-assertion.

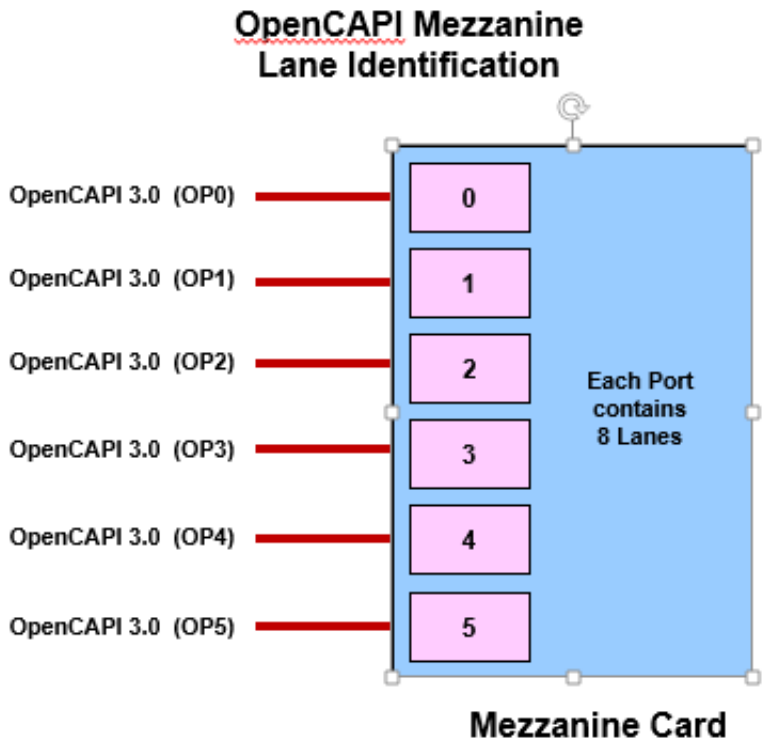
3.4. Adapter device wiring for OpenCAPI 3.0

Figure 3.2. POWER 9 CPU OpenCAPI Lanes



The above diagram illustrates one POWER9 CPU module. There are 32 available lanes for OpenCAPI use. The lanes are divided into four 8 Lane groups. This diagram also details the two 8 lane groups that are not compatible with OpenCAPI.

Figure 3.3. Mezzanine Card Lane Identification



The Mezzanine Card pinout is pre-defined and contains six 8-lane ports (OP0-OP5). These pin assignments are defined in Section 3.6, “Connector 1 Pinout” [20]. The implementation and assignment of the available ports is determined by the developer’s needs and requirements.

Potential use cases would be:

1. Developing the Mezzanine card for a pre-defined backplane. Care must be taken to ensure that the chosen ports and pin assignments of the Mezzanine card match the backplane wiring. The chosen lanes must be attached to the valid OpenCAPI 3.0 lanes from the Power 9 planar.

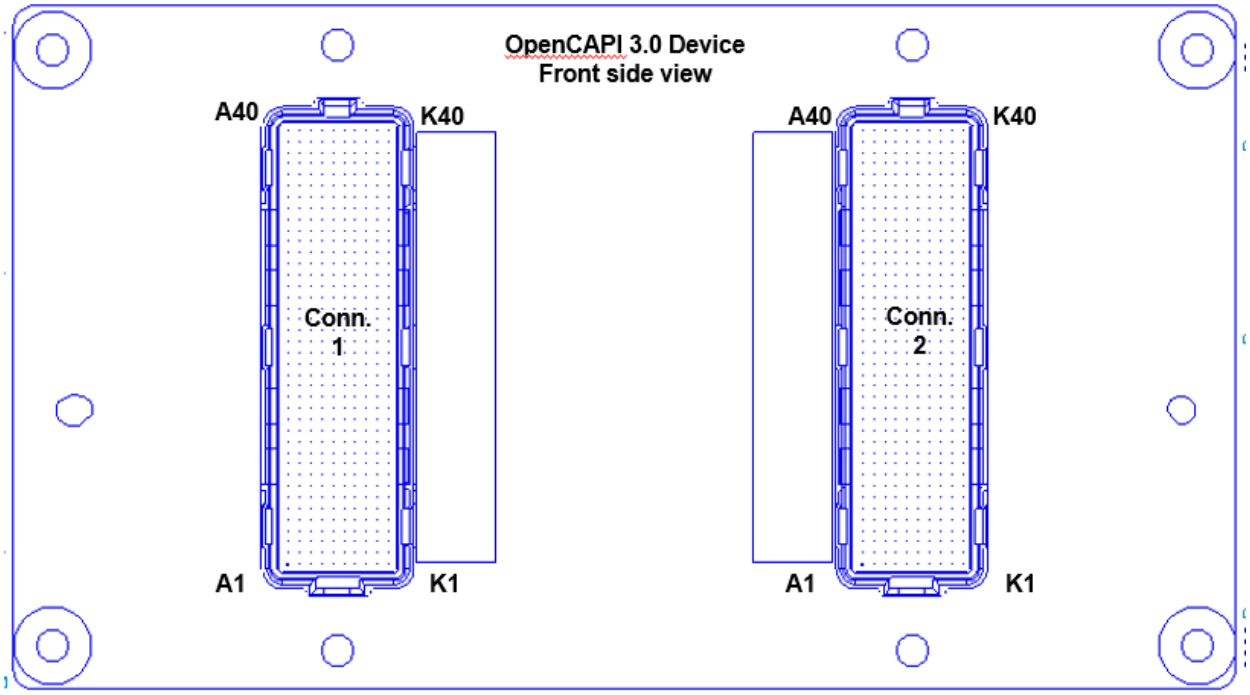
2. Developing the Mezzanine card and the backplane. The developer has more freedom in this case to maximize the lane use and backplane wiring. The port use would be determined by the developer’s application which would potentially factor in system planar component layout, number of CPU’s and wiring plane restrictions. The chosen lanes must attach to the valid OpenCAPI 3.0 lanes from the selected Power 9 module.

Note

There are 6 lanes groups on the Mezzanine card and 4 lane groups from the Power 9. The developer’s use of each group is determined by their application.

3.5. Adapter Physical card diagram

Figure 3.4. Mezzanine Card Physical Diagram



3.6. Connector 1 Pinout

Figure 3.5, “Connector 1 Pinout” [20] details the pinout of connector 1 (see Figure 2.3, “Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters” [6] for placement definition).

Connector Directivity Example: OP3_RX7_N means that the receiver is on the OP card and the transmitter is on the system planar, while OP3_TX7_N means that the transmitter is on the OP card and the receiver is on the system planar.

The source spreadsheet for the Connector 1 pinout is available at https://members.openpowerfoundation.org/document/dl/1159.

Figure 3.5. Connector 1 Pinout

	A	B	C	D	E	F	G	H	J	K
1	OP3_RX7_N	GND	OP3_RX5_N	GND	GND	GND	GND	OP3_TX6_N	GND	OP3_TX5_N
2	OP3_RX7_P	OP3_RX6_N	OP3_RX5_P	OP3_RX4_N	GND	GND	OP3_TX7_N	OP3_TX6_P	OP3_TX4_N	OP3_TX5_P
3	GND	OP3_RX6_P	GND	OP3_RX4_P	GND	GND	OP3_TX7_P	GND	OP3_TX4_P	GND
4	OP3_RX3_N	GND	OP3_RX2_N	GND	GND	GND	GND	OP3_TX3_N	GND	OP3_TX2_N
5	OP3_RX3_P	OP3_RX0_N	OP3_RX2_P	OP3_RX1_N	GND	GND	OP3_TX0_N	OP3_TX3_P	OP3_TX1_N	OP3_TX2_P
6	GND	OP3_RX0_P	GND	OP3_RX1_P	GND	GND	OP3_TX0_P	GND	OP3_TX1_P	GND
7	OP2_RX7_N	GND	OP2_RX5_N	GND	GND	GND	GND	OP2_TX6_N	GND	OP2_TX5_N
8	OP2_RX7_P	OP2_RX6_N	OP2_RX5_P	OP2_RX4_N	GND	GND	OP2_TX7_N	OP2_TX6_P	OP2_TX4_N	OP2_TX5_P
9	GND	OP2_RX6_P	GND	OP2_RX4_P	GND	GND	OP2_TX7_P	GND	OP2_TX4_P	GND
10	OP2_RX3_N	GND	OP2_RX2_N	GND	GND	GND	GND	OP2_TX3_N	GND	OP2_TX2_N
11	OP2_RX3_P	OP2_RX0_N	OP2_RX2_P	OP2_RX1_N	GND	GND	OP2_TX0_N	OP2_TX3_P	OP2_TX1_N	OP2_TX2_P
12	GND	OP2_RX0_P	GND	OP2_RX1_P	GND	GND	OP2_TX0_P	GND	OP2_TX1_P	GND
13	GND	GND	PRSNT1A_N	GND	GND	GND	GND	GND	GND	GND
14	GND	OP5_RX6_N	GND	OP5_RX5_N	GND	GND	OP5_TX7_N	GND	OP5_TX5_N	GND
15	OP5_RX7_N	OP5_RX6_P	OP5_RX4_N	OP5_RX5_P	GND	GND	OP5_TX7_P	OP5_TX6_N	OP5_TX5_P	OP5_TX4_N
16	OP5_RX7_P	GND	OP5_RX4_P	GND	GND	GND	GND	OP5_TX6_P	GND	OP5_TX4_P
17	GND	OP5_RX3_N	GND	OP5_RX2_N	GND	GND	OP5_TX3_N	GND	OP5_TX2_N	GND
18	OP5_RX0_N	OP5_RX3_P	OP5_RX1_N	OP5_RX2_P	GND	GND	OP5_TX3_P	OP5_TX0_N	OP5_TX2_P	OP5_TX1_N
19	OP5_RX0_P	GND	OP5_RX1_P	GND	GND	GND	GND	OP5_TX0_P	GND	OP5_TX1_P
20	GND	GND	GND	GND	REFCLK_156M_N	REFCLK_156M_P	GND	GND	GND	GND
21	OP4_RX7_N	GND	OP4_RX5_N	GND	GND	GND	GND	OP4_TX6_N	GND	OP4_TX5_N
22	OP4_RX7_P	OP4_RX6_N	OP4_RX5_P	OP4_RX4_N	GND	GND	OP4_TX7_N	OP4_TX6_P	OP4_TX4_N	OP4_TX5_P
23	GND	OP4_RX6_P	GND	OP4_RX4_P	GND	GND	OP4_TX7_P	GND	OP4_TX4_P	GND
24	OP4_RX3_N	GND	OP4_RX2_N	GND	GND	GND	GND	OP4_TX3_N	GND	OP4_TX2_N
25	OP4_RX3_P	OP4_RX0_N	OP4_RX2_P	OP4_RX1_N	GND	GND	OP4_TX0_N	OP4_TX3_P	OP4_TX1_N	OP4_TX2_P
26	GND	OP4_RX0_P	GND	OP4_RX1_P	GND	GND	OP4_TX0_P	GND	OP4_TX1_P	GND
27	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
28	GND	OP1_RX6_N	GND	OP1_RX5_N	GND	GND	OP1_TX7_N	GND	OP1_TX5_N	GND
29	OP1_RX7_N	OP1_RX6_P	OP1_RX4_N	OP1_RX5_P	GND	GND	OP1_TX7_P	OP1_TX6_N	OP1_TX5_P	OP1_TX4_N
30	OP1_RX7_P	GND	OP1_RX4_P	GND	GND	GND	GND	OP1_TX6_P	GND	OP1_TX4_P
31	GND	OP1_RX3_N	GND	OP1_RX2_N	GND	GND	OP1_TX3_N	GND	OP1_TX2_N	GND
32	OP1_RX0_N	OP1_RX3_P	OP1_RX1_N	OP1_RX2_P	GND	GND	OP1_TX3_P	OP1_TX0_N	OP1_TX2_P	OP1_TX1_N
33	OP1_RX0_P	GND	OP1_RX1_P	GND	GND	GND	GND	OP1_TX0_P	GND	OP1_TX1_P
34	GND	PRSNT1B_N	GND	GND	SMB_DAT	SMB_CLK	GND	GND	GND	GND
35	OP0_RX7_N	GND	OP0_RX5_N	GND	GND	GND	GND	OP0_TX6_N	GND	OP0_TX5_N
36	OP0_RX7_P	OP0_RX6_N	OP0_RX5_P	OP0_RX4_N	GND	GND	OP0_TX7_N	OP0_TX6_P	OP0_TX4_N	OP0_TX5_P
37	GND	OP0_RX6_P	GND	OP0_RX4_P	GND	GND	OP0_TX7_P	GND	OP0_TX4_P	GND
38	OP0_RX3_N	GND	OP0_RX2_N	GND	GND	GND	GND	OP0_TX3_N	GND	OP0_TX2_N
39	OP0_RX3_P	OP0_RX0_N	OP0_RX2_P	OP0_RX1_N	GND	GND	OP0_TX0_N	OP0_TX3_P	OP0_TX1_N	OP0_TX2_P
40	GND	OP0_RX0_P	GND	OP0_RX1_P	GND	GND	OP0_TX0_P	GND	OP0_TX1_P	GND

3.7. Connector 2 Pinout

Figure 3.6, “Connector 2 Pinout” [21] details the pinout of connector 2 (see Figure 2.3, “Top view of the reference mezzanine card outline with North direction indicator; dimensions in millimeters” [6] for placement definition).

Connector Directivity Example: OP3_RX7_N means that the receiver is on the OP card and the transmitter is on the system planar, while OP3_TX7_N means that the transmitter is on the OP card and the receiver is on the system planar.

The source spreadsheet for the Connector 2 pinout is available at https://members.openpowerfoundation.org/document/dl/1159.

Figure 3.6. Connector 2 Pinout

	A	B	C	D	E	F	G	H	J	K
1	GND	RSVD	GND	RSVD	GND	GND	RSVD	GND	RSVD	GND
2	RSVD	RSVD	RSVD	RSVD	GND	GND	RSVD	RSVD	RSVD	RSVD
3	RSVD	GND	RSVD	GND	GND	GND	GND	RSVD	GND	RSVD
4	GND	RSVD	GND	RSVD	GND	GND	RSVD	GND	RSVD	GND
5	RSVD	RSVD	RSVD	RSVD	GND	GND	RSVD	RSVD	RSVD	RSVD
6	RSVD	GND	RSVD	GND	GND	GND	GND	RSVD	GND	RSVD
7	GND	INT_RST_N	GND	GND	REFCLK_100M_N	REFCLK_100M_P	GND	GND	PRSNT2A_N	GND
8	RSVD	GND	RSVD	GND	GND	GND	GND	RSVD	GND	RSVD
9	RSVD	RSVD	RSVD	RSVD	GND	GND	RSVD	RSVD	RSVD	RSVD
10	GND	RSVD	GND	RSVD	GND	GND	RSVD	GND	RSVD	GND
11	RSVD	GND	RSVD	GND	GND	GND	GND	RSVD	GND	RSVD
12	RSVD	PE0_RX1_P	RSVD	PE0_RX0_P	GND	GND	PE0_TX1_P	RSVD	PE0_TX0_P	RSVD
13	GND	PE0_RX1_N	GND	PE0_RX0_N	GND	GND	PE0_TX1_N	GND	PE0_TX0_N	GND
14	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
15	RSVD	GND	RSVD	GND	GND	GND	GND	RSVD	GND	RSVD
16	RSVD	RSVD	RSVD	RSVD	GND	GND	RSVD	RSVD	RSVD	RSVD
17	GND	RSVD	GND	RSVD	GND	GND	RSVD	GND	RSVD	GND
18	PWR_EN	GND	I2C_ADDR_ID1	GND	PWR_BRAKE_N	PERST_N	GND	I2C_ADDR_ID0	GND	PWR_GOOD
19	TH_OVERT_N	GND	JTAG_SEL	GND	GND	GND	GND	GND	GND	GND
20	JTAG_TRST_N	JTAG_TMS	JTAG_TCK	JTAG_TDO	JTAG_TDI	+5.0V	+5.0V	+5.0V	+5.0V	+5.0V
21	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
22	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
23	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
24	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
25	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
26	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
27	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
28	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
29	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
30	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
31	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
32	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
33	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
34	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
35	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
36	GND	GND	GND	GND	GND	GND	GND	GND	GND	GND
37	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
38	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V
39	GND	GND	GND	GND	GND	GND	GND	GND	GND	PRSNT2B_N
40	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V	+12.0V

3.8. Sideband Signals

The following are the sideband signals present:

REFCLK_100MHz	PCIE reference clk. (from system logic)
REFCLK_133MHz	156.25MHz reference clk. From CPU. Can also be configured to be 133MHz.
I2C_ADDR0/1	I2C bus address LSBs. Pulled high (1.8V only) or pulled low on motherboard. These are used to set the physical address of the adapter device. These pins must be at a steady state voltage before the adapter device is taken out of reset.
SMB_CLK/DAT	SMBUS. Only supports operation at 1.8V. Care must be taken that this signal does not activate before PWR_EN and PWR_GOOD is valid. It is recommended that this signal be ANDED with PWR_EN at minimum since PWR_GOOD is optional
JTAG_SEL	control signal into adapter device to enable JTAG debug interface. 1.8V signal. Pulldown on motherboard for normal operation. Pull high on Mezz. card for JTAG operation.
JTAG_TCK/TMS/TDI/TDO/ TRST_N	JTAG interface for debug. Scan rings can span multiple riser sites. Ring accessible through debug connector on motherboard. 1.8V operation only.
PERST_N	PERST signal to adapter device. Effectively is the master reset. 1.8V level signal.
PWR_EN	Master on/off switch to adapter device power subsystem. 1.8V level signal only (OD, pullup is on motherboard). Signal is generated by logic on system planar and is controlled by power sequencer. This signal indicates to the adapter card that the adapter card power is enabled. The system asserts this signal to power on the module and shall be asserted only after all input rails are stable.
PWR_BRAKE_N	Power control signal to adapter device. 1.8V signal. OD. Pullup on motherboard. This signal may be used to indicate the adapter may go into a lower power state potentially reducing clock rates and power consumption. **Use by system is optional.**
PWR_GOOD	Adapter device PGOOD signal to system. 1.8V OD. Pullup with 10K on motherboard. Indicates to power sequencer that -adapter regs are good. Module power good. The module will assert this signal when all its internal power regulators are stable. **Use by system is optional.**
TH_OVERT_N	Emergency panic signal from the adapter device to the system that catastrophic thermal condition is imminent and device is going to shut down. 1.8V level signal to motherboard. Signals

	for each adapter device gathered together and fed to power sequencer. Care must be taken that this signal does not activate before PWR_EN and PWR_GOOD is valid. It is recommended that this signal be ANDED with PWR_EN at minimum since PWR_GOOD is optional and pulled up when not used.
INT_RST_N	Reset to adapter device controlled by Processor I2C bus. This signal must have a pullup to 3.3V on the Mezzanine card.
PRSNT_1A/2A/1B/2B	presence detect loop. Short PRSNT_1A to PRSNT_1B and short PRSNT_2A to PRSNT_2B on adapter device. Motherboard uses these to form a presence detect loop that is only true when the card is fully and correctly seated. Signal is used by JTAG scan ring bypass logic & pcie device prsnt detect logic

Note

1.8V and 3.3V levels imply +/- 10% inclusive of AC+DC variations.

3.9. PCI-e Express Interface

Each Mezzanine card has a x2 PCI Express (PCIe) interface that can be used in many ways, with one being sideband support for configuration. Section 3.7, “Connector 2 Pinout” [21] shows this interface labeled as PE0. This interface is compliant with PCI Express base Specifications.

Part II. Cabled Interface Extension

This part describes the use of a cabled connection to an adapter card. It uses a PCIe™ card as an example but the cabled extension does not require the adapter card be PCIe.

4. Advanced Accelerator Adapter Cable Interface Guidelines

Table of Contents

4.1. Power9 Advanced Accelerator Cable Interface

Power9 platforms support the optional cabling of the 25 Gbit/s Advanced Accelerator Interface to the advanced accelerator adapter in a riser card plugged into a PCIe slot in the same system. In addition, the adapter could be located in different drawer of the rack. This chapter contains information on topologies, connectivity, and routing guidelines for the advanced accelerator adapter cable interface. Refer to Part I, “Mezzanine Adapter Card” [2] for the non-cabled version of the interface.

4.1.1. Advance Accelerated Cable (AAC) Interconnection

The end-to-end AAC interconnection consists of the host board, the AAC cable, and the carrier card as shown in Figure 4.1, “Mechanical Components of Advanced Accelerator Cable Interface” [25]

Figure 4.1. Mechanical Components of Advanced Accelerator Cable Interface

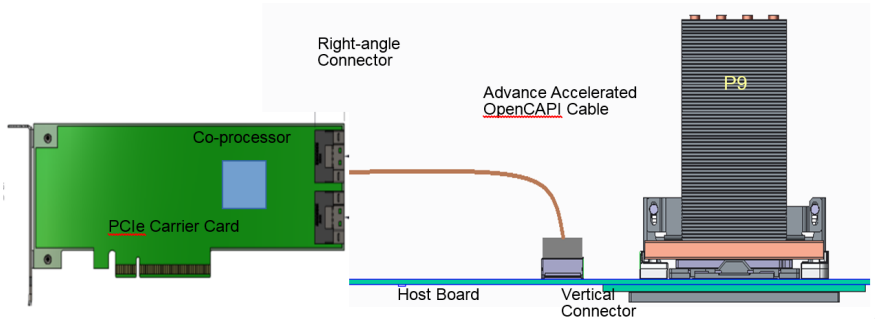


Figure 4.2. Electrical Components of Advanced Accelerator Cable (AAC) Interface

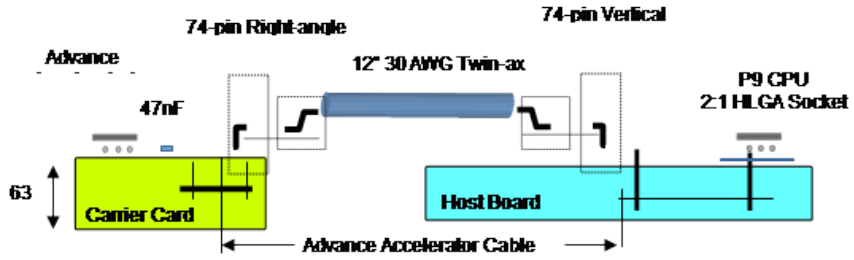


Table 4.1. Loss Budget for AAC Interface

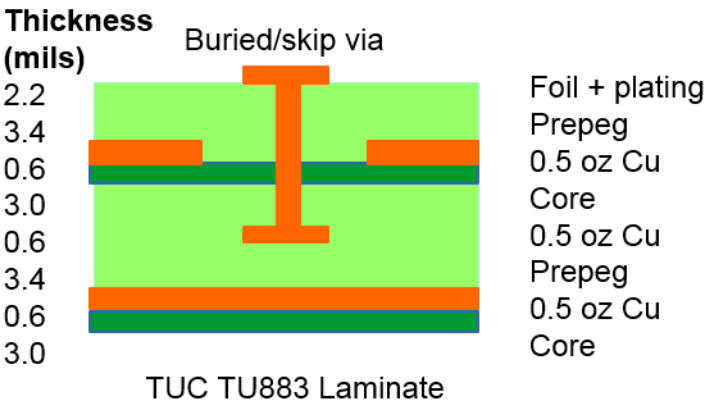
Components	Loss Budget (dB)	Notes
CPU Socket	4.25	34 mm trace, LGA 7-2-7 laminate package
Host Planar Trace	3.8	2.2 inch Meg4, back-drilled via PCB
SlimSAS Cable Assembly	6.5	12.5 inch 30 AWG twin-ax cable and connectors
Carrier Card (PCIe)	3.5	3.0 inch Meg6, 0201 AC cap, blind via PCB
Advance Accelerator Socket	1.9	14mm trace, laminate package
Worst case PCB	0.9	Account for +/-10% PCB tolerance
Total	20.9	POWER9 receiver specification

4.1.2. PCIe Carrier Card Layout Guidelines

4.1.2.1. Example Stack-up

Figure 4.3, “Example of PCIe Carrier Card Stack-up” [26] is an example stack-up for the PCIe carrier card that supports the advanced accelerator cable interface. This example stack-up is used to calculate the trace geometry used in this document. Skip or blind via technology and low loss material are assumed. To minimize via length signals should be routed toward the top layers where the connector resides.

Figure 4.3. Example of PCIe Carrier Card Stack-up



4.1.2.2. Signal Routing Guidelines

Figure 4.4, “PCIe Carrier Routing Topology” [27]shows the routing topology of the PCIe carrier card for AAC interface.

Figure 4.4. PCIe Carrier Routing Topology

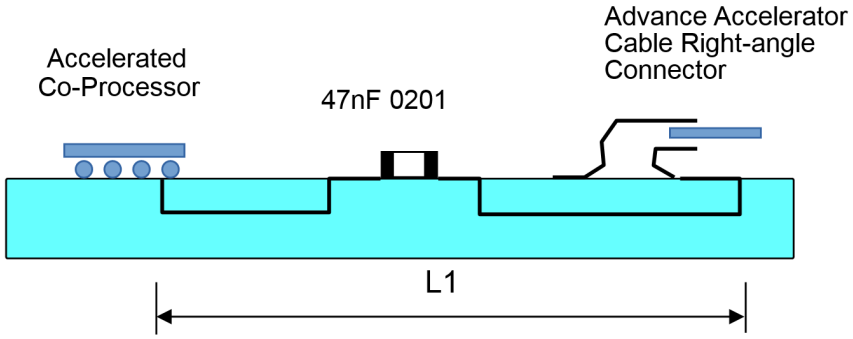


Table 4.2. PCIe Carrier Routing Guidelines

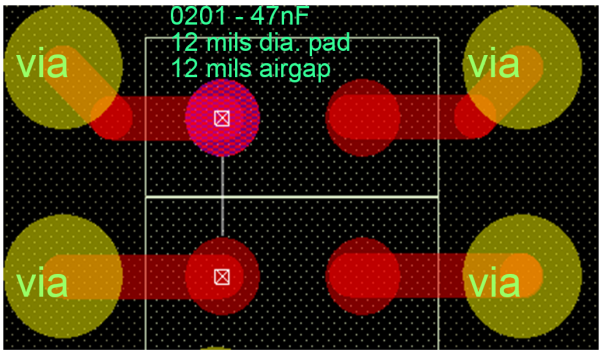
Layout Guideline	L1	Unit	Notes
Reference Plane	GND		Gnd-signal-Gnd
Differential Trace Impedance	85 diff +/- 10%	Ohms	
Trace Width	3.9	mils	Depend on stack-up
Trace Differential Spacing	5.1	mils	Depend on stack-up
Spacing Between Differential Pairs	20	mils	
Spacing to Other Signals	40	mils	Apply to Tx and Rx if in the same layer
Min-Max Trace Length	0 - 3.0	inch	
Differential Pair Phase Matching	5	mils	Include co-processor
Trace Length Tuning	1200	mils	Min-max delta
Blind via PCB	0	mils	No via stub
Maximum Active Via Length	25	mils	

4.1.2.3. Micro-strip Routing Guidelines

This section contains the connector and co-processor break-out routing guidelines for the PCIe carrier card.

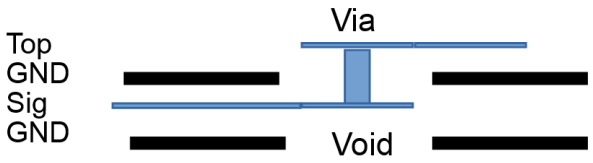
4.1.2.3.1. Guideline 1: AC Coupled Capacitor

- AC coupling capacitor must be located on the PCIE carrier card for both Rx and Tx directions.
- Use 47nF - 200nF capacitor for all Tx and Rx signals
- 12 mils diameter circular pad size
- 12 mils air-gap between pads



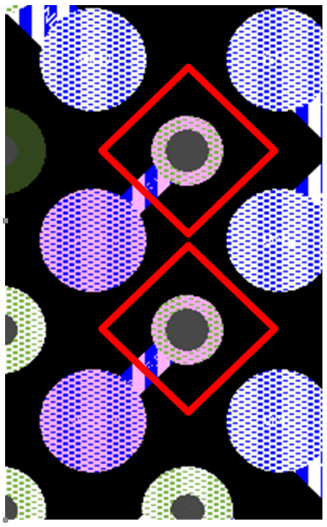
4.1.2.3.2. Guideline 2: Signal Blind Via and Anti-pad

- Signal blind via pads must be voided (anti-pad) in the GND plane below. Minimum void size is equal to pad size in the constraint area of co-processor break-out, but it should be a regular anti-pad size in the open area.

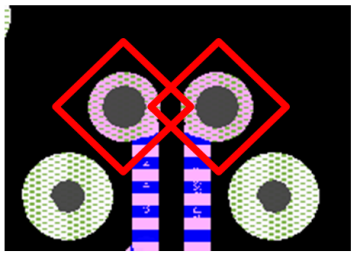


- Avoid signal crossing blind via GND void
- Prefer signal break-out in top and 1 st signal layer for 1-layer-depth via to minimize active via length
- GND return via is required for each signal via, placed symmetrically 0.8 mm away
- Differential via pitch in open-area = 20 mils
- Via drill diameter = 10 mils
- Via finish plated diameter = 8 mils
- Via pad diameter = 15.7 mils
- Via anti-pad in co-processor break-out area = 24x24 mils square
- Via anti-pad in open area = 28x28 mils square

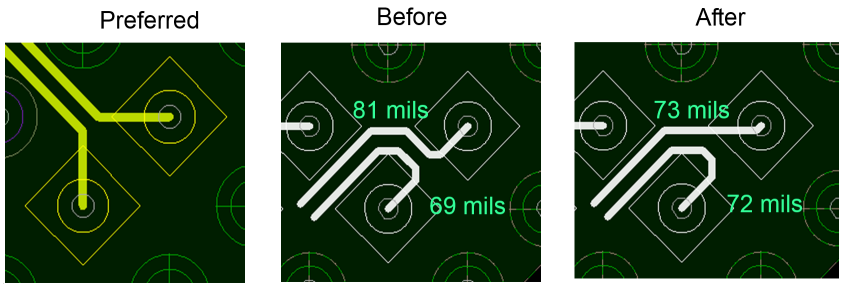
24x24 mils square anti-pad
In co-processor area



28x28 mils square anti-pad
In open area

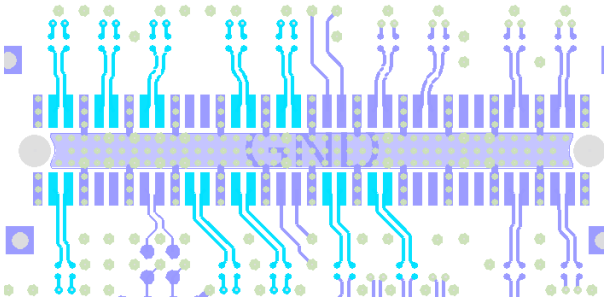


- Immediately match trace length at break-out to minimize serpentine compensation for phase matching rule in the open area. As shown as an example below: the before and after matching reduces the phase tolerance from 12 mils to 1 mils.

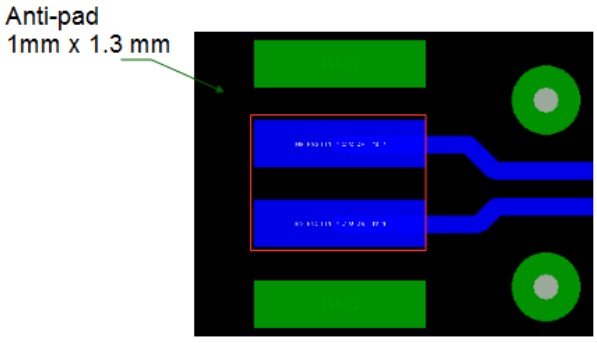


4.1.2.3.3. Guideline 3: 74-pin Connector Break-out

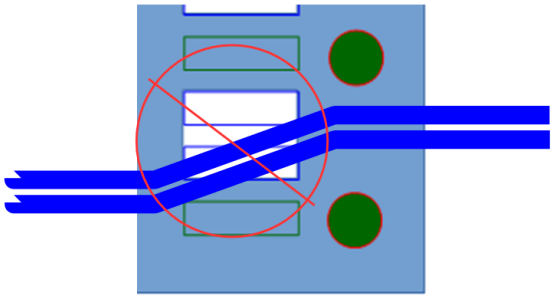
- Maximum micro-stripline length = 200 mils
- Minimum micro-stripline spacing = 40 mils
- Example of connector break-out



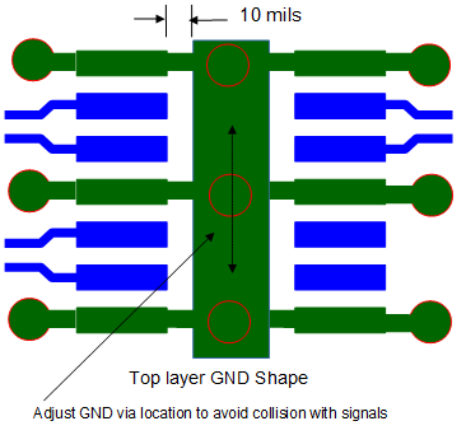
- Void GND reference plane below the signal pads, minimum void dimensions of 1x1.3mm as shown



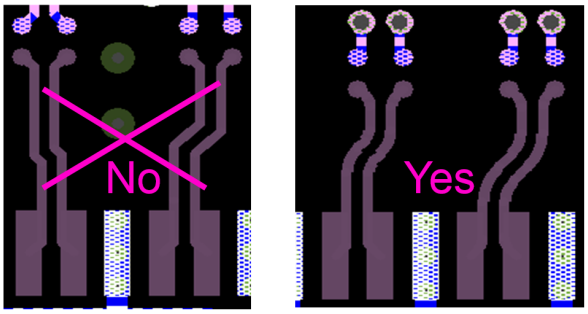
- Avoid signals crossing connector anti-pad void



- Connect GND pads to common shape in the middle of the rows as shown



- Smooth out the bending wire through out the design



4.2. Advanced Accelerator Cable

The advanced accelerator cable connecting the host board to the PCIe carrier card is shown in Figure 4.5, “Internal Cable Connection” [30]. The right angle 74-pin connector should be used and placed at the end of the PCIe card toward the CPU side. If the 74-pin vertical connector is selected, interference with the neighboring PCIe card may result.

Figure 4.5. Internal Cable Connection

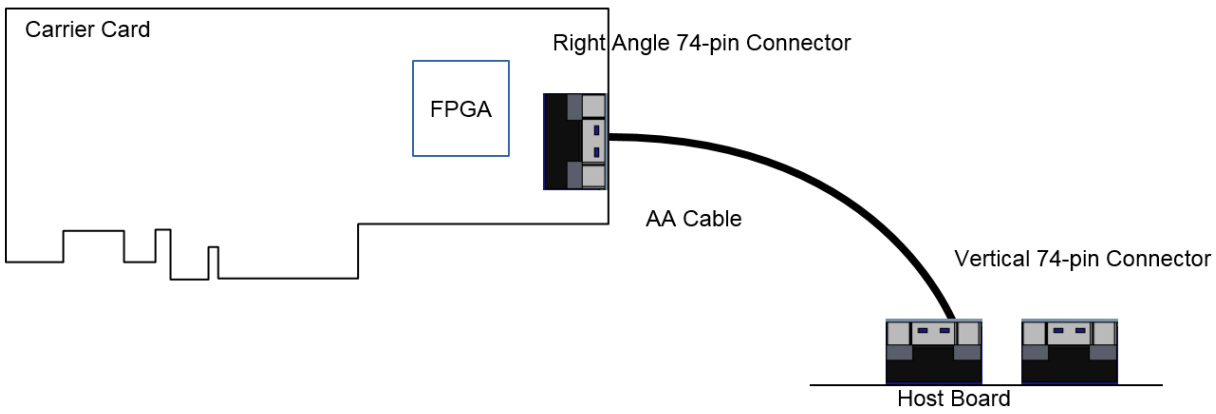
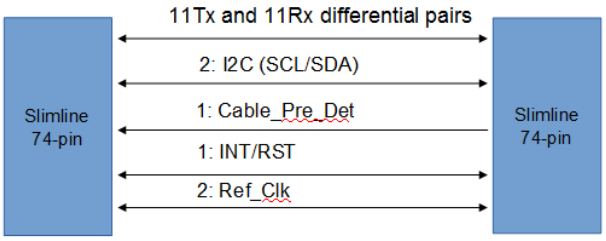


Figure 4.6. Advanced Accelerator Cable Circuit Schematic



I2C (SCL/SDA) and INT/RST voltage level is 3.3V +/- 10%

Cable_Pre_Det must have a 49.9 ohm pull-down to GND resistor located on the adapter card.

4.2.1. Cable/Connector Pin Mapping

The interconnect allows lane and polarity reversal. Pin swapping is not allowed. Host connector pin ordering and list are shown in Figure 4.7, “Host Connector Pin Assignment” [32]. Carrier connector pin ordering and list are shown in Figure 4.8, “Carrier Connector Pin Assignment” [33].

The source spreadsheet for the connector pinouts is available at https://members.openpowerfoundation.org/document/dl/1424.

Note: the pinout of the two connectors has been chosen such that a single twinax cable design can be used for both connections in a differential pair / lane.

Figure 4.7. Host Connector Pin Assignment

Host Logical Net Name	74-pin Vertical	Host Logical Net Name	74-pin Vertical
GND	A1	GND	B1
Host_Transmitted_Tx0_N	A2	Host_Transmitted_Tx1_N	B2
Host_Transmitted_Tx0_P	A3	Host_Transmitted_Tx1_P	B3
GND	A4	GND	B4
Host_Transmitted_Tx2_N	A5	Host_Transmitted_Tx3_N	B5
Host_Transmitted_Tx2_P	A6	Host_Transmitted_Tx3_P	B6
GND	A7	GND	B7
N/C	A8	N/C	B8
N/C	A9	N/C	B9
GND	A10	GND	B10
N/C	A11	Host_Transmitted_Tx4_N	B11
N/C	A12	Host_Transmitted_Tx4_P	B12
GND	A13	GND	B13
Host_Transmitted_Tx5_N	A14	Host_Transmitted_Tx6_N	B14
Host_Transmitted_Tx5_P	A15	Host_Transmitted_Tx6_P	B15
GND	A16	GND	B16
Host_Transmitted_Tx7_N	A17	I2C_SCL	B17
Host_Transmitted_Tx7_P	A18	I2C_SDA	B18
GND	A19	GND	B19
PRE_DETECT	A20	Host_Received_Rx1_N	B20
INT_RST	A21	Host_Received_Rx1_P	B21
GND	A22	GND	B22
Host_Received_Rx0_N	A23	Host_Received_Rx3_N	B23
Host_Received_Rx0_P	A24	Host_Received_Rx3_P	B24
GND	A25	GND	B25
Host_Received_Rx2_N	A26	N/C	B26
Host_Received_Rx2_P	A27	N/C	B27
GND	A28	GND	B28
REFCLK_156MHZ_P	A29	Host_Received_Rx4_N	B29
REFCLK_156MHZ_N	A30	Host_Received_Rx4_P	B30
GND	A31	GND	B31
N/C	A32	Host_Received_Rx6_N	B32
N/C	A33	Host_Received_Rx6_P	B33
GND	A34	GND	B34
Host_Received_Rx5_N	A35	Host_Received_Rx7_N	B35
Host_Received_Rx5_P	A36	Host_Received_Rx7_P	B36
GND	A37	GND	B37

Figure 4.8. Carrier Connector Pin Assignment

Carrier Logical Net Name	74-pin Right angle	Carrier Logical Net Name	74-pin Right angle
GND	B1	GND	A1
Carrier_Received_Rx0_N	B2	Carrier_Received_Rx1_N	A2
Carrier_Received_Rx0_P	B3	Carrier_Received_Rx1_P	A3
GND	B4	GND	A4
Carrier_Received_Rx2_N	B5	Carrier_Received_Rx3_N	A5
Carrier_Received_Rx2_P	B6	Carrier_Received_Rx3_P	A6
GND	B7	GND	A7
N/C	B8	N/C	A8
N/C	B9	N/C	A9
GND	B10	GND	A10
N/C	B11	Carrier_Received_Rx4_N	A11
N/C	B12	Carrier_Received_Rx4_P	A12
GND	B13	GND	A13
Carrier_Received_Rx5_N	B14	Carrier_Received_Rx6_N	A14
Carrier_Received_Rx5_P	B15	Carrier_Received_Rx6_P	A15
GND	B16	GND	A16
Carrier_Received_Rx7_N	B17	I2C_SCL	A17
Carrier_Received_RTx7_P	B18	I2C_SDA	A18
GND	B19	GND	A19
PRE_DETECT	B20	Carrier_Transmitted_Tx1_N	A20
INT_RST	B21	Carrier_Transmitted_Tx1_P	A21
GND	B22	GND	A22
Carrier_Transmitted_Tx0_N	B23	Carrier_Transmitted_Tx3_N	A23
Carrier_Transmitted_Tx0_P	B24	Carrier_Transmitted_Tx3_P	A24
GND	B25	GND	A25
Carrier_Transmitted_Tx2_N	B26	N/C	A26
Carrier_Transmitted_Tx2_P	B27	N/C	A27
GND	B28	GND	A28
REFCLK_156MHZ_P	B29	Carrier_Transmitted_Tx4_N	A29
REFCLK_156MHZ_N	B30	Carrier_Transmitted_Tx4_P	A30
GND	B31	GND	A31
N/C	B32	Carrier_Transmitted_Tx6_N	A32
N/C	B33	Carrier_Transmitted_Tx6_P	A33
GND	B34	GND	A34
Carrier_Transmitted_Tx5_N	B35	Carrier_Transmitted_Tx7_N	A35
Carrier_Transmitted_Tx5_P	B36	Carrier_Transmitted_TRx7_P	A36
GND	B37	GND	A37

4.2.2. Reference Advanced Accelerator Cable Solution Mechanical Description

This section contains information about a reference solution that has been used to implement this specification.

The following Amphenol® part numbers represent a family of connectors, connector hardware, and cables that may be used as a reference solution.

Reference Part Numbers

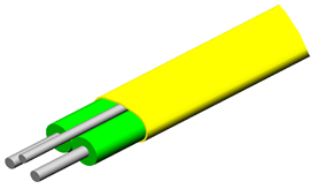
Cable Assembly	RSL74-0540-X X
Vertical 74-pin Connector	U10-K274-26X X
Rigt Angle 74-pin Connector	U10-J074-24X X

Table 4.3, “Raw cable Characteristics” [34] describes the physical and electrical characteristics of the reference cable.

Table 4.3. Raw cable Characteristics

Shielded Parallel Pair

30 AWG (0.25mm) Solid Silver Plated Copper
0.66mm Nom Fluorinated Polymer, Green Tint, Parallel Pair
0.0254mm Aluminized Polyester, Foil In, Yellow Clear Polyester, Heat Sealed



Conformance

RoHS Compliant
AWM Style 22018 30V 80C VW-1

Electrical

AWG: 30
Impedance: 85 Ohm +/- 5 Ohm
Design: 23 GHz
SDD21 (Maximum, 3M Sample)

- 6.71 dB/M @ 12.89 GHz
- 7.04 dB/M @ 14.025 GHz
- 8.47 dB/M @ 19.00 GHz

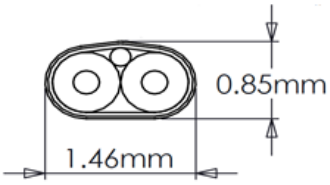


Figure 4.9, “Cable Assembly” [35] shows a drawing of the reference cable assembly.

Figure 4.9. Cable Assembly

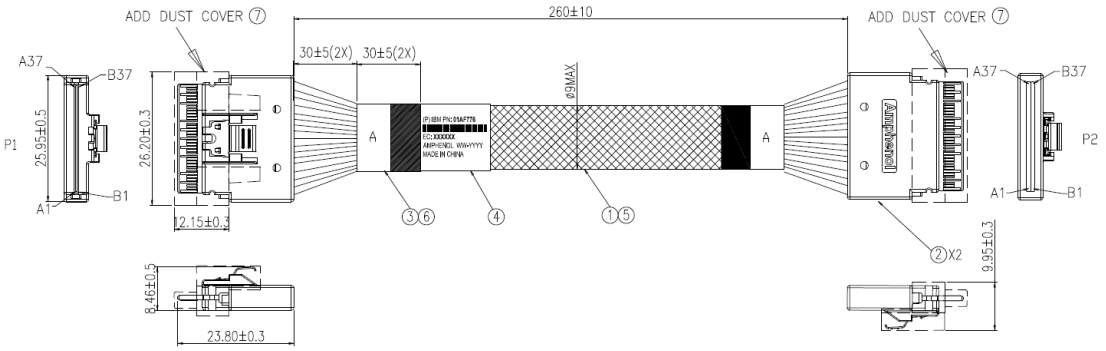


Figure 4.10, “Internal Cable Construction” [35] shows how the cable internal connections are constructed.

Figure 4.10. Internal Cable Construction

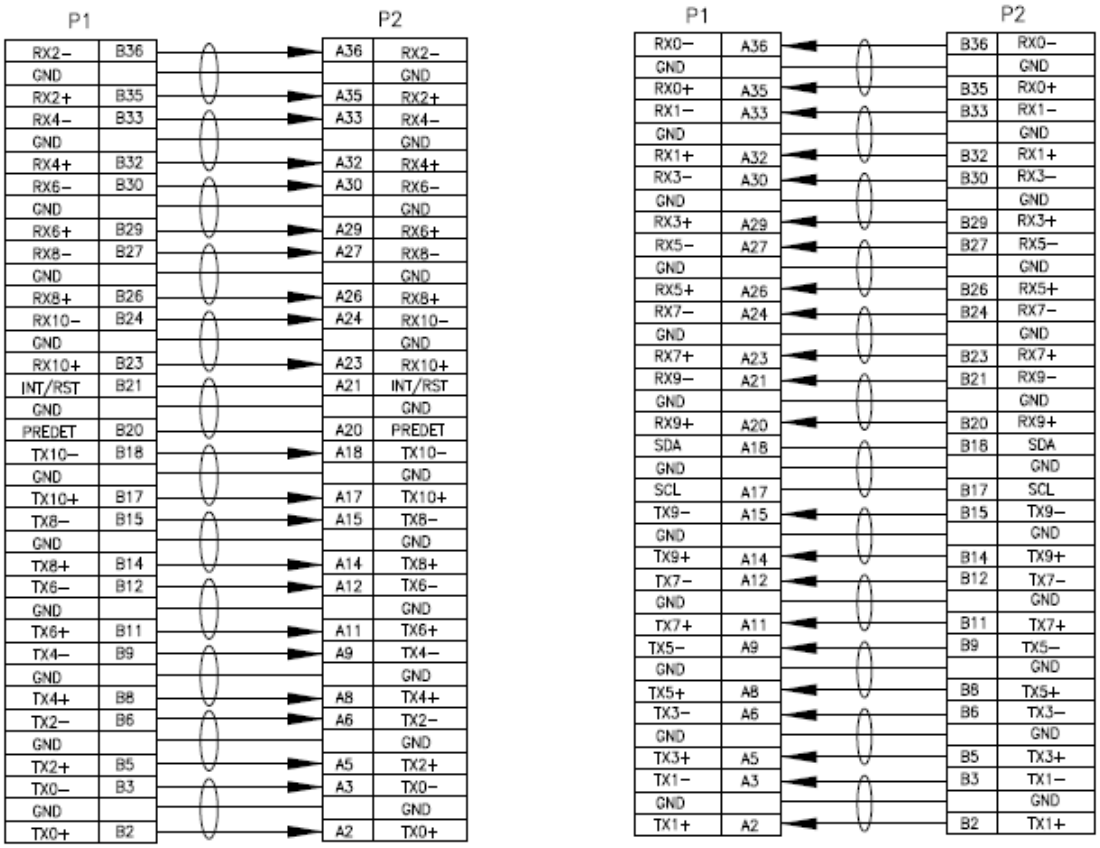


Figure 4.11, “Connector Features” [36] shows the various connector features. A specific cable / connector solution can be created from them. Figure 4.12, “74 Pin Right Angle Receptacle + Straight Plug Solution” [37] and Figure 4.13, “74 Pin Vertical Receptacle + Right Angle Plug Solution” [37] show two examples.

Figure 4.11. Connector Features

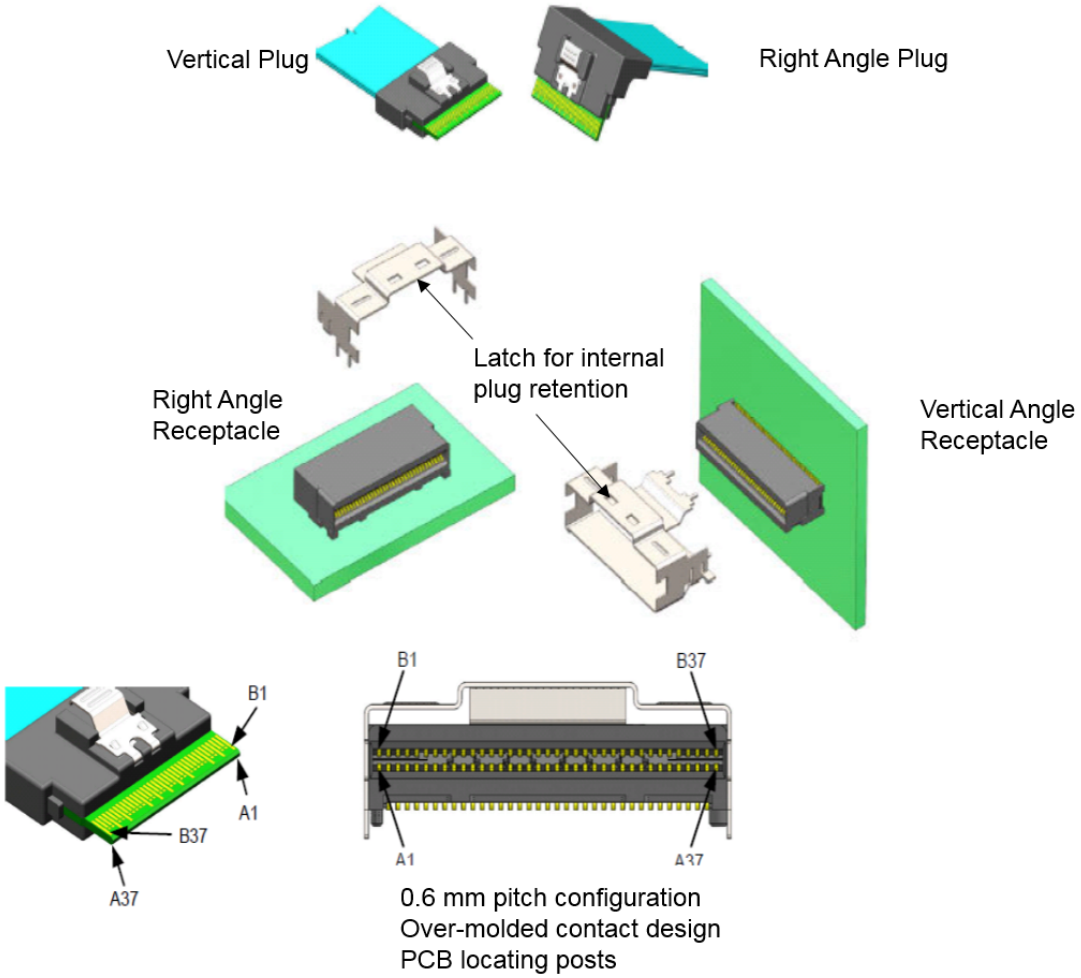
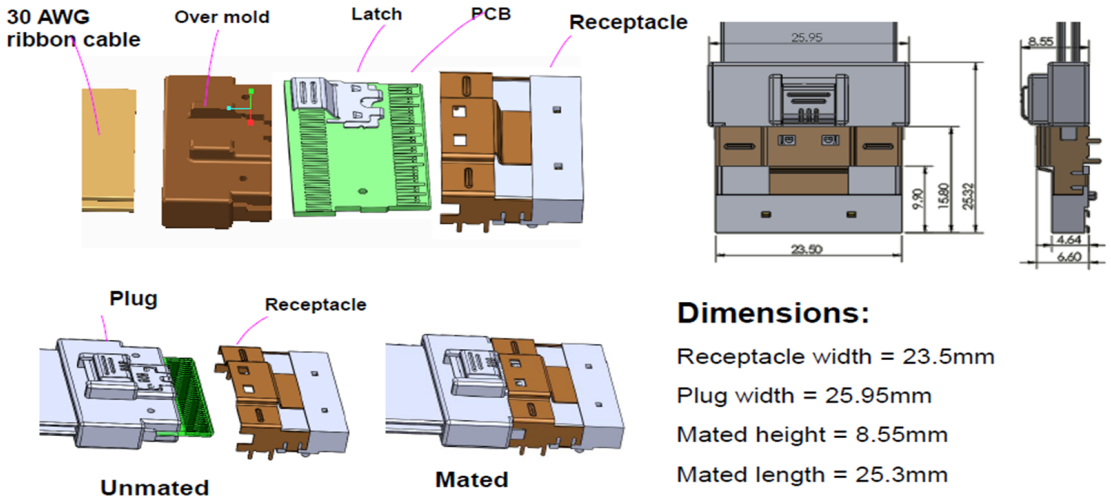


Figure 4.12. 74 Pin Right Angle Receptacle + Straight Plug Solution



Dimensions:

Receptacle width = 23.5mm

Plug width = 25.95mm

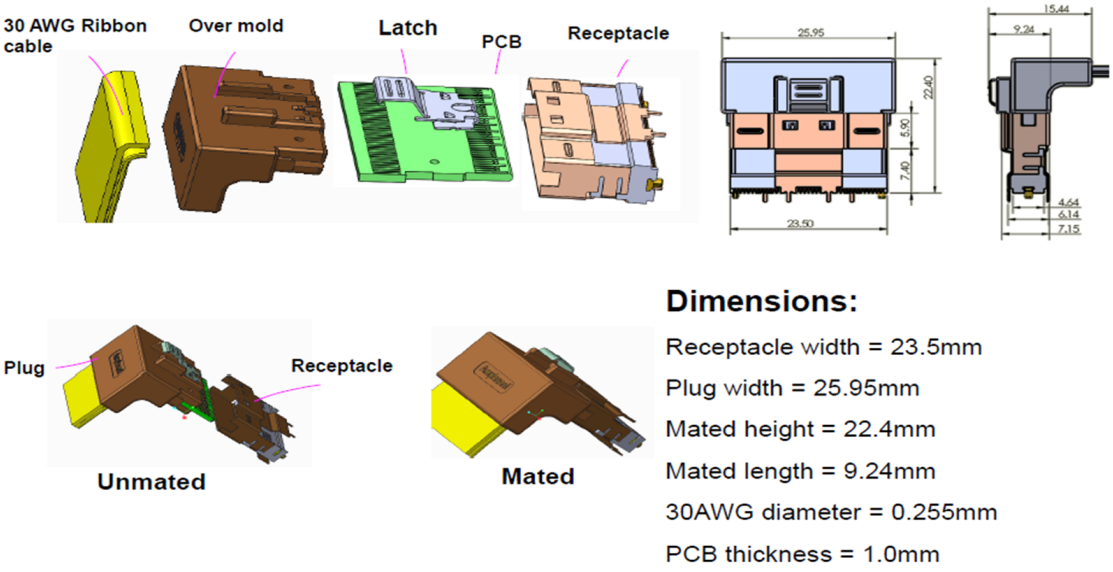
Mated height = 8.55mm

Mated length = 25.3mm

30AWG diameter = 0.255mm

PCB thickness = 1.0mm

Figure 4.13. 74 Pin Vertical Receptacle + Right Angle Plug Solution



Dimensions:

Receptacle width = 23.5mm

Plug width = 25.95mm

Mated height = 22.4mm

Mated length = 9.24mm

30AWG diameter = 0.255mm

PCB thickness = 1.0mm

Part III. 25 Gbit/sec Electrical Channel

This part describes the electrical interface specification for the 25 gbit/sec interface provided by the IBM® POWER9™ processor for POWER9 based OpenPOWER systems. It defines all requirements to allow electrical signaling compatibility with the 25 Gbit/sec ports on the POWER9 microprocessor module. This part identifies requirements on the endpoint PHY and the channel characteristics to ensure electrical compatability.

The endpoint PHY is required to be compatible with the *Optical Internetworking Forum (OIF) CEI 28Gbps SR Specification*. See Section B.1, “OIF CEI 28Gbps SR Specification” [58].

If the endpoint PHY is compatible with that specification all the compliance work done to prove compatability applies and the endpoint will interoperate with POWER9.

Details of the POWER9 PHY are provided in Appendix A, IBM® POWER9™ Specific Information [53] as a reference for the creation of accurate channel models.

This document will focus on the endpoint PHY specifications. Training will be covered in separate document.

5. Channel Overview

Table of Contents

5.1. High Level Description

The POWER9 25Gigabit port is a short channel chip-to-chip differential interface to provide data links between the POWER9 CPU and various specialized components. For example:

- CPU <=> CPU
- CPU <=> GPU
- CPU <=> FPGA
- CPU <=> NIC

The port / channel design also allows cabled connections. The interface is a striped serial design where each lane is required to perform Clock Data Recovery (CDR) and there is no clock-forwarding.

- Short reach chip-to-chip interface.
  - -21 dB insertion loss at Nyquist
    - OIF CEI 28G SR allows 15dB module pin to module pin with module loss allocation of 2(3dB). All 25G power designs close at -21dB C4-C4 to allow margin.
  - Bit Error Rate per lane = 1E-15
  - ~7 inches of Main planar PCB wiring using Mg6
  - Up to 2 meters with half-active electrical cabling and ~6 inches of MG6 per end.
  - Up to 3 meters with full-active electrical cabling and ~6 inches of MG6 per end.
  - Active optical cables
  - Active Copper cables
- IO Protocol
  - 25.78125Gbit, one speed no negotiation
  - Differential signaling with termination.
  - NRZ
  - Scrambled
  - DC or externally AC coupled. (AC coupling is externally located on the add-in card for TX and RX.)
- Link Configuration: Unidirectional lanes *Upstream* and *Downstream* at equal widths.
- Reference clock: External 156.25 MHz crystal clock distributed on-PCB

  The reference clock is always common and forwarded to the endpoint with the possibility of common spread.

In a multi-planar configuration (for example: drawer to drawer) the reference clock is forwarded. Thus making it a common reference clock design as required for OIF CEI compatibility.

5.2. Link Perspective

The channel is constructed from the endpoint PHY, the module package/substrate wiring and C4, the card/planar wiring, connectors, and the Power9 module package/substrate wiring, C4, and socket. From a system perspective the channel is used to establish a communication link between the endpoint logic functions and the Power9 logic functions. This part of the electro-mechanical specification does not concern itself with the following link level information.

Link Power States

Link Power States are defined in link specification.

Power-up Sequence

The PHY requires a sequence of events in order to power-up properly for operation. The power-up sequence of the voltage rails is defined in the electromechanical section. See Section 3.2, “Power Initialization Timing” [15].

Training sequence

The OIE CEI 28G SR specfication defines a "thin" PHY which does not include "thick" PHY functions such as: bit-lane repair, deskew, scrambler/descramble, etc. In addition to the "thick" PHY functions, the Link Layer is responsible for the CRC insertion/checking, replay buffers, and link layer retry protocols.

Datalink Layer (DL) and Transaction Layer (TL) layer materials.

This design invokes 64/66 coding and scrambling and framing which are defined in the link specification.

Note

Details about the use of this PHY and channel for OpenCAPI may be found in the OpenCAPI DL Specification

6. Channel Definition

Table of Contents

6.1. Channel requirements

The channel is a description of the end-end link. It consists of C4 bump pads, package including balls, LGA modules, PCB route, connectors, sockets and other such physical media between the driver and receiver. It does not include on-die termination.

This document defines a frequency domain compliance approach (see Chapter 7, “OIF CEI 28G Compliance” [50]) that specifies different combinations of frequency domain boundary conditions for compliant channels with the I/O properties defined in this document.

Following the channel compliance criteria of OIF CEI 28G SR will result in acceptable performance.

Note

Complementing the OIF CEI 28G SR compliance criteria with end to end time domain simulation with POWER9 root complex and eco-system endpoints is strongly recommended.

Reference channel models are available at: https://members.openpowerfoundation.org/document/dl/455.

S-Parms are available at: https://members.openpowerfoundation.org/document/dl/455.

IBM Engineering support is available. See Section A.5, “Contacting IBM for Assistance” [57].

Table 6.1. Channel Requirements

Symbol	Parameter	Min	Typ	Max	Unit	Notes
T_{CH_SKEW}	Delay difference between lanes within a bundle when FPGA is driving to POWER9			132	UI	Skew between lanes inclusive of on chip consumption (maximum skew of TX is 64UI from endpoint inputers)
T_{CH_SKEW}	Delay difference between lanes within a bundle driven from POWER9, Launch skew, board and package skew and retime skew			10	UI	Skew between lanes inclusive with 1.5UI allocated to cables
IL(f)	Channel insertion loss at Nyquist			21	dB	Includes PCB manufacturing variations and tolerances including effects of humidity/temperature variations on dielectric materials. Ball to Ball is the OIF CEI 28G SR = 14dB at Nyquist +6dB for each endpoint.
ILD(f)	Insertion Loss Deviation Measure of deviation from the insertion loss curve within a specified frequency range			0.45	dB	The insertion loss deviation (ILD) measure is used to quantify the amount of reflections within a channel. In order to calculate ILD, a fit of the insertion loss curve (in dB) is generated between 0Hz and a high frequency point defined at which the insertion loss is 40 dB or the highest frequency in the S-parameter model, whichever is lowest. The S-parameter file is assumed to have a 20MHz start frequency and 20MHz step frequency. The insertion loss curve fit is found using *moving average smoothing*.[a] The moving average smoothing procedure uses a window size of 51 discrete frequency points centered at the point under consideration. This window spans a 1GHz range with a 20MHz step S-parameter model. Once the fitted curve is obtained, the error between the original insertion loss curve and the fitted curve is found at each discrete frequency point in the S-parameter model. The squares of each discrete frequency point error are then summed. The insertion loss deviation (ILD) is the sum of all squares divided by the total number of discrete frequency points considered.
ILDB(f)	Maximum Insertion loss deviation from Insertion Loss Fit below fundamental frequency			1	dB	ILDB is the maximum difference between the fitted line calculated for ILD and the original insertion loss at any frequency between the 20MHz and the fundamental frequency.
NEXT(f)	Power sum NEXT @ Nyquist			-50	dB	Calculated by carrying out the root sum square of all NEXT aggressors on to the victim differential pair. All channels must be designed so that FEXT is the dominant source of crosstalk. Tx/Rx PCB wiring must be done on different layers or sufficiently spaced. Tx/Rx pin and via combinations must have sufficient ground isolation.
FEXT(f)	Power sum FEXT @ Nyquist			-38	dB	Calculated by carrying out the root sum square of all FEXT aggressors on to the victim differential pair.
ICR	ICR @ Nyquist	18			dB	Difference between insertion loss and Crosstalk power sum value. Also known as signal to crosstalk ratio.
$TEMP_{CH}$	Temperature	0		100	^{o}C	Ambient

Symbol	Parameter	Min	Typ	Max	Unit	Notes
RL_{CH_DIFF}	Differential Return Loss @ Nyquist	10			dB	Optimally the minimum differential return loss between 0Hz and Nyquist would occur at Nyquist. This rule can be broken if simulations show passing eye margins. Refer to Section 6.2.7, “Differential Return Loss for both Transmitter and Receiver” [47].
RL_{CH_CM}	Common Mode Return Loss @ Nyquist	10			dB	Optimally the minimum common mode return loss between 0Hz and Nyquist would occur at Nyquist. Refer to Section 6.2.7, “Differential Return Loss for both Transmitter and Receiver” [47].
Z_{CH_DIFF}	Differential impedance	-10%	85	+10%	Ohms	Channel impedance (note this deviates from OIF CEI which is 100+/-)

[a]**Moving Average Smoothing:**

The following steps describe in further detail the moving average smoothing to be used.

1. Select the size of the window of the moving average to be 51 discrete points. This window would span a 1 GHz range with 20 MHz steps.
2. Calculate "moving average" for each frequency point.
   A. For each discrete frequency point, beginning with the 26th discrete point and ending with the highest_frequency_point_minus_25, find the average of all the values ranging between 25 values before the considered point and 25 values after the considered point.
   B. For each discrete frequency point between the first and the 25th, find the average of all the values between all the lower frequency points and as many frequency points higher than the considered point,
   C. For each discrete frequency point between the 25th point to the last and the last, find the average of all the values between all the higher frequency points and as many frequency points lower than the considered point.
3. Generate a smoothed curve of the insertion loss using the averages calculated at each frequency point

6.2. Electrical Specifications

6.2.1. Power Supply

Endpoint power supplies that are required are a system design implementation aspect. Since links can be AC coupled the endpoint power supply is not specified. For DC coupled links the VCM must satisfy the root complex VCM range.

6.2.2. External Reference Clock

The external reference clock system design is common clocking with allowed down spread.

Table 6.2. Specification for Reference Clock

Symbol	Parameter	Min	Typ	Max	Unit	Notes
ref_clk	Mhz	-50ppm	156.25	50 ppm		
L_100K	Phase noise @ 100 kHz from carrier		-120		dBc/Hz	
JIT_{REFCLK}				1	pS,rms	Integrate from 1 kHz 20 MHz
S_SSC	SSC (optional) Spread	-5000		0	ppm	This is for common clock based *spread*. If the endpoints do not have the same oscillator, *spread* cannot be invoked . System design shall require "common clock with possible downspread" Steps shall be 1000, 2000, 3000, 4000, or 5000 ppm

Symbol	Parameter	Min	Typ	Max	Unit	Notes
FM_SSC	Modulation freq	30		33	kHz	
T_{REFCLK}	Routing between Rx/Tx			10	nS	Insertion delay between chips
AMP_{REFCLK}	Differential voltage swing	0.575		0.85	Vppd	POWER9 uses SILabs HSCL driver
VCM_{REFCLK}	Common mode	.35		0.4	V	
F_{AC_REFCLK}	AC coupling (Required)			10	kHz	AC coupling at the HPF corner If 0-1V is incompatible with the endpoint VCM then AC coupling must be added.
TR_{REFCLK}	Rise/Fall time	0.2		1	nS	20-80% and the voltage swing is 0-1V with SST driver
Z_{DIFF_REFCLK}	Differential impedance	-10%	85	+10%	Ohms	

6.2.3. CDR PLL

Table 6.3. Specification for CDR PLL

Parameter	Min	Typ	Max	Unit	Description
LC2 VCO Frequency	-10%	25.78125G	+10%	GHz	Tuning range of VCO_2
POWER9 PLL Lock time			5	mS	Calibration, acquisition and lock from PWR_GOOD [22].
Loop BW	2	3	8	MHz	Programmable BW. Loop BW shall track SSC, if present
Peaking	0		2	dB	Programmable Damping
RJ Jitter (absolute)			0.2	ps	RMS jitter spec of PLL output clock measured from baud/1667 to baud/2

6.2.4. IO Lane

Table 6.4. Electrical Spec: Tx and Rx Termination

Symbol	Parameter	Min	Typ	Max	Unit	Notes
Z_{DIFF_IO}	Differential impedance		85		Ohms	During normal mode and IDLE mode.
Z_{SE_IO}	Single ended impedance		43		Ohms	The two single ended resistances shall match within 10%.
$dV_{GND,TxRx}$	Ground difference			50	mVpp	Ground differences between Tx and far-end Rx
IL(TX)	POWER9 TX has AC boost	-0.5			dB	POWER9 TX at Pad has BOOST
IL_{IO}	BW or loss due to poles		-1	-1.5	dB	A total budget of 3 dB for (Tx + Rx).
RL_{DIFF_IO}	Differential return loss				dB	More details in template in Section 6.2.7, “Differential Return Loss for both Transmitter and Receiver” [47].
RL_{CM_IO}	Common-mode return loss				dB	More details in template in Section 6.2.8, “Common to differential mode and differential to common mode conversion” [48].

6.2.5. Endpoint Transmitter

6.2.5.1. Electrical Output Specification

This section describes the Endpoint Transmitter Electrical Output Specification. This specification is essentially the same as OIF CEI 28G SR and is enumerated in Table 6.5, “Endpoint Transmitter Electrical Output Specification. ” [45]. Key exceptions are identified via italics and highlighted in red.

Table 6.5. Endpoint Transmitter Electrical Output Specification.

Characteristic	Symbol	Condition	MIN.	TYP.	MAX.	UNIT
Baud Rate	T_Baud			*25.78125*		Gsym/s
Output Differential Voltage	T_Vdiff	*Note 4*	800		1200	mVppd
Differential Resistance	T_Rd	*Channels are designed 85 ohms typical (Note 3)*		*85*		Ohm
Differential Termination Resistance Mismatch	T_Rdm				10%	%
Output Rise and Fall Time (20% to 80%)	T_tr, T_tf	Emphasis off (Note 2)	8			ps
Common Mode Noise	T_Ncm	Note 4			12	mVrms
Differential Output Return Loss	T_SDD22	See Section 6.2.7, “Differential Return Loss for both Transmitter and Receiver” [47] curves copied from OIF 28G spec				dB
Common Mode Output Return Loss	T_SCC22	Below 10 Ghz			-6	dB
		10GHz to Baud rate			-4	
Output Common Mode Voltage	T_Vcm	Load type 0	-100		-1700	mV

NOTES:

1. Load type 0 is AC coupled. The AC coupling caps exist near the endpoint devices
2. The transmitter under test is preset such that C0 is its maximum value (C0_max) and all other coefficients are zero. The 20% and 80% values are of the steady state one and zero. The max value is limited by meeting the transmit launch s12 of 0dB
3. Nominal 85 ohm channels are designed from C4-C4. Endpoint package traces of endpoint module should also be defined to be 85 ohms. *If the endpoint is 100ohms in the silicon and the package is 100 ohms then regression is required to insure the ILD is acceptable.*
4. Procedure defined below from OIF CEI 28G specification,

6.2.5.2. Endpoint TX Jitter Models

The jitter terms below are specified at the PAD or C4 the chip carrier organic package is considered part of the channel.

Compliance to the TX model is defined by following the direction of the OIF CEI 28G VSR and SR standards.

The TX jitter test methods follow the test methodologies of OIF CEI-28G-SR. Table 6.6, “Transmitter Output Jitter Specification for endpoint (from CEI 28G SR)” [46] can be used as the jitter parameters for simulating the the channel. Alternately, the modelling may use the POWER9 TX info found in Section A.3.2, “POWER9 TX Jitter Terms” [54] for the transmitter if the endpoint is only intended to connect to Power9.

Table 6.6. Transmitter Output Jitter Specification for endpoint (from CEI 28G SR)

Characteristic	Symbol	Condition	MIN.	TYP.	MAX.	UNIT
Uncorrelated Unbounded Gaussian Jitter	T_UUGJ				0.15	UIPP
Uncorrelated Bounded High Probability Jitter	T_UBHPJ	Note 2			0.15	UIPP
Duty Cycle Distortion (component of UBHPJ)	T_DCD	Note 3			0.035	UIPP
Total Jitter	T_TJ	Note 1			0.28	UIPP

NOTES:

1. T_TJ includes all of the jitter components measured without any transmit equalization.
2. Measured with all possible values of transmitter equalization, excluding DDJ as defined in the Section below.
3. included in T_UBHPJ

6.2.6. Endpoint Receiver

6.2.6.1. Electrical Input Specification

This section describes the Endpoint Receiver Electrical Output Specification. This specification is essentially the same as OIF CEI 28G SR and is enumerated in Table 6.7, “ Endpoint Receiver Electrical Input Specification. ” [46]. Key exceptions are identified via italics and highlighted in red.

Table 6.7. Endpoint Receiver Electrical Input Specification.

Characteristic	Symbol	Condition	MIN.	TYP.	MAX.	UNIT
Baud Rate	R_Baud			*25.78125*		Gsym/s
Input Differential Voltage	R_Vdiff	*Note 1*			1200	mVppd
Differential Input Impedance	R_Rdin	*Channels are designed 85 ohms typical*		*85*		Ohm
Differential Termination Resistance Mismatch	R_Rdm				10%	%
Differential Output Return Loss	R_SDD11	See Section 6.2.7, “Differential Return Loss for both Transmitter and Receiver” [47] curves copied from OIF 28G SR spec (14.3.2.5)				dB
Common Mode Output Return Loss	R_SCC11	Below 10 Ghz			-6	dB
		10GHz to Baud rate			-4	
Input Common Mode Voltage	R_Vcm	Load type 0	-100		1800	mV

NOTES:

1. The receiver shall have a differential input range sufficient to accept a signal produced at point R by the combined transmitter and channel. The channel response shall include the worst case effects if the return losses at the transmitter and receiver.
2. Load type 0 with min. T_Vdiff, AC-Coupling or floating load. For floating load, input resistance shall be >1 kOhm.

6.2.6.2. Endpoint Receiver Input Jitter Specification.

Table 6.8. Endpoint Receiver Input Jitter Specification.

Characteristic	Symbol	Condition	MIN.	TYP.	MAX.	UNIT
Sinusoidal Jitter, Maximum	R_SJ-max	CEI OIF 28G SR section 2.5.4 (Note 1)			5	UIpp

Characteristic	Symbol	Condition	MIN.	TYP.	MAX.	UNIT
Sinusoidal Jitter, High Frequency	R_SJ-hf	CEI OIF 28G SR section 2.5.4 (Note 1)			0.05	UIpp

NOTES:

1. The receiver shall tolerate the sum of these jitter contributions. Total transmitter jitter from table in TX section ; Sinusoidal jitter as defined in receiver section; The effects if the channel compliant to the Channel Characteristics.

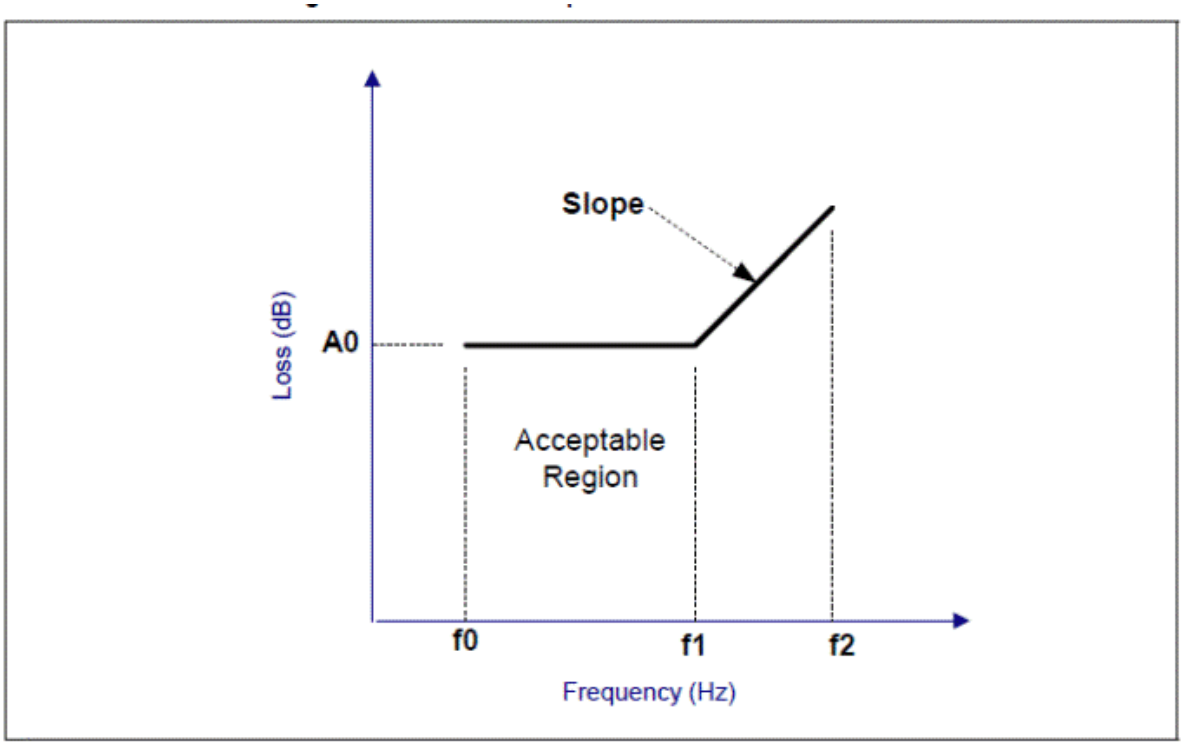
6.2.7. Differential Return Loss for both Transmitter and Receiver

The required Differential return loss curve is the OIF CEI 28G SR specs. The table is recopied below.

The sdd11 and sdd22 are the OIF-CEI-28G compliance curve

Parameter	Value	Units
A0	-12	dB
f0	50	MHz
f1	0.1714 x R_Baud	Hz
f2	R_Baud	Hz
Slope	12.0	dB/dec

Figure 6.1. SDD11 and SDD22 differential return loss at C4(PADS) template for RX and TX.

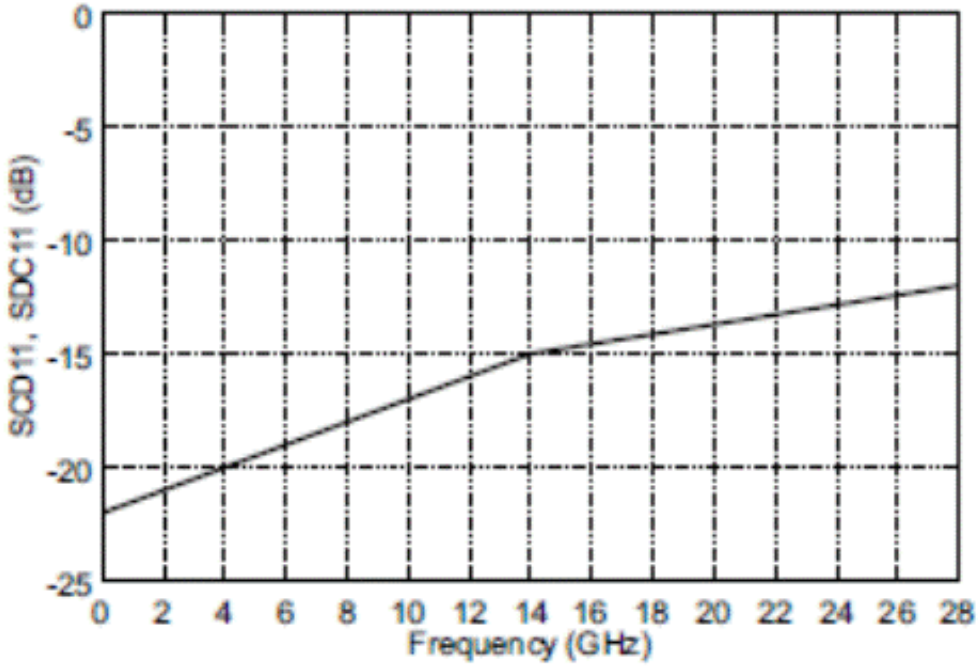


6.2.8. Common to differential mode and differential to common mode conversion

The common to differential mode and differential to common mode conversion specifications are intended to limit the amount of unwanted signal energy that is allowed to be generated due to conversion of common mode voltage to differential mode voltage or vice versa. When measured at the respective input test point, common to differential mode or differential to common mode conversion shall not exceed the limits illustrated in Figure 6.2, “SDC11 and SCD11 for module input (TP1) and host input (TP4a) (for fb = 28 GHz)” [49].

Figure 6.2, “SDC11 and SCD11 for module input (TP1) and host input (TP4a) (for fb = 28 GHz)” [49] is the OIF-CEI-28G compliance curve given by Section 6.2.8, “Equation Y” [49].

Figure 6.2. SDC11 and SCD11 for module input (TP1) and host input (TP4a) (for fb = 28 GHz)



Equation Y

SDC11, SCD11 < -22 + 14 * (f / fb) dB for 0.05 < f < fb/2

SDC11, SCD11 < -18 + 6 * (f / fb) dB for fb/2 < f < fb

6.2.9. Common Mode Noise

The Common Mode Noise is measured in compliance with CEI 28G SR

7. OIF CEI 28G Compliance

Table of Contents

7.1. Approach

The range of channel designs encountered when interfacing to the Power9_25Gigabit PHY are expected to deviate somewhat from the Industry Standard Specs of OIF CEI 28G VSR and OIF CEI 28G SR. These channels will be greater in loss then the Industry Standard Spec of OIF CEI 28G VSR and less in loss than OIF CEI 28G SR. All compliance testing done for OIF CEI 28G SR is directly applicable and does not need to be repeated. If the endpoint PHY is SR compliant the endpoint PHY most likely has more RX equalization then would be optimal but should be adequate. The rational is as follows:

- The channel loss C4 to C4 is specified at 21 dB.
- The OIF CEI 28G VSR specifies the loss at 10dB module-module.
- If a token 3 dB per endpoint (module to c4) is assumed then the equivalent VSR C4 to C4 result would be 16dB. A loss of 16dB is typically CTLE equalizable.
- Assuming SR wire length adds another 5 dB if you use the same token 3dB per end giving the resultant to be 21dB.

Note

To ensure operating margin the endpoint IP provider is strongly suggested to simulate the full channel model in their simulator of choice.

Reference channel models are available at: https://members.openpowerfoundation.org/document/dl/453.

S-Parms are available at: https://members.openpowerfoundation.org/document/dl/455.

IBM Engineering support is available. See Section A.5, “Contacting IBM for Assistance” [57].

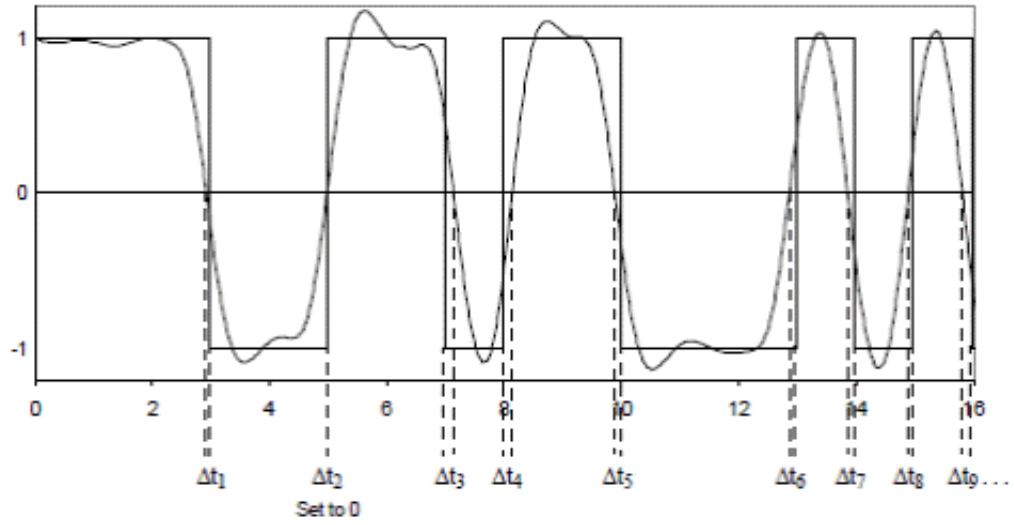
7.2. Data Dependent Jitter (DDJ) measurement from CEI 28G SR

A high-resolution oscilloscope, time interval analyzer, or other instrument with equivalent capability may be used to measure DDJ. Establish a crossing level equal to the average value of the entire waveform being measured.

Synchronize the instrument to the pattern repetition frequency and average the waveforms or the crossing times sufficiently to remove the effects of random jitter and noise in the system. The mean time of each crossing is then compared to the expected time of the crossing, and a set of timing variations is determined. DDJ is the range (max-min) of the timing variations. Keep track of the signs (early/late) of the variations. Note, it may be convenient to align the expected time of one of the crossings with the measured mean crossing. All edges of the repeating pattern that have been averaged need to be included in the measurement.

Figure 7.1, “DDJ Measurement Method” [51] below illustrates the method. The vertical axis is in arbitrary units, and the horizontal axis is plotted in UI. The waveform is AC coupled to an average value of 0, therefore 0 is the appropriate crossing level. The rectangular waveform shows the expected crossing times, and the other is the waveform with jitter that is being measured. Only 16 UI are shown in this example. The waveforms have been arbitrarily aligned with (delta t2 = 0) at 5 UI.

Figure 7.1. DDJ Measurement Method



$$DDJ = \max(\Delta t_1, \Delta t_2, \ldots \Delta t_n) - \min(\Delta t_1, \Delta t_2, \ldots \Delta t_n)$$

7.3. Endpoint compliance TX jitter models for channel simulation

A TX model for compliance will be formulated as follows:

CEI TX Jitter Interactions

The jitter terms below are specified at the PAD or C4 the chip carrier organic package is considered part of the channel.

Using the Dual Dirac jitter model[1], total jitter (TJ) is given by Eq. (1)

$$T_{TJ} = T_{DCD} + T_{DDJ} + 2Q\,(BER)\sqrt{\sigma_{RJ}^2 + \sigma_{BUJ}^2}$$

where Q (BER) is Q-factor, both RJ and BUJ are assumed as Gaussians down to 1e-15 probability; σ_{RJ} and σ_{BUJ} are rms values for RJ and BUJ respectively. We could rewrite Eq (1) in terms of pk-pk values of RJ and BUJ as the following Eq (2)

$$T_{TJ} = T_{DCD} + T_{DDJ} + \sqrt{T_{RJ}^2 + T_{BUJ}^2}$$

For CEI[2], we have T_{DCD} =0.035 UI, T_{RJ} = 0.15 UI, T_{BUJ}= 0.15 UI, T_{TJ} =0.28 UI, and with Eq. (2), we get T_{DDJ} = 0.0329 UI.

By specifying the upper bounds of T_{DCD} , T_{RJ} , T_{BUJ}, T_{TJ} , it implicitly specifies the upper limit for DDJ, as such all the jitter components are bounded and limited.

7.4. Receiver Compliance

This section is taken from section 2.5.4 in the OIF_CEI_28G_SR specification. All references are to the OIF_CEI_28G-SR specification and the reader is refered to that specification for further details.

The following steps shall be made to identify whether a receiver is considered compliant.

1. The DUT shall be measured to have a BER1 better than specified for a stressed signal (see Appendix 2.E.4.2 for a suggested method) with a confidence level of three sigma (see Appendix 2.F.2. for a suggested method) given

    - for non-transparent applications, the defined sinusoidal jitter mask for relative and total wander as per Annex 2.A.1 and Annex 2.A.2, with a high frequency total/relative wander and a maximum total/relative wander as defined in the Implementation Agreement
    - for transparent application, the defined appropriate sinusoidal jitter mask for the specific optical standard
    - the high frequency jitter should be calibrated by either:
        - applying the maximum specified amount of receiver High Probability Jitter and Gaussian jitter2 including CBHPJ

        or

        - applying the maximum specified amount of receiver High Probability Jitter and Gaussian jitter3 excluding CBHPJ
        - cascading with a compliance channel or filter as identified by 2.5.2.
        - applying an additive crosstalk signal of amplitude such that the resulting statistical eye, given the channel, jitter and crosstalk, is as close as feasible in amplitude when compared to the defined minimum amplitude for channel compliance

Appendix A. IBM® POWER9™ Specific Information

A.1. Global Parameters

Table A.1. Global Parameters

Parameter	Specification
Data rate	25.78125 Gbps (other speeds possible)
POWER9 I/O Power Supply	<table><tr><th></th><th>Min</th><th>Typ</th><th>Max</th></tr><tr><td>VIO</td><td>0.95V</td><td>1.0V</td><td>1.125V</td></tr></table> VDN(logic supply) = 0.7V nominal and is adaptive
POWER9 Temperature (junction)	TJ -10C to 85C
Termination	85 Ohms differential at Rx , POWER9 TX is SST with 42.5 Ohms from VIO to PAD and 42.5 Ohms from PAD to GND (85 Ohms is a deviation from the OIF CEI 28G SR specification which defines the nominal impedance as 100 Ohms)
ESD	1000V HBM, 200V CDM (This is outside of the OIF CEI SR spec)
POWER9 PHY Latency	Core to TX driver output: 17 UI; RX sampler to Core interface:17 UI

A.2. IBM POWER9 Nominal Estimated Power Dissipation

The power budget breakdown at 25 Gbps for some of the major blocks within a IOLANE are shown below. Numbers reported below for some of the shared blocks outside of the IOLane, such as the PLL, are amortized, per lane values. The overall link efficiency is targeted around 5-6 pJ/b at 25 Gbps (1Tx + 1Rx + PLL). The PLL amortized over 24 lanes in the POWER9 Processor.

A.3. POWER9 Transmitter

A.3.1. Electrical Spec: POWER9 Tx

Table A.2. Electrical Spec: POWER9Tx

Symbol	Parameter	Min	Typ	Max	Unit	Notes
F_{BAUD}	Baud rate		25.78125		Gbps	
$V_{TX-DIFF-PP}$	Differential p-p Tx voltage swing	900	1000	1100	mV	TX launch into a DC matched 85 ohm system
$V_{TX\_CM\_DC}$	Tx Output CM	0.475	0.5	0.562		CM during normal operation and during RxCAL mode. For DC coupled systems
$V_{TX-CM-AC}$	AC CM ripple			25	mVp	Measured up to Fbaud.
V_{DIFF_RxCAL}		-10	0	10	mV	Differential voltage during RxCAL.
RTX_{RxCAL}	Impedance during RxCAL	95		170	Ohms	
TR_{TX}	Rise/fall time at Tx			TBD	psec	20% - 80 %, In TX Sparms

Symbol	Parameter	Min	Typ	Max	Unit	Notes
TX_S21	Insertion loss at Nyquist	-0.5			dB	POWER9 TX has boost, from simulation.
TX FIR	Tx de-emphasis: C_{-1},C_0, C_{+1}					$\|C_0\| + \|C_{+1}\| + \|C_{-1}\| = 1$ Ci resolution to be better than 1/36. (POWER9 does not implement C_{+1})

A.3.2. POWER9 TX Jitter Terms

The jitter terms are specified at the PAD or C4 the chip carrier organic package is considered part of the channel.

Note

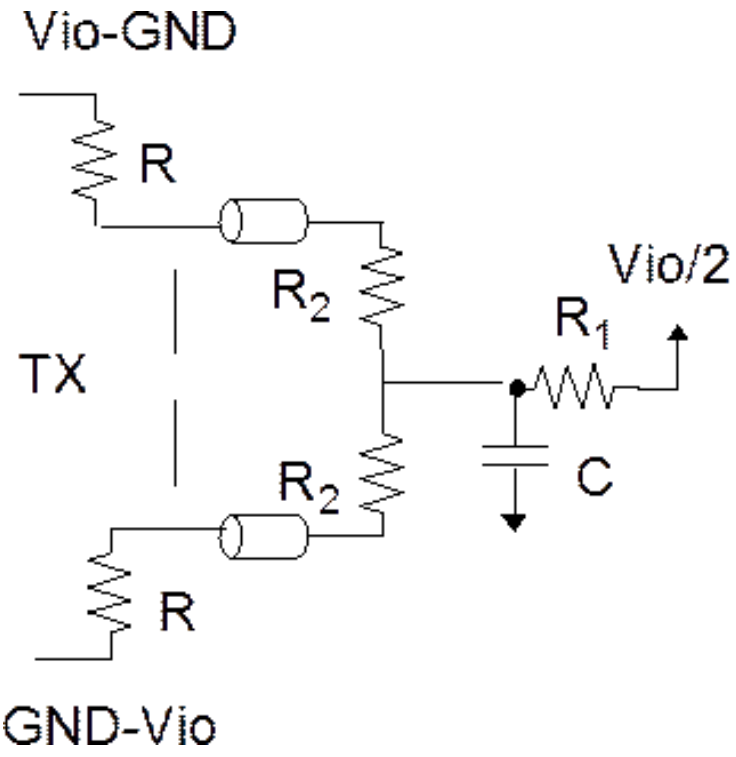
Jitter terms for inclusion in channel sims for POWER9 can be provided.

POWER9 TX silicon S-parameter models are available upon request.

See Section A.5, “Contacting IBM for Assistance” [57].

A.3.3. POWER9 TX SST termination picture excluding Tcoils and ESD affects

Figure A.1. POWER9 TX SST termination picture excluding Tcoils and ESD affects added to the RX above



A.4. POWER9 Receiver

A.4.1. Electrical Spec: Rx

Below is the characteristics of the POWER9 Rx design this does not preclude other equalization method of choice.

The design must be able to equalize the channel without a back channel to TX_EQ like PCI-G3/G4 and the POWER9 TX will only have a pre-cursor.

Table A.3. Rx Electrical Spec

Symbol	Parameter	Min	Typ	Max	Unit	Notes
$V_{RX-DIFF-PP}$	Diff Rx p-p voltage			1100	mVppd	
BER	Target BER			10^{-15}		Extrapolated where needed.
VN_{RX}	Input referred electrical noise			4.0	mVrms	Integrated from 1 MHz
JTOL(f)	Jitter tolerance. F_jitter > Fbaud/200 F_jitter < Fbaud/25000			0.1 10	UIpp UIpp	Measured at 10^{-15} BER More details in template.
V_{RX-SEN}	Receiver input sensitivity			50	mVppd	Measured from breakout board.
$CTLE_{BOOST}$	Rx CTLE boost ratio			12	dB	See curves
H0	DFE Target	50	100	150	mVpp	
Hi i = 1	DFE tap weight H1	0		0.7	*h0	Resolution of Hi to be better than h0/48. (POWER9 RX implements only H1)
$LTEQ_{ZERO}$			200		MHz	Long tail EQ
$LTEQ_{POLE}$			333		MHz	Long tail EQ
LTE_{PEAK}				3.5	dB	Gain range of LTE (LTE will be fixed) DC is at -3.5dB
V_{RX-CM}	Input CM voltage (DC)	0.45	0.5	0.562	V	Floating termination.
CDR BW	CDR BW of first order loop		10		MHz	Some programmability
CDR Peaking	Peaking/overshoot			0.5	dB	
$T_{RX-LATENCY}$	Rx data path latency			24	UI	Latency defined from arrival of bit at Rx input to rising edge of.
P-N skew	In pair skew			0.25	UI	The POWER9 RX can tolerate 25% P-N before additive jitter occurs

A.4.2. POWER9 RX jitter terms

The jitter terms are specified at the PAD or C4 the chip carrier organic package is considered part of the channel.

Note

Jitter terms for inclusion in channel sims for POWER9 can be provided.

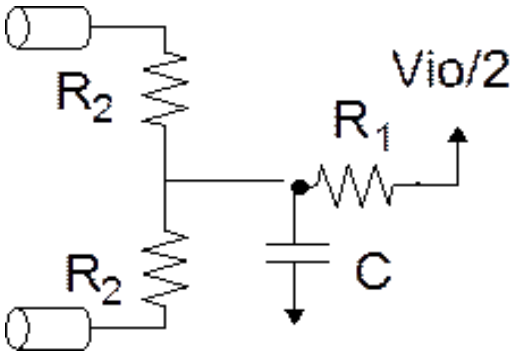
RX silicon S-parameter models are available upon request inclusive of CTLE variances (many S paramters)

See Section A.5, “Contacting IBM for Assistance” [57].

A.4.3. POWER9 Rx termination picture excluding Tcoils and ESD affects.

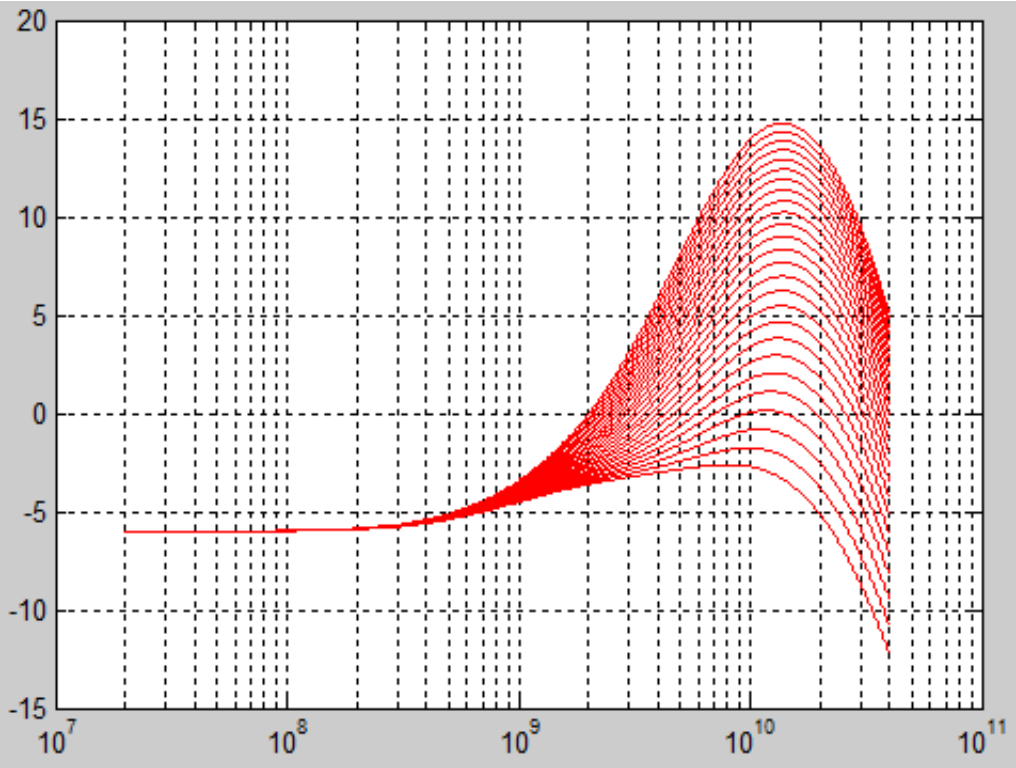
For AC coupling we have a Vcc/2 bias method and the OpenCAOI protocol has 64/66 coding for clock compensation and scrambling to limit baseline wander.For AC coupling we have a Vcc/2 bias method and the OpenCAPI protocol has 64/66 coding for clock compensation and scrambling to limit baseline wander.

Figure A.2. POWER9 Rx termination picture excluding Tcoils and ESD affects.



A.4.4. POWER9 RX CTLE+LFEQ suite of curve in S parameter form

Figure A.3. POWER9 RX CTLE+LFEQ suite of curve in S parameter form



A.5. Contacting IBM for Assistance

Support for developers within the OpenPOWER ecosystem may be obtained from IBM. Information is available on the IBM Portal for OpenPOWER https://www-335.ibm.com/systems/power/openpower/.

To email IBM Support for OpenPOWER: <openpower@us.ibm.com>

Appendix B. References

B.1. OIF CEI 28Gbps SR Specification

The OIF CEI 28Gbps SR Specification was referenced in this specification document and in some cases copied. To include the OIF content in this document, the following notices are required.

Appendix C. OpenPOWER Foundation overview

The OpenPOWER Foundation was founded in 2013 as an open technical membership organization that will enable data centers to rethink their approach to technology. Member companies are enabled to customize POWER CPU processors and system platforms for optimization and innovation for their business needs. These innovations include custom systems for large or warehouse scale data centers, workload acceleration through GPU, FPGA or advanced I/O, platform optimization for SW appliances, or advanced hardware technology exploitation. OpenPOWER members are actively pursing all of these innovations and more and welcome all parties to join in moving the state of the art of OpenPOWER systems design forward.

To learn more about the OpenPOWER Foundation, visit the organization website at openpowerfoundation.org.

C.1. Foundation documentation

Key foundation documents include:

- *Bylaws of OpenPOWER Foundation*
- *OpenPOWER Foundation Intellectual Property Rights (IPR) Policy*
- *OpenPOWER Foundation Membership Agreement*
- *OpenPOWER Anti-Trust Guidelines*

More information about the foundation governance can be found at openpowerfoundation.org/about-us/governance.

C.2. Technical resources

Development resouces fall into the following general categories:

- Foundation work groups
- Remote development environments (VMs)
- Development systems
- Technical specifications
- Software
- Developer tools

The complete list of technical resources are maintained on the foundation Technical Resources web page.

C.3. Contact the foundation

To learn more about the OpenPOWER Foundation, please use the following contact points:

- General information -- <info@openpowerfoundation.org>
- Membership -- <membership@openpowerfoundation.org>
- Technical Work Groups and projects -- <tsc-chair@openpowerfoundation.org>
- Events and other activities -- <admin@openpowerfoundation.org>
- Press/Analysts -- <press@openpowerfoundation.org>

More contact information can be found at openpowerfoundation.org/get-involved/contact-us.