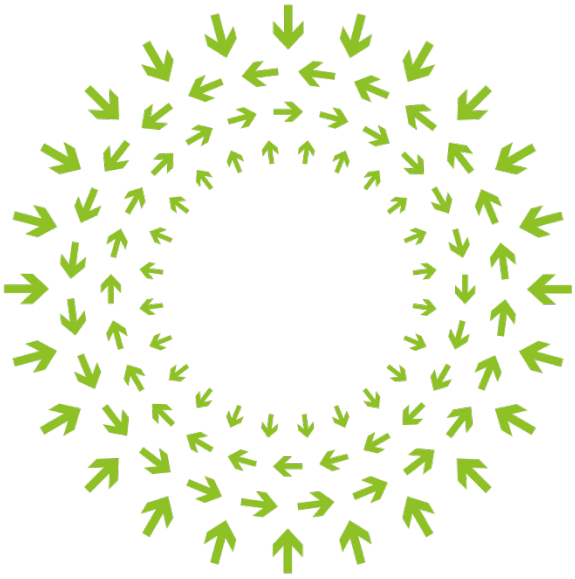

OPEN

Compute Project

Fusion-io
3.2 TB I/O Accelerator
Hardware v0.5

1 Scope

This document defines the technical specifications for the Fusion-io 3.2TB I/O Accelerator used in Open Compute Project servers.

2 Contents

3 Document Overview

3.1 License

As of January 16, 2013, the following persons or entities have made this Specification available under the Open Web Foundation Contributor License Agreement, which is available at http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owf-contributor-license-agreement-1-0---copyright-and-patent

Fusion-io, Inc.

You can review the signed copies of the Open Web Foundation Contributor License Agreement for this Specification at http://opencompute.org/licensing/, which may also include additional parties to those listed above.

Your use of this Specification may be subject to other third party rights. THIS SPECIFICATION IS PROVIDED "AS IS." The contributors expressly disclaim any warranties (express, implied, or otherwise), including implied warranties of merchantability, non-infringement, fitness for a particular purpose, or title, related to the Specification. The entire risk as to implementing or otherwise using the Specification is assumed by the Specification implementer and user. IN NO EVENT WILL ANY PARTY BE LIABLE TO ANY OTHER PARTY FOR LOST PROFITS OR ANY FORM OF INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION OR ITS GOVERNING AGREEMENT, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND WHETHER OR NOT THE OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

3.2 Document Purpose

When data center design and hardware design move in concert, they can improve efficiency and reduce power consumption. To this end, the Open Compute Project is a set of technologies that reduces energy consumption and cost, increases reliability and choice in the marketplace, and simplifies operations and maintenance. One key objective is openness—the project is starting with the opening of the specifications and mechanical designs for the major components of a data center, and the efficiency results achieved at facilities using Open Compute technologies.

One component of this project is a high-density I/O accelerator PCI Express adapter card. This document describes the Open Compute Project Fusion-io I/O accelerator board design, which supports up to 3.2TB memory. The board design is capacity-optimized, designed to provide the highest capacity in a PCI Express card full height, half length form factor. Many features found in traditional PCI Express cards have been removed from this board design.

3.3 Document Scope

This document focuses on the hardware specifications for the 3.2TB I/O Accelerator and does not dictate requirements for any future releases. This version of the document supersedes any previous versions and is subject to revision.

3.4 Reference Documents

PCI EXPRESS BASE SPECIFICATION, Revision 2.0, December 20, 2006

PCI EXPRESS CARD ELECTROMECHANICAL SPECIFICATION, Revision 2.0, April 11, 2007

4 Introduction

3.2TB I/O Accelerator is based on the Fusion-io I/O technology, but provides increased bandwidth and higher memory density while consuming a similar amount of power. It is modular and is constructed of 2 board types: a Controller which provides system management/control and Tall Flashpaks (TPak) which contain NAND flash memory.

The primary distinguishing features of this product are:

- PCIe 2.x compliant
- Consists of 2 board types for NAND package/type flexibility:
  - Controller – receives PCIe signals and contains product logic
  - TPak (tall NAND Module) – contains MLC NAND Flash memory (attaches to Controller)
- Large memory capacity
- Low power consumption

5 Product Specifications

5.1 Features & Technologies

- Usable Capacity: 3.2TB
- PCIe 2.x, 8 lanes
- Modular TPaks for a total of up to 3.8TB NAND (3.2TB usable)
- Compatible with Fusion-io qualified TPaks (NAND Modules)
- 20 NAND configuration for each TPak
- Atmel AVR processors for board management
  - MidPROM (configuration)
  - LEDs
  - Power good
  - Voltage and current measurements
  - Temperature measurements

6 Architectural Block Diagram

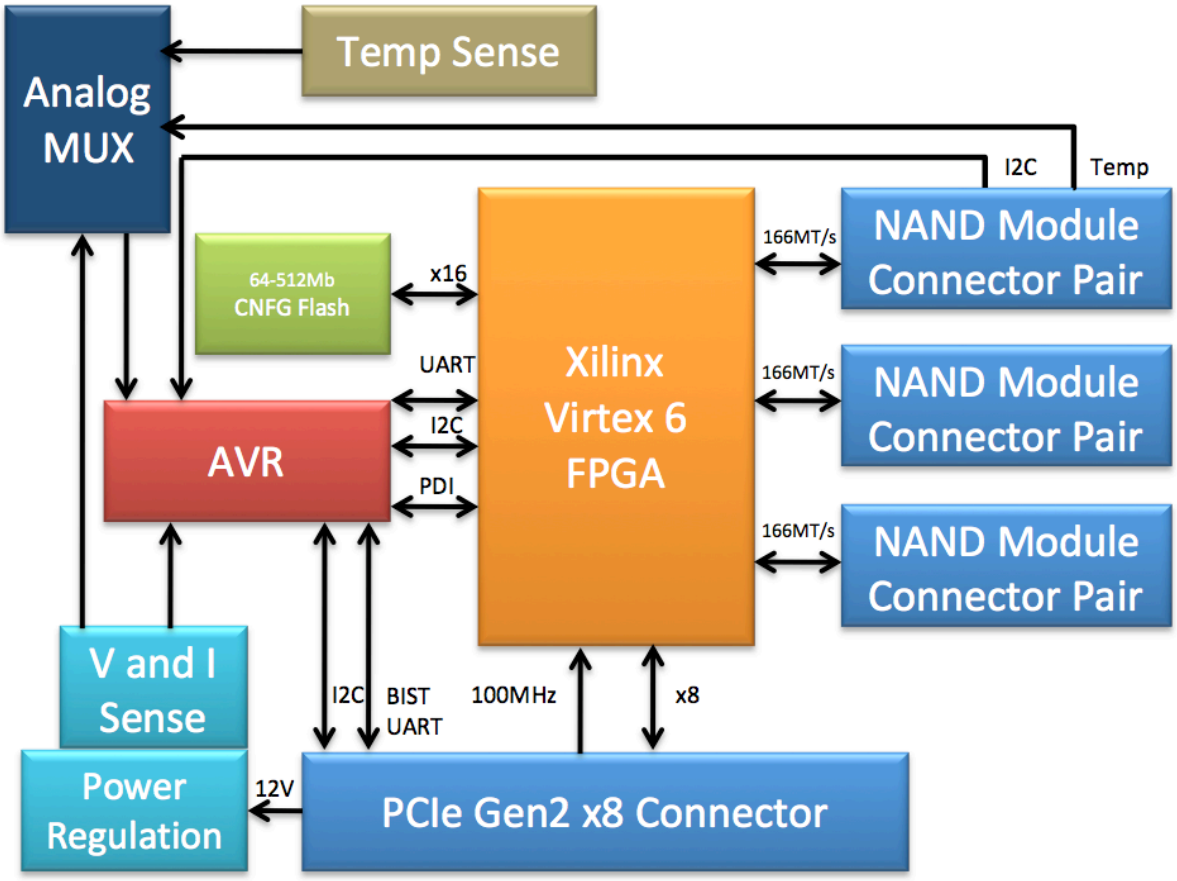

Figure 1 3.2TB I/O Accelerator Block Diagram

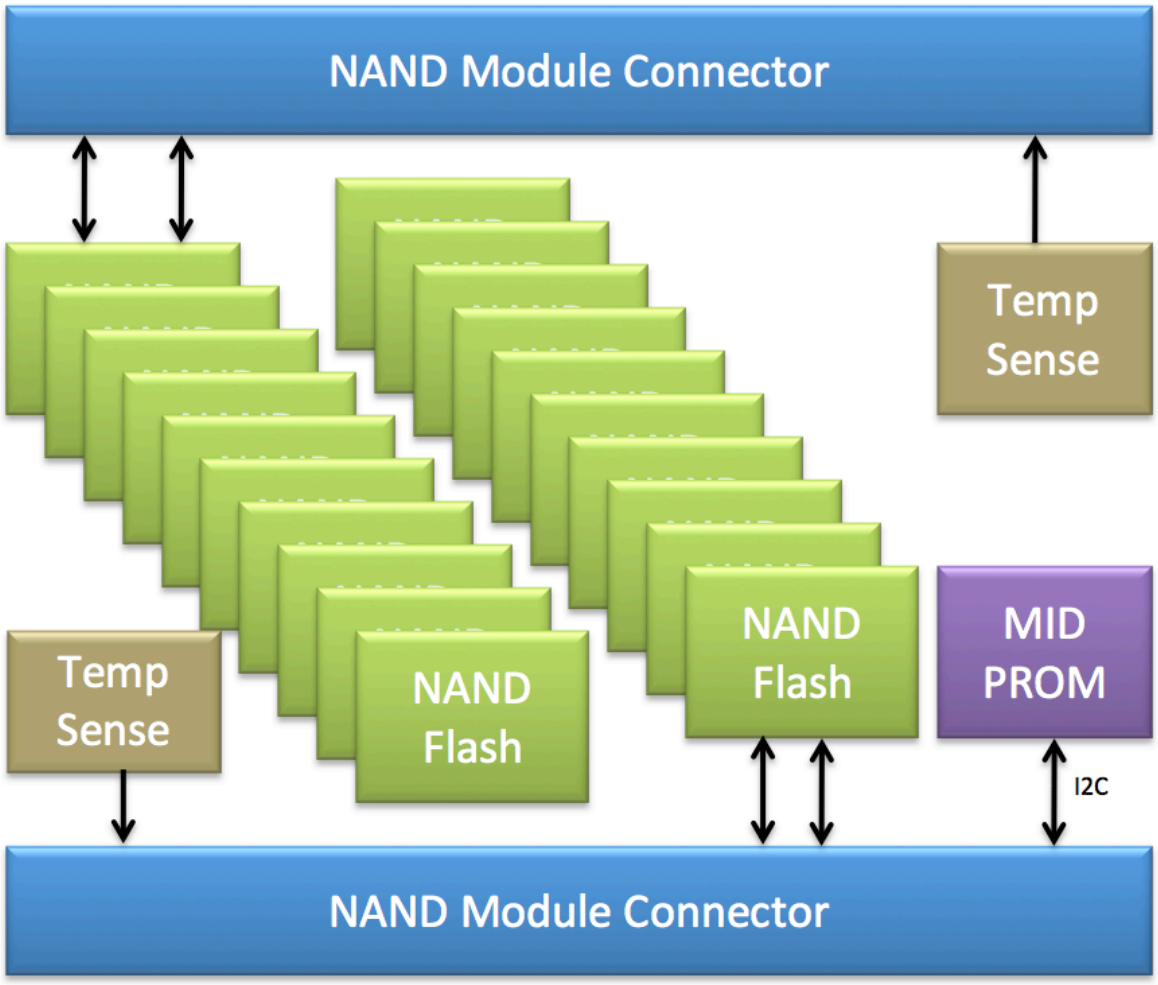

Figure 2 TPak Block Diagram

7 Functional Description

7.1 Board Interfaces

7.1.1 Host Board PCIe Interface

3.2TB I/O Accelerator:

- Operates at up to full bandwidth PCIe Gen2 x8.
- Provides for simultaneous full functionality, including:
  - Full bandwidth PCIe interface at Gen2 x8.
  - Automated firmware self-updates.
  - Management information gathering and reporting: MidPROM, temperature, power, and usage information.
  - Power-cut safe shutdown. Capacitance sufficient to hold power for >6ms following power failure.
- Allows operation with Fusion-io's existing software for cross-platform compatibility.

7.1.2 External Power Cabling

3.2TB I/O Accelerator does not consume more than the specified 25W available per PCIe slot, so an external power cable is not supported.

7.1.3 I2C & PDI Interfaces

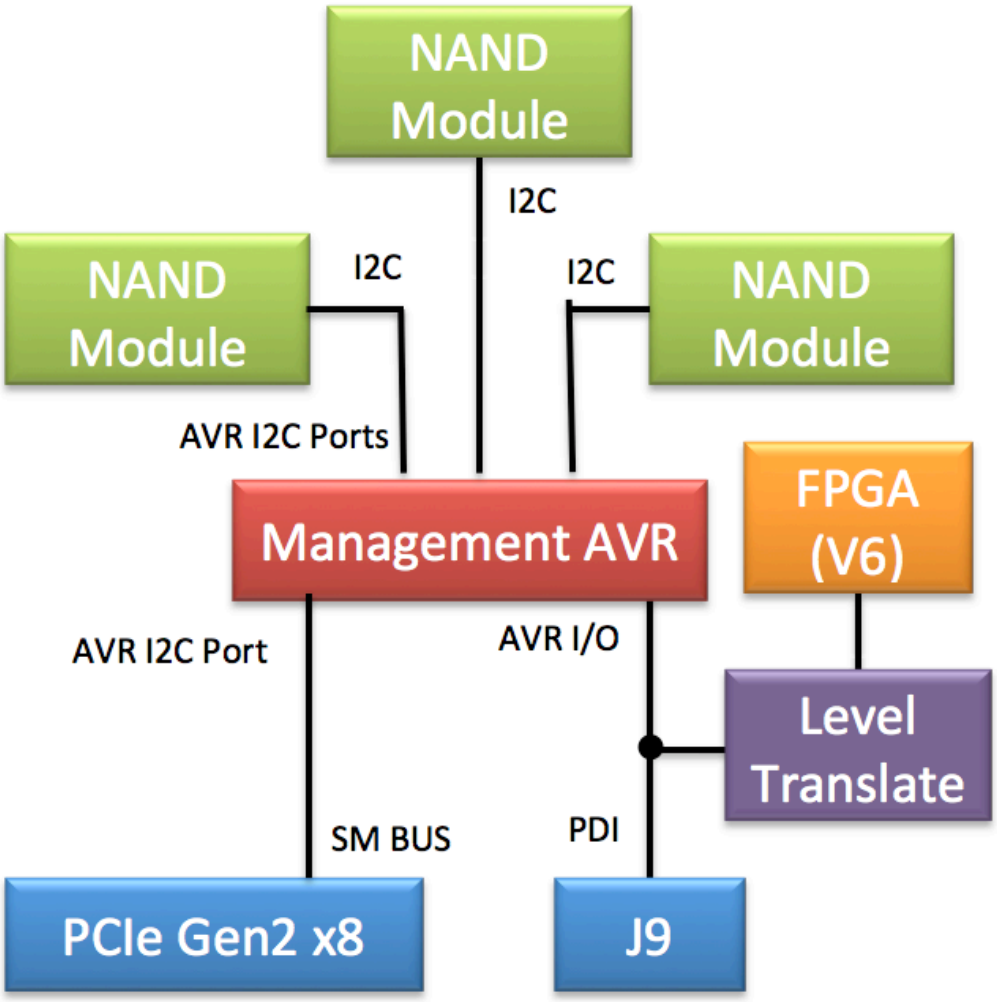

Figure 3 I²C & PDI Block Diagram

- PDI connection through header J9
  - Programming access to the Management AVR
  - Provides communication to the NAND modules
    - Temperatures
    - MidProm information
  - Provides access to controller (FPGA) information
    - Temperatures
    - Voltage levels
    - Current levels
  - Provides communication to FPGA

- I^2C connection through PCIe slot. An isolation circuit cuts off the I^2C if the board is not powered for any reason. When the board is powered, the I^2C automatically connects to the PCIe slot. When connected, the following functionality can be supported.
  - Provides access to ID PROM
  - Provides access to the Management AVR
  - Management AVR can emulate multiple devices

7.1.4 BIST Interface

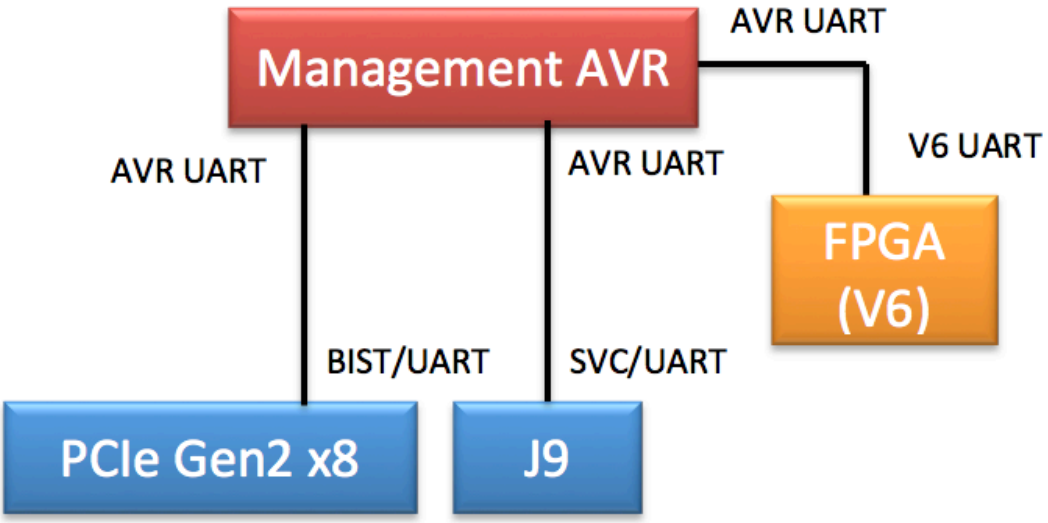

Figure 4 BIST/UART Block Diagram

- BIST/UART connection through header J9
  - Access to the Management AVR
  - Provides access to controller (FPGA) information
  - Provides communication to FPGA

- BIST/UART connection through the PCIe slot. These pins are actually connected to +3.3V on the PCIe connector and cannot be used during normal operation. This control path is only operational when the board is connected to custom hardware created by Fusion-io for factory testing.

7.1.5 JTAG Interface

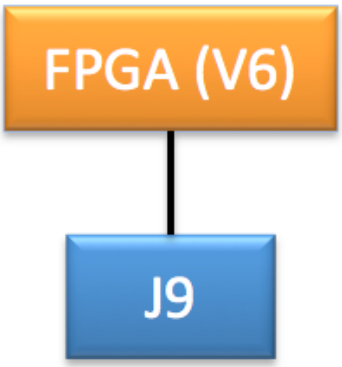

Figure 5 JTAG Interface

1. JTAG connection is made through header J9
   - Provides JTAG communication to the FPGA

7.1.6 Tall Flashpak Interface

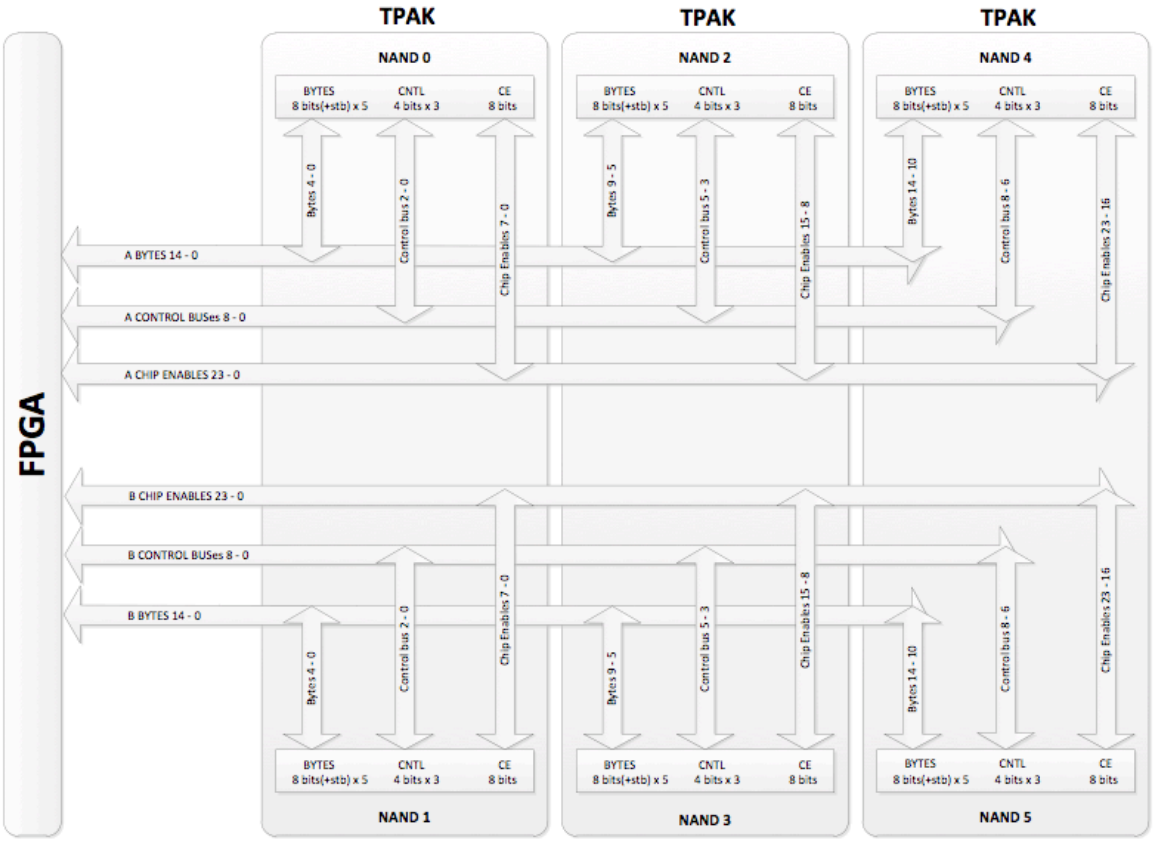

Figure 6 FPGA to Tall Flashpak Interconnects

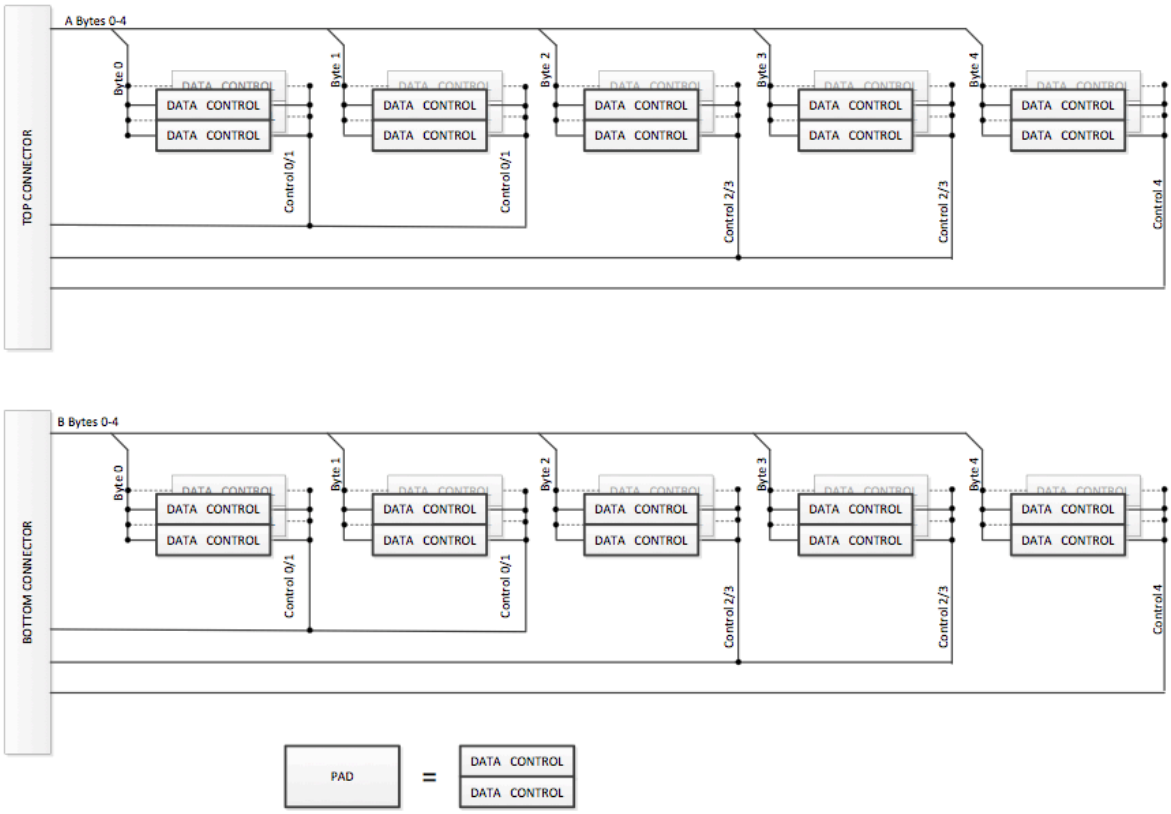

Figure 7 Tall Flashpak Data/Control Distribution

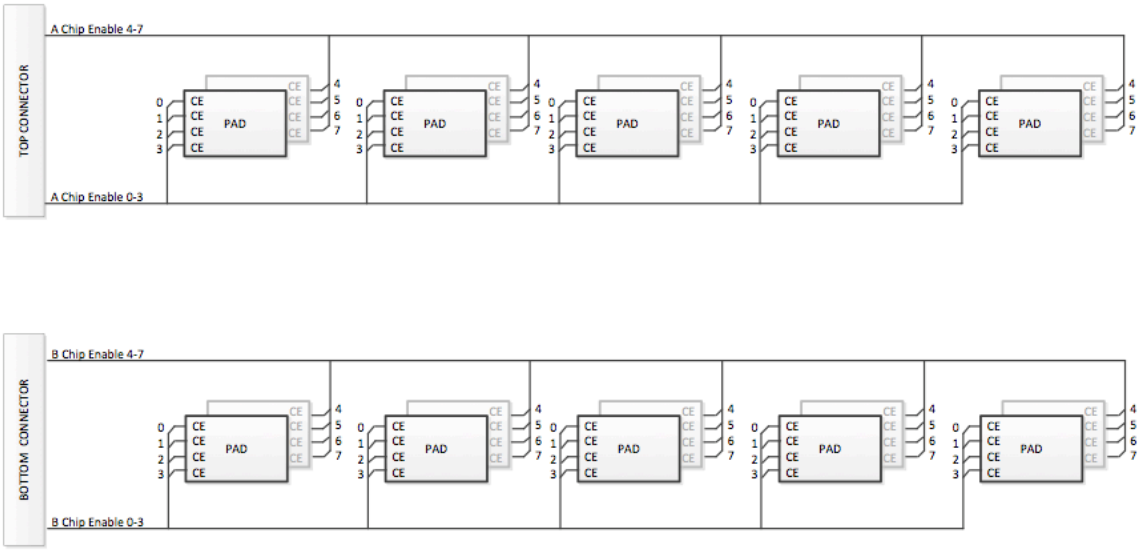

Figure 8 Tall Flashpak Chip Enable Distribution

8 Standard Features

8.1 LEDs :

System Indicators

DS1, DS2, and DS3 (LEDs that can be seen through the PCIe bracket). These are controlled by the Management AVR.

Functions: DS2 (Yellow) – Flashing denotes writes.

DS3 (Green) – Solid : denotes driver working.

Flashing: denotes reads.

DS1 (Amber) – Denotes a problem or driver not loaded.

All three on at the same time denotes "beacon" (used to visually determine card location).

PCIe Link Indicator

Controlled by the Management AVR.

Debug Indicators

These LEDs are controlled by the Management AVR and are used to provide controller status. They are used for factory test and, typically, are not easily visible when the drive is installed.

D1 (Yellow) – Definition varies.

D2 (Green) – Definition varies, but typically means OK if on.

D3 (Yellow) – Definition varies.

D5 (White) – Definition varies, but typically means the Management AVR is operating properly if pulsing (White).

D6 (Blue) – Definition varies, but typically means the FPGA is programmed if on.

8.2 EEPROM:

U16 – Configuration NOR Flash for the Xilinx FPGA.

8.3 Fan Support:

3.2TB I/O Accelerator provides support for up to two fans via connector J4. Connector J4 supports power and tach feedback for each fan. The tach pins are directly connected to the management AVR for monitoring. This is reserved for future use.

8.4 Headers:

J9 – 14 pin Management Service Header. The pads for this header are also used in board testing. The location of this header is compatible with other Fusion-io boards of a similar form factor.

PIN	Function
1	2.5V
2	GND
3	JTAG_TCK
4	PDI_DATA
5	JTAG_TMS
6	PDI_CLK
7	JTAG_TDI
8	GND
9	JTAG_TDO
10	BIST_RX
11	NC
12	BIST_TX
13	3.3V
14	GND

Table 1: J9 Pin Definitions

J4 – 4 pin Fan Connector. TACH_1 and TACH_2 are directly connected to the management AVR to monitor fan rotation.

PIN	Function
1	GND
2	TACH_2
3	TACH_1
4	+12V (FAN)

Table 2: J4 Pin Definitions

8.5 Temperature Sensors:

There are a total of 9 temperatures that can be monitored.

There are two thermistors on the controller: R31 and R111. R31 is located on the primary side by the power supplies, and R111 is on the secondary side near the FPGA.

Each TPak also contains two thermistors, for a total of 6.

The FPGA junction temperature is also available.

8.6 Power-Cut:

3.2TB I/O Accelerator has local capacitance enough to flush all data (>6ms hold up power) in the event of loss of power.

8.7 ECC:

3.2TB I/O Accelerator implements ECC error correction.

8.8 Adaptive Flashback:

3.2TB I/O Accelerator incorporates adaptive flashback for longer life.

8.9 Power Regulation:

Power sequencing is required. The power comes up in the following order:

1. 12V comes from the PCIe slot or external connector
2. 3.3V – AVR and NAND core
3. 1.0V – FPGA core voltage
4. 2.5V – FPGA auxiliary voltage
5. 1.8V – FPGA and NAND I/O
6. 1.03V and 1.2V – MGT

8.9.1 Power Block Diagram

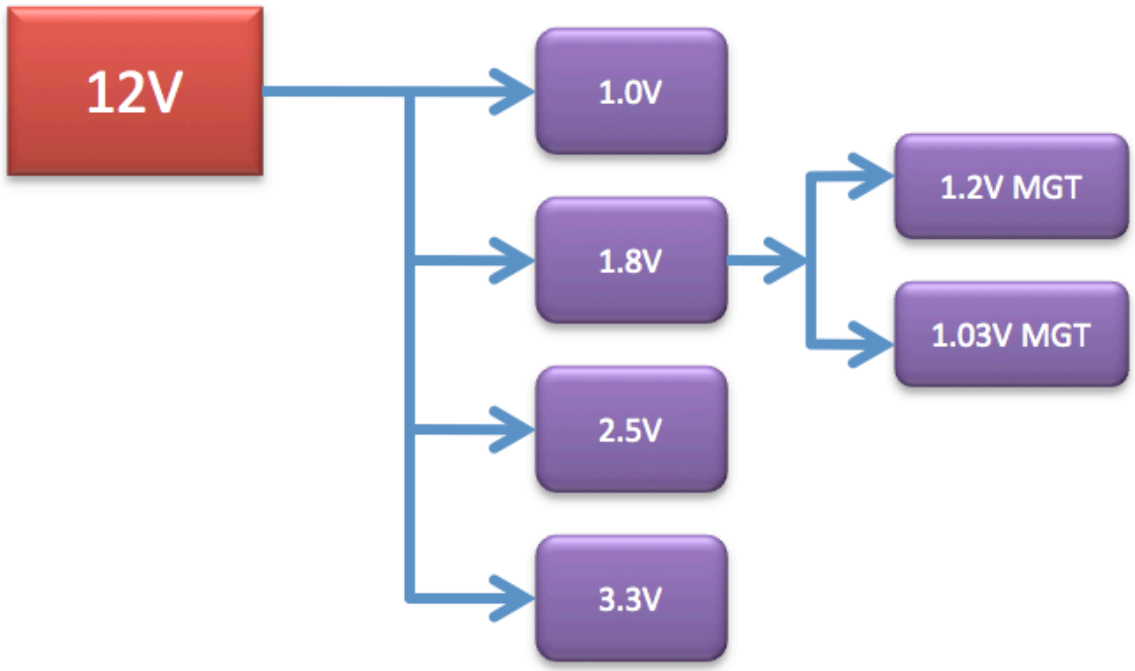

Figure 9 Power

8.10 Management

A management system is implemented on 3.2TB I/O Accelerator which provides:

- Product and revision identification
- Temperature reporting of controller and NAND modules, and of each system chip capable of thermal reporting
- Power reporting of each primary system voltage and current
- Programmable tuning of system variables
- Diagnostics information

9 Layout

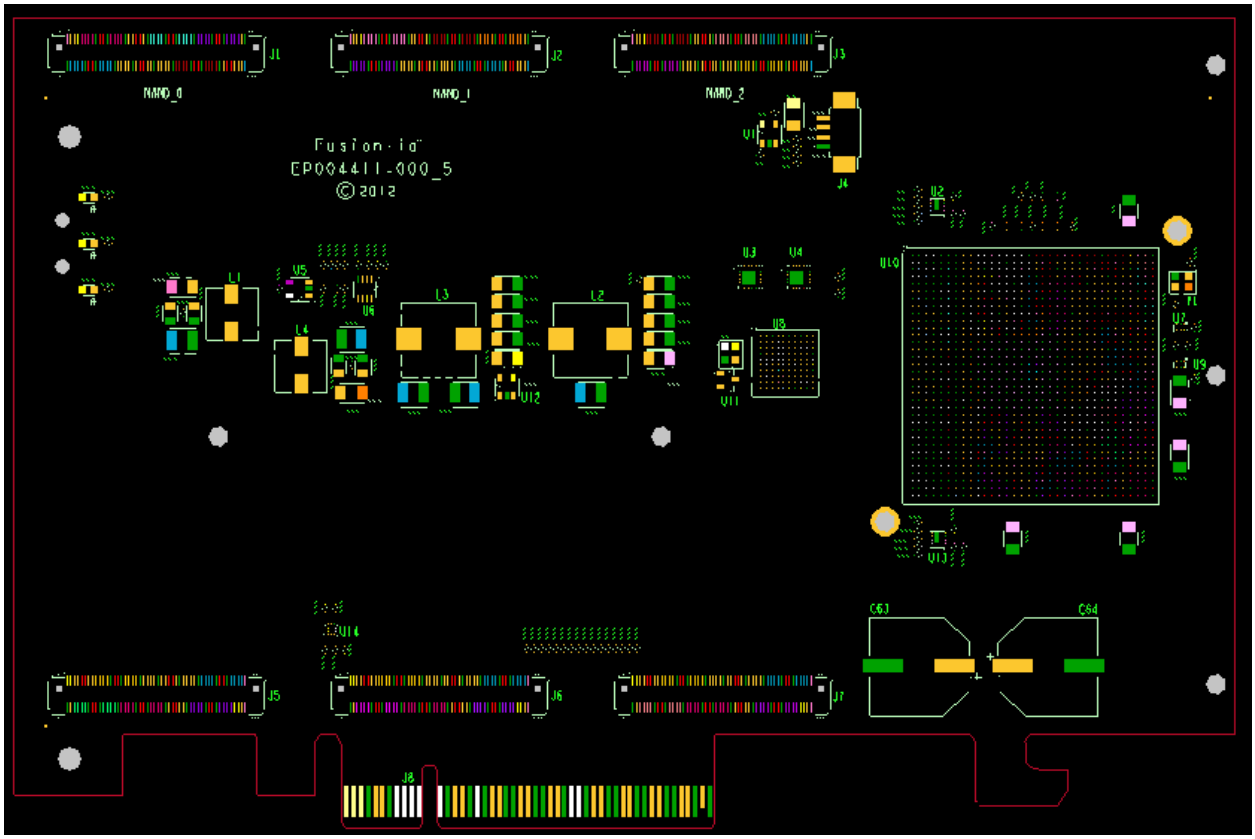

Figure 10 Controller Top Side Layout

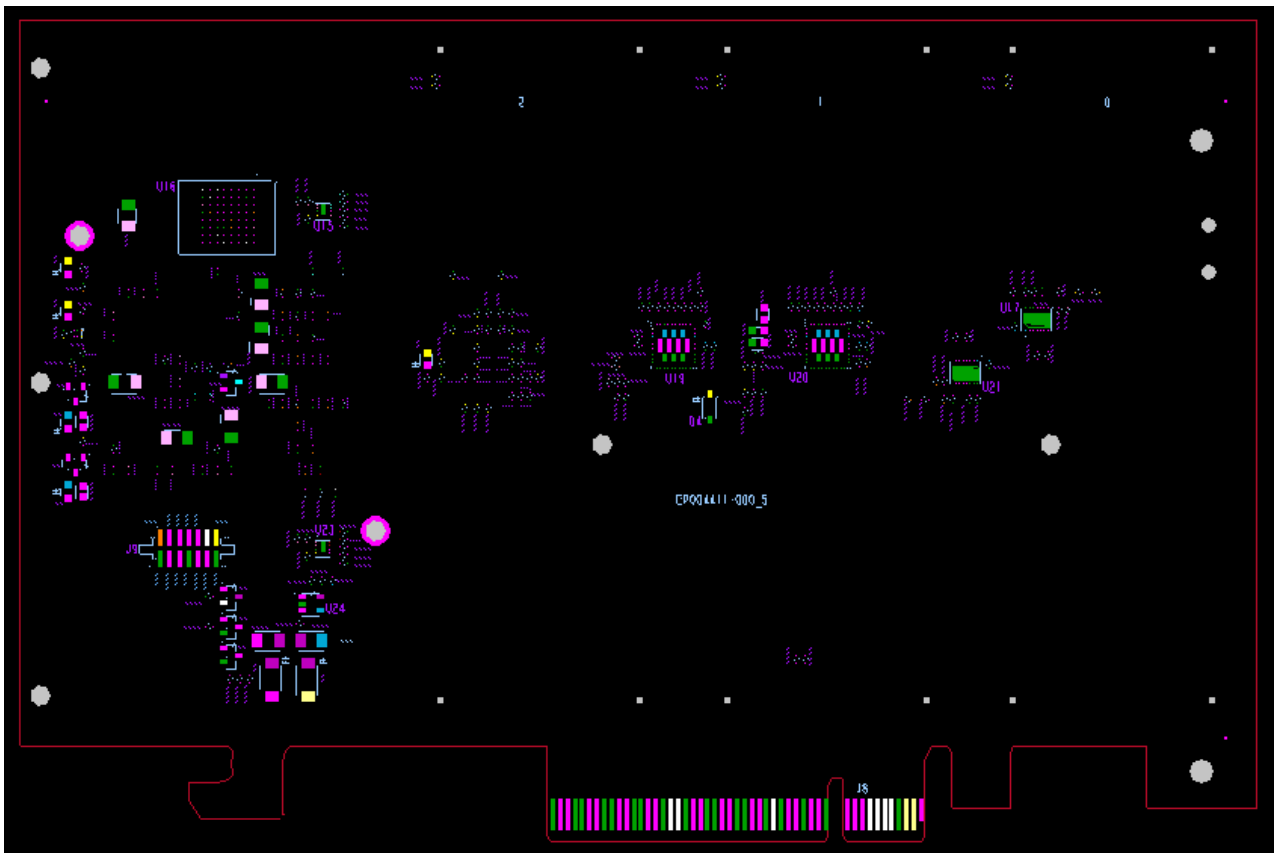

Figure 11 Controller Bottom Side Layout

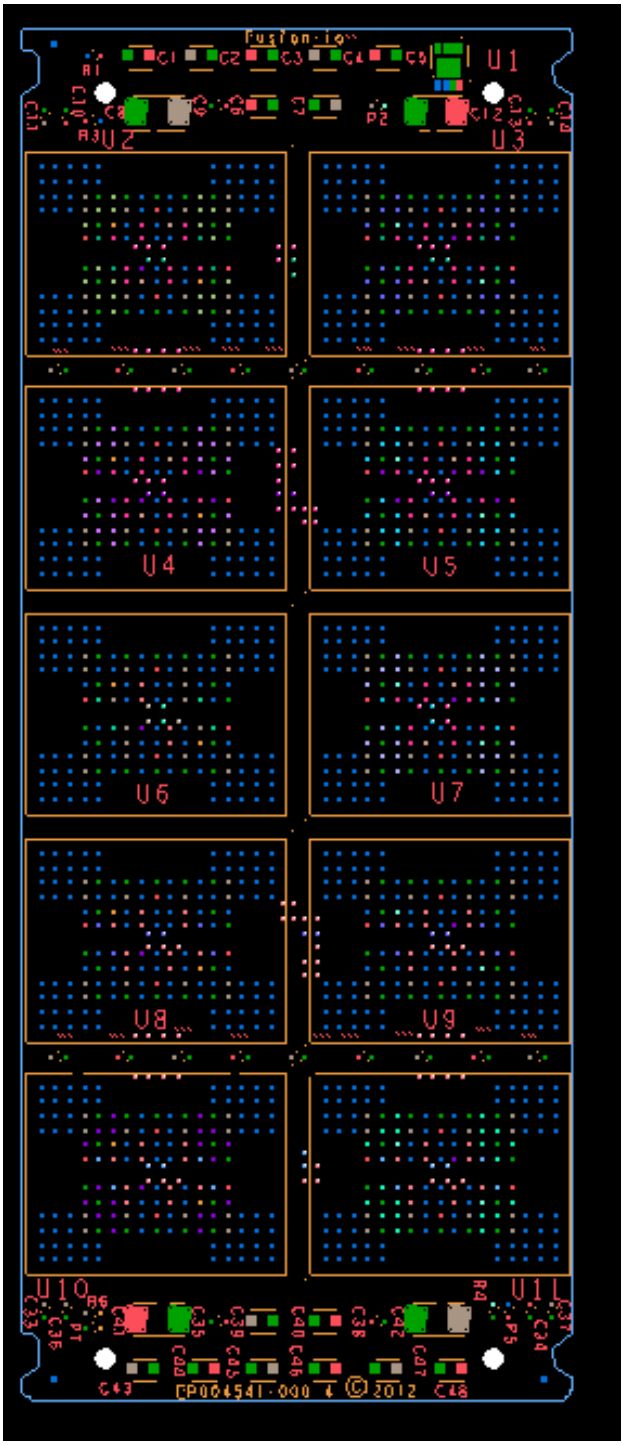

Figure 12 NAND Module Top Side

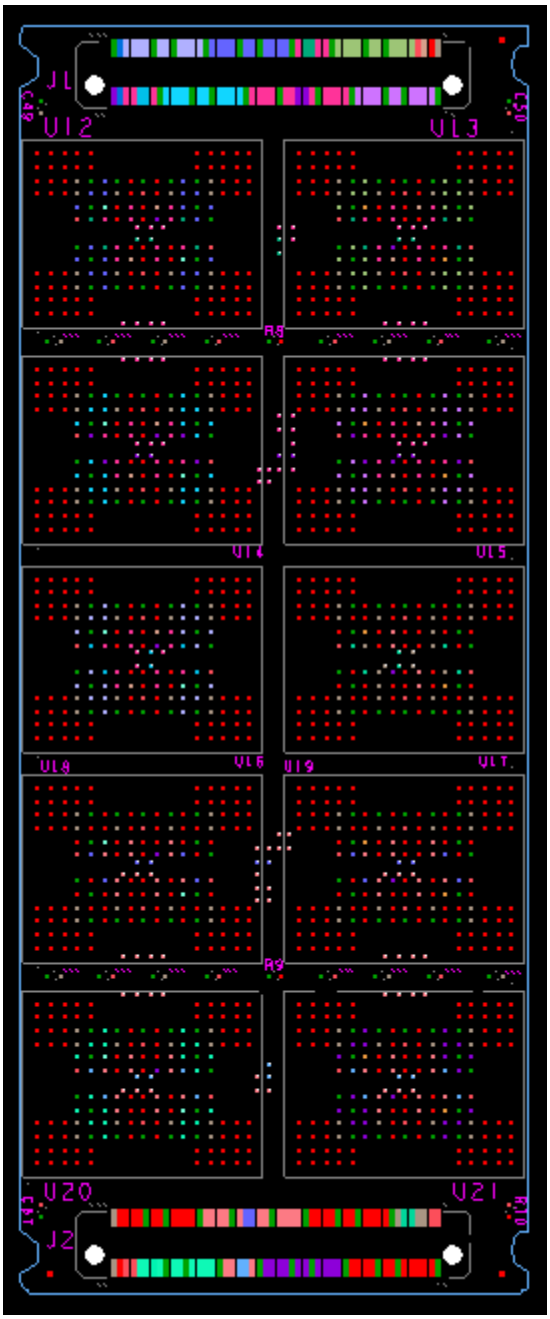

Figure 13 NAND Module Bottom Side

10 Physical

10.1 Conforms to PCI Express Specification

3.2TB I/O Accelerator conforms to the PCIe specification for half-length, full-height cards (see documentation sited). The board's overall dimensions are as defined in the standard Card Electromechanical Specification. Note that this card has the "hockey stick" shaped retention feature used for increased mechanical rigidity when in a socket designed to accept this.

10.2 Thermal Constraints

3.2TB I/O Accelerator operates with inlet temperatures up to 55°C. The NAND on the TPaks are designed to operate at up to 70°C. Cooling is sufficient to keep all components within the card below their thermal limits with 55°C air at 300 LFM. 3.2TB I/O Accelerator relies solely on external airflow provided by the server for cooling.

Temperature variation will not exceed 5 degrees Celsius at rated airflow and maximum inlet temperature.

10.3 Reliability

- MTBF > 2,000,000 hours
- UBER <= 1e-20
- Availability > 99.99% (a.k.a. 4-9's)
- Serviceability Objectives:
  - 90% first fault isolation
  - Mean-time-to-repair <= 15 minutes
  - Ability to hold >= 2 firmware images, with distinct download and activate functions for firmware upgrade
- Proprietary wear-leveling is implemented to maximize NAND life
- Retention >= 3 months
- Endurance >= 20 PBW
- **Important Note: Endurance and Retention rates are preliminary and subject to change**

10.4 Weight

10.9 oz. (310 g) with modules, 6.7 oz. (190 g) without modules

10.5 Logos and Silk

Per marketing and OEM.

11 Environmental

- 0°C to 55°C operating temperature (ambient air @ 300 LFM)
- De-rated 1°C per 1000 feet above sea level
- 5% to 95% operating humidity (non-condensing)
- -40°C to 70°C storage temperature
- 5% to 95% storage humidity (non-condensing)
- Typical power dissipation: 19W (Est.)
- Maximum power dissipation: 25W
- Max operating altitude: 10,000 feet
- Max non-operating altitude: 30,000 feet

12 Certifications and Regulatory Compliance

3.2TB I/O Accelerator complies with the following electrical device standards:

• ANSI C63.4/EN 55022/CNS13438	US/Canada
• Radiated and Conducted Emissions Class B	US/Canada
• EN55024 Immunity, EN55022 Class B	US/Canada
• 2004/108/EC EMC Directive CE	Europe
• IEC 61000 Class B Mark	Europe
• VCCI – V-2/2009.04	Japan
• BSMI – CNS 13438 / EN 55022 Class B	Taiwan
• AS/NZS CISPR22:2006 / 47CFR Part 15	Australia/New Zealand
• Radiated and Conducted Emissions Class B	Australia/New Zealand
• RoHS – EU Directive 2002/95/EC	RoHS

13 Testability

13.1 Environmental Product Testing

MIL-STD-810 Transportation Profile: Shock, vibration and thermal testing, non-operational.