Chipsmall Limited consists of a professional team with an average of over 10 year of expertise in the distribution of electronic components. Based in Hongkong, we have already established firm and mutual-benefit business relationships with customers from,Europe,America and south Asia,supplying obsolete and hard-to-find components to meet their specific needs.

With the principle of “Quality Parts,Customers Priority,Honest Operation,and Considerate Service”,our business mainly focus on the distribution of electronic components. Line cards we deal with include Microchip,ALPS,ROHM,Xilinx,Pulse,ON,Everlight and Freescale. Main products comprise IC,Modules,Potentiometer,IC Socket,Relay,Connector.Our parts cover such applications as commercial,industrial, and automotives areas.

We are looking forward to setting up business relationship with you and hope to provide you with the best service and solution. Let us make a better world for our industry!

Contact us

Tel: +86-755-8981 8866 Fax: +86-755-8427 6832
Email & Skype: info@chipsmall.com Web: www.chipsmall.com
Address: A1208, Overseas Decoration Building, #122 Zhenhua RD., Futian, Shenzhen, China

PI7C9X2G304EL

PCI EXPRESS GEN 2 PACKET SWITCH

3-Port, 4-Lane, ExtremeLo PCIe2.0 Packet Switch

DATASHEET

REVISION 2-2

September 2017

A Product Line of Diodes Incorporated

1545 Barber Lane Milpitas, CA 95035
Telephone: 408-232-9100
FAX: 408-434-1040
Internet: http://www.diodes.com

IMPORTANT NOTICE

DIODES INCORPORATED MAKES NO WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, WITH REGARDS TO THIS DOCUMENT, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE (AND THEIR EQUIVALENTS UNDER THE LAWS OF ANY JURISDICTION).

Diodes Incorporated and its subsidiaries reserve the right to make modifications, enhancements, improvements, corrections or other changes without further notice to this document and any product described herein. Diodes Incorporated does not assume any liability arising out of the application or use of this document or any product described herein; neither does Diodes Incorporated convey any license under its patent or trademark rights, nor the rights of others. Any Customer or user of this document or products described herein in such applications shall assume all risks of such use and will agree to hold Diodes Incorporated and all the companies whose products are represented on Diodes Incorporated website, harmless against all damages.

Diodes Incorporated does not warrant or accept any liability whatsoever in respect of any products purchased through unauthorized sales channel.
Should Customers purchase or use Diodes Incorporated products for any unintended or unauthorized application, Customers shall indemnify and hold Diodes Incorporated and its representatives harmless against all claims, damages, expenses, and attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized application.

Products described herein may be covered by one or more United States, international or foreign patents pending. Product names and markings noted herein may also be covered by one or more United States, international or foreign trademarks.

This document is written in English but may be translated into multiple languages for reference. Only the English version of this document is the final and determinative format released by Diodes Incorporated.

LIFE SUPPORT

Diodes Incorporated products are specifically not authorized for use as critical components in life support devices or systems without the express written approval of the Chief Executive Officer of Diodes Incorporated. As used herein:

A. Life support devices or systems are devices or systems which:

1. are intended to implant into the body, or

2. support or sustain life and whose failure to perform when properly used in accordance with instructions for use provided in the labeling can be reasonably expected to result in significant injury to the user.

B. A critical component is any component in a life support device or system whose failure to perform can be reasonably expected to cause the failure of the life support device or to affect its safety or effectiveness.

Customers represent that they have all necessary expertise in the safety and regulatory ramifications of their life support devices or systems, and acknowledge and agree that they are solely responsible for all legal, regulatory and safety-related requirements concerning their products and any use of Diodes Incorporated products in such safety-critical, life support devices or systems, notwithstanding any devices- or systems-related information or support that may be provided by Diodes Incorporated. Further, Customers must fully indemnify Diodes Incorporated and its representatives against any damages arising out of the use of Diodes Incorporated products in such safety-critical, life support devices or systems.

www.diodes.com

REVISION HISTORY

Date	Revision Number	Description
03/13/14	0.1	Preliminary Datasheet
07/18/14	1.0	Updated Section 3 Pin Description Updated Section 4 Pin Assignments
11/17/14	1.1	Updated Section 7.2 Transparent Mode Configuration Registers Updated Section 8 Clock Scheme
07/16/15	1.2	Updated Section 3.1 PCI Express Interface Signals Updated Section 3.2 Port Configuration Signals Updated Section 5.1 Physical Layer Circuit Updated Section 6.1 EEPROM Interface Updated Section 7.2 Transparent Mode Configuration Registers Updated Section 8 Clock Scheme Updated Table 9-1 Instruction Register Codes Updated Table 9-2 JTAG Device ID Register Updated Table 9-3 JTAG Boundary Scan Register Definition Updated Table 11-2 DC Electrical Characteristics
09/11/15	1.3	Updated Table 11-1 Absolute Maximum Ratings
12/23/15	1.4	Updated Section 3.1 PIN Description Updated Table 11-1 Absolute Maximum Ratings Updated Table 11-2 DC Electrical Characteristics
02/26/16	1.5	Added Section 11 Power Sequence
05/05/16	1.6	Updated Section 3.2 Port Configuration Signals
09/05/17	2-2	Updated Section 1 Features Updated Section 3.2 Port Configuration Signals Updated Section 5.1 Physical Layer Circuit Updated Section 6.1.4 Mapping EEPROM Contents to Configuration Registers Updated Section 7.2 Transparent Mode Configuration Registers Updated Section 12.1 Absolute Maximum Ratings Updated Table 12-2 DC Electrical Characteristics Added Section 12.4 Operating Ambient Temperature Added Section 12.5 Power Consumption Revision numbering system changed to whole number

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

1 FEATURES

- 4-lane PCI Express Gen 2 Switch with 3 PCI Express ports
- Supports "Cut-through"(Default) as well as "Store and Forward" mode for packet switching
- Peer-to-peer switching between any two downstream ports
- 150 ns typical latency for packet routed through Switch without blocking
- Integrated reference clock for downstream ports
- Strapped pins configurable with optional EEPROM or SMBus
- SMBus interface support
- Compliant with System Management (SM) Bus, Version 1.0
- Compliant with *PCI Express Base Specification Revision 2.1*
- Compliant with *PCI Express CEM Specification Revision 2.0*
- Compliant with *PCI-to-PCI Bridge Architecture Specification Revision 1.2*
- Compliant with *Advanced Configuration Power Interface (ACPI) Specification*
- Reliability, Availability and Serviceability
  - Supports Data Poisoning and End-to-End CRC
  - Advanced Error Reporting and Logging
  - IEEE 1149.1 JTAG interface support
- Advanced Power Saving
  - Empty downstream ports are set to idle state to minimize power consumption
- Link Power Management
  - Supports L0, L0s, L1, L2, L2/$L3_{Ready}$ and L3 link power states
  - Active state power management for L0s and L1 states
- Device State Power Management
  - Supports D0, $D3_{Hot}$ and $D3_{Cold}$ device power states
  - 3.3V Aux Power support in $D3_{Cold}$ power state
- Port Arbitration: Round Robin (RR), Weighted RR and Time-based Weighted RR
- Extended Virtual Channel capability
  - Two Virtual Channels (VC) and Eight Traffic Class (TC) support
  - Disabled VCs' buffer is assigned to enabled VCs for resource sharing
  - Independent TC/VC mapping for each port
  - Provides VC arbitration selections: Strict Priority, Round Robin (RR) and Programmable Weighted RR
- Supports Isochronous Traffic
  - Isochronous traffic class mapped to VC1 only
  - Strict time based credit policing
- Supports up to 512-byte maximum payload size
- Programmable driver current and de-emphasis level at each individual port
- Support Access Control Service (ACS) for peer-to-peer traffic
- Support Address Translation (AT) packet for SR-IOV application
- Support OBFF and LTR
- Low Power Dissipation: 650 mW typical in L0 normal mode
- Industrial Temperature Range -40° to 85°C
- 136-pin aQFN 10mm x 10mm package

2 GENERAL DESCRIPTION

Similar to the role of PCI/PCIX Bridge in PCI/PCIX bus architecture, the function of PCI Express (PCIE) Switch is to expand the connectivity to allow more end devices to be reached by host controllers in PCIE serial interconnect architecture. The 4-lane PCIe Switch is in 3-port type configuration. It provides users the flexibility to expand or fan-out the PCI Express lanes based on their application needs.

In the PCI Express Architecture, the PCIE Switch forwards posted and non-posted requests, and completion packets in either downstream or upstream direction concurrently as if a virtual PCI Bridge is in operation on each port. By visualizing the port as a virtual Bridge, the Switch can be logically viewed as two-level cascaded multiple virtual PCI-to-PCI Bridges, where one upstream-port Bridge sits on all downstream-port Bridges. Similar to a PCI Bridge during enumeration, each port is given a unique bus number, device number, and function number by the initiating software. The bus number, device number, and function number are combined to form a destination ID for each specific port. In addition to that, the memory-map and IO address ranges are exclusively allocated to each port as well. After the software enumeration is finished, the packets are routed to the dedicated port based on the embedded address or destination ID. To ensure the packet integrity during forwarding, the Switch is not allowed to split the packets to multiple small packets or merge the received packets into a large transmit packet. Also, the IDs of the requesters and completers are kept unchanged along the path between ingress and egress port.

The Switch employs the architecture of Combined Input and Output Queue (CIOQ) in implementation. The main reason for choosing CIOQ is that the required memory bandwidth of input queue equals to the bandwidth of ingress port rather than increasing proportionally with port numbers as an output queue Switch does. The CIOQ at each ingress port contains separate dedicated queues to store packets. The packets are arbitrated to the egress port based on the PCIe transaction-ordering rule. For the packets without ordering information, they are permitted to pass over each other in case that the addressed egress port is available to accept them. As to the packets required to follow the ordering rule, the Head-Of-Line (HOL) issue becomes unavoidable for packets destined to different egress ports since the operation of producer-consumer model has to be retained; otherwise the system might occur hang-up problem. On the other hand, the Switch places replay buffer at each egress port to defer the packets before sending it out. This can assure the maximum throughput being achieved and therefore the Switch works efficiently. Another advantage of implementing CIOQ in PCIe Switch is that the credit announcement to the counterpart is simplified and streamlined because of the credit-based flow control protocol. The protocol requires that each ingress port maintains the credits independently without checking other ports' credit availability, which is otherwise required by pure output queue architecture.

The Switch supports two virtual channels (VC0, VC1) and eight traffic classes (TC0 ~ TC7) at each port. The ingress port independently assigns packets into the preferred virtual channel while the egress port outputs the packet based on the predefined port and VC arbitration algorithm. For instance, the isochronous packet is given a special traffic class number other than TC0 and mapped into VC1 accordingly. By employing the strict time based credit policy for port arbitration and assigning higher priority to VC1 than VC0, the Switch can therefore guarantee the time-sensitive packet is not blocked by regular traffic to assure the quality of service. In addition, some data-centric applications only carry TC0/VC0 traffic. As a result, there are no packets that would consume VC1 bandwidth. In order to improve the efficiency of buffer usage, the unused VC1 queues can be reassigned to VC0 and enable each of the ingress ports to handle more data traffic bursts. This virtual channel resource relocation feature enhances the performance of the PCIe Switch further.

The Switch provides the advanced feature of Access Control Service (ACS). This feature regulates which components are allowed to communicate with each other within the PCIe multiple-point fabric, and allows the system to have more control over packet routing in the Switch. As a result, peer-to-peer traffic can be facilitated more accurately and efficiently. When the system also implements Address Translation Service (ATS), the peer-to-peer requests with translated address can be routed directly by enabling the corresponding option in ACS to avoid possible performance bottleneck associated with re-direction, which introduces extra latency and may increase link and RC congestion.

The built-in Integrated Reference Clock Buffer of the PCI Express Switch supports three reference clock outputs. The clock buffer is from a single 100MHz clock input, and distributes the clock source to three outputs, which can be used by the downstream PCI Express end devices. The clock buffer feature can be enabled and disabled by strapping pin setting.

3 PIN DESCRIPTION

3.1 PCI EXPRESS INTERFACE SIGNALS

NAME	PIN	TYPE	DESCRIPTION
REFCLKP REFCLKN	A21 A19	I	**Reference Clock Input Pair:** Connect to 100MHz differential clock when integrated reference clock buffer is disabled (CLKBUF_PD=1), or connect to one of the Integrated Reference Clock Output Pairs (REFCLKO_P and REFCLKO_N) of this Switch when integrated reference clock buffer is enabled (CLKBUF_PD=0). The input clock signals must be delivered to the clock buffer cell through an AC-coupled interface so that only the AC information of the clock is received, converted, and buffered. It is recommended that a 0.1uF be used in the AC-coupling.
PERP [3:0]	B10, B26, B32, B4	I	**PCI Express Data Serial Input Pairs:** Differential data receive signals in four ports. Port 0 (Upstream Port) Lane 0 is PERP[0] and PERN[0] Port 0 (Upstream Port) Lane 1 is PERP[3] and PERN[3] Port 1 (Downstream Port) is PERP[1] and PERN[1] Port 2 (Downstream Port) is PERP[2] and PERN[2]
PERN [3:0]	A9, A27, B30, B6,	I	
PETP [3:0]	A11, A25, A31, A5	O	**PCI Express Data Serial Output Pairs:** Differential data transmit signals in four ports. Port 0 (Upstream Port) Lane 0 is PERP[0] and PERN[0] Port 0 (Upstream Port) Lane 1 is PERP[3] and PERN[3] Port 1 (Downstream Port) is PERP[1] and PERN[1] Port 2 (Downstream Port) is PERP[2] and PERN[2]
PETN [3:0]	A13, A23, A29, A7	O	
PERST_L	N1	I	**System Reset (Active LOW):** When PERST_L is asserted, the internal states of whole chip except sticky logics are initialized. Please refer to Table 11-2 for PERST_L spec.
DWNRST_L[2:1]	K2, G1	O	**Downstream Device Reset (Active LOW):** DWNRST_L provides a reset signal to the devices connected to the downstream ports of the switch. The signal is active when either PERST_L is asserted or the device is just plugged into the switch. DWNRST_L [x] corresponds to Portx, where x= 1,2,3.
REXT	B14	I	**External Reference Resistor:** Connect an external resistor (1.43K Ohm +/- 1%) to REXT_GND to provide a reference to both the bias currents and impedance calibration circuitry.
REXT_GND	B16	I	**External Reference Resistor Ground:** Connect to an external resistor to REXT.
REFCLKI_P, REFCLKI_N	AG35 AF34	I	**Integrated Reference Clock Input Pair:** Connect to external 100MHz differential clock for the integrated reference clock buffer.
REFCLKO_P[2:0]	AA35, U35, N35	O	**Integrated Reference Clock Output Pairs:** 100MHz external differential HCSL clock outputs for the integrated reference clock buffer.
REFCLKO_N[2:0]	AC35, W35, R35	O	
IREF	M34	I	**Differential Reference Clock Output Current Resistor:** External resistor (475 Ohm +/- 1%) connection to set the differential reference clock output current.
CLKBUF_PD	AP28	I	**Reference Clock Output Pairs Power Down:** When CLKBUF_PD is asserted high, the integrated reference clock buffer and Reference Clock Outputs are disabled. When it is asserted low, the integrated reference clock buffer and Reference Clock Outputs are enabled. This pin has internal pull-down. If no board trace is connected to this pin, the internal pull-down resistor of this pin is enough. However, if pin is connected to a board trace and not driven, it is recommended that an external 330-ohm pull-down resistor be used.

3.2 PORT CONFIGURATION SIGNALS

NAME	PIN	TYPE	DESCRIPTION
VC1_EN	W1	I	**Virtual Channel 1 Resource Sharing Enable:** The chip provides the capability to support virtual channel 1 (VC1), in addition to the standard virtual channel 0. When this pin is asserted high, Virtual Channel 1 is enabled, and virtual channel resource sharing is not available. When it is asserted low, the chip would allocate the additional VC1 resource to VC0, and VC1 capability is disabled. This pin has internal pull-down resistor. If no board trace is connected to this pin, the internal pull-down resistor of this pin is enough. However, if pin is connected to a board trace and not driven, it is recommended that an external 330-ohm pull-down resistor be used.
RXPOLINV_DIS	AD2	I	**Rx Polarity Inversion Disable:** When RXPOLINV_DIS is asserted high, it indicates to disable Rx Polarity Inversion detection function. Otherwise, it indicates to enable Rx Polarity Inversion detection function. This pin has internal pull-down resistor. If no board trace is connected to this pin, the internal pull-down resistor of this pin is enough. However, if pin is connected to a board trace and not driven, it is recommended that an external 330-ohm pull-down resistor be used.
PL_512B	AP22	I	**Max. Payload Size 512B:** When PL_512B is asserted high, it indicates the max. payload size capability is 512B. Otherwise, it indicates the max. Payload size is 256B. This pin has internal pull-down resistor. If no board trace is connected to this pin, the internal pull-down resistor of this pin is enough. However, if pin is connected to a board trace and not driven, it is recommended that an external 330-ohm pull-down resistor be used.
PRSNT[2:1]	AA1, Y2	I	**Present:** When PRSNT is asserted low, it indicates that the device is present in the slot of downstream port. Otherwise, it indicates the absence of the device. PRSNT[x] is correspondent to Port x, where x=1,2. These pins have internal pull-down resistors.
SLOTCLK	AP6	I	**Slot Clock Configuration:** It determines if the downstream component uses the same physical reference clock that the platform provides on the connector. When SLOTCLK is high, the platform reference clock is employed. By default, all downstream ports use the same physical reference clock provided by platform. This pin has internal pull-down resistor. If no board trace is connected to this pin, the internal pull-down resistor of this pin is enough. However, if pin is connected to a board trace and not driven, it is recommended that an external 330-ohm pull-down resistor be used.
SLOT_IMP[2:1]	AR15, AP14	I	**Slot Implemented:** These signals are asserted to indicate that the downstream ports are connected to slots. SLOT_IMP[x] corresponds to Port x, where x= 1, 2. When SLOT_IMP[x] is asserted, the Port x is connected to slot. Otherwise, it is chip-to-chip connection directly. These pins have internal pull-down resistors. If no board trace is connected to these pins, the internal pull-down resistors of these pins are enough. However, if pins are connected to a board trace and not driven, it is recommended that external 330-ohm pull-down resistors be used.
PORTSTATUS[2:0]	AK34, AN35, AM34	O	**Port Status:** These signals indicate the status of each port. Please connect to pin header for debug used. PORTSTATUS[x] is correspondent to Port x, where x=0, 1, 2.

3.3 MISCELLANEOUS SIGNALS

NAME	PIN	TYPE	DESCRIPTION
EECLK	AL35	O	**EEPROM Clock:** Clock signal to the EEPROM interface.
EEPD	AH34	I/O	**EEPROM Data:** Bi-directional serial data interface to and from the EEPROM. The pin is set to '1' by default.
SMBCLK	AG1	I	**SMBus Clock:** System management Bus Clock. This pin requires an external 5.1K-ohm pull-up resistor.

NAME	PIN	TYPE	DESCRIPTION
SMBDATA	AF2	I/O	**SMBus Data:** Bi-Directional System Management Bus Data. This pin requires an external 5.1K-ohm pull-up resistor.
SCAN_EN	AJ35	I/O	**Full-Scan Enable Control:** For normal operation, SCAN_EN is an output with a value of "0". SCAN_EN becomes an input during manufacturing testing.
GPIO[7:0]	AR13, AP12, AR11, AR9, AP10, AR7, AR5, AP8	I/O	**General Purpose Input and Output:** These eight general-purpose pins are programmed as either input-only or bi-directional pins by writing the GPIO output enable control register. When SMBus is implemented, GPIO[7:5] act as the SMBus address pins, which set Bit 2 to 0 of the SMBus address. **Debug Mode Selection:** In debug mode, GPIO[4:0] are used for Debug Mode Selection.
PWR_SAV	AH2	I	**Power Saving Mode:** PWR_SAV is a strapping pin. When this pin is pulled high when system is reset, the Power Saving Mode is enabled. When this pin is pulled low when system is reset, the Power Saving Mode is disabled. When this pin is pulled low, it should be tied to ground through a 330-ohm pull-down resistor. When this pin is pulled high, a 5.1K-ohm pull-up resistor should be used.
TEST3 TEST5 TEST6	V2 AE1 AP20	I	**Test3/5/6:** These pins are for internal test purpose. Test3, Test5 and Test6 should be tied to ground through a 330-ohm pull-down resistor.
TEST4	AC1	I	**Test4:** The pin is for internal test purpose. It should be tied to ground through a 330-ohm pull-down resistor for normal operation. **Port Status Output Enable: I**n debug mode, it is used to enable Port Status output.
TEST1	L1	I	**Test1:** The pin is for internal test purpose. It should be tied to 3.3V through a 5.1K-ohm pull-up resistor for normal operation. **Debug Mode Enable**: In debug mode, it need be tired to low through a 330-ohm pull-down resistor.
TEST2	U1	I	**Test2:** The pin is for internal test purpose. Test2 should be tied to 3.3V through a 5.1K-ohm pull-up resistor.
NC	A1, A3, A15, A33, A35, B1, B2, B8, B12, B20, B24, B28, B34, B35, C1, C35, D2, D34, E1, E35, H2, J1, AB2, AD34, AE35, AJ1, AK2, AL1, AM2, AN1, AP1, AP2, AP4, AP16, AP24, AP26, AP32, AP34, AP35, AR1, AR3, AR17, AR21, AR23, AR25, AR27, AR29, AR33, AR35		**Not Connected:** These pins can be just left open.

3.4 JTAG BOUNDARY SCAN SIGNALS

Name	Pin	Type	Description
TCK	J35	I	**Test Clock:** Used to clock state information and data into and out of the chip during boundary scan. When JTAG boundary scan function is not implemented, this pin should be left open (NC).
TMS	K34	I	**Test Mode Select:** Used to control the state of the Test Access Port controller. When JTAG boundary scan function is not implemented, this pin should be pulled low through a 330-Ohm pull-down resistor.
TDO	L35	O	**Test Data Output:** When SCAN_EN is high, it is used (in conjunction with TCK) to shift data out of the Test Access Port (TAP) in a serial bit stream. When JTAG boundary scan function is not implemented, this pin should be left open (NC).
TDI	G35	I	**Test Data Input:** When SCAN_EN is high, it is used (in conjunction with TCK) to shift data and instructions into the TAP in a serial bit stream. When JTAG boundary scan function is not implemented, this pin should be left open (NC).
TRST_L	H34	I	**Test Reset (Active LOW):** Active LOW signal to reset the TAP controller into an initialized state. When JTAG boundary scan function is not implemented, this pin should be pulled low through a 330-Ohm pull-down resistor.

3.5 POWER PINS

NAME	PIN	TYPE	DESCRIPTION
VDDC	B, L, R	P	**VDDC Supply (1.0V):** Used as digital core power pins.
VDDR	F2, F34, M2 AP18, AP30, AR31, AR19	P	**VDDR Supply (3.3V):** Used as digital I/O power pins.
CVDDR	P34, T34, Y34	P	**VDDR Supply (3.3V):** Used as reference clock power pins.
VDDCAUX	P2, R1	P	**VDDCAUX Supply (1.0V):** Used as auxiliary core power pins.
VAUX	T2	P	**VAUX Supply (3.3V):** Used as auxiliary I/O power pins.
AVDD	T	P	**AVDD Supply (1.0V):** Used as PCI Express analog power pins.
AVDDH	A17	P	**AVDDH Supply (3.3V):** Used as PCI Express analog high voltage power pins.
CGND	V34, AB34	P	**Ground:** Used as reference clock ground pins.
AGND	B18, B22	P	**Ground:** Used as analog ground pins.
VSS	GND	P	**VSS Ground:** Used as ground pins.

4 PIN ASSIGNMENTS

4.1 PIN LIST of 136-PIN aQFN

PIN	NAME	PIN	NAME	PIN	NAME	PIN	NAME
A1	NC	B35	NC	Y34	CVDDR	AP18	VDDR
A3	NC	C1	NC	AA1	PRSNT[2]	AP20	TEST6
A5	PETP[0]	C35	NC	AA35	REFCLKO_P[2]	AP22	PL_512B
A7	PETN[0]	D2	NC	AB2	NC	AP24	NC
A9	PERN[3]	D34	NC	AB34	CGND	AP26	NC
A11	PETP[3]	E1	NC	AC1	TEST4	AP28	CLKBUF_PD
A13	PETN[3]	E35	NC	AC35	REFCLKO_N[2]	AP30	VDDR
A15	NC	F2	VDDR	AD2	RXPOLINV_DIS	AP32	NC
A17	AVDDH	F34	VDDR	AD34	NC	AP34	NC
A19	REFCLKN	G1	DWNRST_L[1]	AE1	TEST5	AP35	NC
A21	REFCLKP	G35	TDI	AE35	NC	AR1	NC
A23	PETN[2]	H2	NC	AF2	SMBDATA	AR3	NC
A25	PETP[2]	H34	TRST_L	AF34	REFCLKI_N	AR5	GPIO[1]
A27	PERN[2]	J1	NC	AG1	SMBCLK	AR7	GPIO[2]
A29	PETN[1]	J35	TCK	AG35	REFCLKI_P	AR9	GPIO[4]
A31	PETP[1]	K2	DWNRST_L[2]	AH2	PWR_SAV	AR11	GPIO[5]
A33	NC	K34	TMS	AH34	EEPD	AR13	GPIO[7]
A35	NC	L1	TEST1	AJ1	NC	AR15	SLOT_IMP[2]
B1	NC	L35	TDO	AJ35	SCAN_EN	AR17	NC
B2	NC	M2	VDDR	AK2	NC	AR19	VDDR
B4	PERP[0]	M34	IREF	AK34	PORTSTATUS[2]	AR21	NC
B6	PERN[0]	N1	PERST_L	AL1	NC	AR23	NC
B8	NC	N35	REFCLKO_P[0]	AL35	EECLK	AR25	NC
B10	PERP[3]	P2	VDDCAUX	AM2	NC	AR27	NC
B12	NC	P34	CVDDR	AM34	PORTSTATUS[0]	AR29	NC
B14	REXT	R1	VDDCAUX	AN1	NC	AR31	VDDR
B16	REXT_GND	R35	REFCLKO_N[0]	AN35	PORTSTATUS[1]	AR33	NC
B18	AGND	T2	VAUX	AP1	NC	AR35	NC
B20	NC	T34	CVDDR	AP2	NC	T	AVDD
B22	AGND	U1	TEST2	AP4	NC	B	VDDC
B24	NC	U35	REFCLKO_P[1]	AP6	SLOTCLK	R	VDDC
B26	PERP[2]	V2	TEST3	AP8	GPIO[0]	L	VDDC
B28	NC	V34	CGND	AP10	GPIO[3]	GND	VSS
B30	PERN[1]	W1	VC1_EN	AP12	GPIO[6]		
B32	PERP[1]	W35	REFCLKO_N[1]	AP14	SLOT_IMP[1]		
B34	NC	Y2	PRSNT[1]	AP16	NC		

5 FUNCTIONAL DESCRIPTION

Multiple virtual PCI-to-PCI Bridges (VPPB), connected by a virtual PCI bus, reside in the Switch. Each VPPB contains the complete PCIe architecture layers that consist of the physical, data link, and transaction layer. The packets entering the Switch via one of VPPBs are first converted from serial bit-stream into parallel bus signals in physical layer, stripped off the link-related header by data link layer, and then relayed up to the transaction layer to extract out the transaction header. According to the address or ID embedded in the transaction header, the entire transaction packets are forwarded to the destination VPPB for formatting as a serial-type PCIe packet through the transmit circuits in the data link layer and physical layer. The following sections describe these function elements for processing PCIe packets within the Switch.

5.1 PHYSICAL LAYER CIRCUIT

The physical layer circuit design is based on the PHY Interface for PCI Express Architecture (PIPE). It contains Physical Media Attachment (PMA) and Physical Coding Sub-layer (PCS) blocks. PMA includes Serializer/ Deserializer (SERDES), PLL[1], Clock Recovery module, receiver detection circuits, beacon transmitter, electrical idle detector, and input/output buffers. PCS consists of framer, 8B/10B encoder/decoder, receiver elastic buffer, and PIPE PHY control/status circuitries. To provide the flexibility for port configuration, each lane has its own control and status signals for MAC to access individually. In addition, a pair of PRBS generator and checker is included for PHY built-in self test. The main functions of physical layer circuits include the conversion between serial-link and parallel bus, provision of clock source for the Switch, resolving clock difference in receiver end, and detection of physical layer errors.

In order to meet the needs of different application, the drive amplitude, de-emphasis and equalization of each transmitting channels can be adjusted using EEPROM individually. De-emphasis of -3.5 db is implemented by the transmitters when full swing signaling is used, while an offset can be individually applied to each channel.

5.1.1 RECEIVER DETECTION

The physical layer circuits implement receiver detection, which detects the presence of an attached 50 ohm to ground termination as per PCI Express Specification. The detect circuits determine if the voltage levels of the receiver have crossed the internal threshold after a configurable time determined by the Receiver Detection Threshold field in the PHY Parameter 2 Register (offset 7Ch, bit[6:4]) as listed in Table 5-1, which can be configured by EEPROM or SMBUS settings.

Table 5-1 Receiver Detection Threshold Settings

Receiver Detection Threshold	Threshold
000	1.0 us
001	2.0 us
010	4.0 us (Recommended)
011	5.0 us
100	10 us
101	20 us
110	40 us
111	50 us

[1] Multiple lanes could share the PLL.

5.1.2 RECEIVER SIGNAL DETECTION

Receiver signal idling is detected with levels above a programmable threshold specified by Receiver Signal Detect field in the PHY Parameter 2 Register (Offset 7Ch, bit[21:20]) as listed in Table 5-2, which can be configured on a per-port basis via EEPROM or SMBUS settings.

Table 5-2 Receiver Signal Detect Threshold

Receiver Signal Detect	Min (mV ppd)	Max (mV ppd)
00	50	80
01 (Recommended)	65	175
10	75	200
11	120	240

5.1.3 RECEIVER EQUALIZATION

The receiver implements programmable equalizer via the Receiver Equalization field in the PHY Parameter 2 Register (Offset 7Ch, bit[25:22]) as listed in Table 5-3, which can be configured on a per-port basis via EEPROM or SMBUS settings.

Table 5-3 Receiver Equalization Settings

Receiver Equalization	Equalization
0000	Off
0010	Low
0110 (Recommended)	Medium
1110	High

5.1.4 TRANSMITTER SWING

The PCI Express transmitters support implementations of both full voltage swing and half (low) voltage swing. In full swing signaling mode, the transmitters implement de-emphasis, while in half swing mode, the transmitters do not. The Transmitter Swing field in the PHY Parameter 2 Register (offset 7Ch, Bit[30]) is used for the selection of full swing signaling or half swing signaling, which can be configured on a per-port basis via EEPROM or SMBUS settings.

Table 5-4 Transmitter Swing Settings

Transmitter Swing	Mode	De-emphasis
0	Full Voltage Swing	Implemented
1	Half Voltage Swing	Not implemented

5.1.5 DRIVE AMPLITUDE AND DE-EMPHASIS SETTINGS

Depending on the operation condition (voltage swing and de-emphasis condition), one of the Drive Amplitude Base Level fields in the Switch Operation Mode Register (offset 74h) and one of the Drive De-Emphasis Base Level fields in the PHY Parameter 1 Register (offset 7Ah) are active for configuration of the amplitude and de-emphasis.

The final drive amplitude and drive de-emphasis are the summation of the base level value and the offset value. The offset value for drive amplitude is 25 mV pd, and 6.25 mV pd for drive de-emphasis.

The driver output waveform is the synthesis of amplitude and de-emphasis as shown in Figure 5-1. The driver amplitude without de-emphasis is specified as a peak differential voltage level (mVpd), and the driver de-emphasis modifies the driver amplitude.

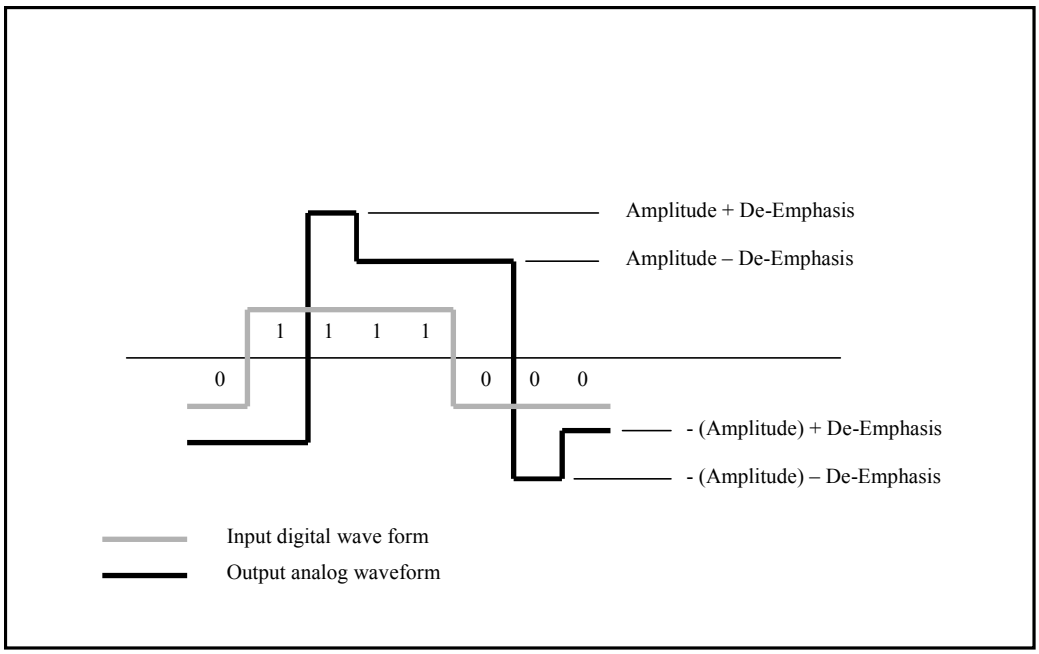

Figure 5-1 Driver Output Waveform

5.1.6 DRIVE AMPLITUDE

Only one of the Drive Amplitude Level field in the Switch Operation Mode Register (offset 74h, bit[20:16], bit[25:21] and bit[30:26]) listed in Table 5-5 is active depending on the de-emphasis and swing condition. The settings and the corresponding values of the amplitude level are listed in Table 5-6, which can be configured by EEPROM or SMBUS settings.

Table 5-5 Drive Amplitude Base Level Registers

Active Register	De-Emphasis Condition	Swing Condition
C_DRV_LVL_3P5_NOM	-3.5 db	Full
C_DRV_LVL_6P0_NOM	-6.0 db	Full
C_DRV_LVL_HALF_NOM	N/A	Half

Table 5-6 Drive Amplitude Base Level Settings

Setting	Amplitude (mV pd)	Setting	Amplitude (mV pd)	Setting	Amplitude (mV pd)
00000	0	00111	175	01110	350
00001	25	01000	200	01111	375
00010	50	01001	225	10000	400
00011	75	01010	250	10001	425
00100	100	01011	275	10010	450
00101	125	01100	300	10011	475
00110	150	01101	325	Others	Reserved

Note:
1. Nominal levels. Actual levels will vary with temperature, voltage and board effects.

2. The maximum nominal amplitude of the output driver is 475 mVpd. Combined values of driver amplitude and de-emphasis greater than 475 mVpd should be avoided.
3. At higher amplitudes, actual swings will be less than the theoretical value due to process variations and environment factors, such as voltage overhead compression, package losses, board losses, and other effects.

5.1.7 DRIVE DE-EMPHASIS

The Drive De-Emphasis Level field in the PHY Parameter 1 Register (Offset 78h, bit[20:16], bit[25:21] and bit[30:26]) listed in Table 5-7 controls the de-emphasis base level. The settings and the corresponding values of the de-emphasis level are listed in Table 5-8, which can be configured globally via EEPROM or SMBUS settings.

Table 5-7 Drive De-Emphasis Base Level Register

Register	De-Emphasis Condition
C_EMP_POST_GEN1_3P5_NOM	-3.5 db
C_EMP_POST_GEN2_3P5_NOM	-3.5 db
C_EMP_POST_GEN2_6P0_NOM	-6.0 db

Table 5-8 Drive De-Emphasis Base Level Settings

Setting	De-Emphasis (mV pd)	Setting	De-Emphasis (mV pd)	Setting	De-Emphasis (mV pd)
00000	0.0	01011	69.0	10110	137.5
00001	6.0	01100	75.0	10111	144.0
00010	12.5	01101	81.0	11000	150.0
00011	19.0	01110	87.0	11001	156.0
00100	25.0	01111	94.0	11010	162.5
00101	31.0	10000	100.0	11011	169.0
00110	37.5	10001	106.0	11100	175.0
00111	44.0	10010	112.5	11101	181.0
01000	50.0	10011	119.0	11110	187.5
01001	56.0	10100	125.0	11111	194.0
01010	62.5	10101	131.0	-	-

Note:
1. Nominal levels. Actual levels will vary with temperature, voltage and board effects.
2. The maximum nominal amplitude of the output driver is 475 mVpd. Combined values of driver amplitude and de-emphasis greater than 475 mVpd should be avoided.
3. At higher amplitudes, actual swings will be less than the theoretical value due to process variations and environment factors, such as voltage overhead compression, package losses, board losses, and other effects.

5.1.8 TRANSMITTER ELECTRICAL IDLE LATENCY

After the last character of the PCI Express transmission, the output current is reduced, and a differential voltage of less than 20 mV with common mode of VTX-CM-DC is established within 20 UI. This delay time is programmable via Transmitter PHY Latency field in the PHY Parameter 2 Register (Offset 7Ch, bit[3:0]), which can be configured by EEPROM or SMBUS settings.

5.2 DATA LINK LAYER (DLL)

The Data Link Layer (DLL) provides a reliable data transmission between two PCI Express points. An ACK/NACK protocol is employed to guarantee the integrity of the packets delivered. Each Transaction Layer Packet (TLP) is protected by a 32-bit LCRC for error detection. The DLL receiver performs LCRC calculation to determine if the incoming packet is corrupted in the serial link. If an LCRC error is found, the DLL transmitter would issue a NACK data link layer packet (DLLP) to the opposite end to request a re-transmission, otherwise an ACK DLLP would be sent out to acknowledge on reception of a good TLP.

In the transmitter, a retry buffer is implemented to store the transmitted TLPs whose corresponding ACK/NACK DLLP have not been received yet. When an ACK is received, the TLPs with sequence number equals to and smaller than that carried in the ACK would be flushed out from the buffer. If a NACK is received or no ACK/NACK is returned from the link partner after the replay timer expires, then a replay mechanism built in DLL transmitter is triggered to re-transmit the corresponding packet that receives NACK or time-out and any other TLP transmitted after that packet.

Meanwhile, the DLL is also responsible for the initialization, updating, and monitoring of the flow-control credit. All of the flow control information is carried by DLLP to the other end of the link. Unlike TLP, DLLP is guarded by 16-bit CRC to detect if data corruption occurs.

In addition, the Media Access Control (MAC) block, which is consisted of LTSSM, multiple lanes de-skew, scrambler/de-scrambler, clock correction from inserting skip order-set, and PIPE-related control/status circuits, is implemented to interface physical layer with data link layer.

5.3 TRANSACTION LAYER RECEIVE BLOCK (TLP DECAPSULATION)

The receiving end of the transaction layer performs header information retrieval and TC/VC mapping (see section 5.5), and it validates the correctness of the transaction type and format. If the TLP is found to contain an illegal header or the indicated packet length mismatches with the actual packet length, then a Malformed TLP is reported as an error associated with the receiving port. To ensure end-to-end data integrity, a 32-bit ECRC is checked against the TLP at the receiver if the digest bit is set in header.

5.4 ROUTING

The transaction layer implements three types of routing protocols: ID-based, address-based, and implicit routing. For configuration reads, configuration writes, transaction completion, and user-defined messages, the packets are routed by their destination ID constituted of bus number, device number, and function number. Address routing is employed to forward I/O or memory transactions to the destination port, which is located within the address range indicated by the address field carried in the packet header. The packet header indicates the packet types including memory read, memory write, IO read, IO write, Message Signaling Interrupt (MSI) and user-defined message. Implicit routing is mainly used to forward system message transactions such as virtual interrupt line, power management, and so on. The message type embedded in the packet header determines the routing mechanism.

If the incoming packet cannot be forwarded to any other port due to a miss to hit the defined address range or targeted ID, this is considered as Unsupported Request (UR) packet, which is similar to a master abort event in PCI protocol.

5.5 TC/VC MAPPING

The 3-bit TC field defined in the header identifies the traffic class of the incoming packets. To enable the differential service, a TC/VC mapping table at destination port that is pre-programmed by system software or EEPROM pre-load is utilized to cast the TC labeled packets into the desired virtual channel. Note that TC0 traffic is mapped into VC0 channel by default. After the TC/VC mapping, the receive block dispatches the incoming request, completion, or data into the appropriate VC0 and VC1 queues.

5.6 QUEUE

In PCI Express, it defines six different packet types to represent request, completion, and data. They are respectively Posted Request Header (PH), Posted Request Data payload (PD), Non-Posted Request Header (NPH), Non-Posted Data Payload (NPD), Completion Header (CPLH) and Completion Data payload (CPLD). Each packet with different type would be put into a separate queue in order to facilitate the following ordering processor. Since NPD usually contains one DW, it can be merged with the corresponding NPH into a common queue named NPHD. Except NPHD, each virtual channel (VC0 or VC1) has its own corresponding packet header and data queue. When only VC0 is needed in some applications, VC1 can be disabled and its resources assigned to VC0 by asserting VC1_EN (Virtual Channel 1 Enable) to low.

5.6.1 PH

PH queue provides TLP header spaces for posted memory writes and various message request headers. Each header space occupies sixteen bytes to accommodate 3 DW or 4 DW headers. There are two PH queues for VC0 and VC1 respectively.

5.6.2 PD

PD queue is used for storing posted request data. If the received TLP is of the posted request type and is determined to have payload coming with the header, the payload data would be put into PD queue. There are two PD queues for VC0 and VC1 respectively.

5.6.3 NPHD

NPHD queue provides TLP header spaces for non-posted request packets, which include memory read, IO read, IO write, configuration read, and configuration write. Each header space takes twenty bytes to accommodate a 3-DW header, s 4-DW header, s 3-WD header with 1-DW data, and a 4-DW header with 1-DW data. There is only one NPHD queue for VC0, since non-posted request cannot be mapped into VC1.

5.6.4 CPLH

CPLH queue provides TLP header space for completion packets. Each header space takes twelve bytes to accommodate a 3-DW header. Please note that there are no 4-DW completion headers. There are two CPLH queues for VC0 and VC1 respectively.

5.6.5 CPLD

CPLD queue is used for storing completion data. If the received TLP is of the completion type and is determined to have payload coming with the header, the payload data would be put into CPLD queue. There are two CPLD queues for VC0 and VC1 respectively.

5.7 TRANSACTION ORDERING

Within a VPPB, a set of ordering rules is defined to regulate the transactions on the PCI Express Switch including Memory, IO, Configuration and Messages, in order to avoid deadlocks and to support the Producer-Consumer

model. The ordering rules defined in table 5-4 apply within a single Traffic Class (TC). There is no ordering requirement among transactions within different TC labels. Since the transactions with the same TC label are not allowed to map into different virtual channels, it implies no ordering relationship between the traffic in VC0 and VC1.

Table 5-9 Summary of PCI Express Ordering Rules

Row Pass Column	Posted Request	Read Request	Non-posted Write Request	Read Completion	Non-posted Write Completion
Posted Request	Yes/No[1]	Yes[5]	Yes[5]	Yes[5]	Yes[5]
Read Request	No[2]	Yes	Yes	Yes	Yes
Non-posted Write Request	No[2]	Yes	Yes	Yes	Yes
Read Completion	Yes/No[3]	Yes	Yes	Yes	Yes
Non-Posted Write Completion	Yes[4]	Yes	Yes	Yes	Yes

1. When the Relaxed Ordering Attribute bit is cleared, the Posted Request transactions including memory write and message request must complete on the egress bus of VPPB in the order in which they are received on the ingress bus of VPPB. If the Relaxed Ordering Attribute bit is set, the Posted Request is permitted to pass over other Posted Requests occurring before it.

2. A Read Request transmitting in the same direction as a previously queued Posted Request transaction must push the posted write data ahead of it. The Posted Request transaction must complete on the egress bus before the Read Request can be attempted on the egress bus. The Read transaction can go to the same location as the Posted data. Therefore, if the Read transaction were to pass the Posted transaction, it would return stale data.

3. When the Relaxed Ordering Attribute bit is cleared, a Read completion must ''pull'' ahead of previously queued posted data transmitting in the same direction. In this case, the read data transmits in the same direction as the posted data, and the requestor of the read transaction is on the same side of the VPPB as the completer of the posted transaction. The posted transaction must deliver to the completer before the read data is returned to the requestor. If the Relaxed Ordering Attribute bit is set, then a read completion is permitted to pass a previously queued Memory Write or Message Request.

4. Non-Posted Write Completions are permitted to pass a previous Memory Write or Message Request transaction. Such transactions are actually transmitting in the opposite directions and hence have no ordering relationship.

5. Posted Request transactions must be given opportunities to pass Non-posted Read and Write Requests as well as Completions. Otherwise, deadlocks may occur when some older bridges, which do not support delayed transactions are mixed with PCIe Switch in the same system. A fairness algorithm is used to arbitrate between the Posted Write queue and the Non-posted transaction queue

5.8 PORT ARBITRATION

Among multiple ingress ports, the port arbitration built in the egress port determines which incoming packets to be forwarded to the output port. The arbitration algorithm contains hardware fixed Round Robin, 128-phase Weighted Round-Robin and programmable 128-phase time-based WRR. The port arbitration is held within the same VC channel. It means that each port has two port arbitration circuitries for VC0 and VC1 respectively. At the upstream ports, in addition to the inter-port packets, the intra-port packet such as configurations completion would also join the arbitration loop to get the service from Virtual Channel 0.

5.9 VC ARBITRATION

After port arbitration, VC arbitration is executed among different VC channels within the same source. Three arbitration algorithms are provided to choose the appropriate VC: Strict Priority, Round Robin or Weighted Round Robin.

5.10 FLOW CONTROL

PCI Express employs Credit-Based Flow Control mechanism to make buffer utilization more efficient. The transaction layer transmitter ensures that it does not transmit a TLP to an opposite receiver unless the receiver has enough buffer space to accept the TLP. The transaction layer receiver has the responsibility to advertise the free buffer space to an opposite transmitter to avoid packet stale. In this Switch, each port has its own separate queues for different traffic types and the credits are sent to data link layer on the fly. The data link layer compares the current available credits with the monitored ones and reports the updated credit to the counterpart. If no new credit is acquired, the credit reported is scheduled for every 30 us to prevent the link from entering retrain. On the other hand, the receiver at each egress port gets the usable credits from the opposite end in a link. The output port broadcasts them to all the other ingress ports to get packet transmission.

5.11 TRANSATION LAYER TRANSMIT BLOCK (TLP ENCAPSULATION)

The transmit portion of transaction layer performs the following functions. They construct the all types of forwarded TLP generated from VC arbiter, respond with the completion packets when the local resource (i.e. configuration register) is accessed, and regenerate the message that terminates at receiver to RC if acting as an upstream port.

5.12 ACCESS CONTROLS SERVICE

Traditionally, the packet routing between the peer-to-peer downstream ports is determined by either the address or ID field embedded in the packet header. ACS provides a mechanism for customer to selectively control access between PCI Express Endpoints attached to the downstream ports of packet switch. If ACS is enabled in the ingress port, the peer-to-peer packet forwarding will follow the rule sets of ACS rather than the destination ID or address. ACS is implemented as a set of capabilities and control registers in the associated hardware component. It brings the following benefits such as preventing the silent data corruption presented in Requests from being incorrectly routed to a peer Endpoint, validating every Request transaction between two downstream components and enabling direct routing of peer-to-peer Memory Requests whose addresses have been Translated when ATS system is being used.

6 EEPROM INTERFACE AND SYSTEM MANAGEMENT BUS

The EEPROM interface consists of two pins: EECLK (EEPROM clock output) and EEPD (EEPROM bi-directional serial data). The Switch may control an ISSI IS24C04 or compatible parts using into 512x8 bits. The EEPROM is used to initialize a number of registers before enumeration. This is accomplished after PRST# is de-asserted, at which time the data from the EEPROM is loaded. The EEPROM interface is organized into a 16-bit base, and the Switch supplies a 7-bit EEPROM word address. The Switch does not control the EEPROM address input. It can only access the EEPROM with address input set to 0.

The System Management Bus interface consists of two pins: SMBCLK (System Management Bus Clock input) and SMBDATA (System Management Bus Data input/ output).

6.1 EEPROM INTERFACE

6.1.1 AUTO MODE EERPOM ACCESS

The Switch may access the EEPROM in a WORD format by utilizing the auto mode through a hardware sequencer. The EEPROM start-control, address, and read/write commands can be accessed through the configuration register. Before each access, the software should check the Autoload Status bit before issuing the next start.

6.1.2 EEPROM MODE AT RESET

During a reset, the Switch will automatically load the information/data from the EEPROM if the automatic load condition is met. The first offset in the EEPROM contains a signature. If the signature is recognized, the autoload initiates right after the reset.

During the autoload, the Bridge will read sequential words from the EEPROM and write to the appropriate registers. Before the Bridge registers can be accessed through the host, the autoload condition should be verified by reading bit [3] offset DCh (EEPROM Autoload Status). The host access is allowed only after the status of this bit is set to '0' which indicates that the autoload initialization sequence is complete.

6.1.3 EEPROM SPACE ADDRESS MAP

15 – 8	7 – 0	BYTE OFFSET
EEPROM Signature (1516h)		00h
Vendor ID		02h
Device ID		04h
Extended VC Count / Link Capability / Switch Mode Operation / Interrupt pin for Port 1 ~ 2		06h
Subsystem Vender ID		08h
Subsystem ID		0Ah
Max_Payload_Size Support / ASPM Support / Role_Base Error Reporting		0Ch
Global PHY TX Margin Parameter for Port 0~2		0Eh
Global PHY Parameter 0 for Port 0~2		10h
Global XPIP_CSR6[0] / Global PHY Parameter 1 for Port 0~2		12h
Global XPIP_CSR6[4:1] / Global PHY Parameter 2/3 for Port 0~2		14h
Global XPIP_CSR4[15:0] for Port 0~2		16h
Global XPIP_CSR4[31:16] for Port 0~2		18h
Global XPIP_CSR5[15:0] for Port 0~2		1Ah
Buffer_ctrl[4:0] / Globe XPIP_CSR6[7:5] for Port 0~2	Global XPIP_CSR5[23:16] for Port 0~2	1Ch
MAC_CTR / Global PHY Parameter 3 for Port 0~2		1Eh