

IBM Packet Routing Switch Serial Interface Converter

**Databook**

**Advance**



<span id="page-1-0"></span>Copyright International Business Machines Corporation 1999

All Rights Reserved Printed in the United States of America July 2000

The following are trademarks of International Business Machines Corporation in the United States, or other countries, or both. IBM IBM Logo PowerPC

Other company, product and service names may be trademarks or service marks of others.

All information contained in this document is subject to change without notice. The products described in this document are NOT intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change IBM product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of IBM or third parties. All information contained in this document was obtained in specific environments, and is presented as an illustration. The results obtained in other operating environments may vary.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED ON AN "AS IS" BASIS. In no event will IBM be liable for damages arising directly or indirectly from any use of the information contained in this document.

While the information contained herein is believed to be accurate, such information is preliminary, and should not be relied upon for accuracy or completeness, and no representations or warranties of accuracy or completeness are made.

IBM Microelectronics Division 1580 Route 52, Bldg. 504 Hopewell Junction, NY 12533-6351 The IBM home page can be found at http://www.ibm.com

The IBM Microelectronics Division home page can be found at http://www.chips.ibm.com

prssi.01 July 10, 2000



# **Contents**







#### **Advance IBM Packet Routing Switch Serial Interface Converter**





IBM Packet Routing Switch Serial Interface Converter Advance





# IBM Packet Routing Switch Serial Interface Converter





# **[List of Figures](#page-11-0)**





### **Advance IBM Packet Routing Switch Serial Interface Converter**





# **[List of Tables](#page-18-2)**





### **Advance IBM Packet Routing Switch Serial Interface Converter**





# <span id="page-10-0"></span>**1. General Information**

# <span id="page-10-1"></span>**1.1 Features**

- Companion to the IBM 28.4 G Packet Routing Switch
- Support for internal (8 ports) and external (16 ports) switch speed expansion mode
- Proprietary 440 Mbps, 8 HSTL pair, data aligned serial link (DASL) switch interface
- 3.52 Gbps aggregate throughput per speed expanded port
- 32-bit Ingress/Egress Protocol Engine Interface Bus (UTOPIA-3 like bus)
- Implements switch plane redundancy system architecture with two independent paths
- 3 packet ingress/6 packet egress shared buffers
- Implements In Band Flow Control (IBFC) via packet header information
- Up to four priority levels packet handling
- Programmable Packet Length (64 to 80 bytes)
- Link liveness packet insertion (Yellow packet)
- 8-bit Processor Interface (with bursting option)
- Internal ABIST
- IEEE 1149.1 standard boundary scan to facilitate circuit board testing
- 1.5 V DASL differential I/Os for DASLs
- 2.5 V Supply Voltage (3.3 V-tolerant I/Os) and 3.3 V LVTTL for the other signal I/Os
- 360 CBGA package
- IBM CMOS 6 SF SA-12E technology

# <span id="page-10-2"></span>**1.2 Description**

The IBM Packet Routing Switch Serial Interface Converter (the converter) is a companion chip to the IBM 28.4 G Packet Routing Switch (the switch), which connects the switch's serial link to a protocol engine (PE) on a 32-bit interface bus. The converter attaches to a switch port operating in speed expansion mode up to 4 Gbps, wired on eight DASL pairs running at up to 500 Mbps per pair.

No synchronization is required between input ports. However, packets on a given port are always received or transmitted at a fixed interval equal to the packet length. The converter ingress/egress packet length is programmable from 64 to 80 bytes in increments of four bytes. Input/Output packets to/ from the switch are mapped into 4x logical units (LU) of 16 to 20 bytes depending upon the packet length.

The converter is composed of two fully independent data paths (X and Y) to provide a resilient switch fabric.These paths are clocked, reset and controlled independently to support independent activation/ deactivation of each switch plane.

Ingress traffic (packets received from a PE interface bus) is routed to both X and Y path switch planes, thereby duplicating packets on both planes. Egress traffic (packets received from switch plane X or Y) is routed to a bus.

The converter normally uses In Band Flow Control (IBFC). The continuous flow of packets transmits necessary information to regulate packet traffic between the PE and the switch planes (in band signalling). Out of Band Flow Control (OBFC) may also be used with IBFC at the interface level, using the traditional operation UTOPIA-3 handshake signals.

The converter extracts output queue grant information from the 16 output queues of the switch, and Memory Grant information corresponding to the four priority watermark levels of the total shared memory. It then provides this congestion control information to the attached PE through an IBFC mechanism. The information is transmitted to the PE in the packet header bytes of either data or idle packets.



<span id="page-11-0"></span>



### <span id="page-11-1"></span>**Figure 2: IBM Packet Routing Switch Serial Interface General Data Flow**







# <span id="page-12-0"></span>**2. Converter Ingress/Egress Data Flow**

The ingress block is the receive path between the protocol engine device (PE) and the rest of the IBM Packet Routing Switch Serial Interface Converter logic. The egress block is the transmit path between the converter transmit logic and the PE. This section provides an overview of functions implemented in the converter ingress and egress data flows. Some functions are duplicated to support two switch planes.

# <span id="page-12-1"></span>**2.1 Ingress Data Flow**

## <span id="page-12-2"></span>**2.1.1 Protocol Engine (PE) Ingress Interface**

The converter connects to the PE via a 32-bit bus. Ingress data packets are simultaneously routed to PATH X and to PATH Y. Idle packets are inserted in the word stream when there is no data to transfer and are used to maintain a synchronous packet operation in the ingress interface. Idle packets are identified by a bit in the packet qualifier byte. A Receive Start of Packet (RXSOP) synchronized with the data packet is used to delineate packets.

Combined with the RXPRTY signal issued from the PE, the ingress interface checks the parity coherency on each incoming RXDATA [31:0]. A specific bit in the configuration table registers can be set so each parity error issued from the parity checker is reported.

Under the control of the configuration table, RXPRTY\_error assertion indicates that the cell that is currently pushed into the ingress reshuffling buffer will be optionally ignored and will not be sent to the ingress FIFO. Cells which are pushed into the ingress FIFOs (X/Y RX FIFOs) are always good and can be treated by the data flow.

### <span id="page-12-3"></span>**2.1.2 Ingress Packet Reshuffling**

The ingress logical unit framing block maps incoming data packets into the IBM 28.4 G Packet Routing Switch (switch) logical units (LUs) by moving the five bytes (the packet qualifier and the bit map fields) selected from the configuration register into the master LU, which becomes the switch header information field. The framing block also extracts the IBFC information, discards idle packets, and performs parity checking on the switch header.

### <span id="page-12-4"></span>**2.1.3 Ingress Receive FIFO**

The ingress receive FIFOs provide packet synchronization between the 50 - 125 MHz PE interface and the 110 - 125 MHz switch core interface.

### <span id="page-12-5"></span>**2.1.4 Ingress Data Aligned Serial Link Interface (IDI)**

The IDI sends packets continuously. Synchronization packets are sent during the DASL synchronization sequence. Data or idle packets are sent once data mode is active. On request, through the configuration table, the LU serializer is filled with a yellow packet and the incoming data packet is buffered while the yellow packet is sent. When there is no data packet to be transmitted to the switch core, the IDI inserts an idle packet, computes the inter idle CRC for each LU, and inserts it as the last byte of each LU.

#### <span id="page-13-0"></span>**2.1.5 Data Aligned Serial Link (DASL) Port Serializer**

The IDI feeds the DASL port serializer with packets compatible with the switch LU format. (The 16 - 20 byte LU width is set in the configuration register.) The DASL port serializer performs a multi-bit serialization for each LU. Each serial DASL interface line represents a nibble of the LU, so there are two DASL links per byte. The converter provides a total of eight serial links per port (one for each 4-bit nibble) representing a 32-bit wide word.

### <span id="page-13-1"></span>**2.1.6 Egress Data Flow**

#### <span id="page-13-2"></span>**2.1.7 Data Aligned Serial Link (DASL) Port Deserializer**

The DASL egress performs 32-bit deserialization on incoming data and builds LUs for the DASL egress interface. It continuously monitors signal quality on the incoming high speed serial link and performs continuous bit positioning adjustment on the incoming data to maintain synchronization.

#### <span id="page-13-3"></span>**2.1.8 Egress Data Aligned Serial Link Interface (EDI)**

The DASL's Receive Data Indicator line triggers the LU deserializer logic block which receives a continuous stream of packets. Packet length is programmable from 64 to 80 bytes and is mapped on 4-bytes word boundaries. Therefore, a new packet is received from the DASL every 16 to 20 cycles.

The LU deserializer extracts the switch packet header from the master LU to determine packet type (idle or data) and priority. Idle packets are discarded. Data packets are forwarded (LU format) to the egress buffer interface. The LU deserializer checks the LU CRC after each idle cell, the parity on the switch header, and the type of packet. When an error or yellow packet is detected, and the checker is enabled, the corresponding converter interrupt line is asserted.

### <span id="page-13-4"></span>**2.1.9 Egress FIFO**

The egress FIFO interface ignores any idle packets issued from the switch. Only data packets are pushed into the transmit FIFOs (TXFIFOs) for a further word formatting packet operation. The TXFIFOs provide packet synchronization between the 110 - 125 MHz converter and the 50 - 110 MHz protocol engine.

#### <span id="page-13-5"></span>**2.1.10 Egress Path Selection**

The egress path selection multiplexes packets coming from the X and Y paths to the protocol engine's egress interface bus.The data is selected from the path in service through two interface lines (X\_InService and Y\_InService) received from the switch core.

#### <span id="page-13-6"></span>**2.1.11 Egress Transmit Framing**

The packet formatter translates custom formatted packet LUs, reversing the operation performed in the ingress path. Header byte swapping moves the header bytes back to their original position. The converter simultaneously takes the latest available output queue grant flow control data from the EDI and moves it into the corresponding byte positions that were used for output port addressing in the ingress path. The most recent shared memory grant and the output queue grant (OQG) priority (related to the bit map field's OQG) are both stored in the packet qualifier byte that is sent to the PE so it can perform virtual output queuing scheduling for the ingress packets. Idle packets are also generated and sent to the PE in order to maintain continuous flow control information.





#### <span id="page-14-0"></span>**2.1.12 Protocol Engine (PE) Egress Interface**

The converter connects to the PE's 32-bit bus via the PE egress interface. Egress transmit framing controls data and idle packet transfer. The PE's egress interface monitors the behavior of TXFULL, controls TXENB accordingly, and generates the Start of Packet to the PE. A parity bit covering the 32-bit word can be generated according to a dedicated bit in the configuration table registers.



# <span id="page-15-0"></span>**3. Functional Description**

This chapter describes each functional block of the IBM Packet Routing Switch Serial Interface Converter (the converter). The description is by layer and covers both the ingress and the egress functions because, in general, they are completely symmetrical.

<span id="page-15-1"></span>





### <span id="page-16-2"></span>**Figure 4: Bit and Byte Notation**



The ingress block is the receive path between the protocol engine device (PE) and the rest of the converter logic. The egress block is the transmit path between the converter transmit logic and the PE.

The timing on the UTOPIA-3 interface is the same for out of band flow control (OBFC) and for in band flow control (IBFC) modes except when operating in IBFC mode, the UTOPIA-3 flow control (RXENB at ingress and TXFULL at egress) is not normally used, as the packet header carries all the required information. However if a PE performs an interface level flow control (OBFC), those interface leads can be used for that purpose. In IBFC, fill-up (idle) packets flow through the interface to maintain a continuous stream of flow control information.

# <span id="page-16-0"></span>**3.1 Data Interface Between IBM Packet Routing Switch Serial Interface Converter and Protocol Engine**

UTOPIA-3 bit/byte notation is used throughout the protocol engine (PE) interface description. After the byte formatting (which reshuffles the byte position to match the IBM 28.4G Packet Routing Switch (switch) DASL interface), the switch DASL bit/byte notation is used.

# <span id="page-16-1"></span>**3.2 Functional Overview**

The IBM Packet Routing Switch Serial Interface Converter (the converter) ingress and egress interfaces are consistent with a subset of the UTOPIA-3 specifications:

- Single PHY interface mode (connection between one converter and one protocol engine)
- Typical operating clock range is 100 111 MHz to match the IBM 28.4 G Packet Routing Switch clock rate (up to 111 MHz), but an interface clock rate as low as 50 MHz is supported
- Only 32-bit data paths
- Packet format is 64 80 bytes (programmed in the converter at configuration time)
- Only packet level handshake mode (use of signals RXPAV in Receive Path and TXPAV in Transmit Path)

- RXENB and TXENB signals can not be used to perform flow control at an octet level, as is allowed in UTOPIA-2 specifications. Each packet transfer initiated in either the ingress or egress direction will continue to flow until current packet transfer completion, eliminating the ability to insert wait states during the current packet transfer
- The assertion of the RXENB signal depends only on the ingress FIFO filling status and so may be asserted while the RXPAV signal is de-asserted
- Wait states insertion on the bus is only allowed between transmission of two different packets
- All input and output signals are registered. Therefore, a device (either the PE or the converter) responds in not less than two clock cycles after the initiating signal is sent across the interface
- All output signals are generated and all input signals are sampled on low-to-high clock transitions
- All signals are active high, unless the name has an overbar  $\overline{(xxx)}$ .

#### <span id="page-17-0"></span>**3.2.1 Ingress Interface**

#### <span id="page-17-1"></span>**3.2.1.1 Bus Protocol**

The ingress block is the receive path between the protocol engine device (PE) and the IBM Packet Routing Switch Serial Interface Converter (the converter). Data is sent by the PE to the converter according to the following protocol:

- The converter provides the receive clock PE\_RXCLK\_OUT
- The PE asserts the signal RXPAV when it is ready to send at least one complete packet on the bus
- The converter asserts the signal  $\overline{\text{RXENB}}$  when it is ready to receive at least one complete packet
- Receive packet transfer can start once the PE detects RXENB asserted and asserts RXPAV
- The assertion of the signal RXSOP during one clock cycle indicates start of a receive packet transfer
- RXDATA[31:0] is transferred on each low-to-high clock transition and the first data word of the packet is transferred simultaneously with the signal RXSOP
- The converter de-asserts the **RXENB** signal two clock cycles before the end of the current packet transfer to indicate that it can not accept an immediate transfer of the subsequent packet from the PE

This protocol applies if a user wishes to use OBFC mechanisms in addition to IBFC. Under IBFC there is no need for the use of RXENB /RXPAV protocol, as all the flow control is performed in band (through the packet header). Also under IBFC, if there is no data packet to be sent by the protocol engine, it will insert an idle packet that will be discarded by the IBM Packet Routing Switch Serial Interface Converter.



## <span id="page-18-2"></span>**Table 1: Ingress I/O Pin Description**



### <span id="page-18-0"></span>**3.2.1.2 Ingress Operation and Timing**

#### <span id="page-18-1"></span>**Figure 5: Ingress Timing for RXENB Deasserted by Converter for 1 Clock Cycle**  Ingress Operation for 64-Byte Packet Flow Control







#### <span id="page-19-0"></span>**Figure 6: Ingress Timing for RXENB Deasserted by Converter for 3 Clock Cycles**  Ingress Operation for 64-Byte Packet Flow Control

<span id="page-19-1"></span>**Figure 7: Ingress Timing for RXPAV Deasserted by Protocol Engine for 1 Clock Cycle**  Ingress Operation for 64-Byte Packet Flow Control





Advance **IBM Packet Routing Switch Serial Interface Converter** 



#### <span id="page-20-0"></span>**Figure 8: Ingress Timing for RXPAV Deasserted by Protocol Engine for 3 Clock Cycles**  Ingress Operation for 64-Byte Packet Flow Control

<span id="page-20-1"></span>**Figure 9: Ingress Packets in Back-to-Back from Protocol Engine**  Receive Operation for 64-Byte Back-to-Back Packets





#### <span id="page-21-0"></span>**3.2.2 Egress Interface**

#### <span id="page-21-1"></span>**3.2.2.1 Bus Protocol**

The egress block is the transmit path between the IBM Packet Routing Switch Serial Interface Converter (the converter) and the protocol engine (PE). The converter sends data according to the following protocol:

- The converter provides the transmit clock called PE\_TXCLK\_OUT
- The PE asserts the TXPAV signal when it is ready to receive at least one packet from the converter
- When the converter is ready to send at least one packet on the bus and the PE device has asserted the signal TXPAV, it starts the transfer by simultaneously asserting the signal TXSOP and TXENB
- TXDATA[31:0] is transferred on each low-to-high clock transition, the first data word of the packet being transferred simultaneously with the signal TXSOP
- The PE device de-asserts the TXPAV signal at least four cycles before the end of the current packet transfer to indicate that it cannot accept an immediate transfer of the subsequent packet

This protocol applies when a user wishes to use OBFC mechanisms in addition to IBFC if the attached PE has two sets of buffers. The small buffer allows the link layer to absorb the lack of synchronization between the UTOPIA-3 interface and the rest of the PE chip. The large buffer absorbs traffic burst. Under pure IBFC operation, because all the flow control is performed in band (through the packet header), there is no need for the use of TXENB / TXFULL (TXPAV) protocols. Also under IBFC, if there is no data packet to be sent by the converter, it will insert an idle packet that will be discarded by the PE.



# <span id="page-22-0"></span>**Table 2: Egress I/O Pin Description**





### <span id="page-23-0"></span>**3.2.2.2 Egress Operation and Timing**

# <span id="page-23-2"></span>**Figure 10: TXFULL Timing**



# <span id="page-23-3"></span>**Figure 11: Egress Timing for Back-to-Back Packets**

Transmit Operation for 64-Byte Back-t0-Back Packet Transmission



## <span id="page-23-1"></span>**3.2.3 TXFULL Timing**

When the PE asserts TXFULL to stop the flow of packets, its effect is extended for 28 PE clock cycles from the end of the current packet. After the first packet following the deassertion of TXFULL by the PE, the UDASL inserts two empty slots of idle (no data) before the traffic resumes. If the PE resasserts TXFULL upon the reception of the first packet during the TXFULL cycle, it will be ignored until the end of the complete cycle.



# <span id="page-24-0"></span>**3.3 Packet Reshuffling**

### <span id="page-24-1"></span>**3.3.1 Ingress Receive Logical Unit Framing**

The Ingress LU formatter is the interface between the PE interface logic and the ingress FIFO (RXFIFO) which translates various types of ingress packet formats into the IBM 28.4 G Packet Routing Switch (switch) LU format is defined in the IBM Packet Routing Switch Serial Interface Converter (the converter) configuration registers.

Data inputs are in a 4-byte format. The ingress LU formatter extracts TxPause flow control information from the incoming packet qualifier byte. The incoming packet is comprised of 16 to 20 32-bit words. Packet Reshuffling stores the packet in the switch format in four LUs which, when fully generated, are forwarded to the RXFIFO.

The Ingress LU Formatter:

- Discard idle packets
- Extract Send Grant flow control information from the incoming PQ byte
- Modify the packet qualifier byte into switch format and compute the associated parity
- Change the position of the bit map bytes in the packet header
- Know of any bus parity error through the PE interface RXPRTY\_error
- May not forward a packet to the RXFIFOs in case of any bus or header error (depending upon a configuration register setting)

### <span id="page-24-2"></span>**3.3.1.1 Header Bytes Reshuffling**

The header bytes are moved according to the content of the "byte positioning in LU formatter" configuration fields. These allow any byte contained in the four words to be moved to any other byte position in the master LU.

### <span id="page-24-3"></span>**3.3.1.2 Input/Output Packet Format**

The following tables highlight the position of the different information fields in an incoming packet. All bytes in the first four words can be repositioned in the switch header according to the formatter field in the configuration registers. The data source is based on a word/byte coordinate in a nibble. The first two bits correspond to the word selection; the last two bits correspond to the byte selection.



Example of Header Bytes in Sequence

When the bitmap field is defined for 32-bits but only 16 are used, only bits 16 through 31 of the BM field are required to identify the 16 switch core ports. The others are considered to be part of the payload. The parity sub-field of the Packet Qualifier field is calculated across the first three or five bytes (depending upon the configuration).

For example, with header bytes in sequence the content of the bit positioning registers (@C8 for ingress and @C4 for the egress) will be 0, 1, 2, 3, 1, as shown in the table below:



<span id="page-25-0"></span>1. This last index is not 4 because during the second move operation this byte was moved from word 1 position 0 to word 0 position 1

The full packet is stored in the ingress LU formatter. Reshuffling starts when the first five words are received. They are stored again in another set of five word buffers. The header information on which the switch acts to route the packet to the appropriate output port is moved to the appropriate byte location in the master LU during the cycles required to store the current packet's remaining words. Byte swapping is done according to the 16 bits stored in the configuration register @C8 (byte positioning in LU formatter).

#### Byte Byte 0 Byte 1 Byte 2 Byte 3 MSB LSB F J ¢  $\sqrt{2}$ ı  $\mathbf{I}$ Bit 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0 Word 0 Packet Qualifier Word 1 BM[31:24] BM1 BM[23:16] BM2 BM[15:8] BM3 BM[7:0] BM4 Word 2 Word 15

Example of Bit Map Field in Single Word

When 16 (or 8)-port mode is used, only bits 31 - 16 of the BM field are required to identify the 16-switch core ports. The others are considered to be part of the payload. The outgoing packet is made of four logical units (LU0, LU1, LU2, and LU3), each 16-bytes wide. The following tables show how the bytes of the incoming packet are rearranged inside the four logical units. Each line represents the content of each LU.

#### 8x8 and 16x16 Switch LU Output Format:





#### **Advance IBM Packet Routing Switch Serial Interface Converter**

#### 32x32 Switch LU Output Format:



When the switch is in 32x32 mode, all bytes of the Bitmap are used (BM1, BM2, BM3, and BM4) for both port addressing and to compute the header parity. When the switch is in 16x16 mode, only BM1 and BM2 are used while BM3 and BM4 are considered as payload and are ignored for the computation of the header parity. When the switch is in 08x08 mode, only BM1 is used while BM2 must be '00' so it does not impact the computation of the parity in packet qualifier.

#### <span id="page-26-0"></span>**3.3.1.3 Port Addressing**

The following table shows the switch output port addressing (a binary 1 in the bit map field indicates that the packet should be sent to the corresponding switch output port).



#### <span id="page-26-2"></span>**Table 3: Output Port Bitmap Fields**

## <span id="page-26-1"></span>**3.3.2 Nested TxPause Extraction**

In IBFC, The bits PQ (26,27 and 30,31) of the packet qualifier are monitored to extract the converter egress flow control information (TxPause). The TxPause information is used by the PE to flow control the switch's output port. The converter interprets those bits and translates them into Sent Grant to the switch. The use of TxPause is dependent upon the setting of register @C8 bit 30 of the configuration table register. When Bit 3 (Send Grant per Priority) of configuration register table @C4 is not enabled, any bit set to '0' in the TxPause field is sufficient to stop transmission to the attached protocol engine.

#### <span id="page-27-0"></span>**3.3.3 Ingress Packets for IBFC**

#### <span id="page-27-1"></span>**Figure 12: Example of Converter Ingress Idle Packet**



### <span id="page-27-2"></span>**Table 4: Packet Qualifier for Ingress Idle Packet**



# <span id="page-27-3"></span>**Table 5: Packet Qualifier for Ingress Data Packet**





# Advance **IBM Packet Routing Switch Serial Interface Converter**



# <span id="page-28-0"></span>**Table 6: Packet Qualifier for Ingress Data Packet**



#### <span id="page-29-0"></span>**3.3.4 Egress Packet Formatter**

The Egress Packet Formatter is the interface between the egress FIFO and the PE interface logic which translates the switch LU format into the egress packet format that matches the format expected by the PE.

The Egress Packet Formatter:

- Changes the position of the proper data bytes in the packet which have been moved in the ingress path
- Inserts the switch Output Queue Grant & Shared Memory Flow control information at the byte position occupied by the bit map and some of the packet qualifier bits
- Inserts idle packets when there is no data packet to be sent

#### <span id="page-29-1"></span>**Figure 13: Example of Converter Egress Idle Packet**



#### <span id="page-29-2"></span>**Table 7: Packet Qualifier for Egress Idle Packet**





#### Advance **IBM Packet Routing Switch Serial Interface Converter**

### <span id="page-30-1"></span>**Table 8: Packet Qualifier for Egress Data Packet**



#### <span id="page-30-0"></span>**3.3.4.1 Building the Egress Packet Qualifier Byte**

- Shared Memory Information
	- The bit PQ(2) is replaced by the shared memory information (SM) whose priority level corresponds to shared memory grant priority levels PQ(0) and PQ(1) that cycle one priority after another.
- Output Queue Grant Priority Indicator
	- The first two bits PQ(3) and PQ(4) are loaded from the Output Queue Grant Priority Level values.
	- The other bits PQ(6) and PQ(7) are unchanged, as they belong to the data information contained in the packet.





#### <span id="page-31-0"></span>**Figure 14: IBM 28.4 Packet Routing Switch (switch) Packet Qualifier Bits Reshuffling**

#### • Output Queue Grant Bytes Description

- The IBM Packet Routing Switch Serial Interface Converter egress block provides the OQG values. There are 16 or 32 bits for each priority.
- For the IBM 28.4 Packet Routing Switch, the first OQG sent is an idle packet's priority level 00. The OQG mechanism gives the priority level in the idle packet's qualifier byte. The other OQGs are sent in data packets (or idle packets) and the process is looped. In every idle packet sent, the flywheel mechanism gives the priority level to check the position in the loop.
- The latest refreshed OQG bits are inserted into the packet header to be sent to the protocol engine.

# Data Packet | Bit 31 | Bit 30 | Bit 29 | Bit 28 | Bit 27 | Bit 26 | Bit 25 | Bit 24 Bit Meaning 0 Header Color/Type Color/Type 0 0 0 Packet Priority Bits Color/Type Red Complemented **1** 1 Red | | | | 1 | 0 Blue | | | | 1 | 1 Packet Priority Highest | | | | | | | | | | | 0 | 0 Medium High 0 1 Medium Low 1 0 Lowest | | | | | | | | | | | | | | | |

#### <span id="page-31-1"></span>**Table 9: Packet Qualifier for Egress Data Packet**



# <span id="page-32-0"></span>**3.4 Packet Buffering**

## <span id="page-32-1"></span>**3.4.1 X and Y Path Receive FIFO (RXFIFO)**

The RXFIFO block:

- Writes the data coming from the ingress formatter (clocked on the PE clock (URXCLK)) to the FIFO
- Reads the data from FIFO to the X and Y paths to the ingress DASL interface (IDI) block, clocked on the X/Y path clock (accounting for the size of the packet in steps of four bytes)
- Detects and reports a FIFO almost full event, whose threshold is programmable through configuration registers
- Depending upon the configuration table, it either de-asserts RXENB (if it is in use), or de-grants Shared Memory priority 0 (highest) to create room in the RXFIFO when there is an indication of FIFO almost full.

The FIFO is 256 bytes wide, or three 80-bytes packets. A 3-packet buffer temporarily stores incoming data packets, thereby allowing the insertion of a yellow packet without requiring any link level flow control. This absorbs any clock difference between the PE and the switch

## <span id="page-32-2"></span>**3.4.2 X/Y Path Transmit FIFO (TXFIFO)**

The X and Y path TXFIFO block:

- Writes the data coming from the X or Y path TX EXTRACTION BLOCK, as clocked on DASL X CLK or DASL\_Y\_CLK respectively, to the FIFO
- Reads the data from the FIFO to the path selection block, clocked on UTXCLK (PE clock)
- Detects and reports a FIFO almost full event, whose threshold is programmable
- Detects and reports a FIFO not empty, whose threshold is programmable
- Stops its data flow when the path selection block decides

Data packets are never written in the unused path's FIFO.

When the FIFO is almost full, send grant to the switch is de-asserted.

The FIFO is 480 bytes wide, i.e. six 80-bytes packets. Six packets of buffer in the TX FIFO take into account the switch's latency in reacting to the send grant (three packets in the worst case). This is true when asserting or de-asserting send grant.

# <span id="page-32-3"></span>**3.5 Path Selection Block**

This block selects the X or Y FIFO depending on the switch control's X inService or Y\_inService inputs. Only the selected path's TXFIFO is filled with packets, the other one is empty. Switching from one plane to the other is done on a packet boundary. However, there can be duplicated or missing packets.



#### <span id="page-33-0"></span>**Figure 15: Path Selection**





The attached microprocessor polls those bits to check the validity of the path selection.



Use this scheme to force any path with the force bit to perform a switch-over at switch core level and to verify that the X\_/Y\_inService lines operate properly.



# <span id="page-34-0"></span>**3.6 Interfacing Data Aligned Serial Link (DASL) Macro**

#### <span id="page-34-1"></span>**3.6.1 Ingress Data Aligned Serial Link (DASL) Interface**

The ingress DASL interface logic (IDI) is the interface between the ingress FIFO and the DASL macro. The IDI feeds the DASL macro with data packets following a DASL packet request. When no packet is available to serve the DASL, the IDI should provide an idle packet to the DASL macro. Data packets sent to the DASL macro come from the ingress FIFO. The IDI is the packet clock provider to the output logic of the ingress FIFO and to the DASL macro. The IDI also triggers the ingress FIFO output scheduler operation.

During the initial training sequence required to synchronize the remote switch DASL, the IDI provides 'synchronization packets'. The IDI is also designed to insert link liveness packets (liveness function covered by yellow packets) into the flow of packets sent to the DASL box following a request from the control interface. When data mode is activated, the IDI is fed with any data packet, otherwise it generates an idle packet (lack of data packet). Idle packets carry CRC (one per LU) to protect logical unit transport media (IBM Packet Routing Switch Serial Interface Converter and IBM 28.4 G Packet Routing Switch).

#### <span id="page-34-2"></span>**3.6.2 Functions**

- Generate Idle Packet: no data available from ingress FIFO
- Generate Synchronization Packet (when DASL training sequence is being run)
- Generate Yellow Packet (when requested by control interface)
- Forward Data Packets
- LU protection CRC insert (within idle packets)
- Provide Packet Clock to ingress FIFO
- Provide Packet Request to ingress FIFO Output Scheduler
- DASL interface LU/Byte/Bit map compatible

#### <span id="page-34-3"></span>**3.6.3 Packets Format**

#### <span id="page-34-4"></span>**3.6.3.1 Idle Packets**

Idle packets are generated by IDI logic when no data packet is available from the ingress path. Idle packets have the following format:



The Master LU must be routed to the switch operating as master.

**Note:** Yellow packets sent to the switch are considered idle packets and therefore contain CRC trailers.

#### <span id="page-34-5"></span>**3.6.3.2 Idle Packet CRC Computing**

Idle packet CRC detects physical media errors, so 4x LUs must be covered. The last byte of every LU within an idle packet carries a CRC byte protecting each LU.

The CRC polynom is X8+X4+X3+X2+1. CRC register is initialized (software configuration) depending on the LU depth.



Considering an LU, the CRC byte when inserted is computed from the end of the previous idle packet sent on the DASL interface (end of cell boundary, CRC byte), and up to the current idle packet last byte (byte preceding CRC location).

#### <span id="page-35-0"></span>**3.6.3.3 Synchronization Packets Format**

Synchronization packets synchronize DASLs. Characteristics of this packet allows the remote DASL operating as receiver to recover bit transition and packet delineation (packet clock recovery). Synchronization packet format is as follows:



#### <span id="page-35-1"></span>**3.6.3.4 Data Packets**

#### • Data Packet Format Switch 8x8 and 16x16



#### • Data Packet Format Switch 32x32



### <span id="page-35-2"></span>**3.6.4 Egress Data Aligned Serial Link (DASL) Interface (EDI)**

The IBM Packet Routing Switch Serial Interface Converter's packets have a fixed length of 64 - 80 bytes (based on the setting in the configuration register) and are mapped on 4-byte words. Therefore a new packet is received every 16 to 20 clock cycles from the DASL. When a complete packet has been received, the 16 to 20 words are transferred to the egress FIFO.

The EDI analyzes the IBM 24.8 G Packet Routing Switch (switch) packet header from the master LU to indicate the presence of an idle packet or data packet and the packet priority. Idle packets are discarded when received, (no packet write request is presented to the egress FIFO). The EDI block also checks the LU CRC (mapped into Idle Packets LU trailer). When an error is detected, the converter interrupt line is asserted (if checker enabled) and the Idle CRC error counter is incremented.

The EDI checks switch header parity. An error count is incremented when an error is detected and the processor interruption line is asserted (may be masked). Optional "error" packet discard function is provided.

The EDI performs switch grant extraction from the switch packet header. A grant information bit map is carried into the switch output packet header and is reported per priority. Therefore grant information is refreshed over four packet periods (if the four priorities are in use) for a given priority. A flywheel counter mapped into the Idle Packet Qualifier synchronizes the converter with the switch counter.




The EDI GRANT extraction mechanism is synchronized on the incoming idle packet grant flywheel counter. When a desynchronization problem is detected between the switch counter and the EDI grant flywheel counter, an error is reported and the EDI flywheel counter is automatically resynchronized. In order to minimize the latency of the switch Grant mechanism, the Grant vector is extracted before the completion of the packet reception.

The EDI performs switch memory grant synchronization. Memory Grant information and Destination Grant Vectors are passed to the in band grant generation block for insertion into the egress packets. The EDI is also the protocol engine interface wrap point.

EDI functions:

- Detect idle packets received from the switch
- Synchronize on switch Output Queue Grant counter
- Output Queue Grant Extraction (one to four priorities traffic)
- Yellow Packet Received Indication (Maskable error Interruption)
- LU CRC checking (Maskable error interruption)
- Header Parity checking (Maskable error interruption) with optional discard
- Data Packet Received Detection
- Wrap Indication Input line (Grant Indication Bypass)

### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance



### **3.6.4.1 Packets Format**

### Idle Packets

Idle packets received from EDI have the following format:

### 16x16 Switch Interface



For idle packets PQ is mapped as follows:



Parity Bit is EVEN parity over 3 bytes PQ, OQG1 and OQG2

### 32x32 SWITCH Interface



In the packet qualifier the parity bit is EVEN parity over 5 bytes PQ, OQG1, OQG2, OQG3 and OQG4

**Note:** The content of PQ byte for yellow packets received from the switch is '01001000'. It does not contain flywheel synchronization information because it is generated by the switch control. In the switch, egress yellow packets are considered data packets and therefore do not contain link CRC information fields. Yellow packets are detected and an interruption is reported when the detection function is enabled. When detection is disabled, the yellow packet is considered a normal time fill packet (idle packet).

### Idle Packet CRC Computing

Idle packet CRC detects physical media errors, so the 4x LUs must be covered. The last byte of each LU for each idle packet carries a CRC byte protecting each LU. The four LU CRC bytes are cumulative between two idle packets.

The CRC polynom is  $X8+X4+X3+X2+1$ . The CRC register is initialized (configuration table address 08 @ 28, bits 24-31) depending on the LU depth.





# **3.6.4.2 Data Packets**

The data packet format can be made of LUs of 16 to 20 bytes long depending on the configuration register.





# Data Packet Format Switch 32x32



# **3.6.5 IBM 28.4G Packet Routing Switch (switch) in band Output Queue Grant Information**

Bytes BM1 and BM2 carry the output queue grant coming from the switch. This information is carried for all 16 output ports simultaneously for a given priority. Consecutive packets, either data or idle, carry a different priority of the output queue grant bits, cycling from 0 to the highest priority value that is enabled. For instance, when two priorities are enabled, it takes two packets to transmit the output queue grant information.

## **Table 10: Output Queue Grant Bit Map Fields**



When a bit of this OQG field is set to '1', data can be sent to the corresponding output queue. When set to '0' no data should be sent to that output queue. The switch sets (8-port mode) OQG1 to 'FF' when operating in internal speed expansion. If a yellow packet is detected when detection is enabled, an interruption is reported. If detection is not enabled the yellow packet is considered an idle packet.



**IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance

# **3.6.6 IBM Packet Routing Switch Serial Interface Converter (the converter) Switch Interface**

### **Figure 16: Converter Interface Lines**



# **3.7 Egress & Ingress Interface Diagnostic Functions**

### **3.7.1 Loopbacks**

Loopbacks are controlled through configuration registers. There are two possible paths, X and Y, for each loopback, but an exclusive choice must be made and a path reset executed before the loopback is performed. The IBM Packet Routing Switch Serial Interface Converter (the converter) supports two loopback modes:

- Protocol engine (PE) X/Y loopback: the PE sends and checks data and the grant mechanism must be bypassed (forced to all ones through configuration register)
	- Protocol engine X domain loopback (register @00 bit 0 set to '1')
	- Protocol engine Y domain loopback (register @20 bit 0 set to '1')
- Switch loopback: data is initiated from the switch and the switch control verifies the overall operation
	- Switch X and Switch Y domain loopbacks: (register @C4 bit 1 set to '1'). Depending on the status of the force path and select X/Y bits of registers A0 bits 21 and 22 respectively, it is possible to use either the inService line or the select X/Y by means of bit 22 to decide which plane is selected. In the latter case, plane X is used if set to '0' and plane Y is used if set to '1'.



### **3.7.1.1 Normal Operating Mode**

Data sent by the PE to the switch core is transferred to the ingress interface where it is parity checked, LU formatted, and queued to both planes for transmission by the DASL. Data transmitted by the switch core is checked for framing consistency, header parity, queued, and word formatted for transmission by the egress interface to the PE.

# **Figure 17: Configuration in Normal Operating Mode**



# **3.7.1.2 Protocol Engine X/Y Loopback**

Protocol engine X/Y loopback provides a connection from the ingress protocol engine bus to the egress protocol engine bus. This loopback mode can be activated on either path X or path Y but not on both simultaneously. The converter in loopback mode is internally disconnected from the DASL macro for both the data and control buses and does not test the DASL macro. The loopback is initiated on a packet boundary.

**Note:** A switch plane (or substitute) clock must be present on the converter's switch interface.

**Figure 18: Protocol Engine Loopback Through Path X or Path Y** 



Perform the following to execute the PE X/Y domain loopback (starting from operational mode):

- 1. Reset the path to be set in loopback.
- 2. Force X or Y switch plane in service (switch control can also decide in this mode).
- 3. Select the clock source (switch or microprocessor) for the plane to be wrapped.
- 4. Set the PE loopback bit corresponding to the plane to be set in loopback.
- 5. Initiate the traffic from the PE.



The data transfer must have been stopped for long enough and in such a way that all buffers are empty before the loopback is initiated. If the buffers are not empty, reset the path on which the loopback is to be performed. To execute the loopback, the converter must be set in OBFC mode to avoid modifying the packet qualifier byte.

# **3.7.1.3 Protocol Engine (PE) External Loopback**

The PE bus is wrapped via the DASL interface. The switch is disconnected from the converter. A wrap plug is connected to the DASL port corresponding to the path to be tested.

**Note:** The switch plane (or substitute) clock must be selected.





The following operational sequence must be performed to execute PE external loopback:

- 1. Reset the path to be set in loopback.
- 2. Force X or Y switch plane in service.
- 3. Select the clock source (switch or microprocessor) for the plane to be wrapped.
- 4. Synchronize the DASL.
- 5. Initiate the traffic from the PE.

The PE external loopback can be performed without switch clock by using the microprocessor clock and properly programming the PLLs (the PLL must be reset to acquire its newly programmed value). In all cases the objective is to be in the frequency range of 100 - 125 MHz to allow DASL operation.

# **3.7.1.4 Switch X/Y Loopback**

Switch X/Y loopback provides a connection from the egress input to the ingress output to the ingress input through either the X or the Y path. During this test RXENB is de-asserted so the PE cannot send data.The result of this test is obtained through the switch control.



# **Figure 20: Switch X Loopback**



Perform the following operational sequence to execute the switch loopback (starting from operational mode):

- 1. Set the converter in OBFC mode.
- 2. Force X or Y switch plane in service or let the switch control decide which switch plane is in service.
- 3. Set the switch loopback bit corresponding to the plane to be set in loopback.
- 4. Initiate the traffic from the switch control.

The data transfer must have been stopped for long enough and in such a way that all buffers are empty before the loopback is initiated. If the buffers are not empty, reset the path on which the loopback is to be performed. To execute the loopback, the converter must be set in OBFC mode to avoid modifying the packet qualifier byte.



### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance

# **3.8 Clocks Generator Description**

# **Figure 21: Clocks Distribution Diagram**



**Table 11: Selecting the Signal That Appears on the TO\_SMOOTH\_PLL\_IN Signal** 





# **Table 12: External Clocks Description**



# **3.8.1 IBM Packet Routing Switch Serial Interface Converter Internal Clocks Description**

The converter clock tree is articulated around three PLLs and the microprocessor (MP\_CLK) clock. Seven clock domains exist to clock the converter. The clock generator acts as a programmable selector. UTXCLK and URXCLK internal clock trees are generated from several clock sources that are selected by programming the "UCLK source  $1\left[1:0\right]$ " and "XY SW source enb  $1\left[1:0\right]$ " in the configuration table registers.

# **3.8.2 IBM Packet Routing Switch Serial Interface Converter External Traffic:**

External traffic from the converter and the IBM 28.4 G Packet Routing Switch (switch) is timed using the two edges of the clock issued from the SWITCH\_X\_PLL or the SWITCH\_Y\_PLL. SWITCH\_X/Y\_PLLs deliver two clocks which are 100 - 125 MHz (DASL\_X125\_CLK) and 200 - 250 MHz (DASL\_X250\_CLK) respectively. The X/Y transmit and the X/Y receive DASLs are clocked by DASL\_X250\_CLK and DASL\_Y250\_CLK respectively. Receive and transmit data is then input and output at a frequency of 400 - 500 MHz. At each occurrence of the DASL\_X/Y250\_CLK clock edge, transition data is sent to or received from the switch.



# **3.9 IBM Packet Routing Switch Serial Interface Converter RESET Scheme Description**

The converter allows total and selective reset. PATHX and PATHY can be reset together or independently. The ingress and egress interfaces and their associated FIFOs are reset independently. External reset, programmable reset, and power-on-reset are implemented in the converter.

# **3.9.1 Reset Strategy**

The converter is reset when power-on-reset (POR) is active. T/S drivers are in Hi-Z state and bidirectional I/ Os are sourced to receiver mode. The up\_interface and configuration table register are reset, and then the PE interface, FIFOs and their associated logic, and PATH\_X/Y are reset.

When RESET\_X is active, PATH\_X is reset. When RESET\_Y is active, PATH\_Y is reset. DASL\_X/Y differential drivers are in Hi-Z state.

The uP interface initiates the reset by a write access to the bit that needs to be reset. A second write access is necessary to restore the bit to its inactive position.

The following table shows the different reset cases:

# **Table 13: Register Reset Settings**



# **3.9.2 Power-On-Reset (POR) Procedure**

The POR action uses the external MP\_CLK microprocessor clock to reset the chip. All drivers are set to high impedance during the POR. The register table controlling those output lines must be reconfigured after POR.

After POR action, the converter must be re-configured via the  $\mu$ P\_interface. The PLLs are switched into BYPASS mode and neither PATH X nor PATH Y are configured.

POR external signal is an asynchronous signal. POR initiates when the POR signal input is going down. POR procedure stops when the POR signal input is held up. POR must be asserted for at least 10 MP\_CLK cycles to insure proper chip reset.

The following table shows pin/pad TEST I/O initialization values (the values during and after reset):

### **Table 14: I/O Initialization Values**



### **3.9.3 Path Reset**

Individual PATH X and PATH Y resets are performed by addressing an appropriated bit in the COMMON\_CONFIG\_REGISTER @A0 (CCR). SWITCH\_X\_PLL or SWITCH\_Y\_PLL must run during the PATH X/Y reset action. PLLs are not affected by PATH X/Y resets, but the PLL must be reset when an external loopback is initiated with a new clock source (PLL needs a delay to lock). Once a reset is established, the software must reconfigure the CCR @A0 bits 0, 1, 2, 3 into the system mode.

### **Table 15: Path Resets**



### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance

### **3.9.4 PLL Reset**

Once a reset is established, the software must reconfigure the SWITCH\_PLL X/Y bit or PLL\_PE bit 15 to '0' (system mode).

## **Table 16: System Mode PLL Resets**



### **3.9.5 Ingress/Egress Interface Reset**

Ingress and egress interfaces are reset on software request. An interface reset restores the internal scheduler to the idle phase. Transfer data is lost and internal registers are swapped to their reset position. Signals that drive the PE outputs are switched so to hold the PE interface drivers in inactive state.

Both ingress and egress PE interfaces are reset simultaneously by programming @A0 CONFIG reg bit 8 at '1' in the configuration table registers. Once the reset is established, the software must reconfigure the @A0 CONFIG\_reg bit 8 at '0' (system mode).

# **3.10 Microprocessor Interface Description**

The IBM Packet Routing Switch Serial Interface Converter (the converter) chip is initialized and controlled via a processor interface which works on an 8-bit data bus and operates in two modes. The external input pin MP\_BURST\_SEL selects the operational mode (low - 8-bits mode and high - 32-bits mode):

- 8-bit mode (byte mode): The converter registers are considered single 8-bits registers and are addressed via MP\_ADD[7:0] address bus signals. Each register access is a single-beat access of one byte.
- 32-bit mode (burst mode): The converter registers are considered as single 32-bits registers and are addressed via MP\_ADD[7:2] address bus signals. Each access is a burst access of four bytes. The burst order is the following: data bits [7:0], data bits [15:8], data bits [23:16], and data bits [31:24].

3.10.1 The microprocessor interface:

- Provides read/write access to all chip registers
- Provides DASL picocode downloading
- Provides error reporting
- Collects the converter interrupts and pass them to the attached processor.
- Monitors all interrupt signals generated by other converter functional blocks and, when one is asserted, latches and holds the value until the interrupt event register is read and reset.
- Provides the necessary handshake protocol to interface the attached processor, including address bus decoding, wait state insertion, data bus drivers control, and optional parity checking on both the data and address busses.



IBM3209K3114

# **3.10.2 Processor Interface Lines**

**Figure 22: IBM Packet Routing Switch Serial Interface Converter Processor Interface Lines** 



# **3.10.3 Processor Interface I/O Lines Description**

The processor interface is synchronized to the external processor clock. This clock operates at a different frequency than the switch fabric clock.







### **3.10.4 32-Bit Mode Processor Interface Timing**

# **Figure 23: Processor Read Access in 32-bit Burst Mode**



**Figure 24: Processor Write Access in 32-bit Burst Mode** 





# **3.10.5 8-Bit Mode Processor Interface Timing**

### **Figure 25: Processor Read Access in 8-bit Byte Mode**



### **Figure 26: Processor Write Access in 8-bit Byte Mode**



**Note:** MP\_SEL can be active one more cycle without impact on the processor access, however the minimum deactivate time must remain two clock cycle times.

### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance



# **4. Converter Configuration Table Registers**

This module contains all the configuration registers that the converter needs to operate.

- Registers are referenced from @00 up to @FF.
- There are two ways to access the configuration table registers:
	- During the POR procedure, some registers can be preset to start the system as predetermined.
	- Using the microprocessor interface, the configuration table register can be either READ or WRITE.
- Write access to a read-only register is not valid and does not change the held value of that register.
- All errors are logged (not first failure data capture).
- All bits in configuration table registers are active binary 1.
- The registers can be accessed either in 8-bit (byte) mode or in 32-bit (burst) mode through bursts of 4x 8 bits modes (for I960 Processors). In 32-bit mode, the least significant byte is sent first, then the second least significant byte, and so forth.
- The registers are based on little endian notation.

### **Figure 27: Register Mapping**







# **4.1 Error Detection, Reporting, and Interrupt Registers**

The converter uses a common strategy for error detection, error reporting, and interrupt generation:

- Each error is detected by an individual checker (for instance parity checker)
- An error must be individually enabled in the CHECKER\_ENABLE\_REGISTER to be reported into EVENT\_REGISTER
- An error must be individually enabled in the INTERRUPT\_ENABLE register to generate an interrupt
- The INTERRUPT\_REGISTER\_INDIRECTION is the first register to read when an interrupt is raised. The interrupt cause can either be present in the register itself or be via an indirection to another EVENT\_REGISTER
- For some specific errors (Parity and CRC), it is possible to discard or not discard the corresponding packet, depending on the setting of the corresponding configuration register

The interrupt routine is as follows:

- 1. The processor reads the INTERRUPT\_REGISTER\_INDIRECTION for the interrupt cause (either directly or indirectly via a second read into the EVENT\_REGISTER flagged by INTERRUPT\_REGISTER\_INDIRECTION)
- 2. After processing the interrupt routine, the processor resets only the EVENT\_REGISTER. The reset is under mask, so the whole byte of a register can be reset by writing x'FF' inside, or it can be reset a single bit at a time by writing a '1' at the bit position in the byte to be reset. However two bits of a byte cannot be reset simultaneously (applicable only to Event registers @ 10, 30, 88)

### IBM Packet Routing Switch Serial Interface Converter Advance



# **4.1.1 Register Map**





IBM3209K3114

### Advance **IBM Packet Routing Switch Serial Interface Converter**

# **4.1.1.1 Setup 1\_X PATH Register**





<span id="page-54-0"></span>





### **4.1.1.2 Setup 2\_X Path Register**





# **Access Type** Read/Write



The settings of Set Up Register 2 depend on the packet size, the PE clock, the switch clock, and the flow control mechanism in use.

Example 1: for a 64-bytes packet and a PE and switch using the same range clock frequency, the setting might be:

- Ingress buffer with out of band flow control (RXENB)
	- Almost empty: x'10' for a 16 word-packet (packet of 64 bytes) or above to avoid underrun
	- Almost full: x'2C' or below to avoid overrun
- Egress buffer
	- Almost empty: x'10' or above to avoid underrun
	- Almost full: x'3D' or below to avoid overrun

Example 2: for an 80-bytes packet and a protocol engine and switch clock at 110 MHz, the setting might be:

- Ingress buffer with out of band flow control (RXENB)
	- Almost empty: x'14' for a 20 word-packet (packet of 80 bytes) or above to avoid underrun
	- Almost full: x'28' or below to avoid overrun
- Egress buffer
	- Almost empty: x'14' or above to avoid underrun
	- Almost full: x'30' or below to avoid overrun



# **4.1.1.3 Control \_X PATH Register**



### **Reset Output Status** (Power-on-Reset, Software Reset Path X) '0000 0C00'

**Address in Word Mode** x'08'

**Address in Byte Mode** X'08 to 0B'

**Access Type** Read/Write



### IBM3209K3114



### IBM Packet Routing Switch Serial Interface Converter Advance





# **4.1.1.4 X Plane Parity and CRC\_Error\_Count\_X Registers**

The microprocessor has the highest priority when writing into this register, so it may at any time overwrite the current value even if the counter is being incremented by the source of an error.

## **Note:** All the counters do not wrap around.



**Reset Output Status** (Power-on-Reset, Software Reset Path X) All '0's



**Address in Byte Mode** X'0C to 0F'

## **Access Type** Read/Write





### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance

# **4.1.1.5 X Plane Events 1 Register (Event 1 \_X)**



# **Reset Output Status** (Power-on-Reset, Software Reset Path X) All '0's

### **Address in Word Mode x**<sup>1</sup>10'

# **Address in Byte Mode** X<sup>'10</sup> to 1B'

# **Access Type** Read/Write



<span id="page-59-1"></span><span id="page-59-0"></span>2.  $S =$  status and is reported through polling.

<span id="page-59-2"></span>3. In case of buffer overrun all packets are discarded.





3. In case of buffer overrun all packets are discarded.





# **Figure 29: DI and Data Aligned Serial Link (DASL) Startup Sequence Path X**

# **Figure 30: Enabling of Data Aligned Serial Link (DASL) Data Transmission and Reception Path X**







# **4.1.1.6 X Plane Event 1 Checker Enable Register (Event 1 Checker Enable\_X )**



# **Address in Byte Mode** X<sup>'14</sup> to 17'

**Access Type** Read/Write



**Note:** To enable the checker corresponding to an event, set '1' in the corresponding bit position.



# **4.1.1.7 Interrupt Enable \_X Register**







**Note:** To enable the interrupt a '1' must be set in the corresponding bit position.



# IBM3209K3114

# Advance **IBM Packet Routing Switch Serial Interface Converter**

# **4.1.1.8 Setup 1\_Y PATH Register**





<span id="page-64-0"></span>





# **4.1.1.9 Setup 2\_Y PATH Registers**



# **Address in Byte Mode** X'24 to 27'

**Access Type** Read/Write





# **4.1.1.10 Control \_Y PATH Registers**



**Reset Output Status** (Power-on-Reset, Software Reset Path Y) '0000 0C00'

**Address in Word Mode** x'28'

**Address in Byte Mode** X'28 to 2B'

**Access Type** Read/Write





### IBM Packet Routing Switch Serial Interface Converter Advance







# **Figure 31: DI and Data Aligned Serial Link (DASL) Startup Sequence Path Y**

# **Figure 32: Enabling of Data Aligned Serial Link (DASL) Data Transmission and Reception Path Y**



### **IBM Packet Routing Switch Serial Interface Converter Advance** Advance

# **4.1.1.11 Y Plane Parity and CRC\_Error\_Count\_Y Registers**

The microprocessor has the highest priority when writing into this register, so it may at any time overwrite the current value even if the counter is being incremented by the source of an error.

## **Note:** All the counters do not wrap around.



**Reset Output Status** (Power-on-Reset, Software Reset Path Y) All '0's **Address in Word Mode** x'2C'

**Address in Byte Mode** X'2C to 2F'

 $\Box$ 

**Access Type Read/Write Read/Write Read/Write Read/Write** 











# **4.1.1.12 Y Plane Events 1 Register (Event 1 \_Y)**



# Reset Output Status (Power-on-Reset, Software Reset Path Y) All '0's

### **Address in Word Mode** x'30'

# **Address in Byte Mode** X'30 to 3B'

## **Access Type** Read/Write



<span id="page-70-1"></span>1.  $I = an$  interrupt is raised because of error or status change.

<span id="page-70-0"></span>2. S = status and is reported through polling.

<span id="page-70-2"></span>3. In case of buffer overrun all packets are discarded.

### IBM3209K3114

### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance**



1.  $I =$  an interrupt is raised because of error or status change.

2.  $S =$  status and is reported through polling.

3. In case of buffer overrun all packets are discarded.



Ē,

Ť


## **4.1.1.13 Y Plane Event 1 Checker Enable Register (Event 1 Checker Enable \_Y)**







**Note:** To enable the checker corresponding to an event, set '1' in the corresponding bit position.



## **4.1.1.14 Interrupt Enable \_Y Register**







**Note:** To enable the interrupt a '1' must be set in the corresponding bit position.



### **4.1.1.15 DASL M3 Picocode X Register**







### Read and Write Operations

The write operation is performed as follows:

- 1. Set data byte in @ 40
- 2. Set data byte in @ 41
- 3. Set address in @ 42
- 4. Set the remaining bits of the address and bit 30 in @ 43 to issue the write operation
- 5. Reset bit 30 in @ 43

Repeat the above operational sequence as long as there is data to be loaded in the DASL. Bit 30 does not need to be reset between two successive write operations. For example, in the sdc.h file provided by IBM, the syntax in cmd file: WRITE (ADDRESS >= x'40', DATA >= x'4000900F'); where '4000' means write enable and 000 is the start address) is Start Address and '900F' is Data.

The read operation is performed as follows:

- 1. Set address byte in @42
- 2. Set the remaining address byte and bit 31 in @43
- 3. Poll @43 bit 29 for read completion. Then the data is valid and can be read from @40 and 41
- 



































### **4.1.1.21 DASL M3 Picocode Y Register**





























7-0 7-0 YSDC\_Add0 Debug Bus Address









### **4.1.1.27 Event 2 Checker Enable\_X and \_Y Registers**



**Reset Output Status** (Power-on-Reset, PE Reset) All '0's

#### **Address in Word Mode** x'80'

**Address in Byte Mode** X'80 to 83'





#### **IBM Packet Routing Switch Serial Interface Converter Advance** Advance

## **4.1.1.28 Event 2 µP Interrupt Enable\_X and \_Y Registers**



Reset Output Status (Power-on-Reset, PE Reset) All '0's

### **Address in Word Mode** x'84'

## **Address in Byte Mode** X'84 to 87'





## **4.1.1.29 Event 2 \_X and \_Y Registers**

Some of these register bits directly reflect the status of some chip input lines, so their value depends on those lines.





### **Address in Word Mode** x'88'

**Address in Byte Mode** X'88 to 8B'

**Access Type** Read/Write



**Note:** Some of the bits in this register are directly forced by the hardware. Even if they are Read/Write, when the microprocessor writes a value different from the current bit content, the hardware will immediately overwrite the microprocessor value.

#### IBM3209K3114



### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance**



**Note:** Some of the bits in this register are directly forced by the hardware. Even if they are Read/Write, when the microprocessor writes a value different from the current bit content, the hardware will immediately overwrite the microprocessor value.



### **4.1.1.30 ABIST Failure Test Status \_X\_Y Registers**

This register is not accessible by the user and the bit should not be set to '1'. It is used only during Manufacturing test.



**Address in Word Mode x'8C'** 

**Address in Byte Mode** X'8C to 8F'

<span id="page-90-0"></span>



#### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance

## **4.1.1.31 Switch X PLL Setting Register**

These settings correspond to a VCO setting of 660 MHz. The PLL observe bits will automatically overwrite what has been written by the microprocessor.



**Reset Output Status** (Power-on-Reset) (2800 C000'

**Address in Word Mode** x<sup>'90'</sup>

**Address in Byte Mode** X'90 to 93'





#### IBM3209K3114

## **4.1.1.32 Switch Y PLL Setting Register**

These settings correspond to a VCO setting of 660 MHz. The PLL observe bits will automatically overwrite what has been written by the microprocessor.



**Reset Output Status** (Power-on-Reset) (2800 C000'

**Address in Word Mode** X'94'

**Address in Byte Mode** X<sup>'94</sup> to 97'





## **4.1.1.33 Chip ID Register**



0-31 ChipID 34423600





## **4.1.1.34 Protocol Engine PLL Setting Register (PE PLL Register)**



## **Reset Output Status** (Power-on-Reset) (2000) (2000'

### **Address in Word Mode** x'9C'

**Address in Byte Mode** X'9C to 9F'



### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance



## **4.1.1.35 Common Control Register**

The PLL observe bits automatically overwrite the microprocessor writes.



## **Reset Output Status** (Power-on-Reset) **1998** (0500413F'

**Address in Word Mode** X'A0'

**Address in Byte Mode** X'A0 to A3'







### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance



### **4.1.1.36 Interrupt Register Indirection**



## **Access Type** Read/Write



**Note:** The PLL Clock missing bits are not implemented.

**Note:** The bits in this register are followers of what is occuring in other triggering registers. To clear a bit in this register, its source must be cleared. For example, if a bit is set in Register 18 (Interrupt Enable \_X), Register 38 (Event 1 Interrupt Enable \_Y), or Register 84 (Event 2 µP Interrupt Enable \_X and \_Y), the corresponding bit in this register will be set. When the bits in those registers are cleared, the corresponding bit in this register is cleared.



## **4.1.1.37 Ingress PE Settings Register (INGRESS\_PE\_INTERFACE (IPI) - Receive)**



## **Reset Output Status** (Power-on-Reset, PE reset) '20018000'

## **Address in Word Mode** x'C0'

## **Address in Byte Mode xicology xicology xicology xicology xicology xicology**

## **Access Type** Read/Write



0 Disable shadow receive clock out driver

Disable PE TX clock out driver

9 1 Shadow\_RXCLK\_Out\_En 1 Enable shadow receive clock out driver<br>Disable shadow receive clock out driver

8 0 PE\_TXCLK\_Out\_En 1 Enable the PE TX transmit clock out driver 1 Enable the PE TX transmit clock out driver







## **4.1.1.38 Egress PE Setting Register (EGRESS\_PE\_INTERFACE - Transmit)**



**Reset Output Status** (Power-on-Reset, PE reset) '38432100'

**Address in Word Mode** X'C4'

**Address in Byte Mode** X'C4 to C7'









## **4.1.1.39 Common PE Setting Register (PE (Common))**



### **Reset Output Status** (Power-on-Reset) '01321008'

### **Address in Word Mode** X'C8'

**Address in Byte Mode** X'C8 to CB'









## **4.1.1.40 Parity and CRC Error Count Register (PARITY\_Error\_Count)**

The microprocessor has the highest priority when writing into this register.



3-0 3-0 RXHPRTY\_err\_count IBM 28.4 G Packet Routing Switch Header Parity Error Count



# **5. IBM Packet Routing Switch Serial Interface Converter Latency**

There are two kinds of flow control latencies to be considered: packet latency and flow control latency.

Packet latency is shown below. Packet reshuffling is the main cause because, on ingress, the entire packet must be received before it can be ascertained if the packet must be discarded due to errors. The re-synchronization layer between the PE clock domain and the switch clock domain is the secondary contributor to packet latency.



## **Figure 33: Converter Latency Diagram**

Flow control information latency has two impacts:

- The shared memory thresholds within the switch have to be set to low sharing (for example: Priority 3 to 16 Packets, Priority 2 to 32 packets, Priority 1 to 48 packets, and Priority 0 to 64 packets).
- The flow control check bit 10 of mode register must be disabled.



# **6. JTAG Description**

The IBM Packet Routing Switch Serial Interface (the converter) is compliant with the IEEE Standard Test Access Port and Boundary-Scan Architecture for testing and debugging components assembled at the board level. (See "IEEE Standard Test Access Port and Boundary-Scan Architecture", doc. IEEE Std 1149.1-1990, and the IBM ASIC Products Application Note "IEEE 1149.1 Boundary-Scan in IBM ASICs", 11/97, version 4.)

JTAG test architecture consists of a Test Access Port (TAP) and associated controller, an Instruction Register, and (on the converter) three Test Data Registers named the Bypass Register, the Idcode Register, and the Boundary Scan Register.

The TAP, a general-purpose port, provides access to the implemented test support functions defined by the JTAG Standard. It consists of the following five ports: the Test Clock Input (TCK), the Test Mode Select Input (TMS), the Test Data Input (TDI), the Test ReSeT input (TRST), and the Test Data Output (TDO). The TAP controller is a synchronous finite state machine that responds to changes at the TMS and TCK signals and controls the sequence of operations of the internal test logic. It generates the clock and control signals required to shift data down the Instruction Register and Test Data Registers. These registers constitute separate shift-register based paths that are connected in parallel between the TDI and TDO.

The Instruction Register is three bits long. It captures the Instruction Opcode desired during a test. The converter supports the following instructions declared in the JTAG Standard: EXTEST, INTEST, SAMPLE, IDCODE, HIGHZ, CLAMP and BYPASS.



## **Table 17: Supported JTAG Instructions**

The Bypass Register is a one bit long shift register and is scannable during the BYPASS instruction. It provides a minimum length serial path to move test data between TDI and TDO. Use the Bypass Register to speed access to JTAG registers in other components on a board-level test data path.

The Boundary Scan Register is a shift register that allows sampling and/or forcing of signals flowing into and out of the system logic through the system ports.

During the JTAG operation the following converter PIs must be at the specific value indicated in the following table. As these PIs are already internally tied up or tied down, they do not need any user force action except the RI line, which must be forced to 1 externally.



## **Table 18: Compliance Pattern**



The IDcode register is a 32-bit SRL 32h'34423600'. Physically the IDcode register is implemented this way:

TDI->Version (31:28)->Location (27:24)->ID -UDASL P/N- (23:12)->Manufacturer ID (11:1)->'1'->TDO converter.

## **Table 19: ID Code Description**




# **7. I/O Definition and Package Pin Assignment**

# **7.1 Signals Description**

All functional signals are 3.3 V LVTTL compatible for drivers and receivers except the following:

- Protocol engine interface
- Protocol engine clocks provided by the IBM Packet Routing Switch Serial Interface Converter
- Back pressure serial link
- LSDD and JTAG test signals

These are based on the tri-state driver/receiver (BP2550 type) that interfaces 2.5 V internal functions with 3.3 V-tolerant 2.5 V CMOS drivers and receivers off chip bidirectional data buses. The driver is 50 Ω sourceterminated.

The switch fabric clocks are balanced HSTL levels.

#### **Table 20: Tests Signals**



<span id="page-108-0"></span>1. An internal pull-up resistor forces the inactive state.

<span id="page-108-1"></span>2. Must be kept LOW during normal PLL operation.

#### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance**

#### **Table 20: Tests Signals**



2. Must be kept LOW during normal PLL operation.

#### **Table 21: JTAG Interface External Signals**



<span id="page-109-1"></span><span id="page-109-0"></span>





#### Name Input/ nput Levels **Description** MP\_SEL Input LVTTL IBM Packet Routing Switch Serial Interface converter select for processor access MP\_WR Input LVTTL Write/Read indicator on processor bus MP\_PRDY | Output | LVTTL | Ready signal for transfer complete indication MP\_INT | Output | LVTTL | Interrupt to processor MP\_ADDR\_7...<br>MP\_ADDR\_0 Inputs | LVTTL Processor bus address MP\_ADD\_PRTY | Inputs | LVTTL | Processor bus parity (optional and can be set through an interface pin) MP\_DATA\_7... MP\_DATA\_0 Input/ LVTTL Processor Data bus MP\_DATA\_PRTY | Input/  $LVTTL$  Processor Data bus parity (optional and can be set through an interface pin)  $\overline{MP}$  PRTY\_ENB  $\overline{RP}$  Input LVTTL Enable generation of and checking of the microprocessor data and address parity MP\_BURST\_MODE | Input | LVTTL Enable the operation in either single byte mode access or in byte burst mode compatible with I960. When tied to ground, the chip operates in byte mode access, when tied to 3.3 V it operates in word mode access.

#### **Table 22: Processor Interface Signals**







#### **Table 23: IBM Packet Routing Switch Serial Interface Converter (the converter) Signals**

**Note:** The electrical level supported by the converter lines is called High Speed Transceiver Logic (HSTL). It is specified in JEDEC's JESD8-6 standard.



# **Figure 34: Detection of Line Card Fully Inserted**



#### **Table 24: Receive PE Interface Signals** See PE interface description (ingress and egress descriptions)for functional description



#### **Table 25: Transmit PE Interface Signals**







#### **Table 26: Clocking/PLL External Signals**

To provide isolation from the noisy internal digital  $V_{dd}$  signal, VDDA is brought to a package pin. If little noise is expected at the board level, then VDDA can be connected directly to the digital  $V_{dd}$  plane. In most circumstances, it is prudent to palce a filter circuit on VDDA as shown below. All wire lengths should be kept as short as possible to minimize coupling from other signals. The impedance of the ferrite bead should be much greater than that of the capacitor at frequencies where noise is expected. Many applications have found that a resistor, instead of a ferrite bead, does a better job of reducing jitter. The resistor should be kept to a value lowr than 2 Ω. Experimentation is the best way to determine the optimal fileter design for a specific application.



IBM3209K3114

## **Figure 35: VDDA Filtering**



• VDDA filtering for each PLL

The impedance of the ferrite bead should be much greater than that of the capacitor at frequencies where noise is expected. Many applications have found that a resistor instead of a ferrite bead does a better job of reducing jitter. The resistor should be kept to a value lower than  $2 \Omega$ . Experimentation is the best way to determine the optimal filter design for a specific application.

Name	Input/Out- put	Level	Description
START GCXFR	Output	<b>LVCMOS</b>	Start of transfer of the switch grant information asserted by the converter to the PE to indicate the position of the start of the overall priority cycle in a superframe. The superframe is made of as many frames as there are priori- ties in use. This signal is Hi-Z when the interface is not selected
ODD OQG	Output	<b>LVCMOS</b>	Contains the information per priority related to the switch output queues 0-7 and 16-23. This signal is Hi-Z when the interface is not selected
EVEN OQG	Output	<b>LVCMOS</b>	Contains the information per priority related to the switch output queues 8-15 and 24-31. This signal is Hi-Z when the interface is not selected.
SHARED GNT	Output	<b>I VCMOS</b>	Indicate in a frame the relevant priority on the ODD/EVEN OQG serial link and carry the global shared memory grant. The last bit of each frame carries the odd parity bit computed on all data bit in the frame. This signal is Hi-Z when the interface is not selected.

**Table 27: Back Pressure Serial Link Signals** 





#### **Table 28: Miscellaneous External Signals**

<span id="page-115-0"></span>1. An internal pull-up resistor forces the inactive state.

## **Figure 36: Switch Present Detection**





## **Table 29: Spares Signals Used to Carry Additional DC Voltages**



## **Table 30: Debug Purpose External Signals**























# **7.2 I/O Timing**

## **7.2.1 IBM Packet Routing Switch Serial Interface Converter A.C. Characteristics**

#### **7.2.1.1 AC parameters characteristics**

The converter's external interfaces (UTOPIA-3 like, Processor, and DASL) specify that all inputs and outputs are registered. Only Set-up times and Hold times for inputs and Clock-to-outputs times for Outputs are needed.

#### **Figure 37: AC parameters timing diagram**



**7.2.1.2 Protocol Engine (UTOPIA-3 like) Interface A.C. Specifications**





The following drawing shows the flexibility in implementing the UTOPIA-3 like interface timing.

#### **Figure 38: Options for Ingress UTOPIA-3 Like Interface Clocking**



#### **7.2.1.3 Microprocessor Interface A.C. Specifications**





# **8. Electrical Specifications**

#### **Table 32: Absolute Maximum Ratings**



**Note:** Permanent device damage may occur if the above absolute maximum ratings are exceeded. Functional operation should be restricted to the conditions as detailed in the operational section of this data sheet. Exposure to absolute maximum rating conditions for extended periods may affect device reliability.

# **8.1 Power Sequencing**

In order to prevent latchup (and destruction) of the chip, the power supplies must be sequenced up and down in a manner that any supply is greater than or equal to the voltage of another supply of lesser value both during "turn on" and "turn off" (including a quick toff - ton sequence).

For example, 3.3 V  $\ge$   $\ge$  2.5  $\ge$  1.5 V during power up and power down. Actually up to 400 mV of negative voltage can be tolerated during the sequence between any two supplies. There is no time delay requirement, only a negative voltage restriction. Ideally, all supplies would rise together until they reach their operating level and all would fall together until they reach zero.

If one could design a 400 mV (max) diode and place it between 3.3 V (cathode) and 2.5 V (anode), and another between 2.5 V (cathode) and 1.5 V (anode), that would ensure that the 400mV diodes internal to the chip did not get forward biased more than the 400 mV, which would keep the chip out of latchup conditions.

Also, the 3.3 V supply should not exceed the 2.5 V supply by more than 2.7 V during power up/down, and 1.3V in constant use.



# **8.2 Recommended Operating Conditions**

## **Table 33: LVCMOS Compatible I/Os**



# **Table 34: LVTTL Compatible I/Os**



#### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance**



## **Table 35: Recommended Operating Conditions for all I/Os**



# **Table 36: Power Dissipation**





# **8.3 Signal Pin Assignments**

# **Table 37: Signal Pins Sorted by Grid Location** (Page 1 of 4)



#### IBM Packet Routing Switch Serial Interface Converter Advance



## **Table 37: Signal Pins Sorted by Grid Location** (Page 2 of 4)





## **Table 37: Signal Pins Sorted by Grid Location** (Page 3 of 4)



#### IBM Packet Routing Switch Serial Interface Converter Advance



## **Table 37: Signal Pins Sorted by Grid Location** (Page 4 of 4)





# **Table 38: Signal Pins Sorted By Signal Name** (Page 1 of 4)



#### IBM Packet Routing Switch Serial Interface Converter Advance



## **Table 38: Signal Pins Sorted By Signal Name** (Page 2 of 4)





## **Table 38: Signal Pins Sorted By Signal Name** (Page 3 of 4)



#### IBM Packet Routing Switch Serial Interface Converter Advance



## **Table 38: Signal Pins Sorted By Signal Name** (Page 4 of 4)





# **8.4 Power Signals**

# **Table 39: Ground Signals**



#### IBM Packet Routing Switch Serial Interface Converter Advance

# Table 40: 2.5 V V<sub>DD</sub> Signals







#### **Table 41: 1.5 V South VDD 2 Signals**



## **Table 42: 1.5 V East VDD 3 Signals**



# **Table 43: 3.3 V North VDD 4 Signals**



# **Table 44: 1.5 V West VDD 5 Signals**



#### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance

# **9. Packaging Information**

CBGA 360 with standard aluminum lid and grease.







# **10. Appendix A: Data Aligned Serial Link (DASL)**

# **10.1 General Description**

The Data Aligned Serial Link interface is intended for high speed point to point inter-chip communication. The interface macro performs multi-bit serialization and de-serialization to reduce the I/O pin count. It is a synchronous device, removing the need for asynchronous interfaces that introduce additional interface latency. DASLs are intended to operate over a back-plane without any additional components. The DASL macro is designed for high levels of integration and uses reduced voltage differential transceivers to reduce power consumption.

#### **Figure 40: Data Aligned Serial Interface Lines**





## **Table 45: Interface Description for DASL Internal Signals** (Page 1 of 3)





## **Table 45: Interface Description for DASL Internal Signals** (Page 2 of 3)



#### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance**

## **Table 45: Interface Description for DASL Internal Signals** (Page 3 of 3)





## **10.2 Resets**

- SDC RESET: (M3 Reset) The DASL processor reset must be active during picocode download. It is recommended to activate SDC\_RESET during DASL reset (CORE\_RESET).
- CORE\_RESET: (DASL Reset) CORE\_RESET signal resets the DASL. To insure proper reset, it is recommended to maintain the CORE\_RESET signal during at least 500 ns.

Both SDC\_RESET and CORE\_RESET signals are synchronous resets, and therefore need the presence of a clock to be performed.

# **10.3 Picocode Download**

The Instruction Memory for the DASL Shared Controller is accessible through the SDC Internal Resource Interface (SDCIRI). The processor is given a separate reset so that it can be disabled until the Instruction Memory is loaded. The SDC\_RESET signal must be held active until Instruction Memory is completely loaded. Once the picocode is completely loaded, release the SDC\_RESET to start the processor.

#### **10.3.1 Picocode Write**

A picocode write operation is performed using the M3 picocode access register @40 for X plane and @60 for Y plane. Data Instruction, Instruction Memory Address, and Write Enable bits must be written in the SDC access register. The hardware handles the transfer in the Instruction Memory. Contiguous operations can be performed until the picocode is completely downloaded.

#### **10.3.2 Picocode Read**

A picocode read operation is performed using the M3 picocode access register @40 for X plane and @60 for Y plane. Instruction Memory Address and Read Enable bits must be written in the SDC access register. The hardware handles the read operation.

The SDC access register shall be polled until the Read Completed bit is detected to be ON (active high). When this bit is ON, the data contained in the SDC access register is valid and correspond to the selected Instruction Memory Address.

# **10.4 SDC\_INTERRUPT Signal**

A high level on SDC\_INTERRUPT (M3 interruption in registers 10 and 30) denotes an interrupt or a parity error condition. If a parity error is detected on the instruction memory or the local data store, an interrupt will be generated and remain active until the Shared DASL Controller is reset.

# **10.5 SDC Debug Interface**

The SDC debug interface allows the application read and write access to any of the resources available to the processor. Access includes local store, sample memory, and any hardware-assist registers addressable by the processor (access to instruction memory is provided through SDC Internal Resource Interface). The inter-



#### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance**

face includes a 32-bit debug control input, a 32-bit debug data output, a 32-bit debug data input, a 16-bit debug address input, and a one bit debug data valid output. The application requests a service offered by the processor via the debug control input. The results are returned by the processor on the debug data output. The debug control input definition is shown below.



The definition of each field is shown in the following table.

#### **Table 46: Debug Control Field Definitions**




#### **Table 46: Debug Control Field Definitions**



The status register contain various information about the processor status. The output of this register is connected to the SDC\_STATUSREG\_OUT output.

#### **Table 47: Status Register Definition**



#### **Table 48: Data Aligned Serial Link (DASL) Signals for Synchronization**



### **10.6 Data Aligned Serial Link (DASL) Initialization and Operation**

Once the chip has been fully configured, but before actual data traffic can take place between an IBM Packet Routing Switch Serial Interface Converter (the converter) and an IBM 28.4G Packet Routing Switch (switch), the DASL interfaces must be initialized to provide bit phase alignment and packet alignment at the data receivers in both directions.

DASL initialization means communication between two chips: a converter and a switch core port.

The port synchronization is under the overall control of the system Control Processor which coordinates the operation between the switch core and the adapters, but synchronization between the switch control and the port adapter can also be performed directly through the interface lines, Fabric Port Available (FPA), and Adapter Port Available (APA).

The registers of interest at both ends of the link are:

• Port enable register

#### **IBM Packet Routing Switch Serial Interface Converter Advance Advance Advance** Advance



- No signal register
- Sync Status and Sync Hunt register
- Sync Packet Transmit register
- Signal detect Interrupt in Status register

The Sync Status register reports the status of the input receiver, and the Sync Hunt register forces the input ports to start the synchronization sequence.

The Sync Packet Transmit register specifies that synchronization packets have to be transmitted in order for the remote chip input port to synchronize. When not transmitting synchronization packets, the output ports transmit normal traffic packets: data packets or idle packets.

The following steps must be taken in order to synchronize input ports, either after reset and initialization of the chip or when the control processor decides to re-synchronize a link due to data errors on the incoming links. Even if the same steps have to be taken on both the switch port and the IBM Packet Routing Switch Serial Interface Converter (the converter) chip, they don't have to be synchronous, but the global sequence of operation must be followed:

- 1. Disable the switch and converter ports.
- 2. Enable the switch and converter ports by writing binary '1' Port Enable register.
- 3. Disable DASL transmission by disabling 'transmit synch enable'.
- 4. Check for valid connectivity of the receiver to a differential transmitter through the No Signal Register. This ensures the integrity of the serial links.
- 5. Enable 'Transmit Sync Packets'.
- 6. Enable DASL transmission by writing a 1 in 'transmit synch enable'.
- 7. Write a 1 into the Sync Hunt register for the enabled ports to start synchronization.
- 8. Poll the Sync Status register to verify completion of synchronization after a Sync Time-Out period. If the port fails to synchronize, its Sync Status bits will be `0'.
- 9. When synchronization has been achieved the local processor at each end reports to the control processor that it is ready for data transfer.
- 10. Clear any CRC error indication that might have been set up during the synchronization period.
- 11. Upon reception of both reports the control processor indicates to both ends of the full duplex link (switch port and converter) to stop transmitting Sync packets. Normal packet transfer (idle or data) on the input ports can then be initiated.
- 12. As the link is now in data mode both the switch control and the adapter controller have to poll the CRC error registers and the No Signal Register to check that the receiver is synchronized and for error free operation.

This operational sequence is mapped in the following flow chart:



IBM3209K3114



#### **Figure 41: Switch Fabric DASL Port Synchronization Sequence**



## **10.7 Line Termination**

## **10.8 DASL and SYS\_CLK**

The DASL data inputs (DASL\_X\_TX, DASL\_Y\_TX) and the clock reference (SWITCH\_X\_CLK and SWITCH Y\_CLK ) use differential HSTL EIA/JEDEC (EIA/JESD8-6) standard compliant I/O books. The following figure gives the recommended termination for those receiver lines.

10.9

#### **Figure 42: DASL Termination**



The termination network must be placed within 2.5 cm of the receiver chip.

**Note:** For switch clock X/Y differential pair, both termination resistors must be directly connected to GND, without the use of the 100 pF capacitor.



## **11. Appendix B: PLL**

## **11.1 PLL Configuration**

During Chip POR, the PLL is held reset and stays in reset mode until a software access releases the PLL RESET bit in the register configuration table. Before starting the PLL (by releasing the PLL RESET bit), the PLL RANGE, PLL MULT and PLL TUNE bits must be written in the PLL register. Once the PLL is started, a poll the PLL register for PLLLOCK bit detection. When the PLLLOCK bit is detected high, the logic clocked by the PLL is ready to work.

### **11.2 PLL Reset**

PLL\_RESET bit [15] is provided in the register configuration table @ 90, 94, 9C. PLL RANGE [bits 5-0] and MULT [bits 31-28] bits are provided in the register configuration table @ 90, 94, 9C. PLL TUNE bits [13-8] are provided in the register configuration table @ 90, 94, 9C. PLLLOCK bit [14] is provided in the register configuration table @ 90, 94, 9C.

The RESET function puts the PLL in bypass mode so the output clock is identical to the input reference clock. RESET should be held active (high) during power-on until all of the following condition are met: All PLL inputs are stable and at their final values. The reference clock is stable. PLL VDDA and Vdd are at their final values. A RESET is also required should the PLL inputs change after power-on.

### **11.3 PLL Range and MULT**

PLL RANGE [bits 5-0] and MULT [bits 31-28] are provided in the register configuration table @ 90, 94, 9C. Use these programmable inputs to choose the output frequency of the PLL. The logic state of these inputs must be stable before the PLL can begin to lock.

### **11.4 PLL Tune**

PLL TUNE bits [13-8] are provided in the register configuration table @ 90, 94, 9C. These programmable inputs are used to optimize the PLL stability and jitter. The logic state of these inputs must be stable before the PLL can begin to lock.

### **11.5 PLL LOCK**

PLLLOCK bit [14] is provided in the register configuration table @ 90, 94, 9C. This signal indicates when the feedback clock is in phase with the reference clock. While RESET is high, PLLLOCK will be low. Following reset, PLLLOCK will stay low until lock is achieved.

The PLLLOCK signal will go high less than 300 us after:

- The reference clock is stabilized at a constant frequency
- The RANGE and MULT inputs are at their final states
- The VDDA input is at the rail

#### **IBM Packet Routing Switch Serial Interface Converter Advance** Advance

PLLLOCK will return to a low state if the PLL loses lock with the reference clock.

## **11.6 PLL Settings**

The following table gives the recommended values for the IBM 28.4G Packet Routing Switch PLL settings:



**Note:** If the PE PLL is used with a 55 MHz reference clock, the same settings are used except for the unused PLL RANGE B bits.



<span id="page-149-0"></span>



# **12. Glossary**

## **12.1 Definitions**



## **12.2 Terms, Abbreviations, and Acronyms**



IBM Packet Routing Switch Serial Interface Converter Advance



# **Revision Log**

<span id="page-151-0"></span>