

IBM Processor for Network Resources Revision 2.5

**Databook**

**Preliminary**



Copyright International Business Machines Corporation 1999, 2000

All Rights Reserved Printed in the United States of America August 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

pnr25.01 August 14, 2000



**IBM Processor for Network Resources** 

# **Contents**





## Preliminary









## Preliminary

















## Preliminary

























## Preliminary









## Preliminary







**IBM Processor for Network Resources Preliminary** 







# **Preliminary IBM Processor for Network Resources**

## <span id="page-20-0"></span>**Features**

- Supports multiple protocols, including ATM, POS, Frame Relay, and 10/100/Gigabit Ethernet
- Has a customizable on-chip 133 MHz PowerPC<sup>™</sup> processor core
- Manages up to 65535 simultaneous logical channels, individually or in groups
- Integrated 155 Mb/s SONET (Synchronous Optical Network) Framer for simpler, low bandwidth designs
- Flexible ATM Forum-compliant UTOPIA II interface with up to four PHYs
- Switch Interface Extensions
- PCI 32/64-bit interface up to 66MHz.
- Configurable for sustained performance through the subsystem:
	- 155Mb/s full duplex internal SONET framer
	- 622Mb/s full duplex using an external SONET framer
	- 622Mb/s across up to four full duplex 155Mb/s links using an external quad framer
- JTAG Test Interface
- Package: 624 lead, 32 mm x 32 mm CBGA
- Power Supply:  $2.6 \text{ V } \pm 2\%$ ;  $3.3 \text{ V } \pm 5\%$ .

## <span id="page-20-1"></span>**Description**

The IBM Processor for Network Resources (IBM3206K0424) is an Asynchronous Transfer Mode (ATM) support device. It is an interface and translator between a Peripheral Component Interconnect (PCI) bus and an ATM Utopia or similar interface to an ATM PHY. The IBM3206K0424 has an integrated Packet/Frame Memory (DRAM controller) and performs Segmentation and Reassembly (SAR) functions for several of the ATM Adaptation Layers (AALs). The IBM3206K0424 functions are illustrated in the [Block Diagram](#page-20-2) on page 21. A Network Interface Card example is shown in [System](#page-33-1)  [Context of an ATM Subsystem](#page-33-1) on page 34.

<span id="page-20-2"></span>**Block Diagram** (See [page 29](#page-28-0) for descriptions of subsystems)





## **Ordering Information**



# <span id="page-21-0"></span>**Conventions**

The bit notation is non-IBM, meaning that bit zero is the least significant bit and bit 31 is the most significant bit for a four-byte word.

The internal addressing view of the IBM3206K0424 registers and memory is big endian. In most cases, a system will wire its PCI bus interface to make the register view transparent, that is, the most significant bit in this specification will be the most significant bit in the register. If registers are read and written 32 bits at a time (which is the only way to access many of the registers), the endian-ness should not be a programming issue with respect to the registers.

The IBM3206K0424 DMA controller can transfer data in either big endian or little endian mode. See General [Purpose DMA \(GPDMA\)](#page-174-2) on page 175 for details.

Numeric notation is as follows:

- Hexadecimal values are usually preceeded by x or X. For example: X'0B00'. For individual registers, Address values are hexadecimal without any special markings. For example, XXXX 1C3C.
- Binary values in text are either spelled out (zero and one) or appear in quotation marks. For example: '10101'.
- Binary values in the Default and Description columns of the register sections are often isolated from text as in this example:
	- 0: No action on read access
	- 1: Auto-reset interrupt request register upon read access



## **Preliminary IBM Processor for Network Resources**

## <span id="page-22-0"></span>**Standards Compliance**

The IBM Processor for Network Resources, part number IBM3206K0424, has been designed with a number of standards in mind. These standards are listed below, grouped according to the area of IBM3206K0424 functionality they address.

- Network (defined by ITU-TS (formerly CCITT), ANSI and ATM Forum)
	- ITU Recommendation I-361 B-ISDN ATM layer specification
	- ITU Recommendation I.362 B-ISDN ATM Adaptation Layer (AAL) functional description
	- ITU Recommendation I.363 B-ISDN ATM Adaptation Layer (AAL) specification
	- ITU Recommendation I.413 B-ISDN user-network interface
	- ITU Recommendation I-432 B-ISDN user-network interface Physical Layer specification
	- ITU Recommendation I-610 OAM principles of B-ISDN access
	- ANSI T1.ATM-199x Draft, Broadband ISDN ATM Layer Functionality and Specification
	- ANSI T1.CBR-199x Draft, Broadband ISDN ATM Adaptation Layer for Constant Bit Rate Service Functionality and Specification
	- ATM Forum 93-620R2 ATM User-Network Interface Specification Version 2.3 (July 27, 1993)
	- Bellcore TA-NWT-001248 Generic Requirements for Operations of Broadband Switching Systems (October 1993)
- System Interface
	- PCI Local Bus Specification, Production Version, Revision 2.1, June 1, 1995. Interface Technical Reference, 11/89, Part number 15F2160
- PHY Interface
	- SATURN User Network Interface, PMC-Sierra, Inc., February 1995
	- ATM Forum 93-727 An ATM PHY Data path interface, Version 2.01, March 24, 1994
	- Am7968/Am7969 TAXIchip(tm) Handbook, Transparent Asynchronous Transmitter/Receiver Interface, published by Advanced Micro Devices, 1994



# <span id="page-23-0"></span>**Environmental Ratings**

# **Absolute Maximum Ratings**



<span id="page-23-1"></span>1. These are the maximum ratings that can be applied to the device without damage. The device function and specifications are valid only within the Recommended Operating Conditions.

## **Recommended Operating Conditions**



## **Power Dissipation**





**Preliminary IBM Processor for Network Resources**

# **Package Diagram**



**IBM Processor for Network Resources Preliminary** 



# **Pinout Viewed from Above**





**Preliminary IBM Processor for Network Resources**

# <span id="page-26-0"></span>**Dataflow**





# <span id="page-27-0"></span>**Functional Description**

The IBM3206K0424 has been designed by breaking the implementation of the various functions and dataflows into separate entities (major functional units).

This processor acts as a conversion unit from a bus memory interface (which is Work Queue oriented) to a PHY level ATM. To accomplish this, the IBM3206K0424 contains the major functional units listed below and shown in the Dataflow [on page 27.](#page-26-0)

Control Processor Bus Interface





## <span id="page-28-0"></span>**Subsystem Blocks**

**The IBM Processor for Network Resources** provides the host bus interfacing, memory management for buffers and control, cell segmentation and reassembly, and PHY hardware control for an ATM adapter.

**External Memory** consists of a number of SRAM modules, or two SDRAM arrays used for the storage of packet data and the control structures used by the IBM3206K0424. Both the Packet and Control Memory arrays consist of two 32-bit wide banks.

When running at 102Mb/s or slower (full duplex aggregate throughput), a single array of memory can be used. Both control and data store are contained in this single array of memory. For a detailed description of the external memory organization refer to [The DRAM Controllers \(COMET/PAKIT\)](#page-183-2) on page 184.

**The PHY (Physical) Layer** interface connects to several available hardware support devices. This layer of hardware converts a parallel data stream into a serial data stream to be shipped to and from the PMD layer.

The PHY and PMD end of a card design can be implemented as one of several encoding schemes and speeds, supporting both copper and fibre optic serial links. The interface will support the ATM Forum "Utopia spec," the PMC chip, and a 25Mb/s serial interface to the IBM UTP solution. (See [Standards Compliance](#page-22-0) on [page 23](#page-22-0) for documents which describe these interfaces.)

**The PMD (Physical Media Dependent) Laye**r interface connects to the line drivers and receivers. This could be either a copper or a fibre optic transceiver.

## **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



## <span id="page-29-0"></span>**External Architecture**

The IBM Processor for Network Resources has four major interfaces:

**A System Bus** which acts as an actively cached memory slave and as a master for the PCI 32-bit bus.

**The Physical (PHY) Interface** which supports several physical layer hardware devices that perform parallel to serial data conversion and the rest of the transmission convergence.

**An External DRAM Interface** that controls one or two arrays of two-bank interleaved DRAM with 60 ns access time for Packet and Control Memory. The interface is direct drive to the DRAM.

**The Control and Configuration Interface** which covers a number of functions. It gives access from the system bus to the PHYs and to EPROM. The EPROM can also be used to hold initial device configuration, up to and including PVC configurations.

**Note:** IBM3206K0424 has built-in, self-test logic.

The four major interfaces allow the IBM Processor for Network Resources to be used in both "deep" and "shallow" adaptors with minimal external logic. (See [Block Diagrams of Possible Systems](#page-32-0) on page 33 for examples.)



## <span id="page-30-0"></span>**Internal Architecture**

## <span id="page-30-1"></span>**Logical Channel Support**

The Logical Channel is the unit of resource allocation in ATM. At one level, the End Station negotiates with the Network Interface to determine the characteristics of each End Station-to-End Station connection. The resources that may be reserved in the network are defined in the ATM UNI (User Network Interface) Specification (see references in [Standards Compliance](#page-22-0) on page 23). These resources include (but are not limited to) the peak and average bandwidth to be used by the logical channel, the maximum burst length that may be transmitted at the burst rate, the latency and variance of the connection, and the loss probability.

The term Logical Channel rather than virtual circuit or VPI/VCI is used in this databook to provide a level of abstraction from these specific instances.

A Switched Virtual Circuit (SVC) can be negotiated with specific characteristics specifically for it.

A virtual path can be negotiated with the network. Several virtual circuits within that path can then be multiplexed, using the VCI on that single VPI, without having to renegotiate for each additional VCI. The Logical channel, with respect to the network, would be the Virtual Path. There would be multiple logical channels internal to the End Station based on the Virtual Circuits used within the path.

Using ATM Adaptation Layers 3 and 4, a Multiplexing IDentifier (MID) can be used to provide multiple Logical Channels across a single VPI/VCI.

All of these Logical Channels are dealt with uniformly in IBM3206K0424. A hierarchy of Logical Channel Descriptors can be built up, and frames or buffers can be queued to each of the LCDs. See Transmit Buffer (CSKED) [on page 273](#page-272-3) for details.

## <span id="page-30-2"></span>**Virtual Memory Support**

The Packet Memory space appears on the bus as a group of up to 128K buffers (configurable size). A level of indirection has been added to the addressing of Packet memory to provide these large frame buffers without requiring memory behind all of them at the same time. This has been done for a number of reasons:

- The frames on the network can be up to 64KB long.
- The receiver does not know how long a frame will be until it is completely received.
- Software generally has a much easier time dealing with contiguous memory.

The memory does not page or swap. There are two major efficiencies used internally:

- The first N bytes of memory in a buffer are directly referenced.
- The blocks that make up the buffers are of multiple sizes.

## **IBM Processor for Network Resources All and Service Preliminary Preliminary**



## <span id="page-31-0"></span>**Queues**

The IBM Processor for Network Resources makes extensive use of cached single memory operation atomic queues:



**Note:** In order to maintain the atomicity of 64-bit atomic transfers, the user must ensure that 64-bit transfers are bus atomic within the particular bus system in which the IBM3206K0424 is being used.

## <span id="page-31-1"></span>**Scheduling**

There is extensive support for transmit scheduling. Please see Transmit Path [on page 37](#page-36-0) and Transmit [Scheduling Capabilities](#page-37-0) on page 38 for details.



**Preliminary IBM Processor for Network Resources**

<span id="page-32-0"></span>



## <span id="page-33-0"></span>**System Environment**

The dataflow context of an ATM subsystem is shown in the diagram below. The purpose of the communications subsystem of any digital device is to allow the application to share data and to adjudicate the flow of control with other devices.

## <span id="page-33-1"></span>**System Context of an ATM Subsystem**



As shown in the figure above, data, in the form of application objects or control structures, are divided into communication frames at the communication stack interface. The stack may further partition the frames to fit reliability, efficiency, latency, and protocol requirements.

In most cases, the communication stack encapsulates the data frame with protocol headers and/or trailers. These header blocks are often located in memory in areas apart from the data frames. A device driver is often given the task of moving this scattered memory to the actual transmission device. Scatter DMA is often used to make this operation efficient.

In the case of the IBM Processor for Network Resources, the data can be DMAed into virtually contiguous buffers connected to and controlled by the IBM3206K0424. It is also possible to write the frame headers directly from the processor to the IBM3206K0424 memory. The fully assembled frame is enqueued for transmission over a particular logical channel. (See more on the richness of logical channels in ATM and the IBM3206K0424 in Data Structures [on page 61\)](#page-60-2).

The logical channels with pending work are serviced by the ATM Segmentation Layer which breaks the enqueued data into 48-byte chunks (depending on the ATM Adaptation Layer (AAL)) and prefixes it with a five-byte header (yielding the prime number 53) in preparation for transmission.

A Transmission Convergence (TC) sublayer appropriate for the Physical Layer (PHY) and Physical Media Dependent (PMD) connection is then exercised, making ATM cells suitable for transmission.

The receiving process is the reverse of the transmission process, except that the scheduling performed during transmission is replaced by an identification-demultiplexing step during the reception of cells.

**Preliminary IBM Processor for Network Resources**

**Note:** Not all of these separate parts or steps described in this section are necessary for a dedicated function system. IBM3206K0424 can easily be used in dedicated systems due to the goal of minimal processor intervention for steady state operations.



# <span id="page-35-0"></span>**Data Flows**

This section describes the data and control flow to and through the IBM3206K0424. In order for cell traffic to flow through an ATM interface, the cells require that Logical Channels be allocated. For information on Logical Channels, please see Data Structures [on page 61.](#page-60-2)

Feature summary

- Virtual memory
- Memory pools
- Register read/write interface for memory allocation
- Transmit path scheduling
- Receive path demultiplexing
- Event queues

Operation summary

- Basic Assurance Tests (BATs)
- Initialize and configure
- Test path to switch
- Permanent Virtual Circuit setup
- Identify LAN servers
- Initialize SVCs
- Run, initializing circuits (Q.93B) and transmitting data


# **Transmit Path**

A typical transmit operation begins with the software requesting a buffer from POOLS and filling it with data via slave DMA, master DMA, or processor writes. If virtual buffers are being used, the data write operation can fail due to lack of physical buffers. In the event of a failure, the header of the packet is updated to indicate the failure. The software can audit the header after the buffer has been completely transferred, and either take action to recover the data immediately or allow CSKED to generate an event later in the transmit cycle for any buffers that have had a data write failure.

Before the data can be transmitted, the buffer header must be updated to contain information required for correct transmission. Information such as data length, starting offset, and Logical Channel (LC) address are just a few of the fields that must be correctly reflected in the buffer header. For a complete list of the fields in the buffer header refer to Packet Header [on page 61.](#page-60-0)

In addition to the fields in the buffer header, the scheduling and segmentation sections of the Logical Channel Descriptor (LCD), such as peak rate, average rate, and AAL type, must also be set up correctly prior to transmission. For a complete list of the fields in the LCD, refer to [Transmit Logical Channel Descriptor Data Struc](#page-65-0)tures [on page 66](#page-65-0).

After the data have been transferred into packet storage and both the buffer header and the LCD structure have been correctly initialized, the buffer address is queued to CSKED. When it receives a buffer, CSKED checks the buffer header (Packet Memory) to make sure that the data transfer operation that filled the buffer completed without error. If it finds an error, CSKED posts an event to software and does nothing further with this buffer. If no error is indicated in the buffer header, CSKED fetches several fields from the LCD (Control Memory) indicated in the buffer header to determine the current state of that LCD. If the LCD is busy sending another buffer, the new buffer is queued to this LCD and will be processed when all previously enqueued buffers have been transmitted. If the LCD is not busy, CSKED updates the LCD based on several fields in the buffer header and queues the LCD to the next time slot on the time wheel (Control Memory).

When CSKED detects a previously enqueued LCD on the time wheel, several fields are retrieved from the LCD. Among other things, these fields are used by CSKED to determine where on the time wheel to reschedule this LCD. The LCD address is then provided to SEGBF for processing.

When CSKED provides an LCD address to SEGBF, the segmentation portion of the LCD is retrieved from Control Memory to determine both the current address at which to continue buffer segmentation and the type of cell to construct. Depending on the AAL type bits in the segmentation portion of the LCD, the cell is constructed in an internal array using data from the LCD as well as data fetched from Packet Memory. When the cell construction is complete, status is raised to LINKC indicating that a new cell is available for transmission.

Transmit opportunities are repeatedly provided to SEGBF by CSKED at the desired rate until all the data in the buffer has been passed to LINKC via the cell buffer array. When SEGBF detects that no more data exists for a buffer, the LCD address is passed back to CSKED, indicating buffer completion. At this point, CSKED removes the LCD from the time wheel if no more buffers are queued in it. If more buffers are queued, the LCD is updated and the segmentation process continues until all buffers on the LCD queue are serviced. A bit in the buffer header generates a transmit complete event when no buffers remain in the queue.



# **Transmit Scheduling Capabilities**





# **Receive Path**

As cells arrive, they pass from LINKC to REASM. REASM uses a portion of the ATM header to look up the LCD address for this cell. The LCD address is then passed to RAALL. RAALL reads the receive portion of the LCD, and then processes the cell based on the LCD information. For example, the LCD specifies what AAL to use and maintains the current reassembly state. Using the current reassembly state, the cell data is written to Packet Memory. While the data is written to Packet Memory, other functions such as CRC generation and verification are performed in parallel. If a packet is complete, all trailer verification is performed. If the packet is good, an event is placed on a receive queue in the RXQUE entity. For error scenarios, see Receive [Queues \(RXQUE\)](#page-365-0) on page 366. At this point, software can dequeue the packet event from RXQUE using the dequeue operation. It can then examine headers, DMA the data into user space, and perform TCP checksums. When these actions are complete, the buffer is returned to the IBM3206K0424 by performing a POOLS-free buffer operation.

**IBM Processor for Network Resources Accessor For Accessor Preliminary** 





# **Input/Output Definitions**

The several interfaces to IBM3206K0424 are described in the following sections. There are 443 active I/O pins, assigned as follows:

- 90 for the PCI bus
- 86 for the NPBUS
- 76 for the PHY bus
- 156 for the Packet Memory interface
- 35 strictly for configuration and testing

# **DRAM Memory Bus Interface**

## **PCI Bus Connections**





## **PCI Bus Interface Pin Descriptions**



<span id="page-41-0"></span>1. S/T/S = a sustained tri-state pin owned and driven by one and only one agent at a time. The agent that drives the S/T/S pin low must drive it high for at least one clock before letting it float. A new agent cannot start driving a S/T/S signal any sooner that one clock after the previous owner tri-states it. A pullup is required to sustain the inactive state until another agent drives it, and must be provided by the central resource.



### **PCI Bus Interface Pin Descriptions** (Continued)



provided by the central resource.



# **DRAM Memory Bus Interface**

One Control Memory and one Packet Memory bus provide the attachment to the external DRAM. Up to two arrays of 32 data bits plus potential error detection bits may be connected to each bus.

#### **DRAM Memory Bus Connections**





# **DRAM Memory Bus Interface Pin Descriptions**







# **Memory I/O Cross Reference By Device Type**

<span id="page-45-0"></span>1. All signal groups marked by an asterisk are active at the same time.

2. xx = CM for Control Memory or PM for Packet Memory.

3. For SDRAMs, the DQM signals are active independently for shared ECC configurations.

4. For SDRAMs with split ECC, the DQMs are usually active unless doing burst length two and the DQM is needed to terminate a burst.





#### **Possible Memory Configurations Using SDRAM With Shared ECC**

**Note:** While it is possible to connect more than five SDRAM modules to each controller on the IBM3206K0424, it is likely the capacitive loading will not allow the interface to work at 7.5ns. The memory interface would need to be slowed down to allow the interface to work.

#### **Possible Memory Configurations Using SRAM**



1. For x18 SRAM modules, half the data bus goes to one module, and the other half goes to a second module. The chip select to the two modules is common. Therefore, two x18 modules can be connected to a single chip select while only one x32 module can. Therefore, given a constant number of chip selects, using pairs of x18 "x" Mb modules results in a memory that is twice as deep as what is possible with x32 modules. Using x18 modules also lowers the overall capacitance on the memory data nets.

2. While it is possible with the number of chip selects available (multiplexed or not) to connect more than four SRAM modules to each controller on the IBM3206K0424, it is likely the capacitive loading will not allow the interface to work at 7.5ns. The memory interface would need to be slowed down to allow the interface to work.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

# **NPBUS**

The NPBUS supports access to either an EPROM or PHY level hardware microprocessor interface. The NPBUS can operate with an eight-bit multiplexed addr/data bus or an eight-bit data bus with 18 separate address pins. Generic transfer control signals work with the PHY level hardware microprocessor interface or EPROM to accomplish data transfers.

## **NPBUS Connections**







# **NPBUS Pin Descriptions**



<span id="page-48-0"></span>1. S/T/S = a sustained tri-state pin owned and driven by one and only one agent at a time. The agent that drives the S/T/S pin low must drive it high for at least one clock before letting it float. A new agent cannot start driving a S/T/S signal any sooner that one clock after the previous owner tri-states it. A pullup is required to sustain the inactive state until another agent drives it and must be provided by the central resource.



#### **NPBUS Pin Descriptions** (Continued)



must drive it high for at least one clock before letting it float. A new agent cannot start driving a S/T/S signal any sooner that one clock after the previous owner tri-states it. A pullup is required to sustain the inactive state until another agent drives it and must be provided by the central resource.



# **ATM PHY Bus Interface**

The PHY Bus consists of a transmit data path, receive data paths, and control signals.

## **PHY Bus Interface Connections**





## **PHY Bus Pin Descriptions** (Page 1 of 3)



**Note:** Because some of the PHY transmit I/Os are used for receive framer functions and vice versa, there are some restrictions on how the interfaces can be used.

- 1. If the transmit path is using an external PHY and the receive path is using the internal framer, FYTPAR(1) will assume the OOF function and not be available as a parity output. This is only a concern if the PHY uses a 16-bit data interface and parity is being used.
- 2. If the receive path is using an external PHY and the transmit path is using the internal framer, FYRPAR(1) will assume the OFPtx-LPow function and not be available as a parity input. This is only a concern if the PHY uses a 16-bit data interface.
- 3. If the transmit path is using an external PHY and the receive path is using the internal framer, and the external PHY has a 16-bit data interface, then the receive HDLC interface cannot be used. The three I/O for the RX HDLC interface will instead take on the function of FYTDAT(15-13).



# **PHY Bus Pin Descriptions** (Page 2 of 3)



**Note:** Because some of the PHY transmit I/Os are used for receive framer functions and vice versa, there are some restrictions on how the interfaces can be used.

- 1. If the transmit path is using an external PHY and the receive path is using the internal framer, FYTPAR(1) will assume the OOF function and not be available as a parity output. This is only a concern if the PHY uses a 16-bit data interface and parity is being used.
- 2. If the receive path is using an external PHY and the transmit path is using the internal framer, FYRPAR(1) will assume the OFPtx-LPow function and not be available as a parity input. This is only a concern if the PHY uses a 16-bit data interface.
- 3. If the transmit path is using an external PHY and the receive path is using the internal framer, and the external PHY has a 16-bit data interface, then the receive HDLC interface cannot be used. The three I/O for the RX HDLC interface will instead take on the function of FYTDAT(15-13).



# **PHY Bus Pin Descriptions** (Page 3 of 3)



**Note:** Because some of the PHY transmit I/Os are used for receive framer functions and vice versa, there are some restrictions on how the interfaces can be used.

1. If the transmit path is using an external PHY and the receive path is using the internal framer, FYTPAR(1) will assume the OOF function and not be available as a parity output. This is only a concern if the PHY uses a 16-bit data interface and parity is being used.

2. If the receive path is using an external PHY and the transmit path is using the internal framer, FYRPAR(1) will assume the OFPtx-LPow function and not be available as a parity input. This is only a concern if the PHY uses a 16-bit data interface.

3. If the transmit path is using an external PHY and the receive path is using the internal framer, and the external PHY has a 16-bit data interface, then the receive HDLC interface cannot be used. The three I/O for the RX HDLC interface will instead take on the function of FYTDAT(15-13).



# **Transmit PHY I/O Cross Reference**



**Note:** Signals marked with an overbar are active low. Inputs listed as N/A should be tied to their inactive Utopia state.

#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



### **Receive PHY I/O Cross Reference**



**Note:** Signals marked with an overbar are active low. Inputs listed as N/A should be tied to their inactive Utopia state.





## **Clock, Configuration, and LSSD Connections**

#### **IBM Processor for Network Resources Preliminary**



# **Clock, Configuration, and LSSD Pin Descriptions**





# **Clock, Configuration, and LSSD Pin Descriptions** (Continued)



**IBM Processor for Network Resources Preliminary** 





<span id="page-60-0"></span>These structures reside in Control Memory for each of the logical channels that are set up for transmission or reception.

# **Packet Header**

Each packet buffer consists of two parts. The first part is the control information used by the IBM3206K0424. The second portion of the packet buffer is used to hold the actual packet data. The following figures show the structure of the transmit and receive packet headers:

# **Transmit Packet Header Structure**

```
struct tx_min_packhead { struct tx_packhead { bit32 next buffer;
                                                 bit32 next_buffer;<br>bit8 AAL5_user_byte;bit8 AAL5_user_byte1; bit8 AAL5_user_byte<br>bit8 buffer_offset; bit8 buffer_offset;
       bit8 buffer_offset; bit8 buffer_offset;<br>bit16 buffer<sup>-</sup>length; bit16 buffer<sup>-length</sup>;
       bit16 buffer\overline{\phantom{a}}length;
       bit25 lc_address; bit25 lc_address;<br>bit1 EFCI_status; bit1 EFCI_status;<br>bit1 reserved; bit1 reserved;
        bit1 EFCI_status; bit1 EFCI_status;
        bit1 reserved; bit1 reserved;
       bit1 dma_on_xmit;<br>bit1 dma_on_xmit; bit1 dma_on_xmit;<br>bit1 generate_CRC10; bit1 generate_CRC10;
       bit1 generate_CRC10;
        bit1 free_on_xmit; bit1 free_on_xmit;
        bit1 queue_on_xmit; bit1 queue_on_xmit;
                                                bit1 free_on_xmit;<br>bit1 queue_on_xmit;<br>bit1 cell_loss_priority;
       \}; \qquad \qquad \qquad \qquad \qquad \qquad bit32 dma_desc_addr;
                                                 bit16 AAL5_user_byte1_2;
                                                 bit16 reserved
                                                 };
```
The minimum transmit packet header size (and transmit offset) is 0xC bytes.

# **Receive Packet Header Structure**

```
struct rx_packhead {
     bit32 rx_label_flags;
     bit8 AAL5 user byte1;
     bit8 buffer offset;
     bit16 buffer_length;
     bit25 lc_address;
     bit1 EFCI status;
     bit5 reserved;
     bit1 cell loss priority;
     bit32 optional_words[7];
     };
```
See [RXAAL Packet Header Configuration](#page-351-0) on page 352 for available word choices and definitions.



#### **Receive Packet Definitions**



The minimum receive packet offset is 0xC bytes. When the optional fields are enabled, the receive packet off-set increases and should be set appropriately in the receive LCD. See [RXAAL Packet Header Configuration](#page-351-0) [on page 352](#page-351-0) for available word choices and definitions.

## **Transmit and Receive Packet Header Field Descriptions** (Page 1 of 3)





# **Transmit and Receive Packet Header Field Descriptions** (Page 2 of 3)



#### **IBM Processor for Network Resources Preliminary**



# **Transmit and Receive Packet Header Field Descriptions** (Page 3 of 3)





#### IBM3206K0424

### <span id="page-64-0"></span>**Logical Channel Data Structure**





### **General LCD Layout**

```
struct lcd_struct {
  tx_lcd_struct tx_lcd ;/* transmit portion of the lc descriptor */
  bit8 fill[N] ;/* the fill area depends on the type of tx lcd */
 rx lcd struct rx lcd \frac{1}{2} receive portion of the lc descriptor */
} ;
```
## <span id="page-65-0"></span>**Transmit Logical Channel Descriptor Data Structures**

[Logical Channel Data Structure](#page-64-0) on page 65 and [Scheduling Portion of a Transmit Descriptor](#page-66-0) on page 67 show the layout of the transmit portion of a Logical Channel Descriptor. When initializing an LCD, any locations that are not written to a specific value should be initialized to zeros. Fields that typically need to be initialized to a non-zero value are flagged with a # in the structure below.

**Note:** This is only one possible layout of the transmit portion of the LCD. Some field locations vary and are further defined later in this section.

Care must be taken when updating fields in the LCD and then immediately causing the updated fields to be accessed by other IBM3206K0424 entities. For example, it is possible, although not likely, under the right conditions, for a normal LCD update followed by a SEGBF cell enqueue operation to actually execute in reverse order. This is due to IBM3206K0424 internal priority levels and could result in SEGBF fetching the LCD data before it has been updated to the new value. For this reason, it is highly advisable to use the LCD update mechanism in [Cell/Packet Re-assembly \(REASM\)](#page-319-0).

The transmit portion of the LCD can be subdivided into three distinct parts based on which chip functions or entities access that particular part of the LCD. The first three 64-bit words are scheduling related and are accessed only by CSKED. The next two 64-bit words (four if using switch fabric header) are related to both scheduling and segmentation and are accessed and shared by both CSKED and SEGBF. The words following these shared locations are related only to segmentation and are accessed only by SEGBF. The number of 64-bit words in this portion of the LCD can vary from one to three. The actual number being used in an LCD is determined by the segmentation processor code entry point field of that LCD. The following figure is the layout of the entire transmit portion of the LCD.

#### **Overall Transmit LCD Layout**



The three 64-bit words containing CSKED scheduling information can have four different formats depending upon whether scheduling is based on virtual path parameters or is configured for ABR.

If LCD-based memory management is used, a 64-bit section is inserted at byte offset '18'X.



On ABR connections, the ABR code running in PCORE will use ten bytes of the LCD for managing the connection, starting at offset 0x46 in the LCD.

### <span id="page-66-0"></span>**Scheduling Portion of a Transmit Descriptor**



.



#### **Transmit Logical Channel Descriptor Structure**

typedef struct { bit32 next\_lcd; bit16 #peak\_interval; bit16 #average\_interval; bit32 #timestamp; bit11 Reserved; bit1 remove\_lcd; bit1 lc\_on\_timewheel; bit3 #alter\_sched; bit2 #transmit\_priority; bit1 #max\_resolution; bit3 #max\_burst\_mult; bit10 #max\_burst\_value; bit26 head\_packet\_pointer; bit1 free\_on\_xmit; bit1 queue\_on\_xmit; bit1 conn\_suss; bit1 #drop; bit26 tail\_packet\_pointer; bit6 reserved; bit16 transmit length; bit8 buffer offset; bit2 xmt cmp evt mod; bit2 reserved; bit4 seg\_prc\_Entry\_point; bit32 #ATM\_header; bit32 segmentation pointer; bit32 current CRC; bit32 xmit stat1; bit32 xmit stat2; } tx\_lcd\_struct, \*tx\_lcd\_struct\_ptr;



.

**Preliminary IBM Processor for Network Resources**

### **Transmit Logical Path Descriptor Structure**

```
typedef struct {
    bit32 next_lcd;
    bit16 #peak_interval;
    bit16 #average_interval;
   bit32 #timestamp;
   bit11 Reserved;
    bit1 remove_lcd;
    bit1 lc_on_timewheel;
    bit3 #alter_sched;
    bit2 #transmit_priority;
    bit1 #max_resolution;
    bit3 #max_burst_mult;
    bit10 #max_burst_value;
   bit26 forward LCD pointer;
   bit6 reserved;
   bit26 backwared LCD pointer;
   bit6 reserved;
```
} tx\_lpd\_struct, \*tx\_lpd\_struct\_ptr;

typedef struct {



#### **Redefinition of Transmit Logical Channel Descriptor for Connections Sharing** (Logical Path Bandwidth)

bit32 next lcd; bit32 reserved; bit32 lcd\_pointer; bit11 reserved; bit1 remove\_lcd; bit1 lc\_on\_timewheel; bit3 #alter\_sched; bit2 #transmit\_priority; bit14 reserved; bit26 head\_packet\_pointer; bit1 free\_on\_transmit; bit1 queue\_on\_transmit; bit1 dma\_on\_transmit; bit1 conn\_suss; bit2 #drop; bit26 tail\_packet\_pointer; bit6 reserved; bit16 transmit length; bit8 buffer offset; bit2 xmt\_cmp\_evt\_mod; bit2 reserved; bit4 seg\_prc\_Entry\_point; bit32 #ATM header; bit32 segmentation\_pointer; bit32 current\_CRC; bit32 xmit\_stat1; bit32 xmit\_stat2;

} tx\_lcd\_struct, \*tx\_lcd\_struct\_ptr;



#### **Redefinition of Shared and Segmentation Portion of Transmit LCD for ABR**

```
typedef struct {
   bit16 transmit length;
   bit8 buffer offset;
    bit2 xmt_cmp_evt_mod;
    bit2 reserved;
    bit4 seg_prc_Entry_point;
   bit32 #ATM_header;
   bit32 segmentation pointer;
   bit32 current_CRC;
   bit32 xmit_stat1;
   bit32 xmit_stat2;
    bit16 reserved;
    bit16 explicit_rate;
    bit16 current_rate;
    bit16 minimum_rate;
   bit32 backward_ptr;
   bit32 reserved;
} tx_lcd_struct, *tx_lcd_struct_ptr;
```
Owing to the use of certain LCD fields, a connection running ABR cannot be set up for segmentation AAL types 0x6 (fixed-sized blocking) or 0x7 (MPEG-2 assist).



**Redefinition of Segmentation Portion of Transmit LCD for Fixed Size AAL5 Blocking** (segmentation type 0x6)

```
typedef struct {
    bit16 transmit length:
    bit8 buffer offset;
    bit2 xmt_cmp_evt_mod;
    bit2 reserved;
    bit4 seg_prc_Entry_point;
    bit32 #ATM header;
    bit32 segmentation pointer;
    bit32 current CRC;
    bit32 xmit stat1;
    bit32 xmit<sup>-</sup>stat2;
    bit8 #Packets per AAL5 frame
    bit8 #Blocking size (4^-bytes x'2F' for MPEG-2)
    bit8 Current_transport_stream_packet
    bit8 Current_Blocking_Count (4 bytes)
    bits2 reserved;
} tx_lcd_struct, *tx_lcd_struct_ptr;
```
**Redefinition of Segmentation Portion of Transmit LCD for MPEG2** (segmentation type 0x7)

```
typedef struct {
   bit16 transmit length:
   bit8 buffer offset;
    bit2 xmt_cmp_evt_mod;
    bit2 reserved;
    bit4 seg_prc_Entry_point;
   bit32 #ATM header;
   bit32 segmentation pointer;
   bit32 current_CRC;
   bit32 xmit_stat1
   bit32 xmit_stat2
   bit8 #Packets_per_AAL5_frame
   bit8 #Blocking size (4 bytes x'2F' for MPEG-2)
    bit8 Current_transport_stream_packet
    bit8 Current_Blocking_Count (4 bytes)
   bit32 reserved;
   bit32 reserved;
   bit32 reserved;
} tx lcd struct, *tx lcd struct ptr;
```


**Preliminary IBM Processor for Network Resources**



## **Redefinition of Scheduling Portion of Transmit LCD for ABR**

typedef struct { bit32 next lcd; bit16 #peak interval; bit3 #Nrm; bit3 #Trm; bit10 #Tadtf; bit8 #Nc; bit8 #Ncrm; bit16 Reserved; bit11 Tlrm1; bit1 remove\_lcd; bit1 lc\_on\_timewheel; bit3 #alter\_sched; bit2 #transmit\_priority; bit1 #max\_resolution; bit13  $T1rm\overline{2}$ ; bit26 head\_packet\_pointer; bit1 free\_on\_xmit; bit1 queue\_on\_xmit; bit4  $reserved;$ bit26 tail packet pointer; bit6 reserved;

### **Redefinition of Scheduling Portion of Transmit LCD for Timers**

```
typedef struct {
   bit32 next_lcd;
   bit32 #timer period;
   bit32 #timestamp;
   bit12 reserved;
   bit1 lc_on_timewheel;
    bit1 reserved;
    bit2 #timer_type;
    bit2 #transmit_priority;
    bit1 #max_resolution;
   bit13 reserved;
   bit32 #dma_desc_addr;
   bit32 reserved;
} tx lcd struct, *tx lcd struct ptr;
```
} tx lcd struct, \*tx lcd struct ptr;

# **Definition of LCD-Based Memory Management of Transmit LCD**

```
typedef struct {
    bit16 #threshold_1;
    bit16 #threshold<sup>-2</sup>;
    bit4 #pool id1;
    bit4 #pool<sup>-</sup>id2;
    bit24 #bytes_queued;
} tx_lcd_struct, *tx_lcd_struct_ptr;
```


# **Definition of ABR Code Variables**

```
typedef struct {
   bit8 #CRM;
    bit8 #iCDF;
    bit16 #iMCR;
   bit16 #PCR;
    bit8 #iRDF;
    bit8 #iRIF;
    bit16 #ICR;
} tx_lcd_struct, *tx_lcd_struct_ptr;
```
# **Field Definitions**

The following is a detailed description of the fields listed above. This data structure should be initialized at connection setup but not modified while transmission is occurring on the connection. Only those fields marked with a # typically need to be initialized to something other than zero.

#### **ABR Code Variables Definitions** (Page 1 of 4)





# **ABR Code Variables Definitions** (Page 2 of 4)



### **IBM Processor for Network Resources Preliminary**



# **ABR Code Variables Definitions** (Page 3 of 4)





# **ABR Code Variables Definitions** (Page 4 of 4)





# **Transmit Data Structure Linkage**



### **Receive LCD Data Structure and Modes**

The format of the receive LCD structure depends on which AAL is being configured and which options are used. It is also dependent on if TCP/IP checksum verification has been enabled. When TCP/IP checksum verification is enabled, 16 additional bytes are added to the LCD format. TCP/IP checksum is enabled in the REASM Mode Register. The following are the basic layouts of the receive LCD:



#### **Preliminary IBM Processor for Network Resources**

# **Basic Receive LCD Layout**



The basic layout is the same for all LCD types. Only the packed Info and misc fields vary between the different LCD types. The following sections detail the receive LCD and the differences from the basic layout for each major option.



# **Raw LCD**

#### **Raw LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {
  bit4 aalType; <br>bit2 ppMode: // 00 - normal
  bit2 ppMode;<br>bit2 state;
                                            \frac{1}{2} 00->down 01->idle/enabled
  bit1 reserved;<br>bit1 reserved;
                                           // set to zero<br>// set to zero<br>// set to zero
  bit1 reserved;<br>bit1 size;
                                           \frac{1}{1} 1->52 byte cell 0->48 byte cell
  bit1 storeCrc10;
  bit3 reserved;
  bit4 rxqNum;
  bit4 rxPoolId;
  bit8 rxOffset;
};
struct Misc {
  bit6 reserved;
  bit2 packHeadSel;
  bit8 reserved;
  bit16 oamMask;
};
struct Misc2 {
  bit32 reserved;
};
```
A raw LCD allows raw ATM cells to be received with no reassembly. The user can select to receive 52- or 48-byte cells. The packet header may or may not contain the ATM header. The cell data is then placed after the packet header at the configured receive offset. The 52-byte mode stores the entire cell minus the HEC, and the 48-byte mode stores only the ATM cell payload. Optional CRC-10 checking is available in raw modes.



#### IBM3206K0424

# **Raw Routed LCD**

## **Raw Routed LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {
 bit4 aalType; <br>bit2 ppMode: // 01 - routed
 bit2 ppMode;<br>bit2 state;
                                   \frac{1}{2} 00->down 01->idle/enabled
  bit1 reserved; // set to zero
  bit1 reserved; // set to zero
 bit1 reserved;<br>
bit1 size;<br>
// 1->52 byte
                                   \frac{1}{1} 1->52 byte cell 0->48 byte cell
  bit1 reserved;
  bit3 reserved;
 bit4 rxqNum;
 bit4 rxPoolId;
 bit8 rxOffset;
};
struct Misc {
 bit6 reserved;
  bit2 packHeadSel;
  bit8 reserved;
 bit16 oamMask;
\}:
struct Misc2 {
 bit32 routedLcd;
};
```
A raw routed LCD receives data in the same way that a raw LCD does. Once received, the cell buffer is routed internally to the scheduler and rescheduled for transmission. Normally, when a cell is received, the receive LCD address is written into the packet header and the buffer is surfaced to the user. When a cell is routed, the routedLcd field is used to fill in the LCD address in the packet header. This allows cells to be routed out the transmit interface with the same or different VPI/VCI.

The low order bits in the routedLcd field should be set correctly to free the buffer on transmission. These bits correspond to the flag bits in the packet header. Raw routing is also called forwarded or fast forward mode.



### **Raw Routed Early Drop LCD**

#### **Raw Routed Early Drop LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {<br>bit4 aalType;
  bit4 aalType; <br>
bit2 ppMode; <br>
bit2 state; <br>
bit2 stat
                                                 \frac{1}{2} 01 - routed
                                                // 00->down 01->idle/enabled 11->error
   bit1 reserved; // set to zero
   bit1 reserved; // set to zero
   bit1 reserved; \frac{1}{2} // set to zero<br>bit1 size; \frac{1}{2} // 1->52 byte
                                                \frac{1}{1} 1->52 byte cell 0->48 byte cell<br>\frac{1}{1} pool id
   bit4 finalPoolId;
   bit4 rxqNum;
   bit4 rxPoolId;
   bit8 rxOffset;
};
struct Misc {
  bit6 reserved;
   bit2 packHeadSel;
  bit8 reserved;
  bit16 oamMask;
};
struct Misc2 {
  bit32 routedLcd;
};
```
A raw routed early drop LCD receives data in the same way that a raw LCD does. Once received, the cell buffer is then routed internally to the scheduler and rescheduled for transmission. Normally when a cell is received, the receive LCD address is written into the packet header and the buffer is surfaced to the user. When a cell is routed, the routedLcd field is used to fill in the LCD address in the packet header. This allows cells to be routed out the transmit interface with the same or different VP/VC.

This mode should only be used when the routed cell stream is actually an AAL5 packet stream. In this mode, a cell being dropped due to resource causes the LCD to go into error mode until the cell that contains the user indicate (UIND) bit is received. All cells received in error mode are dropped, except the final cell which is forwarded. This conserves bandwidth while maintaining the AAL5 integrity. The finalPoolId provides a second poolid for the final cells to use to be sure that these final cells are always forwarded even when resources are low. The finalPoolId is only used if no buffers are available in the normal pool.

The low order bits in the routedLcd field should be set correctly to free the buffer on transmit. These bits correspond to the flag bits in the packet header.

This is also called forwarded or fast forward mode.



# **Raw Scatter/Cut-Through LCD**

### **Raw Scatter/Cut-Through LCD Packed and Miscellaneous Field Layouts**



A raw scatter/cut-through LCD receives data in the same way that a raw LCD does. Once received, the cut-ThruSel field is used to select one of four configurations. Each configuration specifies a receive queue and a DMA queue. The cut-through selector is used to select a cut-through/scatter configuration. The DMA descriptor is then built using the cell buffer address and the data length and the flags specified in the cut-through configuration. After being built, it is enqueued to the DMA queue specified. If there is no DMA descriptor available, then a no descriptor event is enqueued.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



# **AAL5 LCD**

#### **AAL5 LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {
 ruct Packed {<br>
bit4 aalType; // 0101 - aal5<br>
bit2 state; // 00 - norm<br>
bit2 state; // 00->down 01
                                  \frac{1}{2} 00 - normal
                                  bit2 state; // 00->down 01->idle/enabled 10->reasm 11->error
  bit1 reserved; // set to zero
  bit1 rtoTest; // set to zero
  bit1 rtoEnable;
  bit1 tmpCLP;
  bit1 tmpCongestion;
  bit3 reserved;
  bit4 rxqNum;
 bit4 rxPoolId;
 bit8 rxOffset;
};
struct Misc {
  bit6 reserved;
  bit2 packHeadSel;
  bit8 reserved;
  bit16 oamMask;
};
struct Misc2 {
  bit32 reserved;
};
```
An AAL5 LCD allows AAL5 packets to be received with no special processing.



# **AAL5 Routed LCD**

# **AAL5 Routed LCD Layout**

```
struct Packed {
 bit4 aalType; <br>bit2 ppMode: // 01 - routed
 bit2 ppMode;<br>bit2 state;
                            \frac{1}{2} 00->down 01->idle/enabled 10->reasm 11->error
  bit1 reserved; // set to zero
  bit1 rtoTest; // set to zero
 bit1 rtoEnable;
 bit1 tmpCLP;
  bit1 tmpCongestion;
  bit3 reserved;
 bit4 rxqNum;
 bit4 rxPoolId;
 bit8 rxOffset;
};
struct Misc {
 bit6 reserved;
  bit2 packHeadSel;
  bit8 reserved;
 bit16 oamMask;
\}:
struct Misc2 {
 bit32 routedLcd;
};
```
An AAL5 Routed LCD allows AAL5 packets to be received. Once received, the packet buffer is then routed internally to the scheduler and rescheduled for transmission. Normally when a packet is received, the receive LCD address is written into the packet header and the buffer is surfaced to the user. When a packet is routed, the routedLcd field is used to fill in the LCD address in the packet header. This allows packets to be routed out the transmit interface with the same or different VP/VC.

The low order bits in the routedLcd field should be set correctly to free the buffer on transmission. These bits correspond to the flag bits in the packet header.

AAL5 routing is also called forwarded or fast forward mode.

**Note:** Non-user data cells are terminated.

**IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary** 



# **AAL5 Cut-Through/Scatter Mode LCD**

#### **AAL5 Cut-Through/Scatter Mode LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {<br>bit4 aalType;
     bit4 aalType; // 0101 - aal 5<br>bit2 ppMode; // 10 - scattbit2 ppMode; // 10 - scatter<br>bit2 state; // 00->down 01->io
                                 bit2 state; // 00->down 01->idle/enabled 10->reasm 11->error
     bit1 reserved; // set to zero<br>bit1 rtoTest; // set to zero
     bit1 rtoTest; // set to zero<br>bit1 rtoEnable;
     bit1 rtoEnable;<br>bit1 tmpCLP;
     bit1 tmpCLP;<br>bit1 tmpConqo
     bit1 tmpCongestion;<br>bit1 reserved;
              reserved;
     bit2 cutThruSel;
     bit4 rxqNum;<br>bit4 rxPoolI
    bit4 rxPoolId;<br>bit8 rxOffset;
              rxOffset;
   };
   struct Misc {
      bit6 numDesc;
 bit2 packHeadSel;
 bit8 reserved;
      bit16 oamMask;
   };
   struct Misc2 {
 bit5 reserved;
 bit1 useCrcNumHead;
 bit10 numHeadBytes;
      bit16 reserved;
   };
```


# **Packet LCD**

#### **Packet LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {<br>bit4 aalType;<br>bit2 ppMode;<br>bit2 state;
  bit4 aalType; // 0111 - packet
  bit2 ppMode; // 00 - normal
                            bit2 state; // 00->down 01->idle/enabled 10->reasm 11->error
  bit1 reserved; // set to zero
  bit1 rtoTest; // set to zero
  bit1 rtoEnable;
  bit5 reserved;
  bit4 rxqNum;
  bit4 rxPoolId;
  bit8 rxOffset;
};
struct Misc {
  bit6 reserved;
  bit2 packHeadSel;
  bit8 reserved;
 bit16 reserved;
};
struct Misc2 {
 bit5 dropNBytes;
 bit11 reserved;
 bit16 reserved;
};
```
A packet LCD allows packets from the POS-PHY to be received with no special processing. The header-Thresh can be used to allow packet header thresholding events to be surfaced.



# **Packet Routed LCD**

#### **Packet Routed LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {<br>bit4 aalType;
  bit4 aalType; <br>
bit2 ppMode; <br>
bit2 state; <br>
bit2 stat
                                                \frac{1}{2} 01 - routed
                                               \frac{1}{2} 00->down 01->idle/enabled 10->reasm 11->error
   bit1 reserved; // set to zero
   bit1 rtoTest; // set to zero
   bit1 rtoEnable;
   bit5 reserved;
   bit4 rxqNum;
   bit4 rxPoolId;
   bit8 rxOffset;
};
struct Misc {
  bit6 reserved;
   bit2 packHeadSel;
   bit8 reserved;
  bit16 reserved;
};
struct Misc2 {
  bit32 routedLcd;
};
```
A packet routed LCD allows packets to be received from the POS-PHY. Once received, the packet buffer is then routed internally to the scheduler and rescheduled for transmission. Normally, when a packet is received, the receive LCD address is written into the packet header and the buffer is surfaced to the user. When a packet is routed, the routedLcd field is used to fill in the LCD address in the packet header. This allows packets to be routed out the transmit interface with the same or different LCD.

The low order bits in the routedLcd field should be set correctly to free the buffer on transmit. These bits correspond to the flag bits in the packet header.

This is also called forwarded or fast forward mode.



# **Packet Cut-Through Scatter Mode LCD**

#### **Packet Scatter Mode LCD Packed and Miscellaneous Field Layouts**

```
struct Packed {
 bit4 aalType;<br>
bit4 aalType;<br>
bit2 ppMode;<br>
bit2 state;<br>
// 00->down 01
  bit2 ppMode; // 10 - scatter
  bit2 state; // 00->down 01->idle/enabled 10->reasm 11->error
  bit1 reserved; // set to zero
  bit1 rtoTest; // set to zero
  bit1 rtoEnable;
  bit3 reserved;
 bit2 cutThruSel;
  bit4 rxqNum;
  bit4 rxPoolId;
 bit8 rxOffset;
};
struct Misc {
  bit6 numDesc;
  bit2 packHeadSel;
  bit8 reserved;
 bit16 reserved;
};
struct Misc2 {
 bit5 dropNBytes;
  bit1 useCrcNumHead;
  bit10 numHeadBytes;
 bit16 reserved;
};
```


# **Field Definitions**

The following are the definitions of the LCD fields, grouped by major function. All reserved fields should be set to zero.

# **Common Field Definitions**

<span id="page-89-0"></span>



# **Raw Mode Field Definitions**

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

# **IBM Processor for Network Resources Preliminary**



# **Packet/AAL5 Field Definitions**

<span id="page-91-0"></span>



# **Internal Organization: Entity Descriptions**

This part contains detailed descriptions of the entities which, working together, make up the IBM3206K0424. The data flows through the chip have already been described; now the details of the registers and algorithms will be revealed. The entity descriptions are numbered for easy reference.

# **Note on Set/Clear Type Registers**

There are many registers in IBM3206K0424 that operate as a set/clear type. These registers have two addresses. The base address is for clearing bits in the register, and base address +4 bytes is for setting bits in the register. The setting or clearing operations occur only for those bits that have the value of'1' on the write of the register. Either of the addresses can be used for reading the register.

# **Control Processor Bus Interface Entities**

# **Entity 1: The IOP Bus Specific Interface Controller (PCINT)**

This entity provides PCI specific interfacing between the external connection and the internal entities. It will support the following functions:

- PCI memory target
- PCI master
- Address and data latching
- Provide parity error detection and generation
- Provide configuration space registers
- 64-bit data path for master and slave operation
- 64-bit addressing support for master and slave operation
- Auto 64-bit slot detection supported
- 66MHz PCI bus clock operation supported

# **PCI Options Taken**

- Medium address decode design point
- Locking as a memory target supported
- Interrupt A will be supported, with interrupt 2 as a the sideband signal
- Registers will not burst, but cause retries when a burst is attempted
- BIST defaults set at the PCI 2 second maximum

# **PCI Target Response**

- A Target Retry is issued if a burst crosses the end of the IBM3206K0424's memory space.
- A Target Abort will be issued if AD and command bus have bad parity (address phase parity error). Optionally, if SERR is enabled, it will also be returned.
- If enabled, the PERR signal will be driven on bad parity during data write cycles (data phase parity error) when the IBM3206K0424 is the target of the command.
- A Target Retry will be issued by the IBM3206K0424 if internal contention will cause a large bus access delay.

# **PCI Master Response**

- A Master Abort will be issued if DEVSEL is not asserted after five clocks.
- If enabled, the PERR signal will be driven on bad parity during data read cycles (data phase parity error) when the IBM3206K0424 is the initiator of the command.

#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **PCI Master Retry**

• IBM3206K0424 will retry when requested by the slave.

# **1.1: PCINT Config Word 0**

Identifies this device and vendor type, allocated by PCI SIG.









# **1.2: PCINT Config Word 1**

The Status register is used to record status information for the PCI bus related events. Writing '1' to a bit in this register will reset that bit. The Command register provides coarse control over a device's ability to generate and respond to PCI cycles. Access type of the Command register is read/write. See bit definitions.







#### IBM3206K0424



# **IBM Processor for Network Resources Accessor For Accessor Preliminary**





# **1.3: PCINT Config Word 2**

The Class Code is used to identify the generic function for this device. The Revision ID is used to identify the level of function for this device. See bit definitions.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **1.4: PCINT Config Word 3**

This word specifies the system cache size in units of 32-bit words, the value of the Latency Timer for this PCI bus master, the Header Type which identifies the layout of bytes in configuration space, and the register for the control and status of BIST (Built-in self-test). See bit definitions.







# **1.5: PCINT Base Address 1 (I/O for Register)**

This register specifies the base address of where in PCI I/O or memory space the IBM3206K0424 registers will be mapped. When written with '1's and read back, the least significant bits read back as '0' will indicate the amount of I/O space required for this device to operate. For example, when a value of 'FFFFFFFF' is written, a value read of 'FFFFFF00' indicates that 256 bytes of address space is required. See bit definitions.

The programming of this bit depends on whether the IBM3206K0424 is in 64-bit addressing mode or not. When in 64-bit addressing mode, bit 4 of the PCINT 64-bit Controller Register is set to '1', and this register specifies a memory address. When the IBM3206K0424 is not in 64-bit addressing mode because bit 4 of the PCINT 64-bit Control Register is set to '0', this register specifies an I/O address. See bit definitions and [PCINT 64-bit Control Register](#page-115-0) on page 116.



**When in 64-bit Addressing Mode (that is, bit 64 of PCINT 64-bit Control Register is set to '1'):**





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **When not in 64-bit Addressing Mode (that is, bit 64 of PCINT 64-bit Control Register is set to '0'):**







#### **1.6: PCINT Base Address 2 (Mem for Register)**

This register specifies the base address of where in PCI memory space the IBM3206K0424 registers will be mapped. When written with '1's and read back, the least significant bits read back as '0' will indicate the amount of memory space required for this device to operate. For example, when a value of 'FFFFFFFF' is written, a value read of 'FFFFFF00' indicates that 256 bytes of address space this required. See bit definitions.

The programming of this bit depends on whether the IBM3206K0424 is in 64-bit addressing mode or not. When in 64-bit addressing mode, bit 4 of the PCINT 64 bit Controller Register is set to '1', and this register specifies a memory address. When the the IBM3206K0424 is not in 64-bit addressing mode because bit 4 of the PCINT 64-bit Control Register is set to '0', this register specifies an I/O address. See bit definitions and [PCINT 64-bit Control Register](#page-115-0) on page 116.



### **When in 64-bit Addressing Mode (that is, bit 64 of PCINT 64-bit Control Register is set to '1'):**



Base Address



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **When not in 64-bit Addressing Mode (that is, bit 64 of PCINT 64-bit Control Register is set to '0'):**







### **1.7: PCINT Base Addresses 3-6 (Memory)**

This register specifies the base address of where in PCI memory space the IBM3206K0424 memory will be mapped. When written with '1's and read back, the least significant bits read back as '0' will indicate the amount of memory space required for this device to operate. For example, when a value of 'FFFFFFFF' is written, a value read of 'FFFFFF00' indicates that 256 bytes of address space this required. See bit definitions.

The mapping for the base address of registers into IBM3206K0424 memory is one-to-one, assuming a memory windowing option is not set in the PCINT Base Addr Control Register for that base address register (BAR). Multiple BARs are only used to use a given system memory map more efficiently. As required by the BAR, the addresses are size-aligned. For example, that means a 16MB size could be represented with one BAR as one 16MB size aligned on a 16MB boundary. However, four-4MB BARs could represent the same 16MB size but be aligned on any 4MB boundary. The value in any of the BARs does not map directly to any particular IBM3206K0424 memory structure, such as Control Memory. The addresses are mapped using the Virtual, Packet, and Control base address registers in VIMEN.



### **When in 64-bit Addressing Mode (that is, bit 64 of PCINT 64-bit Control Register is set to '1'):**







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **When not in 64-bit Addressing Mode (that is, bit 64 of PCINT 64-bit Control Register is set to '0'):**





**Note:** These registers power up to X'08000000' if accessed little endian.



# **1.8: PCINT CardBus CIS Pointer**

This register contains the offset to where the Card Information Structure (CIS) is located. See bit definitions.



**Power on Reset value** X'00000000'





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **1.9: PCINT Subsystem ID/Vendor ID**

This register contains the Subsystem ID and Subsystem Vendor ID. See bit definitions.

Other possible codes that could be returned for the Subsystem ID are listed below. The correctness of their value is superseded by higher (IOA card) levels of documentation.









# **1.10: PCINT ROM Base Address**

This register specifies the base address of where in PCI memory space the IBM3206K0424 ROM will be mapped. When written with ones and read back, the least significant bits read back as '0' will indicate the amount of memory space required for this device to operate. For example, when a value of 'FFFFFFFF' is written, a value read of 'FFFFFF00' indicates that 256 bytes of address space is required. See bit definitions.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **1.11: Capabilities Pointer**

This register contains the Capabilities Pointer. See bit definitions.



Capabilities Pointer






# **1.12: PCINT Config Word 15**

This register is used to communicate interrupt line routing information, tell which interrupt pin this device uses, and specify the desired setting for Latency Timer values. See bit definitions.







#### **IBM Processor for Network Resources All and Service Construction Construction Preliminary**



## **1.13: PCINT Endian Control Register**

This register allows control and status to the big/little endian address selection. It controls the byte order across the PCI bus. See bit definitions.



**Power on Reset value** X'00000000'







# <span id="page-110-0"></span>**1.14: PCINT Base Address Control Register**

This register controls all the base address registers that map to memory. See bit definitions.





## **IBM Processor for Network Resources Preliminary**







## **1.15: PCINT Window Offsets for Base Addresses 3-6**

These registers specify the amount of memory space required for this device to operate. See bit definitions.



**Restrictions** Can be written or read during configuration cycle, memory cycle when enabled (see [PCINT Base Address Control Register on page 111](#page-110-0)), or an I/O cycle. This register is documented as big endian, but how data is presented on the PCI bus depends on how the controls are set in the PCINT Endian Control Register.







# **1.16: PCINT Count Timeout Register**

This register holds the count limit of PCI slave retry cycles. See bit definitions.







## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**





# <span id="page-115-0"></span>**1.17: PCINT 64-bit Control Register**

This register contains miscellaneous control bits.









## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



### **1.18: PCINT 64-bit Enable Register**

See the [PCINT 64-bit Control Register](#page-115-0) on page 116 for the bitwise description that the corresponding bit in this register will enable (a value of '1' means enabled). Any bit in this register ANDed with bit 0 of PCINT 64-bit Control Register will determine if the other bits in PCINT 64-bit Control Register are set.





# <span id="page-118-0"></span>**1.19: PCINT Perf Counters Control Register**

This register contains control bits for the PCINT performance Counter 1 and PCINT Performance Counter 2.





## **IBM Processor for Network Resources Accessor For Accessor Preliminary**



T





**Preliminary IBM Processor for Network Resources**

# **1.20: PCINT Perf Counter 1**

This register contains PCI performance counter 1.







## **1.21: PCINT Perf Counter 2**

This register contains PCI performance counter 2.







## **1.22: PCI Master Options Control**

This register contains the control register when the IBM3206K0424 is the PCI master.



**Restrictions** None

**Power on Reset Value**  X'00000000' **(Big Endian)**

**Power on Reset Value**  X'00000000' **(Little Endian)**





## **IBM Processor for Network Resources Preliminary**







# **1.23: Power Management Program Control**

This register contains the control register for power management signalling.





## IBM3206K0424



## **IBM Processor for Network Resources Accessor For Accessor Preliminary**





# **1.24: Message Signaled Interrupts-Word 1**

This register contains the part of the Message Signaled Interrupts structure. See bit definitions.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **1.25: Message Signaled Interrupts-Word 2**

This register contains the part of the Message Signaled Interrupts structure. See bit definitions.



**(Little Endian)**





# **1.26: Message Signaled Interrupts-Word 3**

This register contains the part of the Message Signaled Interrupts structure. See bit definitions.





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **1.27: Message Signaled Interrupts-Word 4**

This register contains the part of the Message Signaled Interrupts structure. See bit definitions.

**Length** 16 bits

**Type** Read/Write **Address** XXXX 00CC

**Restrictions**

**Power on Reset Value**  X'00000000' **(Big Endian)**

**Power on Reset Value**  X'00000000' **(Little Endian)**







# **1.28: Power Management Interface-Word 1**

This register contains the part of the Power Management Interface structure. See bit definitions.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **1.29: Power Management Interface-Word 2**

This register contains the part of the Power Management Interface structure. See bit definitions.



**(Little Endian)**

Power Management Capabilities







## **1.30: Vital Product Data Interface-Word 1**

This register contains the part of the Vital Product Data Interface structure. See bit definitions.



**(Little Endian)**





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **1.31: Vital Product Data Interface-Word 2**

This register contains the part of the Vital Product Data Interface structure. See bit definitions.

**Length** 32 bits **Type** Read/Write **Address** XXXX 00DC **Restrictions Power on Reset Value**  X'00000000' **(Big Endian) Power on Reset Value**  X'00000000'

**(Little Endian)**





# **Entity 2: Interrupt and Status/Control (INTST)**

This entity contains the masking registers that choose which interrupt/status source will be gated onto one of the two available interrupt I/O pins. A new delayed interrupt function has been added. This function allows IBM3206K0424 status registers to be read and placed in system memory before the interrupt signal is raised. For details, see [DMA QUEUES \(DMAQS\)](#page-153-0) on page 154.

A bus timer function is provided in this entity that times a single bus access to make sure that the cycle is terminated before the system timer times out. This allows the user code an opportunity to recover from the error as opposed to the subsystem common code.

Below is a summary of this entity's functions:

- Interrupt Prioritized Status Registers
- Interrupt Source Register
- Interrupt Enable Registers
- Bus timer function
- Control Processor error register with enable register

## **2.1: INTST Interrupt 1 Prioritized Status**

Used to help quickly parse which interrupting entity of the IBM3206K0424 is active.

read back.



31-0 Prioritized Status



will be a hex number equal to bit number  $n + 1$ . For example, if bit 31 is on, X'20' will be

 $\sqrt{ }$ 

#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**



J

## **2.2: INTST Interrupt 2 Prioritized Status**

Used to help quickly parse which interrupting entity of the IBM3206K0424 is active.



**Power on Reset value** X'00000000'

Prioritized Status









## **2.3: INTST Control Register**

This register is used to control various IBM3206K0424 functions. See [Note on Set/Clear Type Registers on](#page-92-0)  [page 93](#page-92-0) for more details on addressing.







## **IBM Processor for Network Resources Accessor For Accessor Preliminary**







## <span id="page-138-0"></span>**2.4: INTST Interrupt Source**

This register indicates the source(s) of the interrupt(s) pending. It can also be used as a status register when the bits are enabled. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing. Note that bits in this register always reflect the state of the source register bit: Writing a value will have no effect. Reserved bits will not take on the written value. The delay of running through a latch has been removed. For the delayed interrupts feature, writing this register at the end of an interrupt handling routine will guarantee that interrupt1 and interrupt2 (if enabled) will pulse off, allowing the logic to get ready for the next interrupt DMA.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**





## **2.5: INTST Enable for Interrupt 1 (MINTA)**

This register serves as an enable for interrupt 1. See the [INTST Interrupt Source](#page-138-0) register on [page 139](#page-138-0) for the bitwise description that the corresponding bit in this register will enable. See [Note on Set/Clear Type Regis](#page-92-0)[ters on page 93](#page-92-0) for more details on addressing.





### **2.6: INTST Enable for Interrupt 2 (MINT2)**

This register serves as a enable for interrupt 2. See the [INTST Interrupt Source](#page-138-0) Register on [page 139](#page-138-0) for the bitwise description that the corresponding bit in this register will enable. See [Note on Set/Clear Type Regis](#page-92-0)[ters on page 93](#page-92-0) for more details on addressing.



## **2.7: INTST Interrupt Source without Enables**

This register is used to help quickly parse which interrupting bit of INTST Interrupt Source is active. It does not matter what state the Enable registers are set to because the value returned does not depend on them.



Prioritized Status ſ 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



### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# <span id="page-141-0"></span>**2.8: INTST CPB Status**

This register holds the status bits for errors on the Control Processor bus. These bits, when disabled, will set a bit in the INTST Interrupt Source register. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



### <span id="page-143-0"></span>**2.9: INTST CPB Status Enable**

This register serves as an enable for the INTST CPB Status register. See Note on Set/Clear Type Registers [on page 93](#page-92-0) for more details on addressing. See the [INTST CPB Status on page 142](#page-141-0) for the bitwise description that corresponding bit in this register will enable. This enable will initialize to the disabled state.



## **2.10: INTST IBM3206K0424 Halt Enable**

This register serves as an enable for the INTST CPB Status register and gates which errors will reset bit 4 (Master chip enable), bit 5 (Master chip enable for Transmitting), and bit 6 (Master chip enable for Receiving), all in the INTST Control Register register. This allows selected bits to disable the IBM3206K0424, especially in the case of severe hardware detected errors. See the [INTST CPB Status on page 142](#page-141-0) for the bitwise description that corresponding bit in this register will enable. This enable will initialize to the disabled state. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



## **2.11: INTST CPB Capture Enable**

This register serves as an enable for the INTST CPB Status that will determine on which error type the INTST CPB Captured Address register will be updated. See the [INTST CPB Status on page 142](#page-141-0) for the bitwise description that corresponding bit in this register will enable. See Note on Set/Clear Type Registers on page [93](#page-92-0) for more details on addressing.




 $\omega$ 

### **2.12: INTST CPB Captured Address**

This information can be used to attempt a retry in the exception handling microcode. This register holds the value of the IBM3206K0424 register address on the PCI during a bus error condition. This only latches values from sources that are enabled in the INTST CPB Capture Enable register.







### **2.13: INTST General Purpose Timer Pre-scaler**

This is the pre-scaler for the INTST General Purpose Timer Compare.This register will hold the value of the pre-scale count. The default value is 1 tick every 10.02uS, assuming a 33MHz or 66MHz PCI Bus clock, producing a 66MHz system clock (count is system clock). The pre-scale count value is n-1, where n is the desired increment count.

Owing to a physical design problem, the function of this register was lost. It should be set to a non-zero value, so that the INTST General Purpose Timer Counter can be used with a prescale of only the default clock (one tick every 30ns, assuming a 33-MHz system clock).



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **2.14: INTST General Purpose Timer Compare**

This is the compare value for the general purpose timer. This register holds the value of the data that is compared to the count value in the INTST General Purpose Timer Counter, setting the INTST General Purpose Timer Status bits. See [INTST General Purpose Timer Mode Control](#page-147-0) on page 148 for details on the operation of this register.



# **2.15: INTST General Purpose Timer Counter**

This is the general purpose timer counter.

This register holds the value of the counter. It always counts up. See INTST General Purpose Timer Mode Control [on page 148](#page-147-0) for details on operation of this register.





# **2.16: INTST General Purpose Timer Status**

This is the status of the general purpose timer counter. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



**Power on Reset value** X'0'





# <span id="page-147-0"></span>**2.17: INTST General Purpose Timer Mode Control**

This register controls the operating modes of the general purpose timer counter. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing.



**Power on Reset value** X'4'







# **2.18: INTST Enable for PCORE Normal Interrupt**

This register serves as an enable for the PCORE normal interrupt input. See INTST Interrupt Source on page [139](#page-138-0) for the bitwise description that the corresponding bit in this register will enable. See Note on Set/Clear [Type Registers on page 93](#page-92-0) for more details on addressing.



# **2.19: INTST Enable for PCORE Critical Interrupt**

This register serves as an enable for the PCORE critical interrupt input. See INTST Interrupt Source on page [139](#page-138-0) for the bitwise description that the corresponding bit in this register will enable. See Note on Set/Clear [Type Registers on page 93](#page-92-0) for more details on addressing.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **2.20: INTST Debug States Control**

This register serves as the control for external debug states.



**Type** Read/Write

**Address** XXXX 0490

**Restrictions** None

**Power on Reset value** X'38030201'







# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

### **2.21: INTST Delayed Interrupts DMA System Address 1**

This register serves as the Delayed Interrupts DMA System Address. This register holds the value of the system DMA address to which the delay interrupt status data will be moved.



### **2.22: INTST Delayed Interrupts DMA System Address 2**

This register serves as the Delayed Interrupts DMA System Address. This register holds the value of the system DMA address to which the delay interrupt status data will be moved.



### **2.23: Current PCI Master Address Counter for Debug**

This register holds the current PCI Master address counter value. This register holds the value of the PCI Master DMA address. The function of this register is primarily for debug when a severe error occurs that stops the DMA engine from running.





# **2.24: External Entity States Read**

This register will get a snapshot value of the enstates pin I/O. This register will return whatever is on the output of the enstates mux output. It is strictly for debug and a convenient way to look at the current state of IBM3206K0424 internal logic. It is controlled by INTST Debug States Control.





# <span id="page-153-0"></span>**Entity 3: DMA QUEUES (DMAQS)**

DMAQS provides the interface to the IBM3206K0424's DMA master capability (described in General Purpose [DMA \(GPDMA\)](#page-174-0) on page 175). It provides three DMA queues that hold DMA descriptor chains that are executed in a multiplexed fashion. Together with GPDMA, a very powerful interface is provided to software to complete complex tasks including TCP/IP checksumming for transmit and receive packets. The following sections describe the features of DMAQS, how to set up DMAQS, and some troubleshooting tips.

# **DMA Descriptors**

DMA descriptors can reside in either PCI/system memory space or the IBM3206K0424 memory space. Certain types of descriptors, called cut-through DMA descriptors, must be located in the IBM3206K0424 memory space. DMA descriptors that are located in the IBM3206K0424 memory space are more efficient to process because they do not need to be moved across the PCI bus. However, it is more costly for software to update across the bus. The best option is to mix descriptors in both locations. DMA descriptors that are infrequently changed should reside in the IBM3206K0424 memory, while dynamic descriptors should be placed in system memory. Descriptors located in the IBM3206K0424 memory space must fall in a definable address range. See [DMAQS Local Descriptor Range Registers](#page-172-0) on page 173.

# **DMA Descriptor Layout**





# **DMA Types and Options**

The DMA descriptor is very versatile and can perform many actions. The following list shows some examples andpossible flags to use. Other combinations are possible: see [GPDMA Transfer Count and Flag Register](#page-178-0) [on page 179.](#page-178-0)

# **DMA Types and Flags**



**Note:** These are not the only options. Some of the above can be ORed together also.

Using the above, you can efficiently do TCP checksumming, place user events in receive queues, do register reads/writes, free buffers, and get buffers.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



### **Descriptor Based DMAs**

This is the recommended approach to processing DMAs. A single descriptor or a descriptor chain is built that describes the actions to take. The descriptor is then enqueued to the proper DMA Queue. The number of the descriptor in the DMA chain is placed in the lower six bits of the descriptor address as it is enqueued.

### **Register Based DMAs**

While register based DMAs can be enabled and used, they are not recommended because they are not as efficient and they do not leave a debug trail as the descriptors do in the DMA queue. These should not be used concurrently with descriptor-based DMAs for a particular queue, but register-based and descriptor-based DMAs can be used on different queues. One possible use for register-based DMAs is doing DMAs from the core.

# **Polling, Interrupts, or Events**

There are several choices for handling DMA completion. First, the status register can be polled. While not very efficient, it is the easiest option. Second, you can use interrupts to tell when a DMA is done. Again, not very efficient. However, interrupts should be used to tell when a DMA error has occurred.

One way to deal with DMA completes is to use the RXQUE event mechanism. By generating events, the user can dump in DMA descriptor and clean up at a later time when it is convenient. The user can use the automatic DMA events using the queue on DMA complete flag, or the user can place a user event on an arbitrary queue by writing a DMA descriptor that does an explicit RXQUE enqueue with user data.

### **Error Detection and Recovery**

Ideally, there should not be any errors. Errors are usually user errors in the DMA descriptor which need to be fixed and are not recoverable. Errors on the PCI bus (that is, parity) should not occur in a normal working system, and typically you do not want to recover them. However, if recovery is desired, the current DMA must be recovered in GPDMA. Upon successful completion of the recovered DMA, DMAQS will resume operation.

### **DMA/Queue Scheduling Options**

There are three DMA queues. Queue 0 is higher priority than the other two. This high priority queue is always scheduled to go if the current descriptor is ready. The other two queues (Q1 and Q2) are of equal priority and are scheduled in a round robin fashion when the descriptor is ready. This is meant to provide a transmit DMA queue, receive DMA queue, and a high priority DMA queue. However, these queues can be used for any purpose by setting the routing registers properly.

The queues can be arbitrated after each DMA request length operation, after complete DMA descriptor chains complete, or after a single DMA descriptor in a chain completes. The queues can also be placed in true round robin mode, where all three queues have equal priority.

### **Address Size**

DMAQS can be operated with either 32- or 64-bit System Addresses. See PCINT 64-bit Control Register. All DMAQS address registers are 64 bits wide. In 32-bit addressing mode, the high order portion of address registers are initialized at reset to '0', and cannot be modified. In 32 bit addressing mode, word four of the DMA Buffer Descriptor is ignored.



### **Data Width**

DMAQS recognizes 64 bit writes to 64 bit internal registers. DMAQS internal 64-bit registers may be written either as a 64-bit entity, or by two 32-bit writes. All DMAQS registers are memory mapped on a 64-bit boundary (address bits 2:0 = 0). When in 64-bit addressing mode, an address register is updated with 32 bit writes (atomicity of update cannot be guaranteed). The user should use semaphores to assure the integrity of the operation.

# **Initialization of DMAQS**

DMAQS is very simple to set up by following these steps:

1. Set up each of the three DMA queues.

To do this, you need to know the size of each queue (see [DMAQS Upper Bound Registers on page 159](#page-158-0) for choices). Given this information, the DMA queue is set up with two register writes in diagnostic mode (see [DMAQS Control Register](#page-163-0) on page 164).

 $dmags->lowerBound[g] = baseAddress / / should be aligned with size of queue$  $dmags \rightarrow upperBound[g] = encodedSize; // et encoded size of dmag$ or dmags->bounds[q] = baseAddress + encodedSize; // if 64 bit data is enabled

The data structure for the DMA queue is now set up.

- 2. Set up the queue thresholds if they are being used: dmaqs->threshold[q]=threshold //setthresholdsizetobeinterrupted on //may also need to set int mask
- 3. Set up the local DMA descriptor range if local descriptors are being used:

```
dmaqs->localDescriptorLowerBound = localDescriptorBase // set base addr of local desc in 
IBM3206K0424 memory
dmaqs->localDescriptorUpperBound = localDescriptorEnd; // set ending addr of local desc in 
IBM3206K0424 memory
```
4. Set up any options that are being used in the DMAQS Control Register:

```
dmaqs->control[set] = ENABLE_DMA_QUEUES | CLR_CHECKSUM_TO_FOXES; // set options/modes
```
5. Finally, clear the diagnostic bit:

dmags->control $[clr] = DIGMODE$  // clear the diag mode bit

6. Need to set up memory bank selection if necessary, but normally Control Memory is used.

#### **IBM Processor for Network Resources All and Service Control of All and Service Preliminary**



#### **3.1: DMAQS Lower Bound Registers**

These registers specify the lower bound of the corresponding DMA queue data structure. These registers specify the lower bound of the corresponding DMA queue data structure. The head, tail, and length of the DMA queue are initialized when this register is written. When the DMA queue wraps past the upper bound, it wraps back to the value in the lower bound register, thus implementing the DMA queue as a circular buffer.

When this register is written, the corresponding DMA queue is essentially reset. This is because the head, tail, and length of the queue are all reset.



The low order nine bits are not writable and read back '0'.



# <span id="page-158-0"></span>**3.2: DMAQS Upper Bound Registers**

These registers specify the encoded size/upper bound of the corresponding DMA queue data structure. The actual upper bound is calculated by adding the decoded queue size to the lower bound. When the DMA queue wraps past the upper bound, it wraps back to the lower bound register, thus implementing the DMA queue as a circular buffer.



These registers can only be written when the diagnostic bit has been set in the DMAQS Control Register.





#### **3.3: DMAQS Head Pointer Registers**

These registers point to the head element of the corresponding DMA queue. During normal operations, these registers do not need to be read or written; they are used by the IBM3206K0424 to implement the DMA queues. These registers are initialized when the DMAQS Lower Bound Registers for the corresponding DMA queue is written.



# **3.4: DMAQS Tail Pointer Registers**

These registers point to the next free element of the corresponding DMA queue.





# **3.5: DMAQS Length Registers**

These registers specify the length in bytes of the corresponding DMA queue. This register is cleared when the corresponding DMAQS Lower Bound Registers is written.



# **3.6: DMAQS Threshold Registers**

These registers specify a queue length threshold at which the corresponding status bit is generated.

These registers should be set equal to the queue length that should cause status to be generated. For example, if the value was set to five, then no interrupt would be generated until the queue was length five or more for the corresponding DMA queue. The threshold is level sensitive, so as long as the length is greater than or equal to the threshold, the corresponding status bit is set. When this register is set to '0', no thresholding occurs.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# <span id="page-161-0"></span>**3.7: DMAQS Interrupt Status**

This register indicates the source(s) of the interrupt(s) pending.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **3.8: DMAQS Interrupt Enable**

This register serves as a mask for [DMAQS Interrupt Status](#page-161-0). See DMAQS Interrupt Status on page 162 for the bitwise description that the corresponding bit in this register will mask. See Note on Set/Clear Type Registers [on page 93](#page-92-0) for more details on addressing.



# <span id="page-163-0"></span>**3.9: DMAQS Control Register**

Used to set options for DMAQS. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









### **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Community Preliminary**





### **3.10: DMAQS Enqueue DMA Descriptor Primitive**

This register enqueues a DMA descriptor chain to the corresponding DMA queue. The write data is the address of the descriptor chain that describes the DMA transfers. The low six bits contain a count of the number of DMA descriptors in this chain. After the DMA descriptors are enqueued by writing to this register, the chain of descriptors is fetched from system memory and the DMA transfers described by the chain of descriptors are performed. In 32 bit addressing mode, the low-order 32 bits of the register are written, and the high-order 32 bits are reset when the register is loaded.



### **3.11: DMAQS Source Address Register**

This register is used to set and keep track of the Source Address during a DMA transfer. This is the source for the current DMA transfer. A bit in the Transfer Count and Flag Register determines whether the source address is internal to the IBM3206K0424 or is a system address. In 32-bit addressing mode, the low-order 32 bits of the register are written, and the high-order 32 bits are reset when the register is loaded.





# **3.12: DMAQS Destination Address Register**

This register is used to set and keep track of the destination address during a DMA transfer. This is the Destination address for the current DMA transfer. In 32-bit addressing mode, the low-order 32 bits of the register are written, and the high-order 32 bits are reset when the register is loaded. A bit in the Transfer Count and Flag Register determines whether the destination address is internal to the IBM3206K0424 or is a system address.



# **3.13: DMAQS Buffer Address Register**

This register is used to set and keep track of the POOLS Buffer address during a DMA transfer. When the DMA Descriptor directs that a new buffer address be obtained from POOLS, this is Buffer Address for the current DMA transfer. A bit in the Transfer Count and Flag Register determines whether a buffer address has been obtained for this descriptor. This register can be written to an RXQUE queue. The low-order seven bits should be set to 0x2a, the event code for Assign Pool Buffer events.

This is the Destination address for the current DMA transfer. In 32-bit addressing mode, the low-order 32 bits of the register are written, and the high-order 32 bits are reset when the register is loaded. A bit in the Transfer Count and Flag Register determines whether the destination address is internal to the IBM3206K0424 or is a system address.





# <span id="page-167-0"></span>**3.14: DMAQS Transfer Count and Flag Register**

This register specifies the type and number of bytes transferred during a DMA transfer. The lower 16 bits are a counter of the number of bytes transferred during a DMA transfer. The upper 16 bits specify the type of transfer.







# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



# **IBM Processor for Network Resources Accessor For Accessor Preliminary**







### **3.15: DMAQS System Descriptor Address**

The upper 57 bits contain the address of the current descriptor block and the lower seven bits contain the number of descriptors in the chain that remain to be processed. When doing register-based DMAs, the low six bits are set to "000001" when the DMAQS Transfer Count and Flag Register is written. If DMA descriptors are used for DMA transfers, this register will contain the system address of the current descriptor block and the number of descriptors that remain to be processed. This address may be queued on DMA completion to correlate DMA transfers with system control blocks. In 32-bit addressing mode, the low-order 32 bits of the register are written, and the high-order 32 bits are reset when the register is loaded.



# **3.16: DMAQS Checksum Register**

This register contains the accumulated checksum. This register contains the accumulated checksum value. It can also be used to initialize the checksum with a seed value. The most significant bit contains the alignment state  $(1 = \text{odd}, 0 = \text{even alignment})$ . The alignment state is significant between subsequent checksummed DMAs.

This register can be read at four different addresses. The base address returns the unmodified accumulated checksum. The base address +4 returns the inverted accumulated checksum. The base address + 8 returns the byte-swapped accumulated checksum. The base address + 12 returns the inverted byte-swapped accumulated checksum.



# **IBM Processor for Network Resources Preliminary**







### <span id="page-172-0"></span>**3.17: DMAQS Local Descriptor Range Registers**

These registers specify the lower and upper bounds respectively of the memory range for local DMA descriptors.

These registers contain the address of the lower and upper bound of the memory range of descriptors that are in the IBM3206K0424. If a descriptor block is enqueued, it is compared to these registers. If it falls within this range, only the descriptor address is placed on the queue. When the descriptor is to be loaded into the DMA registers, and it falls within this range, it will not be taken from the queue but loaded directly from the descriptor address. These registers are 4K aligned. The upper bound register contains the address of the last 4K block in the local descriptor address range.



**Restrictions** Can be written in diagnostic mode only. Upper bound is 32 bits. The upper 32 bits are internally generated, and are not different from the upper 32 bits of the lower bound register. The last four addresses in this range are reserved, and should not be used to hold descriptors.

### **3.18: DMAQS Event Queue Number Register**

This register specifies which DMAQS queue should be used when DMA descriptors are enqueued from CSKED (DMA on transmit comp). This register also indicates the RXQUE Event Queue to which events should be enqueued for each DMAQS queue register.



#### **IBM Processor for Network Resources And American Structure III and American Structure Preliminary**



#### **3.19: DMAQS DMA Request Size Register**

This register specifies the maximum request size for DMA descriptor scheduling.

This is the amount of data that DMAQS will request GPDMA to move in a single request. For example, if a descriptor wants to move 2K of data and the request size is set to 512 bytes, then DMAQS will request 512 bytes to be moved and then rearbitrate the DMA queues. A value of '0' is the same as 0xffff.



# **3.20: DMAQS Enq FIFO Register**

This register is for diagnostic use only. It holds DMA descriptor waiting to be placed on a DMA queue. Reading this register is destructive. The oldest entry is read on each read. If it is desired to re-dispatch the dequeued entries, they will have to be re-enqueued.

DMA Descriptors which reside in System Storage are not immediately copied into the queue, but space is reserved in the queue for the descriptors, and a descriptor is built which will copy the descriptor into the queue when needed.





# <span id="page-174-0"></span>**Entity 4: General Purpose DMA (GPDMA)**

This entity provides DMA control between System Memory and IBM3206K0424 Packet Memory.

DMA transfers must be enabled in the GPDMA control registers for transmit and/or receive. There are two ways to initiate DMA transfers. The first is by directly writing the Source Address, Destination Address, and Transfer Count and Flag Registers. The second is by using DMA descriptors and enqueueing them using DMAQS. These two methods should not be used simultaneously. If using descriptors, refer to the DMAQS section beginning on [DMA QUEUES \(DMAQS\)](#page-153-0) on page 154 for more information.

DMA transfers to system I/O space are not allowed.

# <span id="page-174-1"></span>**4.1: GPDMA Interrupt Status**

This register indicates the source(s) of the interrupt(s) pending, or is used as a status register when the bits are enabled. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







#### IBM3206K0424







# **4.2: GPDMA Interrupt Enable**

This register serves as a mask for GPDMA Interrupt Status. See [GPDMA Interrupt Status](#page-174-1) on page 175 for the bitwise description that the corresponding bit in this register will mask. See Note on Set/Clear Type Registers [on page 93](#page-92-0) for more details on addressing.





# **4.3: GPDMA Control Register**

Used to set options for DMA operations. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.













# **4.4: GPDMA Source Address Register**

Used to set and keep track of the Source Address during a DMA transfer. This is the system address that increments during a DMA transfer. A bit in the Transfer Count and Flag Register determines if the source address is internal to the IBM3206K0424 or is a system address.





#### **4.5: GPDMA Destination Address Register**

Used to set and keep track of the Destination address during a DMA transfer. This is the Destination address that increments during a DMA transfer. A bit in the Transfer Count and Flag Register determines if the destination address is internal to the IBM3206K0424 or is a system address.



### <span id="page-178-0"></span>**4.6: GPDMA Transfer Count and Flag Register**

Specifies the type and number of bytes transferred during a DMA transfer. The lower 16 bits are a counter of the number of bytes transferred during a DMA transfer. It is a count down counter; when zero is reached, the transfer ends. Writing a non-zero value to the lower 16 bits starts the DMA transfer. The upper 16 bits specify the type of transfer as follows.







# **IBM Processor for Network Resources Accessor For Accessor Preliminary**






# **4.7: GPDMA DMA Max Burst Time**

Used to limit the number of cycles a master can burst on the PCI bus. When a DMA burst is started, a counter is loaded with the value in this register. When the counter expires and the current access completes, the PCI bus is released for use by another bus master. Writing a non-zero value to this register enables this function.



# **4.8: GPDMA Maximum Memory Transfer Count**

Used to limit the size of data requests to the Control/Packet Memories. This register defines the maximum number of bytes to be transferred in a single storage request to IBM3206K0424 Storage.



# **4.9: GPDMA Checksum Register**

This register contains the accumulated checksum value. It can also be used to initialize the checksum with a seed value. The most significant bit contains the alignment state  $(1 = odd, 0 = even$  alignment). This register can be read at four different addresses. The base address returns the unmodified accumulated checksum. The base address +4 returns the inverted accumulated checksum. The base address + 8 returns the byte-swapped accumulated checksum. The base address + 12 returns the inverted byte-swapped accumulated checksum.



### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



# **4.10: GPDMA Read DMA Byte Count**

This register counts the bytes transferred into the IBM3206K0424 by the DMA controller. Descriptor bytes can optionally be included. (See the [GPDMA Control Register](#page-176-0) on page 177 for details.)



# **4.11: GPDMA Write DMA Byte Count**

This register counts the bytes transferred out of the IBM3206K0424 by the DMA controller.



# **4.12: GPDMA Array Read Address**

This register is used to read the GPDMA internal array. The internal array is used to hold data for the DMA. The array is organized as 64 32-bit words. The GPDMA Array Read Address is written with the address of the word that is to be read. The GPDMA Array Data Register is read to obtain the contents of the addressed word. The GPDMA Array Read Address is incremented each time the GPDMA Array Data Register is read, causing repeated reads of the GPDMA Array Data Register to obtain sequential words from the array.





# **4.13: GPDMA Array Write Address**

This register is used to write the GPDMA internal array. The internal array is used to hold data for the DMA. The array is organized as 64 32-bit words. The GPDMA Array Write Address is written with the address of the word that is to be written. The GPDMA Array Data Register is written to write the addressed word. The GPDMA Array Write Address is incremented each time the GPDMA Array Data Register is written, causing repeated writes of the GPDMA Array Data Register to write sequential words in the array.



# **4.14: GPDMA Array**

Reads the contents of the internal array. The internal array is used to hold data for the DMA.





# **Memory Controlling Entities**

# **Entity 5: The DRAM Controllers (COMET/PAKIT)**

This section describes the function of the COMET/PAKIT entities. COMET is the memory controller for Control Memory, and PAKIT is the memory controller for Packet Memory.

Each controller can support the following types of memory:

- Synchronous DRAMs running at 133MHz (7.5 ns cycle time) with a CAS latency of two or three and a burst length of one or two. Memory sizes of 4MB, 8MB, 16MB, and 32MB are supported. Please note that the cycle time of the SDRAM clock is a constant on the IBM3206K0424. Any SDRAM part selected must be capable of running at 133MHz or faster at CAS latency 2 or 3.
- Synchronous SRAM running at 133MHz (7.5 ns cycle time) with a read latency of two and a write latency of zero or two. Memory sizes of 1MB, 2MB, 4MB, and 8MB are supported.

**Note:** For any memory configuration, modules must be selected such that the loading on any memory net (including card wiring) does not exceed 120pF.

The number of column address lines is programmable, allowing both DRAMs with symmetric address (same number of row and column address lines) and asymmetric address (typically having more row than column address lines).

If using SDRAM, the memory may be operated as having one or two arrays. The arrays are differentiated by their chip selects. If the memory is configured to have two arrays, the memory's address range is split equally between the two arrays.

Memory checking can be enabled/disabled, and the method of checking selected can be either ECC or parity. IF ECC is selected, seven data bits are used for ECC over the 32 data bits. If parity is selected, four data bits are used to provide parity over the 32 data bits.

COMET/PAKIT are designed so that memory contents are preserved over a reset. If the IBM3206K0424 is reset while a memory write cycle is in progress, the cycle is completed in an orderly fashion to ensure that valid ECC/parity is written. Memory timings are not violated when reset goes active. Refresh is maintained during the reset.



# **Memory Reset Sequence**

After a reset, onboard ROM or external firmware must properly configure the control registers for COMET/PAKIT.

If using SRAM, the reset sequence is complete. If using SDRAM, bit 3 of the memory controller's SDRAM Command and Status Register must be written to a '1' to initiate forcing the SDRAMs out of the self refresh state and performing the POR sequence. When bits five and four of this register are '00', the SDRAMs are ready for use.

**Note:** Memory configuration errors occur if an attempt is made to use memory that is configured incorrectly or, if an attempt is made to use SDRAM before the POR sequence is completed.

Accesses to the first 0x20 bytes of memory (Control or Packet) are not allowed unless bit 26 of the corresponding memory control register is set. With this restriction in place, accesses with zero-valued pointers will cause the zero address error bit in the memory controller's status register to be set.

### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **5.1: COMET/PAKIT Control Register**

This register contains the information that controls the functions of the entity. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing. Before this register can be altered, writing it must be enabled in [COMET/PAKIT Memory Controller Write Enable Register](#page-196-0) [\(](#page-196-0)described on [page 197\)](#page-196-0).









# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**







# **5.2: COMET/PAKIT Status Register**

This register contains status information for COMET/PAKIT. See Note on Set/Clear Type Registers on page [93](#page-92-0) for more details on addressing.









# **5.3: COMET/PAKIT Interrupt Enable Register**

This register contains bits corresponding to the bits in the COMET/PAKIT Status Register. If a bit in this registe[r](#page-92-1) is set and the corresponding bit is set in the COMET/PAKIT Status Register, an interrupt is generated. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



# **5.4: COMET/PAKIT Lock Enable Register**

This register contains bits corresponding to the bits in the COMET/PAKIT Status Register. If a bit in this register is set and the corresponding bit is set in the COMET/PAKIT Status Register, a signal is sent to VIMEM indicating that memory should be locked. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.





# **5.5: COMET/PAKIT Memory Error Address Register**

This register holds the address at which the last memory error occurred.







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



### **5.6: COMET/PAKIT SDRAM Command and Status Register**

This register is used to issue various commands to and control the timing operation of Synchronous DRAMS when they are attached to the IBM3206K0424. If the IBM3206K0424 is not configured for SDRAMs, any writes to bits 3-0 of this register are ignored. This register is also used to reflect the status of the Synchronous DRAMs. When a command bit in this register is set (bits 3-0 only), the command executes and resets the bit upon completion. Only one bit (3-0 only) may be set during any write. Software should poll this register to make sure the previous command has completed before issuing another write to this register. If more than one bit at a time is written to this register (3-0 only), the results may be unpredictable.











# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **5.7: COMET/PAKIT DRAM Refresh Rate Register**

This register holds the value of a counter used to control the rate of refresh for the DRAM.









# **5.8: COMET/PAKIT Syndrome Register**

This register holds the syndrome bits that can be used to isolate the data or check bit in error when ECC is used. When parity is used, this register indicates which of the four bytes of the memory bus had a parity error.







# **IBM Processor for Network Resources Preliminary**



# **ECC Syndrome Bits**





# **5.9: COMET/PAKIT Checkbit Inversion Register**

This register can be used for diagnostic purposes to invert the ECC/parity check bits that are written to memory.







# <span id="page-196-0"></span>**5.10: COMET/PAKIT Memory Controller Write Enable Register**

This register must be written to a specific pattern before the Memory Control Register can be written.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **5.11: COMET/PAKIT Memory Configuration Error Sense Register**

This register can be read to help determine the source of a memory configuration error. The bits in this register reset automatically once the configuration error is resolved.









# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**





# **Entity 6: ATM Virtual Memory Logic (VIMEM)**

This entity is responsible for adjustment of all addresses provided to the memory control entities. All addresses can be categorized into three distinct types, based entirely upon the location of the requested address with respect to the three base registers defined in this entity. The three types of addresses are referred to as control, real packet, and virtual packet addresses.

All memory requests arriving on the Control Memory bus are handled as Control Memory accesses, and simply have the contents of the Control Memory Base Register subtracted from them before being passed on to the Control Memory Entity. When the processor accesses memory, the cache controller compares the requested address to the Real Packet Memory Base Register and if the address is less than the base register, the request is routed to the Control Memory bus; otherwise it is routed to the Packet Memory bus. All requests arriving on the Packet Memory bus are compared to the Virtual Memory Base Address Register. If the address of the request is less than the base register, the contents of the Real Packet Memory Base Register are subtracted from the address and this address is passed on to the Packet Memory Control Entity. If the requested address is greater than or equal to the base register, a more complex, but flexible scheme is used to determine the real address to provide to the Packet Memory Control Entity. For a detailed explanation of the virtual address generation scheme refer to [Virtual Memory Overview](#page-247-0) on page 248 and the accompanying figures.

# **6.1: VIMEM Virtual Memory Base Address**

This register defines the starting address of the virtual address space used to manage incoming and outgoing frames. Any time an access is made to Virtual Memory that falls within the defined bounds of Virtual Memory, the contents of this register are subtracted from the virtual address to derive the true offset into Virtual Memory. This true offset, along with the known length of all virtual buffers, allows the index of the specific virtual buffer to be derived by the Virtual Memory access hardware. This index can then be used to access the real buffer map associated with this virtual buffer.







# **6.2: On-Chip Memory Base Address**

This register is used by various entities to generate the base address of the On-Chip Memory (OCM).



# **6.3: VIMEM Control Memory Base Address**

This register defines the starting address of the Control Memory address space. Any time an access is made to Control Memory, the contents of this register are subtracted from the address before an access to memory occurs.







### **6.4: VIMEM Packet Memory Base Address**

This register defines the starting address of the Packet Memory address space. Any time an access is made to Packet Memory, the contents of this register are subtracted from the address before an access to memory occurs.







# **6.5: VIMEM Virtual Memory Total Bytes**

This register defines the total number of bytes in the address space being allocated for Virtual Memory. The contents of this register, divided by the configured size of virtual buffers, yields the total number of virtual buffer indices that should be used to initialize POOLS. The value of the indices should range from this calculated value minus one, down to zero. If an address is determined to be above or equal to the Virtual Memory Base Register it is assumed to be a virtual access. If the virtual buffer index derived from the requested address indicates that the virtual buffer space being accessed is above the limit defined by this register, an error is generated.





### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **6.6: VIMEM Virtual/Real Memory Buffer Size**

This register defines the total number of bytes to be occupied by each of the virtual or real buffers as well as the spacing from one buffer to the next.











# **6.7: VIMEM Packet Memory Offset**

This register contains the number that will be added by the VIMEM access logic to all accesses of real Packet Memory that occur. In a high performance configuration (separate control and packet store), this register should be written to all zeros to indicate that all accesses of real Packet Memory do not require any additional offset to be added. In a medium performance configuration (combined control and packet store), this register should be loaded with a value that indicates the logical partitioning between control and packet storage. If for instance, a single bank of 2 meg was configured and this register was loaded with X'00100000' (1 meg), then all accesses to real Packet Memory would be forced into the 1-meg to 2-meg range.



# **6.8: VIMEM Maximum Buffer Size**

This register is used by the Virtual Memory logic to determine if an access to a virtual buffer falls into the region of the buffer that can be accessed. If a virtual buffer read or write accesses an offset in a virtual buffer that is greater than the contents of this register, the Virtual Memory logic can be configured to halt and generate an interrupt. The power up value of all ones causes this check to be disabled. This register is intended to provide the user with a means of providing additional protection to accesses of the virtual buffers. For example, if this register is loaded with X'FF8', all memory access up to and including the byte at address X'FFF' are allowed. Any access of offset X'1000' or above will cause an exception.





# **6.9: VIMEM Access Control Register**

The bits in this register control the configurable features of the Virtual Memory logic. See Note on Set/Clear [Type Registers on page 93](#page-92-0) for more details on addressing.









# **6.10: VIMEM Access Status Register**

This register contains information regarding the current status of the Virtual Memory logic mainly with respect to detected error access conditions. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







# **IBM Processor for Network Resources Accessor For Accessor Preliminary**







### **6.11: VIMEM Access Status Interrupt Enable Register**

This register allows the user to enable interrupts for each of the conditions reported in the VIMEM Access Status Register. Each bit corresponds to the same bit in the status register and when set to '1' generates an interrupt to the processor if the condition is detected. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



# **6.12: VIMEM Memory Lock Enable Register**

This register allows the user to selectively allow each of the conditions reported in the VIMEM Access Status Register to force a memory lock condition in the memory controller. Each bit corresponds to the same bit in the status register and when set to '1' causes a memory lock if the condition is detected. See Note on [Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



## **6.13: VIMEM State Machine Current State**

This register provides feedback to the user regarding the current status of the state machines in VIMEM. One use of this register is to make sure that the required initialization time has expired after loading the segment size register. This is accomplished by reading this register repeatedly until the initialization state machine is in the idle state.









# **6.14: VIMEM Last Processor Read Real Address**

This register provides information to the user about the last read access of virtual Packet Memory by the processor. If a virtual address was accessed, this register contains the real address generated by the Virtual Memory logic that can be used to access the same location. This register is intended mainly as an aid in debugging to make virtual address translation easier. To perform the translation, the processor must read from the desired virtual address: after the read is complete, this register contains the real address that was accessed. The address contained in this register is an offset from the beginning of physical Packet Memory.





#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



# **6.15: VIMEM Virtual Buffer Segment Size Register**

This register, along with the lower four bits of the real buffer base registers, defines the size of the second through 16th real buffers that are concatenated to make up a virtual buffer. Two bits of this register are associated with each real buffer segment and indicate one out of four possible associations. The associative possibilities are shown in the bit table below. Every two bits defines the connection between a particular buffer segment and the real buffer base registers.



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



# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources And American Structure III and American Structure Preliminary**



### **6.16: VIMEM Buffer Map Base Address**

This register contains the address in Packet Memory at which the buffer map table starts. The buffer map table consists of a variable number of eight-byte entries for each buffer that is allocated in the system. The first 16 bits of each eight-byte entry contains the POOL ID and various status flags associated with this buffer, thus this base register is used in both real and Virtual Memory modes.

In Virtual Memory mode, each of the three subsequent 16 bits contains an index which is associated with a buffer size base register using the Buffer Segment Limit Register. The index and buffer size base register are used to determine a real buffer address. If the map size is set to eight bytes, only one eight-byte entry is used for each buffer. If the map size is set to 16 bytes, two eight-byte entries are used for each buffer. If the map size is set to 32 bytes, four eight-byte entries are used for each buffer. If the map size is set to 64 bytes, five eight-byte entries are used for each buffer, and the remaining 24 bytes of the map are unused by the hardware.





 $\tilde{E}$ 

Size  $\mathbf{\Sigma}$ 





# **6.17: VIMEM Real Buffer Base Addresses**

These registers contain the address in Packet Memory at which a block of memory begins that is used to provide a given size buffer. In general, the block allocated must be large enough to contain as many buffers as will be freed to POOLS on initialization. However, for Real Buffer Base 4, the size of the block reserved must be large enough so that one buffer is available for each of the virtual buffers freed to POOLS. These buffers must not be freed to POOLS because they are implicitly used as the first real buffer segment for each of the virtual buffers. If a given base register (and associated buffer size) is not used, the low four bits of the register should be set to X'F' to ensure that accesses of this buffer size are detected and flagged as an error. When using real memory mode (controlled in POOLS), all of these base registers are unused with the exception of base register zero, which contains the base address for all real memory buffers. In real mode, the low four bits of base register zero are of no significance. The size of the real buffers is controlled through the Buffer Size Register.



**Restrictions** The base address for any given buffer size must begin on a boundary that is equal to the buffer size. For example, the base address for 128-byte buffers must be on a 128-byte boundary, and the base address for 4096-byte buffers must be on a 4096-byte boundary.

> When a base register is written, the hardware performs an automatic adjustment to the address using the contents of the Packet Memory real base register and the Packet Memory offset register. This results in the actual value being stored, not being the value that is written by the program. This is done to make the memory accesses that use the base register execute quicker. The reverse adjustment is made when the read operation is performed, so that it appears to the program no different than a normal operation. Care must be taken however to ensure that the Packet Memory Real Base Register and the Packet Memory Offset Register are set-up before any of the base registers are written. If the Packet Memory Base Register or the Packet Memory Offset Register is changed, Packet Memory should not be accessed until all the base registers have been written again. The power on value of these registers is actually the power on value of the Packet Memory Real Base Register added to the contents of the Packet Memory Offset Register added to the original contents of these registers (X'0000000F').

# **IBM Processor for Network Resources Accessor For Accessor Preliminary**






# **Entity 7: ATM Packet/Control Memory Arbitration Logic (ARBIT)**

This section contains descriptions of the registers used by the arbiter logic.

## **7.1: ARBIT Control Priority Resolution Register High**

The bits in this register define the priority of requesting entities to Control Memory.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **7.2: ARBIT Control Priority Resolution Register Low**

The bits in this register define the priority of requesting entities to Control Memory.







### **7.3: ARBIT Control Error Mask Register**

The bits in this register control whether ARBIT detected error conditions on an entity's interface will lock the Control Memory subsystem. Bits in this register also control the locking of the Control Memory subsystem based on Control Memory, Packet Memory, Virtual Memory, and BCACH detected error conditions. Resetting the appropriate bit will force errors from that source to be ignored.







## **7.4: ARBIT Control Error Source Register**

The bits in this register provide feedback to indicate the source of errors that have been detected by the memory subsystem.







# **7.5: ARBIT Control Winner Register**

The bits in this register indicate which entity currently owns Control Memory.







#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**

### **7.6: ARBIT Control Address Register A**

If latch bank A is active, the bits in this register indicate the last address that was used to access Control Memory.



Last Address Provided by Arbiter



### **7.7: ARBIT Control Address Register B**

If latch bank B is active, the bits in this register indicate the last address that was used to access Control Memory.



Last Address Provided by Arbiter



31-0 These bits contain the last address provided by the arbiter to the Control Memory controller.



# **7.8: ARBIT Control Length Register**

The bits in this register indicate the last length that was used to access Control Memory.









#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**

### **7.9: ARBIT Control Lock Entity Enable Register**

The value programmed in this register controls what entity, if any, has access to Packet Memory immediately after memory has locked. This register powers up to a value that will not allow any entity to access memory after a lock condition until the lock condition has been properly cleared.



Bit Map Value

3 2 1 0





# **7.10: ARBIT Control Config Register**

The bits in this register control the operation of the Control Memory arbiter.







# **7.11: ARBIT Packet Priority Resolution Register High**

The bits in this register define the priority of requesting entities to Packet Memory.



#### **IBM Processor for Network Resources Preliminary**







# **7.12: ARBIT Packet Priority Resolution Register Low**

The bits in this register define the priority of requesting entities to Packet Memory.





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



#### **7.13: ARBIT Packet Entity Error Mask Register**

The bits in this register control whether ARBIT detected error conditions on an entity's interface will lock the Packet Memory subsystem. Bits in this register also control the locking of the Packet Memory subsystem based on Control Memory, Packet Memory, Virtual Memory, and BCACHE detected error conditions. Resetting the appropriate bit will force errors from that source to be ignored.







# **7.14: ARBIT Packet Error Source Register**

The bits in this register provide feedback to indicate the source of errors that have been detected by the memory subsystem.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **7.15: ARBIT Packet Winner Register**

The bits in this register indicate which entity currently owns Packet Memory.









ı

### **7.16: ARBIT Packet Address Register A**

If latch bank A is active, the bits in this register indicate the last address that was used to access Packet Memory.





### **7.17: ARBIT Packet Address Register B**

If latch bank B is active, the bits in this register indicate the last address that was used to access packet memory.



Last Address Provided by Arbiter





 $\Gamma$ 



# **7.18: ARBIT Packet Length Register**

The bits in this register indicate the last length that was used to access Packet Memory.









### **7.19: ARBIT Packet Lock Entity Enable Register**

The value programmed in this register controls what entity, if any, has access to Packet Memory immediately after memory has locked. This register powers up to a value that will not allow any entity to access memory after a lock condition until the lock condition has been properly cleared.



Bit Map Value

ſ 3 2 1 0



### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **7.20: ARBIT Packet Config Register**

The bits in this register control the operation of the Packet Memory arbiter.









### **7.21: ARBIT Performance Counter Control**

The bits in this register determine what events are counted by the memory performance counters. This 32-bit register is divided into four 8-bit values, one value for each of the counters. The eight bits determine what memory event is counted by the associated counter.



### **IBM Processor for Network Resources Accessor For Accessor Preliminary**







### **7.22: Arbit Memory Performance Counter**

These registers count memory events as defined in the ARBIT performance counter control register.



#### Number of times the event selected in the performance counter control register has occurred





# **Entity 8: The Bus DRAM Cache Controller (BCACH)**

This entity provides the caching function for data transfers on the Control Processor bus. The array is organized in four logically separate cache lines, any of which can be used for processor accesses or master/slave DMA accesses. The cache is accessible on byte boundaries on the Control Processor side; access of this entity to COMET is performed on 64-bit (word) boundaries. The address tags of each of the four 32-byte cache lines are compared to the requesting address to select the bank to be used to satisfy the Control Processor bus operation.

Streaming accesses of the cache use a predictive look-ahead scheme to fill the cache for read operations from Packet Memory. Under normal conditions, a single cache miss will be expected at the start of each DMA read operation. This cache miss will initiate a read operation from Packet Memory to fetch the requested data and enough additional data to fill the remainder of the cache line. If the requested data is in the last N bytes (N is programmable via the BCACH control register) of the cache line, the read operation to COMET will be extended to fill the next cache line with sequential data as well. This same programmable value is used to determine when to initiate the next sequential cache line fill operation during a DMA read operation. During non-aligned write operations to Packet Memory, BCACH will perform read/modify/write cycles to PAKIT.

Processor accesses operate without predictive caching. When a cache miss occurs, a COMET read operation will be initiated to fetch the 32-byte block of data that contains the requested data. The data read from COMET will be loaded into the 'Least Recently Used' cache line.

This section contains descriptions of the registers used by the Bus Cache logic.

# **8.1: BCACH Control Register**

The bits in this register control the various functions provided by the cache logic. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing.





### **IBM Processor for Network Resources Accessor For Accessor Preliminary**









### **8.2: BCACH Status Register**

The bits in this register reflect the current status of the cache. See Note on Set/Clear Type Registers on page [93](#page-92-0) for more details on addressing.







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **8.3: BCACH Interrupt Enable Register**

The low eight bits in this register allow the user to selectively determine which bits in the BCACH status register will cause processor interrupts. A '0' in a bit position masks interrupts from the corresponding bit location in the BCACH status register. A '1' in a bit position allows interrupts for the corresponding bit in the BCACH status register. The high eight bits in this register allow the user to selectively determine which bits in the BCACH status register will lock the cache. A '1' in any bit position forces the cache to lock if the corresponding bit is set in the BCACH status register. If the cache locks, all status regarding the cache lines is maintained until the cache enable bits in the control register are turned off.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



### **8.4: BCACH High Priority Timer Value**

This register defines the number of 15 ns cycles that will pass from the time that a valid PCI bus request is raised to BCACH until BCACH will raise its high priority request to the memory controllers. A value of '0' in this register disables this function completely.



Number of 15 ns Cycles







### **8.5: BCACH Line Tag Registers**

These registers are useful only in diagnostic testing of the cache logic. Each register will contain the tag value for the data contained in that particular cache line.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



#### **8.6: BCACH Line Valid Bytes Register**

These registers are useful only in diagnostic testing of the cache logic. Each register will contain a bit significant flag indicating which bytes in the 32-byte cache line are valid. All of these bits will be active after a cache line fill operation has occurred, but any combination of these bits can be valid after the processor has performed a write operation to memory.



Cache Line Fill Operation Set/Reset





### **8.7: BCACH Line Status Register**

These registers are useful only in diagnostic testing of the cache logic. Each register will contain a bit significant flag indicating the current status of the associated cache line.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**





# **8.8: BCACH Cache Line Array**

This array is divided into four 32-byte buffers used as cache lines 0, 1, 2, and 3.







# **Entity 9: Buffer Pool Management (POOLS)**

POOLS acts as a memory manager for the IBM3206K0424. Memory buffers are checked out and checked in via two operations (primitives) supported by POOLS: the get pointer primitive and the free pointer primitive. These primitives can be performed explicitly by accessing specified addresses within the POOLS entity, and they may also be done by hardware. CSKED can free a buffer upon transmission if specified by the corresponding packet header (see *[ECC Syndrome Bits](#page-195-0)* on page 196), and RAALL gets buffers to store received data. In addition, POOLS contains mechanisms to control resource utilization and supports a Real Memory Mode and a Virtual Memory Mode.

### **Basic Operation in Real Memory Mode**

If memory is viewed as a series of buffers, POOLS maintains a circular list of available buffers. There are pointers (the head and tail) to the start and the end of the list. When a get pointer primitive is executed, the buffer at the head of the list is checked out, the head pointer is advanced and the correct resource group(s) is debited. When a free pointer primitive is executed, the freed buffer is checked in at the end of the list, the tail pointer is advanced, and the correct resource group is credited.

### **Basic Operation in Virtual Memory Mode**

With the addition of Virtual Memory, POOLS must maintain five sets of head and tail pointers, thresholds, and active counts: one for the virtual buffers themselves and the rest for the four regions of real buffers that constitute the virtual buffers. In this case the base virtual address is the item returned from a get pointer operation and returned during a free pointer operation. When the get buffer primitive is executed, POOLS creates an active buffer map (page table) for the virtual address. As the virtual address is used and buffer (page) boundaries are crossed, VIMEM will request buffers from POOLS when a buffer (page) fault occurs. VIMEM then places the buffer index in the buffer map. When the virtual buffer is no longer needed and a free pointer primitive is issued with the starting virtual address, POOLS takes the contents of the buffer map and frees the resources that were assigned to the buffer map.

### **Resource Controls**

POOLS adds another layer of service by creating "pools" of buffers (currently a maximum of 16 pools). For each pool, a maximum number of allowable buffers can be specified. The intent is to make it possible for several applications to use an IBM3206K0424 at once without one or more applications starving the remaining applications for memory buffers. Particular pools buffers are divided into "guaranteed" and "common" buffers. All the guaranteed buffers are considered to be dedicated to their respective pool and are therefore not available for general use. The common buffers are all the memory buffers remaining after the guaranteed buffers are subtracted from the total buffers. To maintain the buffer limits on each pool, every pool has a guaranteed threshold, total threshold, and an active count. When a request is made for a buffer from a particular pool, the guaranteed threshold is first checked. If the active count of the pool is less than the guaranteed threshold, the buffer is provided. If the guaranteed threshold has been reached, then the total threshold is checked. If the active count is equal to the total threshold, no buffer is provided. If the active count is less than the total threshold, and a common buffer is available, a buffer is provided. If there are no common buffers available, a buffer cannot be provided and a null index is returned. To determine if a common buffer is available, a count is maintained for each size of buffer.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



### **Virtual Memory Overview**

Each virtual buffer consists of a number of real buffers. For each virtual buffer there is a buffer map that defines the size and number of real buffers that may be allocated to the virtual buffer. Each map is built from a common template (the VIMEM Virtual Buffer Segment Size Register) that associates 1 to n buffer indexes in the map to a real buffer in one of the four real buffer regions defined in VIMEM. In VIMEM, the Buffer Map Base Address Register defines the size of the map and therefore also the number of buffer indexes in the virtual buffer map. Each eight-byte entry of the map contains the pool ID of the pool to which the buffer is allocated plus space for three real buffer segment indexes. This implies the smallest map yields a virtual buffer of one to four real buffer segments (three real buffer segments plus the implicit real buffer that all virtual buffers are allocated). The biggest map defines a virtual buffer of 1-16 real buffer segments (15 plus the implicit one).

The intention of this structure is to allow the user to customize the value in the Virtual Buffer Segment Size Register to utilize memory in an efficient manner relative to network data traffic. For example, if network traffic contained 50% packets of < 512 bytes, 35% packets of < 1K bytes, and the rest was < 5K bytes, the user could set up Virtual Memory to use three real segments of 512 bytes, 512 bytes, and 4K bytes respectively. The incoming data would neatly fit into the segments and minimize wasted memory.

POOLS and VIMEM maintain the maps for the virtual buffers. On a write that crosses a real buffer boundary into an as yet an unresolved region of a virtual buffer, a page fault occurs. When a page fault occurs, POOLS determines whether or not a real buffer can be assigned. If it can be assigned, the index of the real buffer relative to the base address of the particular buffer size is placed by VIMEM into the buffer map. The first buffer is implicitly associated with the Virtual Memory address for a particular virtual buffer and enough real memory must be available to support the first real buffer of each virtual buffer at initialization time. There is not necessarily enough real storage for all the possible real buffers associated with a virtual buffer.



# **Virtual Address Buffer Map**



All real buffers of a particular size are stored in a contiguous region of memory. The buffer index, in conjunction with the base address for this real buffer size, points to a particular real buffer. The implicit buffers are also stored in a data structure of this type.



# **Buffer/Virtual Memory Allocation Structure in Memory**



The following example illustrates these concepts.



# **Virtual Address Buffer Map**



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

The lower seventeen bits of the virtual address are used in conjunction with the segment template in the VIMEM Virtual Buffer Segment Size Register to determine from which portion of the buffer map the buffer index is retrieved. Once the buffer index is retrieved, it is combined with the appropriate base address for that particular buffer size. The offset into the buffer is then added to get the real 32-bit address that is used in physical memory.

POOLS uses the data structures above to manage Packet Memory resources. Each LCD is associated with a particular POOL and multiple different LCDs may be associated with that same POOL. Within a POOL, there are five different resource categories and two variables to go with each resource.

### **Resources and Variables Example**



### **9.1: POOLS Get Pointer Primitive**

The POOLS Get Pointer Primitive returns a pointer to the requester. The request to the virtual packet/buffer size 4 address will always return a memory address. If in virtual mode, the address will be virtual. Requests made for buffer sizes 0 to 3 will not return an address but rather a buffer index in bits 15-0. The real address associated with this index can be generated by shifting the index by the buffer size (for example, six bit positions for a 64-byte buffer) and adding the result to the base address for this size buffer. Access to buffer sizes 0 to 3 is not permitted in operational mode.

The address of the primitive also selects the pool ID. The pool ID is contained in address bits 5-2, and it selects which pool will be charged for the pointer. The buffer size is selected with address bits 8-6.

If there are no more pointers available in the specified pool, a null pointer is returned. The active pointer count for that pool is incremented if a non-null pointer is returned. If the guaranteed threshold has been exceeded and a buffer from the common pool is returned, the common pools count for that size is decremented by 1.


#### **9.2: POOLS Free Pointer Primitive**

The POOLS Free Pointer Primitive returns the pointer to the proper free list. If it is a Virtual Memory address, the Virtual Memory Buffer Map is traversed to free the indexes associated with the Virtual Memory address. In the case where it is a real memory buffer, the single index is freed.

This primitive uses address mapping to select the size of the object to be freed. The size is contained in address bits 4-2. During normal operation, only frees to buffer size four are relevant. During initialization mode, buffer sizes 0 to 3 can be used to load indexes. The indexes are loaded into bits 31-16.

In normal operations it is not necessary to read this "register".



#### **9.3: POOLS Common Pools Count Registers**

The POOLS Common Pools Count Registers indicates the number of pointers in the particular common pool.

The bits are a 16-bit count. The Get Pointers that exceed the guaranteed allocation decrement this count by one assuming that the count is non-zero. When the count is zero, the Get Buffer operation will fail. The Free Pointers that operate beyond the guaranteed threshold for a particular client and free the pointer(s) increment this count by one. The microcode should initialize these registers to the value of the respective common pool that it desires to have.



#### **IBM Processor for Network Resources All and Security Construction Construction Preliminary**



#### **9.4: POOLS Client Thresholds Array**

The POOLS Client Thresholds Array holds the guaranteed and total threshold values for the 16 pools and the four (five) pointer sizes. This array contains the guaranteed and total thresholds for the managed POOLs.

When a Get Pointer primitive is processed, the values in this array are used to determine if a primitive can return a pointer. The active count from the Active Packet Count Array is used with these registers to determine if a threshold has been exceeded. If the guaranteed threshold has been exceeded and the total not exceeded and there is a common pointer available, then the common count will be incremented. If there are no common buffers available or the request will cause the total threshold to be exceeded, the request will be rejected. During a Free Pointer primitive processing, the pointer is returned to the free list and these thresholds are used to determine if a common count should be credited.







### **9.5: POOLS User Threshold and Client Active Packet Count Array**

The POOLS User Threshold and Client Active Packet Count Array holds the user thresholds and active pointer counts for each of the 16 managed pools and four (five) pointer sizes.

When a Get Pointer primitive is processed, the active count is retrieved and compared with the threshold counts. If it falls within bounds and a pointer is available, the active count will be incremented by one reflecting the additional buffer charged to that queue.

When a Free Pointer primitive is processed, the active count is retrieved, and, when the pointer is returned to the free list, the active buffer count is decremented by one.

The user threshold may be used to check on resource utilization as opposed to resource allocation. The Guaranteed and Total Thresholds are used when allocating resources to make decisions. The User Threshold is not used to govern resource allocation directly. One such use is for high water mark indication. When a Free Pointer primitive is processed or a Get Pointer is processed The active packet count is compared to the user threshold. If the event interface is enabled and a boundary condition is crossed an event is issued to the event interface.



User Threshold **Active Packet Count** 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



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **9.6: POOLS Pointer Queues DRAM Head Pointer Offset Address Register**

The POOLS Pointer Queues DRAM Head Pointer Offset Address Register indicates the address in DRAM where the head of the queue starts. This address, however, is only relative to the DRAM portion of the queue. Unless the head of the queue portion of the cache is locked out and needs two frames, the actual head of the queue is in the cache.

These 19 bits on write represent the offset to the address in DRAM of the head of the queue relative to the DRAM base address. On a read, the address in DRAM of the pointer is returned. This pointer is adjusted every time a cache frame boundary is crossed and a cache update cycle is completed to write through the additional queue elements. Because each memory reference contains four indices, this allows for a possible 128K index locations in the queue.



**Restrictions** During normal operations this register is to be used as a read only register. This register defaults to zero at initialization. It is assumed that the queues start on a maximum size queue boundary. These registers should be set up at initialization time. This register is cleared when the POOLS Pointer Queues DRAM Lower Bound Address Register is written to.



#### **9.7: POOLS Pointer Queues DRAM Tail Pointer Offset Address Register**

The POOLS Pointer Queues DRAM Tail Pointer Offset Address Register indicates the offset address in DRAM where the tail of the queue starts. This address, however, is only relative to the DRAM portion of the queue. Unless a 'no cache frames to be written through' state is in effect, the actual tail of the queue is in the cache.

These 19 bits on write represent the offset to the address in DRAM of the tail of the queue relative to the DRAM base address. On a read, the address in DRAM of the pointer is returned. This pointer is adjusted every time a cache frame boundary is crossed and a cache update cycle is completed to write through the additional queue elements. Since each memory reference contains four indices this allows for 128K index locations possible in the queue.



**Restrictions** During normal operations this register is to be used as a read only register. This register defaults to zero at initialization. It is assumed that the queues start on the maximum size queue boundary. These registers should be setup at initialization time. This register is cleared when the POOLS Pointer Queues DRAM Lower Bound Address Register is written to.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **9.8: POOLS Pointer Queues DRAM Lower Bound Address Register**

The POOLS Pointer Queues DRAM Lower Bound Address Register indicates the address in DRAM where the queue data structure is initially started. When the queue reaches the maximum address allowed for in the upper bound register, it wraps back around to the address specified in this register. This implements the queue in a circular buffer.

These 32 bits represent the address in DRAM where the queue begins and eventually wraps to. At initialization, this register and the POOLS Pointer Queues DRAM Tail Pointer Offset Address Register and the POOLS Pointer Queues DRAM Head Pointer Offset Address Register must be equal.



**Restrictions** During normal operations, this register is to be used as a read only register. This register should be setup at initialization time. The size of the DRAM queue storage which is formed with the lower and upper bounds is constrained in its size. It can be written when the diagnostic mode bit is set, otherwise the write is ignored. Note that if the maximum queue length exceeds the space available in the circular buffer, data corruption will occur when the actual queue length exceeds the maximum queue space available.



#### **9.9: POOLS Pointer Queues DRAM Upper Bound Register**

The POOLS Pointer Queues DRAM Upper Bound Register indicates the max queue length in DRAM of the queue data structure. When the queue reaches this address, it wraps back to the address specified by the lower bound register. This implements the queue in a circular buffer. This upper bound is to be provided as an encoded field. The encoded field represents the number of eight-byte addresses that can be contained by the queue.

These four bits represent the encoded maximum queue length in DRAM which, when matched, trigger the queue to wrap back to the address contained in the DRAM Lower Bound Address Register.



**Restrictions** During normal operations, this register is to be used as a read only register. This register should be setup at initialization time. The size of the DRAM queue storage which is formed with the lower and upper bounds is constrained in its size. It can be written when the diagnostic mode bit is set, otherwise the write is ignored. Note that if the maximum queue length exceeds the space available in the circular buffer, data corruption will occur when the actual queue length exceeds the maximum queue space available.

# **IBM Processor for Network Resources Accessor for Network Resources** Preliminary







#### **9.10: POOLS Pointer Queues Length Registers**

The POOLS Pointer Queues Length Registers indicates the length of the queue. The bits are a 16-bit count. A primitive that adds to the queue increments this counter. Primitives that remove items from the queue decrement this counter.



### **9.11: POOLS Interrupt Enable Register**

This register is used to enable bits from the POOLS Status Register and potentially generate interrupts to the control processor. When both a bit in this register and the corresponding bit(s) in the POOLS Status Register are set, the POOLS interrupt to PCINT will be enabled.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing. See [POOLS Status Regis](#page-264-0)ter [on page 265](#page-264-0) for the bit descriptions.



#### **IBM Processor for Network Resources All and Service Construction Construction Preliminary**



# **9.12: POOLS Event Enables**

This register is used to enable an event based on bits from the corresponding primitive transaction. If the bits are set in the enable and a transaction occurs that matches the event, an event will be sent to the RXQUE.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



# **9.13: POOLS Event Hysteresis Register**

The POOLS Event Hysteresis Register provide the capability for hysteresis on threshold checking.



Hysteresis Value







# **9.14: POOLS Event Data Register**

The POOLS Event Data Register provides the data that was sent on the last event.







### IBM3206K0424



# **IBM Processor for Network Resources Accessor For Accessor Preliminary**



# <span id="page-264-0"></span>**9.15: POOLS Status Register**

The POOLS Status Register provides status information about pools operations. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing.















# **9.16: POOLS Control Register**

The POOLS Control Register provide status information about POOLs operations. See Note on Set/Clear [Type Registers on page 93](#page-92-0) for more details on addressing.













### **9.17: POOLS Buffer Threshold Registers 0-4**

The POOLS Buffer Threshold Registers 0-4 is the threshold set by the software to set the threshold crossed bit in the POOLS Status Register. This register is used to compare with the queue length register.

This register consists of a 16-bit count match. The threshold count is compared to the queue length count. If the queue length is less than the value in this register, the appropriate bit is set in the status register respective to this queue.



# **9.18: POOLS Index Threshold Registers 0-4**

The POOLS Index Threshold Registers 0-4 provide error checking. These are the thresholds set by the software or hardware to set the index threshold crossed bit in the POOLS Status Register. This register is used to check indexes during free operations to look for an out of bounds index.

Each register consists of a 16-bit compare value. The threshold count is compared to the index while being processed. If an index is greater than the value in this register, the appropriate bit is set in the status register.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **9.19: POOLS Last Primitive Trap Register**

The POOLS Last Primitive Trap Register provide debug assistance. It contains the 32-bit last primitive address, and it is the last primitive address to POOLS, as selected in the POOLS Control Register, while in operational mode.



#### **9.20: POOLS Last Buffer Map Read on Free Register**

The POOLS Last Buffer Map Read on Free Register provide debug assistance. It contains the 32-bit address of the buffer map used in the last free operation, and it is the address of the last buffer map read on a free.



#### **9.21: POOLS Error Lock Enable Register**

The POOLS Error Lock Enable Register provides the ability to halt pools when the corresponding status bit in the status register are set.

When a bit in this register that corresponds to a bit that is set in the status register, the state machines in pools will be held in idle state until the lock is disabled.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.





#### **9.22: POOLS Packet and Control Memory Access Threshold**

The POOLS Packet and Control Memory Access Threshold timers are used to help limit the amount of time that pools can be held off from its respective memory.

The bits are a 12-bit count. When the proper bit in the POOLS Control Register is set and a request is made to the requisite memory, a counter is loaded with this value. The counter will then count down to zero in 30 ns ticks. When it hits zero, it forces the request to high priority.



### **9.23: POOLS Buffer Map Group**

The POOLS Buffer Map Group holds the buffer map of the packet that is in the process of being freed. From this map, the pool ID and the indexes that have been used are returned to their correct queue.

This register consists of a 16-bit flag field and 16-bit indexes.

The flag field contains the pool id and the valid bit. When a packet is freed, the valid bit is set to '0'. When a get operation occurs, the valid bit is then set. This helps to find address duplicates and other address related problems that software can generate.



**IBM Processor for Network Resources Accessor For Accessor Preliminary** 





# **Transmit Data Path Entities**

# **Entity 10: Transmit Buffer (CSKED)**

The transmit cell scheduler entity is responsible for receiving a packet from the processor, determining when cells from the packets need to be transmitted, and passing this information to the segmentation buffer entity.

The logic consists of timers and counters for determining transmit opportunities and interfaces to ARBIT (for accessing the timing data and descriptors), PCINT (for register accesses), RXQUE (for queuing events), POOLS (for returning buffers when finished transmitting), and SEGBF (for getting the data from memory to transmit).

# **Scheduling Overview**

This entity provides traffic shaping to ensure that traffic sent by the IBM3206K0424 conforms to the Quality of Service (QoS) parameters as defined by the ATM Forum. CSKED provides support for the following QoS parameters:

- Peak Cell Rate (PCR) The maximum number of cells per second that the connection can transfer into the network.
- Sustained Cell Rate (SCR) The average number of cells per second that the connection can transfer into the network. The burst tolerance determines the length of time over which the network measures this average.
- Burst Tolerance The maximum length of time that the user can transfer at the peak cell rate. Burst Tolerance can be measured in number of cells, a measurement known as maximum burst size (MBS).

CSKED will send out cells at the SCR. If transmit opportunities are missed, as is the case when there is no data to send, the actual rate will become less than SCR. When data becomes available to send, CSKED will transmit up to MBS cells at the PCR, until the transmit rate returns to SCR.



# **Operational Description**

#### **LCD Initialization**

A Logical Channel Data Structure (LCD) containing scheduling parameters for the circuit must be initialized before segmentation can be started. The parameters that are important to the operation of this entity are:



See [Transmit Logical Channel Descriptor Data Structures](#page-65-0) on page 66 for further information.

# **A Scheduling Example**

If a connection is to have an SCR of 50 Mbps and a PCR at the line rate of 150 Mbps and a MBS of 10 cells, the LCD needs to be initialized as follows:

- average\_interval =  $150$  Mbps/50 Mbps =  $3$
- peak interval = 150 Mbps/150 Mbps = 1
- max\_burst\_value =  $10*(3-1) = 20$

The following example uses a timeline to show how a connection with these parameters is scheduled. Cells are sent every third slot while there is data to send. After the first two cells are sent there is no more data to



send until another packet is enqueued. For each missed transmit opportunity, a cell can be sent at the peak interval, which is one. In the 15 timeslots after the first cell, five cells are sent for an SCR of 50 Mbps. The burst size is three, which is less than MBS. The unfilled slots can be used by other connections.

# **Timeline Example of Scheduling**



# **CSKED Initialization**

Before packets are enqueued for transmission, in addition to initializing the above scheduling parameters in the LCD, the following registers need to be set up.

- Timeslot Prescaler Register The amount of time for one timeslot is defined by this register. It defaults to 707 ns which is one cell time on a 622 Mbps Sonet interface.
- CSKED Control Register Additional scheduling options such as number of physical drops needs to be set in this register.

#### **IBM Processor for Network Resources All and Security Construction Construction Preliminary**



#### **Packet Initialization**

Packets to be segmented are written to Packet Memory, which has been allocated by POOLS. The address of the LCD describing the channel that this packet is to be transmitted on must be written to the header of the packet. Packet segmentation is started by issuing the transmit enqueue primitive to this entity. This entity will schedule segmentation of the packet according to the parameters set up in the LCD.

#### **Scheduling Options**

# **ABR Scheduling**

CSKED has logic to assist in the processing of ABR connections. If the connection is ABR, the LCD will have a different configuration, as specified in [Transmit Logical Channel Descriptor Data Structures](#page-65-0) on page 66. The following fields need to be initialized before the packets are sent on the connection.

- **Scheduling type** This field must be set to the value specifying an ABR connection.
- **Nrm** This field should specify the maximum number of cells a source may send for each forward RM-cell. Number of cells = (2∗∗Nrm)+1.
- **Trm** This field provides an upper bound on the time between forward RM-cells for an active source. Time = 100∗(2∗∗-Trm) msec.
- **ADTF** The ACR Decrease Time Factor is the time permitted between sending RM-cells before the rate is decreased to ICR. Time =  $ADTF* 0.01$  msec.
- All other ABR fields should be initialized to '0'.

#### **Frame Scheduling**

CSKED has logic to support frame-based scheduling. It is enabled whenever the PHY type is configured for POS-PHY in LINKC. In frame-based scheduling the packet is sent out at the line rate, but the start time of the next frame is determined by multiplying the peak interval by the number of 64-byte blocks in the packet. The average portion of the bandwidth used by a connection will be 1/(peak interval).

For example, if a connection is to use 1/4 of the bandwidth, the peak interval should be set to four. The frames will be sent out at line rate, but the spacing between the start of each frame will be four timeslots for each cell sent from the packet, so on average 1/4 of the bandwidth will be used by the connection. The following timeline example depicts sending two packets, the first contains four cells and the second contains three cells. The unfilled slots can be used by other connections.





# **Timeline Example of Frame Scheduling**

```
4 cells 3 cells<br>x 4 slots/cell 3 x 4 slots
                                  x 4 slots/cell x 4 slots/cell
       = 16 timeslots = 12 timeslots
|<----------------------------->|<--------------------->|
 S S S S S S S S ...
|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|->time
 |-| represents 1 time slot (the time it take to send 1 cell).
 S represents a cell being sent.
```
The above example assumes that one slot time is initialized to the time it takes to send 64 bytes out on the line. The term "cells" was used in this example to mean a 64-byte block of packet data. In frame mode, the ATM header is not prepended to the data being sent.

Weighted fair queueing on a frame basis is supported on the low priority queue by setting bit 17 in the CSKED Control Register. When using frame-based scheduling and weighted fair queuing together, the average interval will be used to limit the spacing between packets, not cells.

# **Path Scheduling**

CSKED has logic to support sharing scheduling parameters between multiple connections. In path scheduling, an LPD is set up to contain the scheduling parameters for the group of connections in the same way it is done for LCDs. All connections that wish to share this bandwidth set the alter sched field in their LCD to indicate this VC is on a VP, and initialize the lpd pointer field to point to the LPD. The segmentation portion of the LPD is not used since the segmentation parameters are taken from the LCD. The scheduling parameters in the LCDs are not used as they are in the LPD. The bandwidth is shared on a packet or cell basis depending the value of the alter\_sched field in the LCD. Since both the LPD and LCD need to be fetched for each transmit opportunity, this scheduling method should not be used where the performance boundaries are being pushed, as in 622 Mbps. See [Transmit Logical Channel Descriptor Data Structures](#page-65-0) on page 66 for further information on LPD descriptors.

#### **Primitives**

#### **Enqueue**

After a packet has been written to memory and the packet header updated with the offset and length of the data and the LCD address of the connection, an enqueue primitive needs to be issued to the Transmit Enqueue Primitive address.

#### **Close Connection**

When no more traffic is to be sent on a connection, this primitive can be executed to cause an event to be generated when segmentation has stopped on this connection. Segmentation will be stopped immediately, or stopped after all packets on this connection have been transmitted as specified in the CSKED Control Register.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



### **Start/Stop Timer**

When this primitive is executed, a timer is started or stopped whose parameters are contained in the specified LCD. When the timer pops, a DMA descriptor specified in the LCD will be executed.

#### <span id="page-277-0"></span>**10.1: Transmit Enqueue Primitive**

Enqueues a buffer for transmission.





#### **10.2: Resume Transmission Primitive**

Resumes transmission on an ABR connection that has been suspended. On an ABR connection, ADTF, CRM, and CCR=0 events will cause the transmission to be suspended until a rate conversion is completed, normally by the internal processor. This primitive will resume transmission on those connections, once the rate conversion is completed.









Start - Stop

# **10.3: Start/Stop Timer Primitive**

Start or stop a timer with the parameters in the specified LCD address. When this primitive is executed, a timer is started or stopped whose parameters are contained in the specified LCD. Bit 0 specifies whether to start (0) or stop (1) the timer. When the timer pops, a DMA descriptor specified in the LCD will be executed.







### **10.4: Close Connection Primitive**

Transmission is complete on a connection specified by an LCD address. When no more traffic is to be sent on a connection, this primitive can be executed to cause an event to be generated when segmentation has stopped on this connection. Segmentation will be stopped immediately, or stopped after all packets on this connection have been transmitted, as specified in the CSKED Control Register.







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **10.5: Timeslot Prescaler Register**

This register determines the length of time for one timeslot. This controls the rate that the cell scheduling counters are incremented. Each clock cycle, the value in this register is added to a 24-bit counter. When the upper bit of the counter changes state, the Current Timeslot Counter is incremented. This should normally be set to the time it takes to transmit one cell. It will be initialized to the cell time for a 622 Mb/s SONET connection (599.04 Mb/s payload). The following formula should be used to determine the value to load in this register: Timeslot Prescaler = (clock interval/timeslot interval) x 2∗∗23.



Timeslot Counter Rate



#### **10.6: Current Timeslot Counter**

This counter contains a count of how many prescaled intervals have elapsed. It is used to determine if scheduling needs to be done or credits exist.



Elapsed Interval Count





# **10.7: CSKED Control Register**

This register is used to control the actions of CSKED. See [Note on Set/Clear Type Registers on page 93](#page-92-1) for more details on addressing.







# **IBM Processor for Network Resources Accessor For Accessor Preliminary**







### **10.8: Transmit Segmentation Throttle Register**

This register contains the number of cycles to wait between successive requests to transmit a cell. Its purpose is to slow segmentation on all VCIs if it is determined by software that the network can not handle the generated load. The value in this register will be loaded into the Transmit Segmentation Counter each time a cell is accepted for transmission. For normal operation the value in this register should be '0'.











#### **10.9: Transmit Segmentation Throttle Counter**

This register is loaded with the value in the Transmit Segmentation Throttle Register after each cell is accepted for transmission and counts down until it reaches '0'. A new cell transmission will not be requested until this counter reaches '0'.



Transmit Segmentation Throttle Counter





# **10.10: MPEG Conversion Register**

This register is used to convert MPEG time units into timeslot time units. If MPEG traffic is configured in the LCD, the data stream will be monitored for PCRs. If a PCR is detected, it will be scheduled at the time specified in the PCR. A conversion factor needs to be written into this register to convert the MPEG time units into timeslot units. It will be initialized to a value that converts the MPEG time units (90 KHz) into the timeslot units (353.2 KHz, assuming one timeslot is the time it takes to send one cell over a SONET connection). The lower 12 bits of this register contain the fractional portion of this conversion factor.

Example: 353.2076 KHz / 90 KHz = 3.924528 = 3.ECB hex



MPEG Time Conversion Factor 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0 Bit(s) Description 14-0 Contains the conversion factor.



# **10.11: ABR Timer Prescaler Register**

This register determines the length of time for a tick of the RM Cell Timer. This controls the rate that the cell scheduling counters are incremented. Each clock cycle, the value in this register is added to a 24-bit counter. When the upper bit of the counter changes state, the RM Cell Timer is incremented. This should be set to value of 0.78 ms. It will be initialized to 0.78 ms assuming a 30-ns clock (as set up in SCLOCK). The following formula should be used to determine the value to load in this register:

ABR Timer Prescaler = (clock interval/0.78 ms) x 2∗∗23.



ABR Counter Rate





# **10.12: RM Cell Timer**

This register is used to keep track of the last time that an ABR RM cell was sent. Its period should be 0.78 ms.



RM Timer Value



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **10.13: CSKED LCD Update Data Registers**

Used to specify data to write into the LCD on the update LC operation. These registers contain the data used in the LC Update Operation. For more information on its use, see the [CSKED LCD Update Operation Regis](#page-286-0)ters [on page 287](#page-286-0).

This register changes to contain the updated data written to the LC word while the operation is completing.

The second set of LCD update registers is meant for the core to use, but is available for general use.



#### **10.14: CSKED LCD Update Mask Registers**

Used to specify data to write into the LCD on the update LC operation. These registers contain the mask used in the LC Update Operation. For more information on its use, see the [CSKED LCD Update Operation Regis](#page-286-0)ters[on page 287](#page-277-0).

The second set of LCD update registers is meant for the core to use, but is available for general use.





#### <span id="page-286-0"></span>**10.15: CSKED LCD Update Operation Registers**

Used to specify the LCD word to update. This operation is used to update a portion of an LCD. If this operation is not used, software or IBM3206K0424 updates of the LCD may be lost because the LCD is cached in IBM3206K0424 while cells are being processed.

This register is written with the address of the LCD word to update. Once this register is written the update operation starts. All subsequent reads or writes to the data, mask, or update registers are held off until the operation completes. A read-modify-write will occur to update the portion specified by the mask with the masked value in the data register.

Normally this register would not be read. However, if it is read then the low order bit is read as '0' and the next lowest order bit (bit 1) is read as the busy bit. This signifies whether an operation is still going on. If an operation is still going on, then a new write to any of the data, mask, or update operation registers is held off until the original operation is complete.

The second set of LCD update registers is meant for the core to use, but is available for general use.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



#### **10.16: Drop Access Control Register**

Each drop(4) has registers that can be used for debugging purposes and bandwidth limiting. To conserve address space, this register determines the drop for the register access. These registers are Fast Serviced Counters, Slow Serviced Counters, and Priority Bandwidth Limit Registers. This register must be rewritten whenever values for a different drop need to be read or written.



Drop






# **Performance Registers**

This section contains registers that are for performance purposes.

### **10.17: High Priority Bandwidth Limit Register**

This register can be used to limit the bandwidth used by high priority connections. The upper eight bits of this register specifies the number of high priority cells that can be sent in a window specified by the lower eight bits of this register. If no data needs to be sent by lower priority connections, high priority connections will not be limited.















#### **10.18: Medium Priority Bandwidth Limit Register**

This register can be used to limit the bandwidth used by medium priority connections. The upper eight bits of this register specify the number of medium priority cells that can be sent in a window specified by the lower eight bits of this register. If no data needs to be sent by low priority connections, medium priority connections will not be limited.





15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0



### **10.19: Low Priority Bandwidth Limit Register**

This register can be used to limit the bandwidth used by low priority connections. The upper eight bits of this register specify the number of low priority cells that can be sent in a window specified by the lower eight bits of this register.



Cells Transmitted Cell Times in Window







 $\sqrt{2}$ 

#### **10.20: High Priority Cells Transmitted Counter**

This register contains the number of cells transmitted from high priority connections.



Number of High Priority Cells Transmitted





#### **10.21: Medium Priority Cells Transmitted Counter**

This register contains the number of cells transmitted from medium priority connections.



Number of Medium Priority Cells Transmitted



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



 $\overline{\phantom{a}}$ 

### **10.22: Low Priority Cells Transmitted Counter**

This register contains the number of cells transmitted from low priority connections.



Number of Low Priority Cells Transmitted







#### **10.23: Bytes Queued Counters**

These registers (12) contain the number of bytes queued for transmission for each priority (3) on each drop (4). The addresses are assigned to the range in the following order:



Number of bytes Waiting to be Transmitted



31-0 This value represents the number of bytes waiting to be transmitted for each priority on each drop.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



# **Debugging Register Access**

This section contains registers that are for debug purposes only. These registers need not be written or read during normal operations.

#### **10.24: Fast Serviced Counters**

There are three fast serviced counters, one for each transmit priority: high, medium, and low. These registers contain the value of the last fast time slot that has been serviced. When this count differs from the current timeslot count, at least one fast slot needs servicing. Each time the fast slot is serviced, this counter will increment.



Fast Serviced Counters





#### **10.25: Slow Serviced Counters**

There are three slow serviced counters, one for each transmit priority: high, medium, and low. These registers contain the value of the last slow time slot that has been serviced. When this count differs from the current timeslot count, at least one slow slot needs servicing. Each time the slow slot is serviced, this counter will increment.



Slow Serviced Counters







### **10.26: Timer Serviced Counters**

In addition to the counters above, there is an additional counter for processing timer requests. These registers contain the value of the last timer slot that has been serviced. When this count differs from the current timeslot count (bits 22-15), at least one slow slot needs servicing. Each time a timer slot is serviced, this counter will increment.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



### **10.27: CSKED Status Register**

This register is used to control the actions of CSKED. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









#### **10.28: CSKED Interrupt Enable Register**

This register is used to enable interrupts from CSKED. If a bit is on in the status register and the corresponding enable bit is on in this register then an interrupt will be generated if enabled in INTST. See Note on [Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









٦

#### **10.29: CSKED Timing Data Array Pointer**

The CSKED Timing Data Array contains data relevant to scheduling cells. It contains 96 32-bit words. It should only be written for diagnostic purposes to test the array. It will power up to all zeros and should be rewritten to zeros after the array has been tested. This register points to an offset in the CSKED timing data array for accesses from the PCI bus. This register must be loaded with the correct offset before the desired data can be read from the CSKED timing data array data register. This register will auto-increment after each access to the associated data register. When the last address is accessed, this register wraps to zero.







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **10.30: CSKED Timing Data Array Data**

This register is used to access the array pointed to by the CSKED Timing Data Array Pointer.





#### **10.31: CSKED Time Wheel Array Pointer**

The CSKED Time Wheel Array contains data relevant to scheduling cells. It contains 16K 19-bit words. It should only be written for diagnostic purposes to test the array. It will power up to all zeros and should be rewritten to zeros after the array has been tested. This register points to an offset in the CSKED time wheel array for accesses from the PCI bus. This register must be loaded with the correct offset before the desired data can be read from the CSKED Time Wheel Array Data register. This register will auto-increment after each access to the associated data register. When the last address is accessed, this register wraps to zero.









### **10.32: CSKED Time Wheel Array Data**

This register is used to access the array pointed to by the CSKED Time Wheel Array Pointer.



#### **10.33: CSKED LCD Cache Array Pointer**

The CSKED LCD Cache Array contains the transmit portion of LCDs used by CSKED and SEGBF. It contains 64 32-bit words. It should only be written for diagnostic purposes to test the array. It will power up to all zeros and should be rewritten to zeros after the array has been tested. This register points to an offset in the LCD Cache array for accesses from the PCI bus. This register must be loaded with the correct offset before the desired data can be read from the LCD Cache Array Data register. This register will auto-increment after each access to the associated data register. When the last address is accessed, this register wraps to zero.







 $\downarrow$ 



 $\downarrow$ 

### **10.34: CSKED LCD Cache Array Data**

This register is used to access the array pointed to by the CSKED LCD Cache Array Pointer.





#### **10.35: CSKED Congestion Control Register**

CSKED can halt scheduling on up to 160 groups of connections. A connection specifies which group it belongs to by encoding the QNR(0-39) and DP(0-3) bits in the LCD. This register contains 160 bits which specify which of the 160 groups should be halted. Bit xx in the CSKED control register must be set in order to write to these bits from software. This register is read in eight groups of 20 bits. The group is specified by bits four through two of the address, with increasing addresses corresponding to increasing group numbers.









# **10.36: State Machine Variables**

This register contains the current state of the three main state machines in this entity.



State Machine Variables





#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



# **Entity 11: ATM Transmit Buffer Segmentation (SEGBF)**

The segmentation buffer entity (SEGBF) accepts frames from the cell scheduler (CSKED) or software, and then generates ATM cells to send out over the external physical interface. This entity knows or cares nothing about scheduling cells over time; it will simply construct a cell when it is provided an address of a logical circuit descriptor to operate on. All rate and scheduling concerns must be addressed by the CSKED logic or software prior to queueing a frame to SEGBF.

The SEGBF logic consists of four input LCD address latches, two specialized processors and the associated ROM/RAM for each, a 16-cell array buffer, and various support logic. The input latches and cell buffers are logically divided into four different drops, with each drop consisting of an input latch and four cell buffers. Under normal operation (CSKED providing LCD), the drop is defined by the drop field in the transmit LCD. When an LCD is enqueued by software, data bits five and six of the enqueued address define which drop the enqueued LCD is associated with. The four logical drops in SEGBF can be mapped to any of the physical link level addresses via registers in LINKC (LINKC Map Transmit Ports to Configuration). The processors fetch instructions from ROM/RAM and handle normal segmentation activity such as LCD update operations and cell generation functions. The type of cell that is generated by the segmentation logic is determined by the initial instruction pointer that is contained in the LCD structure. For example, software can enqueue an LCD that has the initial instruction pointer field set for normal AAL5 cells, and SEGBF will generate a single AAL5 cell as a result of the enqueue operation. If, however, software enqueues an LCD that has the initial instruction pointer field set for POS-PHY operation, then SEGBF will continue to generate buffers to pass to the link level until all the data has been exhausted. For a more complete description of the segmentation entry points, refer to the seg\_prc\_Entry\_point field in the transmit LCD data structure section of this document. After the [buffers/cells are built in a 16-by-64 byte array, they are marked as available to the link level layer \(LINKT\) in](#page-302-0)  the IBM3206K0424. A simplified block diagram is shown in the SEGBF Block Diagram.



**Preliminary IBM Processor for Network Resources**

#### <span id="page-302-0"></span>**SEGBF Block Diagram**



The sequence of events that happens when an AAL5 frame is enqueued to SEGBF is as follows:

- 1. An initial check is made to determine if there is space available in the cell buffer. If no space is available, processing stops for this drop until a cell buffer is freed by the link logic (LINKT).
- 2. When buffer space becomes available, the enqueued LCD address is requested from the transmit LCD cache (there are four entries in the LCD cache). The segmentation logic waits for a valid indication from the cache; at this time all LCD information is available to the segmentation logic on the LCD cache interface.
- 3. The initial instruction pointer (IP) is fetched from the LCD and loaded for both of the segmentation processors. Software controls what type of cells are generated by the segmentation logic by initializing this field in the LCD to the entry points defined for different types of cell generation. For a more complete description of the segmentation entry points, refer to the seg\_prc\_Entry\_point field in the transmit LCD data structure section of this document.
- 4. Assuming an AAL5 entry point is setup in the LCD, the segmentation logic will first initiate a memory fetch of the data required to build the cell.
- 5. While the data fetch is in process, the segmentation processors will update various fields in the LCD including statistics and the next segmentation pointer. Cell construction will be started using the ATM

#### **IBM Processor for Network Resources All and Service Construction Construction Preliminary**



header from the LCD.

- 6. When the payload data is available from memory, it is written to the cell buffer following the header.
- 7. If the segmentation logic determines that the current cell being assembled will be the last cell of this frame, the AAL5 trailer is appended to the cell buffer, with any unused bytes being padded with zeros.
- 8. If the current cell is not the last of the frame, the partial CRC is written back out to the LCD.
- 9. After the cell has been completed, it is marked as available to the link level logic



#### **11.1: SEGBF Software LCD Enqueue**

This register provides a mechanism for software to transmit a single cell or a group of cells making up a buffer that can contain any user-defined data at any time. To cause a cell/buffer to be transmitted, the software must write the address of a valid LCD control block to this register. The segmentation hardware will then construct a cell to match the AAL type defined in the LCD control block, using the segmentation pointer contained in the LCD to fetch data and present this cell to the next lower level of hardware to transmit. This method of cell transmission bypasses the cell scheduler completely, so it is the responsibility of the software to ensure that peak and average rates are not violated. When the segmentation logic has completed building the cell/frame and queued it for transmission, the LCD address will be loaded into the software LCD complete register. This method of cell transmission is not designed for high performance and, as such, there is only a single level of queueing underneath the complete register. It is recommended that only a single software LCD be queued to the segmentation logic at any one time to prevent hanging the segmentation logic as it attempts to queue a complete software LCD to the complete queue.











#### **11.2: SEGBF Control Register**

This register provides a mechanism to control the various programmable features of SEGBF. See Note on [Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









### **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**







#### **11.3: SEGBF Status Register**

This register provides feedback to the user on the current status of SEGBF. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing.









### **11.4: SEGBF Invalid LCD Register**

This register provides feedback to the program when the segmentation logic detects an invalid LCD. If multiple invalid LCDs are being processed, this register will contain the address of the last one that was processed by the segmentation logic. There are several invalid LCD situations that the segmentation logic checks for: The first is the LCD address not being on the correct boundary. For example, if the chip is configured to have all LCDs on 128-byte boundaries and an LCD is encountered that is not on a 128-byte boundary. Another invalid condition is when the transmit length configured in the LCD plus the offset in the LCD when added together exceed the maximum overall packet size configured in the chip. It is up to the program to determine which of the possible conditions caused the error to be reported.





#### **IBM Processor for Network Resources All and Service Control of All and Service Preliminary**



#### **11.5: SEGBF Software LCD Complete**

This register provides feedback to the program when the segmentation logic completes cell generation for an LCD that was enqueued by the software. After the segmentation logic has updated the LCD, the address of the LCD is copied into this register providing any previous LCD addresses written to this register have been read by the software. If multiple software queued LCDs are outstanding to the segmentation logic at any time, the segmentation process can be delayed when multiple software enqueued LCDs complete without the software getting a chance to read the LCD addresses from this register. To guarantee that the segmentation logic never has to wait for the software to read this register, it is recommended that only one software LCD be enqueued at any one time.









#### **11.6: SEGBF Interrupt Enable Register**

This register allows the user to selectively determine which bits in the SEGBF status register will cause processor interrupts. A '0' in a bit position masks interrupts from the corresponding bit location in the SEGBF status register. A '1' in a bit position allows interrupts for the corresponding bit in the SEGBF status register. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



#### **11.7: SEGBF Programmable Counters**

This register provides the user with feedback on the number of times that a particular event or condition has occurred in the segmentation logic. The event or condition that causes this counter to increment is defined by the associated SEGBF Programmable Counter Source Specification register. When the counter wraps, an event is generated.



Count of Occurrences





#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **11.8: SEGBF Transmit LCD Size**

This register should be loaded with the number of eight-byte words that are needed for the maximum-sized LCD that will be setup by software. Refer to the previous section describing LCD layout to determine the number of words required to support the different modes. The minimum value is six. This is the correct value when running only AAL5 mode; it includes three words of scheduling information, two words shared between CSKED and SEGBF, and one word for SEGBF to maintain statistics. Setting this register to a value that is too small will likely cause the chip to function improperly. Setting this register to a value that is too large will adversely affect performance.









### **11.9: SEGBF Cell Queue Status**

This register indicates the number of cells queued up for transmission over the media. SEGBF can have a maximum of 16 cells queued up for transmission: four cells on each of four drops. When a bit is set to '1', the corresponding cell buffer contains a cell that has not been completely processed by the link logic.



Incomplete Cell Incomplete Cell Incomplete Cell Incomplete Cell





#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

#### **11.10: SEGBF Processor 1 Control/Status**

Reading this register provides feedback to the user on the current state of segmentation processor 1. Writing the appropriate bits in this register causes processor 1 to begin executing at a new location specified in the data that was written.









#### **11.11: SEGBF Processor 2 Control/Status**

Reading this register provides feedback to the user on the current state of segmentation processor 2. Writing the appropriate bits in this register causes processor 2 to begin executing at a new location specified in the data that was written.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **11.12: SEGBF Programmable Counter Source Specification**

This register determines what event or condition will cause the associated counter to increment.







#### **11.13: SEGBF Cell Staging Array Pointer**

This register points to an offset in the SEGBF cell staging array for accesses from the PCI bus. This register must be loaded with the correct offset before the desired data can be read from the SEGBF Cell Staging Array Data Register. This register will auto-increment after each access to the associated data register. When the last address is accessed, this register wraps to zero.







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **11.14: SEGBF Cell Staging Array Data**

This array is divided into 16 64-byte buffers used to assemble cells that are ready for transmission on the physical interface.



#### **11.15: SEGBF Instruction SRAM Pointer**

This register points to an offset in the SEGBF SRAM that is used to store processor instructions. This register must be loaded with the correct offset before the desired data can be read/written from/to the SEGBF instruction SRAM data register. This register will auto-increment after each access to the associated data register. When the last address is accessed, this register wraps to zero. The first 256 locations in the array should be loaded with the instructions for processor 1, and the second 256 locations should be loaded with the instructions for processor 2.









### **11.16: SEGBF Instruction SRAM Data**

This register address can be used to read/write the instruction data that is needed by the segmentation processors. All instructions must be written before the processors can be brought out of the halt state.



#### **11.17: MPEG-2 PCR Increment Register**

Each tick of the time base will add the contents of this register to the MPEG PCR Reference Register. This register contains a fixed point number with 27 bits of fraction and five bits of units. This means that the external reference clock can range in speed from 22.5 Khz to the maximum speed of this entity which is TBD. Assuming that the entity will run with a 50 Mhz clock, the conversion to 720Khz can be done with an accuracy of 1.1 parts in two million. (A clock of 19.4 Mhz will give a conversion accuracy of one part in 4.9 million.) If the input clock is 19.4 Mhz, then the value to put in the increment register is (720,000 / 19,400,000) \* 2\*\*27 or 4,981,277. If the input clock is 33 Mhz, then the value to put in the increment register is (720,000 / 16,666,666) \* 2\*\*27 or 5,798,206.









# **Receive Data Path Entities**

# **Entity 12: Cell/Packet Re-assembly (REASM)**

REASM is the top level receive entity that encapsulates all of the receive sub-entities.

**Note:** The receive portion of the chip is very different from previous versions of the processor. REASM no longer does HEC correction or detection, since everything we interface to already does this function.

The following figure shows how REASM interacts with the other entities:

### **REASM Entity Interfaces**





REASM is made up of a number of sub-entities to provide the overall receive functionality. REASM contains the following sub-entities:



# **REASM Sub-Entity Block Diagram**



RXXLT, RXCRC, and RXAAL form a cell processing pipeline. Each stage is allowed up to a cell time to process the cell it is currently working on. This means each cell has the potential to have a three-cell time latency through the REASM entity. The overall performance should be at its maximum since a single cell completes processing every cell time. The stages are allowed to run faster if there is no blocking condition. The only blocking condition is the latency of memory accesses and the total length of the nano-program that needs to run.

The following sections provide a brief description of each sub-entity and the function it provides.

**IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary** 



#### **Miscellaneous Reassembly Functions**

#### **ATM OAM Cell Processing**

[When processing an ATM data stream, OAM cell processing may be necessary. This function needs to be](#page-329-0)  [enabled in the R](#page-329-0)EASM Reassembly Modes Register. REASM performs OAM discrimination when enabled. This function can be enabled on a per port basis. As cells arrive for processing, the nano-code works together with some dedicated logic to discriminate between different types of OAM cell traffic. Each type that is discriminated produces a different event when traffic is surfaced to a receive queue. The different types that are discriminated are:



Both the RXCRC and RXAAL nano-code understand and process OAM traffic differently when enabled. In the receive LCD there are control bits that allow the user to specify on a per connection basis how the OAM traffic should be handled. There are two types of flows when processing OAM traffic. A connection can be either "routed" or "terminated." A terminated connection is any connection that terminates either cells or packets. AAL5 and packet LCDs that are fast forwarded are considered "terminated" because the packet is terminated before the packet is fast forwarded. A "routed" connection is a raw LCD that is fast forwarded including raw mode with early packet discard. When OAM traffic is processed on a "terminated" connection, the OAM traffic can be terminated or dropped based on the discrimination configuration bits in the receive LCD. A mask value of '0' specifies to terminate the cell, and a mask value of '1' specifies to drop the cell. When OAM traffic is processed on a "routed" connection, the OAM traffic can be fast forwarded or dropped based on the discrimination configuration bits in the receive LCD. A mask value of '0' specifies to fast forward the cell, and a mask value of '1' specifies to drop the cell. This allows the user to filter the OAM traffic based on the type of OAM traffic on a per connection basis.

The mask bit is determined by using the sub-type mask bit if pertinent. For example, an F5 Segment Pm Cell will use the "Segment Pm Cell" mask bit. If one of the sub-types does not match, then the basic types are



used (that is, F4 End-to-end). There is a 16-bit mask field in the receive LCD. A description of each bit is in the table below:



When an OAM cell is processed, the configuration information from the gpmtagE is used to make reassembly decisions like which buffer pool or receive queue to use.

### **TCP/IP Receive Checksum Verification**

When enabled, the IBM3206K0424 is able to perform verification of the IP header checksum and the associated protocol checksum for the user. The final checksum status is raised to the user in the packet header flags. This function is accomplished by the RXCRC entity using the CRC nano-program.

To enable and use the function, several things must be done:

- TCP/IP checksum function must be enabled in the REASM Mode Register. This enables the overall function and includes the checksum state words in the receive portion of the LCD.
- Specify a frame type in the frameType field of the receive LCD. This field specifies an entry point into the checksum verification nano-code. The entry point points to a piece of code that understands how to locate the IP header based on the current frame type. For example, a connection might be native IP over ATM or LANE ETH. For the currently supported types, see the crc.txt file provided with the RXCRC code load.

**Note:** Other types can be added if needed.

Once properly enabled, TCP/IP checksums are verified as packets arrive. The status is raised to the user in the packet header using four bits of status (two bits for IP and two bits for protocol). The format and meanings of the bits are specified in Packet Header. For connections that have mixed traffic, the checksum operation is only run on packets that are recognized as TCP/IP. For example, on a LANE ETH connection packets of different protocols can be interleaved.

#### **IBM Processor for Network Resources All and Security Construction Construction Preliminary**



#### **Scatter/Cut Through Receive Processing**

The scatter/cut through support is very versatile and easy to use. Scatter typically refers to scattering pages from a contiguous IBM3206K0424 buffer into multiple non-contiguous host pages. Cut through refers to moving contiguous data from IBM3206K0424 buffers into contiguous storage in host memory. The functions are similar, and the terms scatter and cut through may be interchanged in the following text since cut through is really a special case of scatter. The correct term is used when context requires it.

The scatter processing is performed by RXAAL along with RXQUE and DMAQS. The following items need to be setup:

- DMAQS should be setup and ready to run.
- RXQUE should be setup and ready to run. Each queue that is used should have the proper event size selected, the direction should be set correctly, and timestamps and bcach advice should be disabled.
- Page buffers or descriptors should be placed on the configured receive queues.
- The scatter configurations need to be set in the RXALL Scatter/Cut Through Info Registers.
- The scatter flags need to be set in the RXALL Scatter/Cut Through Flag Registers.
- Each LCD that is to use scatter, needs to have the ppMode field set to do scatter.
- A scatter configuration needs to be selected for each connection by setting the cutThruSel field in the receive LCD.

Once set up, the scatter mechanism is performed by RXAAL for the user. There is user intervention required to process the packets when the scatter is complete, and for page recovery in error scenarios.

There are many options when setting up scatter, and the selection of the type of scatter processing to perform depends on the users environment. The basic scatter mechanism will be described followed by a discussion of the different options that can be used.

When packets are received and scattered, the following structure shows the main components of the received packet:

#### **General Layout of a Received Packet in Scatter Mode**

```
struct ScatterRxPacket {
  packetHeader; // Charm packet header
 dmaList; // Scatter dma list<br>padbytes; // Pad out to recei
                // Pad out to receive offset from lcd
  packetData; // Actual receive packet
};
```
The DMA list is maintained immediately after the IBM3206K0424 packet header, and before the receive packet data. For this reason, the user should allocate enough space between the packet header and the packet data to accommodate a maximum-sized DMA list based on the maximum number of pages that could be DMAed. The rxOffset field in the receive LCD is where this offset is specified. The maximum receive offset is 256 bytes, which is specified with an rxOffset value of '0'. The following is the layout of the DMA list:


### **General Layout of a Scatter DMA List**

```
struct dmaList {<br>bit16 numHeadbytes:
                                        \frac{1}{16} number of bytes included with header dma
     bit16 numTailbytes; // number of bytes in last dma page
    bit1 deqLocked; // error status<br>bit1 deqInvalid; // "
    bit1 deqInvalid; \frac{1}{2} // "<br>bit1 headerTruncated: // "
    bit1 headerTruncated; //<br>bit1 badDmaI.ist: //"
    bit1 badDmaList;
    bit4 DeqLockedQueueNum; // "
    bit16 reserved;<br>bit2 cutThruSel;
    bit2 cutThruSel; // copy of cutThruSel used from receive lcd<br>bit6 numPages; // number of page entries that follow (max=
                                    \frac{1}{10} number of page entries that follow (max=63)
    bit32 pageList[N]; // Actual dma descriptors or page addresses
                           // Note: each entry can be 3\overline{2}, 64, or 1\overline{2}8 bits wide
  };
```
The first eight bytes are always present and are filled in when the packet is complete or when an error occurs. The actual page list is filled in as the data is DMAed. The first location is initially skipped and is filled in later when the header is DMAed. The second and subsequent entries are filled as each page is DMAed. Each page list entry contains either physical page addresses, IBM3206K0424-based DMA descriptor addresses, or user data. Whether a physical page address or DMA descriptor address is present depends on the cut through configuration. From here on, the term page address and DMA descriptor are used interchangeably as either is valid based on the configuration, but the correct term is used when the context requires it.

A page list entry can contain one or two pieces of information. Each page list entry can be 32, 64, or 128 bits long. If using 32 bit addresses, then the entry is 32 or 64 bits. If using 64-bit addresses, then the entry is 64 or 128 bits. The first piece is always the page address or the DMA descriptor address. Each cut through configuration allows the user to specify an optional second deque operation from the receive queue being used. This allows the user to place user information associated with the particular page address in the receive queue and the page list. Thus, the corresponding virtual address of a page could be surfaced along with the physical page address. It is very important to enque information to the receive queue in the proper order if using this mode. The physical page address is first, followed by the user information.

As the receive packet data is being received, it is DMAed as soon as a page crossing occurs if there is a DMA descriptor available on the receive queue being used. If no descriptor is available, then no DMA takes place until the next receive cell is received, at which time the receive queue availability is checked again. This catch-up process continues until the packet is complete. When the packet is complete, all DMAs are scheduled if page addresses are available. If a page address is not available, then a "no DMA descriptor available" is surfaced to the user so the packet can be used and the DMA list recovered.

As each data page is DMAed to the user, a DMA descriptor is formed and enqueued to the DMAQS DMA queue specified in the cut through configuration. The DMA descriptor is formed based on if page addresses or DMA descriptor are provided on the receive queue being used. If IBM3206K0424-based DMA descriptors are being used, then the receive queue contains DMA descriptor chains where the low order bits of the DMA descriptor address specify how many descriptors are in the chain. Typically, this chain length is one, but more can be used. The first descriptor in the chain provides the destination address (page address) which is filled in by the user. The source address, the length, and the flags are filled in by the IBM3206K0424 for the first descriptor in the chain. The source address is filled in with the beginning address of the page within the current receive buffer. The length is filled in with either the page size (if a full page is present) or the number of bytes in the last page for the last page. The flags are filled in using the flags from the appropriate RXALL - Scatter/Cut Through Flag Registers. If physical page addresses are contained in the receive queue, then a

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



single DMA descriptor is formed by DMAQS directly in DMA queue storage using the page address and the same information that would have been filled into the DMA descriptor. Generally, using page addresses performs better, but is less versatile, which is usually a good trade off. Normally no DMA event is generated in the flags when pages are DMAed.

When the last cell of a packet is received, and all the data pages have been DMAed, the packet header and DMA list are updated and DMAed into a header buffer. The mechanism and configuration of the header buffer is similar to the pages, but separate configuration is usually necessary for correct functionality. For example, different flags are normally used in order to get an event for the header DMA so the user can process the packet. The header DMA normally frees the IBM3206K0424 buffer; another difference is the page (or buffer) size used for headers is normally different than the normal page size. Some of the optional features described later also drive having different configuration for the header DMAs.

Once the user gets the event for the header DMA, the user processes the received packet using the DMA list and the packet header. Once the packet has been processed, the pages or descriptors need to be returned to the proper receive queue when the pages can be reused, thus completing the scatter processing.

There are several ways of getting that important event to start the receive processing. Usually the event is generated when the last header DMA is complete. If DMA descriptors are being used, then there are two choices. First, the normal DMA flag that generates an event can be used. This generates an event with the DMA descriptor address in the significant bits. While this works and may be desirable in some environments, the DMA descriptor needs to be read in order to get access to the host header buffer address. The second way to generate an event when using descriptors is to provide a second DMA descriptor in the DMA descriptor chain that enqueues the header buffer address and a user-defined event in the lower order bits. Generating an event in this manner provides the user with the buffer address; the header DMA descriptor address is available in the header buffer as part of the DMA list.

When using page addresses, the event source in the cut through configuration should be set to use the destination address as the event data. When this is done, the event data contains the header buffer address as in the second case above.

### **Error Recovery**

There are two types of error that need to be handled. First, if there is no error on the receive packet and a page address was not available, then the packet is surfaced to the user with a "no DMA descriptor available" event in the event type field and the IBM3206K0424 buffer address in the event information field. The user has a choice at this point. The packet is good so it can be used or freed by the user. In either case, the DMA list in the packet header must be recovered. So if the DMA list in the header is not used, it should be recovered by returning it to the proper receive queues. The event surfaced specifies which type of page address failed so the user can parse the DMA list properly. The event will specify whether a normal page descriptor, optional header descriptor, or packet header descriptor. The same descriptor recovery must be performed when an error event is surfaced (that is, CRC error). When there is an error event, no packet header DMA is performed so no packet header descriptor is in the DMA list.

### **Scatter Options**

Most of the optional scatter features have to do with the header bytes and how the header DMA is performed. How the number of header bytes is determined is explained below. For now, just assume there are some number header bytes. The numHeadbytes field in the DMA header specifies the number of header bytes that are available. The location of the header bytes in the DMA buffers can be configured. The default is for them to be kept with the packet header and DMA list in the packet header buffer. For this case, the user should be sure the packet header buffer size is large enough to contain all of this data.

Alternatively, the header bytes can be placed in a separate buffer. To do this, split header mode should be



enabled in the cut through configuration. This buffer is referred to as the optional header page and uses its own page size and receive queue to allow for better storage utilization. When enabled, the second entry in the DMA list becomes the optional header page entry, and should be treated accordingly for page recovery. In split header mode, the header bytes are DMAed as soon as they are all received. This mode is useful if the user environment requires the header bytes to be in a separate buffer from the packet header and DMA list.

Another optional feature is to DMA the header only. This feature is enabled in the cut through configuration. When enabled, only the packet header, DMA list, and header bytes are DMAed. Either a single DMA or two DMAs occur based on if the optional header feature is enabled. This feature can be useful when a routing decision needs to be made for a packet and the entire packet does not need to be brought into host storage. Another possible scenario is the header bytes may determine how the user wants to DMA the packet data to the host.

For smaller packet sizes, it may be more efficient to perform a single DMA and keep the packet header, DMA list, and packet data in a single buffer. To do this, single page mode should be enabled in the cut through configuration. When enabled, the header page size is used to determine the DMA behavior. If a packet completes and the total length of the packet and headers will fit in the header page size, then a single DMA is performed. If the data length exceeds the single page size as it is being received, the data is scattered using the normal options (might need to catch up). This feature can be useful for optimizing user processing for smaller packets. For example, all packets less than 2K might be a candidate for this feature.

# **Head Bytes**

There are two ways to set the number of head bytes. They are set on a per connection basis in the receive LCD using the numHeadbytes and useCrcNumHead fields. When the useCrcNumHead field is set to '0', the numHeadbytes field provides a fixed number of bytes that is used as the number of header bytes. When useCrcNumHead field is set to '1', then the RXCRC nano code will calculate the number of header bytes using the frameType field to index an IP procedure. Currently, RXCRC will only set the number of header bytes for recognized TCP/IP headers. Other headers can be recognized. Contact IBM technical support to discuss requirements.

#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **12.1: REASM Logical Channel Descriptor Base Register**

The REASM Logical Channel Descriptor Base Register indicates the starting address of the logical channel descriptor table. This register defines where the Logical Channel Descriptors are located.











# **12.2: REASM Mode Register**

Used to set REASM and sub-entity modes. This register contains the mode bits that specify how REASM is to operate. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









## **12.3: REASM Reassembly Modes Register**

Used to set reassembly modes. This register contains the mode bits that specify different reassembly modes. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









# <span id="page-330-0"></span>**12.4: REASM Status Register**

Used to surface REASM and sub-entity status. This register contains the status bits for the REASM sub-entities. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









## **IBM Processor for Network Resources All and Service Construction Construction Preliminary**

# **12.5: REASM Interrupt Enable Register**

Used to enable interrupts for REASM status conditions. When set, the corresponding status condition generates an interrupt from REASM to INTST. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.





## **12.6: REASM DEBUG State Selector Register**

Selects which entity states are surfaced. This register specifies which entity states are surfaced. Use this register only under the advice of IBM technical support.





## **RXBUF Block Diagram**



## **RXBUF Functional Description**

RXBUF provides the cell/packet buffering mechanism between LINKR and the rest of REASM. The buffering provides enough storage for 16 64-byte cells and the associated control information for each cell buffer. LINKR writes the cell/packet data and control information into the buffers. The buffering mechanism is then exposed to the remainder of the REASM sub-entities for reading.

The cell buffers are sequenced through in sequential order regardless of which port received cells arrive on. Thus, a total of 16 cells of buffering exists and is used as a shared 16-cell FIFO.

### **12.7: RXBUF Cell Data Buffer Address**

Provides the read/write address for accessing cell data buffer. This register provides the read/write address for accessing cell data buffer. When RXBUF Cell Data Buffer Read/Write Port is read/written, this register is auto-incremented. When the last address is read/written, the address rolls over to zero.





#### **12.8: RXBUF Cell Data Buffer Read/Write Port**

Provides read/write access to receive cell data buffer. The array is divided into 16 64-byte buffers used to buffer/assemble cells received from the line. When this register is read/written, the address provided by RXBUF Cell Data Buffer Address is used to select the array word to be accessed. RXBUF Cell Data Buffer Address is auto-incremented on each read/write. Thus, this port can be read/written multiple times to read/write the entire array.



## **12.9: RXBUF Cell Info Buffer Address**

Provides the read/write address for accessing cell information buffer. This register provides the read/write address for accessing cell information buffer. When RXBUF Cell Info Buffer Read/Write Port is read/written, this register is auto-incremented. When the last address is read/written the address rolls over to zero.



# **12.10: RXBUF Cell Info Buffer Read/Write Port**

Provides read/write access to the Receive Cell Info buffer. The array is divided into 16 areas used to provide information about each received cell from LINKR. When this register is read/written, the address provided by RXBUF Cell Info Buffer Address is used to select the array word to be accessed. RXBUF Cell Info Buffer Address is auto-incremented on each read/write. Thus, this port can be read/written multiple times to read/write the entire array.





### **12.11: RXBUF Receive Buffer Threshold**

Provides a method to monitor number of cell buffers in use. The value of this register is used to compare against the number of active cell buffers. When this threshold is exceeded, the status is raised in the REASM status register.





## **RXXLT Block Diagram**



## **RXXLT Functional Description**

RXXLT is the first stage in the cell processing pipeline. RXXLT provides general LCD translation facilities and link level statistics. These are provided via a nano-processor and nano-programs.

Up to four unique drops/ports can be supported, each with a unique configuration. Each port is able to run a unique nano-program to do LCD address translation. For example, one port may be a packet-based PHY using 64-byte segments with PPP LCD translation, and another port may be an ATM cell based PHY. The drop/config number (0-3) along with the port id is passed to RXXLT from RXBUF/LINKR. The drop number (0-3) is used to specify which nano-program is executed by RXXLT to translate cell/packet information into an LCD address. This is different from previous versions of the processor which had a fixed LCD translation mechanism. RXXLT instruction formats are not defined here and are IBM Confidential.

RXXLT uses the resulting LCD address to load the LCD cache for the next stage of the pipeline. The LCD address and cell buffer address are passed to RXCRC upon completion of processing.

There are a number of degrees of freedom in the LCD translation. Generally, the nano program performs the following steps:

- Gather some bytes from the cell buffer
- Do some error checking and default LCD checking
- Select and shift appropriate bits to form an LCD translate table index
- Read the LCD translate table to get an LCD index
- Generate the LCD address from the LCD index and the LCD base address

There are eight general purpose registers per PHY drop and four drops. These eight registers contain values that the nano-code uses, and are available on a per port basis. For example, the LCD translate table addresses would reside in these registers (this is different from previous versions of the processor where these registers were at fixed addresses). Other items that might reside in these registers are default/error LCD addresses, compare values, and masks. Because these are general purpose registers, multiple LCD tables or multiple default LCDs can be specified. One quarter of the total registers are available to each port, so there can be an LCD translate table (etc.) for each port.

The total LCD translate tables and LCD table sizes are only limited in size by the amount of memory that is available to the IBM3206K0424. The LCD indexes are limited to 16 bits, and LCD addresses are always 128-byte aligned.

The LCD translation code must execute in one cell time in order to run at full bandwidth. The only variable portion of the code execution time is the reads to IBM3206K0424 memory when reading the LCD translation



table. Thus, while a double lookup is possible, it may not meet the time constraints of running at full bandwidth.

The following pseudo code shows some of the types of LCD translations that are possible using the nano-processor.

**Note:** The following do not imply types/numbers of instructions needed, but the different variable names imply different general purpose registers are being used, so variables are specified on a per port basis.

In the following code there are shift, mask, and compare values that are not specified. These are values that would be customizable at run time. These will be customized via the general purpose registers.

The following code uses nomenclature that matches the IBM3206K0424 (zero-based big endian). So bit(1) in a comm spec might be bit(0) in the following code. The comments use spec nomenclature so you can match the two up.

# **Standard ATM**

 $atmH = read 4 bytes from rxbuf at offset 0 - read the atm header from cell$ 

```
if non_user_data { -- non-user data as specified in pti<br>lcdAddr = defaultLcd -- field is handled by the deafult le
                                                   -- field is handled by the deafult lcd
return
}
tblIndex = (atmH and vciMask) >> 4 -- gather pertinent vci bits<br>tblIndex = ((atmH and vpiMask) >> X) -- gather pertinent vpi bits
tblIndex | = ((atmH and vpiMask) >> X)-- X depends on num of vci bits used
if non-masked bits on {<br>
-- check for out of range<br>
out of range counter++<br>
-- update counter
out of range counter++<br>if mode(X) {
if mode(\overline{X}) { -- if configured to flush out of flush cell<br>flush cell -- range cells flush and we are don
                                        -- range cells flush and we are done
}
else {
lcdAddr = errorLcdAddr
}
return
}
if tblIndex == 0 { -- check for zero id zero id counter ++ --- update counter
zero id counter++<br>if mode(X) {
if mode(X) {\qquad \qquad \qquad - if configured to flush zero id flush cell \qquad \qquad - cells flush and we are done
                                        -- cells flush and we are done
}
else {
lcdAddr = errorLcdAddr
}
return
}
lcdIndex = Lookup(lcdTable, tblIndex) -- read index from cm using tableIndex lcdAddr = lcdBase + lcdIndex*128 -- calc lcd addr from index
lcdAddr = lcdBase + lcdIndex * 128
```
#### **IBM Processor for Network Resources All and Security Construction Construction Preliminary**



### **PPP**

```
label = read 4 bytes from rxbuf at offset 0 -- read the ???<br>if label == 0xff030021 {<br>-- what is this field
if label == 0xff030021 {<br>read 1 bytes from rxbuf at offset 4 -- read the ip header version
                                                 -- read the ip header version (5th byte)
if IPVER_HEADERLGT == 0x45 {
tblIndex = read byte from offset 5 - read OOS field (6th byte)
lcdIndex = Lookup(lcdTable, tblIndex) -- read index from cm using tableIndex \text{ledAddr} = \text{ledBase} + \text{ledIndex} * 128 -- calc lcd addr from index
lcdAddr = lcdBase + lcdIndex * 128}
else {
lcd\hat{A}ddr = defaultPortLcd -- use port default lcd}
}
else {
lcdAddr = defaultPortLcd -- use port default lcd
}
```
# **Q.922 2 Byte Addressing**

```
head = read 2 bytes from rxbuf at offset 0 -- read the header (network bytes 0-1)
if (bit(0) == 1) \& (bit(8) == 0) { -- EA bits indicate two byte addr
- byte 0 bit 1 = 0 & & byte 1 bit 1 == 1tblIndex = (head and dlciMask0) >> X - gather 6 of 10 dlci bits from byte 0
tblIndex |=(\text{head and } \text{dlciMask1}) \gg X) -- gather 4 of 10 dlci bits from byte 1
lcdIndex = Lookup(lcdTable, tblIndex) -- read index from cm using tableIndex lcdAddr = lcdBase + lcdIndex * 128 -- calc lcd addr from index
lcdAddr = lcdBase + lcdIndex * 128}
else {
lcdAddr = errorLcdAddr -- not a two byte addr
}
```
# **Q.922 4 Byte Addressing**

```
head = read 4 bytes from rxbuf at offset 0 -read the header (network bytes 0-3)
if (bit(0) == 1) & \& (bit(8) == 0) & \& -- EA bits indicate four byte addr
(bit(16) == 0) & & (bit(24) == 0) { -- byte 0 bit 1 == 0 & & byte 1 bit 1 == 0 & &
-- byte 2 bit 1 == 0 & & byte 3 bit 1 == 1<br>if dcBit == 1 { -- check f
if d\tilde{c}Bit = 1 {<br>- check for dc bit being set lcdAddr = defaultPortLcd<br>-- and surface on default
                                     -- and surface on default port lcd
}
else { -- gather 16 least significant dlci bits
tblIndex = (head and dlciMask1) >> X - gather 3 dlci bits from byte 1
tblIndex = ((head and dlciMask2) >> X) -- gather 7 dlci bits from byte 2
tblIndex = ((head and dlciMask3) >> X) - gather 6 dlci bits from byte 3
lcdIndex = Lookup(lcdTable, tblIndex) -- read index from cm using tableIndex
lcdAddr = lcdBase + lcdIndex*128 -- calc lcd addr from index
}
}
else {
lcdAddr = errorLcdAddr -- not a four byte addr
}
```




**Preliminary IBM Processor for Network Resources**

## **FUNI 2.0 2 Byte Addressing**

```
head = read 2 bytes from rxbuf at offset 0 -- read the header (network bytes 0-1) if (bit(0) == 1) & \& (bit(8) == 0) { -- EA bits indicate two byte addr
                                                -- EA bits indicate two byte addr
- byte 0 bit 1 = 0 & & byte 1 bit 1 = 1
if (bit(2) == 0) && (bit(9) == 0) { \longrightarrow FID1 & FID2 == 0 (byte 0 bit 2 & byte 1 bit 3)
tblIndex = (head and faMask1) >> X - gather 6 of 10 FA bits from byte 0
tblIndex |=(\text{(head and } \text{faMask0}) \gg X) -- gather 4 of 10 FA bits from byte 1
lcdIndex = Lookup(lcdTable, tblIndex) -- read index from cm using tableIndex lcdAddr = lcdBase + lcdIndex*128 -- calc lcd addr from index
lcdAddr = lcdBase + lcdIndex * 128}
else {<br>lcdAddr = defaultPortLcd
                                           -- otherwise surface on port default lcd
}
}
else {
lcdAddr = errorLcdAddr -- not a two byte addr
}
```
# **FUNI 2.0 4 Byte Addressing**

```
head = read 4 bytes from rxbuf at offset 0 -- read the header (network bytes 0-3)
if (bit(0) == 1) && (bit(8) == 0) && -- EA bits indicate four byte addr
(bit(16) == 0) & & (bit(24) == 0) { -- byte 0 bit 1 = 0 & & byte 1 bit 1 = 0 & &
-- byte 2 bit 1 == 0 & & byte 3 bit 1 == 1<br>if (bit(18) == 0) & (cit(25) == 0) {
                                             \text{- FID1} \& \text{FID2} == 0 \text{ (byte 0 bit 2 } \& \text{ byte 1 bit 3)}-- gather 16 vpi/vci bits
tblIndex = (head and vciMask0) >> 1 -- gather X vci bits from byte 3
tblIndex = (head and vciMask1) >> X -- gather X vci bits from byte 2
tblIndex = (head and vciMask2) >> X - gather X vci/vpi bits from byte 1
tblIndex = (head and vpiMask3) >> X -- gather X vpi bits from byte 0
lcdIndex = Lookup(lcdTable, tblIndex) -- read index from cm using tableIndex lcdAddr = lcdBase + lcdIndex*128 -- calc lcd addr from index
lcdAddr = lcdBase + lcdIndex * 128}
else {
lcdAddr = defaultPortLcd -- otherwise surface on port default lcd
}
}
else {<br>lcdAddr = errorLcdAddr-- not a four byte addr
}
```




### **12.12: RXXLT Register Array Address Port**

Provides the read/write address for accessing register array. This register provides the read/write address for accessing register array. When RXXLT Register Array Read/Write Port is read/written, this register is auto-incremented. When the last address is read/written the address rolls over to zero.



### **12.13: RXXLT Register Array Read/Write Port**

Provides read/write access to the register array. The array is divided into four groups of eight registers. Each group is associated with a receive port (0-3). So each nano-program has eight registers to use. These registers have intended uses and also serve as the link level counters, so they are not general purpose registers.

When this register is read/written, the address provided by RXXLT Register Array Address Port is used to select the array word to be accessed. RXXLT Register Array Address Port is auto-incremented on each read/write. Thus, this port can be read/written multiple times to read/write the entire array.





## <span id="page-340-0"></span>**12.14: RXXLT Processor State Selector**

Allows user to select which data should be accessed with RXXLT Processor State Read/Write Port. This register provides the encoded selector for accessing internal processor registers and state via reads/writes to the RXXLT Processor State Read/Write Port.





# **12.15: RXXLT Processor State Read/Write Port**

Provides read/write access to the internal state of the processor. The internal processor state is externalized for debug and testing reasons. See the description of [RXXLT Processor State Selector](#page-340-0) for definitions on the addresses.





#### **12.16: RXXLT Instruction Array Address Port**

This register provides the read/write address for accessing the instruction array. When RXXLT Instruction Array Read/Write Port is read/written, this register is auto-incremented. When the last address is read/written, the address rolls over to zero.



### **12.17: RXXLT Instruction Array Read/Write Port**

Provides read/write access to the instruction array. The instruction array is divided into four groups of 32 entries. Each group is associated with a receive port (0-3) so a nano-program can be loaded for each active port. When this register is read/written, the address provided by RXXLT Instruction Array Address Port is used to select the array word to be accessed. RXXLT Instruction Array Address Port is auto-incremented on each read/write. Thus, this port can be read/written multiple times to read/write the entire array.





# **12.18: RXXLT Last LCD Index Register**

These registers provide the previous LCD index that was used for the corresponding port.





## **RXCRC Block Diagram**



## **RXCRC Functional Description**

RXCRC is the second stage in the cell processing pipeline. It performs the ATM CRC-32 (Ethernet FCS) and ATM CRC-10 functions if necessary. RXCRC gets LCD type, state, and seed information from the LCD cache, and updates the cache on completion. The results are passed to RXAAL upon completion, along with the LCD address.

## **12.19: RXCRC Instruction Array Address Port**

This register provides the read/write address for accessing the instruction array. When RXCRC Instruction Array Read/Write Port is read/written, this register is auto-incremented. When the last address is read/written, the address rolls over to zero.





## **12.20: RXCRC Instruction Array Read/Write Port**

Provides read/write access to the instruction array. The instruction array contains a single nano-program. When this register is read/written, the address provided by RXCRC Instruction Array Address Port is used to select the array word to be accessed. RXCRC Instruction Array Address Port is auto-incremented on each read/write. Thus, this port can be read/written multiple times to read/write the entire array.



## <span id="page-344-0"></span>**12.21: RXCRC Processor State Selector**

Allows user to select which data should be accessed with RXCRC Processor State Read/Write Port. This register provides the encoded selector for accessing internal processor registers and state via reads/writes to the RXCRC Processor State Read/Write Port.







### **12.22: RXCRC Processor State Read/Write Port**

Provides read/write access to the internal state of the processor. The internal processor state is externalized for debug and testing reasons. See the description of [RXCRC Processor State Selector](#page-344-0) on page 345 for definitions on the addresses.



### **12.23: RXCRC Last LCD Index Register**

These registers provide the previous LCD index that was used for the corresponding port.



### **12.24: RXCRC Checksum Protocol Registers**

These registers provide additional protocol bytes for which checksum calculations should be enabled. The first register allows up to four protocols to be enabled with headers that are similar to UDP and TCP which use a pseudo-header. The second register allows up to four protocols to be enabled with headers that are similar to ICMP (no pseudo header). IP, UDP, TCP, V4 ICMP, and V6 ICMP are automatically recognized and should not be specified again in these registers.

Each byte specifies a different protocol.





# **RXAAL Functional Description**

The following is a block diagram of RXAAL:

## **RXAAL Block Diagram**



RXAAL performs the cell and packet reassembly functions. This includes the AAL processing and moving cell data and packet headers to Packet Memory.

The nano-program uses state and configuration informations from the LCD cache to perform the necessary function for each cell. The nano-processor is capable of executing programs to run the following types of reassembly:

- AAL5
- AAL3/4
- Raw cells
- Non-user data (this might be the same as raw)
- Packets
- MPEG FIFO Mode

RXAAL also performs the cell/packet post processing step. This includes event generation to RXQUE, cut through processing, scatter processing, and DMA enqueues.



## **12.25: RXAAL Instruction Array Address Port**

This register provides the read/write address for accessing the instruction array. When RXAAL Instruction Array Read/Write Port is read/written, this register is auto-incremented. When the last address is read/written, the address rolls over to zero.



### **12.26: RXAAL Instruction Array Read/Write Port**

Provides read/write access to the instruction array. The instruction array contains a single nano-program. When this register is read/written, the address provided by RXAAL Instruction Array Address Port is used to select the array word to be accessed. RXAAL Instruction Array Address Port is auto-incremented on each read/write. Thus, this port can be read/written multiple times to read/write the entire array.





## <span id="page-348-0"></span>**12.27: RXAAL Processor State Selector**

Allows user to select which data should be accessed with RXAAL Processor State Read/Write Port. This register provides the encoded selector for accessing internal processor registers and state via reads/writes to the RXAAL Processor State Read/Write Port.





# **12.28: RXAAL Processor State Read/Write Port**

Provides read/write access to the internal state of the processor. The internal processor state is externalized for debug and testing reasons. See the description of [RXAAL Processor State Selector](#page-348-0) on page 349 for definitions on the addresses.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **12.29: RXAAL Last LCD Index Register**

This register provides the previous LCD index that was used.









## **12.30: RXAAL Transmit Queue Length Compression Configuration**

This register allows the user to configure how the transmit queue lengths should be compressed for use in the receive packet header. CSKED provides twelve transmit queue lengths specified in bytes. A 32-bit register (the [Bytes Queued Counters](#page-292-0) in CSKED) is available for the high, medium, and low priority queues for each of the four PHY ports. Using the full counts in the receive packet header generally uses too much room. This register allows the user to configure how this information should be compressed for use in the receive packet header.





#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **12.31: RXAAL Packet Header Configuration**

Allows user to configure the contents of each optional packet header word, and specify how many optional packet words are used.

This register configures the contents of each optional packet header word in the optional portion of the packet header. There are four possible configurations, and the configuration used is selected with the packHeadSel field in the receive LCD.

The first three words of the packet header are fixed, and up to seven additional words can be configured. The low nibble of this register specifies how many optional packet header words are used, and the remaining nibbles of the register configure each packet header word if used.

**Note:** The base receive and transmit packet headers must be compatible if internal cell or packet routing is being used. The receive packet header becomes the transmit packet header in this scenario.

User 0 and user 1 values can be used to place non-standard values in the packet header. These values are built by the nano-code, and are then placed in the packet header. In order to use these values, the nano-code must be customized. Do this only under the advisement of IBM technical support.







### **Preliminary IBM Processor for Network Resources**



# **12.32: RXAAL Error Count Register**

Maintains a count of detected errors. This register maintains a count of error conditions that are detected. For example, CRC errors and other protocol types of errors are counted. This count is useful when the chip is configured to only surface good packets.

When this counter overflows, a counter overflow event is generated to RXQUE.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **12.33: RXAAL Dropped Count Register**

This register maintains a count of packets that are dropped due to a lack of resource. For example, if no POOLS buffer (real or virtual mode) is available, the packet is dropped and this counter is incremented. This count is useful when the chip is configured to only surface good packets.

When this counter overflows, a counter overflow event is generated to RXQUE.



#### **12.34: RXAAL Maximum SDU Length Register**

Specifies the maximum SDU size for a packet. This register contains the maximum SDU size for a packet. This size includes only the protocol data length of the packet. For example, this length would be compared with the AAL5 length field in the AAL5 trailer. When a packet is completely received, this register is used to make sure it does not exceed the MSDU specified.



# **12.35: RXAAL OAM LCD Information Register**

This register specifies the reassembly information for OAM cells. The format of this register is equivalent to word zero of the receive LCD. The following fields are valid: ppMode, size, storeCrc10, rxqNum, rxPoolId, rxOffset, and cutThruSel. Refer to the format of LCD word zero for the Raw receive LCDs.





### **12.36: RXALL - Scatter/Cut Through Info Registers**

These registers specify the scatter/cut through configurations. A configuration is selected in the LCD via the cut through selector field when doing cut through/scatter mode. A configuration consists of three registers. The first two registers, described here, define the four possible configurations for scatter/cut through. The third register, RXALL - Scatter/Cut Through Flag Registers, is described in ["](#page-357-0)RXALL - Scatter/Cut Through Flag Registers.

The first register, one of cti(0-3), is defined as follows:







### IBM3206K0424



## **IBM Processor for Network Resources Accessor For Accessor Preliminary**









#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

## <span id="page-357-0"></span>**12.37: RXALL - Scatter/Cut Through Flag Registers**

Use to specify the scatter/cut through flags. These registers specify the scatter/cut through flags. A flag register is selected in the LCD via the cut through selector field when doing cut through/scatter mode. The flags are used when building DMA descriptors for scatter pages and the first scatter buffer (packet header, DMA list, etc.).







# **RXLCD Block Diagram**



# **RXLCD Functional Description**

RXLCD provides an LCD cache for REASM sub-entities. This cache holds the last four receive LCDs. The sub-entities can request to load an LCD, read the LCD words, and update parts of the LCD.

# **12.38: RXLCD Cache Data Array Address Port**

This register provides the read/write address for accessing the LCD cache data array. When the RXLCD Cache Data Array Read/Write Port is read/written, this register is auto-incremented. When the last address is read/written, the address rolls over to zero.

The cache is organized as 64 32-bit words. Each 16 words comprises an LCD. The first four words and the last word of each LCD do not contain valid data and can not be written. These locations return zero on reads.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **12.39: RXLCD Cache Data Array Read/Write Port**

Provides read/write access to the LCD cache data array. When this register is read/written, the address provided by RXLCD Cache Data Array Address Port is used to select the array word to be accessed. RXLCD Cache Data Array Address Port is auto-incremented on each read/write so this port can be read/written multiple times to read/write the entire array.

The cache is organized as 64 32-bit words. Each 16 words comprises an LCD. The first four words and the last word of each LCD do not contain valid data and can not be written. These locations return zero on reads.



#### **12.40: RXLCD Cache Line Info Registers**

These registers provide the cache line tags, valid, and dirty bits. There is a register for each of the four cache lines.








**Preliminary IBM Processor for Network Resources**

# **12.41: RXLCD Mode Register**

This register provides a means to control cache operation.







## **RXRTO Block Diagram**



## **RXRTO Functional Description**

RXRTO performs periodic reassembly timeout processing and LCD update operations.

## <span id="page-361-0"></span>**Reassembly Timeout (RTO) Processing**

Reassembly timeout processing is generally an AAL5 operation. It is supported for other LCD types as well. It can be enabled on an LCD basis by turning on the RTO enable bit in the LCD. The following registers also need to be properly set up to run RTO processing:

- [R](#page-363-0)XRTO RTO LCD Table Bound Registers
- [R](#page-364-0)XRTO Reassembly Timeout Value Register
- [R](#page-364-1)XRTO Reassembly Timeout Pre-Scaler Register

See the register descriptions for more register details.

The LCD table registers define the LCD table that the RTO process examines. The value register is used as a compare value against a counter that counts based on a pre-scaler. Each time the registers compare, RTO processing is started for a single LCD and the time base is reset. RTO processing checks the RTO test and set bit. If it is reset, it sets it and continues. If it is set, then a timeout occurs and the LC is placed in error state and the current packet is surfaced to the user via an event. Any resource associated with the packet must be recovered by software. For example, if the LCD is setup to use scatter mode, then there may be scatter DMA pages in the DMA list that need to be returned to the proper receive queue. The RTO bit is reset with each inbound cell received. An LCD needs to be touched twice to cause a timeout (once to set it and once to detect that it is already set).

The time base starts running as soon as the RTO processing is complete. Thus, RTO processing is a low priority task.

#### **Shutting Down an LCD**

To shut down a receive LCD, the following steps should be followed:

- Clear the entry for this LCD in the LCD table (to stop receiving cells for this LCD)
- Do an LCD update operation that sets LCD state to down
- Read the LCD REASM ptr
- If REASM ptr is non-zero and LCD is set up to do cut through, be sure to free any DMA descriptor that was added with cut through operation
- If REASM ptr is non-zero and LCD is set up to do scatter, be sure to free any pages in the DMA list
- If REASM ptr is non-zero, free it to POOLS



#### **12.42: RXRTO LCD Update Data Registers**

These two registers are used to specify data to write into the receive LCD on the update LC operation. They contain the data used in the LC update operation. For more information on their use, see the [RXRTO LCD](#page-363-1)  [Update Op Registers](#page-363-1) on page 364.

The Update Date Register changes to contain the updated data written to the LC word while the operation is completing.

The second set of LCD update registers is meant for the core to use, but is available for general use.



#### **12.43: RXRTO LCD Update Mask Registers**

These two registers are used to specify which data to write into the LCD on the update LC operation. They contain the mask used in the LC update operation. For more information on their use, see the [RXRTO LCD](#page-363-1)  [Update Op Registers](#page-363-1) on page 364.

The second set of LCD update registers is meant for the core to use, but is available for general use.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### <span id="page-363-1"></span>**12.44: RXRTO LCD Update Op Registers**

Used to specify the LCD word to update. This operation is used to update a portion of the receive LCD. If this operation is not used, then software or the IBM3206K0424 updates of the LCD may be lost because the receive LCD is cached in the IBM3206K0424.

This register is written with the address of the LCD word to update. Once this register is written the update operation starts. All subsequent reads or writes to the data, mask, or update registers are held off until the operation completes. A read-modify-write will occur to update the portion specified by the mask with the masked value in the data register.

Normally this register would not be read. However, if it is read then the low order bit is read as '0' and the next lowest order bit (bit 1) is read as the busy bit. This signifies whether an operation is still going on. If an operation is still going on, then a new write to any of the data, mask, or update operation registers is held off until the original operation is complete.

The second set of LCD update registers is meant for the core to use, but is available for general use.



#### <span id="page-363-0"></span>**12.45: RXRTO RTO LCD Table Bound Registers**

Used to specify the lower/upper bounds of the LCD table. The lower bound should be initialized to the LCD index of the first LCD in the LCD table if reassembly timeout processing is to be done. The upper bound should be initialized to the LCD index of the last LC in the LC table if reassembly timeout processing is to be done.





#### <span id="page-364-0"></span>**12.46: RXRTO Reassembly Timeout Value Register**

Used to specify the time interval used for reassembly timeout processing. This register is the number of pre-scaler intervals between reassembly processing. The pre-scaler interval is determined by RXRTO Reassembly Timeout Pre-Scaler Register. A single LC is checked for reassembly timeout during each reassembly processing interval.

When this register is set to '0', reassembly timeout processing is disabled.

For more information on how reassembly timeout conditions are processed see Reassembly Timeout (RTO) Processing [on page 362.](#page-361-0)



## <span id="page-364-1"></span>**12.47: RXRTO Reassembly Timeout Pre-Scaler Register**

Used to specify the time interval of each RTO timer tick. This register determines the number of 15 ns intervals between RTO timer ticks. The value in the register plus 1 is the number of 15 ns intervals between RTO timer ticks. Thus, the default value of '0' means that the RTO timer ticks every 15 ns. If a value of nine is placed in this register, the RTO timer ticks every 150 ns (10 \* 15 ns).

[For more information on how reassembly timeout conditions are processed, s](#page-361-0)ee Reassembly Timeout (RTO) Processing [on page 362.](#page-361-0)





# **Entity 13: Receive Queues (RXQUE)**

### **Functional Description**

RXQUE has a single function: to manage the receive queues for software by providing an easy to use primitive interface. When talking about the receive queues, the term rxq is used to talk about a receive queue, and the term deq is used to refer to a dequeue operation, and the term enq is used to refer to an enqueue operation.

## **Receive Queue Interface**

A group of sixteen receive queues is available for software use. The receive queues hold events or user specified data.

Each queue entry (event) is either 32 or 64 bits and contains two fields: event-identifier and event-information. The seven least significant bits in the entry contain the event-identifier field. The most significant bits in the entry comprise the event-information field.

**Warning:** In order to maintain the atomicity of 64-bit atomic transfers, the user must ensure that 64-bit transfers are bus atomic within the particular bus system in which IBM3206K0424 is being used.

The event information typically contains a pointer (when low order bits are zeroed) to a packet buffer, a cell buffer, or an LCD. It can also contain a DMA descriptor address or user-specified data. Event Summary and [Routing Information](#page-366-0) on page 367 lists the different event types.



#### IBM3206K0424



# <span id="page-366-0"></span>**Event Summary and Routing Information** (Page 1 of 3)

## **IBM Processor for Network Resources Preliminary**



# **Event Summary and Routing Information** (Page 2 of 3)





#### **Preliminary IBM Processor for Network Resources**

# **Event Summary and Routing Information** (Page 3 of 3)



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **AAL5 Packet Events**

For AAL5 packet events, the event specifies the packet buffer address, and the event type field specifies the type of packet event. The following event types are defined:





## **Cell Events**

For AAL0 events, the event specifies a cell buffer address, and the event type field specifies type of AAL0 event. The following event types are defined:



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



#### **LC Events**

For LC events, the event specifies a LC, and the event type field specifies what happened on the LC. The following event types are defined:



## **ABR Events**

The following events are used for ABR processing and are routed to the receive queue specified in the ABR event routing register.





## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary P**

# **Miscellaneous Events**



## **IBM Processor for Network Resources Preliminary**



## **Miscellaneous Events** (Continued)





**Preliminary IBM Processor for Network Resources**

## **Frame-Based Events**



## **PCORE Events**



## **System-Receive-Queue Events**



The receive queues are maintained by RXQUE with the following operations being available to software:

- **dequeue** Remove the entry at the head of the queue
- **enqueue**  Add arbitrate entry at the tail of the queue

The following figure shows how events in a receive queue link to other data structures including LC control blocks, packet buffers, and cell buffers.

#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**









**Preliminary IBM Processor for Network Resources**

## **RXQUE Structure**

Each queue has a number of registers that define the queue and its behavior:



## **RXQUE Initialization**

To set up a receive queue, at least two pieces of information are needed. The first is the receive queue's set of properties, and the second is its base address.

The following restrictions should be taken into account when setting up a queue:

- The properties register must be set up before the lower-bound and next-lower-bound registers can be set up.
- The lower bound and next lower bound must be at least 1K aligned. The low order 10 bits of these registers are not writable, so the minimum physical size of a receive queue is 1024 bytes (256 32-bit entries). The alignment should correspond to the size specified in the properties register.
- The head and tail pointers are initialized when the lower bound register is written. These registers are only writable for diagnostic purposes.
- The threshold is level sensitive, so as long as the queue length is greater than or equal to the threshold, the appropriate status bit is driven.
- All registers, except the threshold and next-lower-bound registers, can only be written in diagnostic mode and are intended to only be written once when they are set up.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



### **RXQUE Event Routing**

Events are routed to a receive queue based on the current event type and mode of the chip. Events fall into these categories:

- Normal Events
- Error Events
- Counter Events
- Transmit Complete Events
- DMA Events (tx/rx)
- ABR Events
- POOLS Status Events

See [Event Summary and Routing Information](#page-366-0) on page 367 to see how the different events are categorized.

All events other than normal, DMA, and error events are routed using the corresponding RXQUE Event Routing Registers.

Normal events are always routed to the receive queue specified in the receive portion of the LCD.

DMAQS specifies the route for all DMA events.

Error events are special, in that they are routed based on the values of the "receive bad frame mode" bit and the "always route error events" mode bit in RXQUE Control Register. If the "always route error events" bit is on, then the error events are always routed to the error queue. Otherwise, if the receive bad frame mode is on, then the error events are routed to the receive queue specified in the receive portion of the LCD just like a normal event would be. When receive bad frames is off, then the error events are routed to the error queue.



#### **RXQUE Normal Operation**

This section describes how to use the receive queues (rxq) and the rxq operations.

The receive queue contains events for the end user to process. These events are obtained by the user by executing the rxq deq operation. The user can be notified of new events by setting up the threshold and interrupt enable registers appropriately. Otherwise, the rxq length register can be polled to check for events.

The deq operation is executed by reading the deq register address for the appropriate rxq. The event at the head of the queue is returned and the event is removed from the queue. Some events have a packet/cell buffer associated with it. This buffer is owned by the user, and it is the users responsibility to free this buffer.

The following pseudo code illustrates how an rxq could be processed:

## **RXQUE Dequeue Event Loop**

// rxq was polled or int ocurred to get here

Event =  $RXQUE$ - $Deq$ ; // read an event from rxq if (Event neq  $0$ ) {  $\frac{1}{2}$  // need to check for null event EventType = Event &  $0x7f$ ; // calc event type Event = Event  $\& \sim (0x7f)$ ; // calc lc or buffer ptr switch (EventType) { case(Event1): ProcessSimpleEvent1(Event); break; case(Event2): ProcessSimpleEvent2(Event); break; case(EventX): ProcessSimpleEventX(Event); break; } }

#### **RXQUE Queue Full Operation**

When a receive queue is full (length is equal to maximum length), the appropriate status bit is set. When a queue is full, all subsequent events are flushed until room is available in the receive queue. If a buffer was associated with the event and the RXQUE-Properties-Register bit Disable Auto-Free is not set, then that buffer is freed back to POOLS.

When an event is dropped, the event dropped status bit is set and the event data that was dropped can be found in RXQUE Last Event Dropped Register. The RXQUE Last Event Dropped Register will not be changed until the event dropped status bit is cleared.

It is not good to let a receive queue become full.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **RXQUE Event Timestamping**

When timestamp mode is set in the RXQUE Control Register, events are timestamped. When timestamping is enabled, a timestamp event is placed in the corresponding rxq followed by the actual event. The event information of the timestamp event carries the timestamp. The timestamp is determined from the RXQUE Timestamp Register, RXQUE Timestamp Pre-Scaler Register, and the [R](#page-393-0)XQUE Timestamp Shift Register.

If the corresponding rxq is full, both events are dropped. It is possible to lose only the timestamp event or lose the actual event depending on the length of the queue and the timing of the dequeue operations.

#### <span id="page-379-0"></span>**RXQUE System Receive Queues**

To set up a system receive queue, set the "Diagnostic-Mode" bit in the RXQUE Control Register. Next, set the system receive queue bit in the RXQUE Properties Register. Load the upper bound and size of event fields, as well. After this, allocate two identical buffers in system memory. Let each buffer be large enough to contain N events (the upper bound field prescribes the value N) and initialize both to all zeros. Write one buffer's starting address into the RXQUE Lower Bound Register (LOBR), and write the other's starting address into the RXQUE Next Lower Bound Register (NLBR). Finally, reset the diagnostic mode bit in the RXQUE Control Register. System-receive-queue setup is complete.

The first event enqueued to a system receive queue causes RXQUE to begin filling the buffer which LOBR references. RXQUE fills the buffer's first entry with a "System-Receive-Queue Start-of-Buffer" event whose information is the value '0'. After this, RXQUE fills the buffer's second entry with the enqueued event and writes the value '2' into the RXQUE Length Register (LENR).

The second event enqueued to the system receive queue causes RXQUE to fill the buffer's third entry and write the value '3' into LENR

The third event enqueued to the system receive queue causes RXQUE to fill the buffer's fourth entry and write the value '4' into LENR.

This continues until LENR contains the value 'N-1'. At that time, RXQUE fills the buffer's Nth entry with a System-receive-queue end-of-buffer event whose information is the value in NLBR. After this, RXQUE begins filling the buffer which NLBR references. RXQUE fills its first entry with a "System-Receive-Queue Start-of-Buffer" event whose information is the value in LOBR. After this, RXQUE copies the contents of NLBR into LOBR and writes the value '1' into LENR. Finally, RXQUE writes the value '0' into NLBR. Subsequent events enqueued to the system receive queue fill the buffer which LOBR references. To prevent the system receive queue from becoming full, write a non-zero value into NLBR.

The system receive queue becomes full when the value in LENR becomes 'N-1' while NLBR contains the value '0'. At that time, RXQUE fills the buffer's Nth entry with a system-receive-queue end-of-buffer event whose information is the value '0'. RXQUE writes the value N into LENR and preserves the contents of LOBR while dropping subsequent enqueued events. To restart the system receive queue from the "full" state, write a non-zero value into NLBR. After restarting, the next event enqueued to the system receive queue causes RXQUE to begin filling the buffer which NLBR references. RXQUE fills the buffer's first entry with a system-receive-queue start-of-buffer event whose information is the value preserved in LOBR. RXQUE fills the buffer's second entry with the enqueued event. After this, RXQUE copies the contents of NLBR into LOBR and writes the value '2' into LENR. Finally, RXQUE writes the value '0' into NLBR. Subsequent events enqueued to the system receive queue fill the buffer which LOBR references. To prevent the system receive queue from becoming full, write a non-zero value into NLBR.

The RXQUE Queues' Status Register bit "Threshold Exceeded" indicates that the value in LENR is greater than or equal to the value in the RXQUE Threshold Register while NLBR contains the value '0'. To reset this



status bit, write a non-zero value into NLBR.

Events enqueued to a system receive queue may not be dequeued via the RXQUE Dequeue Register. To dequeue from a system receive queue, poll system memory directly. A buffer's entry is filled if its value is non-zero.

RXQUE synchronizes its internal-register operations with the initiation, rather than completion, of its system-memory operations. Therefore, the state of system memory lags the state of RXQUE.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### <span id="page-381-0"></span>**13.1: RXQUE Lower Bound Registers**

These registers specify the lower bound of the corresponding receive queue data structure. The head and tail of the receive queue are initialized when this register is written. When the receive queue wraps past the upper bound, it wraps back to the value in the lower bound register, thus implementing the receive queue as a circular buffer.

When this register is written, the corresponding receive queue is essentially reset. This is because the head, tail, and length of the queue are all reset.

The length of the RXQUE Lower Bound Register is 64 bits if all three conditions, below, are met; otherwise, the length is 32 bits.

- **System Receive-Queue** in the RXQUE Properties Register is set.
- **System-Memory Select** in the RXQUE Properties Register indicates "PCI Memory."
- **Enable Master 64-bit Addressing** in the PCINT 64bit Control Register is set.



The lower bound registers must be at least 1K aligned (low order 10 bits not writable). The alignment should also correspond to the size specified in the upper bound register. For example, it should be 4K aligned if the upper bound specifies 4K size.



# **13.2: RXQUE Properties Registers**

These registers specify the properties of the corresponding receive queue.



pnr25.chapt05.01 August 14, 2000









## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **13.3: RXQUE Head Pointer Registers**

These registers point to the head element of the corresponding receive queue. During normal operations, these registers do not need to be read or written, as they are used by the IBM3206K0424 to implement the receive queues. These registers are initialized when the lower bound register for the corresponding receive queue is written.





## **13.4: RXQUE Tail Pointer Registers**

These registers point to the next free element of the corresponding receive queue. During normal operations, these registers do not need to be read or written, as they are used by the IBM3206K0424 to implement the receive queues. These registers are initialized when the lower bound register for the corresponding receive queue is written.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **13.5: RXQUE Length Registers**

These registers specify the length (number of valid entries) of the corresponding receive queue. They can be used to query the status of a receive queue.

This register is cleared when the corresponding lower bound is written.





## **13.6: RXQUE Threshold Registers**

These registers specify a queue length threshold at which the corresponding status bit is generated. These registers should be set equal to the number of queue entries that should cause status to be generated. For example, if the value was set to five, then no interrupt would be generated until five or more events were queued on the corresponding receive queue. The threshold is level sensitive, so as long as the length is greater than or equal to the threshold, the corresponding status bit is set. When this register is set to '0', no thresholding is done.

When the direction bit is set for a receive queue, the threshold has the opposite polarity. For example, as long as there are more events in the queue than specified in the threshold register, no status would be raised.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **13.7: RXQUE Dequeue Registers**

These registers are used to retrieve the event at the head of the corresponding receive queue. These registers are used to retrieve the event at the head of the corresponding receive queue.

The length of an RXQUE Dequeue Register is 64 bits if Size of Event is set in its corresponding RXQUE Properties Register; otherwise, the length is 32 bits.



**Restrictions** This is a read only register, and all writes will be ignored. Events are only returned when the diagnostic bit is reset in the control register, otherwise zero will be returned.



## **13.8: RXQUE Enqueue Registers**

These registers are used to enqueue user events at the tail of the corresponding receive queue.

The length of a RXQUE Enqueue Register is 64 bits if Size of Event is set in its corresponding RXQUE Properties Register; otherwise, the length is 32 bits.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **13.9: RXQUE Next Lower Bound Registers**

These registers specify the next lower bound of the corresponding system receive queue data structure. See [RXQUE System Receive Queues](#page-379-0) on page 380 for instruction about managing system receive queues.

[The length of the RXQUE Next Lower Bound Register is the same as the length of the RXQUE Lower Bound](#page-381-0)  [Register. See](#page-381-0) [RXQUE Lower Bound Registers](#page-381-0) on page 382 for conditions which determine the length.





## **13.10: RXQUE Last Event Dropped Register**

This register contains the last event that was dropped. It holds its value until the event dropped status bit is cleared.

The length of the RXQUE Last Event Dropped Register is 64 bits if Size of Dropped Event is set in the RXQUE Status Register; otherwise, the length is 32 bits.



#### **13.11: RXQUE Timestamp Register**

Used to specify the current timestamp measured using the timestamp pre-scaler ticks. It counts based on the value in the RXQUE Timestamp Pre-Scaler Register. It can be read or written at any time. It is cleared when the pre-scaler register is written.



#### **13.12: RXQUE Timestamp Pre-Scaler Register**

Used to specify the time interval of each timestamp timer tick. This register determines the number of 15 ns intervals between timestamp timer ticks. The value in the register plus one is the number of 15 ns intervals between timestamp timer ticks. So, the default value of '0' means that the timestamp timer ticks every 15 ns. If a value of four is placed in this register, the timestamp timer ticks every 75 ns (5 x 15 ns).



#### **IBM Processor for Network Resources All and Service Control of All and Service Preliminary**



#### <span id="page-393-0"></span>**13.13: RXQUE Timestamp Shift Register**

This register determines the number of bits that the timestamps are shifted. For example, if a value of '0' is placed in this register, then the timestamp is not shifted, and the low order seven bits are lost. If a value of '2' is placed in this register, then the timestamps are shifted two places and only the low order five bits of the timestamp are lost. This allows the user to control what portion of the timestamp is lost due to the low order event bits.



#### **13.14: RXQUE Event Routing Registers**

Used to specify which receive queue different types of events should be routed to. These registers contain the receive queue that different types of events should be routed to. See [Event Summary and Routing Infor](#page-366-0)mation [on page 367](#page-366-0) for event type mappings.





## **13.15: RXQUE Event Latency Timer Register**

Used to specify the event latency time interval. This register is specified in 15 ns intervals. When a new event is placed on an rxq, the event latency timer is started (if not already started). When this timer expires, the event latency timer expired status bit is set, and the timer is stopped. The status bit must be reset before the timer is started again. Every time the status register (or prioritized status) is accessed, the timer is stopped. If this register is written while the timer is running, the new value takes effect immediately. If this register is set to '0', the latency timer does not run.





## **13.16: RXQUE Queues Status Register**

Indicates the status for all receive queues.








### **13.17: RXQUE Interrupt Enable Registers**

Used to specify which status register bits should be used to generate interrupts. Each mask register is used to drive a different RXQUE status bit in intst. The different masks and status bits allow two RXQUE interrupts on both the interrupt A and B pins. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing. See [RXQUE Status and Enabled Status Registers](#page-397-0) on page 398 for the bit descriptions.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

### <span id="page-397-0"></span>**13.18: RXQUE Status and Enabled Status Registers**

This register contains the status bits used to relay RXQUE status information. The enabled version of these registers provides a version of the status register that is masked with the corresponding interrupt enable register. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







# IBM3206K0424





#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

# **13.19: RXQUE Control Register**

Used to set RXQUE modes. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing. This register contains the mode bits that specify how RXQUE is to operate.









### **Preliminary IBM Processor for Network Resources**



# **Debugging Register Access**

This section is a very brief documentation of access that has been put in for the internal registers of RXQUE. These addresses need not be written or read during normal operations.

### **13.20: RXQUE RXQ State Machine Variable Register**

Main state variable for RXQUE processing state machine.



# **13.21: RXQUE RXQ ENQ State Machine Variable Register**

Main state variable for RXQUE processing state machine.





#### **13.22: RXQUE Enq FIFO Head Ptr Register**

Used to maintain the enqueue FIFO. Points to the head FIFO entry in the FIFO array. The MSB bit is used to determine if the head is chasing the tail, and is inverted each time the head pointer wraps.



### **13.23: RXQUE Enq FIFO Tail Ptr Register**

Used to maintain the enqueue FIFO. Points to the next free FIFO entry in the FIFO array. The MSB bit is used to determine if the head is chasing the tail, and is inverted each time the tail pointer wraps.



#### **13.24: RXQUE Enq FIFO Array**

Holds events waiting to be placed on an rxq. Array is organized as a 16x36 array. To access the upper four bits of each word (holds the receive queue number for event), the array word number should be used as the address. To access the low order 32 bits (the event portion), the array word number times two plus four should be used. For example, address zero accesses the receive queue portion of array word zero, and address four accesses the event portion of the array word zero.

**Note:** The most significant bit is not used. Only the three bits are needed for the receive queue number.





# **PHY Level Interfaces**

# **Entity 14: The PHY Interface (LINKC)**

# **Functional Description**

LINKC provides the interface between the IBM3206K0424 and either an ATM PHY device or, when the internal framer is selected, a serializer/deserializer device. LINKC is composed of three pieces. LINKX, which contains all the registers described below, is clocked with the same clock as other parts of the chip. LINKT, the transmit logic, is clocked on the transmit clock, which is selected via the Clock Control Register (described in [Clock Control Register \(Nibble Aligned\)](#page-515-0) on page 516). LINKR, the receive logic, is clocked on the receive clock, which is also selectable via the Clock Control Register. Transmit and receive transfers are synchronized via their respective interface transfer clock. The data path size is 8- or 16-bits wide and is selectable through bit 3 of the control register. The PHY devices that the IBM3206K0424 interfaces to are:

- PMC SIERRA PM5346 SUNI LITE FOR SONET STS-3c 155.52 MB/s
- UTOPIA 8 or 16 bit interface
- PMC SIERRA POS-PHY

New features added to LINKC are:

- Multi-drop Utopia support
- PMC SIERRA POS-PHY interface support

# **Multi-Drop**

When the IBM3206K0424 is in multi-drop Utopia mode, it supports four external PHY devices. Each port is associated with a configuration. Four configurations are provided so up to four different types of PHYs can be connected to the IBM3206K0424. This allows the user to mix cell and POS-PHY devices on the transmit and/or receive interface.

The multi drop PHY devices supported are Utopia Level 2 (cell based) and PMC Sierra POS-PHY (packet/frame based). The IBM3206K0424 will select which PHY device will transfer data next by polling each of the devices to determine which PHYs can transfer data. A round-robin switching scheme is used to determine which PHY has the priority if more then one wants to transmit/receive data. The IBM3206K0424 will switch to a new drop when a cell has been received/transmitted (for a cell-based PHY) or when 64 bytes or EOP has been received/transmitted (for POS-PHY PHYs). The transmit and receive sides of LINK are separately configurable for multi-drop mode (bits 1 and 0 of the global control register).

# **POS-PHY**

The POS-PHY interface complies with the PMC Sierra POS-PHY Level 2 Specification. The IBM3206K0424 polls each POS-PHY device to determine its status on both the receive and transmit side. It looks to switch to a different port when 64 bytes or EOP (End of Packet) have been transferred between the POS-PHY and itself. The IBM3206K0424 does not support direct status indication or byte-level transfers. Therefore, the PHY must be programmed to always be able to send/receive at least 64 bytes of information. The RMOD signal will only be looked at when REOP is b'1'; at all other times it will be ignored. POS-PHY devices should be configured so that they will only signal they are ready for a transfer if they have 64 bytes free in their receive buffer and 64 bytes or EOP in their transmit FIFO.

#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

#### **Moving Cells To and From the IBM3206K0424**



# **14.1: LINKC Global Control Register**

This register contains the information which controls the operation of LINKC. These controls affect all config-urations. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.











# **IBM Processor for Network Resources Preliminary**



# <span id="page-405-0"></span>**Legal Loopback Configurations**





# **14.2: LINKC Configuration 0 Transmit & Receive Control Register**

This register contains the information which controls the operation of Configuration 0 on the transmit and receive. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









## **IBM Processor for Network Resources Accessor For Accessor Preliminary**









#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **14.3: LINKC Configuration 1 Transmit & Receive Control Register**

This register contains the information which controls the operation of Configuration 1 on the transmit and receive. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.













## **IBM Processor for Network Resources Accessor For Accessor Preliminary**





## **14.4: LINKC Configuration 2 Transmit & Receive Control Register**

This register contains the information which controls the operation of configuration 2 on the transmit and receive. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.





### IBM3206K0424



## **IBM Processor for Network Resources Accessor For Accessor Preliminary**







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **14.5: LINKC Configuration 3 Transmit & Receive Control Register**

This register contains the information which controls the operation of configuration 3 on the transmit and receive. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







# IBM3206K0424



## **IBM Processor for Network Resources Accessor For Accessor Preliminary**







# **14.6: LINKC Map Transmit Configurations to Port Addresses**

This register contains the port address for each of the transmit configurations. If the port address is B"11111" then the configuration is unused. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **14.7: LINKC Map Receive Configurations to Port Addresses**

This register contains the port address for each of the receive configurations. If the port address is "11111" then the configuration is unused. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









### **14.8: LINKC Transmitted HEC Control Byte**

When the IBM3206K0424 is transmitting to a 53-byte Utopia PHY the HEC byte (byte five of the ATM header) will be sent as x'00', if bit 17 of the transmitting configuration is a '1'. Otherwise the value of the LINKC Transmitted HEC Control Byte Register will be sent.



HEC Byte Value





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# <span id="page-421-0"></span>**14.9: LINKC Interrupt/Status Register**

This register reports the status of LINKC. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.











### **14.10: LINKC Interrupt Enable Register**

This register enables the LINKC interrupt to INTST. When a bit is set in this register and the corresponding bit is set in the Interrupt Status Register, the LINKC interrupt will be driven. See [Note on Set/Clear Type Regis](#page-92-0)[ters on page 93](#page-92-0) for more details on addressing. See the [LINKC Interrupt/Status Register](#page-421-0) on page 422 for the bitwise description of this register.



#### **14.11: LINKC Prioritized Interrupts**

Used to access the prioritized encoding of LINKC interrupts. Reading this location will give a decimal number that is the prioritized encoding of bits 7-0 in COMET/PAKIT Status Register (seven being the most significant bit) assuming the corresponding enable bit is on.





**Preliminary IBM Processor for Network Resources**

### **14.12: LINKC Transmit State Machine Register**

This register indicates the state of the transmit sequencer.



Unassigned Transmit





## **14.13: LINKC Receive State Machine Register**

This register indicates the state of the receiver sequencer.







## **14.14: LINKC LAN Address Register**

If using an IBM built adapter that utilizes CRISCO, this register contains the ROM level in bits 63-48 and the LAN address of the adapter in bits 47-0. The lower address selects bits 63-32 and the higher address selects bits 31-0.



# **14.15: LINKC Canonical LAN Address Register**

This register contains the same data as the LINKC LAN Address Register except each byte is bit reversed. This allows the user to obtain the LAN address in canonical format.





# **14.16: LINKC Passed TX Data Register**

The bits in this register are passed over PHY transmit data I/O 15-8 when using an eight bit wide PHY data bus.







# <span id="page-427-0"></span>**Entity 15: Nodal Processor Bus Interface (NPBUS)/CRISCO Processor for Register Initialization from EPROM Data**

This entity controls the signals of the NPBUS. The PHY registers are accessible to the processor by way of the address space of the IBM3206K0424. In addition, the operation of the front end is affected by the NPBUS Status Register.

This entity also contains the CRISCO processor which can initialize chip registers at boot time by reading a data stream from EPROM which specifies the address of registers and data values to which the registers are to be initialized. In general, the data stream consists of a series of single-byte instruction operation codes, followed by an address and data values.

# **15.1: NPBUS Control Register**

This register is used to report PHY level hardware errors and interrupts. See [Note on Set/Clear Type Regis](#page-92-0)[ters on page 93](#page-92-0) for more details on addressing.







#### IBM3206K0424



#### **Preliminary IBM Processor for Network Resources**









# <span id="page-430-0"></span>**15.2: NPBUS Status Register**

This register is used to report PHY level hardware errors and interrupts. See [Note on Set/Clear Type Regis](#page-92-0)[ters on page 93](#page-92-0) for more details on addressing.







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **15.3: NPBUS Interrupt Enable Register**

This register is used to mask bits from the NPBUS Status Register and potentially generate interrupts to the control processor. When both a bit in this register and the corresponding bit in the NPBUS Status Register are set, the NPBUS status bit will be set in INTST Interrupt Source[.](#page-427-0) See Note on Set/Clear Type Registers on [page 93](#page-92-0) for more details on addressing. See [NPBUS Status Register](#page-430-0) on page 431 for the bit descriptions.




## **15.4: NPBUS EPROM Address/Command Register**

Used to accessed a maximum of 2K external serial EPROM or 16 Meg of parallel EPROM or the Sonet Framer Core. This register is used access bytes from the external EPROM or the Sonet Framer Core.









# **15.5: NPBUS EPROM Data Register**

Used to accessed a maximum of 2K external serial EPROM or 16 Meg of parallel EPROM.







## **15.6: PHY 1 Registers**

This address range provides access to the PHY 1 hardware. The details of the registers can be found in the specific publication for the PHY hardware.



# **15.7: PHY 2 Registers**

This address range provides access to the PHY 2 hardware. It should be noted that not all applications of IBM3206K0424 will use this access port. The details of the registers can be found in the specific publication for the PHY hardware.





# **Hardware Protocol Assist Entities**

# **Entity 16: On-chip Checksum and DRAM Test Support (CHKSM)**

# **Functional Description**

The CHKSM entity has two functions: First, it is capable of initializing and/or testing packet and Control Memory; second, it can perform TCP checksums (Two's complement, 16-bit sum with "end-around-carry").

# **16.1: CHKSM Base Address Register**

The CHKSM Base Address Register indicates the starting address of a test operation or checksum calculation. The base address can be set up with any alignment and valid address.

This register increments with each read or write to memory. On a test mode error, the register holds the address of the 64-bit word which was read in error.



#### **IBM Processor for Network Resources All and Service Construction Construction Preliminary**



## **16.2: CHKSM Read/Write Count Register**

The CHKSM Read/Write Count Register indicates the count of remaining bytes of a checksum operation. This register keeps track of how many bytes remain in the current checksum operation and is decremented with each read or write operation.

Any length can be set in the 30 lower significant bits.

The upper two bits of this register can be written when starting a checksum operation instead of writing the control register.









#### **16.3: CHKSM TCP/IP Checksum Data Register**

The CHKSM TCP/IP Checksum Data Register collects the 16-bit, two's complement, end-around-carry sum of the bytes. The source data is zero padded if it starts/ends on an odd byte boundary. The CHKSM TCP/IP Checksum Data Register collects the 16-bit, two's complement, end-around-carry sum of the bytes. It can be seeded with an initial value. If it is not cleared before running a checksum, the previous value will act as a seed. This register can be cleared when starting a checksum operation by writing the CLIP bit in the CHKSM Control Register or by writing upper bits of CHKSM Read/Write Count Register.

See [CHKSM Control Register](#page-439-0) on page 440) for description of how to get/set current checksum alignment.



## **16.4: CHKSM Ripple Base Register**

This register is used as base of a ripple pattern when in test mode. This register forms the base for a ripple pattern. Each consecutive byte will be incremented by one or eight in the pattern. The ripple mode must be set in the Control Register to use this feature. The value of this register will change during the test operation. If a write and compare operation are being performed, this register needs to be setup again for the second operation.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **16.5: CHKSM Ripple Limit Register**

This register is used to determine when the ripple base register overflows to zero. This register forms the compare value for the ripple base register. When the value of the ripple base register is greater than or equal to this register, the base register will overflow to zero. For example, when this register is set to four, the ripple base register would count from zero through four if counting by one.

When set to 0x00, no overflows to zero occur. For example, when the ripple value is 0xff, and you are counting by eight, the next value would still be seven. If counting by one, then the next value would be zero.

This register should be written before the ripple base.



## **16.6: CHKSM Interrupt Enable Register**

Used to specify which status register bits should be used to generate interrupts. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing. See[CHKSM Control Register](#page-439-0) on page 440) for the bit descriptions.





# <span id="page-438-0"></span>**16.7: CHKSM Status Register**

Used to relay CHKSM status information. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# <span id="page-439-0"></span>**16.8: CHKSM Control Register**

The various bits in this register control the mode in which the checksum entity operates. See Note on [Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







#### IBM3206K0424



## **Preliminary IBM Processor for Network Resources**



# **Debugging Register Access**

## **16.9: CHKSM Internal State**

Internal state of checksum.

**Note:** This register should not be written unless specifically directed to do so by IBM technical support.



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **Software Use of CHKSM**

This section outlines some ways CHKSM can be set up and used.

## **Test Mode Possible Patterns**

In test mode, a 64-bit pattern is written/compared to/with memory. There are several different patterns that can be used:







# **Initializing Packet/Control Memory**

The following list shows the steps to use CHKSM to initialize Packet or Control Memory:

- Make sure CHKSM is in diagnostic mode, and other mode bits are reset
- Set the start address by writing the base address
- Set up the read/write count with number of bytes to initialize
- Set up the test pattern register (ripple pattern register) with pattern to use
- Set up the Control Register to enable test mode, enable checksum entity, and set the memory select bit correctly based which memory is to be initialized
- Now busy wait until operation is done (or set up Interrupt Enable Register and wait for interrupt)

# **Testing Packet/Control Memory**

The following list shows the steps to use CHKSM to test packet or Control Memory:

- First initialize memory with a pattern using above sequence
- Make sure CHKSM is in diagnostic mode, and other mode bits are reset
- Set the start address by writing the base address
- Set up the read/write count with number of bytes to test (same as initialization value)
- The test pattern register (ripple pattern register) already contains the pattern
- Set up the Control Register to enable test mode, turn on RW bit, enable checksum entity, and set the memory select bit correctly based which memory is to be initialized
- Now busy wait until operation is done (or set up interrupt Enable register and wait for interrupt)
- When done, check the status register for any errors

# **Using Ripple Pattern Generation/Checking in Packet/Control Memory**

The procedures to use the ripple pattern generation and checking are the same as using test write/read modes. The only difference is that the use ripple pattern mode bit must be set and the ripple pattern base register must be set up.

# **Running a TCP/IP Checksum in Packet/Control Memory**

The following list shows the steps to use CHKSM to generate/verify a TCP/IP checksum:

- Make sure CHKSM is in diagnostic mode (not enabled)
- Set the start address by writing the base address
- Set up the read/write count with number of bytes to run checksum over, and set the upper two bits of the read/write count register. Writing these upper two bits assumes other mode bits are set correctly (that is, memory bank select).
- Now busy wait until operation is done (or set up interrupt Enable register and wait for interrupt)

#### **IBM Processor for Network Resources All and Security Construction Construction Preliminary**



# **Entity 17: Processor Core (PCORE)**

PCORE contains the on board processor and its local subsystems. The primary intent is to run Available Bit Rate(ABR) control software and user application code such as protocol termination/assist code.

# **DCR Interface**

The Device Control Register (DCR) interface is a special processor bus to access local registers. These include serial port registers and various other registers.

## **Interrupt Controller**

This logic manages the interrupts that are passed on to the Cobra Core. There are two levels of interrupt for the core: Critical Interrupts and Normal Interrupts. Interrupts can be taken from both on chip and off chip sources. PCORE has a variety of interrupt source and enable registers.

## **Bridge-Address Translation**

Cobra Core can access a variety of memory subsystems. Facilities are provided that allow multiple subsystem accesses. Processor address space is translated into target memory system addresses. For the most part, this allows the processor to be unaware of target memory address space considerations while running mainline code.

## **OCM SRAM**

OCM is provided for the use of the processor. Typical access time to the OCM is a single cycle, the same as for cache. Address translation facilities have been added to make more efficient use of this memory via additional and extended BAT registers.

## **Control Memory**

Control Memory can be accessed by the processor. This memory may be mapped into the processor space in a number of different ways.

## **Packet Memory**

Packet memory can be accessed by the processor. This memory may be mapped into the processor space in a number of different ways. Packet memory space also includes the virtual memory space of the IBM3206K0424.

## **PCI Master Interface-External**

The processor can access the PCI bus through this interface. Parts of PCI space are mapped into processor space. There are a number of different ways that this can be mapped into processor space.

## **Processor Register Space**

This access mode of the PCI master interface allows access to the internal IBM3206K0424 registers. This access is handled internally and does not affect the external PCI bus.



**Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Community Preliminary** 

# **Address Translation Examples**

TBD

# **Cobra Structure**

# **Cobra Core Data Side**



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **Cobra Core Instruction Side**



## **Cobra Core "Glossy" Description**

The Cobra Core is a 32-bit PowerPC RISC embedded controller. It is fully compatible with the PowerPC User Instruction Set Architecture. Details about the exact instruction set are described below. The Cobra Core has a PowerPC instruction execution complex, separate 32k instruction and data caches, separate instruction and data 604-style MMUs (not supported this pass), and 401-style interrupts, timers and debug facilities. The Cobra Core has a direct connection to 96k of on-chip memory which can be used for both instruction and data storage, as well as interfaces to the IBM3206K0424's PCI and register buses and both of the IBM3206K0424's memory controllers. The DCR bus provides fastand private access to specific performance-sensitive registers.

## **Features**

## **Instruction Execution**

- Compatible with PowerPC User Instruction Set Architecture (UISA)
- Separate Branch, Condition Register, Integer, and Load/Store units for super-scalar execution
- Support for limited out-of-order execution
- Dispatches/Executes up to 2/4 instructions per clock cycle
- Four stage pipeline allowing single cycle execution for most instructions
- 32x32-bit general purpose registers (GPRs)
- Single cycle loads and stores
- Byte, halfword, word, and string accesses to any byte alignment supported in hardware
- Hardware multiply and divide (multiply up to 10 cycles, divide up to 32 cycles)
- No FPU hardware FPU instructions result in interrupts



# **Preliminary IBM Processor for Network Resources**

# **Timers**

- 32-bit decrementer
- 64-bit time base
- Programmable and fixed interval timers
- Watchdog timer

# <span id="page-446-0"></span>**Cobra Core Exceptions**

- Two priority levels
	- Normal uses SRR0/1 (603)
	- Critical uses SRR2/3 (401 extension)
- Exception Types
	- System Reset (boot vector)
	- Machine Check
	- Data Storage Interrupt
	- Instruction Storage Interrupt
	- External
	- Alignment
	- Trap
	- Invalid Opcode
	- Privilege Violation
	- Floating Point Unavailable
	- Decrementer
	- System Call
	- Trace
	- System Management (603)
	- Programmable Interval Timer (401)
	- Fixed Interval Timer (401)
	- Watch Dog Timer (401)
	- Critical Interrupt (401)
	- Debug Interrupt (401)

#### **IBM Processor for Network Resources All and Security Construction Construction Preliminary**

- **Optional Architecture Extensions**
	- Programmable boot address (system interrupt vector)
	- Interrupt enhancements
		- Individually re-locatable interrupts
		- Individually programmable interrupt level (normal/critical)

#### **Caches**

- 32-byte cache lines (eight words/line)
- Four-way set associative write back 32k instruction and data caches

## **Memory Management**

• Real Flat Address Mode supported

# **Interrupt Controller**

- Two interrupt levels external to COBRA: normal and critical
- Three-way interrupts
	- From IBM3206K0424 to COBRA
	- From COBRA to PCI
	- From PCI to COBRA

# **Debug Support**

- PowerPC JTAG debugger support (401 RiscWatch)
- 401 debug instructions
- Serial Port Debugger support
- PCI Debug access to JTAG debug facilities

# **Interfaces**

## **On-Chip Memory**

- 96k of on-chip memory
- Can be used simultaneously by instruction and data accesses
- OCM Basic DMA controller provides bulk data moves to/from OCM

## **IBM3206K0424 Registers**

- Read/Write access to the IBM3206K0424 register bus
- Some critical registers are mapped to COBRA Core DCR register space

## **PCI Bus**

- Read/Write master access to PCI Bus
- Currently no actual streaming/bursting supported
- Pseudo bursting supported (multiple back to back single transfers)
- Interrupt sink
- No arbitration supported (we are not a complete bridge)

## **Comet/Pakit Memory**

• Able to use both memory controllers for both instruction and data accesses





# **Preliminary IBM Processor for Network Resources**

# **Performance**

- 133 MHz operating frequency
- Performance estimates unavailable

# **Instruction Set**

# **Instruction Set (with 401 as a base)**

- Supports ALL 401 instructions with the following additions/subtractions:
- Added from 603
	- mfsr, mfsrin, mtsr, mtsrin, tlbie, tlbld, tlbli, tlbsync (**Note**: virtual memory not supported this pass)
	- mftb (for 603 style time base support)
- Removed from 401
	- dcread, icread (replaced with SPR access to the caches)
	- Changed mfspr, mtspr (to accommodate new SPRs)

# **Instruction Set (with 603 as a base)**

- Supports ALL 603 instructions with the following additions/subtractions
- Added from 401
	- dccci, icbt, iccci, mfdcr, mtdcr, rfci, wrtee, wrteei
- Removed from 603
	- fXXXXX, lfXXX, stfXXX, mffXXX (floating point instructions),
	- eciwx, ecowx (PowerPC I/O addressing not supported only memory space addressing)
	- Changed mfspr, mtspr (to accommodate new SPRs)

# **Cobra Instruction Overview**

Cobra Core supports all of the instructions in either the PowerPC 603 User's Manual (MPR603UMU-01,MPC603UM/AD) or the Power PC 401 Core User's Manual (v0.07 1/28/978) with the following changes. If an instruction does not exist in both user's manuals, it is described below.



# **Cobra Instruction Changes**



## **Cobra Facilities Overview**

The following registers are patterned after the registers in either the PowerPC 603 User's Manual (MPR603UMU-01,MPC603UM/AD) or the Power PC 401 Core User's Manual (v0.07 1/28/978). Differences between the two implementations and the Cobra Core implementations are detailed below.

The registers are split into these functional groupings: Machine Control/Status Registers, Branch Control Registers, Debug Control Registers, Special Purpose Facilities, Interrupt and Exception Registers, Timer Registers, Cache Control Registers, and Translation Control Registers.

Source Key:

- BOTH same bit definitions as 603 and 401
- 401 same bit definitions as 401
- 603 same bit definitions as 603
- PPC same bit definitions as general Power PC architecture
- COBRA COBRA Configuration



#### **Preliminary IBM Processor for Network Resources**

# **Machine Control/Status Registers**



# **Branch Control Registers**



# **Debug**



# **Special Purpose Facilities**





# **Interrupt and Exception Registers**



# **Timer Registers**





## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Community Preliminary**

# **Cache Control Registers**



# **Translation Registers**



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **Translation Registers (Continued)**



# **Exception Vector Override Registers**





## **Preliminary IBM Processor for Network Resources**

## **Exception Vector Override Registers (Continued)**



# **Internal Debug Access Address Map**



# **Cobra Specific Register Definitions**

For most of the registers named above, either a 40x or 60x spec will give the bit definition. There are a few registers which are different enough that they are described in detail below.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

## <span id="page-455-0"></span>**17.1: Hardware Implementation Detail 0 Register (HID0)**

Enables caches and controls architecture specific functionality. This register controls whether Cobra Core acts as a 40x or 60x series processor and allows the different features of each architecture to be individually selected.



**Type** Read/Write

**Address** 1023

**Power on Reset value** X'00 00 00 FC' (40x mode)

X'80 00 00 00' - Alternative value (60x mode suggested value)

# **Restrictions**







# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **17.2: Machine State Register (MSR)**

Controls the run time state of the Cobra Core.









# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **17.3: Exception Status Register (ESR)**

When an exception occurs, this register is updated to indicate which condition caused the exception. If an exception vector can only be reached by one exception, this register is cleared.









# **17.4: Machine Check Enable Register (MCHK)**

When an exception occurs, this register is updated to indicate which condition caused the exception. If an exception vector can only be reached by one exception, this register is cleared.







# IBM3206K0424



# **IBM Processor for Network Resources Accessor for Network Resources** Preliminary





# **PCORE Register Definitions**

# **17.5: PCORE Control Register**

The PCORE Control Register provides control information about PCORE operations. This is the PCORE control register. It is used to control operation.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.





## IBM3206K0424











# **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources All and Service Construction Construction Preliminary**



# **17.6: PCORE Reset Control Register**

The PCORE Reset Control Register provides control information about PCORE reset operations. See [Note](#page-92-0)  [on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









**Preliminary IBM Processor for Network Resources**

# **17.7: PCORE Status Register**

The PCORE Status Register provides status information about PCORE operations. See Note on Set/Clear [Type Registers on page 93](#page-92-0) for more details on addressing.







#### IBM3206K0424



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **17.8: PCORE User Status Register**

The PCORE User Status Register provides user-defined status information about PCORE software opera-tions. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.






## **17.9: PCORE Cobra Core External Status Register**

The PCORE Cobra Core External Status Register provides Cobra Core-defined status information about PCORE. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







### IBM3206K0424



## **IBM Processor for Network Resources Accessor For Accessor Preliminary**





## **17.10: PCORE Cobra Core External Machine Check Status Register**

The PCORE Cobra Core External Machine Check Status Register provides Cobra Core machine check status information about PCORE. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







### IBM3206K0424



## **IBM Processor for Network Resources Accessor For Accessor Preliminary**





## **17.11: PCORE JTAG Debug Control Register**

The PCORE JTAG Debug Control Register enables the JTAG port or a PCI interface processor debugger to control the processor core.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **17.12: PCORE JTAG Debug Status Register**

The PCORE JTAG Debug Status Register returns core status.







 $\mathbf{r}$ 

# **17.13: PCORE JTAG Instruction Stuff Buffer**

The PCORE JTAG Instruction Stuff Buffer is used to insert an instruction into the processor for execution.



User Defined







#### **17.14: PCORE JTAG Debug Data Register**

The PCORE JTAG Debug Data Register enables passing data between the JTAG/PCI debug port and the general purpose registers of the processor core.







## **17.15: PCORE Cobra Core Boot Address**

The PCORE Cobra Core Boot Address provides Cobra Core its boot time address.



This is the PCORE register. It is used to provide the address that is used to fetch the first instruction.







## **17.16: PCORE Cobra Core Access Priority Control Register**

The PCORE Cobra Core Access Priority Control Register provides Cobra Core defined status information about PCORE. It is used to control the order and priority of access to the various memory subsystems that the Cobra Core has access to. It is to be set up at initialization time an not dynamically changed.









## IBM3206K0424

## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

## **17.17: PCORE Transaction Dead Man Timer Value Registers**

These registers are used to load timers that count to zero from the value loaded in this register. The maximum wait for an I/O transaction is about 2ms when this is set to X'FFFF'. The value of this register is written into the corresponding timer after the transaction is initiated with the target. It continues to then count down until the target responds or zero is reached in the timer. When the timer reaches zero, a status bit is set and action can be taken from there.





## **17.18: PCORE High Priority Access Timer Value Registers**

These registers are used to load timers that count to zero from the value loaded in this register. The maximum wait for an I/O transaction is about 2ms when this is set to X'FFFF'. The value of this register is written into the corresponding timer after the transaction is initiated with the target. It will continue to then count down until the target responds or zero is reached in the timer. When the timer reaches zero, a status bit is set and action can be taken from there.



## **17.19: PCORE Transaction Dead Man Timer Register**

These timers are used to time transactions that are valid but the target does not respond right away. The timer counts on a 7.5 ns time base. The maximum wait for an I/O transaction is about 2ms when the timer counts down from X'FFFF'. This timer counts down to zero from the values set in the value register. When zero is reached, the transaction is considered broken and the request will be acknowledged back to the requestor.



### **17.20: PCORE IBM3206K0424 Shadow Status Register**

This register is used to shadow the INTST Interrupt Source. The purpose of this register is to allow polling for IBM3206K0424 interrupts without having to use the PCI bus.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **17.21: PCORE IBM3206K0424 Packet Last Write with Error Address**

This register is used to shadow the INTST Interrupt Source. The purpose of this register is to store the address associated with the previous virtual write error.



### **17.22: PCORE IBM3206K0424 RXQUE Master Status Register**

This register is used to shadow the RXQUE Master Status Register. The purpose of this register is to allow fast access to RXQUE's Master Status Register without having to use the regular register interface.



# **17.23: PCORE IBM3206K0424 RXQUE Enabled Status Register 1**

This register is used to shadow the RXQUE Enabled Status Register 1. The purpose of this register is to allow fast read access of the RXQUE Enabled Status Register 1 without having to use the normal IBM3206K0424 register interface.





## **17.24: PCORE IBM3206K0424 RXQUE Enabled Status Register 2**

This register is used to shadow the RXQUE Enabled Status Register 1. The purpose of this register is to allow fast read access of the RXQUE Enabled Status Register 2 without having to use the normal IBM3206K0424 register interface.



## **17.25: PCORE IBM3206K0424 RXQUE Upper Queues Status Register**

This register is used to shadow the RXQUE Upper Queues Status Register. The purpose of this register is to allow fast read access of the RXQUE Upper Queues Status Register without having to use the normal IBM3206K0424 register interface.



## **17.26: PCORE IBM3206K0424 RXQUE Lower Queues Status Register**

This register is used to shadow the RXQUE Lower Queues Status Register. The purpose of this register is to allow fast read access of the RXQUE Lower Queues Status Register without having to use the normal IBM3206K0424 register interface.





### **17.27: PCORE DMAQS Master Status Register**

This register is used to shadow the DMAQS Master Status Register. The purpose of this register is to allow fast read access of the DMAQS Master Status Register without having to use the normal IBM3206K0424 register interface.



## **17.28: PCORE DMAQS Enabled Status Register**

This register is used to shadow the DMAQS Master Status Register. The purpose of this register is to allow fast read access of the DMAQS Master Status Register without having to use the normal IBM3206K0424 register interface.



## **17.29: PCORE RXQUE Queue Length Registers**

The PCORE RXQUE Queue Length Registers provide event enqueue queue lengths to the Cobra Core. Reads from this address will return event queue lengths from RXQUE.





## **17.30: PCORE DMAQS Queue Length Registers**

The PCORE DMAQS Queue Length Registers provide DMAQS queue lengths to the Cobra Core. Reads from this address will return DMAQS queue lengths from DMAQS.



### **17.31: PCORE Interrupt Enable Register**

This register is used to enable bits from the PCORE Status Register and potentially generate interrupts to the control processor. When both a bit in this register and the corresponding bit(s) in the PCORE Status Register are set, the PCORE interrupt to PCINT will be enabled. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing. See [PCORE Status Register](#page-466-0) on page 467 for the bit descriptions.



### **17.32: PCORE User Interrupt Enable**

[This register is used to enable an interrupt based on bits from the corresponding PCORE User Status Regis](#page-467-0)ter and potentially generate interrupts to the control processor. When both a bit in this register and the corre[sponding bit\(s\) in the \[TBD\] register are set, the PCORE status bit\(s\) will be set in the corresponding P](#page-467-0)CORE User Status Register. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing. See [PCORE User Status Register](#page-467-0) [on page 468](#page-467-0) for the bit descriptions.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **17.33: PCORE Cobra Core Interrupt Enable Register**

This register is used to enable bits from the PCORE Cobra Core External Status Register and generate interrupts to the Cobra Core processor. When both a bit in this register and the corresponding bit(s) in the PCORE Cobra Core External Status Register are set, the Cobra Core interrupt to the Cobra Core core will be enabled. See [Note on Set/Clear Type Registers on page 93](#page-92-0) [for more details on addressing. See](#page-466-0) [PCORE Status Regis](#page-466-0)ter [on page 467](#page-466-0) for the bit descriptions.



### **17.34: PCORE Cobra Core External Machine Check Enable Register**

This register is used to enable bits from the PCORE Cobra Core External Machine Check Status Register and generate machine checks to the Cobra Core processor. When both a bit in this register and the corresponding bit(s) in the PCORE Cobra Core External Machine Check Status Register are set, the requisite Cobra Core Machine Check to the Cobra Core core will be enabled. See Note on Set/Clear Type Registers [on page 93](#page-92-0) [for more details on addressing. See](#page-466-0) [PCORE Status Register](#page-466-0) on page 467 for the bit descriptions.



### **17.35: PCORE Error Lock Enable Register**

The PCORE Error Lock Enable Register provides the ability to halt PCORE when the corresponding status bit in the status register is set and locking is enabled. When a bit in this register corresponds to a bit that is set in the status register, the state machines in PCORE will be held in idle state until the lock is disabled.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.





### **17.36: PCORE User Error Lock Enable Register**

The PCORE User Error Lock Enable Register provides the ability to halt PCORE when the corresponding status bit in the User Status Register is set and locking is enabled. When a bit in this register corresponds to a bit that is set in the Status Register, the state machines in PCORE will be held in idle state until the lock is disabled.

See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



### **17.37: PCORE RXQUE Event Interface Enqueue Register**

The PCORE RXQUE Event Interface Enqueue Register provides event enqueue interface for the Cobra Core. Writes to this address will enqueue an event to an RXQUE queue.





### **17.38: PCORE DMAQS DMA Enqueue Register**

The PCORE DMAQS DMA Enqueue Register provides DMAQS enqueue interface for the Cobra Core. Writes to this address will enqueue an event to an RXQUE queue.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



### **17.39: PCORE RXQUE Event Interface Deque Register**

The PCORE RXQUE Event Interface Deque Register provides event deque interface for the Cobra Core. Reads from this address return an event.



### **17.40: PCORE Cobra SPR Read Data Access Register**

The PCORE Cobra SPR Read Data Access Register stores the data from the requested Cobra facility on a read. These are message passing facilities. They are used for inter-device communication.

These facilities, with their control register bits, allow for either interrupt or polling-based message passing from the Cobra Core to a PCI bus device.



### **17.41: PCORE Cobra SPR Write Data Access Register**

This register stores the data from the PCIrequested Cobra facility on a read. These are message passing facilities. They are used for inter-device communication.

These facilities with their control register bits, allow for either interrupt or polling-based message passing from the Cobra Core to a PCI bus device.





## **17.42: PCORE Cobra SPR Access Address Register**

This is the PCORE SPR Access Address Register. It is used to access the internal facilities in Cobra. This includes SPR/DCR and Debug facilities. The address is for a facility and represents a four-byte access.







#### **IBM Processor for Network Resources All and Service Construction Construction Preliminary**



## **17.43: PCORE Address Translation Offset Address Facilities**

The PCORE Address Translation Offset Address Facilities provides the offset that is added to the Cobra Real Address to create the target subsystem address.

When an address is issued from the Cobra Core core it is accompanied by four target translation bits. The translation bits indicate which translation facility is to be used to translate the processor "real" physical address into a target system actual address. This grouping provides for the offset addresses for each target memory system. The offset is added to the Cobra Real Address to create the target system address. The following is a list of targets, each with their own translation facilities.



PCORE Address Translation Offset Address Facilities:



IBM3206K0424



## **Preliminary IBM Processor for Network Resources**



# **17.44: PCORE PCI 64 Bit Address Translation Facilities**

The PCORE PCI 64 Bit Address Translation Facilities provide the upper thirty-two bits of address in 64-bit addressing mode. When an access is issued to the PCI Master Interface in 64-bit addressing mode, these registers are used to create the upper 32 bits of the 64-bit address.





## **17.45: PCORE PCI Master Target Tag Controls**

The PCORE PCI Master Target Tag Controls contains the control for each PCI Tag/View. This register contains bits for each of the four PCI Master Views.









## IBM3206K0424

## **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **17.46: PCORE Last Packet Address Register**

The PCORE Last Packet Address Register is the last address to the Packet Memory bus at the time of the hang condition. When the system locks up, this register holds the last Packet Memory address that was or is currently being presented to the Packet Memory subsystem.



## **17.47: PCORE Last Control Address Register**

The PCORE Last Control Address Register is the last address to the Control Memory bus at the time of the hang condition. When the system locks up, this register holds the last Control Memory address that was or is currently being presented to the Control Memory subsystem.



## **17.48: PCORE Last PCI Lower Address Register**

The PCORE Last PCI Lower Address Register is the last address to the PCI bus at the time of the hang condition. When the system locks up, this register holds the last PCI bus address that was or is currently being presented to the Control Memory subsystem.





## **17.49: PCORE Last Register Address Register**

The PCORE Last Register Address Register is the last address to the PCI bus at the time of the Hang Condition. When the system locks up, this register holds the last PCI bus address that was or is currently being presented to the Control Memory subsystem.



# **17.50: PCORE SRAM Base Address**

The SRAM Base Address register is used to select the base address of the 4K byte window to access the SRAM.





#### **17.51: PCORE Read Data Transfer Buffers**

The PCORE Read Data Transfer Buffers hold the read data that is being transferred from one of the target subsystems and the Cobra Core. Eight bytes are buffered on the interfaces except for the IBM3206K0424 register interface which buffers four bytes.



## **17.52: PCORE Write Data Transfer Buffers**

The PCORE Write Data Transfer Buffers hold the data that is being transferred between the Cobra Core and one of the target subsystems. Eight bytes can be stored for each target subsystem, with the exception of the IBM3206K0424 Register Target which holds just four bytes.





ិ ន

# **17.53: PCORE Polling Register**

The PCORE Polling Register provides status information to PCORE about IBM3206K0424 operations. It allows PCORE to poll specific IBM3206K0424 status without using PCI bus bandwidth.







## **17.54: PCORE Integer Input Rate Conversion Register**

This register is the integer input port for the rate conversion logic. An integer rate is placed in this register. The on board logic converts it to an ABR rate format.





## **17.55: PCORE ABR Output Rate Register**

This register is the output port of the rate conversion logic. An integer rate was placed in the Integer Input Register. The logic converts it to an ABR rate and places the result in this register.



## **17.56: PCORE Debug States Control**

This register serves as the PCORE control for external debug states. The INTST Debug states control for the address range desired must be set to select these PCORE state bits. If that is done, then this register acts to select the four ranges. See bit descriptions below.









# **17.57: PCORE Debug States Config**

This register serves as the PCORE configuration for external debug states. The INTST Debug states control for the address range desired must be set to select these PCORE state bits. If that is done, then this register acts to select the characteristics according to the bit descriptions below.









# **Entity 18: PowerPC On-Chip Memory (PPOCM) Entity**

The PPOCM entity is comprised of several SRAM arrays that provide a xxxK memory that may be used by the internal processor or by the IBM3206K0424. Also included in PPOCM is a DMA controller that the processor may use to do bulk data moves between the SRAM arrays and Control Memory, Packet Memory, or memory on an external PCI device. The PPOCM arrays will subsequently referred to as on-chip memory and the three external memories (control, packet, PCI) just mentioned will subsequently be referred to collectively as off-chip memory.

## **DMA Controller**

The DMA controller moves data in eight byte aligned, eight byte portions. In real addressing mode, up to 64K bytes may be transferred at once. In virtual addressing mode, there are more restrictions. The DMA must remain within the virtual 4K page for both the PPOCM array address and the off-chip memory address.

## **18.1: PPOCM Control Register**

This register contains information which controls the functions of the entity. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing.









#### **Preliminary IBM Processor for Network Resources**



## **18.2: PPOCM Status Register**

This register contains status information that can be used to generate interrupts. See Note on Set/Clear Type [Registers on page 93](#page-92-0) for more details on addressing.







#### IBM3206K0424



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **18.3: PPOCM Interrupt Enable Register**

This register enables the bits of the Status Register to generate an interrupt. The bits of this register corre-spond to the bits of the status register. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.



## **18.4: PPOCM DMA Off-Chip Effective Address Register**

This register provides the DMA controller the effective address of the off-chip portion of the DMA.









# **18.5: PPOCM DMA On-Chip Effective Address Register**

This register provides the DMA controller the effective address of the on-chip portion of the DMA.







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

# **18.6: PPOCM DMA Length Register**

This register provides the DMA controller the length of the DMA. The maximum DMA length in real addressing mode is X'00010000' (64K). The maximum DMA length in virtual addressing mode is X'00001000' (4K).





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

### **18.7: PPOCM DMA Timeout Timer Register**

This register is compared to a timer that begins running when bit 0 of the control register is set to '1'. When the timer reaches the value in this register, the DMA is terminated and a status bit is set. The default value of X'FFFFFF' results in a timeout value of 125 ms.








inal<br>rinal

# **Entity 19: RS-232 Interface Logic (RS-232)**

The RS232 entity provides a means by which an external debugger and the processor core can communicate. The RS-232 operates a one-or four-byte wide basis.

# **RS-232 Interface Logic Registers**

## **19.1: RS-232 Control Register**

This register controls the operation of the RS-232 logic. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.







#### IBM3206K0424



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **19.2: RS-232 Status Register**

This register controls the operation of the RS-232 logic. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









# **19.3: RS-232 Interrupt Enable Register**

This register contains bits corresponding to the bits in the RS-232 status register. If a bit in this register is set and the corresponding bit is set in the RS-232 status register, an interrupt is generated. Bits six through four generate a transmit interrupt and bits three through zero generate a receive interrupt. See Note on Set/Clear [Type Registers on page 93](#page-92-0) for more details on addressing.



#### **19.4: RS-232 Transmit Buffer**

This register contains the data to be sent over the RS-232 connection.



Transmit Data









#### **19.5: RS-232 Receive Buffer**

This register contains the data received over the RS-232 connection. Once this buffer is full, software has four byte receive times minimum to read it before an overrun condition can occur.







## **19.6: RS-232 Baud Rate Register**

This register contains the value used to determine the baud rate. The value to place in this register can be determined by this formula: BaudRate =  $133$ MHz/(8<sup>\*</sup>(Baud Rate Register + 1)).









### **19.7: RS-232 CTS/DSR Glitch Timer Rate**

This register contains the number of (baud rate/8) clocks CTS/DSR must be active/inactive before the state of CTS/DSR is considered valid. Transitions of shorter duration are assumed to be glitches.





# **19.8: RS-232 Reset Register**

This register resets the port. See [Note on Set/Clear Type Registers on page 93](#page-92-0) for more details on addressing.









# **19.9: RS-232 Error Forcing Register**

This register can be used by diagnostics in a wrap environment (external or internal) to force frame and parity errors. Overrun errors can be generated by sending four bytes, not reading the receive buffer, and sending four more bytes.









# **Entity 20: Reset and Power-on Logic (CRSET)**

This entity performs BIST and flush operations. Chip software resets can be controlled by this entity, as well as the chip clock control.

# **Reset and Power-on Logic Registers**

# **20.1: Reset Status Register**

This register is used to reflect the last type of reset was. A hardware reset will clear software reset status bits, but a software reset will not have an affect on the hardware status bits.







#### **IBM Processor for Network Resources All and Service Construction Construction Preliminary**



#### **20.2: Software Reset Enable Register**

This register protects the Software Reset Register. If this register is not set, then a reset will not occur. Write a X'B4' to this register to enable software reset. A software reset will clear this register.



#### **20.3: Software Reset Register**

This register generates a scan path flush reset of the chip, or software initiated run of BIST, with the exception of the registers in the reset entity.









# **20.4: Memory Type Register**

This register indicates the type of memory used for control and Packet Memory so that reset hardware will know how to properly preserve it during a reset.









# **20.5: CRSET PLL Range Debug**

Used to debug the PPL operation.







# **20.6: CRSET Control Register**

Used to control PCI frequency detection logic.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **20.7: Clock Control Register (Nibble Aligned)**

Used to disable clocks for power conservation and provide the Select A Clock function for MPEG and front end support. To change a nibble field in this register, always set it to '0' first, and then to the new value.









#### **Preliminary IBM Processor for Network Resources**



# <span id="page-516-0"></span>**"Select A Clock" Selection Matrix**





# **20.8: CBIST PRPG Results**

This is the PRPG results register, updated after BIST has run. It is used by the BIST function for chip test.



# **20.9: CBIST MISR Results**

This is the MISR results register, updated after BIST has run. It is used by the BIST function for chip test.



#### **20.10: CBIST BIST Rate**

This register holds a counter value that separates the time between when the A clock and the B clock are launched during BIST. This allows finer tuning to how much power BIST uses versus how much testing gets done within the time allowed. It is used by the BIST function for chip test.



# **20.11: CBIST PRPG Expected Signature**

This is the PRPG signature register, which should be written by CRISCO code with the expected value of signature, based on the value in CBIST CYCT Load Value and the clock selected for BIST to run from. It is used by the BIST function for chip test.





# **20.12: CBIST MISR Expected Signature**

This is the MISR Signature Register, which should be written by CRISCO code with the expected value of signature, based on the value in CBIST CYCT Load Value and the clock selected for BIST to run from. It is used by the BIST function for chip test.



# **20.13: CBIST CYCT Load Value**

This register is the loaded value for the CBIST BIST Rate Register. The time for BIST to run can be computed by the following equation: (shift count) \* (c30 clock\*2) \* (cycle time). It is used by the BIST function for chip test.





# **Entity 21: JTAG Interface Logic (CJTAG)**

The CJTAG entity contains logic to support a test access port (TAP) controller compliant with the IEEE 1149.1-1993 standard. The TAP controller is accessed via the following five pins:



The proper operation of these signals and the TAP controller is defined in the IEEE 1149.1-1993 standard.

#### **Scanning**

The TAP controller supports two types of scans: instruction scans and data scans. Instruction scans control the type of operation and select which (if any) scan chains are involved in the operation. Data scans generally clock the data on TDI into the selected scan chain.



# **21.1: Instruction Format**

The processor's JTAG logic supports 32-bit instructions in one of two formats: the first format uses opcodes compliant with the IEEE standard; the other supports opcodes as defined by the Walnut chip that are compatible but not compliant with the IEEE standard. As an instruction is scanned in, status for the previous instruction is presented on TDO.



#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**

#### **Instructions**

The following instructions are supported:

# **21.2: IDCODE**

Returns a 32-bit identification code when a data scan is performed. The IDCODE has the following structure:

# **Opcode** X'0300XXXX'





# **21.3: SAMPLE/PRELOAD**

Captures the state of the boundary scan I/O. As the values captured are scanned out, new values can be loaded into the boundary scan latches. This operation will not affect functional operation.

**Opcode** X'0402XXXX'

# **21.4: EXTEST**

Drives the values in the boundary scan latches onto their respective I/O. This function can be used in conjunction with SAMPLE/PRELOAD to perform card wire tests.

**Opcode (Compliant)** X'00000000' **Opcode (Compatible)** X'0600XXXX'

#### **21.5: BYPASS**

Selects the single bit bypass register for data scans.

**Opcode (Compliant)** X'FFFFFFFF'

**Opcode (Compatible)** X'FFFFXXXX' or X'0000XXXX'





**Preliminary IBM Processor for Network Resources**

### **21.6: RUNBIST**

Causes built in self test (BIST) to execute.

**Opcode** X'0770XXXX'

### **21.7: BIST\_RESULTS**

Returns a 64-bit value when a data scan is performed. Bits 63-32 are the PRPG and bits 31-0 are the MISR from the BIST logic.

**Opcode** X'1F02XXXX'

#### **21.8: WALNUT\_MODE**

This command enables Walnut compatible mode.

**Opcode** X'3000'

#### **21.9: COMPLIANT\_MODE**

This command enables JTAG compliant mode.

**Opcode** X'33010000'

#### **21.10: STOP**

This command halts the functional clocks of IBM3206K0424 in anticipation of a scan. After the STOP command is scanned in, a data scan that takes the TAP controller through the Capture-DR, Exit1-DR, and Update-DR states should be performed. This will capture the state of the I/O so that they can be held in a known state if a scan command is issued.

**Opcode** X'2002'

#### **21.11: SCAN**

This command causes TDI to be clocked into the scan chain during a subsequent data scan. The scan out of the scan chain is placed on TDO. This command will not work unless a STOP command is sent down immediately before the SCAN command is issued.

**Opcode** X'0802'

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



This command causes TDI to be clocked into the scan chain during a subsequent data scan. TDO is forced to '0'. This command will not work unless a STOP command is sent down immediately before the SCAN\_IN command is issued.



# **21.13: SCAN\_OUT**

This command causes the scan out of the scan chain to be placed on TDO. Data is recirculated through the scan chains. TDI is ignored. This command will not work unless a STOP command is sent down immediately before the SCAN\_OUT command is issued.

**Opcode** X'0A00'

## **21.14: Private\_RW1**

This command is used by RISCWATCH.

**Opcode** X'0500'

## **21.15: Private\_RW2**

This command is used by RISCWATCH.

**Opcode** X'0582'

**21.16: Private\_RW3**

This command is used by RISCWATCH.

**Opcode** X'05C0'



# **[S](#page-524-0)onet Framer Core (FRAMR Chiplet Address Mapping)**



# <span id="page-524-0"></span>**FRAMR Chiplet Address Mapping**

# **GPPINT Architecture**

# <span id="page-524-1"></span>**[O](#page-539-0)verview**

The General Purpose Processor INTerface (GPPINT) provides direct access to registers located in the GPPINT module; it provides delayed access to registers and counters located in the GppHandler modules of the various chiplets of the SONET core. GPPINT controls the handshaking with the external microprocessor as well as the handshaking with the GppHandlers at the asynchronous chiplet interfaces. Address decoding is done to the chiplet level in GPPINT. In addition, addresses are decoded to the register level for the local GPPINT registers.

# **Reset Register**

Each chiplet is controlled by one reset bit. At power-on, all reset bits are active and the chiplets are disabled. They can be released by the General Purpose Processor (GPP) only after all global configuration parameters have been set and the clocks to the chiplets have been established. In addition, there are reset bits for the chiplets that do not have their own GppHandler.

# **Interrupt Registers**

The interrupt register is used as a pointer to the chiplet interrupt registers with pending requests: the clock status error register, and the handshaking error register. An active bit of the interrupt register is reset by removing the cause for the request in the corresponding chiplet or by masking the active IRQ bit(s) in the chiplet; therefore, the interrupt registers (including the pointer) are read only. All interrupt and pointer registers have a corresponding MASK register (R/W). Every unmasked, active interrupt bit causes an active pointer bit. Every unmasked, active pointer bit causes activation of the interrupt signal to the microprocessor.

#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



#### **Handshaking Error Registers**

Each bit of the handshaking error registers indicates a locked interface to one of the chiplet GppHandlers. Two additional bits indicate various timeout events. To reset an individual bit of the handshaking error register, the cause for the request must be removed and a one must be written into the bit location of the register (R/W). Reading the register will reset the whole (eight-bit) register if the corresponding "clear-register" option is set in the configuration register. The handshaking error indication register has a corresponding MASK register (R/W). Every unmasked, active handshaking error bit causes activation of the pointer bit in the GPPINT interrupt register.

## **Clock Monitor Status Registers**

The clock monitor status register bits indicate the loss of a specific chiplet's clock. They are set whenever a difference between the clock test signal and the individual chiplet clock acknowledge signal occurs after one clock monitor test period. To reset an individual bit of the clock monitor status registers, the clock of the corresponding chiplet must be restored and a one must be written into the bit location of the register (R/W). Reading one of the registers will reset the whole (eight-bit) register if the corresponding "clear-register" option is set in the configuration register. The clock monitor status register has a corresponding MASK register (R/W). Every unmasked, active clock monitor status bit causes activation of the pointer bit in the GPPINT register.

# **Local Gppint Configuration Registers**

There are registers (R/W) for the Clock Monitor Test Period, the Watchdog Timer Period and the "clear-register" option. A read-only register provides the Vital Product Data (VPD).

# **Global Static Configuration Registers**

These are configuration parameters that are shared by many chiplets or that are needed by chiplets that have no GppHandler. The initial values can be modified by the microprocessor after power-on, but should not be changed later on. All global static configuration registers are R/W.

#### **Status Registers**

These registers provide status information from chiplets that have no GppHandler and are read only. Presently, there is only one status register for the SIM chiplet (PLL lock status).



#### **Preliminary IBM Processor for Network Resources**

<span id="page-526-0"></span>

# **GPPINT Chiplet Address Mapping Overview: Base Address = x'C00'**



# **22: GPPINT Register Description**

# **22.1: Chiplet Reset Register (RESGP)**

The bits of the chiplet reset register control the resetting (enabling/disabling) of complete chiplets.

For each bit position:

 $0 =$  Reset inactive for this chiplet

1 = Reset active (chiplet is disabled; DEFAULT).









# **22.2: Chiplet Interrupt and Mask Registers (IRQGP1 (IRMGP1))**

The chiplet interrupt request register indicates pending interrupt requests from individual chiplets. An active bit of this register is reset by removing the cause for the request in the corresponding chiplet or by masking the active IRQ bit(s) in the chiplet; therefore, this register is read only.

For each bit position:

0 = No chiplet interrupt request pending.

1 = Chiplet has pending interrupt request(s). The chiplet interrupt request mask register bits control the propagation of a chiplet interrupt request to the Sonet Macro Interrupt output pin. The mask registers allow read and write access.

For each bit position:

0 = The corresponding interrupt request bit is masked (DEFAULT).

1 = The corresponding interrupt request bit is active (for IRMGP1, the corresponding interrupt request bit activates the Sonet Macro Interrupt).







#### **IBM Processor for Network Resources All and Service Construction Construction Construction Preliminary**



# **22.3: Handshaking Error Indication and Mask Registers (HShake1)**

The local handshaking error indication register indicates pending handshaking error requests from the GPPINT chiplets.

For each bit position:

 $0 =$  Normal operation of the corresponding chiplet.

1 = The corresponding chiplet did not deassert its DTACK signal.

**Exception:** The signals TOError and IntError (HShake2(1-0)) have the following meaning: Normal operation GPP deasserts Strobes without waiting for DTACK assertion Watchdog Timeout in REST state Watchdog Timeout in REQ state. An active bit of the handshaking error indication register is reset by removing the cause for the malfunctioning of the chiplet and by writing a one into the corresponding bit position. Reading one register will reset all bits of this register if the "clear-register" option is set in ConfGP1(2). The handshaking error indication mask register bits control the propagation of the GPPINT handshaking error requestof the register HShake1. HSMask1 controls propagation to the signal FElocHS (bit 0 of IRQGP1 register). The mask registers allow read and write access.

For each bit position:

0 = The corresponding handshaking error indication bit is masked (DEFAULT).

1 = The corresponding request bit is active (for HSMask1, the corresponding request bit activates signal FElocHS (bit 0 of IRQGP1 register). "Clear-register" option set in ConfGP1(2).









#### **22.4: Clock Monitor Status and Mask Registers (ClkStat1 (ClkMask1))**

The clock monitor status register bits indicate the loss of a specific island's clock. They are set whenever a difference between the clock test signal and the individual island's clock acknowledge signal occurs after the clock monitor test period.

For each bit position:

0 = Normal operation of the corresponding clock island

1 = The corresponding island clock is lost. An active bit of this register is reset by restoring the clock of the corresponding clock island and by writing a one into the corresponding bit position. Reading one register will reset all bits of this register if the "clear-register" option is set in bit ConfGP1(3). The clock monitor mask register ClkMask1 controls the propagation of active clock monitor status signals. ClkMask1 controls propagation to the signal FElocCS (bit 1 of IRQGP1 register). The mask registers allow read and write access.

For each bit position:

 $0 =$ The corresponding clock status bit is masked (DEFAULT).

1 = The corresponding clock status bit is active (for ClkMask1, the corresponding bit activates the signal FElocCS (bit 1 of IRQGP1 register).









#### **22.5: Clock Monitor Test Period Register (CMonGP1)**

Divider ratio to derive the clock monitor test period from the GPPCLK clock. Clock monitoring is disabled if equal x'00' (DEFAULT).



CMonGP1



# **22.6: Watchdog Timer Period Register (WDTGP1)**

Divider ratio to derive the interface timeout period from the GPPCLK clock. This register is reset to x'FF' whenever a timeout occurs; it has to be reconfigured by a GPP write access.



WDTGP1







# **22.7: GPPINT Local Configuration Registers (ConfGP1)**

The bits of this local configuration register control the resetting of complete registers upon read access ("clear register" option).

For each bit position:

- $0 = No$  action upon read access.
- 1 = The corresponding register is reset upon read access (DEFAULT).









# **22.8: Vital Macro Data Register (VPD)**

This read-only register displays the macro identification.







# **22.9: Static Configuration Register (GATMCS)**

Common static configuration data, providing control signals that are distributed to multiple chiplets. Set once by the GPP before the individual chiplets get enabled and not changing during normal operation.









**Preliminary IBM Processor for Network Resources**

**22.10: GCasc**







# **22.11: GLoopTx**

Transmit loopback control.

For each bit position:

0 = ACH Loopback disabled (DEFAULT).

1 = ACH Loopback enabled.







#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**

# **22.12: GLoopRx**

Receive loopback control.

For each bit position:

0 = ACH Loopback disabled (DEFAULT).

1 = ACH Loopback enabled.







# **22.13: GExtRes**

External clock recovery circuit reset signal. Delivered to external circuit (deserializer) via device pins. The active level depends on the external circuit used. Default value at power-on-reset is LOW.











**Preliminary IBM Processor for Network Resources**

# **22.14: OFPTXGP**

Static configuration data, providing control signals for chiplet OFP\_Tx. Set once by the GPP before the individual chiplets are enabled and not changing during normal operation.







#### **22.15: OFPRXGP1**







**IBM Processor for Network Resources All and Service Preliminary Preliminary** 



# **22.16: OFPRXGP2**







### **22.17: PIMRConf2**

Static configuration data, providing control signals for chiplets PIM\_Tx/PIM\_Rx. Set once by the GPP before the individual chiplets are enabled and not changing during normal operation.









#### **22.18: SIMStat**

Status register, providing the GPP with information from the SIM chiplet via PIM. Either SIM-internal or external PLL lock status. "Clear-register" option set in ConfGP1(1).









# **23: GPPHandler Architecture**

#### <span id="page-539-0"></span>**[O](#page-524-1)verview**

All GPP handlers for the various chiplets have the following general register structure.

#### **GPPHandler Architecture**



#### **Counter Registers**

Every counter has an enable bit in the counter enable register (addr 2 or 3), and optionally up to two programmable thresholds. Each counter has an interrupt bit for overflow and up to two interrupt bits for threshold crossing in the counter interrupt registers. For all counters in one handler there is one common 'read-on-the-fly register' that is used to store the higher order bytes to obtain a correct readback value for counters larger than eight bits. Counters are read-only registers; the count enable registers are read/write. Note: COUNTER reading is independent of the counter length, given that a counter has address n as base, reading address n or address n-1 both yield the least significant byte of the counter. Reading address n has no influence on the counter, but reading address n-1 will reset the counter after the read. Reading address n or n-1 will always latch the higher order bytes into the read on the fly register (before the optional automatic reset). Counters can only be read and not written to. For a 16-bit counter, the most significant byte should be read from ROFmid (address 0). For a 24-bit counter, the most significant byte is read from ROFhi (address 1), the next byte from ROFmid (address 0). To completely read a 24-bit counter: first read least significant byte from counter address n or n-1, then read ROFmid and ROFhi (address 0; address 1).

#### **Reset Registers**

Each handler has a two-bit reset register. Bit 0 is the chiplet reset control. This bit is active high after power on reset, causing the chiplet to be disabled. Bit 1 is the chiplet halt signal, which for selected chiplets freezes the state machines for diagnostic purposes. This is a read/write register.

#### **Command Registers**

The optional command register(s) will generate events to the chiplet. When a bit is written high by the microprocessor, it will remain high for one chiplet clock cycle. Therefore, reading back a command register will always read back zeroes. This is a read/write register.


## **Event Latch Registers**

The optional event latch register(s) remember one ore more occurances of events that happen in a chiplet. This may be considered as a one-bit saturating counter. Each bit in the register corresponds to an event in the chiplet. Such bits remain high after the event happened until the microprocessor implicitly or explicitly resets the bit. This is configurable: implicit reset is done by writing a high value to the bit that is to be reset. Explicit will reset all bits of one register when the register is read. This is a read/write register.

# **Interrupt Registers**

When there are counters, user interrupts, or fatal bits in a chiplet, a MAIN INTERRUPT register will be present. Bit 0 always is the fatal interrupt bit, which is set as soon as any of the fatal interrupt events occur. The other bits refer to counters or user interrupt registers to allow easy determination of the interrupt cause. Each Interrupt register has an interrupt MASK register to enable or disable interrupt. After power on Reset, interrupts are disabled. The interrupt registers are the same as the event latch registers, with the addition that when an interrupt register bit is set, and the corresponding mask register bit is set, the interrupt signal to the GPPINT chiplet is activated. The same mechanism to reset the interrupt register bits is used as for the event latch registers. The interrupt MASK registers are only changed by the microprocessor. The interrupt and interrupt mask registers are read/write.

# **Configuration Registers**

These registers are programmed by the microprocessor with setup information, and are read/write. The first configuration register reserves bit 1 and seven to configure explicit or implicit reset of the event latch registers and interrupt registers respectively (when such registers are present).

# **Register Types**

- **F** Read-On-The-Fly register (auto-generated)
- **N** Counter register
- **R** Reset register
- **I** Interrupt register (auto-generated)
- **C** Configuration register
- **X** Control or mask register (auto-generated)
- **S** Status (event latch) register
- **O** Command register



# **ATM Cell Handler Architecture : Transmit Direction**

# **ACH\_Tx GPP Handler Address Mapping** Base Address = x'100'



<span id="page-541-1"></span>1. Defaults according ITU I.432

<span id="page-541-0"></span>2. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation.



# **24: ACH Tx Register Description**

# **Counter Registers**

# **24.1: ROFmid**

Read-on-the-fly register, middle significant byte.



ROFmid





# **24.2: ROFhi**

Read-on-the-fly register, most significant byte.









# **24.3: ACBC**

Number of cells read from external FIFO (24-bit counter). Overflow leads to an interrupt request.







# **24.4: IUC**

Number of transmitted idle and unassigned cells (24-bit counter). Overflow leads to an interrupt request.









**Preliminary IBM Processor for Network Resources**

# **24.5: ACBE**

Number of errors (corrupted cell read from external FIFO). Eight-bit counter overflow leads to an interrupt request.



ACBE 7 6 5 4 3 2 1 0





# **24.6: ACBETh11**

Threshold for number of errors. Threshold overstep leads to an interrupt request.







# **24.7: CntEn1**

Counter On/Off control register for ACH\_Tx.

- $0 =$  Counter is disabled.
- $1 =$  Counter is enabled.









# **24.8: Reset Register (RESET)**

Reset/Halt chiplet control register. This register is automatically preset to the default value by the reset signal ResHT from the GPPINT.

- $0 =$  Reset/Halt not active.
- 1 = Reset/Halt active.







## **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **Status Registers**

## **24.9: STAT1**

Status register #1 of this chiplet. This is an event latch register.









# **24.10: IUCSTAT1**

Status register #2 of this chiplet. This is an event latch register.



**Power On Value** 





# **Interrupt Request and Mask Registers**

# **24.11: MainIRQ**

Register to indicate fatal interrupt events and to point to user IRQ registers with active requests.

- $0 = No$  interrupt request pending.
- 1 = Interrupt request pending.









# **24.12: M\_MainIRQ**

Register to mask pending interrupt requests. A masked request will not generate an outgoing IRQ to the GPPINT.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates signal IRQHT1 to GPPINT.









# **24.13: CntrIRQ1**

Register to indicate active counter interrupt requests of this chiplet.

For each bit position:

0 = No interrupt request pending.

1 = Interrupt request pending.







# **24.14: M\_CntrIRQ1**

Register to mask pending counter interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.











# **Configuration Registers**

# **24.15: CELLTENABLE**

Register to control various modes of operation of this chiplet.









# **24.16: ACBTXTHRPAE**

Threshold for Programmable Almost Empty flag of external FIFO in transmit direction.







# **24.17: SDBTXTHRPAF**

Threshold for Programmable Almost Full flag (SDB\_Tx).









**Preliminary IBM Processor for Network Resources**

# **24.18: HEADERBYTE1**

Idle/Unassigned cell header byte #1. Default pattern according to ITU I.432.



HEADERBYTE1





# **24.19: HEADERBYTE2**

Idle/Unassigned cell header byte #2. Default pattern according to ITU I.432.



HEADERBYTE2







# **24.20: HEADERBYTE3**

Idle/Unassigned cell header byte #3. Default pattern according to ITU I.432.



HEADERBYTE3





# **24.21: HEADERBYTE4**

Idle/Unassigned cell header byte #4. Default pattern according to ITU I.432.



HEADERBYTE4

7 6 5 4 3 2 1 0





**Preliminary IBM Processor for Network Resources**

# **24.22: HEADERBYTE5**

Idle/Unassigned cell header byte #5. Default pattern according to ITU I.432.



HEADERBYTE5





# **24.23: PAYLOADBYTE**

Idle/Unassigned cell payload byte.



PAYLOADBYTE





## **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



HEC processing control configuration register.











# **24.25: HECOFFSET**

HEC offset pattern register for the byte pattern used in the ATM cell header HEC calculation as base offset according to ITU I.432.



HECOFFSET



# **24.26: HECMASKAND**

HEC mask pattern register for the byte pattern used in the ATM cell header HEC calculation as dedicated (ANDing) HEC error corruption mask.



HECMASKAND





## **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



# **24.27: HECMASKOR**

HEC mask pattern register for the byte pattern used in the ATM cell header HEC calculation as dedicated (ORing) HEC error corruption mask.



HECMASKOR





## **Preliminary IBM Processor for Network Resources**

# **ATM Cell Handler Architecture: Receive Direction**

## **ACH\_Rx GPP Handler Address Mapping** Base Address = x'200'



<span id="page-560-0"></span>1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation.



# **ACH\_Rx Register Description**

# **Counter Registers**

# **24.28: ROFmid**

Read-on-the-fly registers, middle significant byte.







# **24.29: ROFhi**

Read-on-the-fly registers, most significant byte.



ROFhi 7 6 5 4 3 2 1 0





# **24.30: FHR**

Number of ATM cells written into external FIFO (24-bit counter). Overflow leads to an interrupt request.



FHR (16:23)





# **24.31: IHR**

Number of idle cells received from OFP\_Rx (24-bit counter). Overflow leads to an interrupt request.



IHR (16:23) 7 6 5 4 3 2 1 0





## **24.32: EHR1**

Number of detected HEC errors (16-bit counter). Overflow leads to an interrupt request.







# **24.33: EHR1Th11**

Threshold for number of HEC cells, most significant byte.



EHR1Th11 7 6 5 4 3 2 1 0





# **24.34: EHT1Th12**

Threshold for number of HEC errors (least significant byte). Threshold overstep leads to an interrupt request.







# **24.35: BHR**

Number of discarded cells because of FIFO full condition (16 bit counter). Overflow leads to an interrupt request.



BHR (8:15)







## **24.36: BHRTh11**

Threshold for number of discarded cells, most significant byte.







# **24.37: BHRTh12**

Threshold for number of discarded cells (least significant byte). Threshold overstep leads to an interrupt request.



BHRTh12







**Preliminary IBM Processor for Network Resources**

# **24.38: CntEn1**

Counter On/Off control register for ACH\_Rx.

- $0 =$  Counter is disabled.
- $1 =$  Counter is enabled.









# **24.39: Reset Register (RESET)**

Reset/Halt chiplet control register. This register is automatically preset to the default value by the reset signal ResHR from the GPPINT.

For each bit position:

- $0 =$  Reset/Halt not active.
- $1 =$  Reset/Halt active.







# **24.40: Command Register (CMD1)**

Command register for this chiplet. Single-cycle active if '1' is written into bit position.









**Preliminary IBM Processor for Network Resources**

# **24.41: Status Register (STAT1)**

Status register of this chiplet. This is an event latch register.



**Power On Value** 





### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



## **Interrupt Request and Mask Registers**

# **24.42: MainIRQ**

Register to indicate fatal interrupt events and to point to user IRQ registers with active requests.

- $0 = No$  interrupt request pending.
- $1 =$  Interrupt request pending.







# **24.43: M\_MainIRQ**

Register to mask pending interrupt requests. A masked request will not generate an outgoing IRQ to the GPPINT.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates signal IRQHR1 to GPPINT.







## **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **24.44: CntrIRQ1**

Register to indicate active counter interrupt requests of this chiplet.

For each bit position:

0 = No interrupt request pending.

1 = Interrupt request pending.



**Power On Value** 









**Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre** 

# **24.45: M\_CntrIRQ1**

Register to mask pending counter interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







## **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

# **Configuration Registers**

# **24.46: CONF5**

Register to control various modes of operation of this chiplet.











# **24.47: CONF6**

Register to control ATM cell synchronization in this chiplet.







# **24.48: CONFC**

Threshold for Programmable Almost Full flag of the external FIFO.









# **24.49: H1CONF**

Header pattern #1 to identify idle/unassigned cells.









# **24.50: H2CONF**

Header pattern #2 to identify idle/unassigned cells.



H2CONF






## **24.51: H3CONF**

Header pattern #3 to identify idle/unassigned cells.



H3CONF





## **24.52: H4CONF**

Header pattern #4 to identify idle/unassigned cells.



H4CONF





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

## **24.53: H5CONF**

Dummy byte to align the Payload in the ACB\_Rx buffer.



H5CONF 7 6 5 4 3 2 1 0







## **Overhead Frame Processor Architecture: Transmit Direction**



# **OFP\_Tx GPP Handler Address Mapping** Base Address = x'400' (Page 1 of 3)

<span id="page-578-0"></span>1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation

2. Address range 100-17F located in 128x8 GRA. Address range 180-1BF located in 64x8 GRA.





#### **OFP\_Tx GPP Handler Address Mapping** Base Address = x'400' (Page 2 of 3)

1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation

2. Address range 100-17F located in 128x8 GRA. Address range 180-1BF located in 64x8 GRA.





#### **OFP\_Tx GPP Handler Address Mapping** Base Address = x'400' (Page 3 of 3)

1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation

2. Address range 100-17F located in 128x8 GRA. Address range 180-1BF located in 64x8 GRA.

**IBM Processor for Network Resources All and Service Preliminary Preliminary** 



## **OFP\_Tx Register Description**

#### **Counter Registers**

## **24.54: PTRINC**

Number of pointer increment events (eight-bit counter). Overflow leads to an interrupt request.







#### **24.55: PTRDEC**

Number of pointer decrement events (eight-bit counter). Overflow leads to an interrupt request.



PTRDEC







## **24.56: ND\_EVCNT**

Number of new data events (eight-bit counter). Overflow leads to an interrupt request.



ND\_EVCNT





## **24.57: JUSCNT**

Number of justification errors detected (eight-bit counter). Overflow leads to an interrupt request.



JUSCNT 7 6 5 4 3 2 1 0





## **24.58: JUSCNTTh11**

Threshold for number of justification errors. Threshold overstep leads to an interrupt request.



JUSCNTTh11





#### **24.59: CntEn1**

Counter On/Off control register for OFP\_Tx.

For each bit position:

- $0 =$  Counter is disabled.
- $1 =$  Counter is enabled.









#### **24.60: Reset Register (RESET)**

Reset/Halt chiplet control register. This register is automatically preset to the default value by the reset signal ResOT coming from GPPINT chiplet.

For each bit position:

- $0 =$  Reset/Halt not active.
- 1 = Reset/Halt active.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

## **24.61: Command Register (CMD1)**

Command register for the chiplet. Single-cycle active if b'1' is written into bit position.











## **Status Registers**

## **24.62: STAT1**

Status register #1 of the chiplet. This is an event latch register.







## **24.63: STAT2**

Status register #2 of the chiplet. This is an event latch register.







#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**



#### **Interrupt and Mask Registers**

#### **24.64: MainIRQ**

Register to indicate fatal interrupt events and to point to user IRQ registers with active requests.

For each bit position:

- $0 = No$  interrupt request pending.
- 1 = Interrupt request pending.









## **24.65: M\_MainIRQ**

Register to mask pending interrupt requests. A masked request will not generate an outgoing IRQ to the GPPINT.

For each bit position:

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates signal IRQOT to GPPINT.







#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**

## **24.66: CntrIRQ1**

Register to indicate active counter interrupt requests of this chiplet.

For each bit position:

0 = No interrupt request pending.

 $1 =$  Interrupt request pending.











## **24.67: M\_CntrIRQ1**

Register to mask pending counter interrupt requests.

For each bit position:

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

#### **24.68: IRQ3**

Register to indicate active user interrupt requests of this chiplet.

For each bit position:

 $0 = No$  interrupt request pending.

1 = Interrupt request pending.



**Power On Value** 









#### **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**

## **24.69: M\_IRQ3**

Register to mask pending user interrupt requests.

For each bit position:

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**

## **Configuration Registers**

## **24.70: CONF1**

Configuration register #1. General OFP\_Tx configuration signals A.











## **24.71: CONF2**

Configuration register #2. General OFP\_Tx configuration signals B.







## **24.72: CONF3**

Configuration register #3.





#### IBM3206K0424

#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**



## **24.73: CONF4**

Configuration register #4.







## **24.74: CONF5**

Configuration register #5.









## **24.75: CONF6**

Configuration register #6. Frame scrambling control register.







#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**

## **24.76: CONF7**

Configuration register #7. DCC control register.







## **24.77: CONF8**

Configuration registers #8. Low water FIFO threshold register.











## **24.78: CONF9**

Configuration registers #9. Normal water FIFO threshold register.







### **24.79: CONF10**

Configuration registers #10. High water FIFO threshold register.









## **Overhead Frame Processor Architecture: Receive Direction**

#### **OFP\_Rx GPP Handler Address Mapping** Base Address = x'800' (Page 1 of 4)



<span id="page-599-1"></span>1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation.

<span id="page-599-0"></span>2. Address range 100-17F located in 128x8. GRA Address range 180-1BF located in 64x8 GRA.





#### **OFP\_Rx GPP Handler Address Mapping** Base Address = x'800' (Page 2 of 4)

1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation.

2. Address range 100-17F located in 128x8. GRA Address range 180-1BF located in 64x8 GRA.





#### **OFP\_Rx GPP Handler Address Mapping** Base Address = x'800' (Page 3 of 4)

1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation.

2. Address range 100-17F located in 128x8. GRA Address range 180-1BF located in 64x8 GRA.





#### **OFP\_Rx GPP Handler Address Mapping** Base Address = x'800' (Page 4 of 4)

1. Independent of the counter width, given that a counter has chiplet address N as a base. Reading address N or address N-1 both yield the least significant byte of the counter. Reading address N has no affect on the counter, but reading address N-1 resets the counter after read operation.

2. Address range 100-17F located in 128x8. GRA Address range 180-1BF located in 64x8 GRA.

#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**



## **Counter Registers**

## **24.80: ROFmid**

Read-on-the-fly registers.





## **24.81: B1BITCNT**

Number of BIP-8 B1 bit errors counted since last counter reset (16-bit counter). Overflow leads to an interrupt request.



B1BITCNT (8:15)







## **24.82: B1BITCNTTh11**

Threshold for number of BIP-8 B1 bit errors.



#### B1BITCNTTh11





#### **24.83: B1BITCNTTh12**

Threshold for number of BIP-8 B1 bit errors. Threshold overstep leads to an interrupt request.



B1BITCNTTh12





#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**



## **24.84: B1BLKCNT**

Number of BIP-8 B1 block errors counted since last counter reset (16-bit counter). Overflow leads to an interrupt request.



B1BLKCNT (8:15)



#### **24.85: B1BLKCNTTh11**

Threshold for number of BIP-8 B1 block errors.



B1BLKCNTTh11







## **24.86: B1BLKCNTTh12**

Threshold for number of BIP-8 B1 block errors. Threshold overstep leads to an interrupt request.



B1BLKCNTTh12 7 6 5 4 3 2 1 0



## **24.87: B2BITCNT**

Number of BIP-24 B2 bit errors counted since last counter reset (16-bit counter). Overflow leads to an interrupt request.



B2BITCNT (8:15)







#### **24.88: B2BITCNTTh11**

Degradation threshold for number of BIP-24 B2 bit errors.







#### **24.89: B2BITCNTTh12**

Degradation threshold for number of BIP-24 B2 bit errors. Threshold overstep leads to an interrupt request.



B2BITCNTTh12







## **24.90: B2BITCNTTh21**

Failure threshold for number of BIP-24 B2 bit errors.



B2BITCNTTh21





#### **24.91: B2BITCNTTh22**

Failure threshold for number of BIP-24 B2 bit errors. Threshold overstep leads to an interrupt request.



B2BITCNTTh22





#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



#### **24.92: B2BLKCNT**

Number of BIP-24 B2 block errors counted since last counter reset (16-bit counter). Overflow leads to an interrupt request.



B2BLKCNT





## **24.93: B2BLKCNTTh11**

Degradation threshold for number of BIP-24 B2 block errors.



B2BLKCNTTh11

$$
\downarrow
$$
\n
$$
\uparrow
$$





## **24.94: B2BLKCNTTh12**

Degradation threshold for number of BIP-24 B2 block errors. Threshold overstep leads to an interrupt request.



B2BLKCNTTh12





## **24.95: B2BLKCNTTh21**

Failure threshold for number of BIP-24 B2 block errors.



B2BLKCNTTh21

$$
\begin{array}{c|cccc}\n\downarrow & & & & \\
\hline\n7 & 6 & 5 & 4 & 3 & 2 & 1 & 0\n\end{array}
$$





#### **24.96: B2BLKCNTTh22**

Failure threshold for number of BIP-24 B2 block errors. Threshold overstep leads to an interrupt request.







#### **24.97: B3BITCNT**

Number of BIP-8 B3 bit errors counted since last counter reset (16-bit counter). Overflow leads to an interrupt request.



B3BITCNT






**Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary P** 

### **24.98: B3BITCNTTh11**

Threshold for number of BIP-8 B3 bit errors.



### B3BITCNTTh11





#### **24.99: B3BITCNTTh12**

Threshold for number of BIP-8 B3 bit errors. Threshold overstep leads to an interrupt request.



B3BITCNTTh12





#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**



#### **24.100: B3BLKCNT**

Number of BIP-8 B3 block errors counted since last counter reset (16-bit counter). Overflow leads to an interrupt request.



B3BLKCNT



# **24.101: B3BLKCNTTh11**

Threshold for number of BIP-8 B3 block errors.



B3BLKCNTTh11







### **24.102: B3BLKCNTTh12**

Threshold for number of BIP-8 B3 block errors. Threshold overstep leads to an interrupt request.



B3BLKCNTTh12





#### **24.103: MSREICNT**

Multiplex Section Remote Error Indication counter (16-bit counter). Overflow leads to an interrupt request.



MSREICNT (8:15)







# **24.104: MSREICNTTh11**

Threshold for number of Multiplex Section Remote Errors.







### **24.105: MSREICNTTh12**

Threshold for number of Multiplex Section Remote Errors. Threshold overstep leads to an interrupt request.



MSREICNTTh12



Page 616 of [676](#page-675-0)





## **24.106: HPREICNT**

Higher-order Path Remote Error Indication counter (16-bit counter). Overflow leads to an interrupt request.



HPREICNT (8:15)





### **24.107: HPREICNTTh11**

Threshold for number of Higher-order Path Remote Errors.



HPREICNTTh11







#### **24.108: HPREICNTTh12**

Threshold for number of Higher-order Path Remote Errors. Threshold overstep leads to an interrupt request.



HPREICNTTh12 7 6 5 4 3 2 1 0



#### **24.109: PJ\_EVCNT**

Positive Justification Event counter (eight-bit counter). Overflow leads to an interrupt request.



PJ\_EVCNT 7 6 5 4 3 2 1 0





# **24.110: NJ\_EVCNT**

Negative Justification Event counter (eight-bit counter). Overflow leads to an interrupt request.







# **24.111: ND\_EVCNT**

New Data Event counter (eight-bit counter). Overflow leads to an interrupt request.



ND\_EVCNT



# **24.112: CntEn1**

Counter On/Off control register #1 for OFP\_Rx.

- $0 =$  Counter is disabled.
- $1 =$  Counter is enabled.











#### **24.113: CntEn2**

Counter On/Off control register #2 for OFP\_Rx.

- $0 =$  Counter is disabled.
- $1 =$  Counter is enabled.









#### **24.114: Reset Register (RESET)**

Reset/Halt chiplet control register. This register is automatically preset to the default value by the reset signal ResOT coming from GPPINT chiplet.

For each bit position:

- $0 =$  Reset/Halt not active.
- 1 = Reset/Halt active.







### **Status Registers**

#### **24.115: STAT1**

Status register #1 of the chiplet. OFP\_Rx Mode status information. This is an event latch register.









# **24.116: STAT2**

Status register #2 of the chiplet. AU pointer status information of OFP\_Rx. This is an event latch register.



**Power On Value** 





#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**



#### **24.117: STAT3**

Status register #3 of the chiplet. Section Overhead (SOH) status of OFP\_Rx. This is an event latch register.



**Address** 835

**Power On Value** -







# **24.118: STAT4**

Status register #4 of the chiplet. Path Overhead (POH) status of OFP\_Rx. This is an event latch register.



**Power On Value** 







#### **Interrupt and Mask Registers**

#### **24.119: MainIRQ**

Register to indicate fatal interrupt events and to point to user IRQ registers with active requests.

- $0 = No$  interrupt request pending.
- $1 =$  Interrupt request pending.









**Preliminary IBM Processor for Network Resources**

#### **24.120: M\_MainIRQ**

Register to mask pending interrupt requests. A masked request will not generate an outgoing IRQ to the GPPINT.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates signal IRQOR to GPPINT.







# **24.121: CntrIRQ1**

Register #1 to indicate active counter interrupt requests of this chiplet.

For each bit position:

 $0 = No$  interrupt request pending.

1 = Interrupt request pending.











**Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre** 

#### **24.122: M\_CntrIRQ1**

Register to mask pending counter interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







#### **24.123: CntrIRQ2**

Register #2 to indicate active counter interrupt requests of this chiplet.

For each bit position:

 $0 = No$  interrupt request pending.

1 = Interrupt request pending.













**Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre** 

Register to mask pending counter interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







#### **IBM Processor for Network Resources All and Service Preliminary Preliminary**

### **24.125: CntrIRQ3**

Register #3 to indicate active counter interrupt requests of this chiplet.

For each bit position:

0 = No interrupt request pending.

 $1 =$  Interrupt request pending.











**Preliminary IBM Processor for Network Resources**

#### **24.126: M\_CntrIRQ3**

Register to mask pending counter interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.









## **24.127: IRQ6**

Register to indicate active user interrupt requests of this chiplet.

For each bit position:

- $0 = No$  interrupt request pending.
- 1 = Interrupt request pending.



**Power On Value** 







#### **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**

### **24.128: M\_IRQ6**

Register to mask pending user interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







#### **24.129: IRQ7**

Register to indicate active user interrupt requests of this chiplet.

For each bit position:

 $0 = No$  interrupt request pending.

 $1 =$  Interrupt request pending.



 $\begin{array}{r}\n\leftarrow & 00F \\
\leftarrow & 109 \\
\leftarrow & 0F \\
\leftarrow & \leftarrow &$  $\overline{\mathsf{L}}$  $\downarrow$  $\downarrow$  $\downarrow$  $\downarrow$  $\downarrow$ 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0







#### **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Preliminary Preliminary Community Preliminary Preliminary Preliminary Preliminary Preliminary Preliminary Pre**

#### **24.130: M\_IRQ7**

Register to mask pending user interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







### **24.131: IRQ8**

Register to indicate active user interrupt requests of this chiplet.

For each bit position:

 $0 = No$  interrupt request pending.

1 = Interrupt request pending.



**Power On Value** 







**Preliminary IBM Processor for Network Resources**

#### **24.132: M\_IRQ8**

Register to mask pending user interrupt requests.

- 0 = The corresponding pending request bit is masked (DEFAULT).
- 1 = The corresponding pending request bit activates the pointer bit in MainIRQ register.







# **Configuration Registers**

### **24.133: CONF1**

Configuration register #1. General OFP\_Rx configuration signals.











# **24.134: CONF2**

Configuration register #2. SOH processing configuration signals.









# **24.135: CONF3**

Configuration register #3. POH byte processing configuration signals.









# **24.136: CONF4**

Configuration register #4. APS processing configuration signals.









# **24.137: CONF7**

Configuration register #7. Miscellaneous OFP\_Rx configuration signals.









**Preliminary IBM Processor for Network Resources**

### **24.138: CONF8**

Configuration register #8. Pattern register signals.



FSCRrx 7 6 5 4 3 2 1 0



#### **24.139: CONF9**

Configuration register #9. Pattern register signals.



Slexpct 7 6 5 4 3 2 1 0



**IBM Processor for Network Resources Preliminary** 





# **Memory Map for Registers and Arrays**





# **Signal Pin Listing By Signal Name** (Page 1 of 5)




### **Signal Pin Listing By Signal Name** (Page 2 of 5)





### **Signal Pin Listing By Signal Name** (Page 3 of 5)





### **Signal Pin Listing By Signal Name** (Page 4 of 5)





### **Signal Pin Listing By Signal Name** (Page 5 of 5)



PMDATA(8) AE11 C PMDATA(9) AC10 C



#### **AC Timing Characteristics**

#### **PHY Timing**



#### **IBM Processor for Network Resources And American Control of American Control of American Preliminary**



#### **NPBUS Timing**



#### **I/O PCI Bus Timing** (Page 1 of 2)





#### IBM3206K0424

#### **Preliminary Community Preliminary Community Preliminary Community Preliminary Preliminary Preliminary Community Processor for Network Resources**

#### **I/O PCI Bus Timing** (Page 2 of 2)



IBM3206K0424 IBM3206K0424

**IBM Processor for Network Resources IBM Processor for Network Resources**

# **Synchronous DRAM Timing Diagrams Synchronous DRAM Timing Diagrams**

SDRAM Read Cycle (1 of 4) **SDRAM Read Cycle (1 of 4)** 

**CAS Latency=2, Burst Length=1** T8 T9 T1 T2 T3 T4 T5 T6 T7 CMCLK (4:0) CMSYNRAS  $(1:0)$ CM0CS (1:0)  $\overline{1}$ CMSYNCAS CMWE (1:0)  $\overline{\phantom{a}}$ CM0DQM (3:0) CMADDR XXXXXXXX XXXXXX Col XXXXXXXXXXRow CM0CS Bank Address(3:2) CMDATA ZZZZZZZZZZZZZ ZZZZZZZZZZ Data Out  $\overline{\phantom{a}}$  $\top$ 

Synchronous DRAM Timing Diagrams<br>Page 656 of 676 Synchronous DRAM Timing Diagrams Page 656 of [676](#page-675-1)

pnr25.chapt07.01<br>August 14, 2000 pnr25.chapt07.01 August 14, 2000

 $\|m\|$ 

 **Preliminary**

Preliminary



**Preliminary**

Preliminary

 **IBM Processor for Network Resources IBM Processor for Network Resources** 

IBM3206K0424

IBM3206K0424

# SDRAM Read Cycle (2 of 4) **SDRAM Read Cycle (2 of 4)**



IBM3206K0424 IBM3206K0424

## $\|$ ini $\|$

 **Preliminary**

Preliminary

**IBM Processor for Network Resources**

**IBM Processor for Network Resources** 



Synchronous DRAM Timing Diagrams<br>Page 658 of 676 Synchronous DRAM Timing Diagrams Page 658 of [676](#page-675-1)

pnr25.chapt07.01<br>August 14, 2000 pnr25.chapt07.01 August 14, 2000

 $\left\Vert \mathbf{m}\right\Vert$ 

**Preliminary**

Preliminary

IBM3206K0424

 **IBM Processor for Network Resources**

**IBM Processor for Network Resources** 

IBM3206K0424



IBM3206K0424 IBM3206K0424

 $\|$ ini $\|$ 

 **Preliminary**

Preliminary

# SDRAM Write Cycle (1 of 4) **SDRAM Write Cycle (1 of 4)**

**IBM Processor for Network Resources**

**IBM Processor for Network Resources** 



Synchronous DRAM Timing Diagrams<br>Page 660 of 676 Synchronous DRAM Timing Diagrams Page 660 of [676](#page-675-1)

pnr25.chapt07.01<br>August 14, 2000 pnr25.chapt07.01 August 14, 2000



**Preliminary**

Preliminary

 **IBM Processor for Network Resources IBM Processor for Network Resources** 

IBM3206K0424

IBM3206K0424

# SDRAM Write Cycle (2 of 4) **SDRAM Write Cycle (2 of 4)**



## $\|m\|$

 **Preliminary**

Preliminary

**IBM Processor for Network Resources**

**IBM Processor for Network Resources** 



Synchronous DRAM Timing Diagrams<br>Page 662 of 676 Synchronous DRAM Timing Diagrams Page 662 of [676](#page-675-1)

pnr25.chapt07.01<br>August 14, 2000 pnr25.chapt07.01 August 14, 2000

 $\|m\|$ 

**Preliminary**

Preliminary

 **IBM Processor for Network Resources IBM Processor for Network Resources** IBM3206K0424

IBM3206K0424





Synchronous DRAM Timing Diagrams<br>Page 664 of 676 Synchronous DRAM Timing Diagrams

Page 664 of [676](#page-675-1)

 **Preliminary** Preliminary  $\|$ ini $\|$ 

# SDRAM Write of 64-byte Burst with CAS Latency=2 **SDRAM Write of 64-byte Burst with CAS Latency=2**

**IBM Processor for Network Resources**

**IBM Processor for Network Resources** 



pnr25.chapt07.01<br>August 14, 2000 pnr25.chapt07.01 August 14, 2000

IBM3206K0424 IBM3206K0424

pnr25.chapt07.01<br>August 14, 2000 August 14, 2000 pnr25.chapt07.01

XX

 $\pm$ 

 $\mathbf{I}$ 

 $\overline{1}$ 

 $\mathbf{I}$ 

 $\mathbf{I}$ 

CM0DQM

(3:0)

CMADDR

CM0CS

CMDATA

(3:2)

CMWE

(1:0)

CMSYNCAS

CMCLK

(4:0)

(1:0)

CM0CS

(1:0)

CMSYNRAS



Data  $14 \times$  Data 15

 $\overline{1}$ 

 $\overline{\phantom{a}}$ 

 $\mathbf{I}$ 

 $\overline{1}$ 

Data 0ZZZZZZZZ Data 1 ZZZZZZZZ

 $\mathbf{I}$ 

 $\mathbf{I}$ 

 IBM3206K0424 IBM3206K0424

 **IBM Processor for Network Resources**

**IBM Processor for Network Resources** 



**Preliminary**

Preliminary

IBM3206K0424 IBM3206K0424

# 

 **Preliminary**

Preliminary

# **IBM Processor for Network Resources IBM Processor for Network Resources**

# **SRAM Timing Diagrams SRAM Timing Diagrams**

## **SRAM Read Cycle SRAM Read Cycle**



SRAM Timing Diagrams<br>Page 666 of 676 SRAM Timing Diagrams Page 666 of [676](#page-675-1)



# **IBM Processor for Network Resources** IBM Processor for Network Resources IBM3206K0424

IBM3206K0424

## Preliminary **Preliminary**

## SRAM Write Cycle **SRAM Write Cycle**



# 

 **Preliminary**

Preliminary

# **IBM Processor for Network Resources IBM Processor for Network Resources**

# **SRAM Read Cycle with Byte Enables SRAM Read Cycle with Byte Enables**





 **IBM Processor for Network Resources**

**IBM Processor for Network Resources** 

## Preliminary **Preliminary**

# **SRAM Write Cycle with Byte Enables SRAM Write Cycle with Byte Enables**



EPROM Timing Diagrams<br>Page 670 of 676 Page 670 of [676](#page-675-1) EPROM Timing Diagrams

PINTCLK

:H or L



T1 T3 T5 T7 T9

┍

IBM3206K0424 IBM3206K0424

IBM Processor for Network Resources **IBM Processor for Network Resources**



**Parallel EPROM Read Parallel EPROM Read** 

T19

T11 T13 T15 T17

П

' 1 '

Hi-Z

 **Preliminary** Preliminary

pnr25.chapt07.01<br>August 14, 2000 pnr25.chapt07.01 August 14, 2000



# IBM Processor for Network Resources IBM3206K0424

 **IBM Processor for Network Resources**

IBM3206K0424

### Preliminary **Preliminary**

# Parallel EPROM Write **Parallel EPROM Write**



# **IBM Processor for Network Resources IBM Processor for Network Resources**

### **Preliminary** Preliminary

# Serial EPROM Read **Serial EPROM Read**



EPROM Timing Diagrams<br>Page 672 of 676 Page 672 of [676](#page-675-1) EPROM Timing Diagrams



IBM3206K0424

# Serial EPROM Write **Serial EPROM Write**



IBM3206K0424 IBM3206K0424

**IBM Processor for Network Resources IBM Processor for Network Resources**

## **Preliminary** Preliminary

# **PHY Timing Diagrams PHY Timing Diagrams**

### **PHY Read PHY Read**



PHY Timing Diagrams<br>Page 674 of 676 Page 674 of [676](#page-675-1) PHY Timing Diagrams



# IBM Processor for Network Resources IBM3206K0424 IBM3206K0424

 **IBM Processor for Network Resources**

### Preliminary **Preliminary**

## **PHY Write PHY Write**





#### <span id="page-675-1"></span><span id="page-675-0"></span>**Revision Log**

