PM7351, PM7350

S/UNI-VORTEX & S/UNI-DUPLEX

S/UNI-VORTEX™ S/UNI-DUPLEX™

TECHNICAL OVERVIEW

PRELIMINARY

ISSUE 2: JUNE 1999

REVISION HISTORY

Issue No.	Issue Date	Details of Change
2	June 1999	Changed the confidentiality notices for the document's public release.
1	February 1999	Document created.

CONTENTS

FIGURES

1 PURPOSE AND SCOPE OF THIS DOCUMENT

The S/UNI-VORTEX and S/UNI-DUPLEX are highly integrated analog/digital devices used to create high performance, fault tolerant (1:1 redundant) “serial backplane” connectivity between printed circuit boards in communications equipment such as DSLAMs, access multiplexers. and switches. Interconnection of cards is achieved via 4-wire serial interconnect over backplane traces and, in multi-shelf architectures, via serial cabling between shelves. The S/UNI-VORTEX and S/UNI-DUPLEX can also be used to interconnect stackable “pizza box” style equipment via low cost serial cabling.

Working together, the S/UNI-VORTEX and S/UNI-DUPLEX will carry traffic between several thousand PHY or modem devices spread across dozens of line cards to a single ATM switching device such as the S/UNI-APEX. The S/UNI-VORTEX and S/UNI-DUPLEX implement flow-control (i.e. back-pressure) across their serial links without relying on intermediate stages of external buffering. This eliminates all extraneous data path devices (cell buffers, traffic management devices, etc.), leaving only the modems/PHYs and a single, tiny 160 pin BGA (the S/UNI-DUPLEX) on each line card. Fully integrated support for hot swap, 1:1 protection switching, an embedded inter-card communications channel, and system timing distribution complete the architecture. **This results in the lowest cost and most functionally complete high fan-in intershelf/intrashelf interconnect architecture available on the market today**.

This Technical Overview describes how the S/UNI-VORTEX and S/UNI-DUPLEX satisfy the system level requirements of a wide range of communication equipment. The document answers the following questions:

- What are the primary applications and equipment that the S/UNI-VORTEX and S/UNI-DUPLEX are designed for?
- What function do the S/UNI-VORTEX and S/UNI-DUPLEX fulfill in the overall system architecture?
- What are the advantages of using these devices? Why?
- What PMC-Sierra products and other components interwork with the S/UNI-VORTEX and S/UNI-DUPLEX?

This document is a companion of, but subordinate to, the S/UNI-VORTEX Data Sheet and the S/UNI-DUPLEX Data Sheet. If there appear to be differences, contradictions, or omissions in this Technical Overview the reader is advised that the Data Sheets take precedence.

2 APPLICATION OVERVIEW

When designing communication equipment such as access switches and multiplexers the equipment architect is faced with a common problem: how do I efficiently connect a large number of lower speed ports to a small number of high speed ports? Typically, a number of line-side ports (modems or ATM PHYs) are terminated on each line card. Numerous line cards are then slotted into one or more shelves and backplane traces or inter-shelf cables are used to connect the line cards to a centralized (often 1:1 protected) common card, hereafter referred to as the core card. The core card normally includes one or more high speed WAN up-link ports that transport traffic to and from a high speed broadband network. Optionally, the WAN up-link ports can be housed on a separate card from the core card, hereafter called the WAN up-link card.

A block diagram of a 1:1 redundant system is shown in Figure 1.

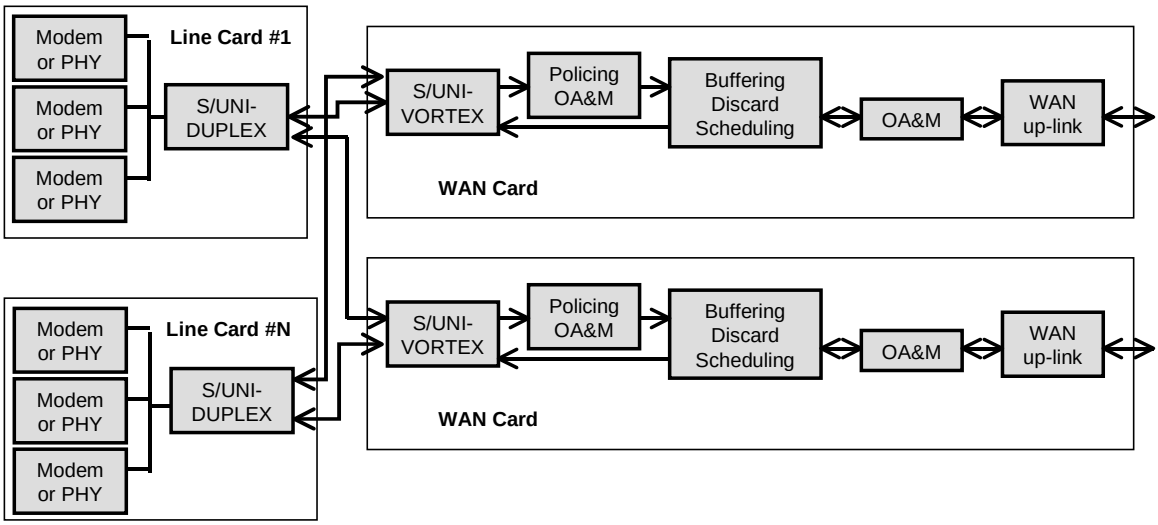

Figure 1 - Typical Target Application

In this type of equipment the majority (perhaps all) user traffic goes from WAN port to line port, or from line port to WAN port. Although the individual ports on the line cards are often relatively low speed interfaces such as T1, E1, or xDSL, there may be many ports per line card and many line cards per system, resulting in hundreds or even thousands of lines terminating on a single WAN up-link. In the upstream direction (from line card to WAN up-link), the equipment must have capacity to buffer and intelligently manage bursts of upstream traffic simultaneously from numerous line cards.

In the downstream direction the equipment must handle a similar issue, the “big pipe feeding little pipe” problem. When a large burst of traffic destined for a single line port is received at the high speed WAN port it must be buffered and managed as it queues up waiting for the much lower speed line port to clear.

The line cards are always the most numerous cards in this type of equipment. An individual line card, even if it terminates a few dozen low speed ports, does not generate or receive enough traffic to justify putting complex buffering and traffic management devices on it. The ideal architecture has low cost "dumb" line cards and a feature rich, "smart" core card. In order to enhance fault tolerance, the architecture should also inherently support 1:1 protection using a redundant core card and WAN up-link without significantly increasing line card complexity. In this type of architecture there are often three stages of signal concentration or multiplexing, as shown in Figure 2.

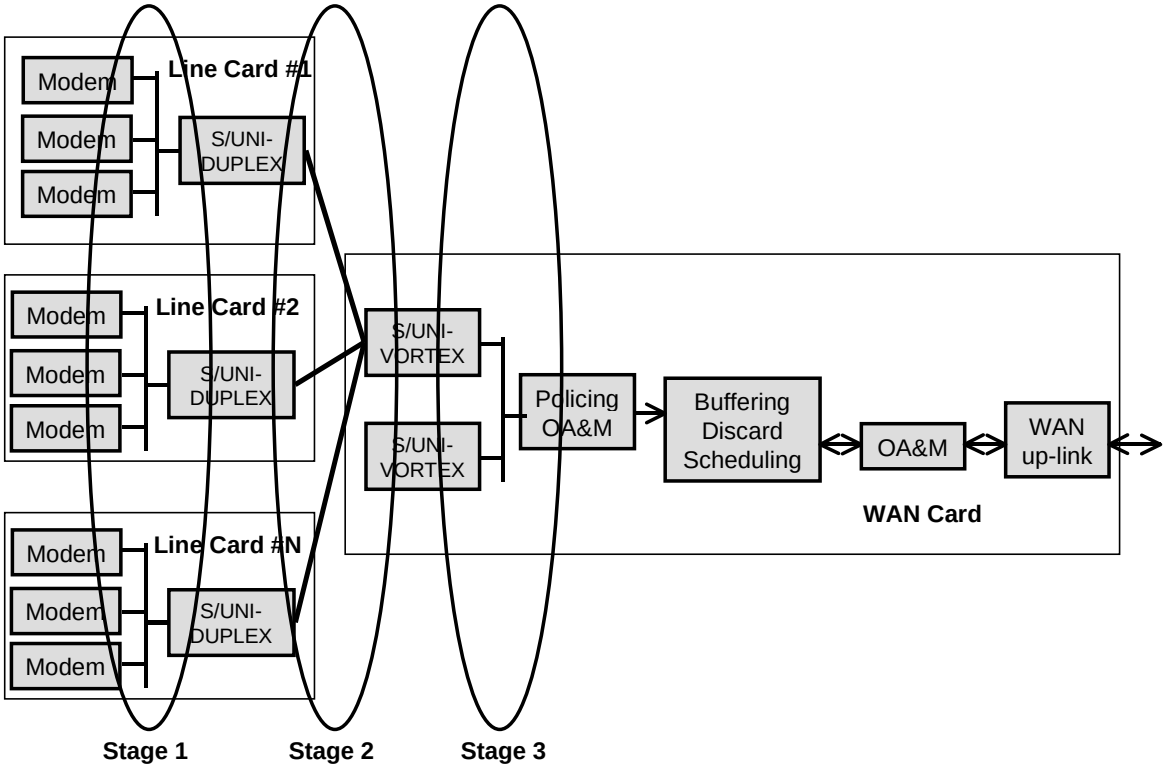

Figure 2 - Three stage multiplex architecture.

The first stage resides on the line card and spans only those ports physically terminated by that card. Since it is confined to a single card, this first stage of multiplexing readily lends itself to a simple parallel bus based multiplex topology. The second stage of concentration occurs between the core card(s) and the line cards, including line cards that are on a separate shelf. As will be shown, this second stage is best served by a redundant serial point-to-point topology. The third stage of multiplexing is optional and resides on the core card. This third stage is used in systems with a large number of line cards that require several devices to terminate the second stage of aggregation. Since the third stage of aggregation is confined to the core card, it lends itself readily to a parallel bus implementation. This three stage approach is implemented directly by the S/UNI-DUPLEX and S/UNI-VORTEX architecture.

The S/UNI-DUPLEX acts as the line card's bus master. It implements the first stage of multiplexing by routing traffic from the PHYs and transmitting the traffic simultaneously over two high speed (200 Mbit/s) serial 4-wire LVDS links. One serial link attaches to the active core card, the other to the standby core card.[1] In the downstream direction the S/UNI-DUPLEX demultiplexes traffic from the active core card's LVDS serial link and routes this traffic to the appropriate PHYs. If the active core card (or its LVDS link) should fail, protection switching commands embedded in the spare LVDS link will direct the S/UNI-DUPLEX to start receiving its traffic from this spare link.

The S/UNI-VORTEX resides on the core card and terminates up to 8 LVDS links connected to 8 S/UNI-DUPLEX devices. The S/UNI-VORTEX implements the second stage of multiplexing. More than one S/UNI-VORTEX will be required if more than 8 links are required – as will be the case for a system with more than 8 line cards. The S/UNI-VORTEX device(s) share a high speed parallel bus with the core card's traffic management and OA&M layers, as implemented by devices such as PMC-Sierra's S/UNI-APEX and the S/UNI-ATLAS. This is the third stage of multiplexing.

When evaluating system interconnect architectures, one must consider the full set of system level requirements:

- Transporting user data: i.e. implementing the data-path.
- Inter-processor communication: i.e. an inter-card control channel.
- Exchanging flow control information or backpressure for the user data as well as the control channel.
- Implementing discrete command and control signals: system reset, error indications, protection switching commands, etc.
- Clock/timing distribution (system clocks as well as reference clocks such as 8 kHz timing references).
- Fault tolerance, protection switching, redundancy, and inserting/removing line card while the system is running (hot swap).

The remainder of this Technical Overview discusses these requirements and describes how each is satisfied by the S/UNI-VORTEX and S/UNI-DUPLEX architecture.

[1] A single core card implementation is also supported, of course.

3 THE PARALLEL BUS SPECIFICATIONS

Throughout this document references are made to three related bus interface specifications.

UTOPIA Level 2 – The industry standard ATM Forum PHY to ATM layer bus specification. Also referred to as UTOPIA L2. It is available from the ATM Forum web site at www.atmforum.com.

SCI-PHY Level 2 – Also referred to as SCI-PHY L2. PMC-Sierra's superset extension to the UTOPIA bus specification to allow the following enhancements:

- 32 PHY maximum instead of 31.
- Extended cells: SCI-PHY supports user defined bytes prepended or postpended to the standard 53 byte cells allowed by UTOPIA. This particularly useful when switching tags are to be attached to the cell by an ATM layer device such as the S/UNI-ATLAS. SCI-PHY allows cells to be up to 64 bytes long.

ANY-PHY – PMC-Sierra's extension to the UTOPIA bus specification to allow the following enhancements:

- Addressing for an unlimited number of logical PHYs.
- Fixed cell or arbitrary length packet transfers. Only cell mode transfers are used by the S/UNI-VORTEX and S/UNI-DUPLEX.
- Extended cells: This includes optional prepend or postpend cell extensions plus the additional overhead bytes used for in-band PHY addressing and selection.
- In the transmit direction, separation of PHY status polling (using bus pins) from PHY selection during start of cell or packet transfer (using in-band addressing).
- In the receive direction, cell overhead bytes can be used to identify the source PHY of the cell.
- Relaxed decode response time to simplify device status polling circuitry.

A comparison of the bus signals for these three bus types is shown in Figure 3.

UTOPIA Level 2 Bus Slave	SCI-PHY Level 2 Bus Slave	Any-PHY Bus Slave
TxClk	TFCLK	TCLK
TxEnb*	TWRENB	TENB
TxAddr[4:0]	TADDR[4:0]	TADR[4:0]
n/a	TAVALID	TADR[5]
TxData[15:0]	TDAT[15:0]	TDAT[15:0]
TxPrty	TPRTY	TPRTY
TxClav	TCA	TPA
TxSOC	TSOC	TSOP
n/a	n/a	TSX

UTOPIA Level 2 Bus Slave	SCI-PHY Level 2 Bus Slave	Any-PHY Bus Slave
RxClk	RFCLK	RCLK
RxEnb*	RWRENB	RENB
RxAddr[4:0]	RADDR[4:0]	RADR[4:0]
n/a	RAVALID	RADR[5]
RxData[15:0]	RDAT[15:0]	RDAT[15:0]
RxPrty	RPRTY	RPRTY
RxClav	RCA	RPA
RxSOC	RSOC	RSOP
n/a	n/a	RSX

Figure 3 - Comparison of Bus Signals

4 DATA PATH AND FLOW CONTROL

This section describes how the S/UNI-VORTEX and S/UNI-DUPLEX implement a flow controlled data path between the line cards and the core card(s). The section is organized around the three stage multiplex architecture shown in Figure 2.

4.1 Multiplexing on the Line Card (Stage 1 Multiplexing)

As long as the PHY devices are UTOPIA L2 compliant, there will be no glue logic or external circuitry needed to interface the S/UNI-DUPLEX to up to 31 logical PHY devices. Using the SCI-PHY UTOPIA extension defined by PMC-Sierra, up to 32 PHY devices can be served. Because the S/UNI-DUPLEX is a mixed analog/digital device with integrated clock synthesis unit (CSU) and clock recovery unit (CRU) there is no external circuitry needed to support the high speed serial interface between the S/UNI-DUPLEX and S/UNI-VORTEX. In this way, parts count and board area on the line card is kept to an absolute minimum.

4.1.1 Upstream (to the Core Card) Traffic

A single parallel data bus (e.g. UTOPIA L2) is used to connect all PHY or multi-PHY devices to the S/UNI-DUPLEX. The PHYs are slave devices, the S/UNI-DUPLEX is the UTOPIA bus master. The bus is normally 8 bits wide with a maximum clock rate of 25 MHz (i.e. maximum 200 Mbit/s bus bandwidth), although a 16 bit 33 MHz bus is also supported.

As bus master, the S/UNI-DUPLEX continuously polls the Receive Cell Available (RCA) status lines of each PHY, reading cells from the PHYs as they become available. Once a cell is brought into the S/UNI-DUPLEX the cell is tagged (via a prepend byte) with the appropriate PHY ID (0:31) and sent simultaneously to both core cards (active and standby) over the S/UNI-DUPLEX's two high speed LVDS links.

For the applications being targeted by this architecture, the 200 Mbit/s bandwidth capacity of the UTOPIA bus and the LVDS serial link is greater than the aggregate maximum upstream bandwidth of the PHYs.[1] Hence, there is no need for extensive buffering in the S/UNI-DUPLEX.[2] The internal four cell FIFO required on all UTOPIA compliant PHYs ensures that each modem can buffer cells while it waits for the S/UNI-DUPLEX to service it.

[1] For example, a line card with 32 E1 ports would, worse case, cause an upstream burst of only 64 Mbit/s. For DSLAM applications based on ADSL modems, the worse case upstream bandwidth requirement is even less.

[2] The S/UNI-DUPLEX actually does have minimal internal cell buffer in the upstream direction. This buffering improves throughput by ensuring that cell transfers from the UTOPIA bus to the LVDS link can occur back-to-back without idle periods on the link.

Since the LVDS link to the core card is not shared by any other line cards there is no need to conserve bandwidth by implementing port-to-port switching on the line card. Leaving all switching functions on the core card greatly simplifies the hardware and software requirements of the line card. This is one of the fundamental advantages of using a point-to-point interface between the line card and core card rather than, for example, a shared bus architecture where multiple line cards share a high speed bus.

In summary, the upstream data path on the line card is:

- Upstream cells are received by the PHY device and buffered until the S/UNI-DUPLEX reads them.
- The connection between the PHYs and the S/UNI-DUPLEX is standard UTOPIA L2.
- PHYs are bus slaves, the S/UNI-DUPLEX is bus master.
- After a cell is transferred from the PHY to the S/UNI-DUPLEX, it is tagged with the PHY ID and sent simultaneously to the active and standby core card via the two point-to-point high speed serial links.

4.1.2 Downstream (to the Line Card) Traffic

Downstream traffic is sent to the line card from the core card over the LVDS link. The S/UNI-DUPLEX will only accept cells from the LVDS link that has identified itself as "active" via an embedded control signal[1]. This allows the core cards to control when protection switching between LVDS links (and hence core cards) occurs.[2]

All downstream cells will have been tagged on the core card (by the ATM layer device) with a prepend (0:31) that identifies which PHY the cell belongs to. For each cell it receives, the S/UNI-DUPLEX strips off the prepend, leaving just the original cell. It temporarily buffers the cell in a shallow (four cells per PHY) internal FIFO[3].

As bus master, the S/UNI-DUPLEX continuously polls the Transmit Cell Available (TCA) status lines of all PHYs for which it has a downstream cell buffered. When a PHY has room (i.e. TCA is asserted), the S/UNI-DUPLEX sends the next buffered cell to the PHY over the UTOPIA bus.

To prevent its internal buffers from overflowing the S/UNI-DUPLEX implements per-PHY back-pressure. Back-pressure signaling is sent to the S/UNI-VORTEX via embedded overhead on the upstream LVDS link. The S/UNI-DUPLEX indicates back-pressure for a PHY if the corresponding internal FIFO is holding two or more cells. Because buffering

[1] Users cells on the inactive link are monitored for errors and then discarded. Any embedded control channel cells received on the inactive link are sent to the microprocessor port. This is discussed further in Section 5.

[2] If a distributed control architecture is desired, the line card's microprocessor can override the active channel selection via a control register in the S/UNI-DUPLEX.

[3] This creates full "cell rate decoupling" between the UTOPIA bus and the LVDS. Therefore the bus and the link can be clocked asynchronously.

and back-pressure are implemented independently on a per-PHY basis, no head of line blocking occurs.

In summary, the downstream data path on the line card is:

- The S/UNI-DUPLEX receives downstream cells on the active LVDS link, strips off the PHY ID and places the cells in an internal per-PHY buffer.
- The S/UNI-DUPLEX sends per-PHY back-pressure to the S/UNI-VORTEX via the LVDS link whenever the corresponding internal buffer begins to fill.
- For each PHY with a non-empty buffer, the S/UNI-DUPLEX continuously polls the PHY's TCA line and sends a cell over the UTOPIA bus once the PHYs TCA status is asserted (i.e. when the PHY has room in its internal transmit cell buffer).

4.2 Connecting Line and Core Cards (Stage 2 Multiplexing)

All inter-card communication between the line cards and the core cards is carried on the 4-wire high-speed LVDS links. No additional wires, clocks, or signals are required between cards. The LVDS transceivers are designed to connect directly to backplane traces or 4-wire (preferably shielded twisted pair) cables up to 10 meters in length. No external drivers are required. The high speed internal transmit clock of the LVDS link is synthesized from a lower (1/8) speed reference clock. The LVDS receiver recovers its clock from the incoming data, so both the Tx and Rx clocking of the LVDS links can be fully asynchronous to the clocking of the UTOPIA bus.

Here we discuss the data path and flow control aspects of the inter-card communications. Later sections discuss the inter-card communication channel, clock, timing, and other signals that are carried by the 4-wire serial connection.

The serial links carry user cells "clear channel" by appending extra bytes to each cell in order to carry all system information. Idle cells are automatically injected when no user data is present to ensure that continuous link error monitoring is available. Loss of receive signal (LOS) and loss of frame (LOF) interrupts are provided, as is far end notification of LOS and LOF conditions. Errored cell and user data cell counters are also provided to assist in link performance monitoring.

In redundant systems, each S/UNI-DUPLEX will be connected to two S/UNI-VORTEX devices, one on the active core card and one on the standby core card. Each S/UNI-VORTEX is capable of terminating 8 LVDS links from 8 S/UNI-DUPLEX devices. Each of the 8 links is independently configurable as active or standby, so a core card can simultaneously act as the active path for one line card but the inactive path for another line card. This allows load sharing between the two core cards and their associated WAN up-links[1].

[1] Load sharing allows each of the two WAN up-links to, under normal operating conditions, carry full traffic loads. Under failure conditions the total up-link bandwidth would be reduced by 50%.

The LVDS links have been designed to support the removal and insertion of cards while the equipment is powered up and carrying traffic. This, for example, allows all traffic to be moved to one core card while the second, standby core card is upgraded or serviced. This "hot swap" capability is a key feature of the point-to-point interconnect architecture.

For systems with greater than 8 line cards, several S/UNI-VORTEX devices will be required on each core card. They share a 16 bit wide, 50 MHz ANY-PHY bus with the ATM layer devices such as the S/UNI-APEX and the S/UNI-ATLAS. The S/UNI-VORTEX devices are always bus slaves. To the bus master, each S/UNI-VORTEX looks like a multi-PHY device supporting up to 264 logical PHYs (8 links times 32 PHYs per link, plus 8 control channels, one per link).

4.2.1 Upstream (to the Core Card) Traffic

In a typical access concentrator the WAN up-link will operate at an OC-3 rate or below (i.e. <155 Mbit/s) while the aggregate upstream burst bandwidth of the PHYs can be much higher. In order to smooth out the upstream traffic bursts with no or minimal loss of traffic there must be buffering and traffic management somewhere in the access concentrator. This is typically handled by a traffic management device such as the PM7326 S/UNI-APEX.

This issue of where upstream traffic buffers are placed in the system architecture is an important one that deserves further discussion. In its simplest form, the system designer of an access multiplexer or switch is faced with three choices:

- put buffering on the line card and only pull upstream traffic off each line card when the WAN up-link can take it, or
- pull the upstream traffic off the line card immediately and buffer it on the core card, such as done by the S/UNI-VORTEX and S/UNI-DUPLEX, or
- put buffering on every card (often done when there is a separate switch fabric between cards).

The first approach requires significant over-engineering of the total amount of buffer space required system-wide and drives these costs onto the line card. It cannot take advantage of the statistical gain made by accumulating traffic bursts across a large number of line cards.[1] As well, under extreme load conditions – situations where upstream traffic must be discarded intelligently – it is very difficult to design the system so that traffic discard is handled optimally and fairly across all line cards. This is especially true when QoS (Quality of Service) issues need to be addressed.

[1] As was mentioned previously, in access applications each line may be idle for significant portions of the time. Further, access speeds vary widely depending on the services being offered. Rate adaptive services such as ADSL vary their speed based the loop conditions of individual customers. Taken together, this wide degree of per line variability lends itself to significant statistical gain if the traffic buffering is centralized.

The second approach eliminates the buffers on the line card,[1] but few things in life are free and this is no exception. The second approach requires that traffic be moved to its buffering point before PHY buffer overflow occurs on the line card. Hence the entire upstream data-path from PHY to first significant buffering point (i.e. the ATM traffic management device) must have sufficient capacity to ensure that it isn't a bottleneck.

For the S/UNI-VORTEX and S/UNI-DUPLEX architecture a good rule of thumb is that the maximum aggregate upstream bandwidth of the PHYs on a single line card should be less than 155 Mbit/s. This eliminates the individual serial links as potential bottlenecks. On the core card the ANY-PHY bus sits between the S/UNI-VORTEXs and the ATM layer, so the 800 Mbit/s bus speed must also be taken into consideration. Hence another rule of thumb is that the maximum sustained upstream data rate, when taken in aggregate from all active PHYs, should be less than 800 Mbit/s. Since the WAN up-link is typically OC-3 or less in speed this is not a significant restriction. **Note also that systems implemented with balanced load sharing between duplicated core cards can essentially double the upstream buffering capacity and the aggregate upstream burst tolerance of the system.**

In very large access multiplexers such as DSLAMs one normally assumes that not every PHY is active at all times. However, in theory it may be physically possible for the aggregate upstream burst bandwidth to exceed the 800 Mbit/s bus capacity unless steps are taken to prevent this error condition. It is impractical to sustain 800 Mbit/s bursting for even a short period of time due to the massive amount of cell buffering that would be required to buffer cells while waiting for a DS-3 or OC-3 WAN up-link to clear the traffic. Hence the multiplexer's call setup software, also known as the Connection Admission Control (CAC) software, will need to prevent potential upstream traffic overflow by simply refusing additional connections if the multiplexer becomes overloaded.

Now that we've discussed why centralized buffering is the preferred approach we can proceed with our description of how the second stage of multiplexing occurs for upstream traffic.

As described on page 7, the S/UNI-DUPLEX multiplexes all upstream traffic into a 200 Mbit/s serial stream and sends it simultaneously to the active and inactive core cards[2] via the 4-wire LVDS serial links. The S/UNI-VORTEX on the active core card terminates the serial link, monitors/counts transmission errors, discards idle cells, extracts the user cells, and places them on the core card's parallel ANY-PHY bus as described in Section 4.3.

The S/UNI-VORTEX schedules upstream traffic from its eight serial links into the ANY-PHY bus. It implements a user programmable weighted round robin polling scheme.

[1] Except for the very small internal buffers used by the S/UNI-VORTEX and S/UNI-DUPLEX to keep the serial links operating efficiently with back-to-back transfers.

[2] The term "active core card" is used rather loosely here since load balancing designs do not have a strictly active and strictly inactive card. However for simplicity we'll continue to use the term "active core card" rather the longer but more accurate "core card active for the line card being discussed".

The per-link weights can usually be left to their default status, which is equal weight on all links. However, a higher weigh may be desirable on a specific link if, for example, the link is being used to connect to a co-processor card or perhaps a high speed PHY card that will be generating a large amount of traffic relative to the other line cards. Weights are relative and linear; a link with a weight two times another will be polled twice as often.

Each S/UNI-VORTEX has a raw upstream bandwidth capability of 1.6 Gbit/s (8 links at 200 Mbit/s each) while the core card ANY-PHY bus is typically a 800 Mbit/s bus (16 bit wide 50 MHz parallel bus). Therefore the S/UNI-VORTEX implements back-pressure on the serial links to prevent the upstream traffic from overflowing at the parallel bus. Each serial link is provided a small internal receive cell buffer. As the receive buffer approaches the full mark the S/UNI-VORTEX sends a back-pressure indication to the S/UNI-DUPLEX via the embedded system overhead channel in the downstream LVDS serial link. The S/UNI-DUPLEX, after receiving the buffer full indication, immediately begins sending idle cells until the back-pressure indication is deasserted by the S/UNI-VORTEX[1].

In systems with duplicated core cards the state of an upstream link is determined (logically) by the state of the active bit on the corresponding downstream link. Upstream links corresponding to inactive downstream links continue to monitor for errors, but can be programmed to stay in the FIFO reset state so that upstream traffic on the spare link is discarded by the S/UNI-VORTEX[2]. Optionally, the inactive links can be programmed to function normally, thereby leaving the spare core card's ATM layer responsible for processing cells quickly enough to prevent buffer overflow in the S/UNI-VORTEX. Buffer overflow on an inactive upstream link is possible because the S/UNI-DUPLEX ignores flow control (back-pressure) information from the inactive link. This must be taken into consideration when determining how or if upstream user cells are handled on spare links.

In some situations less than 8 of the S/UNI-VORTEX LVDS links will be equipped. This occurs if line cards are not equipped, or if a shelf does not have a multiple of 8 line card slots. In this situation the unequipped LVDS links can be disabled through software. The LVDS signal pins for disabled links are left in tri-state.

In summary, the upstream data path in the second stage of multiplexing is:

- Upstream cells are sent by the S/UNI-DUPLEX as soon as they are received from the PHYs. In protected systems the cells are sent simultaneously on both the active and inactive links.
- The S/UNI-DUPLEX will temporarily suspend sending data from all PHYs when the active link is asserting back-pressure. Back-pressure from the inactive link is ignored.

[1] In the upstream direction it is sufficient to use a single back-pressure indication for all channels. Head of line blocking is not an issue because all traffic is being directed to a single port -- the ATM layer device.
[2] On the spare core card embedded control channel cells that are passed through to the ATM layer should normally be left to pass through (i.e. the FIFO should not be held in reset). .

- The serial link and core card bus bandwidths are sufficient to ensure that, in a properly engineered system, any back-pressure is temporary and will not result in buffer overflow at the PHYs.
- The S/UNI-VORTEX services its eight serial links in a simple weighted round robin fashion.
- As each cell is received by the S/UNI-VORTEX it is tagged with a link ID (0..7) and a S/UNI-VORTEX ID (0..31) and made available to the ANY-PHY bus as discussed in *Section 0* - Multiplexing on the Core Card (Stage 3 Multiplexing). Figure 4 shows the port ID information being provided as a cell prepend. It can also be inserted in a standard length cell in the HEC/UDF field.

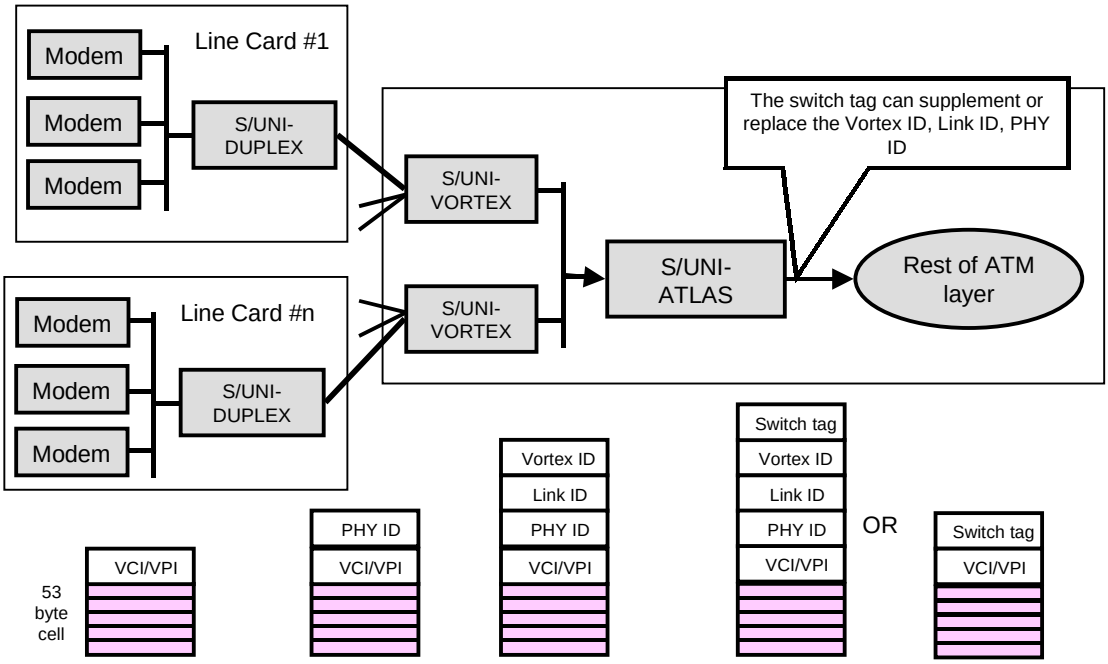

Figure 4 - Upstream Cell Overhead

4.2.2 Downstream (to the Line Card) Traffic

In a typical system the WAN port will operate at a much higher rate than any of the individual PHYs on the line cards. Under burst conditions we want to buffer downstream traffic until the respective PHY is able to receive it. As with the upstream direction, the system designer of an access multiplexer or switch is faced with three choices:

1. put buffering on the line card and send cells to the line card as soon as the cell is received from the WAN port, or

2. only send cells to a PHY when the PHY can accept them, leaving cells buffered on the core card until the PHY clears. This is the approach used in the S/UNI-VORTEX and S/UNI-DUPLEX architecture, or
3. put buffering on all port cards, separated by a switching fabric.

As is the case in the upstream direction, the first approach requires significant over-engineering of the total amount of buffer space required system-wide and drives these costs onto the line card. It cannot take advantage of the statistical gain made by centralizing the buffering of traffic bursts destined to a large number of line cards.[1] To prevent internal bottlenecks the downstream data-path to every line card must be at least as fast as the WAN port if a simple downstream broadcast is used. Otherwise the core card will need to perform demultiplexing on the data stream and only send the cell to its appropriate line card.

Remembering that nothing is free in this world, one potential negative of the approach implemented by the S/UNI-VORTEX and S/UNI-DUPLEX, is that the core card's traffic management device must function across a greater number of PHYs than a traffic management device on a single line card. However, PMC-Sierra's PM7326 S/UNI-APEX Traffic Management device has been designed to perform a full featured traffic and buffer management function across 2048 PHYs while interfacing directly to any number of S/UNI-VORTEX devices.

As described on page 8, the S/UNI-DUPLEX sends the S/UNI-VORTEX the logical equivalent of each PHY's FIFO TCA status[2]. The information is sent on both the active and inactive links, but for the remainder of this section, we will only be discussing the behavior of active serial links. In effect, each S/UNI-VORTEX appears to the traffic manager as a 256 port multi-PHY. This is depicted in Figure 5.

[1] As was mentioned previously, in access applications each line may be idle for significant portions of the time. Further, access speeds vary widely depending on the services being offered. Rate adaptive services such as ADSL vary their speed based the loop conditions of individual customers. Taken together, this wide degree of per line variability lends itself to significant statistical gain if the traffic buffering is centralized.

[2] To be precise, the S/UNI-DUPLEX indicates the status of an internal cell deep buffer associated with each PHY. However, since the status of this buffer (full or empty) depends ultimately on the corresponding PHY's transmit FIFO it is accurate to characterize this back-pressure as a "proxy TCA" indication.

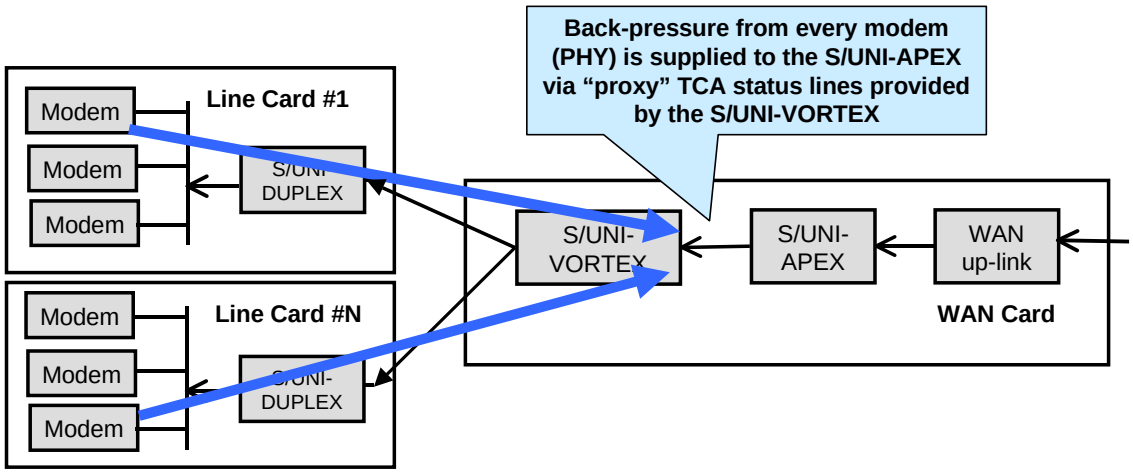

Figure 5 - PHY Back-pressure

For each of its 8 serial links the S/UNI-VORTEX provides an internal cell buffer for each of the maximum 32 PHYs supported by the downstream S/UNI-DUPLEX. This allows cell transfers to occur on different serial links simultaneously. The ANY-PHY bus is four times faster than the serial link, so this ensures the full 800 Mbit/s can be used. Each link schedules cells from its 32 cell buffer (one per downstream PHY) on a strictly round robin basis, where PHYs without downstream cells are skipped over of course.

It should be noted that the maximum effective downstream bandwidth for a single PHY[1] is between ½ and ¼ the bandwidth of the LVDS link, depending on the timing of the ATM layer device acting as bus master. This is due to buffer "high water mark" levels and internal back-pressure signal timing constraints. S/UNI-VORTEX registers can be configured in software to adjust the buffer fill level for line cards with very fast PHYs. An exception to this is when a single PHY is connected to the S/UNI-DUPLEX. In that case the full bandwidth of the LVDS link (up to approximately 90%) is available.

In summary, the downstream data path in the second stage of multiplexing is:

- The S/UNI-VORTEX provides 256 "proxy TCA" signals that can be used directly by the ATM traffic management device to safely schedule cells directed to the PHYs on the line card.
- The serial link bandwidth, the core card bus bandwidth, and the internal buffering in the S/UNI-VORTEX and S/UNI-DUPLEX devices are all sufficient to ensure that downstream traffic destined to one PHY does not block downstream traffic destined to another PHY regardless of where the two PHYs are located (i.e. whether they are on the same or different line cards).

[1] The aggregate bandwidth of all the PHYs on multi-PHY line card can be about 90% of the LVDS link rate. The ½ to ¼ link rate restriction being discussed here applies only to each individual PHY.

- Each link's active indication (in the downstream direction) can be set individually by software. Inactive links (on the S/UNI-VORTEX) still present TCA information to the ATM layer device, but any cells sent on these links will be discarded by the S/UNI-DUPLEX.
- As the S/UNI-VORTEX receives a downstream cell from the ATM layer the in-band prepended address is decoded and used to route the cell to the appropriate link and the appropriate internal PHY buffer. This is discussed further in *Section 6.2, Downstream Addressing.*

4.3 Multiplexing on the Core Card (Stage 3 Multiplexing)

As was discussed in the previous section, in many ways the S/UNI-VORTEX acts as a proxy for all the PHYs attached to the corresponding 8 S/UNI-DUPLEX devices. Since each S/UNI-DUPLEX can interface to a maximum of 32 PHY devices, each S/UNI-VORTEX can represent up to 8 x 32 = 256 PHY devices. Further, there can be up to 31 S/UNI-VORTEX devices directly addressed on a single bus, although electrical limits on the bus will restrict the number of devices to be less. Regardless, this is a lot of PHYs!

In this section we discuss how the traffic from this many PHY devices is handled and addressed by the ATM layer. For reasons that will become obvious shortly, the upstream and downstream directions are handled differently. The remainder of this section focuses on how traffic on the active core card's data path is handled.

4.3.1 Upstream (to the Core Card) Traffic

The ATM layer needs to know which port each cell came from. There are two obvious choices, rely on a unique field within the original cell to identify the source, or add a small, unique PHY ID tag to each cell.

Although it might be tempting to use the existing VPI/VCI address field present in every ATM cell to uniquely identify its port, this is not something that we can guarantee unless we restrict the VPI/VCI values – not a particularly elegant solution. What works better is to "tag" each cell with a prepended physical port ID, and then send the cell to the core card for processing. By adding a short tag to each cell the data path becomes fully "protocol neutral" and makes no assumptions about the contents or address fields within the user cells.

As was described in Section 4.2.1 on page 10, each S/UNI-DUPLEX and S/UNI-VORTEX multiplexes and tags its upstream traffic into a single stream of cells, which are offered one by one to the ATM layer by the S/UNI-VORTEX. Every upstream cell is tagged with prepend bits that uniquely identify the S/UNI-VORTEX ID (0..31), link ID (0..8), and PHY ID (0..31)[1] to which the cell belongs. Since the physical source (i.e. the port) of each upstream cell is self-identified by its tag, the S/UNI-VORTEX can act as if

[1] There is also a control channel cell identification, as discussed in Section 5, The Embedded Inter-device Communication Channel.

the stream of upstream cells is coming from a single PHY. Put another way, in the upstream direction each S/UNI-VORTEX appears as a single PHY slave device supplying a stream of expanded length cells[1] to the ATM layer.

To accommodate more than one S/UNI-VORTEX on the bus, the ATM device need only poll the Receive Cell Available (RCA) status line of each of the S/UNI-VORTEX devices. Simple round robin polling of the devices will normally suffice since back-pressure is used for flow control on the individual serial links, and each S/UNI-VORTEX performs a user programmable weighted round robin polling of its 8 serial links.

In summary, the upstream data path on the core card is simply:

- Each S/UNI-VORTEX services its eight serial links in a simple weighted round robin fashion.
- As each cell is received the S/UNI-VORTEX tags it with a link ID (0..7) and a S/UNI-VORTEX ID (0..31) and makes it available to the ANY-PHY bus. Since the S/UNI-DUPLEX will have added a PHY ID tag, the accumulated tag uniquely identifies the cell's source port. This is shown in Figure 4.
- The ATM layer device, acting as bus master, polls the RCA status of each of the S/UNI-VORTEX devices present on the bus. Each S/UNI-VORTEX looks like a single PHY to the ATM layer device. When a S/UNI-VORTEX indicates that it has a cell available it is read in by the ATM layer device for processing.
- The ATM traffic management device is responsible for buffering upstream cells until they can be forwarded to the WAN link, sent back to a line card (for line to line switching), or otherwise processed.

4.3.2 Downstream (to the Line Card) Traffic

To the ATM layer each of the S/UNI-VORTEX devices on the core card's ANY-PHY bus appears, in the downstream direction, as if it were a 256+8 port multi-PHY[2]. Further, there can be numerous S/UNI-VORTEX devices on the bus, limited mainly by electrical loading. In order to schedule cell traffic into these numerous PHYs the ATM layer device must efficiently perform three related functions: TCA status polling, PHY selection, and cell transfer.

On each bus cycle the ATM layer device, which is bus master, can poll the status of an individual PHY's proxy TCA status. It does this by presenting a PHY address on the ANY-PHY bus address lines. If the S/UNI-VORTEX acting as proxy for the polled PHY has room for the cell in its internal buffer then it will raise its TCA line two bus clock cycles after the polling address is presented. All other S/UNI-VORTEX on the bus will tri-state their TCA line. As discussed below, PHY polling uses a different addressing mechanism than PHY selection for cell transfer. This gives guaranteed and deterministic access to polling bandwidth over the transmit address lines.

[1] Optionally the PHY id tag can be passed in the UDF/HEC word without expanding the cell length.
[2] The extra 8 channels are the embedded control channel per link.

None of the S/UNI-VORTEX devices on the bus will respond to a NULL address (all ones) poll. The ATM layer device will insert a NULL address between valid PHY addresses to give the previously polled S/UNI-VORTEX time to tri-state the TCA line. If the TCA line is not shared among slave devices then a PHY addresses can be presented on every bus cycle.

For PHY selection (to initiate cell transfer) in the downstream direction in-band addressing is used. With this scheme the ATM layer device prepends the selected PHY address to the transmitted cell. The TSX (transmit start of transfer) bus signal is asserted by the ATM layer device during the first cycle of a data block transfer (coinciding with the PHY address) to mark the start of a block transfer period.

All S/UNI-VORTEX devices on the bus receive all cells. It is up to the appropriate S/UNI-VORTEX to recognize the PHY address and latch in the remainder of the cell. The other S/UNI-VORTEX devices will simply ignore the cell. The mapping of the PHY address space of each S/UNI-VORTEX is programmed by software at device initialization time. This is described in the S/UNI-VORTEX Virtual PHY Address Space on page 26.

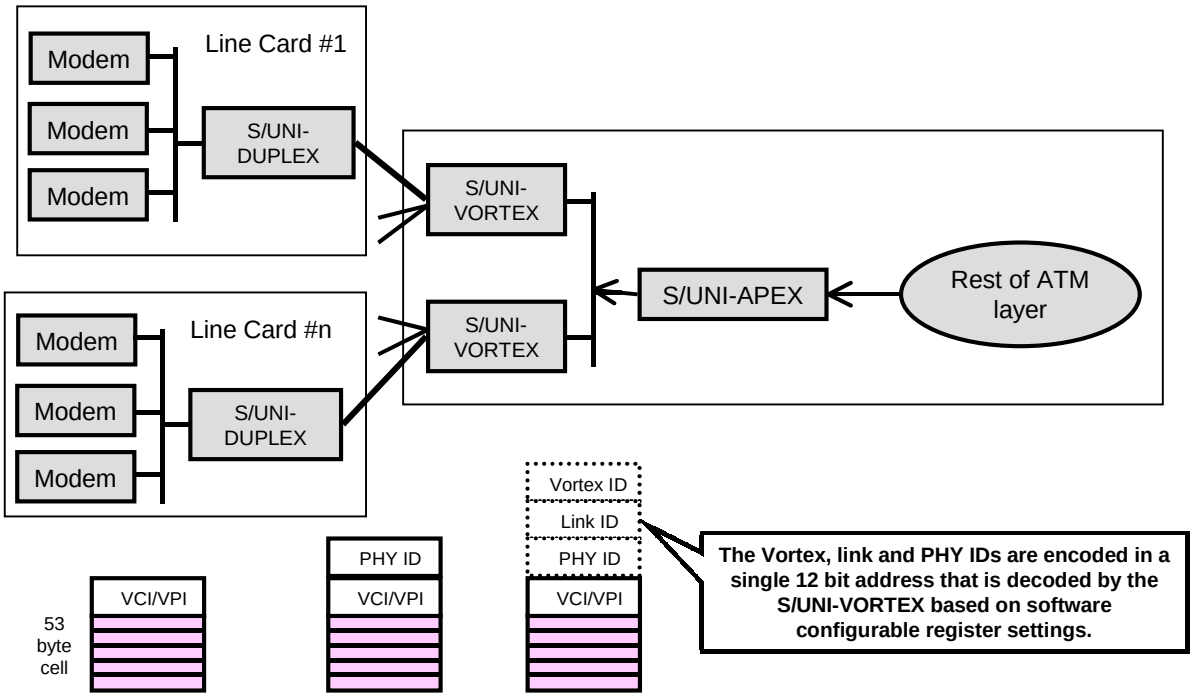

Figure 6 - Downstream Cell Overhead

In summary, the downstream data path on the core card is as follows:

- Traffic arrives from the WAN link and is buffered by the ATM layer.
- The ATM layer prepends each cell with a 12 bit PHY address that identifies the S/UNI-VORTEX, link, and PHY (or identifies the S/UNI-VORTEX, link, and the

embedded control channel). This information is used by the S/UNI-VORTEX to determine whether a cell that is placed on the ANY-PHY bus by the ATM layer should be read in or ignored.

- For each PHY to which the S/UNI-VORTEX is connected (via its 8 S/UNI-DUPLEXs) the S/UNI-VORTEX provides a proxy TCA signal that is polled by the ATM layer via external address lines.

4.3.3 The ATM Layer on the Core Card

ATM layer devices designed for this architecture should take full advantage of the large address space supported by the S/UNI-VORTEX. PMC-Sierra has implemented such a device: the S/UNI-APEX[1]. The S/UNI-APEX provides per VC congestion control (early packet discard, partial packet discard), buffering, and queuing across 64K VCs and multiple classes of service. It implements a highly efficient traffic polling and scheduling algorithm for up to 2048 logical PHYs. It also provides switching and traffic shaping into four high-speed WAN up-link ports.

In the following figure an additional PMC-Sierra device called the S/UNI-ATLAS is also shown. This device manages ingress and egress traffic policing and OAM cell flows[2]. The S/UNI-PLUS is one of PMC-Sierra's OC-3 PHY devices. In this example the S/UNI-VORTEX must be configured to place the upstream PHY ID in the UDF/HEC field because the S/UNI-PLUS cannot handle extended length cells.

[1] Refer to the S/UNI-APEX datasheet. Contact your PMC-Sierra sales representative or PMC-Sierra at info@pmc-sierra.com for details.

[2] Refer to the S/UNI-ATLAS datasheet. Contact your PMC-Sierra sales representative or PMC-Sierra at info@pmc-sierra.com for details.

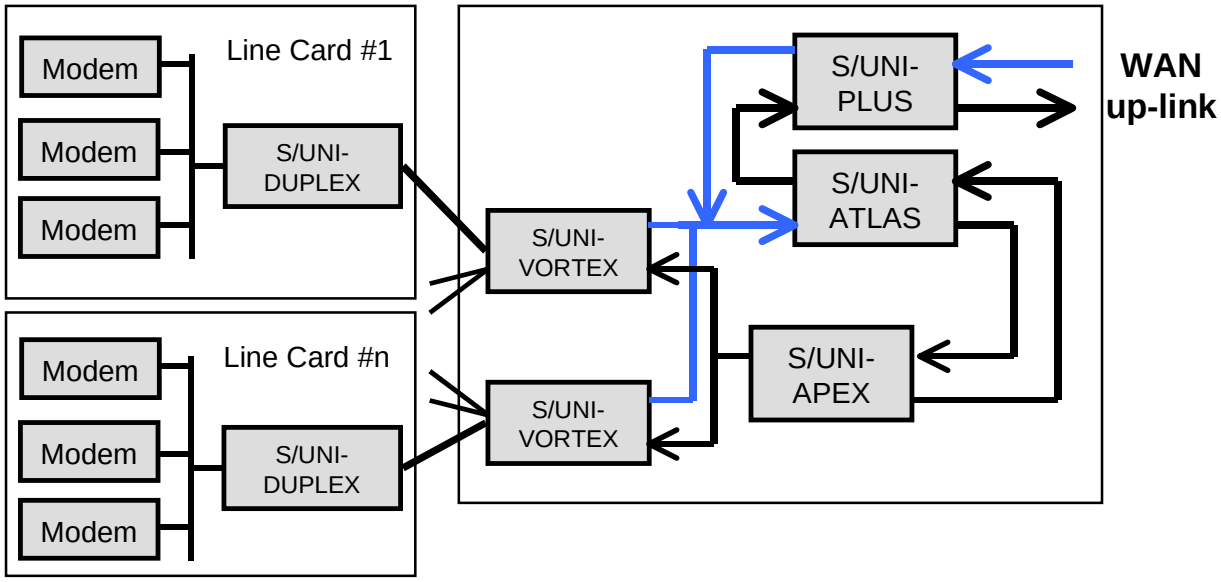

Figure 7 - Typical DSLAM Data Flow

5 THE EMBEDDED INTER-DEVICE COMMUNICATION CHANNEL

Many systems require a microprocessor on each card to perform initialization and status reporting functions. To maximize the value of the serial backplane architecture, the S/UNI-VORTEX and S/UNI-DUPLEX also provide an embedded high speed inter-device communication channel that can be used by the system's embedded control system.

The S/UNI-VORTEX and S/UNI-DUPLEX both provide a microprocessor interface compatible with most industry standard microcontrollers and microprocessors. The microprocessor interface provides access to the device's internal registers, which are primarily used to initialize and monitor the operation of the device. However, there is also a set of registers and associated cell buffers available via the microprocessor interface that can be used to send packets of control system information between cards. This is shown in Figure 8.

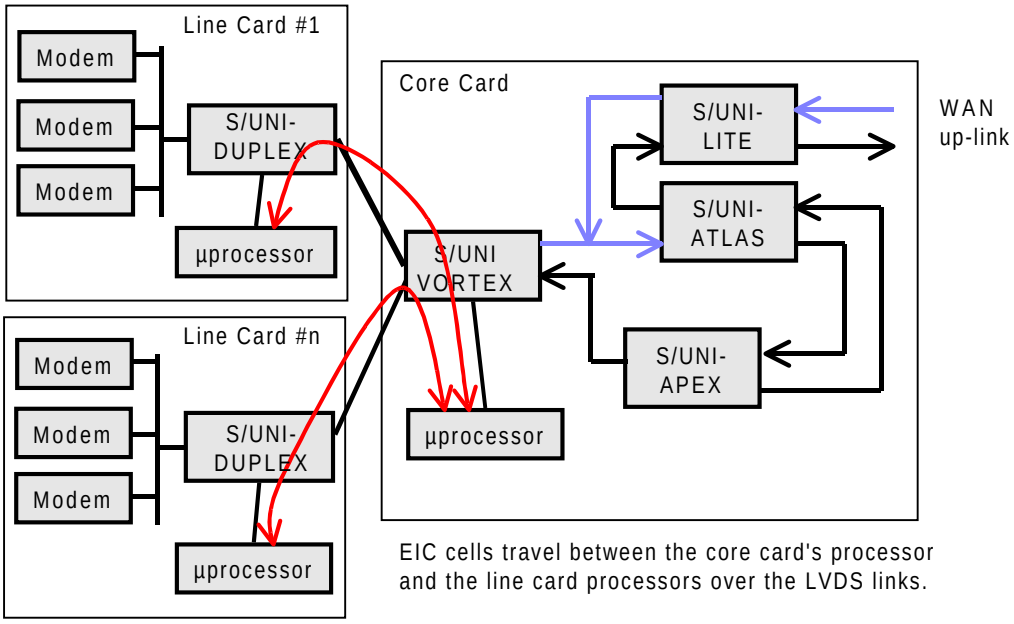

Figure 8 - Embedded Com Channel: S/UNI-VORTEX to S/UNI-DUPLEX

Internally, this embedded inter-device communication channel (hereafter called the EIC channel) works like a "data packet aware" 33rd ATM channel carried over the serial link. As will become clear shortly, by defining the EIC channel this way we create a very flexible system communications path.

5.1 The S/UNI-VORTEX and S/UNI-DUPLEX Communication Channel

The fundamental role of the EIC channel is to carry microprocessor defined, variable length packets between the microprocessor on the line card and the microprocessor on the core card. The S/UNI-VORTEX and S/UNI-DUPLEX do not handle the EIC channel in exactly the same way. The S/UNI-DUPLEX must talk to both the active and the inactive serial links simultaneously, while the S/UNI-VORTEX must handle the 8 independent EIC channels coming from its 8 links. These differences are explained in the following sections.

5.1.1 Line Card to Core Card Communication

When the microprocessor on the line card wishes to communicate with the microprocessor on the core card it will proceed as follows:

- Software running on the line card microprocessor will compose an arbitrary length packet to be sent to the core card.
- The software will divide the packet into N byte segments, where N <= 53[1]. The standard 5 byte ATM header fields (e.g. VCI, VPI, HEC, etc.) are passed clear channel and hence can contain arbitrary information. Normally the higher layer software will include a packet length field in the first segment, or perhaps embed a Start of Message, End of Message indication somewhere in the cell header. If bit error protection is enabled (as is the case in this example) the last segment must reserve the last 4 bytes for the CRC-32 inserted by the S/UNI-DUPLEX.
- To initiate the packet transfer the line card micro writes a CRC-32 reset command to the appropriate S/UNI-DUPLEX's control register, and then writes the first cell, a byte at a time, into the S/UNI-DUPLEX's EIC channel transmit register.
- Once the entire cell has been written into the device, the S/UNI-DUPLEX broadcasts the cell on both the active and inactive channels. As is the case with PHY originated cells, each EIC cell is tagged with system prepend bytes before being sent on the serial link. A reserved PHY ID in the system prepend indicates to the receiver S/UNI-VORTEX that this is an EIC cell. The S/UNI-DUPLEX calculates the CRC-32 across all subsequent EIC cells until the end of packet is indicated.
- When the core card's S/UNI-VORTEX receives an EIC cell it places it into a FIFO and indicates to the core card microprocessor (via interrupt or polled status) that a cell is available. CRC-32 is calculated independently across all EIC cells received on the 8 links.
- To prevent receive buffer overflow the S/UNI-VORTEX asserts back-pressure if the control channel cell FIFO begins to fill. Back-pressure is carried from the S/UNI-VORTEX to the S/UNI-DUPLEX via overhead bits in the downstream

[1] If the extended cell bus option is enabled the cell should match the length of the extended cell. Refer to the data sheet for details.

LVDS link. Back-pressure is provided independently on each of the 8 links by the active channel's S/UNI-VORTEX[1].

- The core card micro reads the received EIC channel cell from the S/UNI-VORTEX register a byte at a time and begins the packet reassembly process.
- The line card micro can write the next 53 byte packet segment once the S/UNI-DUPLEX indicates (via interrupt or status polling) that it is ready to accept another EIC cell. This indication will be deferred if back-pressure has been indicated by the S/UNI-VORTEX.
- Cell transfers continue in this fashion until, on the last packet segment, the line card micro indicates "end of packet" by writing to the appropriate control register. The S/UNI-DUPLEX will insert the CRC-32 value into the last 4 bytes of the cell before sending it to the core card.
- The core card micro will determine the end of packet based on the segmentation protocol being implemented by the software. Once it receives the last segment of the packet the core card micro can check the CRC-32 error indication provided by the S/UNI-VORTEX.
- Requests for retransmission of errored packets is the responsibility of higher layer protocols implemented in software.

5.1.2 Core Card to Line Card Communication

In the downstream direction, the S/UNI-DUPLEX maintains independent embedded communication channels for the active and inactive serial links. Hence, both the active and inactive core card's microprocessor can send packets to any line card. Unlike the upstream direction, back-pressure from S/UNI-DUPLEX to S/UNI-VORTEX is always provided on both the active and inactive links.

When the microprocessor on the core card wishes to communicate with the microprocessor on the line card it will proceed as follows:

- Software running on the core card microprocessor will compose an arbitrary length packet and divide it into cells. It will include a packet length field or BOM/EOM indications, and leave four empty bytes in the last cell.
- To initiate the packet transfer the core card micro must first select the link number (0-7) it wishes to address. It does this by setting the appropriate value in the link selection register of the S/UNI-VORTEX. It then writes the CRC-32 reset command to the appropriate S/UNI-VORTEX's control register, and writes the first cell, a byte at a time, into the S/UNI-VORTEX's EIC channel transmit register.

[1] To prevent the inactive core card from locking out communication with the active core card, only the active serial link's back-pressure is used. However, this does require the inactive core card to "keep up" with the active core card. If this is an issue then the S/UNI-DUPLEX to ATM layer communication channel described in Section 5.2 can be used.

- Once the entire cell has been written, the S/UNI-VORTEX sends the cell on the selected serial link. The S/UNI-VORTEX calculates and stores internally the CRC-32 across all subsequent EIC cells until the end of packet is indicated.
- When the line card's S/UNI-DUPLEX receives an EIC cell it places it into the FIFO associated with that serial link (active or inactive) and indicates to the line card microprocessor (via interrupt or polled status) that a cell is available. If desired the devices can calculate CRC-32 across EIC cells received on the active and inactive links.
- The line card micro will read the received EIC cell from the S/UNI-DUPLEX register a byte at a time and begin the packet reassembly process.
- The core card micro can write the next 53 byte packet segment once the S/UNI-VORTEX indicates (via interrupt or status polling) that it is ready to accept another EIC cell. This indication will be deferred if back-pressure has been indicated. Double buffering is used to improve throughput.
- Cell transfers continue in this fashion until, on the last packet segment, the core card micro indicates "end of packet" by writing to the appropriate control register. The S/UNI-VORTEX will insert the CRC-32 value into the last 4 bytes of the cell before sending it to the line card.
- Once it receives the last segment of the packet the line card micro can check the CRC-32 error indication provided by the S/UNI-DUPLEX.
- Requests for retransmission of errored packets is the responsibility of higher layer protocols implemented in software.

5.2 The S/UNI-DUPLEX to ATM Layer Communication Channel

On core cards with multiple S/UNI-VORTEX devices it may be desirable to allow the embedded inter-device communication channel to pass right through the S/UNI-VORTEX and be terminated by the ATM layer. To accommodate this "pass-through" mode, the S/UNI-VORTEX can be programmed such that the 8 EIC channels are treated in a similar fashion to the 256 PHYs for which the S/UNI-VORTEX is acting as proxy. Hence the S/UNI-VORTEX can be programmed to act as proxy for 264 channels, 8 of which are the EIC channels on each link. This is shown in Figure 9. The addressing scheme implemented by the S/UNI-VORTEX to accommodate this is discussed in Section 6.

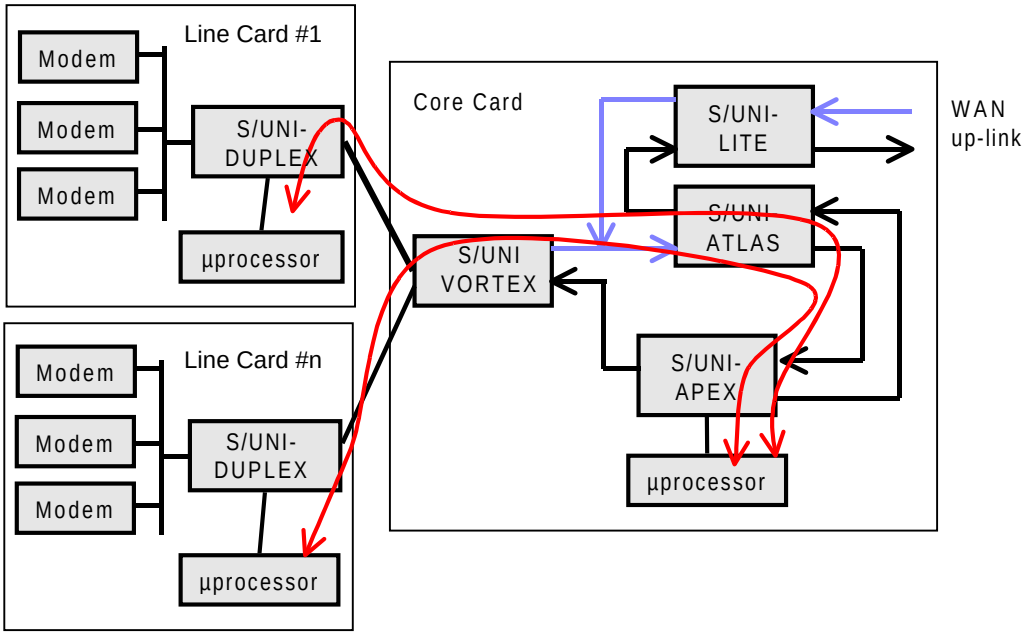

Figure 9 - Embedded Com Channel: S/UNI-APEX to S/UNI-DUPLEX

6 THE S/UNI-VORTEX VIRTUAL PHY ADDRESS SPACE

For systems with less than 32 PHYs per line card the address space occupied by the PHYs (and optionally the embedded inter-device communication channels) will have "gaps" where the unequipped PHYs reside. Depending on the system architecture, this could create some complexity for the ATM layer devices. The S/UNI-VORTEX has implemented an address mapping capability to simplify a sparsely populated address space. As is discussed below, this address mapping works differently in the upstream and downstream directions.

6.1 Upstream Addressing

In the upstream direction each S/UNI-VORTEX appears as a single slave device to the ATM layer. To determine if there are upstream cells available for reading, the ATM layer must poll the various S/UNI-VORTEX devices to check their single RCA status. To assist in the device selection for polling and cell transfers the S/UNI-VORTEX implements an integrated address decode function. When the value of the RADR[4:0] input pins matches the value of the (static) base address input pins VADR[4:0] the device is being addressed for upstream polling or to initiate cell transfer.

Section 4.3.1 describes how, in the upstream direction, the S/UNI-VORTEX and S/UNI-DUPLEX work together to add source PHY address to every cell, either as a prepend or in the UDF/HEC field within the ATM cell. The following 14 bit address map is used:

- 5 bits: Specifies the S/UNI-VORTEX ID. This is always set by the S/UNI-VORTEX to equal the value of the device base address pins VADR[4:0].
- 3 bits: Specifies one of the 8 links.
- 6 bits: Specifies the 32 PHYs (000000:0111111) or the EIC channel (address 111110). The remaining channels are reserved for future use.

Once an upstream cell is read in, the ATM layer will normally use the 14 bit PHY address plus the cell's VPI/VCI field to perform a lookup function[1] that results in a short connection tag being attached to the cell before passing it to the ATM traffic management device. The tag is used by the other ATM layer devices to directly address the context and control information[2] pertaining to this cell. This mapping function permits the ATM layer to act across a smaller, fully populated connection address space even though the physical PHY address space combined with the VCI/VPI address space presents a much larger, sparsely populated address space.

As is discussed in the next section, in the downstream direction the S/UNI-VORTEX can be programmed to perform partial address compression. The address compression

[1] For example, PMC-Sierra's RCMP and S/UNI-ATLAS devices perform this function.

[2] Since the connection tag is normally small (8 or 16 bits for example) it can be used directly as the memory location for the context information used by the switching and multiplexing devices to determine how to process the cells.

implemented in the downstream direction has no impact on the upstream PHY addressing scheme.

6.2 Downstream Addressing

Unlike the upstream direction, in the downstream direction the virtual PHYs for which the S/UNI-VORTEX is acting as proxy must be individually addressed. This permits the ATM layer to scheduling traffic directly into the individual PHYs on each line card.

In systems that have less than 32 PHYs per S/UNI-DUPLEX there would be “gaps” in the PHY address space if every link coming into the S/UNI-VORTEX had the full 32 PHY addresses reserved for it. To simplify the addressing requirements placed on the ATM layer[1] the S/UNI-VORTEX has a programmable address mapping capability that applies to the downstream direction only. Address mapping is used to create a single contiguous address space when line card are equipped with less than 32 PHYs.

Each S/UNI-VORTEX has 8 base address registers and 8 PHY range registers. The combination of these registers permits, for each link, a block of 8, 16, 24 or 32 PHYs to be positioned on an 8 byte boundary anywhere in the 12 bit address range supported by the S/UNI-VORTEX. There is also a single Embedded Inter-device Communication channel base register used to position all 8 EIC channels (one per link) on an 8 byte boundary in the 12 bit downstream address range.

As was discussed in Section 4.3.2, the ATM layer uses hardware (bus based) polling to determine the status of an individual PHY’s proxy TCA status. Polling is performed by presenting PHY addresses on the ANY-PHY bus’s address lines and monitoring the status of the TCA lines from the S/UNI-VORTEXs. To address a specific PHY the address presented must take into account the value of the base address registers of the corresponding link in the corresponding S/UNI-VORTEX.

PHY selection to initiate cell transfer is based on in-band addressing. The address prepended to each cell must also take into account the value of the base address registers. In other words, the 12 bit in-band address for PHY X must match the 12 bit address used to poll the TCA status of PHY X.

[1] The main simplification is that the ATM layer device need not be capable of directly polling full address space supported when numerous S/UNI-VORTEX devices share the bus.

7 DISCRETE COMMAND AND CONTROL SIGNALS

The transport of Bit Oriented Codes (BOC) and a RESET signal are two additional features implemented by the S/UNI-VORTEX and S/UNI-DUPLEX. These features are particularly valuable in multi-shelf designs since they eliminate the need for additional inter-shelf wiring.

Bit Oriented Codes (BOCs) are carried in system overhead embedded in the high speed serial link between the S/UNI-VORTEX and the S/UNI-DUPLEX . 63 possible codes carry predefined or user defined signaling. For maximum fault tolerance the BOC is transmitted as a repeating 16-bit sequence consisting of 8 ones, a zero, 6 code bits, and a trailing zero (111111110xxxxxx0). The code to be transmitted is programmed by writing to the appropriate control registers. Codes can be sent independently by the S/UNI-VORTEX and the S/UNI-DUPLEX.

The receiver can be enabled to declare a received code valid if it has been observed for 8 out of 10 times or for 4 out of 5 times. A maskable interrupt is generated to signal when a detected code has been validated, or a status register can be polled periodically. The value of the received BOC is made available to the microprocessor through a control register.

Function	Codeword
Remote Defect Indication (RDI)	11111111 00000000
Loop-back activate	11111111 00000010
Loop-back deactivate	11111111 00000100
Remote reset activate	11111111 00000110
Remote reset deactivate	11111111 00001000
Reserved	11111111 00001010
. . .	. . .
Reserved	11111111 00100000
User Defined	11111111 00100010
. . .	. . .
User Defined	11111111 01111100
Idle Code	11111111 01111110

The valid codes are shown above. The Customer Defined codes may be used without restriction. Four codes (loop-back activate/deactivate, remote reset activate/deactivate) cause autonomous action by the device, and one code (RDI) is generated by the device automatically. RDI takes precedence over the programmed codes. These are explained as follows:

Loop-back Activate/Deactivate Once the Loop-back Activate code is validated the received signal is looped back onto the transmitted sign on the corresponding high speed serial link. The loop-back is cleared upon the validation of the loop-

back deactivate code. For example, this command might be used by the control system to perform power up testing or fault isolation functions.

Remote Reset Activate/Deactivate These code words are used to control the RSTOB output pin on the S/UNI-DUPLEX. They have no effect in the S/UNI-VORTEX (they are treated the same as user codes). Once the remote reset activate code is validated the S/UNI-DUPLEX asserts the RSTOB output pin low. If the remote reset deactivate code is validated the RSTOB output becomes high impedance. For example, this command might be used by the core card's control system to force a hardware reset on a line card microprocessor that has stopped responding to the embedded inter-device communication channel.

Remote Defect Indication (RDI) RDI is sent in the transmit direction whenever Loss of Signal (LOS) or Loss of Cell Delineation (LCD) is declared by the high speed serial link receiver. For example, the core card microprocessor might initiate protection switching on the failed link whenever the RDI code is received.

8 CLOCK AND TIMING DISTRIBUTION

To fully implement the inter-card communications requirements for as wide a range of equipment as possible the S/UNI-VORTEX and S/UNI-DUPLEX provide two additional features: low speed reference clock distribution, and LVDS serial link receive clock extraction.

Low Speed Reference Clock Distribution

Both devices provide an input and output pin which can be used to distribute low speed reference clocks between cards. In communications equipment this is typically a 8 kHz clock used for timing PCM voice circuitry. The reference clock's rising edge is encoded and carried through each of the serial links in the system overhead. A simple byte marker is included in the system prepend byte in each cell carried over the serial link. This marker identifies the nearest byte boundary that aligns with the most recent rising clock edge. The receiving device pulses its corresponding clock output pin at the appropriate byte boundary.

This approach allows the reference clock to be an arbitrary frequency and yet be carried in-band independent of the serial link's bit rate. However it has limitations that must be accounted for. For example, timing jitter will occur due to the edge encoding scheme. For systems with protection switching, the reference clock on the line card will experience a temporary perturbation as the S/UNI-DUPLEX switches active links. Also, although they will be very rare, cell errors over the serial link could cause a clock edge to be missed. Hence for most applications the designer will want to consider using an external PLL to clean up the reference clock.

Serial Link Receive Clock

In both the S/UNI-VORTEX and S/UNI-DUPLEX, the LVDS receiver recovers its clock from the incoming data. Hence the Tx and Rx clocking of the LVDS links can be fully asynchronous to the clocking of the UTOPIA bus. On the S/UNI-DUPLEX device the clock recovered by the LVDS receiver (divided by 8) is made available on an output pin for optional use on the line card. The clock recovery circuitry deployed by the S/UNI-DUPLEX does not guarantee a jitter or glitch free reference clock, so the appropriate jitter attenuation circuitry should be used to "clean up" the recovered clock.

For example, the recovered, jitter attenuated LVDS receive clock could be used on the line card to supply the LVDS transmit clock. This would allow the line card to be loop timed from the core card.

9 INTERFACING TO ATM PHY DEVICES WITHOUT UTOPIA

Physical layer framer or modem devices without integrated I.432 processing[1] normally support a clock + data interface, and rely on external circuitry to detect and generate ATM cell framing and overhead. To support these devices, the S/UNI-DUPLEX provides a clock+data mode[2]. In this mode, the input/output pins that normally interface to the UTOPIA bus are configured to support up to 16 clock + data serial interfaces. This type of line card is shown in Figure 10.

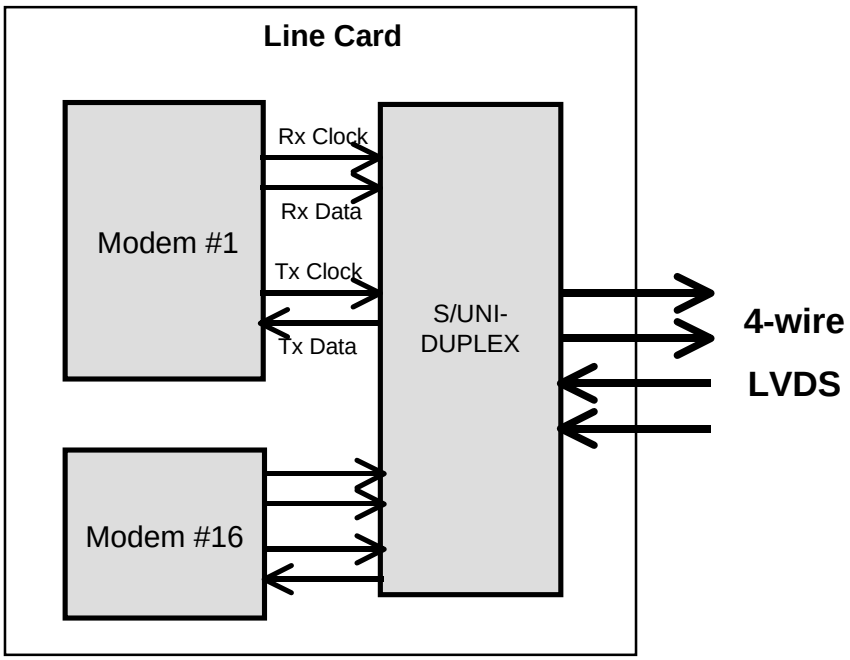

Figure 10 - Clock and Data PHY Interface

Some PHY devices provide a 3-line interface consisting of clock, data, and overhead indication. For these PHYs external circuitry can be used to adapt to the S/UNI-DUPLEX's 2-line interface.

In the clock + data receive direction the S/UNI-DUPLEX performs bit level ATM cell delineation function. In the transmit direction the S/UNI-DUPLEX can operate in either bit or frame aligned mode. In frame mode (also called byte aligned mode) the two wire transmit interface continuously monitors for gaps in the transmit clock to determine where the frame or byte alignment should occur. The circuitry assumes that when a gap in the transmit clock is detected this is either the framing bit position (e.g. the DS-1

[1] Cell delineation, payload scrambling-descrambling, idle cell generation/discard, etc..
[2] Either UTOPIA mode or clock+data mode can be selected, but not both at once.

framing bit) or an overhead byte (e.g. ADSL modem). In either case the next clock period after the gap is assumed to represent the byte alignment position.

10 USING THE S/UNI-DUPLEX TO TAKE THE UTOPIA BUS OFF-CARD

As described in this document the S/UNI-DUPLEX operates as an 8 or 16 bit UTOPIA bus master in both the transmit and receive directions. However, the S/UNI-DUPLEX can be configured in several configurations, as shown below. Note that the configuration of the bus input port is independent of the configuration of the bus output port.

Port	Bus Mode	Bus Format	Features
Input	Master	UTOPIA	31 PHYs, standard length cells
Input	Master	SCI-PHY	32 PHYs, cell prepends allowed.
Input	Master	ANY-PHY	Same as SCI-PHY (decoding of in-band port ID not supported)
Output	Master	UTOPIA	31 PHYs, standard length cells
Output	Master	SCI-PHY	32 PHYs, cell prepends allowed
Output	Master	ANY-PHY	Same as SCI-PHY (generation of in-band port ID not supported)
Input	Slave	UTOPIA	31 PHYs, standard length cells
Input	Slave	SCI-PHY	32 PHYs, cell prepends allowed
Input	Slave	ANY-PHY	32 PHYs, cell prepends allowed, decoding of in-band port ID supported
Output	Slave	UTOPIA	Single PHY, port ID encoded in UDF field
Output	Slave	SCI-PHY	Single PHY, port ID encoded in UDF field, cell prepends allowed
Output	Slave	ANY-PHY	Single PHY, port ID encoded in system prepend, cell prepends (user) allowed

For example, in the configuration shown in Figure 11, two S/UNI-DUPLEX devices can be connected to provide a simple UTOPIA bus extension capability. As a bus slave the S/UNI-DUPLEX functions much like a single link S/UNI-VORTEX. Indeed the S/UNI-DUPLEX in slave mode is bus compatible with the S/UNI-VORTEX (which is always a slave).

Another possibility is to operate both S/UNI-DUPLEX devices in the same mode, thereby acting as a “bus bridge” between two devices that are both slaves, or both masters.

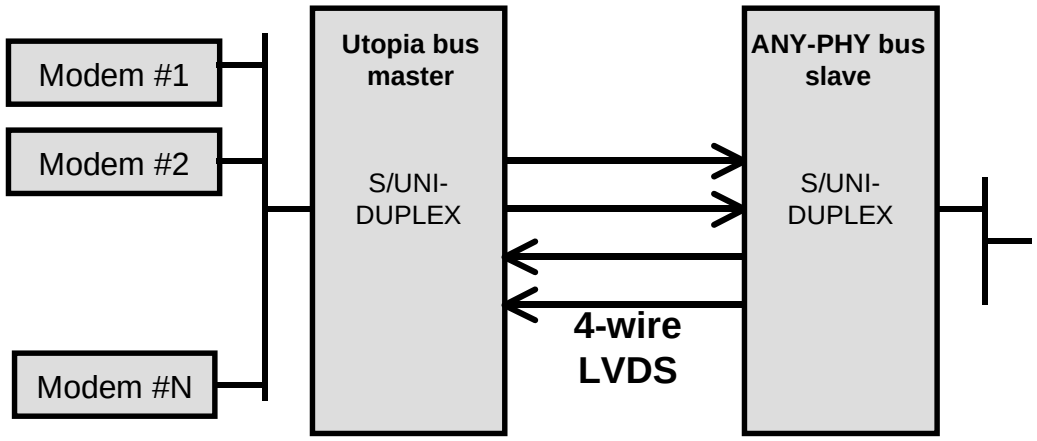

Figure 11 - S/UNI-DUPLEX to S/UNI-DUPLEX Connection

11 GLOSSARY

back-pressure	In this document the term back-pressure refers to an indication from the receiver to the transmitter that the receiver's buffer (normally a cell FIFO) is becoming full and the transmitter should hold off transmitting any more cells until the back-pressure indication is deasserted. For example, on the UTOPIA bus back-pressure from the PHY to the bus master is indicated via the TCA bus signal. Back-pressure from S/UNI-DUPLEX to S/UNI-VORTEX (or visa versa) is indicated by predefined bits embedded in the cell overhead of the LVDS cell format.
BOM	Beginning of Message. Normally used to mark the first segment or cell of a multi-cell packet.
cell	A fixed length unit of data transfer. When used in ATM systems cells are standardized as 53 bytes long – 48 bytes of user data and 5 bytes of overhead. The S/UNI-VORTEX and S/UNI-DUPLEX can operate on 52, 53, 54, 56, or 58 byte cells, and there is no restriction on the content of these cells. Hence there is no requirement that the cell overhead or length be restricted to ATM standards.
core card	A printed circuit board consisting of the circuitry necessary to multiplex and/or switch data traffic to and from the line cards and the up-link port. Also called the WAN card in this document.
DSLAM	Digital Subscriber Line Access Multiplexer. A C.O. located access concentrator terminating local loops.
EIC channel	Embedded Inter-device Communications channel. A communication channel accessed via the microprocessor ports of the S/UNI-VORTEX and S/UNI-DUPLEX, or via the parallel bus of the S/UNI-VORTEX. It is used to send packets of information between the devices via an embedded channel in the high speed serial link.
EOM	End of Message. Normally used to mark the last segment or cell of a multi-cell packet.
head of line blocking	Head of line blocking occurs when a cell at the head of a queue cannot proceed, and it is stopping cells behind it from proceeding even though in theory those cells could proceed if the blocking cell was not in the way. Head of line blocking is an undesirable condition that is avoided completely in the S/UNI-VORTEX and S/UNI-DUPLEX data path.

line card	A printed circuit board on which resides the circuitry necessary to terminate one or more lower speed transmission interfaces.
LVDS	Low Voltage Differential Signal. An IEEE standard defining a 4 wire serial transmission format suitable for backplane and short cable transmission.
PHY	A layer 1 (physical or transmission layer) device capable of transmitting and receiving a signal carrying cell structured traffic. An example of a PHY is an ADSL modem.
QoS	In its most general sense, Quality of Service means that some user traffic is handled differently than the rest of the traffic. The impact that QoS requirements have on system implementation is often most significant under heavy load … the type of situation where the traffic is piling up waiting for a slow modem or a congested WAN up-link.
RCA	Receive Cell Available. This is a standard UTOPIA bus signal that defines the status of a bus slave's receive FIFO. Remember that UTOPIA defines all signals with respect to the bus master, so RCA asserted means the slave has at least one cell for the bus master to receive.
TCA	Transmit Cell Available. This is a standard UTOPIA bus signal that defines the status of a bus slave's transmit FIFO. TCA asserted means the slave has room for at least one more cell. The bus master is responsible for polling the TCA line and sending the PHY a cell only when it has room to accept it.
TSX	Transmit Start of Cell. This ANY-PHY signal is asserted to denote the first byte or word of a cell transfer. When in-band PHY selection is used the first word is a cell prepend that identifies the S/UNI-VORTEX device, the serial link (i.e. line card), and PHY ID to which the cell belongs.
UTOPIA	A parallel bus specification and standard defined by the ATM forum. It is available from the ATM Forum web site at ftp://ftp.atmforum.com/pub/approved-specs/af-phy-0017.000.pdf.
WAN up-link	A high speed transmission interface used to transport traffic to and from the switch or multiplexer.
WAN card	A printed circuit board consisting of the circuitry necessary to multiplex and/or switch data traffic to and from the line cards and the up-link port. Also called the core card in this document.

NOTES

CONTACTING PMC-SIERRA, INC.

PMC-Sierra, Inc.
105-8555 Baxter Place Burnaby, BC
Canada V5A 4V7

Tel: (604) 415-6000

Fax: (604) 415-6200

Document Information: document@pmc-sierra.com
Corporate Information: info@pmc-sierra.com
Application Information: apps@pmc-sierra.com
Web Site: http://www.pmc-sierra.com

PM-981025 (P2) Issue date: June 1999