ATM drop cable interface for the DEC Turbochannel. The Olivetti Research 'Yes V2' option module.

DJ Greaves and Leslie French.

February 20, 1992

The 'Yes' Turbochannel option module is an ATM networking card for the systems which use the DEC Turbochannel option bus. Figure 1 is a block diagram of the option module. The 'Yes' option provides physical layer transmission and reception facilities for the sending and receiving of ATM cells over twisted pair cable at 100 Mbit/second. The modulation technique is 4B5B/NRZI using the AMD TAXI parts. The option module contains FIFO buffering for cells just received or ready to be transmitted. Host access through both polled and DMA access is supported. There are a number of possible programmable interrupt conditions. A delayed-receive interrupt mode is available, whereby the host is not interrupted until a complete, multi-cell message, including the End-Of-Message cell, has accumulated in the reception FIFO. There is no cell-level protocol support, for instance for header generation, SAR operations or demultiplexing.

The ATM cell format used consists of a 48 byte payload with a 5 byte header. The header check byte, which is the fifth byte in the header, is optionally inserted by the option module, if the appropriate bit in the control register is set. The header is checked on reception, if present, indicated by the particular cell start-of-cell physical layer delimiter.

Transmission order is either 'big-endian' or 'little-endian' according to a bit in the control register. Little-endian means that the byte from bits 0 through 7 of a 32 bit word is transmitted first.

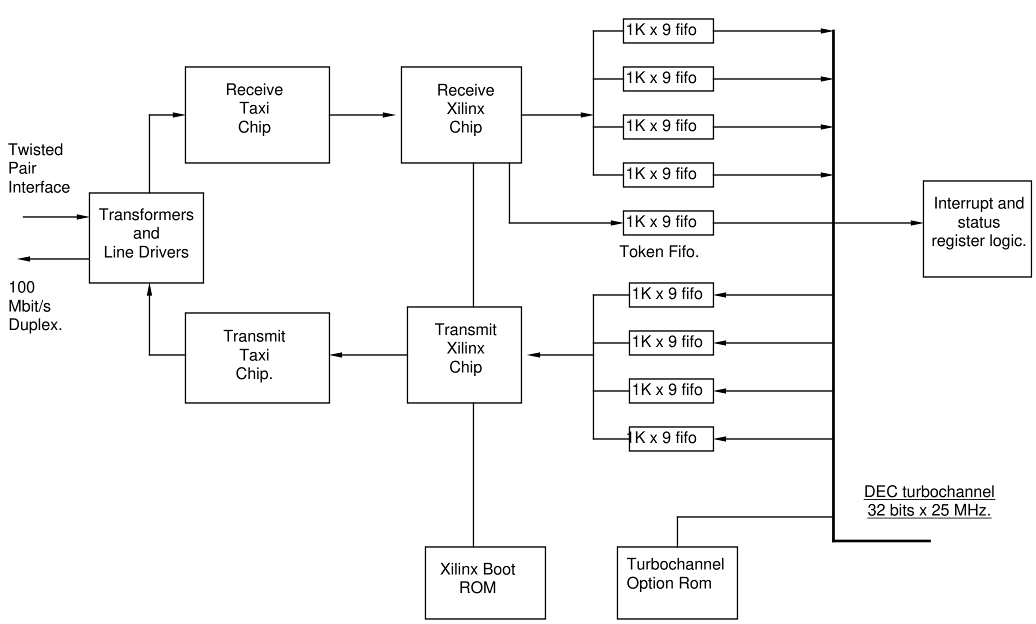

Figure 1: Block diagram of the 'Yes V2' option module.

Name	Width	Cycle types	Byte Offset
Length counter	8	read or write	0x0010
Reset command	n/a	write	0x0014
XReset command	n/a	write	0x0018
Start DMA Tx	32 bit address	write	0x0004
Start DMA Rx	32 bit address	write	0x0008
Control register	8	write only	0x011C
Status register	16	read only	0x001C
Rx FIFO	32	read only	0x0030
Tx FIFO	32	write only	0x0130

Table 1: Register allocation within the 'Yes' option module.

1 Programming Model

Table 1 shows the register offsets within the interface. The base address for the registers is equal to the system and option slot dependent turbochannel programmed IO base plus an offset internal to the 'Yes' board. The internal offset is 0x200000 (0x20 << 16).

The area from 0x000000 to 0x01FFFC (0x7FFF << 2) contains the onboard option EPROM, which is one byte wide, positioned in bits 0 to 7.

The 'Yes' option module contains its own status and control registers. Their bit assignment and meaning is described in section 2.

Writing any value to the 'reset' location causes a reset of the option module. This clears the control register and empties the FIFOs. The control register must be reprogrammed after any 'reset' command.

The 'xreset' location performs the same functions, but in addition, resets the two TAXI transmission chips. This will result in violation flags being set at both ends of the link, which can be cleared by reading the status register once, a sufficient time afterwards. A typical requirement is 500 microseconds. Xreset is more difficult to use than the 'reset', but is the only option if the TAXI devices have become upset, as indicated by repeated violations.

Polled, or programmed, IO may be performed by explicitly writing or reading each 32 bit word of each cell to or from the Tx and Rx FIFOs respectively.

DMA transfers are requested by writing the base physical address of the buffer to one of the DMA start locations. Note that the physical address should be rotated left three bits before writing to the option module, as required by the Turbochannel specification.

'YES V2' OPTION MODULE STATUS REGISTER (LO)

Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit2	Bit 1	Bit 0
Tx FIFO Available	Rx cell Ready	Eom Cell Received	DMA Complete	Tx Flow Blocked	Rx Idle	Rx bad Fault	Tx Idle

'YES V2' OPTION MODULE STATUS REGISTER (HI)

Bit 15	Bit 14	Bit 13	Bit 12	Bit 11	Bit 10	Bit 9	Bit 8
Spare	Tx FIFO Not Full	(Rx FIFO Not Empty)	Tx FIFO Overrun	Rx Hec Bad	Rx Line Violation	Rx FIFO Underrun	Rx FIFO Overrun

Table 2: Bit positions in the status register.

2 Status and control registers.

Table 2 shows the status register bit assignments. They are now described.

Bit 14 - Tx FIFO Not Full. This bit is clear if the Tx FIFO is competely full and any further writes to it would cause an overrun.

Bit 13 - (Rx FIFO Not Empty.) Clear when the Rx FIFO is empty and reading the Rx FIFO would cause an underrun. This flag should be ignored for simple drivers. Use Bit 6 instead.

Bit 12 - Tx FIFO Overrun. Error. Set if Tx FIFO written when full, either by DMA or programmed IO. This bit is cleared by any read of this register.

Bit 11 - Rx Hec Bad. Error. A received cell was received with a bad header check byte. This bit is cleared by any read of this register.

Bit 10 - Rx Line Violation. Error. Latched indication that the TAXI receiver chip strobed its violation output. This indicates that invalid bit sequences have been received down the drop cable. This bit is cleared by any read of this register.

Bit 9 - Rx FIFO Underrun. Error, the Rx FIFO was read when empty. This bit is cleared by any read of this register.

Bit 8 - Rx FIFO Overrun. Error. Too much data has been received without being read out of the interface. Partial cells have most probably been written to the FIFO. This bit is cleared by any read of this register. This fault should be cleared with a write to the reset location.

Bit 7 - Tx FIFO Available. Set if the Tx FIFO is less than half full. This bit is available as an interrupt.

Bit 6 - Rx Cell Ready. Set if the token FIFO is non empty, indicating that at least one cell is available in the interface. This bit is available as an interrupt.

Bit 5 - EOM Cell Received. A cell with its push bit set has been received. The push bit normally indicates last cell of a block. This bit is available as an interrupt. This bit is cleared by any read of this register.

Bit 4 - DMA Complete. A DMA transfer has ended. This bit is available as an interrupt. This bit is cleared by any read of this register.

Bit 3 - Rx Idle. Set if the receiver finite state machine is not in the middle of a cell reception. Faulty (short) cells received on the link can leave the receiver in a non-idle mode. If this bit is clear for an undue length of time without receiving valid data, interface reset is required.

Bit 2 - Tx Flow Blocked. The interface has received an Xoff character from the switch and will not send any further data until re-enabled by Xon or reset.

Bit 1 - Rx Bad Fault. Set if any of the above receive error flags is set. This bit is available as an

'YES V2' OPTION MODULE CONTROL REGISTER

Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit2	Bit 1	Bit 0
Tx FIFO Available I-Enable	Rx Cell Ready I-Enable	Eom Cell Received I-Enable	Dma Complete I-Enable	Tx Flow Control Enable	Transmit Header Check	Big Endian Mode	Rx Bad Fault I-Enable

Table 3: Bit positions in the control register.

interrupt. This bit is cleared by any read of this register.

Bit 0 - Tx Idle. This flag becomes unset when midway through writing cells into the interface. If this bit is clear for an undue length of time without receiving valid data, interface reset is required.

Table 2 shows the control register bit assignments. All control register bits are cleared by hardware reset or the reset command.

Bits 7, 6, 5, 4 and 0 are interrupt enables. When set, they enable interrupt when the flag in the status register of the same name is set.

Bit 3 - Tx Flow Control. When set, the transmitter will reflect the state of the receive data FIFOs by sending Xoff and Xon command messages.

Bit 2 - Transmit Header Check. When set, an extra byte is inserted after the four byte VCI field containing the header check. Different start of cell out-of-band characters are sent depending on whether this bit is set or clear.

Bit 1 - Big Endian Mode. This bit affects receive and transmit in the same way. When set, data from bits 24 to 31 of a 32 bit word are put in the first byte of a cell. When clear, the byte from bits 0 to 7 is put first. This flag should (normally) be set for 68000 systems and clear for DEC Stations.

3 Length register

The length register is an 8 bit counter, incremented after each 32 bit word is transmitted or received from the interface by DMA. It may be read and written via programmed IO with any value.

During DMA transmission, where data is read from the host and sent as cells on the link, the last word will be transferred when the value in the counter reaches 0x02. After this transfer, the counter will contain 0x03. The highest starting value is 0xFE, making a five word transfer the smallest possible for transmit. (The counter wraps around fromm 0xFF to 0x00 during every DMA transfer.) To transfer a 12 word cell payload, the starting value should be 0xF7.

In order to transmit, the 13 words of a whole cell by DMA, the initial counter value needs to be 0xF6 and the final value will be 0x03, as always.

It is only sensible to transmit an integral number of cells, implying an integral multiple of 13 words. Since the Turbochannel can copy data into the 'Yes' option at 800 Mbit/second, while data is only transmitted at 100 Mbit/second, the Tx FIFO can overflow if too long or too frequent DMA transfers are performed. The FIFO sizes fitted may be read from the option ROM at locations TxFIFOSIZE and RxFIFOSIZE. They are typically 512, 1024 or 2048 words of 32 bits. Note also that the host computer has a system-dependent maximum DMA transfer size owing to DRAM page-mode line size.

For DMA reception, the length counter is available to limit the total amount of data transferred and to indicate the amount actually transfered. A receive DMA will normally terminate when there are no further *complete* cells in the receive interface. That is, when the token FIFO goes empty. It will also stop if the length counter reaches 0x02. This is one less than for transmit.

The initial setting of the length counter must be set by the programmer to a value such that the maximum DMA transfer count set by the system is not exceeded and such that the buffer provided for the received data is not overrun.

4 LEDs, Links and Connectors

Link L1 and connector J1 are for remote booting the Xilinx devices. All pins of these links should be left open for normal operation.

One green LED is on during physical layer cell reception and the other is on during physical layer cell transmission. The yellow LED is on when the violation flag in the status register is set.

The SMB connector sites are available for coax drop cable implementations. J2 is the output and J3 is the input. The receiving termination resistor needs changing to 50 ohms when using J3 and on V2.0 boards only, other changes are also required.

The drop cable J4 connector is a 9 pin D male. Pins 1 and 6 are transmit data and pins 5 and 9 are receive data. Drop cables should cross over the pairs. Pin 3 is local ground. To avoid ground loops, the cable screen should be connected to pin 3 at one end only.

5 Drop cable flow control option.

Hardware transmit flow control is available if bit 2 of the control register is set. When enabled, receipt of command FDDI symbol TS from the drop cable will cause transmissions from the 'Yes' board to be blocked and bit 3 of the status register to be set. Any cell currently being transmitted will be completed. Transmissions will recommence on receipt of an II symbol and bit 2 of the status register will return to zero.

6 VME mezzanine system.

The 'CBG turbochannel vme mezzanine card' enables two turbochannel cards, such as the 'Yes' card, to be plugged into a double-height, standard depth VME rack. Several mezzanine cards may be plugged into a single A32/D32 VME system and are distinguished by their base address. The base address of the card may be selected by links for any multiple of 0x2000.0000, give a set of eight possible base addresses.

The two turbochannel slots of each card are termed 'slot 0' and 'slot 1' and are separated in card address space by 0x1000.0000.

For wanda, the prefered mapping into the VME physical address map, as follows

```
Board Number       Base Address      End address
------------       ------------      -----------
Card 0 Slot 0     0x8000.0000         0x8FFF.FFFF
Card 0 Slot 1     0x9000.0000         0x9FFF.FFFF
Card 1 Slot 0     0xA000.0000         0xAFFF.FFFF
Card 1 Slot 1     0xB000.0000         0xBFFF.FFFF
```

For example, the status register of a 'Yes' card in slot zero appears as the bottom sixteen bits of the 32 bit word at address 0x8020001C.

From a programming point of view, the mezzanine card is designed to be transparent. It has no programmable state.

Only VME D32 longword access is supported.

DMA addresses are rotated by three bits, as is the case for all turbochannel systems. When the mezzanine card is bus master there is no provision for bus dtack time-out, so if an option performs a DMA transfer to a non-existent VME address, the VME bus will hang until the VME bus arbiter times out.

DMA transfer between the two option modules plugged into the same mezannine card will NOT work.

Hardware reset of the VME rack causes hardware reset of the two turbochannel slots.

6.1 Interrupts and DMA

A turbochannel option expects interrupt vectors to be generated by its system module. The mezannine card must adopt this function, since the VME requires that VME cards generate their own vectors. Interrupts are generated at an interrupt level and vector set with links on the mezzanine card. The vector for slot 1 is the increment of the vector for slot 0. (There are 7 links for the vector, three for the level in binary and 7 for the level in unary). Interrupts can be disabled by setting the level links for level 0.

Bus requests for DMA are generated at a level set with links on the mezzanine card. Internally, the mezzanine card performs arbitration between the two slots, with transfers from the turbochannel modules to the vme (DMA write) having priority over the reverse direction and round-robin arbitration between the two slots within the same direction.

7 TAXI physical layer specification (Mark Hayter).

Cells are delineated on the physical medium using a 4B5B non-data, command symbol. This is the FDDI TT symbol, which precedes each cell.

There can be any number of synch characters inserted at any byte boundary, inside or outside a cell body. On transmission, at least one synch character is inserted between cells. Additional synchs may appear within cells if programmed IO is being used and the software is slow or pre-empted.[1]

Discussions between the Computer lab and Olivetti have suggested a few changes and extensions to the ATM drop cable interface proposed by Sun/Xerox. These include extensions to the commands used between the Taxi chips to allow hardware handshaking. In addition to the use of optical fiber, as proposed in the Sun/Xerox document, we propose to use twisted pair links for shorter distances. In both cases the link distance is short, with an upper limit of 500 metres for a fiber link and 25 metres for a twisted-pair link. The TAXIchip family allow both systems to be used, having compatibility between the FOXI and TAXI devices.

7.1 Cell Format

The cell format is essentially the same as for B-ISDN, except nothing is said about the header data format. Thus a cell has a 4 octet header, followed by a single octet Header Error Check (HEC) and a 48 octet payload. The HEC may be omitted since there is dispute about its usefulness on a Taxi

[1] This physical layer format conforms to the Sun, Xerox, ORL, Cambridge Consortium ATM workstation physical layer specification.

```
;  YES TURBOchannel Option Rom Stub
;
;

Size        EQU     32768

            ORG     0x03E0>>2

            DEFB    1               ; rom width 1 byte
            DEFB    4               ; rom stride
            DEFB    Size>>13        ; rom size in 8k units
            DEFB    1               ; memory in 4M units

; signature

            DEFB    0x55,0x00,0xAA,0xFF

; Naming and mapping

            ORG     0x400>>2
            DEFB    'V0.001  '      ; firmware version
            DEFB    'ORL/DEC '      ; manufacturer
            DEFB    'YES-CBG '      ; Option Name
            DEFB    'YES0'          ; firmware type

;  Start of modules

            DEFB    0,0,0,0,0,0,0,0 ; none present
```

Figure 2: Example option rom contents.

link. If present the HEC is formed from the CCITT CRC polynomial $x^8 + x^2 + x + 1$, initialised to zero and exclusive-ored with 01010101 before insertion.

The start of a cell that includes the HEC is indicated by transmission of a TAXI TT command, the 53 bytes of the cell will follow this. A cell that does not contain a HEC is started with a TAXI SR character which is followed by 52 data bytes. Thus reciever can determine whether to check for the HEC based on the start mark, and receivers which only accept one form can drop cells of the other type. At least one TAXI Sync character should be inserted between cells[2]. Note that the TAXIchips cannot be told to transmit a Sync character, instead a cycle with no transmission should be used.

7.2 TAXI command characters

The following TAXIchip command characters are used:

Name	TAXI Mnemonic	TAXI code	Use
Sync	JK	0000	Pad when no transmission, line synchronisation
Xon	II	0001	continue transmission (see text)
Soc	TT	0010	Start of Cell with HEC - precedes 53 data octets
Xoff	TS	0011	Stop transmission (see text)
Held	TR	0101	Transmission has been stopped (see text)
Sno	SR	0110	Start of Cell with no HEC - precedes 52 octets

With the exception of Soc/Sno, commands may be sent at any time. Receiving hardware should be able to continue assembling a cell even if the data component is interupted by a command. This is useful for several reasons. Firstly it enables data to be injected onto the link at a slower speed than the TAXIchip is being clocked – the chip will automatically insert Sync commands between data items in this case. Secondly it enables the handshake commands (which apply to the other side of the link) to be sent in a timely manner even if the sender is also transmitting data.

The handshake commands Xoff and Xon are generated by a receiving node to request the transmitter to stop or continue sending. Since the link distance is short this simple protocol can be used, for links with delays of several cells in the transmission media other techniques would be required. The Xoff command requests that the transmitter stops sending at the end of the current cell or before. Thus a transmitter may choose to respond to an Xoff by stopping immediately, but the cell seems the natural unit to use for this activity. A transmitter which has been stopped may continue to send all commands except Soc/Sno. When a transmitter is stopped it should periodically send Held commands to prevent loss of an Xon from stalling the system. A receiver should send the Xon command when it has sent an Xoff and is now prepared to receive another cell, or if it sees a Held command and is prepared to receive a cell. To allow direct working with the Sun/Xerox proposal the Xon/Xoff handshaking could be made advisory – thus Xoff would be a suggestion that the transmitter stop to avoid data loss, but it could be ignored by simple implementations.

The TAXI use of 4B/5B and NRZI encoding ensures that the transmitted data will always have the same average ammount of high and low times, so the DC balance of the line is ensured. The TAXI command set has this property when averaged over the transmission of a single character. However some of the command characters have few transitions, so their use is not recomended – these characters have been avoided when selecting the codes used above.

[2] This is not needed by our hardware, but maintains compatibility with the Sun/Xerox proposals.