LHCb DAQ 2003-021 Rev.2

GiGabit Ethernet mezzanines

for DAQ and Trigger links of LHCb

LHCb Technical Note

Issue: 1.0
Revision: 2

Reference: LHCb 2003-21
Created: 17 March 2003
Last modified: 28 April 2003

Prepared By: Hans Muller, Francois Bal , Angel Guirao

Abstract

The LHCb experiment uses IEEE 802.ab (GiGabit ethernet) technology for the data transport and readout network layer between the L1 data buffers[1] and the CPU farm. The common way to interface the L1 buffers in LHCb to the physical layers of GiGabitEthernet (GBE) is specified in this document. The SPI-3 industry standard is used as a narrow, FiFo-like interface which can be easily implemented in the FPGAs of the L1 boards. The SPI-3 protocol on the mezzanine cards is handled by commercial chips. Up to 4 bidirectional GBE channels can be implemented on a 149 * 74 mm mezzanine of stackheight 11 mm. These are interfaced through one 150 pin, high-speed connector which is defined here as LHCb standard.

Document Status Sheet

Table 1 Document Status Sheet

1. Document Title: GiGabit Ethernet mezzanines			
2. Document Reference Number: LHCb DAQ 2003-021 Rev.2			
3. Issue	**4. Revision**	**5. Date**	**6. Reason for change**
0	1	24.03.2003 H.M	Motherboard connector and dimensions added
1	1	4.04.2003 H.M	Complete Revision of first distributed draft. Intro, Terminology, SPI- Interface details, new pinouts with routed example, chapter on mezzanine features, References
1	2	28.04.2002 H.M.	Acrobat bookmark Index corrected, Rx Block 2 pinout correction F.Bal , Interface to PLX9030 by A.Guirao

1. formerly called ODE for Off Detector Electronics

Introduction

The LHCb experiment uses IEEE 802.3ab (GiGabit ethernet) for the data transport and for the readout network layer between the L1 data buffers[1] and the CPU farm. The common way to interface the LHCb Level-1 buffers to the physical layers of GiGabitEthernet (GBE) is specified in this document. Since the L1 buffer controllers in LHCb are implemented in (programmable) hardware, the interface should be as simple and "narrow" as possible. Like Slink, a FiFo-like I/O scheme has been chosen: the SPI-3[2] industry standard for OC48c industry devices. SPI-3 is specified for receiving and transmitting up to 2.488 Gbit/s over single or multiple GiGabit links per Medium access Controller (MAC). Using such a MAC chip on a mezzanine card, up to 32 bit @ 104 MHz can be transferred in both directions between an FPGA ("serving as link layer device ") and the PHYsical GBE links. Rather than implementing separate mezzanines for both Data Acqusition and for Trigger channels, SPI-3's "inband address selection" mechanism allows using a single connector for transmitting or receiving over several links per mezzanine card.

SPI-3 is a FiFo-like protocol with low pincount (~50) for each of the FiFo directions Rx or Tx, of which only one direction may be used. Refitted on a standard LHCb connector, three transfer types between GBE links and FPGA-embedded logic are possible: (Tx), (Tx + Rx), (Tx + Tx). Since Rx is available by default on a (Tx) card, there are 2 types of mezzaine cards: type A = (Tx +Rx) and type B = (Tx + Tx). Type A is referred here as LHCb Reference Design for 2 channels.

The LHCb connector for GBE mezzanines consists of 3 blocks (50 pins each) for Rx, Tx and Control. The control interface is a generic 16 bit bus for configuring control and status registers of the MAC. A commercial interface chip with SPI-3 interface is used on the Reference Design, allowing the use of two PHYsical GBE channels with the option for twisted pair cables or for optical links. Using the same approach, a 4-channel type B mezzanine can be implemented.

This document uses SPI-3 and POS-PHY-Level3 terminology wherever un-avoidable. Care has been taken to replace abstract terms (like "higher layer link device") by specific terms (like FPGA) which apply for the case of LHCb.

1. formerly called ODE for Off Detector Electronics
2. SPI-3 = System Packet Level Interface level 3, also referred to as POS-PHY-level3

Terminology

- **GBE**: Gigabit ethernet = IEEE 802.3
- **MAC**: Medium Access Control, hardware layer between SPI-X and Physical devices
- **EGMAC**: Enhanced Gigabit Media Access Control
- **Tx**: Transmit to the link or egress
- **Rx**: Receive from the link or ingres
- **SERDES**: Serializer / Deserializer logic
- **SFI**: SERDES Framer Interface
- **GMII**: Gigabit Media Independent Interface
- **LVTTL**: Low voltage TTL compliant signalling, JEDEC standard 8.1
- **SPI-3**: OC-48 / STM-16 up to 2.48832 Gbps
- **SPI-4**: OC-192 / STM-64 up to 10 Gbps
- **SPI-5**: OC-768 / STM-256 up to 40Gbps
- **OC**: Optical Carrier (SONET terminology)
- **STM:** Synchronous Transport Module (SDH terminology)
- **IFG:** Inter-Frame Gap
- **OIF**: Optical Interconnect Forum
- **POS-PHY:** SATURN compatible Packet over SONET interface up to 2.488 Gbit / s.

1 Link requirements in LHCb

The level-1 boards [1] of LHCb store subevents in Level-1 buffers until a Level-1 Trigger decision is recieved [2] upon which they transmit to their downstream Readout Units (RU) for Level2 / Level3 processing. For some detectors, also unbuffered level-1 trigger data are transmitted from the same level-1 boards. Both links are implemented in GBE technology:

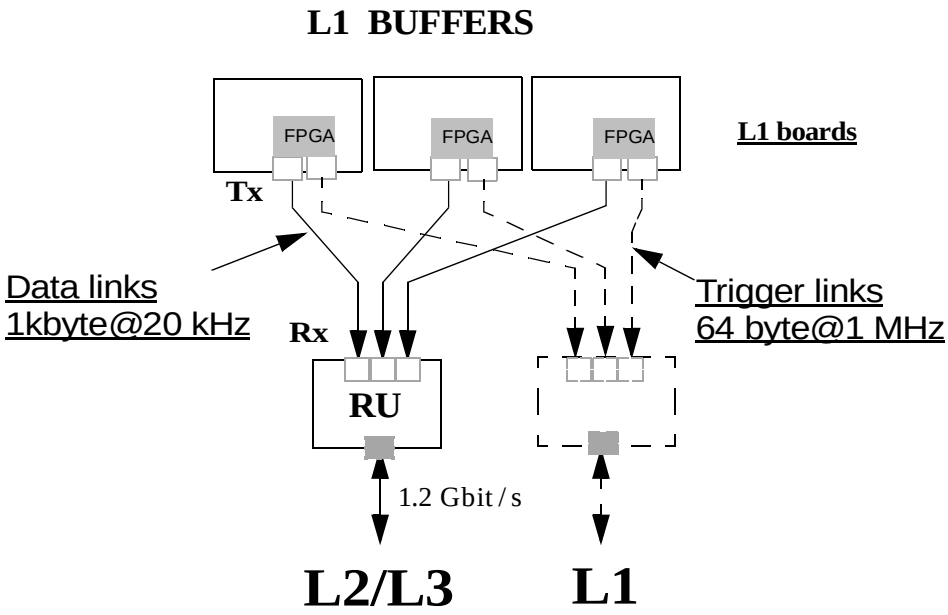


Figure 1 data and trigger links in LHCb

The GBE transmitters (Tx) are designed such that they can transmit two types of data:

a. Detector data which are stored during the Level-1 trigger latency and which are to be transmitted to the L2/L3 trigger. The nominal link output (fragment size @ output-rate$_{L1}$) is: 1 kbyte @ 20 kHz i.e. 20 Mbyte / s[1].

b. Trigger data are instantaneously transferred to the L1 trigger processors. The nominal output is: 64 byte$_{min}$ @ 1 MHz i.e 64 Mbyte / s[2]

Whilst the detector link payload is convenient for GBE, trigger data payloads may reach the transport limit of GBE (less than 100 Mbyte / s), hence for detectors with more than 64 byte average trigger payload it is required to use two or more GBE channels in parallel.

1. For non-zero suppressed detector data runs one would increase fragment size and reduce trigger rate proportionally
2. 64 byte$_{min}$ is the minimum payload due to the padding of GBE links.

1.1 Physical implementation

A common approach to interface one or more GBE links to programmable hardware on the level-1 electronics should be conceptually simple, not excessive in pincount and nevertheless capable of transmitting a maximum of data and trigger payloads, distributed over 1..4 links.

The SPI-3 industry standard is a bi-directional FiFo like standard which explicitly allows using only one direction (Tx or Rx) with a pincount of less than 50 pins for each. It was chosen since ASICs for the Medium Access Control (MAC) to GBE are commercially available as well as SPI-3 IP compatible cores for programmable logic.

The need for future upgrade (different physical links, higher bandwidth etc.) has favored a mezzanine implementation with an LHCb-specific connector with 3 blocks for the 3 signal types: Rx, Tx and Control.

Interfaced via the standard LHCb connector, two variants in the implementation are shown in Figure 2.: type A.) represents the reference design and type B.) represents a future 4-channel design, shown here as an example with optical links.

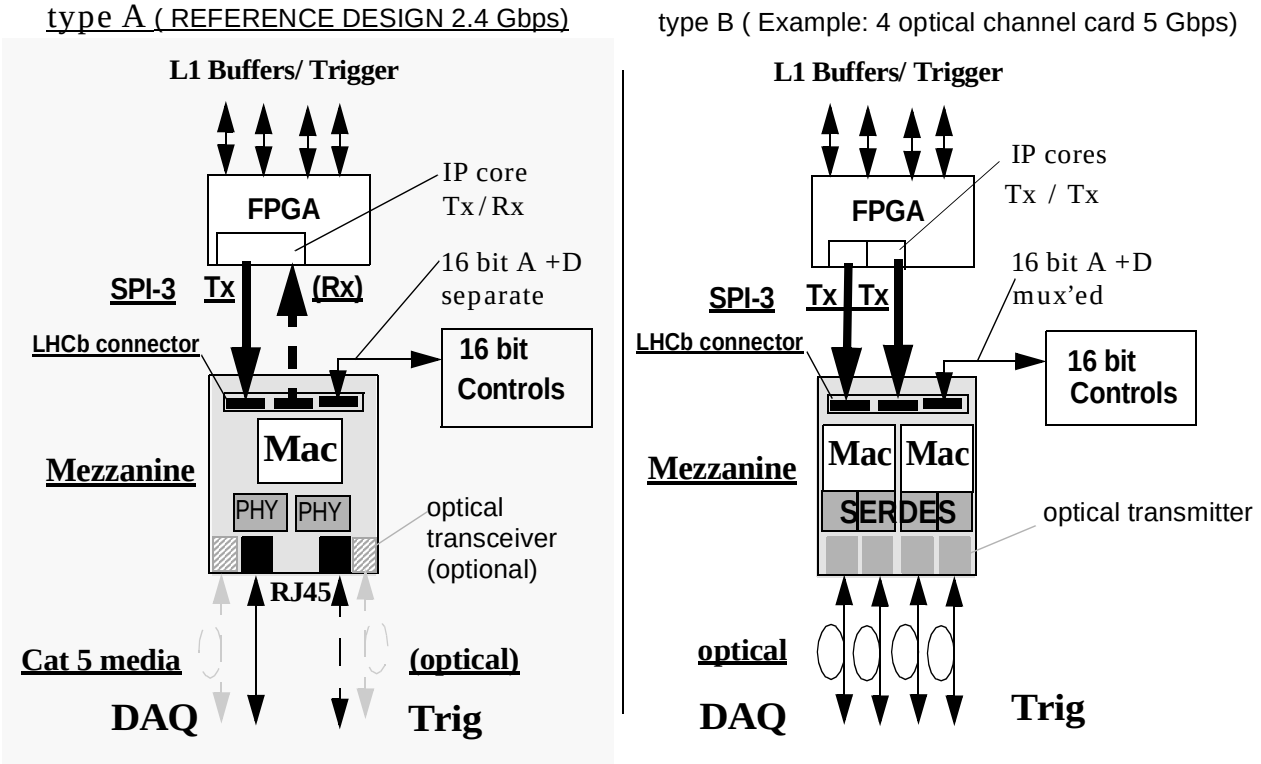


Figure 2 two variants of mezzanine implementation using the LHCb connector

The Reference Design is a 2-channel Rx / Tx mezzanine, based on one dual MAC and one dual PHY chip with integrated SERDES option for optical fibers (see Figure 2 left side).

A type-B (4-channel) mezzanine card with up to 5 Gbps Tx capability would require two dual MAC devices. Shown in see Figure 2 right side is an implementation example with 4 optical transmit links driven by the MAC-internal SERDES (hence no PHY chips are needed).

1.2 LHCb interface for GBE

The LHCb approach to interface copper or fiber via standard industry ASICS is based on the

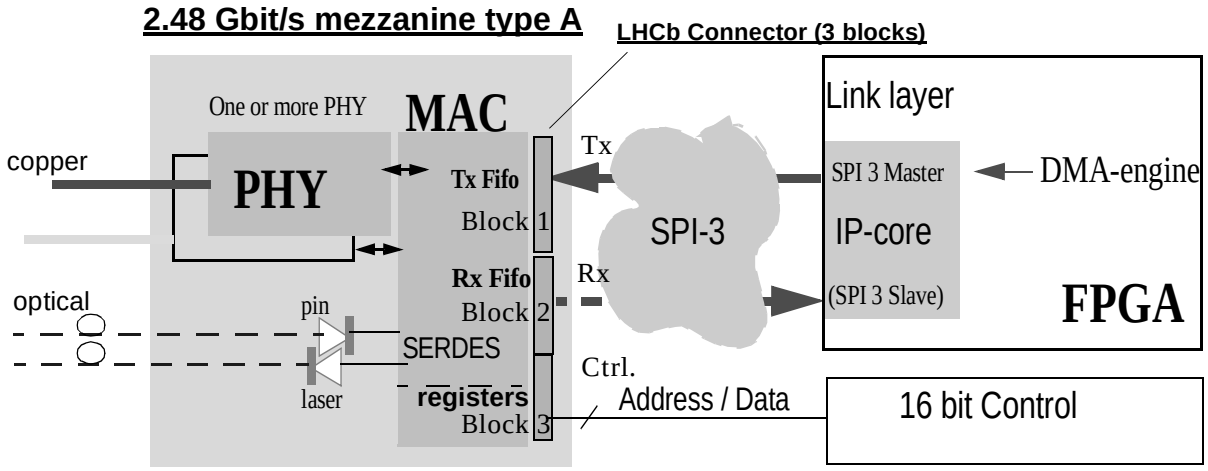


Figure 3 LHCb interfacing to a GBE mezzanine of type A (dual)

definition of the LHCb connector which carries three blocks of protocols: Tx, Rx and Control. These blocks can be used for both type A or type B mezzanines.

The interface principle is shown in Figure 3 with Rx/ Tx protocols for a type A (Tx/ Rx) mezzanine. The SPI-3 industry protocol between the FPGA and the MAC / PHY devices is independent for Rx and Tx. As shown in this example, a type A mezzanine may be operated (and connected) only in the Tx direction.

A mezzanine of type B.) requires that two Tx master protocols are implemented in the FPGA

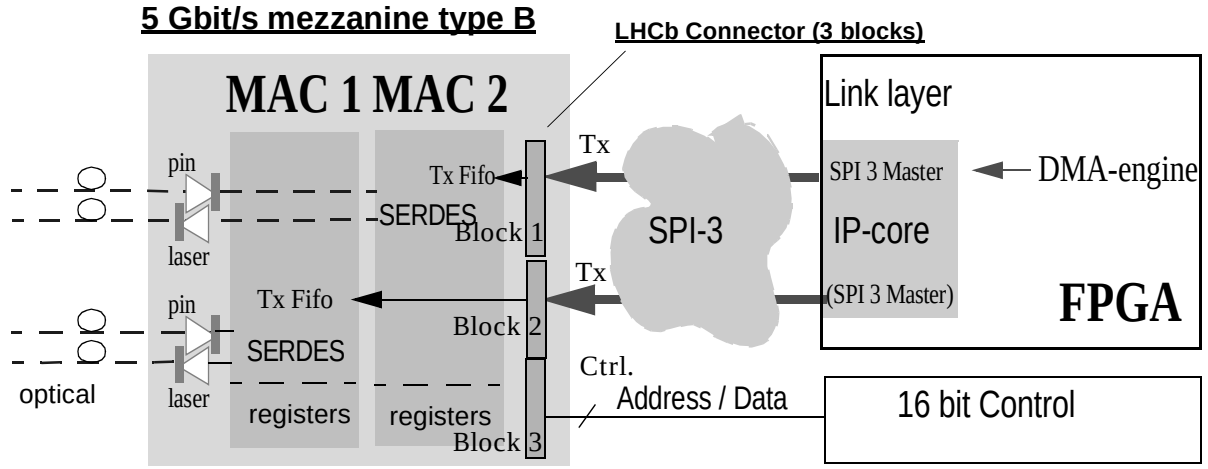


Figure 4 LHCb interfacing to GBE mezzanine of type B (quad)

and that the MAC registers can be addressed separately by the 16 bit control host.

1.3 SPI-3 industry standard

The POS-PHY-Level3 (“Packet over Sonet” = POS) interface was developed at the SATURN® Development Group in 1998. It has since been standardized at the ATM Forum (AF-PHY-0143.000) in March 2000 and at the Optical Internetworking Forum (OIF2000.008.3) in June 2000.

The original industry standard POS-PHY-Level3 is a byte-oriented interface to physical link devices. Sonet is the standard for Synchronous Optical Networks for the Industry where class OC-48 means 2.488 Gbps.There are several versions of POS-PHY level-3 for 622 Mbit/s and 2.48 Gigabit/s.

The Optical Interconnect Forum OIF re-defines POS-PHY-Level3 as **SPI-3 industry standard [5].**

There are several companies which produce chips with the SPI-3 interface, for example the quad channel Intel Chip IXF1104, or the dual channel PMC Sierra 3386 [3] of which the latter was chosen for the LHCb reference design of a type A mezzanine.

SPI-3 defines a Tx and Rx protocol for multi-PHY devices, however Tx or Rx may be used totally independently.

SPI-3 signals are transmitted at nominally 100 MHz over point-point transmission lines of 50 OHM impedance, usually based on LVTTL[1] signalling. FPGAs or ASICs with 3.3V LVTTL drivers and 2.5V receivers can be used. The PMC Sierra dual MACcontroller 3386 has 3.3 Volt drivers which require the use of an external serial resistors for the output drivers. FPGAs like Altera Stratix have 50 OHM compatible LVTTL drivers of 3.3 Volt.

SPI-3 is byte-oriented and depending on the bus width, there are 2 choices[2] for TDAT[31:0],

- 8 bit @ 104 MHz for 622 Mbit/s
- 32 bit @ 104 MHz for 2.488 Gbit/s

In the LHCb Reference Design, which is based on the PMC Sierra chip, only the 32 bit version is used.

The SPI-3 standard supports variable size packets up to 9.6 kbyte (jumbo frames). The FiFo-Fill granularity for Rx or Tx is byte-based. Since SPI-3 supports multiple PHY devices, the PHY selection is address-based and selected via the “inband address selection” protocol of SPI-3 (see Chapt. 5 . for details) The FiFo-Fill levels for hardware Receive and Transmit actions are programmable via the embedded registers.

1. LVTTL 3.3Volt is defined by EIA/JEDEC Standard (JESD8-B)
2. 104 MHz is the maximum frequency for SPI-3, 60 MHz is the minimum ! for the PM 3386 chip , 75 is the frequency at which one channel operates at full 1.2 GBps

2 Dual channel, Rx/Tx reference design (type A)

As a reference design, CERN ED Group[1] specifies and builds a first (type A) mezzanine with 2 GBE channels and an optical/copper option.

It is based on the PMC Sierra chip PM3386 [3] with SPI-3 interface [5], which is grouped into 3 blocks on the LHCb connector : Rx, Tx and Control (Figure 5).

SAMTEC's [6] high speed connector pairs QSS-075-01-F-D-A-xx and QTS-075-03-F-D-A-XX for 11 mm stackheight (see chapt 3) are used and defined for a common LHCb pinout (see Chapt 7 .)

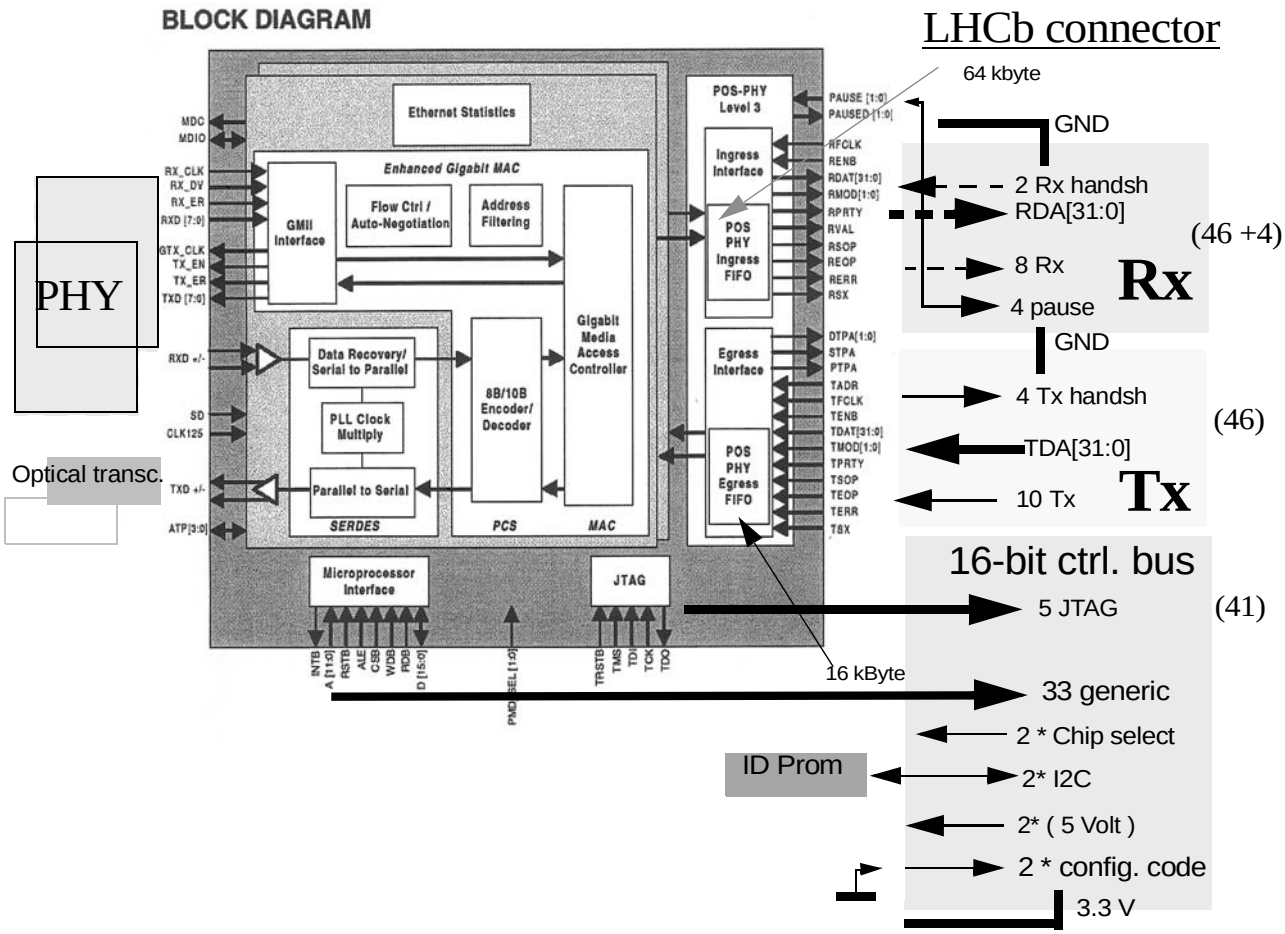


Figure 5 Use of LHCb connector with PMC-Sierra 3386 in a type-A reference design

1. http://web.cern.ch/ep-div-ed/

2.1 PMC-Sierra Interface chip 3386

The PMC-Sierra chip PM3386 was selected for the LHCb Reference design since it is relatively low in power consumption (< 2 Watt) and avalable in a very flat package, ideal for use on a mezzanine. It provides dual standard IEEE 802.3 Gigabit Ethernet MACs for frame verification and is designed for 2.4 Gbit/s OC-48 bandwidth. It's generic POS-PHY Level 3 (= SPI-3) interface provides an Rx/Tx interface to either two external PHY devices (copper) or two integrated SERDES devices (fiber).

The PM3386 chip includes a two port full-duplex Gigabit Ethernet Controller and provides direct connection to optics via dual SERDES, compatible to IEEE 802.3 1998 PMA physical layer specification. The connection to copper Gigabit Ethernet physical layer devices is via two IEEE 802.3 -1998 GMII/TBI interfaces. The PM 3386 provides on-chip data recovery and clock synthesis.

The internal FiFos are 16 kbyte Tx and 64 kbyte Rx

A 16 bit generic microprocessor interface is available for device initialization, control, register and per port statistics access. There is also a standard 5 signal P1149.1 JTAG test port for boundary scan board test purposes.

Other features:

- Verification of frame integrity (FCS and length checks).
- Errored frames can be filtered or passed to a higher layer device.
- Automatic Base Page Autonegotiation, extended Autonegotiation (Next Page) supported via host.
- Egress Ethernet frame encapsulation (pad to minimum size, add preamble, IFG and CRC generation).
- Supports Ethernet 2.0, IEEE 802.3 LLC and IEEE 802.3 SNAP/LLC encoding formats, and VLAN tagged frames.
- Minimum frame size 64 bytes. Supports jumbo frames up to 9.6 kbytes.
- Supports big endian data formats.
- Programmable inter-packet gap (IPG).
- Loopback for diagnostic capability through GMAC.
- Flow Control
- Option to support IEEE 802.3-1998 flow control at each Ethernet port.
- Programmable watermarks for full/empty FIFO conditions.
- Automatic generation of pause frames based on FIFO fill levels.
- Upper layer device can flow control Ethernet ports using side-band or host signaling to cause generation of a Pause frame.
- Provides per-port side-band Pause state indication for upstream devices.
- Loss-less flow control on all valid frames up to 9.6 kbytes.

2.1.1 Statistics registers of the PM3386

- 40 bit counters are used to ensure rollover compliance with IEEE 802.3-1998.
- Minimum 58 minutes before rollover.
- Provides statistic counters to support SNMP and RMON implementations.

2.1.2 Power and Package[1]

- Packaged in a 352-pin UBGA.
- Package height 1.8 mm
- Low power 1.8 V CMOS technology with 3.3 V compatible I/O.
- Industrial temperature range (-40 °C to +85 °C).
- 1.3 Watt in GMII mode (PHY)
- 1.9 Watt in SERDES mode (optical)

2.2 Serializer-Deserializer (SERDES)

The PM3386 has two internal serializer-deserializer transceivers. The SERDES are IEEE 802.3-1998 Gigabit Ethernet compatible supporting gigabit data transfer flows. The SERDES is based on the X3T11 10 Bit specification. The PM3386 receives and transmits Gigabit Ethernet streams using a bit serial interface for direct connection to optical transceiver devices. The SERDES performs data recovery and serial to parallel conversion for connection to the Enhanced Gigabit Media Access Control block.

2.3 Gigabit Media Independent Interface (GMII)

For Gigabit Ethernet over copper support, the PM3386 provides dual standard GMII interfaces. A copper Gigabit Ethernet physical layer device can be connected to the PM3386 via this interface.

1. note: PMC-SIERRA 3386 chip requires RC decoupling filters using exclusively ceramic capactors

2.4 Enhanced Gigabit Media Access Control (EGMAC)

The Enhanced Gigabit Media Access Control (EGMAC) block provides an integrated IEEE 802.3-1998 Gigabit Ethernet Media Access Control (MAC) supporting high performance 1000Base capability. The EGMAC has line side interfaces for connection to internal (SERDES) and external Gigabit PHY via GMII on each Gigabit Ethernet port. The Enhanced Gigabit MAC (EGMAC) incorporates all of the Gigabit Ethernet MAC functions including Auto-Negotiation, statistics, and the MAC Control Sub-layer that adheres to IEEE 802.3-1998 providing support for PAUSE control frames. The EGMAC provides basic frame integrity checks to validate incoming frames. The EGMAC also provides simple line rate ingress address filtering support via 8 exact-match MAC address and VID unicast filters, one 64-bin hash-based multicast filter, and the ability to filter or accept matched frames on a per instance programmable fashion. All inquires for filtering are done at line rate with no system latency introduced for look up cycles.

2.5 Microprocessor Interface

The PM3386 supports a standard 16-bit microprocessor interface which can be used in a multiplexed fashion with both address and data being present on the board system bus or in a de-multiplexed fashion with the address and data on separate busses upon the system board

The interface is an asynchronous protocol which may require adapation on the main board to a synchronous microprocessor protocol like the PLX 9030.

Table 2 Microprocessor Interface signals

Name	Type	Function
CSB	Input	Active-low chip select The CSB signal is low during PM3386 register accesses. If CSB is not required (i.e., registers accesses are controlled using the RDB and WRB signals only), CSB must be connected tied low.
RDB	Input	Active-low read enable The RDB signal is low during PM3386 register read accesses. The PM3386 drives the D[15:0] bus with the contents of the addressed register while RDB and CSB are low.
WRB	Input	Active-low write strobe The WRB signal is low during a PM3386 register write accesses. The D[15:0] bus contents are clocked into the addressed register on the rising WRB edge while CSB is low.

Table 2 Microprocessor Interface signals

Name	Type	Function
D[15:0]	I/O	The bi-directional data bus D[15:0] is used during PM3386 register, read and write accesses.
A[10:0]	Input	Address bus A[10:0] selects specific registers during PM3386 register accesses.
ALE	Input int. pul-lup	Address latch enable ALE is active-high and latches the address bus A[10:0] when low. When ALE is high, the internal address latches are transparent. It allows the PM3386 to interface to a multiplexed address/data bus. ALE has an integral pull-up resistor.
INTB	Ouptput open Drain	Active-low interrupt INTB is set low when a PM3386 interrupt source is active and that source is unmasked. The PM3386 may be enabled to report many alarms or events via interrupts. INTB is tri-stated when the interrupt is acknowledged via an appropriate register access. INTB is an open drain output.

Figure 6 represents the PM3386 microprocessor interface during a multiplexed read access. The signal system bus is used to illustrate the use of a shared system bus that might be implemented on the system board. The host presents a valid read address at points A and F. This address is latched into the PM3386 on the falling edge of ALE at point G. The host then turns the bus control over to the PM3386 by asserting RDB at point I. At point J the PM3386 starts to drive the bus with invalid data. At point K valid data is presented to the D bus and the System Bus. Valid data will continue to be present on the D bus until the host removes the D bus control from the PM3386 by de-assertion of the RDB signal at point L. At point M the PM3386 no longer drives the D or System Bus.

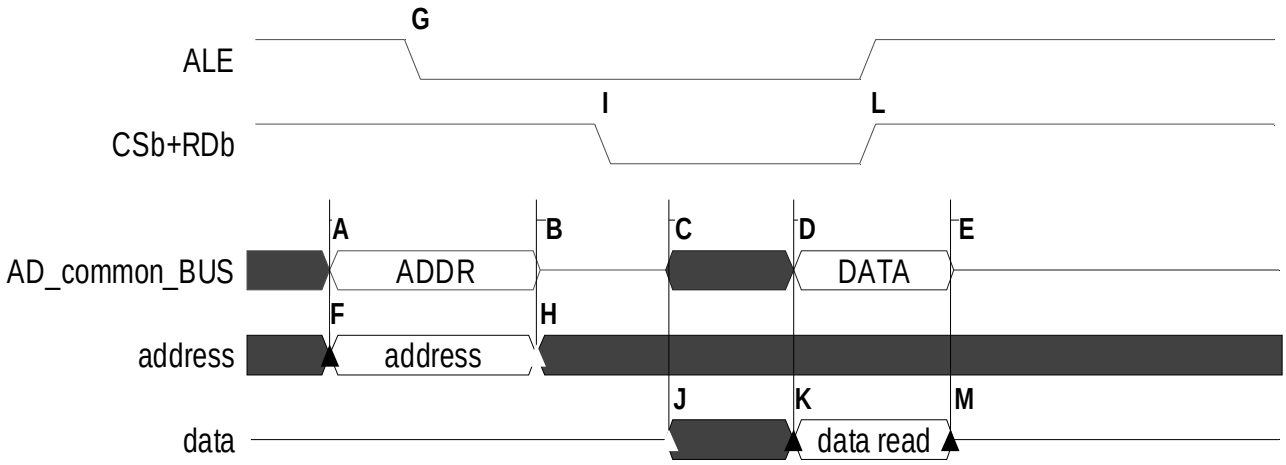


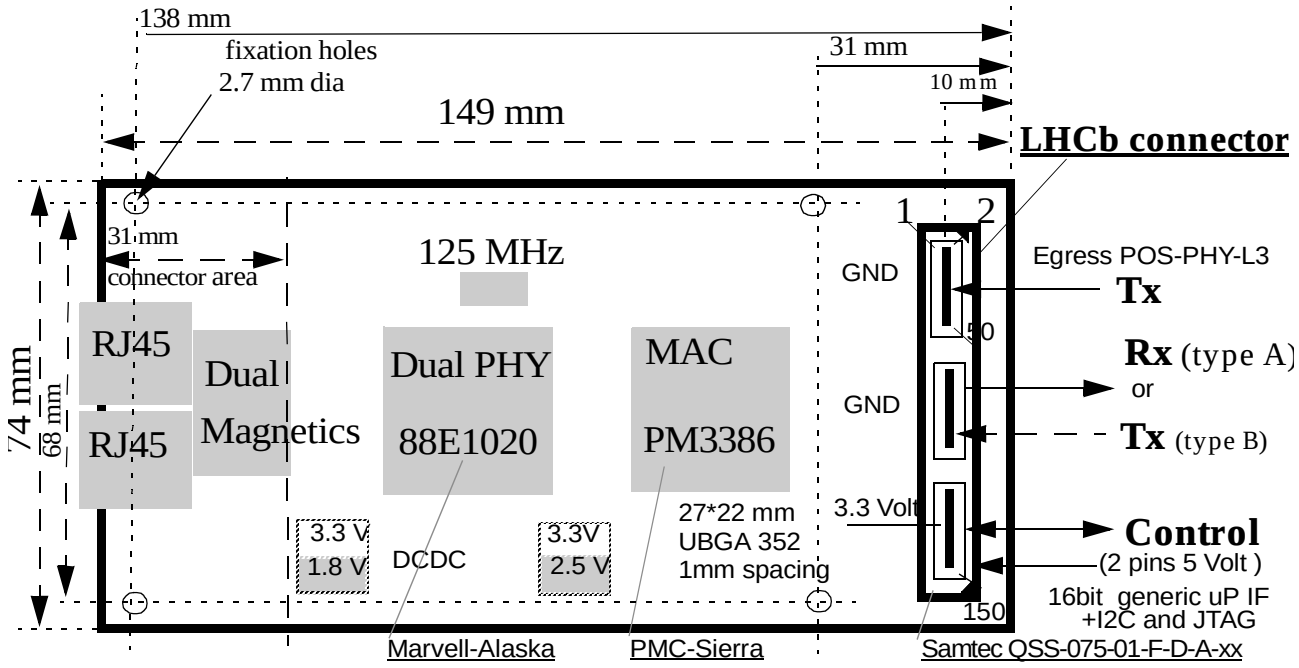
Figure 6 Microprocessor Multiplexed read

3 Mezzanine card outlines (type A and type B)

We use the same outlines 149*74 mm as the common PMC standard [4], however there is only one high speed 150 pin connector, containing 3 blocks of signals: Egress (Tx), Ingres (Rx) and Control. The stackheight (empty space beetween mezzanine card and motherboard) is 11 mm in order to allow the use of low profile[1] RJ45 connectors on the front panel.

3.1 Mezzanine card geometry (bottom view)

The geometry of the mezzanine (Reference design) is shown in Figure 7 as **bottom view**. The



Bottom view, on top only 1 mm components (decoupling + DC/DC) allowed

Figure 7 Mezzanine card outlines of the bottom view. Outlines are valid for both type A and B cards, Type B cards may use different components

PM3386 is as close as possible to the LHCb connector which contains 3 blocks of signals and each containing a power-bus bar. Two blocks contain the high-speed unidirectional SPI-3 lines for which 50 OHm impedance is important between motherboard, connector and the mezzanine. Hence these groups have the middle bus-bar connected to the GND plane. The middle connector block transmits slow control signals, hence it's bus bar is used for providing 3.3 Volt power from the motherboard to the mezzanine.

All other voltages (1.8 and 2.5 Volt) have to be generated on the mezzanine card.

1. low profile 11.56 mm

3.2 Mezzanine card geometry (side view)

The side view of the GBE mezzanine is shown with details for the VME crate environment

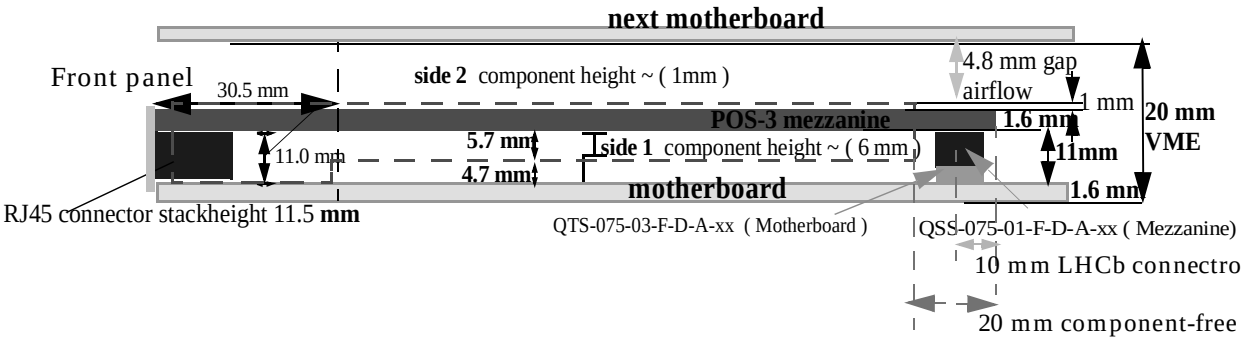


Figure 8 Component areas for GEB mezzanine

which has motherboards spaced every 20 mm. With the chosen stackheight of 11 mm, the space between the neighboring motherboards is divided in such a way that 4.8 mm free space for airflow remains and there is no touch with the next motherboard.

There are 2 sides of the mezzanine:

- **side 1** is the component side on the bottom (connector) side of the mezzanine. It is divided in 3 parts: a.) On the Front panel , VME defines a non-component zone of 30.5 mm on the motherboard. This space can therefore be used by the mezzanine for connectors up to 11 mm height. b.) on the LHCb connector side, there is an area of 20 mm where no components can be mounted. c.) The rest (middle) of side 1can be equipped with components of a max. height of 5.7 mm.
- **side 2** is the component side on the top (visible) side of the mezzanine. Here only components of 1mm height are allowed over the full area.This area is best used for decouplig and very flat DC regulator devices. Note also that screwheads muts not be hiher than 1 mm.

3.3 Mainboard geometry

The mainboard receives a GBE mezzanine in such a way that it's boarder is aligned with the frontpanel of the mezzanine. There are four fixation holes in the mainboard emplacement with 11 mm spacers (see Figure 7 for geometrical details).

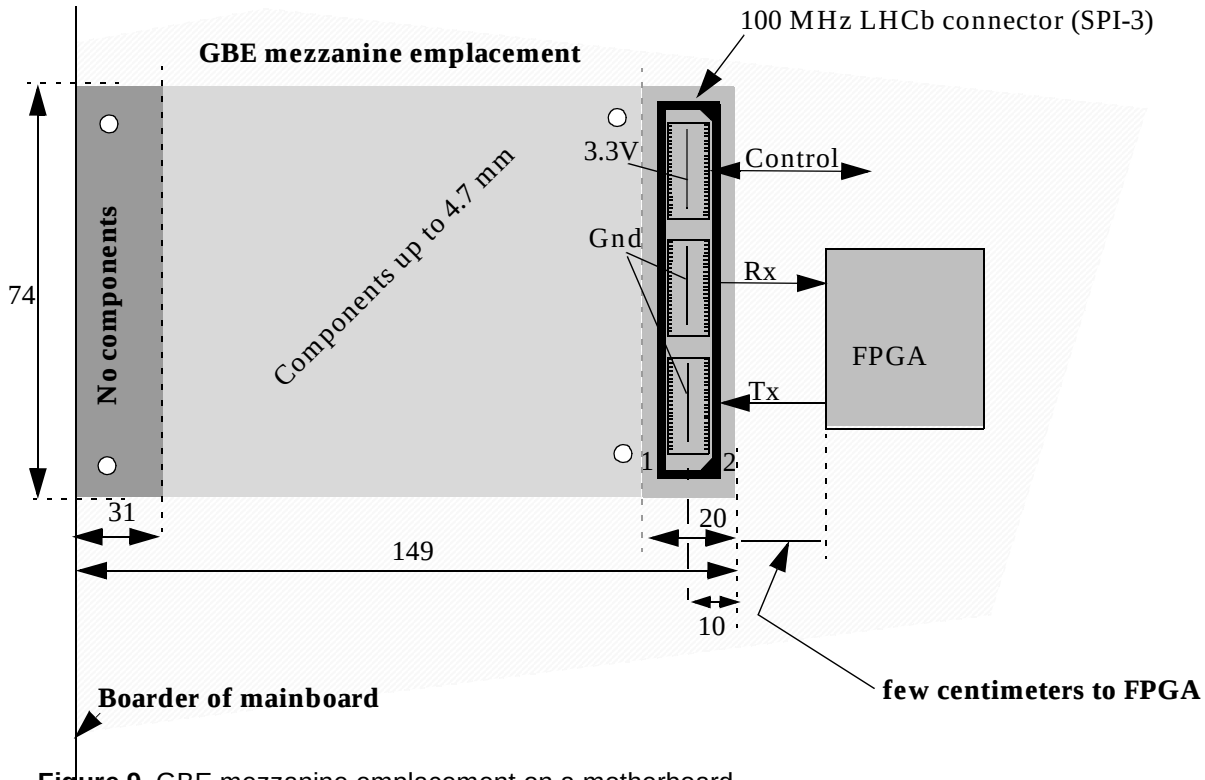
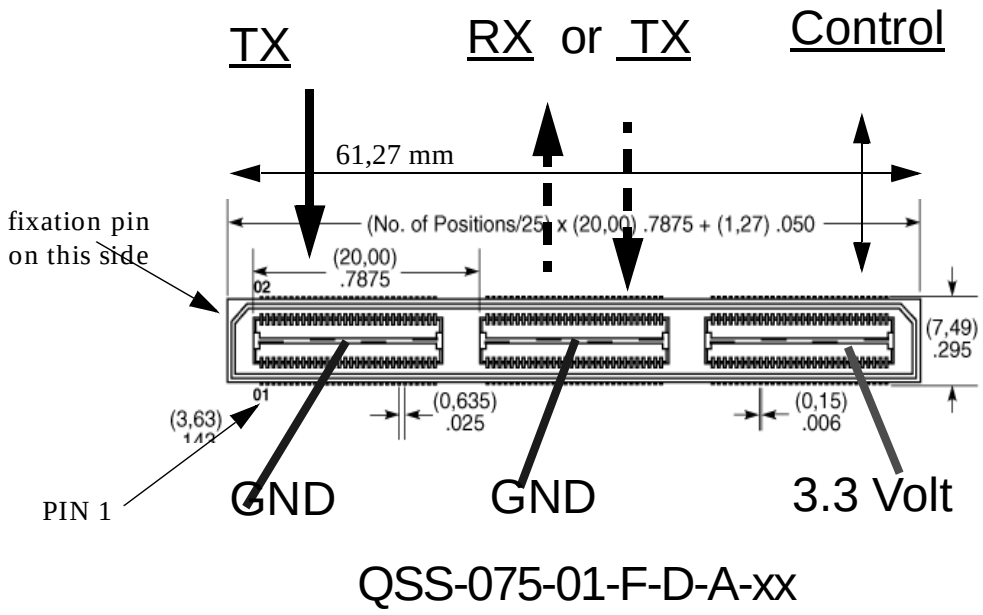


Figure 9 GBE mezzanine emplacement on a motherboard

4 LHCb connector

The SAMTEC[1] high speed connectors QSS privede a 50 OHM transfer impedance in connectors banks of 2* 25 positions centered by a zero-imedance bus bar. Signals beyond 600 MHz are transmitted at low loss. These surface mount connectors are therefore well suited for the SPI-3 signal path of 32 bit @ 104 MHz. Since we need 3 groups of up to 50 pins, a QSS connector with 0.635 mm spacing has a total length of 61.27 mm and hence fits on the 74 mm

Mezzanine: QSS



Note: units in ( ) are in mm

Figure 10 QSS surface mount mezzanine connector 3 * 50 pins and 3 bus bars

wide mezzanine.The QSS-075-01-F-D-A-xx mezzanine connector (Figure 10) mates with the

1. http://www.samtec.com

QTS connector of Figure 11 for a combined stackheight of 11 mm.

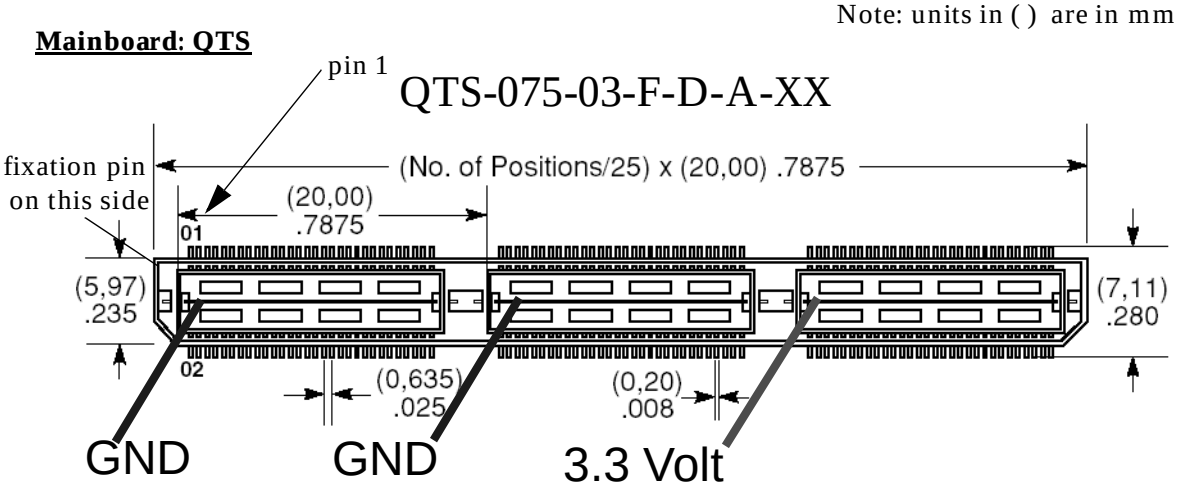


Figure 11 QTS surface mount mainboard connector of 11 mm stackheight (mating QSS)

4.1 Impedances

The SPI-3 signalling is a single-ende, serially terminated scheme, using LVTTL signalling in FPGAs and requireing a 50 ohm impedance between the FPGA pins and the SPI-3 chip. Hence both the motherboard, the connector and the mezzanine must provide a uniform 50 OHM impedance to the signals. FPGA's like Altera Stratix provide SPI-3 compatible 50 OHM

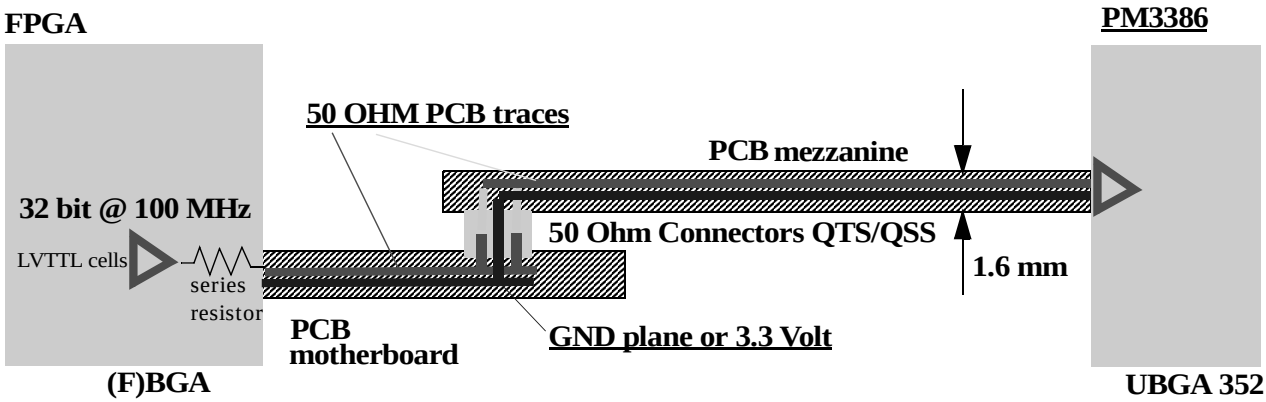


Figure 12 50 OHM transmission between motherboard and GBE mezzaninen

drivers. Note that the PMC Sierra 3386 requires external 33 OHM series resistors on all SPI-3 outputs.

5 Read and Write basics for SPI-3

This chapter describes the read / write basics for the SPI-3 interface using the PM3386 chip as example. For much more detail, refer to the SPI-3 specification [5] and to the PM3386 user manual [3].

Note that signals ending with a “B” are active low. All signals are fully synchronous with the TFCclock (60.. 104 MHz). The transmit interface of the PHY is selected using a so-called “**in-band address**” that is provided on the same bus transferring the packet data. An “inband” selection means that the FPGA selects first the PHY address on the TDAT[0] data bus line[1] via the TSX signal, and followes it by data on the same bus TDAT[31:0] by asserting the TENB signal (low).

5.1 Transmit interface operation (Egress)

There are two transmit modes: **byte-level** mode and **packet-level** mode of which one has to be selected.

In byte-level transfer, FIFO status information is presented on a cycle-by-cycle basis. With packet-level transfer, the FIFO status information applies to segments of data.

With packet level transfer, the FiFo is able to do status polling on the transmit direction. The FPGA can use the transmit port address TADR[] to poll individual PHY ports, which all respond on a common polled (PTPA) signal.

Since the variable size nature of packets does not allow any guarantee as to the number of bytes available, in both transmit and receive directions, a “Selected PHY Transmit Packet Available” is provided on signal STPA and a “Receive Data Valid on signal RVAL. STPA and RVAL always reflect the status of the selected PHY to or from which data is being transferred.

5.1.1 *Byte-level-mode*

When using byte level mode, direct status indication must be used. In this case, the PMC Sierra chip provides the transmit packet available status (STPA). It transitions high when a predefined minimum number of bytes are available in the transmit FIFO specified by the inband address on TDAT[0] (which designates the selected PHY channel). When STPA is high, it indicates the Transmit FiFo is not full. Also two DTPA lines provide direct access to the transmit packet available status of the PHY channels. If DTPA0 or DTPA1 goes high, a predefined number of bytes is free in the corresponding transmit FiFo. Hence both port's FIFO

1. For the PMC Sierra chip with 2 PHY devices, TDAT[0] designates their two addresses, 0 and 1

status'es are present on the DPTA and STPA lines. The available transmit FiFo gets selected by the FPGA via an "inband address" selection, followed by data. For this it places the selected PHY address on the TDAT[0] line and asserts the TSX signal, keeping TENB unasserted (high). This is interpreted as Command (cmd).

Figure 13 shows the example of PHY port 0 selection. In order to start the data transfer, the FPGA asserts the TENB low and drives the first word on the TDAT[31:0] data bus together with the Transmit Start of Packet signal TSOP (high).

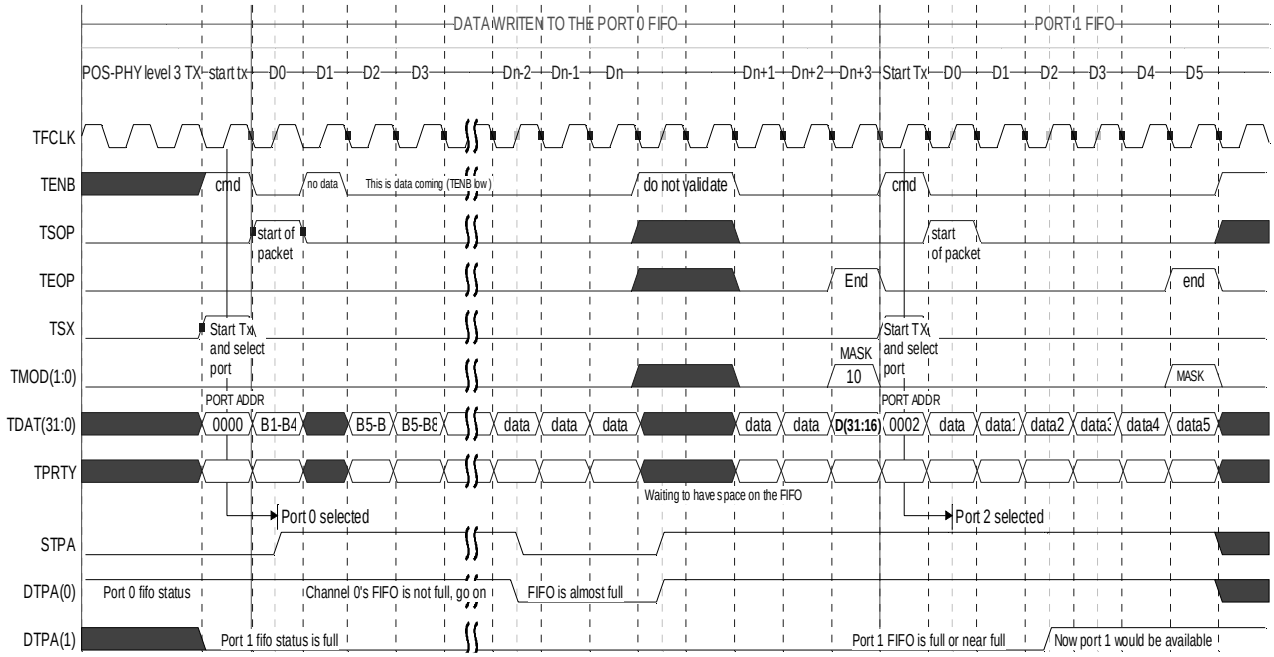


Figure 13 Basic Transmit operation

Data on the SPI-3 interface are valid per TFCLK state when the TENB enable signal is asserted low by the FPGA. At the same time the FPGA has to provide data parity on the Transmit bus Parity line (TPRTY). Odd or even parity is only provided as information and shall not interfere with the transferred data. Wait states are created by deasserting TENB (high), indicated as "no data" in the example. This is required when the FiFo gets full, as indicated on the DTPA0 line.

To complete the transmission, the FPGA asserts the Transmit End of Packet signal (TEOP) and masks out in the same clock state the invalid bytes on the last data word using TMOD(1:0).

If no other packet follows, the FPGA deasserts the enable line TENB (low).

The FPGA can either perform further operations to the same port if there is available enough space for the next packet, or monitor the status of the PHY's port FIFOs and switch to another PHY port if there is enough space in it's FIFO.

The FPGA can also initiate a new packet to the previous PHY port by asserting TSX at the same time as TENB with data on TDAT[31:0].

5.1.2 *Packet level mode*

In this mode, the transmitter polls the PHY ports for the FIFO status according to it's own order, i.e the PMC-Sierra chip selects a PHY port to request the status of its FIFO by driving on TADR the port address (on the PM3386, one line in state 0 or 1). The PHY device answers at the next rising edge of the transmit clock with the FIFO status on the Polled Transmit Packet Available signal (PTPA).

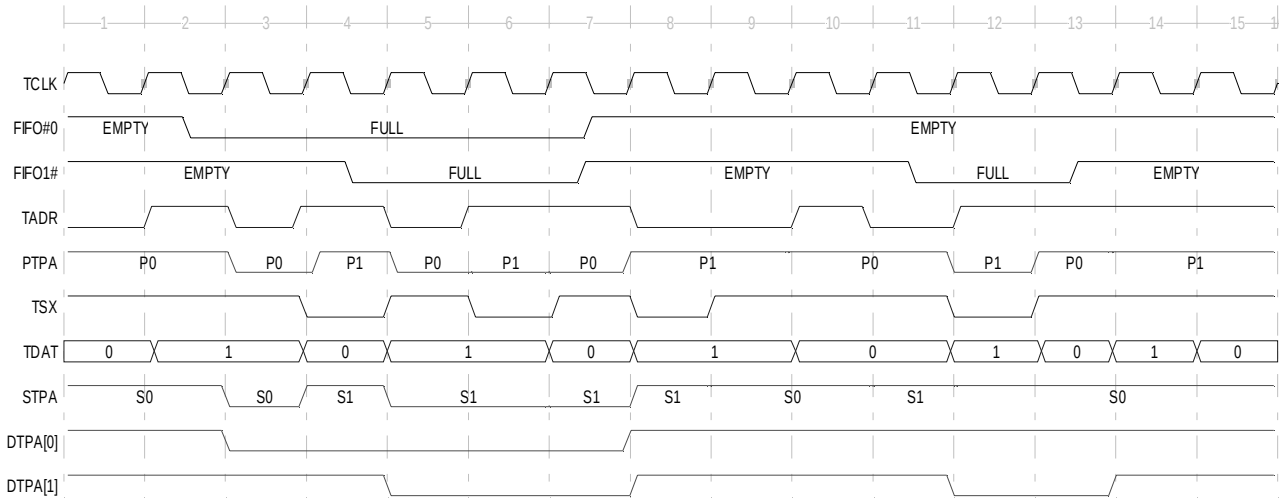


Figure 14 Transmit polling

Eventhough, since the variable size of packets does not allow any guarantee as to the number of bytes available in both transmit and receive directions , the selected PHY transmit packet available STPA should be used to monitor the current status of the port in use.

It appears that the packet level polling is more complicated and hence less recommended.

5.2 Receive Interface (Ingress)

RENB is an active low signal used to control the flow of read data from the PM3386. The FPGA may de-assert RENB at anytime if it is unable to accept data from the PM3386.

When RENB is sampled low by the PM3386, the FPGA is signaling that it can receive data. RSX may then be asserted to indicate a new address on the RDAT[0] bus pin or RVAL may be asserted indicating validity of read data and control on the RDAT[31:0], RPRTY, RMOD[1:0], RSOP, REOP, and RERR signals.

Note that these signals will be updated on the following rising edge of the RFCLK. When RENB is sampled high by the PM3386, the upper level device is signaling hat it can no longer accept data.

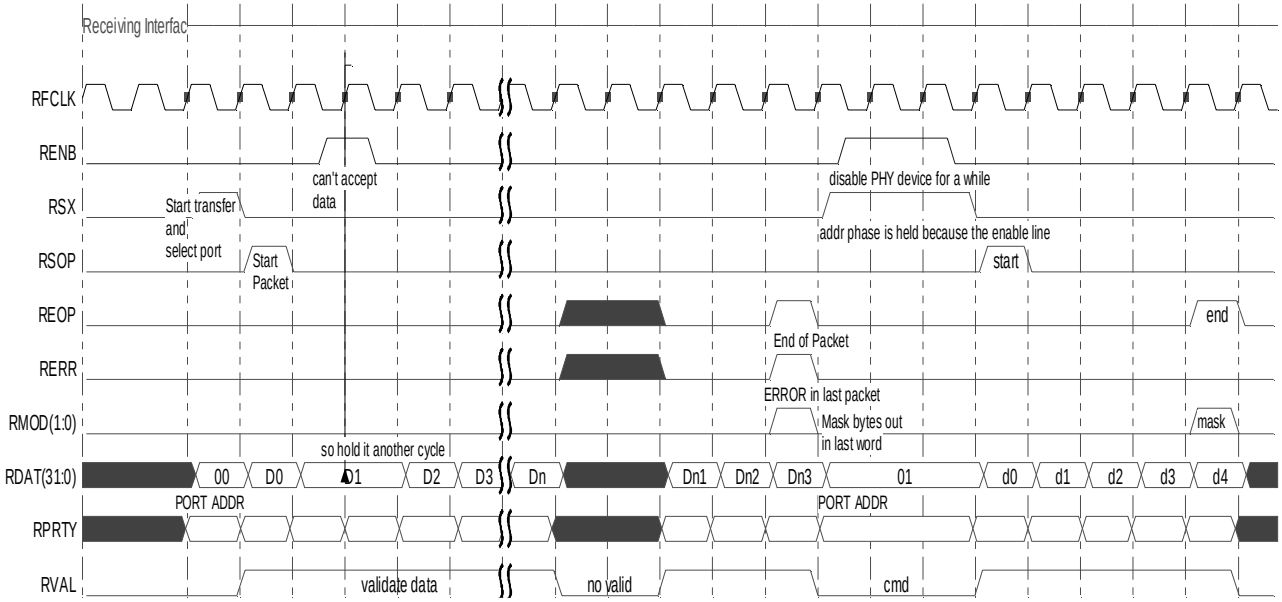


Figure 15 Basic Receive operation

6 GBE mezzanine card features

The two different mezzanine designs envisaged so far to be use with the LHCb connector are:

type A: 2 channel Rx / Tx with a maximum throughput of 2.488 GBps

type B: 4 channel Tx /Tx with a maximum throughput of 5 GBps

The 2 channel mezzanine is the Reference design using a single PMC sierra chip with dual channel Rx /Tx support. In order to allow the use of two PMC Sierra chips, the chip select CSB signals must be available twice on the LHCb connector .the CSB1 and CSB2 are active-low chip selects for the PM3386 register accesses. The CSB1 is chip select for the reference design (type A)

6.1 Card type coding

There are two coded pins Conf0, Conf1 in the Controls interface block which define the type of the card. Note that on the **main board**, both signals have to be pulled high via a resistor of nominally 1 kOHM

Table 3 Mezzanine card codes

Conf0	Conf1	
0	0	**Type A**: Tx / Rx mezzanine card (2 Channel Reference card) 2.4 GBps
0	1	**Type B**: Tx + Tx card (quad channel) 5 GBps
1	0	reserved
1	1	No mezzanine card

6.2 Revision number I2C Prom

The GBE type and revision number is stored in a serial Prom[1] on the GBE mezzanine card. The information can be read out via the I2C bus interface available on the Controls block.

The device is programmed via I2C preferably in the laboratory.

1. a suitable EEprom is the 24LC00 from Microchip

6.3 Interrupts

The PM3386 interrupts the host processor via the use of the INTB active low signal. When INTB is asserted the host processor can interrogate the PM3386 for the source of the interrupt by reading the Interrupt Status register. The resulting information will provide the programmer with the block from which the interrupt originated.

The INTB signals of the PM3386 are open drain interrupt outputs pulled up via a resitor on the GBE mezzanine. Mezzanine cards of type B with two PMC Sierra chip are deemed to connect the two signals INTB signals together.

6.4 Mezzanine card power supplies

The main power supplies is 3.3V which is connected through the bus bar of the Controls block of the LHCb connector. The Rx and Tx bloacks contain the bus-bars for GND.

Two additional +5Volt pins are provided on the controls block for a limited use of max. 1.5 A of mezzanine card devices.

The maximum total power consumption on 3.3V and 5 V is 7.5 Watt.

6.5 SERDES / PHY selection

The PMD_SEL_0 and PMD_SEL_1 signals provide the physical medium selection of the two PHY channels. These signals are not available on the LHCb connector and meant to be configured permanent on the mezzanine cards.

These active high signals select between using the on-board SERDES or external transceiver via the GMII pins. A low (tied to VSS) will select internal SERDES. A high (tied to VDDO) will select external transceiver via the GMII pins. These pins are required to be tied to VDDO or VSS prior to device power up.

6.6 Master Reset

Each PMC Sierra chip has an RST master Reset input (schmitt trigger with internal pull-down) Master Reset. An active low signal input provides an asynchronous reset to the device. When RSTB is forced low, all device registers are forced to their default states.

In case of the two PMC Sierra chips these will always be reset together via a single RST on the LHCb connector.

The RST signals is sourced normally from the FPGA on power-up and on specific request of tyhe slow control system.

RSTB has a **minimum reset pulse width of 1 ms**[1]. Prior to the de-assertion of RSTB the PMD_SEL pins[2] must be in a stable state (strapped high or low) and all clocks for the device are required to be present for a minimum of 1ms. Internally when the RSTB signal is de-asserted the analog portion of the device will start to lock on to the various reference clocks. The digital portion of the device will be held in reset for 10 ms more by an internal timer. System status of analog training and progress can be viewed via the top level Device Status register.

6.7 PMC Sierra Flow Control (Pause)

The Flow control on the PMC3386 is not part of the SPI-3 protocol. The corresponding pins (PAUSE[0:1] and Paused [0:1]) are nevetheless available on the LHCb connector's Rx block.

Flow control handled in the EGMAC block. When a PAUSE control frame is received, the PM3386 will optionally terminate transmission (after the current frame is sent) and assert the appropriate channel side band flow control output to indicate the paused condition.

Assertion of the PAUSE0 or PAUSE1 signals may cause (programmed option) the PM3386 on a per channel basis to transmit 802.3-1998 PAUSE frames and either drop at the MAC layer or pass to the POS-PHY L3 client any further incoming frames (programmed option). De-assertion of the PAUSE0 or PAUSE1 signal can cause the removal of the PAUSE condition on a per channel basis. Due to the programmability options for these pins please see below.

PAUSE0 and PAUSE1 are active high signals input signals to the 3386 with an internal pulldown resistor. PAUSE0 and PAUSE1 are sampled on the rising edge of the RFCLK.

The PAUSED0 and PAUSED1 output signals indicate the reception and execution of 802.3-1998 PAUSE control frames on the given port of the PM3386. 01:22:56 An asserted (high) PAUSED0 or PAUSED1 pin indicates that the corresponding channels ingress PAUSE timer is non-zero. This also typically indicates (if enabled via the FCRX bit in the EGMAC GMACC1-Config Register) that the given channel is in a paused state. De-assertion of the PAUSED0 or PAUSED1 pin indicates that the corresponding channels PAUSE counter is now zero. This also typically indicates that the given channel is no longer pausing on that channel. Please refer to the FCRX bit definition for more information.

PAUSED0 and PAUSED1 are updated on the rising edge of RFCLK.

1. This implies that if RST is driven by an embedded microprocessor, the RESET must be a programmed GPIO function of minim. 1 ms duration. An address-based Reset must be streched by FPGA logic to 1ms.
2. these are not on the LHCb connector, static pins for selection of PHY or SERDES

7 Pinout of LHCb connector

The pinout of the LHCb connector has been assigned (as function of the routing A.for the PMC-Sierra chip) for the 150 pin SAMTEC connectors QSS-075-01-F-D-A & QTS-075-03-F-D-A . The three 50-pin connector blocks are designated as **BLOCK1** (Tx only), **BLOCK2** (Rx or Tx) and **BLOCK3** (Control only). Type B card (quad Tx) must implement both block 1 and block 2 as Tx.

Table 4: LHCb connector: Tx (BLOCK1 and BLOCK2 with GND busbars)

Pin No Block 1 TypeA+B	Pin No Block 2 (type B)	I/O relative to PM3386	Signal	Pin Function
1	51	Input	**TDAT<30>**	POS-PHY Transmit Data bus (32 bits) When a 32-bit interface is used, data must be transmitted in big endian order on TDAT[31:0]. It is sampled on the rising edge of TFCLK.
2	52	Input	**TDAT<28>**	POS-PHY Transmit Data bus (32 bits)
3	53	Input	**TDAT<27>**	POS-PHY Transmit Data bus (32 bits)
4	54	Input	**TDAT<31>**	POS-PHY Transmit Data bus (32 bits)
5	55	Input	**TDAT<29>**	POS-PHY Transmit Data bus (32 bits)
6	56	Input	**TDAT<25>**	POS-PHY Transmit Data bus (32 bits)
7	57	Input	**TDAT<24>**	POS-PHY Transmit Data bus (32 bits)
8	58	Input	**TDAT<21>**	POS-PHY Transmit Data bus (32 bits)
9	59	Input	**TDAT<26>**	POS-PHY Transmit Data bus (32 bits)
10	60	Input	**TDAT<19>**	POS-PHY Transmit Data bus (32 bits)
11	61	Input	**TDAT<22>**	POS-PHY Transmit Data bus (32 bits)
12	62	Input	**TDAT<16>**	POS-PHY Transmit Data bus (32 bits)
13	63	Input	**TDAT<23>**	POS-PHY Transmit Data bus (32 bits)
14	63	Input	**TDAT<13>**	POS-PHY Transmit Data bus (32 bits)
15	65	Input	**TDAT<20>**	POS-PHY Transmit Data bus (32 bits)
16	66	Input	**TDAT<11>**	POS-PHY Transmit Data bus (32 bits)
17	67	Input	**TDAT<18>**	POS-PHY Transmit Data bus (32 bits)
18	68	Input	**TDAT<8>**	POS-PHY Transmit Data bus (32 bits)
19	69	Input	**TDAT<17>**	POS-PHY Transmit Data bus (32 bits)
20	70	Input	**TDAT<5>**	POS-PHY Transmit Data bus (32 bits)
21	71	Input	**TDAT<15>**	POS-PHY Transmit Data bus (32 bits)
22	72	Input	**TDAT<2>**	POS-PHY Transmit Data bus (32 bits)

Table 4: LHCb connector: Tx (BLOCK1 and BLOCK2 with GND busbars)

Pin No Block 1 TypeA+B	Pin No Block 2 (type B)	I/O relative to PM3386	Signal	Pin Function
23	73	Input	**TDAT<14>**	POS-PHY Transmit Data bus (32 bits)
24	74	Input	**TDAT<1>**	POS-PHY Transmit Data bus (32 bits)
25	75	Input	**TDAT<12>**	POS-PHY Transmit Data bus (32 bits)
26	76	Input	**TENB**	POS-PHY Transmit Write Enable. When TENB is high, the TDAT[31:0], TMOD, TSOP, TEOP, TPRTY and TERR signals are invalid and are ignored . However, the TSX signal if asserted is valid and is processed only when TENB is high. When TENB is low, the TDAT[31:0], TMOD, TSOP, TEOP, TPRTY and TERR signals are valid and are processed. The TSX signal is ignored when TENB is low. TENB is sampled on the rising edge of TFCLK.
27	77	Input	**TDAT<10>**	POS-PHY Transmit Data bus (32 bits)
28	78	Input	**TEOP**	POS-PHY Transmit End Of Packet. TEOP is required to be present at the end of every packet and is considered valid only when TENB is asserted.
29	79	Input	**TDAT<9>**	POS-PHY Transmit Data bus (32 bits)
30	80	Input	**TMOD1**	POS-PHY Transmit Modulo. See TMOD0
31	81	Input	**TDAT<7>**	POS-PHY Transmit Data bus (32 bits)
32	82	Input	**TMOD0**	POS-PHY Transmit Modulo. TMOD[1:0] indicates the number of valid bytes of data in TDAT[31:0]. The TMOD bus should always be all zero, except during the last double-word transfer of a packet on TDAT[31:0].
33	83	Input	**TDAT<6>**	POS-PHY Transmit Data bus (32 bits)
34	84	Input	**TERR**	POS-PHY Transmit Error Indicator. Active high signal used to indicate that the current packet must be aborted. TERR should only be considered valid when TENB and TEOP are simultaneously asserted.
35	85	Input	**TDAT<4>**	POS-PHY Transmit Data bus (32 bits)
36	86	Input	**TADR**	POS-PHY Transmit PHY Address. The TADR signal is used with the PTPA signal to poll the transmit FIFOs packet available status.
37	87	Input	**TDAT<3>**	POS-PHY Transmit Data bus (32 bits)
38	88	Output	**STPA**	POS-PHY Selected PHY Transmit Packet aVAILABLE. STPA transitions high when a predefined (user programmable) minimum number of bytes are available in the transmit FIFO specified by the in-band address on TDAT bus. Once high, STPA indicates the transmit FIFO is not full.
39	89	Input	**TDAT<0>**	POS-PHY Transmit Data bus (32 bits)

Table 4: LHCb connector: Tx (BLOCK1 and BLOCK2 with GND busbars)

Pin No Block 1 TypeA+B	Pin No Block 2 (type B)	I/O relative to PM3386	Signal	Pin Function
40	90	Output	**PTPA**	POS-PHY Polled-PHY Address. PTPA transitions high when a predefined (user programmable) minimum number of bytes are available in the polled transmit FIFO. Once high, PTPA indicates that the transmit FIFO is not full.
41	91	Schmitt Input	**TFCLK**	POS-PHY Transmit FIFO Write Clock. 60 MHz min, 104 MHz max. Duty cycle 40..60%
42	92	Output	**DTPA1**	POS-PHY Direct Transmit Packet Available (chan. 1). See DTPA0
43	93	Input	**TSX**	POS-PHY Transmit Start of Transfer. Active high signal indicating when the in-band port address is present on the TDAT[31:0] bus. When TSX is high and TENB is high (not asserted), the value of contained within TDAT[7:0] is the address of the transmit FIFO to be selected.
44	94	Output	**DTPA0**	POS-PHY Direct Transmit Packet Available (chan. 0). Active high signals that provide direct status indication for the corresponding ports in the PM3386. DTPA[1:0] transitions high when a predefined (user programmable) minimum number of byes are available in the transmit FIFO. Once high, the DTPA[1:0] signals indicate that its corresponding transmit FIFO is not full.
45	95	Input	**TSOP**	POS-PHY Transmit Start Of Packet. TSOP is required to be present at the beginning of every packet and is considered valid only when TENB is asserted.
46	96	Input (pull-down)	**PAUSE1**	PAUSE control (ch.1) See PAUSE0
47	97	Input	**TPRTY**	POS-PHY Transmit Bus Parity. The transmit parity (TPRTY) signal indicates the parity calculated over the TDAT bus. TPRTY is considered valid only when TENB or TSX are asserted. (By default the PM3386 uses odd parity.)
48	98	Output	**PAUSED0**	PAUSE status (ch.0). The PAUSED0 and PAUSED1 signals indicate the reception and execution of PAUSE control frames on the given port of the PM3386.
49	99	Input (pull-down)	**PAUSE0**	PAUSE control (ch.0) . Assertion of the PAUSE0 or PAUSE1 signals may cause (programmed option) the PM3386 on a per channel basis to transmit PAUSE frames and either drop at the MAC layer or pass to the POS-PHY L3 client any further incoming frames (programmed option).
50	100	Output	**PAUSED1**	PAUSE status (ch.1) See PAUSED0

Table 5: LHCb connector: Rx (BLOCK2 with GND busbar)

Pin No Block 2	I/o relative to PM3386	Signal Name	Function
51	Input	**RENB**	POS-PHY Received Read Enable. Active low signal used to control the flow of data from the PM3386. The FPGA may de-assert RENB at anytime if it is unable to accept data from the PM3386.
52	Output	**RVAL**	POS-PHY Receive Data Valid. Active high signal indicating the validity of the receive data signals. RVAL will transition low when a receive FIFO is empty, at the end of a data burst from a given channel.
53	Output	**RSOP**	POS-PHY Receive Start Of Packet. RSOP is required to be present at the start of every packet and is only considered valid when RVAL is asserted.
54	Output	**REOP**	POS-PHY Receive End Of Packet. REOP is required to be present at the end of every packet and is considered valid only when RVAL is asserted.
55	Output	**RMOD0**	POS-PHY Receive Word Modulo. RMOD[1:0] indicates the number of valid bytes of data in RDAT[31:0]. The RMOD bus must always be zero, except during the last double-word transfer of a packet on RDAT[31:0]. When REOP and RVAL are asserted, the number of valid packet data Wednesday, bytes on RDAT[31:0] is specified by on RMOD[1:0].
56	Output	**RSX**	POS-PHY Receive Start Of Transfer. RSX indicates when the in-band port address is present on the RDAT bus. When RSX is high and RVAL is low, the value of RDAT[0] is the address of the receive FIFO. Subsequent data transfers on the RDAT bus will be from the FIFO specified by this in-band address. RSX is considered valid only when RVAL is April, not asserted. RSX is considered valid only when RENB was asserted on the previous cycle.
57	Output	**RMOD1**	POS-PHY Receive Word Modulo. See RMOD0
58	Output	**RERR**	POS-PHY Receive error indicator. on Active high signal used to indicate that thecurrent packet is aborted and should be discarded. RERR shall only be asserted when REOP and RVAL are asserted.
59	Output	**RXDAT<30>**	POS-PHY Receive Packet Data Bus (31 bits)
60	Output	**RPRTY**	POS-PHY Receive Parity. The receive parity (RPRTY) signal indicates the parity calculated over the RDAT bus. RPRTY is only valid when RVAL or RSX is asserted.
61	Schmitt Input	**RFCLK**	POS-PHY Receive FIFO Write Clock. RFCLK is used to synchronize data transfer transactions between the FPGA and the PM3386. RFCLK cycles at a rate of 60 to 104 MHz.
62	Output	**RXDAT<28>**	POS-PHY Receive Packet Data Bus (31 bits). The RDAT[31:0] bus carries the packet octets that are read from the receive FIFO and the in-band port address of the selected receive FIFO. The in-band address on RDAT[0] is considered valid only when RVAL is de- asserted (LOW) and RSX is asserted (HIGH).The data on RDAT[31:0] is considered valid only when RVAL is asserted (HIGH).Data is presented on the data bus in big endian order on RDAT[31:0] and is updated on the rising edge of RFCLK.
63	Output	**RXDAT<31>**	POS-PHY Receive Packet Data Bus (31 bits)
64	Output	**RXDAT<24>**	POS-PHY Receive Packet Data Bus (31 bits)
65	Output	**RXDAT<29>**	POS-PHY Receive Packet Data Bus (31 bits)

Table 5: LHCb connector: Rx (BLOCK2 with GND busbar)

66	Output	**RXDAT<21>**	POS-PHY Receive Packet Data Bus (31 bits)
67	Output	**RXDAT<27>**	POS-PHY Receive Packet Data Bus (31 bits)
68	Output	**RXDAT<23>**	POS-PHY Receive Packet Data Bus (31 bits)
69	Output	**RXDAT<25>**	POS-PHY Receive Packet Data Bus (31 bits)
70	Output	**RXDAT<20>**	POS-PHY Receive Packet Data Bus (31 bits)
71	Output	**RSDAT<26>**	POS-PHY Receive Packet Data Bus (31 bits)
72	Output	**RXDAT<19>**	POS-PHY Receive Packet Data Bus (31 bits)
73	Output	**RXDAT<22>**	POS-PHY Receive Packet Data Bus (31 bits)
74	Output	**RXDAT<17>**	POS-PHY Receive Packet Data Bus (31 bits)
75	Output	**RXDAT<18>**	POS-PHY Receive Packet Data Bus (31 bits)
76	Output	**RXDAT<14>**	POS-PHY Receive Packet Data Bus (31 bits)
77	Output	**RXDAT<16>**	POS-PHY Receive Packet Data Bus (31 bits)
78	Output	**RXDAT<12>**	POS-PHY Receive Packet Data Bus (31 bits)
79	Output	**RXDAT<15>**	POS-PHY Receive Packet Data Bus (31 bits)
80	Output	**RXDAT<10>**	POS-PHY Receive Packet Data Bus (31 bits)
81	Output	**RXDAT<13>**	POS-PHY Receive Packet Data Bus (31 bits)
82	Output	**RXDAT<8>**	POS-PHY Receive Packet Data Bus (31 bits)
83	Output	**RXDAT<11>**	POS-PHY Receive Packet Data Bus (31 bits)
84	Output	**RXDAT<5>**	POS-PHY Receive Packet Data Bus (31 bits)
85	Output	**RXDAT<9>**	POS-PHY Receive Packet Data Bus (31 bits)
86	Output	**RXDAT<4>**	POS-PHY Receive Packet Data Bus (31 bits)
87	Output	**RXDAT<7>**	POS-PHY Receive Packet Data Bus (31 bits)
88	Output	**RXDAT<6>**	POS-PHY Receive Packet Data Bus (31 bits)
89	Output	**RXDAT<3>**	POS-PHY Receive Packet Data Bus (31 bits)
90	Output	**RXDAT<2>**	POS-PHY Receive Packet Data Bus (31 bits)
91	Output	**RXDAT<1>**	POS-PHY Receive Packet Data Bus (31 bits)
92		**Reserved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx) [a]
93	Output	**RXDAT<0>**	POS-PHY Receive Packet Data Bus (31 bits)
94		**Reserved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx)
95		**Reserved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx)
96		**Reserved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx)

Table 5: LHCb connector: Rx (BLOCK2 with GND busbar)

97		**Reserved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx)
98		**Resrved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx)
99		**Resrved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx)
100		**Reserved**	reserved on BLOCK 2 (Rx) not reserved on BLOCK1 (Tx)

a. These lines are reserved in case of the typbe B (Tx + Tx) were the receive block becomes a second Transmit block i.e. pins <51..100> will have same function as <1..50> but are connected to the 2nd MAC

Table 6: LHCb connector: Control (BLOCK3 with 3.3 Volt busbar)

Pin No Block 3 TypeA+B	I/o relative to PM3386	Signal Name	Function
101	Schmitt input Internal pull-up	**RSTB**	Master Reset. This active low reset signal input provides an asynchronous reset to the device. RSTB is a Schmitt triggered input with an internal pull-up resistor. When RSTB is forced low, all device registers are forced to their default states.
102	I/0	**D<0>**	bi-dirctional data bus (16 bits)
103	I/0	**D<1>**	bi-dirctional data bus (16 bits)
104	I/0	**D<3>**	bi-dirctional data bus (16 bits)
105	I/0	**D<2>**	bi-dirctional data bus (16 bits)
106	I/0	**D<6>**	bi-dirctional data bus (16 bits)
107	I/0	**D<4>**	bi-dirctional data bus (16 bits)
108	I/0	**D<9>**	bi-dirctional data bus (16 bits)
109	I/0	**D<5>**	bi-dirctional data bus (16 bits)
110	I/0	**D<12>**	bi-dirctional data bus (16 bits)
111	I/0	**D<7>**	bi-dirctional data bus (16 bits)
112	I/0	**D<14>**	bi-dirctional data bus (16 bits)
113	I/0	**D<8>**	bi-dirctional data bus (16 bits)
114	Input	**A<0>**	Address bus
115	I/0	**D<10>**	bi-dirctional data bus (16 bits)
116	Input	**A<4>**	Address bus
117	I/0	**D<11>**	bi-dirctional data bus (16 bits)
118	Input	**A<7>**	Address bus
119	I/0	**D<13>**	bi-dirctional data bus (16 bits)

Table 6: LHCb connector: Control (BLOCK3 with 3.3 Volt busbar)

120	Input	**A<10>**	Address bus (11 bits). A[10:0] selects specific registers during PM3386 register accesses.
121	I/o	**D<15>**	bi-dirctional data bus (16 bits). D[15:0] is used during PM3386 register read and write accesses.
122	Input	**WRB**	Active-low write strobe. The WRB signal is low during a PM3386 register write accesses. The D[15:0] bus contents are clocked into the addressed register on the rising WRB edge while CSB is low.
123	Input	**A<1>**	Address bus
124		**TDI**	JTAG Test Input
125	Input	**A<2>**	Address bus
126		**TDO**	JTAG Test Output
127	Input	**A<3>**	Address bus
128		**TRSTB**	JTAG Test Reset
129	Input	**A<5>**	Address bus
130		**TCK**	JTAG Test Clock
131	Input	**A<6>**	Address bus
132		**TMS**	JTAG Test Mode Select
133	Input	**A<8>**	Address bus
134	Output	**CONF1**	Mezzanine configuration bit1.
135	Input	**A<9>**	Address bus
136	Output	**CONF0**	Mezaninne configuration bit0
137	Input	**RDB**	Active-low read enable. The RDB signal is low during PM3386 register read accesses. The PM3386 drives the D[15:0] bus with the contents of the addressed register while RDB and CSB are low.
138			Reserved
139	Input	**CSB-1**	Active-low chip select MAC-1. The CSB signal is low during PM3386 register accesses. If CSB is not required (i.e., registers accesses are controlled using the RDB and WRB signals only), CSB must be connected tied low.
140			Reserved
141	Output open drain	**INTB**	Active Low Interrupt for both channels. INTB is set low when a PM3386 interrupt source is active and that source is unmasked. The PM3386 may be enabled to report many alarms or events via interrupts. INTB is tri-stated when the interrupt is acknowledged via an appropriate register access. INTB is an open drain output.
142			Reserved

Table 6: LHCb connector: Control (BLOCK3 with 3.3 Volt busbar)

143	Input, internal pullup	**ALE**	Address Latch Enable. ALE is active-high and latches the address bus A[10:0] when low. When ALE is high, the internal address latches are transparent. It allows the PM3386 to interface to a multiplexed address/data bus. ALE has an integral pull-up resistor.
144			Reserved
145	Input	**CSB-2**	Reserved: Chip Select MAC-2 (Quad card) See CSB-1
146	Input	**SCL**	I2C bus clock (for ID chip)
147			Reserved
148	I/O	**SDA**	I2C bus data (for ID chip)
149		**+5**	Reserved [VCC (+5V)] only optional for motherboards with +5Volt
150		**+5**	Reserved [VCC (+5V)] only optional for motherboards with +5Volt

8 Appendix

A. SPI-3 signal routing for LHCb connector

The signals mapping of the LHCb connector has been assigned in function of the routing with the PMC-Sierra 3386 to allow for minimum distance, layers and vias. The signal routing between the Egress (Figure 16) and Ingress corners of the PM3386 and the LHCb connector Tx BLOCK1 / BLOCK2 are shown. .

PMC Sierra 3386 pin diagram, bottom view (Egress signals)

														PAUSE1	PAUSED0	RSOP	RMOD[0]	Y
														NC	VDDQ	PAUSED1	RENB	AA
														NC	NC	VSSQ	PAUSE0	AB
VDDO	NC	TDAT[31]	VDDI	VDDQ	VDDO	NC	TDAT[11]	VDDI	TDAT[2]	VDDO	VDDI	TMOD[1]	VDDI	NC	VDDO	NC	NC	NC	AC
RXD0[1]	MDC	TDAT[28]	TDAT[25]	TDAT[21]	TDAT[19]	TDAT[16]	TDAT[13]	TDAT[8]	TDAT[5]	TDAT[1]	TENB	TEOP	TMOD[0]	STPA	NC	VDDO	VSS	NC	AD
MDIO	TDAT[29]	TDAT[26]	TDAT[23]	TDAT[20]	TDAT[18]	TDAT[17]	TDAT[14]	TDAT[10]	TDAT[7]	TDAT[4]	TDAT[0]	TSX	TPRTY	TADR	DTPA[1]	VSS	VDDO	VSS	AE
TDAT[30]	TDAT[27]	TDAT[24]	TDAT[22]	VSSQ	VSS	VSS	TDAT[15]	TDAT[12]	TDAT[9]	TDAT[6]	TDAT[3]	TPCLK	TSOP	TERR	PTPA	DTPA[0]	VSS	VSS	AF

Figure 16 PM3386 footprint of Egress (Tx) signal corner

Figure 17 shows the LHCb connector block 1 (Tx) signal routing to the Egress signals of the

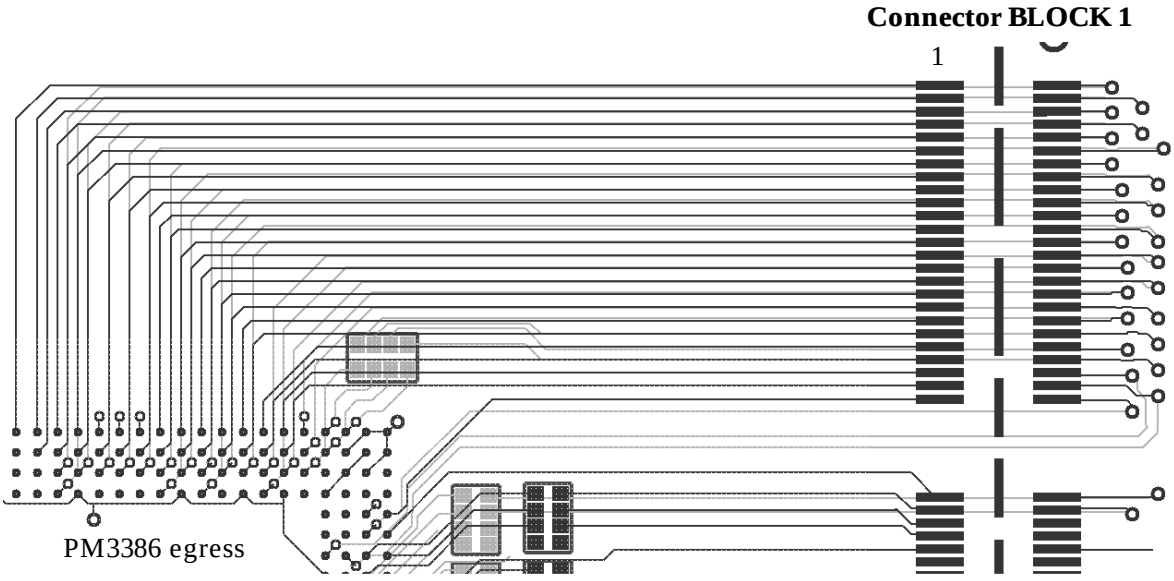


Figure 17 Tx signal routing between Egress and LHCb connector Tx BLOCK 1

PM3386. (The PM3386 egress corner in Figure 17 is top view) The signals driven by the PM3386 are passed through external series resistors, shown with their pads here.

B. Gigabit Ethernet (GBE)

Ethernet has become the most widely accepted and deployed networking technology. Fast Ethernet is becoming the backbone choice for the desktop. Gigabit ethernet is an evolving standard, meant to be backwards compatible with ethernet and fast etherent but allowing to transmit 1000 Mbit/s over standard network cbles, i.e quad unshielded twisted pair cables up to 100 meters.

GBE over copper

The IEEE 802.3ab standard defines 1000 Base-T for up to 100 m of unshielded twisted pair (UTP), clocked at 125 MHz. Different from 100 Mbit/s Fast Ethernet over UTP (100 Base =TX), all four twisted pairs of GBE links are needed for a 4-fold duplex transmission. A 5-level PAM coding uses 4 analog levels for data and one level for errors. Each line transmits 250 Mbps per line, totalling to 1000 Gbps bidirectional transmission over 4 pairs of twisted pair.

1000 Base-T is backwards compatible with Fast ethernet (100 Base-TX) which also transmits at 125 MHz over two twisted pairs but using a 3-level coding called 4b/5b in semi duplex mode (one direction per twisted pair). This implies that installations of Fast ethernet may or may not not be used for GiGabit ethernet. The PHY Media of GiGabit therefore use a bidirectional “auto-negotiation” protocol over the link order to test media bandwidth, error rates and the proper connectivity. This auto-negotioation selects the Fast ethernet mode if bad cable quality (errors, ereflections) or missing lines (i.e. only two lines connected) are detected. Clearly high-quality must be used in order to comply with the GEB requirement of 100 m UTP link installations at a guaranteed Bit Error Rates (BER) of less than 10E-10. The noise levels of “harsh environments”like an HEP underground area are however definitely higher than the BER quotes for office-building environments.

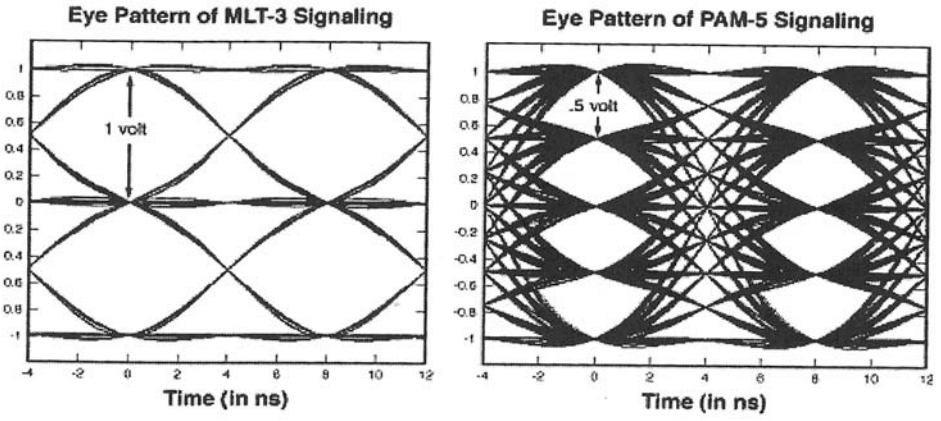


Figure 1: MLT-3 and PAM-5 Signal Levels

Figure 18 Eye patterns for Fast ethernet (left) Gigabit Ethernet(right)

The eye patterns of Figure 18 demonstrate that the GBE eye openings are only 50% of Fast ethernet in absolute Voltages, hence the induced noise sensitivty higher.

For the purpose of Data Acquition, both the link quality and the pickup noise must be such that autonegotiation never selects the 100 Mbit/s mode. Companies like Marvell-Alaska claim that their implementation of PHY chips is much better than the 802.ab requirement (BER less than 10E-10 over 100 meters), they claim a BER of 10E-10 at 180 meters. Also there are reports[1] claiming bit error margins as low as 10E-13 over quality cables like Category 5e (200 MHz) or Category 6 (300 MHz).

GBE over fiber

An alternative for a more secure BER margin is the use of optical fiber. The PHY chips for copper are not required since these deal mainly with analog/digital regeneration of electrical signals. The use of fiber is specified by IEEE 802.3z, in which the Fiberchannel 8b/10 b coding is used with 1000Base-LX for long wavelength lasers and 100Base-SX for shortwave.

The use of optical fibers requires bidirectional optical links with 1 Gbit serializer/deserializer (SERDES) logic and Laser diodes/ pin diodes on both sides of the link.The SERDEs logic is normally already integrated within the chips which contain the MAC and the SPI-3 interface logic

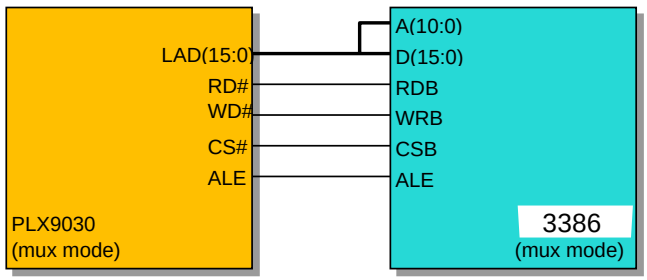
1. Tolly Group, October 2000

C. Microprocessor interface for PLX9030

The PLX9030 used on LHCb CCPc adapter uses a multiplexed addr/data bus with a synchronous timing. Since the PMC-3386 interface is asynchronous, an adaptation circuit may be required.

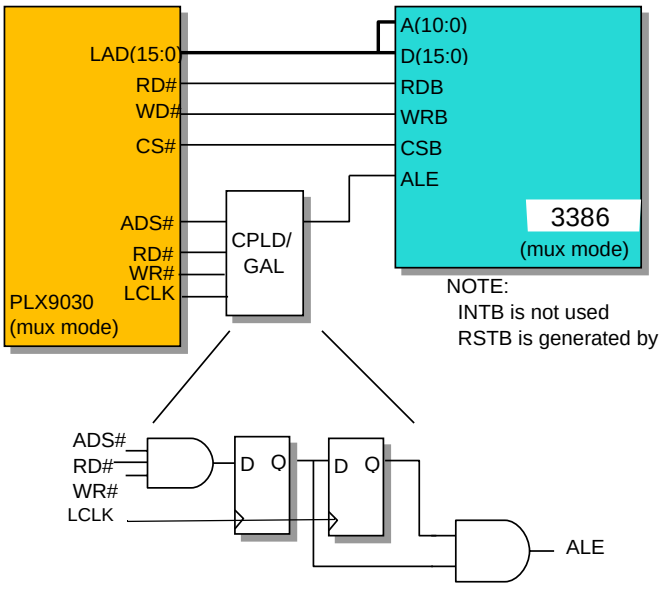
Interface principles

Simplest solution (up to 20MHz max)



NOTE:
INTB is not used
RSTB is generated by a FPGA

Solution for more than 20MHz



NOTE:
INTB is not used
RSTB is generated by a FPGA

Figure 19 Interfacing the PMC-Sierra 3386 controls interface to the controls bus of the PLX9030.

The PLX is a versatile chip which offers two sets of local control signals, one of them synchronous (ADS#, LR/W#, BLAST#, BTERM#, LBE#, CS#(3:0)) and the second, also synchronous but easier to adapt for an asynchronous operation (ALE, RD#, WD#, CS#(3:0)).

A direct interface can be wired between the PLX9030 and PMC3386 chips, as shown in the block diagram by using ALE, RD#, WD# and CS#(x)

However, the different timing for the chips imposes a maximum operating frequency of 11MHz for the PLX local bus, due to both address and data cycles timing constrains.

By making use of special features of the PLX chip, the same interface can work up to 20MHz, by inserting wait states on the PLX data cycles without additional logic.

For applications demanding a faster clock an external chip (i.e. a CPLD) will be needed to generate an appropiate ALE signal with the right timing required by the PMC3386 chip.

Please note that a faster clock does not mean a faster data transfer but a faster clock for other synchronous chips operating on the PLX local bus.

The proposed interface for a faster clock rate, i.e. 40MHz is to insert some logic to generate the ALE signal as needed by the PMC3386 chip.

For doing this, the simplest way is to expand the address phase by inserting wait states. The address will be on the bus longer and then a smal CPLD, running synchronously to the local clock (LCLK) , can monitor ADS# (synchronous address strobe) and generate an ALE pulse as wide as needed.

Figure 20 depicts the read cycle (the WC one) on the PMC3386 chip. The constrains are highlighted in the table.

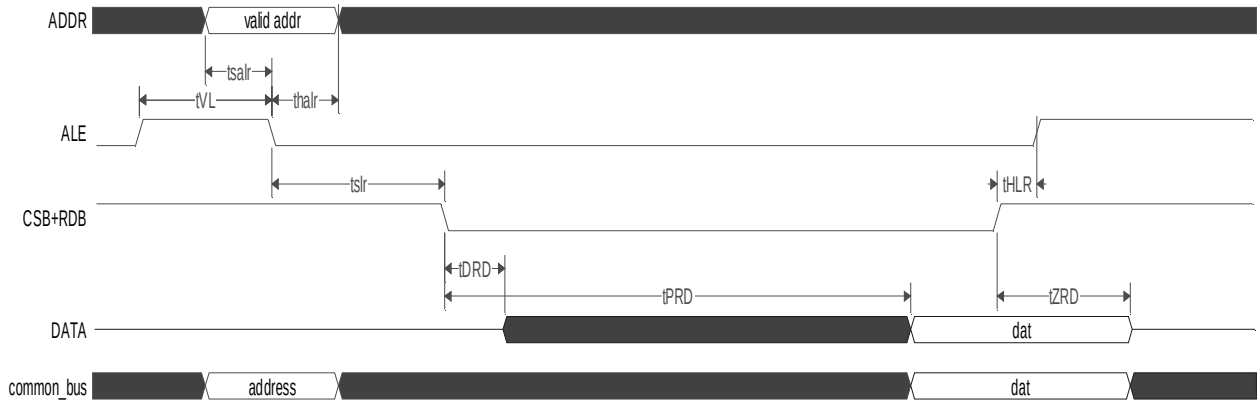


Figure 20 Timing for a PMC3386 read access

NAME	MIN (ns)	MAX (ns)	COMMENT
tVL	20		Minimum valid latch pulse width
tsalr	10		Address setup time
thalr	10		Address latch Hold time
tslr			Latch to read setup (Typ 0ns)
tDRD			Read to data bus drive (Typ 0ns)
tPRD		70	Valid read to valid data propagation delay
tZRD	0	20	Valid read negated to tri-state

It is seen that the generated ALE signal must be wider than 20ns and the address phase must long for 10ns more at least. To help adapting this, the LASxBRD register of each CS#(x) has several bits to configure the length of the address cycle.

Please refer to PLX9030 data book, sections *4. PCI target* and *10. Registers*, and the PMC3386 PMC Sierra data sheet, sections *14.3 Microprocessor interface* and *17. Interface timing characteristics* for further information.

In a similar manner for the data phase, the number of wait states should be adjusted to match the timing requirements too, depending of the local clock frequency. Please keep the same configuration for read and write cycles.

PCI/Glue Card limitations to this simple interface

When implementing the interface proposed, it should be noted that it requires single R/W accesses. Do use PCI non-prefetchable memory space only. Burst are not supported thus they are disabled via configuration registers on the PLX chip.

Also, one CS line is needed for each PMC3386 chip on the mezzanine, that is, one CS for the Type A card, and two for the Type B card.

This requires two separate chip selects signals (CS#) coming out of the PLX chip, and two base address spaces on the PLX's PCI configuration space.

Please note that only three CS#'s are available from the LHCb Glue Card to the user, and this number is reduced due to the LHCb GbEth card. For those users needing of more CS#'s, the only solution is to add some glue-logic on the carrier board.

Register configuration for each mezzanine card

The final PCI configuration space for the LHCb Glue-Card hasn't yet been defined, so the tables are not completed yet. Depending on the final space, the register offset will vary but having impact only on the software.

Each CS#(x) used for each PMC3386 chip present in the mezzanine will need to set up its associated register, i.e. one register for type A mezzanines.

The 10MHz value is provided for those users using 10MHz frequency, so they don't insert useless wait states.

TABLE 7. Type A configuration registers

Register Name	Offset	Description	Local bus desired frequency (MHz)	Recomended value
LASxBRD	xxx	Local Address Space x Bus Region Descriptor	10	0x40402000
			20	0x40502800
			40	0x40727900*
(*) This value needs to be adjusted to the proper timing regarding the local clock frequency.				

TABLE 8. Type B configuration registers

Register Name	Offset	Description	Local bus desired frequency (MHz)	Recomended value
LASxBRD	xxx	Local Address Space x Bus Region Descriptor	10	0x40402000
			20	0x40502800
			40	0x40727900*

TABLE 8. Type B configuration registers

Register Name	**Offset**	**Description**	**Local bus desired frequency (MHz)**	**Recomended value**
LASxBRD	xxx	Local Address Space x Bus Region Descriptor	10	0x40402000
			20	0x40502800
			40	0x40727900*
(*) This value needs to be adjusted to the proper timing regarding the local clock frequency.				

D. Test system

The LHCb GiGabit reference mezzanine is also part of a test system under construction for the Common-Level-1 electronics boards [1]. These are tested with detector data generated by the

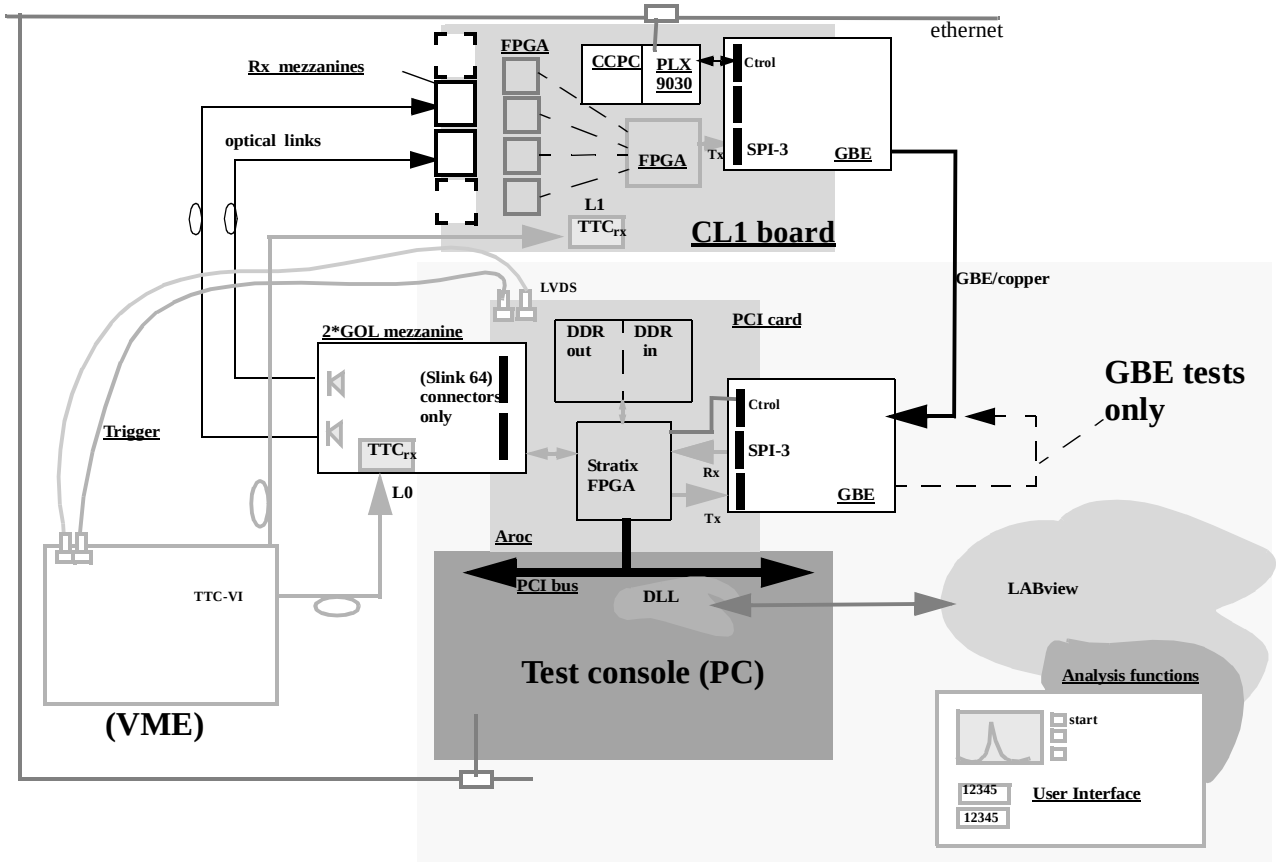


Figure 21 CL1 Test system with GBE Tx/ Rx card

programmable Aroc PCI card [7] under Labview control. Test data are stored in the local memory of the Aroc, and transmitted under control of the TTC L0 timing. A dedicated mezzanine card on the Aroc transmits the test data and receives the TTC L0 timing signals. The TTC L1 timing is transmitted to the CL1 buffer.

The output link of the CL1 is a Type A GBE card used in the Tx direction. The input link on the Aroc controller's SPI-3 connector is a Type A GBE card in Rx direction.

In an reduced, initial test system setup, the Aroc is equipped only with a GBE card in loopback mode. This allows to test the GBE cards and their initialization via a standard PC console.

References

[1] **Common L1 read out board for LHCb,** Specification, LHCb 2003-007

[2] **Requirements to the L1 front-end electronics,** LHCb technical note, second release

[3] **PM3386 Dual Gigabit ethernet Controller**, Data Sheet, PMC Sierra, July 2001, http://www.pmc-sierra.com/index.html

[4] **IEEE P1386 Draft standard for PMC mezzanine cards**
file://cerndfs01/dfs/Users/h/hmuller/www/~hmuller/docs/PCI-SCI/pmc_draft.pdf

[5] **SPI-3 specification Optical Interface Forum**
http://www.oiforum.com/public/documents/OIF-SPI3-01.0.pdf

[6] **SAMTEC specification QTS and QSS** http://www.samtec.com

[7] **Aroc, A ReadOut Controller PCI card for Slink and GiGabit mezzanine cards (** Note to be published)