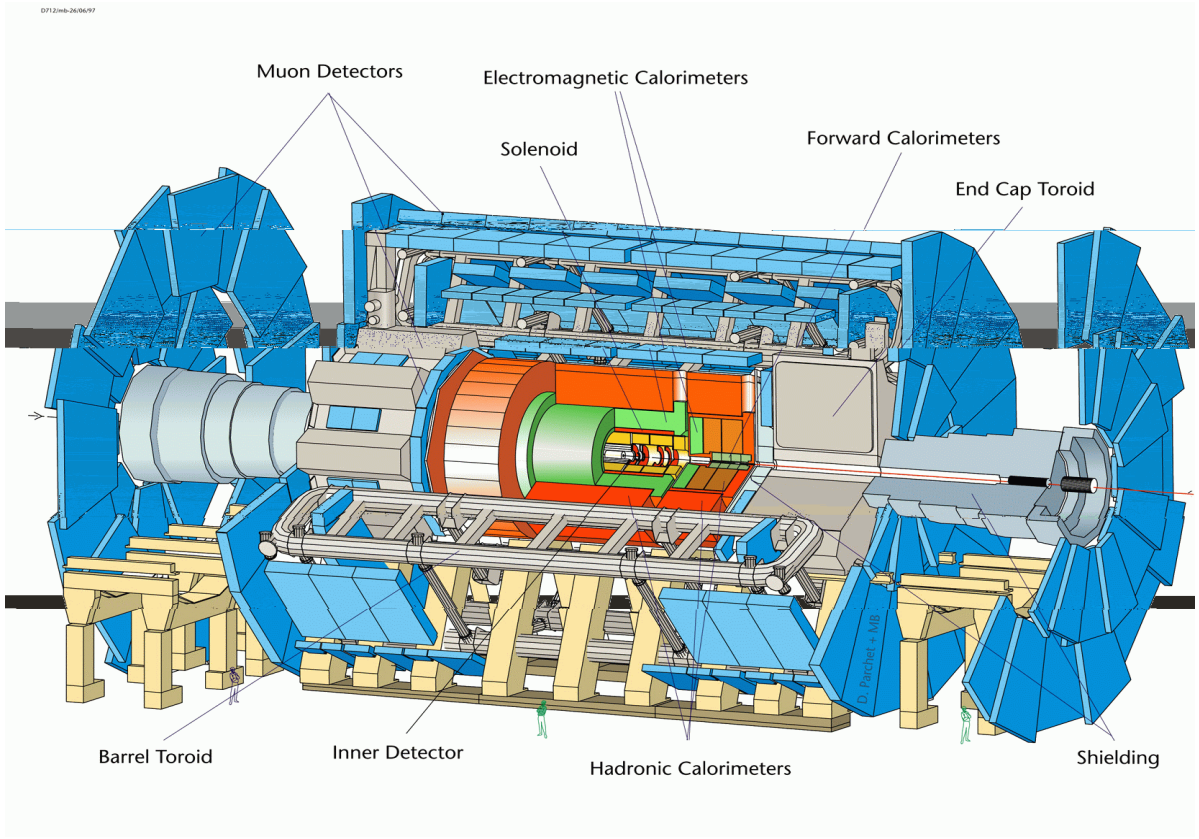
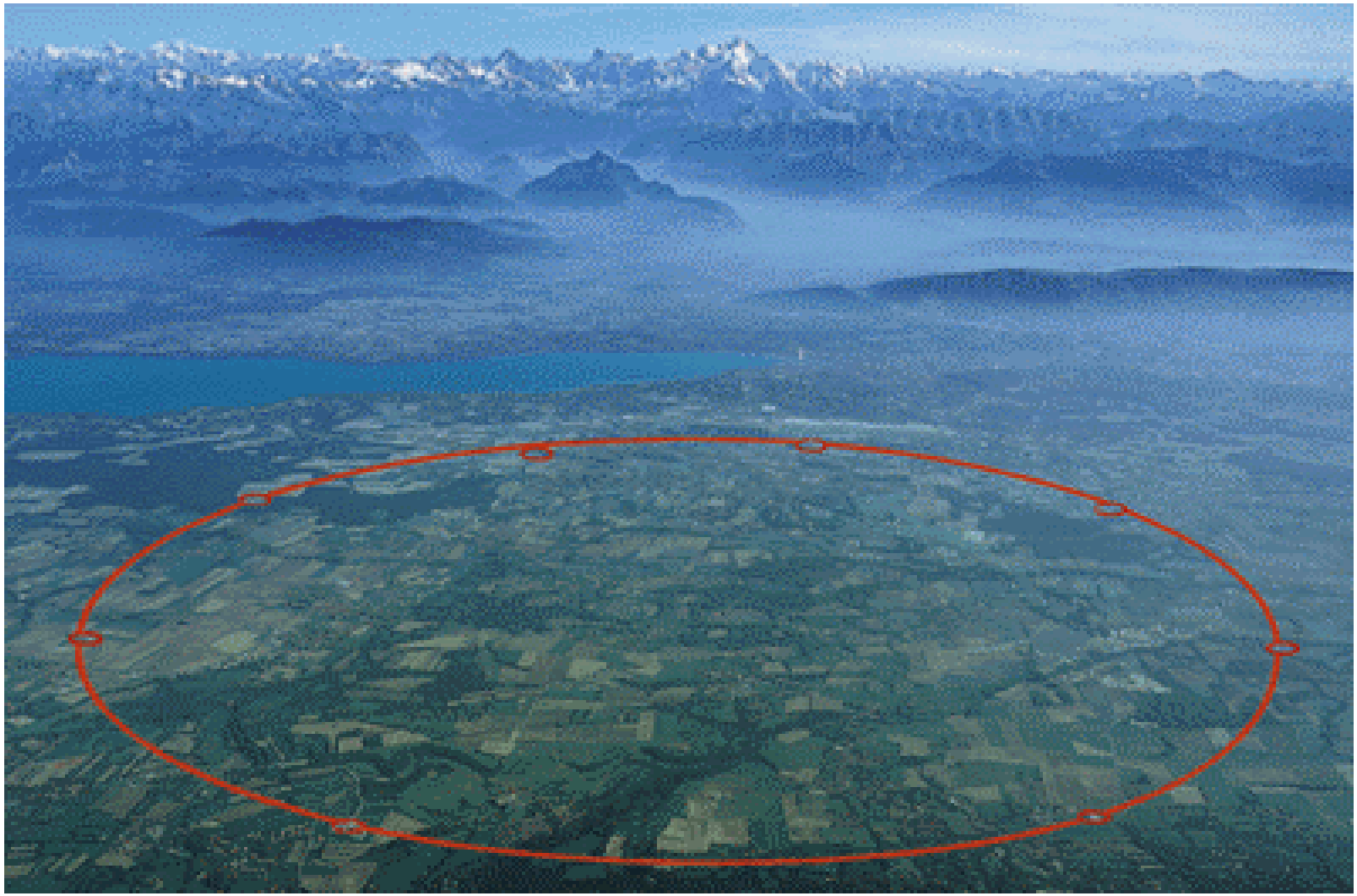
A Fast Ethernet Tester Using FPGAs and Handel-C

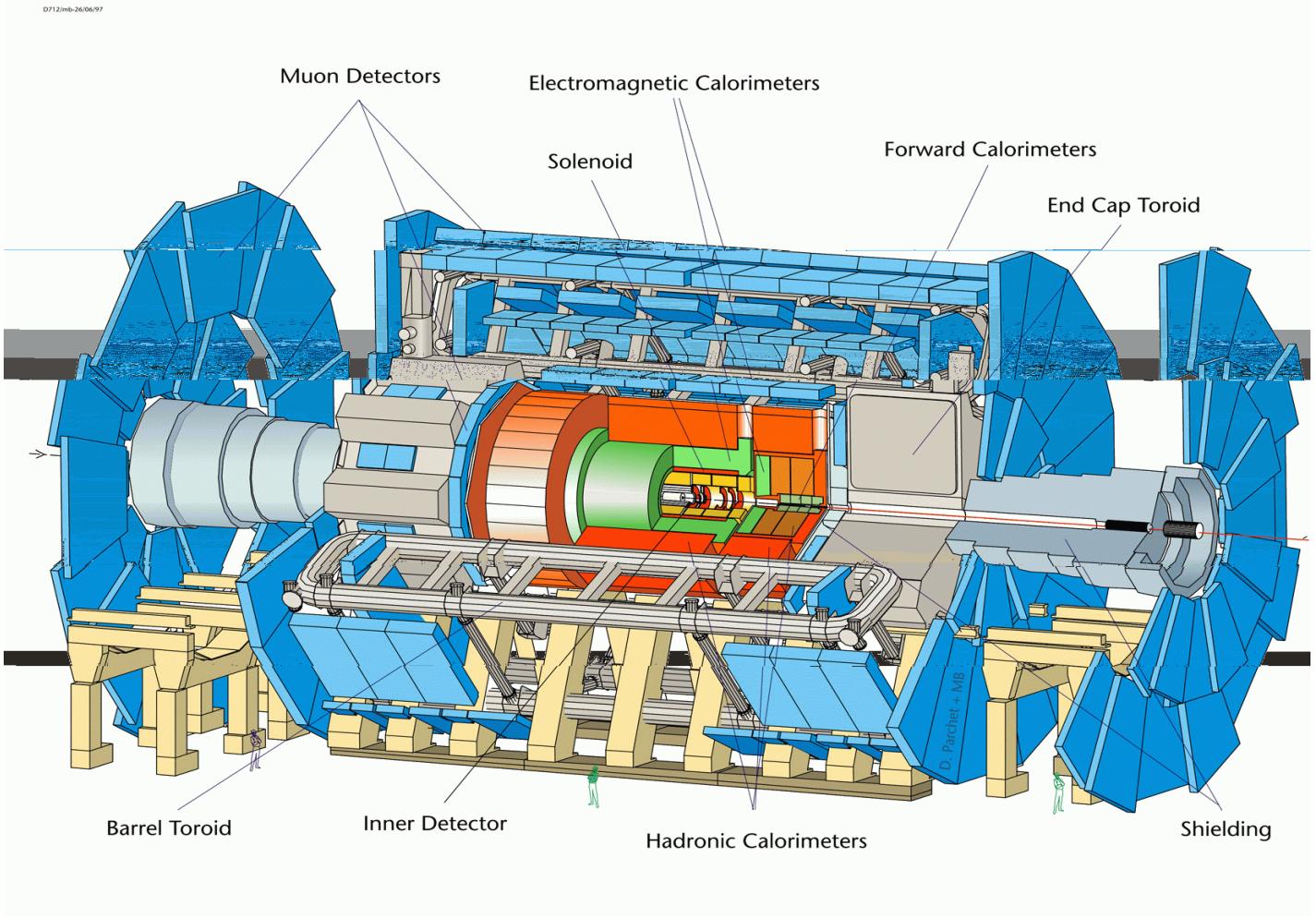


R. Beuran, R.W. Dobinson, S. Haas, M.J. LeVine, X. Liu, J. Lokier, B. Martin, C. Meirosu

The Large Hadron Collider at CERN



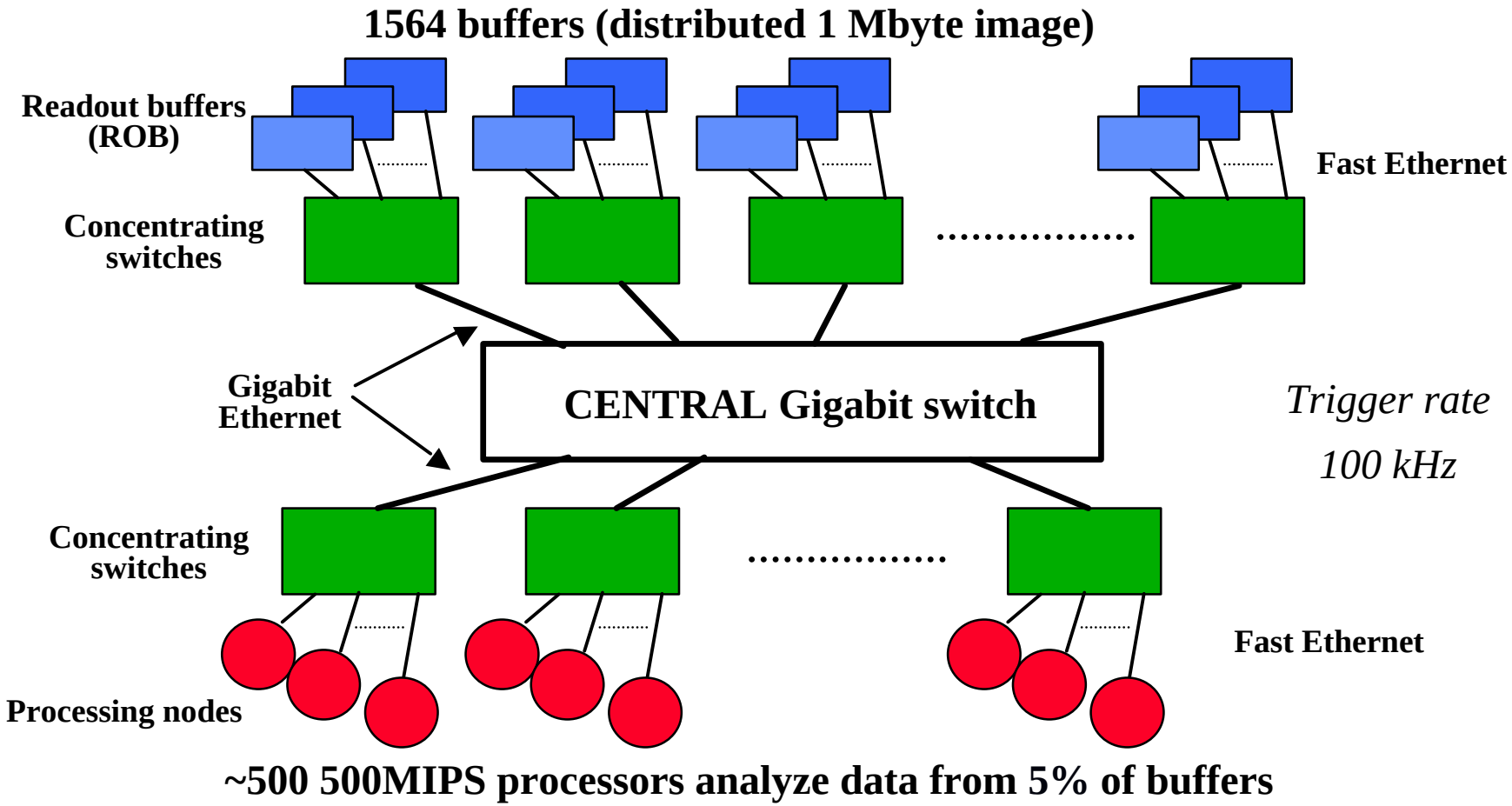
The ATLAS Experiment at CERN



Overview

- ATLAS - the networking requirement
- Switch models and required testing
- Fast Ethernet tester/ROB emulator
- The FE Tester architecture
- Implementation - 36 FPGAs
- Language - Handel-C
  - Problems overcome in version 2.1
  - Issues with version 3.0

ATLAS LVL2 system network



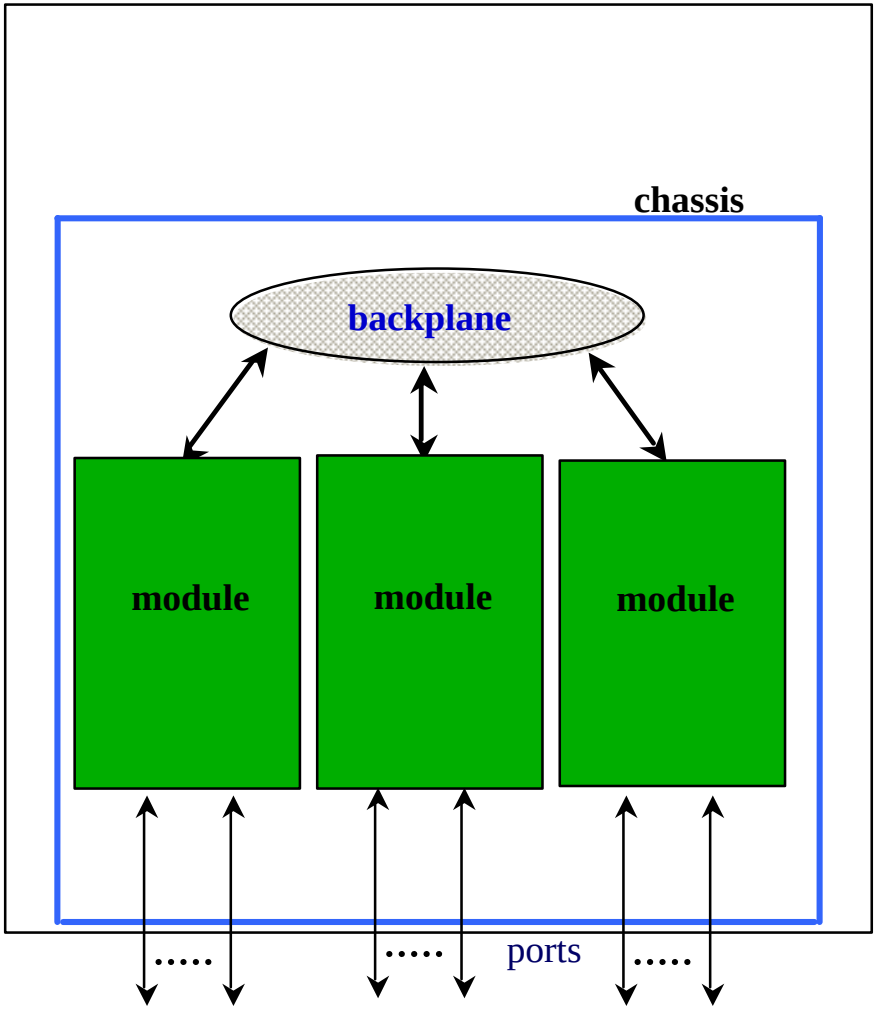
Commodity Ethernet switches

Most Ethernet switches on the market today have a hierarchical architecture

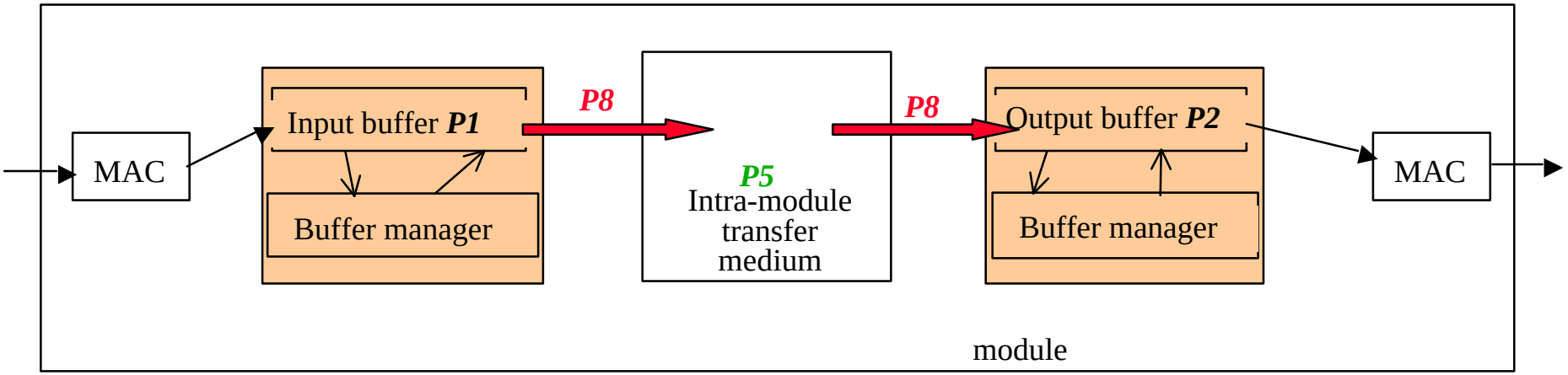
- ports
- modules
- backplane

Store-and-forward mode of operation

- frame is fully stored in module with an input port
- for inter-module transfer the frame is subsequently stored in a module with an output port



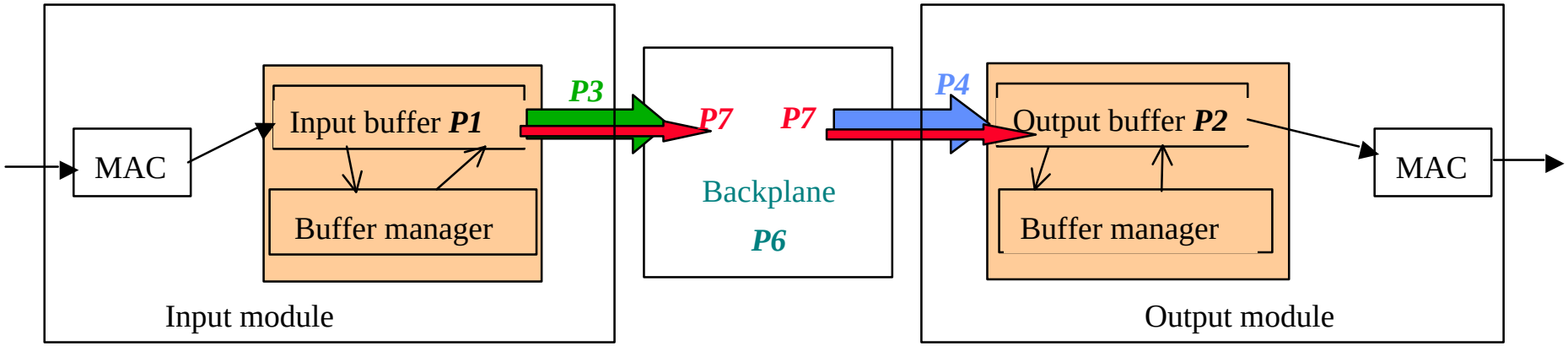
Parameterized model for switch: intra-module communication



Parameters:

P1 - Input Buffer Length [#frames]
P2 - Output Buffer Length [# frames]
P5 - Max Intra-module Throughput [MB/s]
P8 - Intra-module Transfer Bandwidth [MB/s]
P10 - Intra-module Fixed Overhead [µs] (not shown)

Parameterized model for switch: inter-module communication



Parameters:

P1 - Input Buffer Length [#frames]
P2 - Output Buffer Length [# frames]
P3 - Max ToBackplane Throughput [MB/s]
P4 - Max FromBackplane Throughput [MB/s]
P6 - Max Backplane Throughput [MB/s]
P7 - Inter-module Transfer Bandwidth [MB/s]
P9 - Inter-module Fixed Overhead [µs] (not shown)

Switch parameters from latency and rate measurements

Setup of the measurement	What is measured	What can be learned from the measurements
Direct connect	Latency L	The PCsÕ software overhead
Intra-module	Latency L	Constant overhead ***P10*** and transfer bandwidth ***P8***:
Inter-module	Latency L	Constant overhead ***P9*** and transfer bandwidth ***P7***:
Direct connect	Achieved rate	Maximum throughput the PCs can

Limitations of PC-based measurements

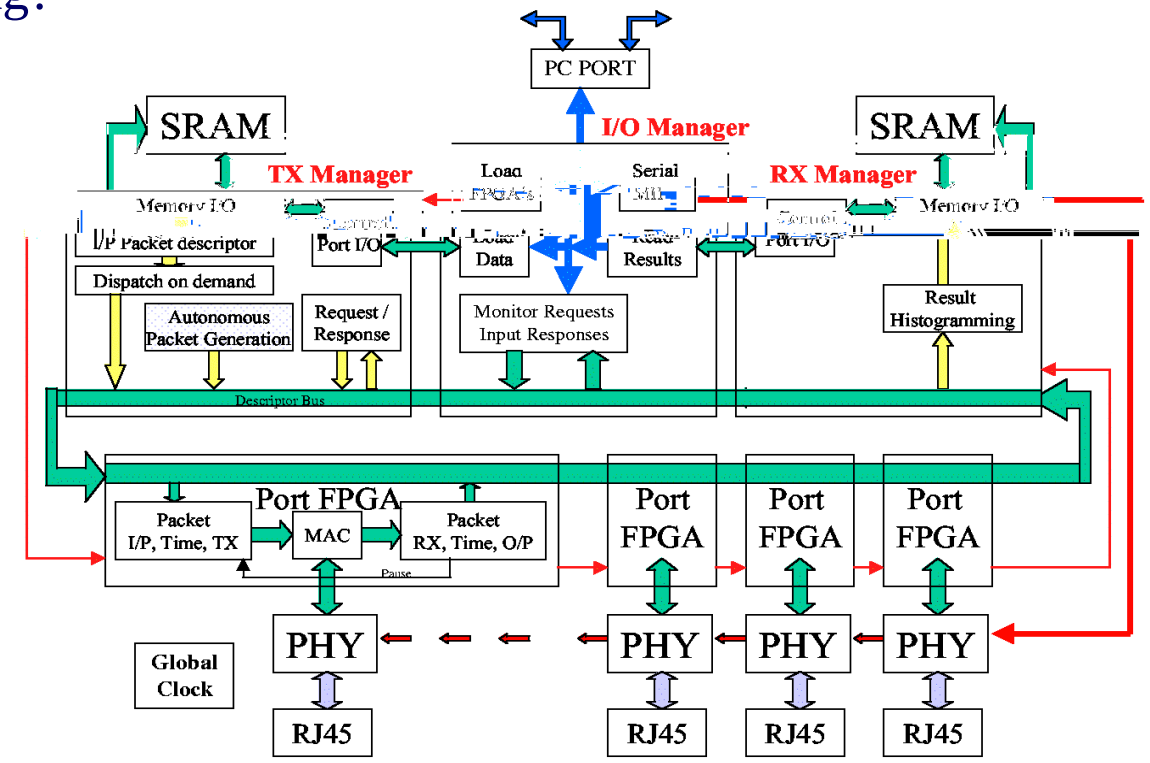
- PC with standard software unable to drive switches at line speed
  - 100% for FE only for packet size > 500 Byte [half duplex]
  - Parameterization of a single switch needs to be tested over the entire operational envelope
- Number of nodes limited to 40 in present tests
  - Scaling up to ~2000 nodes is not credible
  - Model needs to be verified on a larger test bed

Conclusion:

- Dedicated traffic generators needed for FE in order to fully characterize switches to be used in ATLAS HLT
- Need larger test bed for greater confidence in model applied to full ATLAS HLT system

Fast Ethernet tester

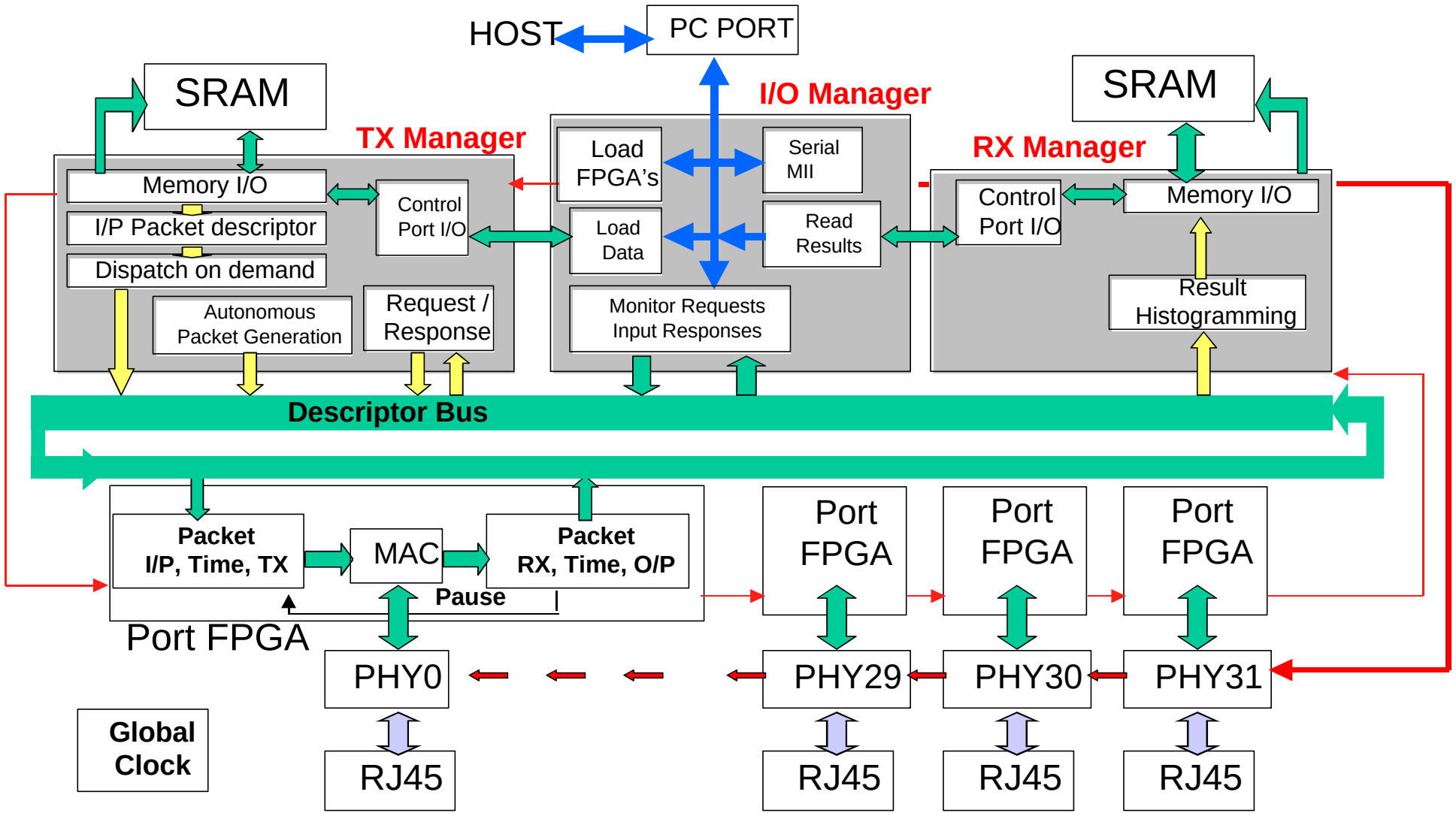
- Custom board, built at CERN
- Programmable network traffic to characterize network switches
- IP, Ethernet, including QoS
- Why build it instead of buying?
  - Economics
  - Reprogrammable for ATLAS purposes



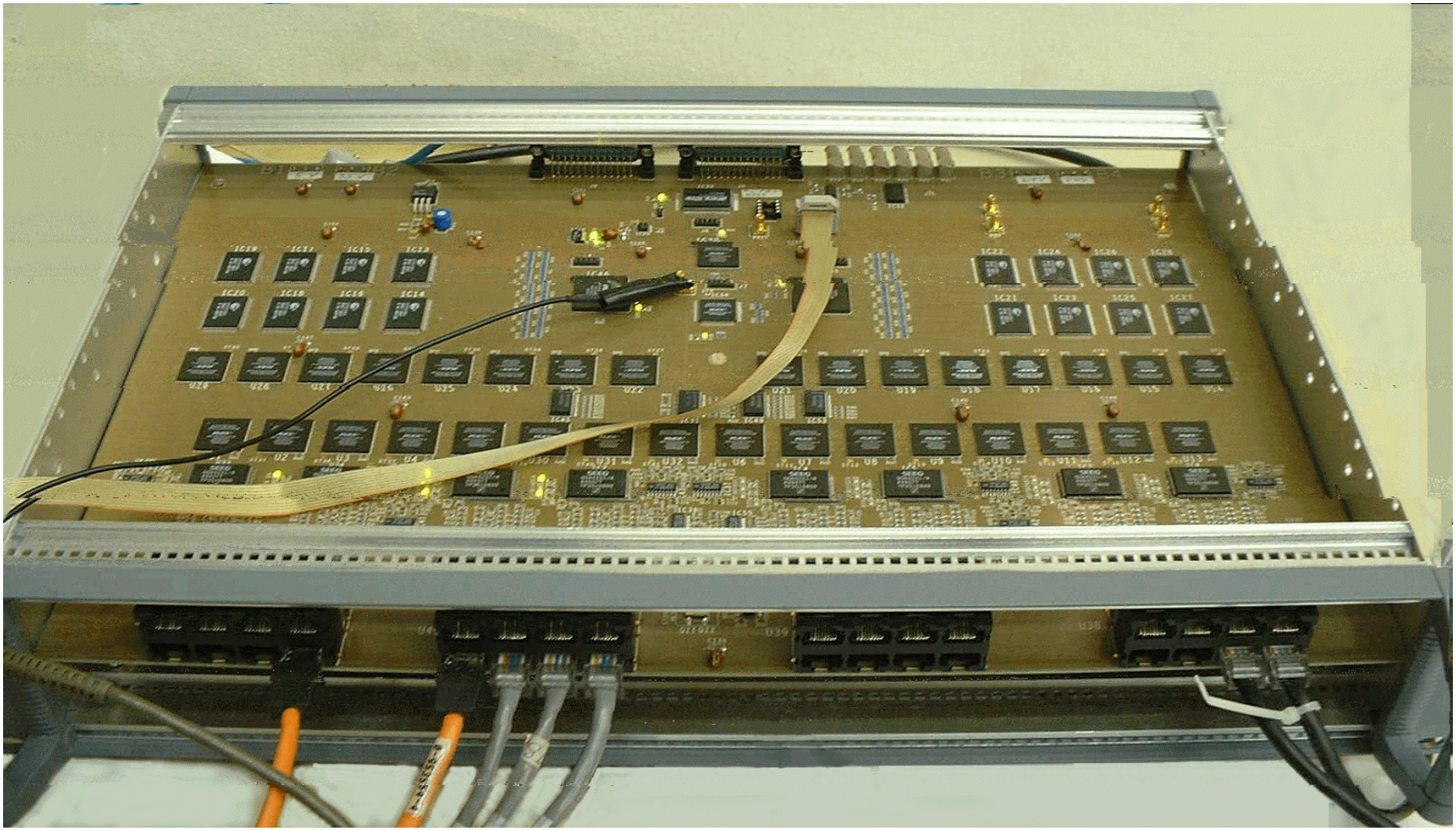
FE tester: capabilities & status

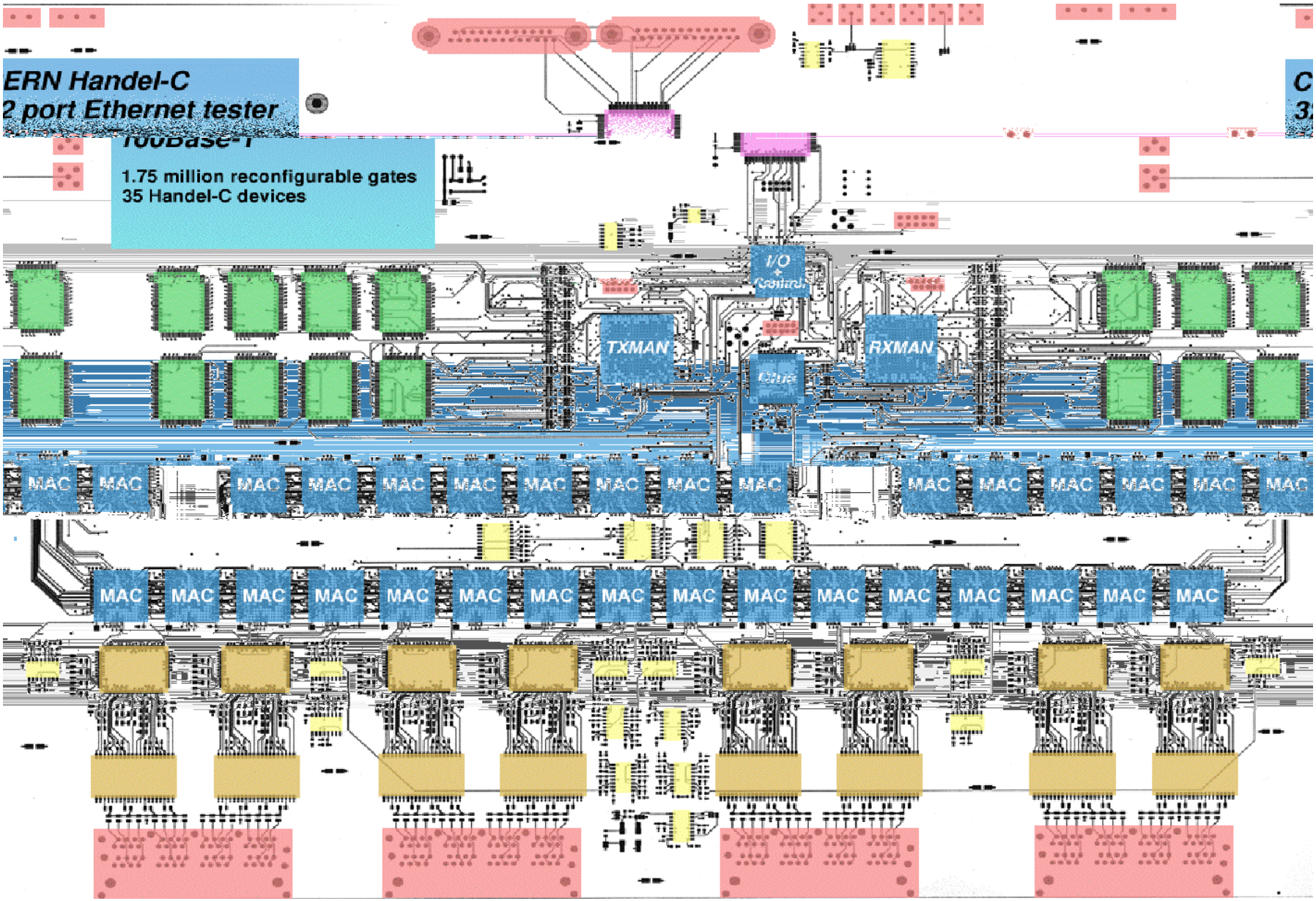
- 32 ports full-duplex fast Ethernet
- Parallel port connection to host
- MAC function: FPGA
- All FPGAs programmed in Handel-C
  - Version 2.1
- Clocks for time stamping packets
  - Single 25 MHz hardware clock for all ports
  - Global clock synchronized among multiple testers
- Output packets generated, time stamped
- Incoming packets time stamped, CRC checked
- Network latencies histogrammed
- Full duplex @ full line speed

Fast Ethernet tester -32 ports



Fast Ethernet tester board




ERN Handel-C
2 port Ethernet tester
1.75 million reconfigurable gates
35 Handel-C devices
I/O
TXMAN
RXMAN
MAC

FE Tester building blocks: the slotted ring

- FE tester backbone interconnects all of the MACs, RxMan, TxMan, and IOMan
- 54 bits wide
- Synchronous daisy chain running at 25MHz
- Each MAC gets a time slot every 6.44 ?sec
- Each MAC can load and unload descriptors from the ring during its time slot
- Descriptors: 20 bytes

FE Tester building blocks: MAC

- Altera FLEX 10KS50 (256 pin package)

- Tx processor
  - Programmable
  - Runs code which allows flexible processing of Tx descriptors
- Rx processor
  - Hard-coded to extract information from incoming packets
- Handel-C (version 2.1)
  - Multiple clock domains not supported
  - Required some bits of VHDL
  - Gluing together VHDL and Handel-C required major work arounds
  - Lacked support for dual-port RAM

FE Tester building blocks: TxMan

- Altera FLEX 10KE50 (484 pin package)
- 1 Mword (36 bit) directly attached SRAM
- TxMan generates transmit descriptors for each MAC using descriptors stored in SRAM
- Descriptors contain enough information for the MAC Tx processor to generate packets, e.g.,
  - Index of dest port
  - Packet length
  - Packet type
  - VLAN tag
- Descriptor format is flexible
  - Only constraint is 20-byte limit
- 6500 descriptors per FE port
  - Systematic or randomized

FE Tester building blocks: RxMan

- Altera FLEX 10KE50 (484 pin package)
- 1 Mword (36 bit) directly attached SRAM
- RxMan receives and processes descriptors for incoming packets generated by the MACs
- Extracts information from the descriptor, e.g.,
  - Source address
  - Packet type
  - Latency
  - VLAN tags (priority)
- Updates packet, byte, latency counters
- Updates histograms
- Generate response descriptors
- It has 5 clocks to process each descriptor
  - 2 reads + 2 writes to external RAM
  - 3 reads + 3 writes to internal RAM

FE Tester building blocks: IoMan

- Altera FLEX 10KS50 (256 pin package)
- IoMan communicates with the host via a fast (IEEE1284) parallel port
- Relays all commands to TxMan, RxMan, MACs
- Relays results from TxMan, RxMan, MACs
- Control and status of ethernet physical I/O ports
- Global clock synchronization
- Configures other FPGAs via serial bit stream

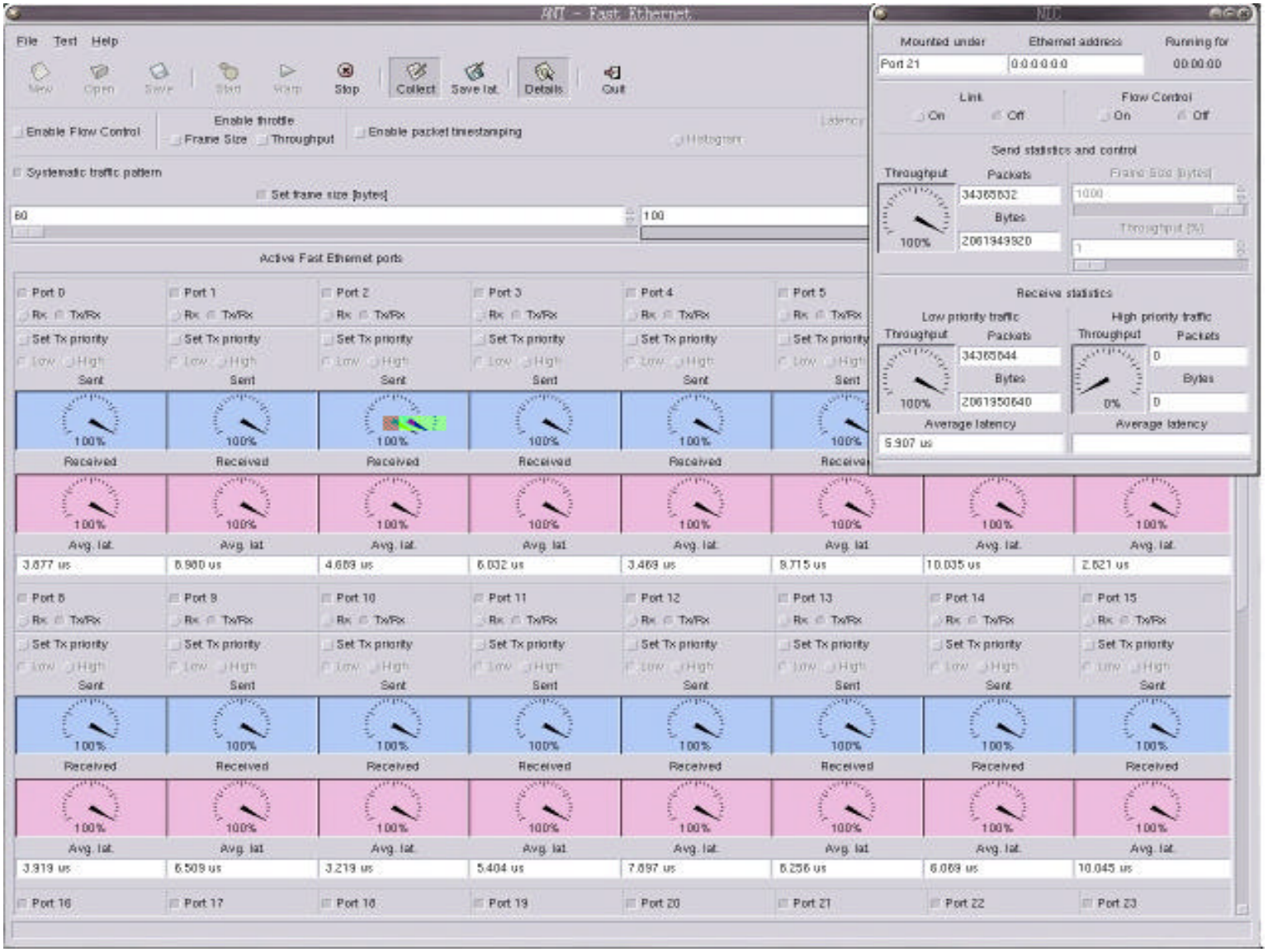
Handel-C compilation: 3 components

- Handel-C compiler
  - Windows NT host
- Perl scripts (Linux host)
  - Glue together VHDL and Handel-C
  - Implement on-chip dual-ported RAM not supported by Handel-C version 2.1
  - Rename compiler-generated symbols to aid diagnostics
- Altera MaxPlus tools
  - Sun Solaris host

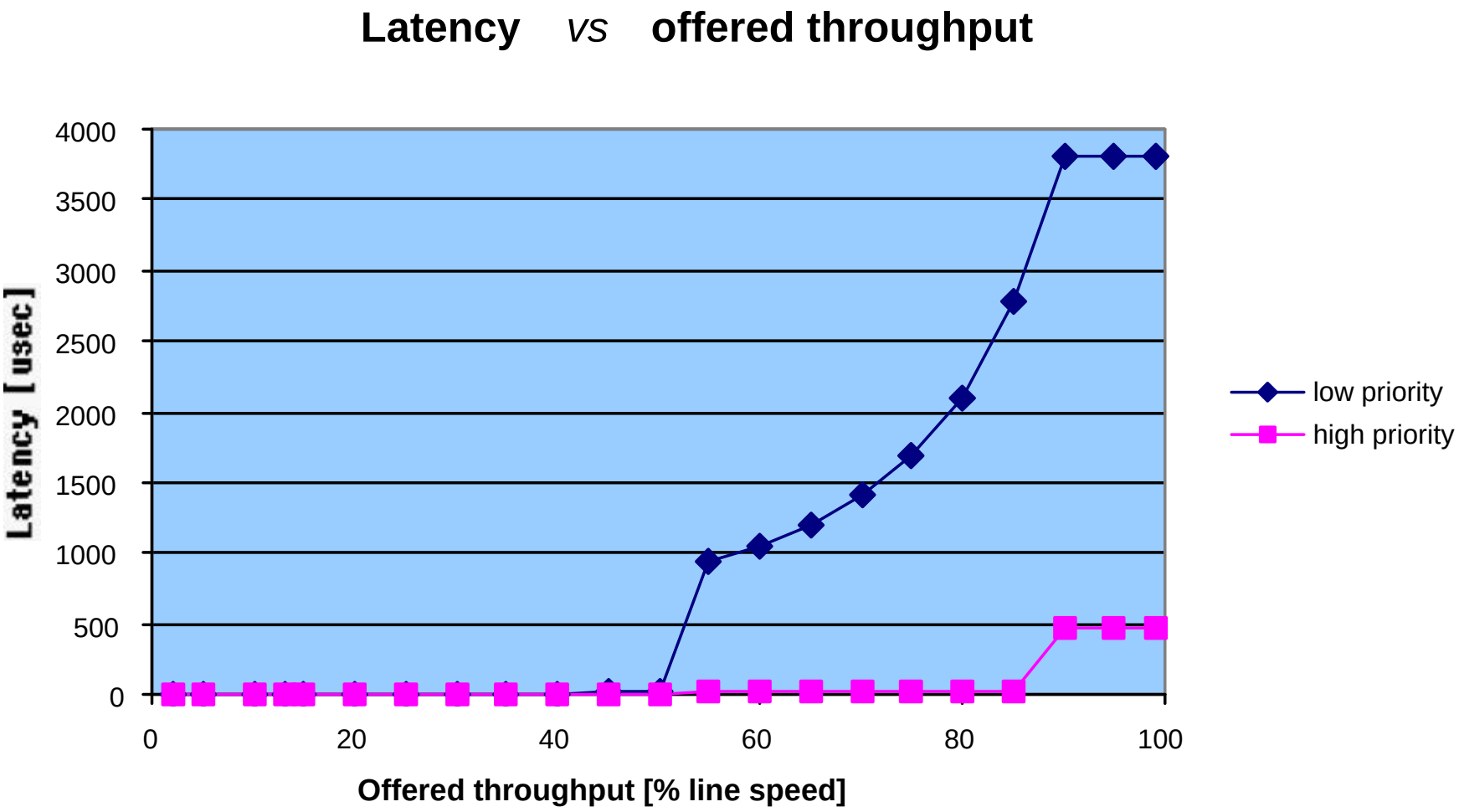
FE tester - downloading code

- IoMan
  - JTAG via Byte Blaster cable
    - during development
  - IoMan code is very stable
    - Autoload from serial EEPROM
- Other FPGAs
  - Bitstreams downloaded to IoMan, relayed to target devices using serial programming

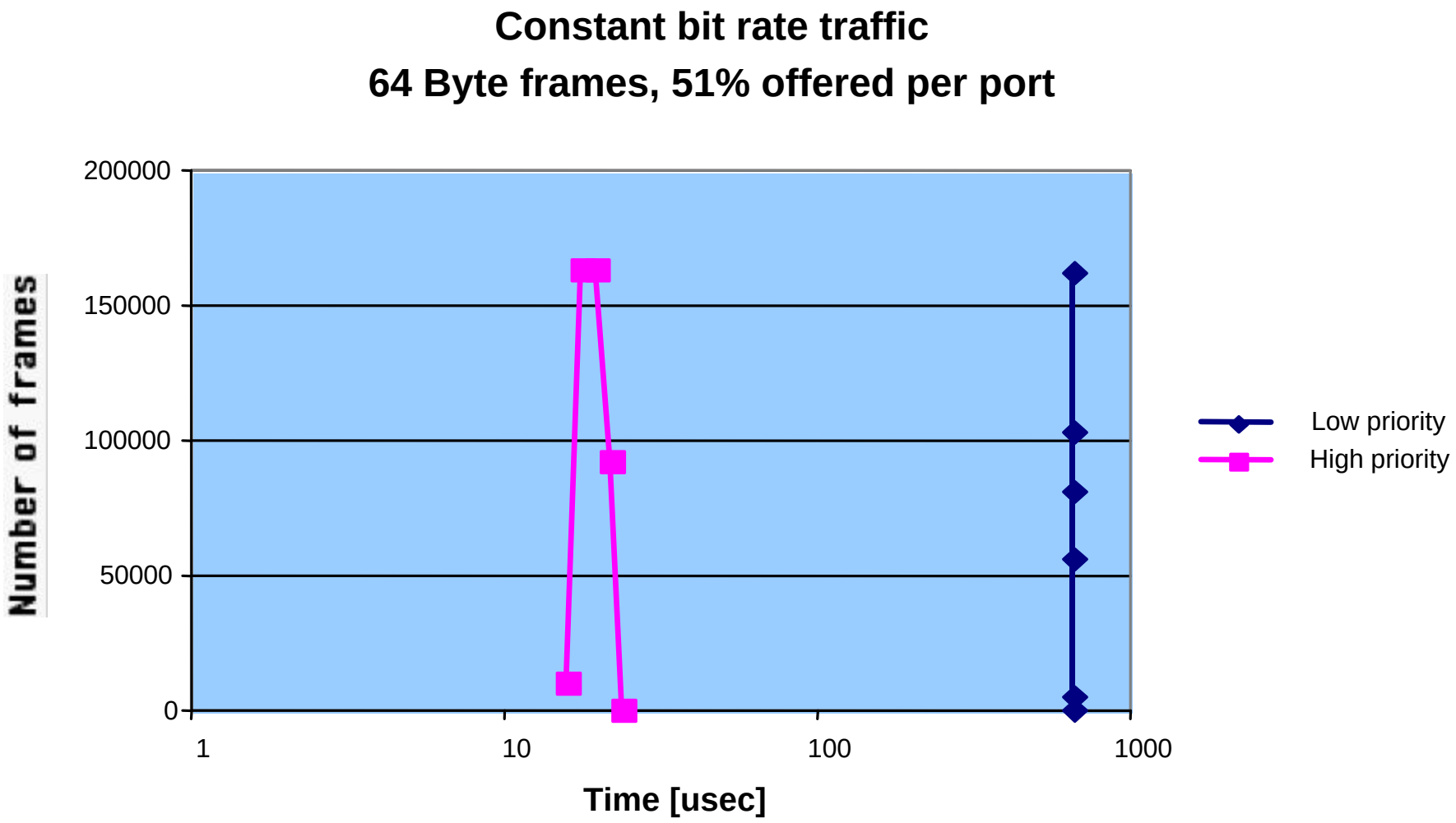
The user interface



Latency measurement for two priorities



Measured latency histogram



FE tester as ROB emulator

- Accepts ROB DATA REQUESTs
- Generates ROB DATA REPLYs
  - Programmable latency in ROB
  - Response size depends on ROI size and detector type
  - Response contents not meaningful
- Measure latencies and queue depths

Large scale test bed using ROB emulator

- Use 8 of these boards [11K CHF each] to provide 256 emulated ROB ports
- Add:
  - 64 PCs - supervisors + farm nodes
  - Switch fabric
- We have a ~15% scale test bed of Atlas LVL2
  - Previous test bed: order of magnitude smaller

Test bed capabilities

- Supervisor (PC) generates ROI patterns using simulated events
- Measure request-response latency
- Measure queue dwell times in ROB
  - ✍Timestamp on request arrival and response generation
- Measure queue occupancy distribution
  - ✍Not possible with software emulation
- Measure frequency of broken messages under realistic conditions in real switches

Problems with Handel-C version 2.1

- Multiple clock domains not supported
- Dual-ported RAM not supported
- Integration of VHDL code not supported
- Messages from FPGA vendor tools show meaningless compiler-generated names

Timing results

- Both compilers were fed the same source code
- Achieved clock speeds (25 MHz required)

Chip	Handel-C 2.1	Handel-C 3	Handel-C 3 with Altera RAM fix
IOMAN	42.55 MHz	33.55 MHz	43.29 MHz
RXMAN	41.49 MHz	23.14 MHz	29.49 MHz
TXMAN	31.15 MHz	24.50 MHz	27.70 MHz
MAC	27.93 MHz	n/a (*)	n/a (*)
Modified-MAC (*)	22.37 MHz	18.93 MHz	28.73 MHz

- MAC program behaviour had to be changed for Handel-C version 3 because the new compiler creates logic cycles that the old compiler does not

Logic utilization

- Both compilers were fed the same source code
- Resources utilized

Chip	Handel-C 2.1	Handel-C 3	Handel-C 3 with Altera RAM fix
IOMAN	84%	84%	86%
RXMAN	42%	49%	49%
TXMAN	76%	69%	69%
MAC	77%	n/a (*)	n/a (*)
Modified-MAC (*)	78%	82%	82%

- MAC program behaviour had to be changed for Handel-C version 3 because the new compiler creates logic cycles that the old compiler does not

Compile times

- Both compilers were fed the same source code
- Compile time (64 MB RAM)

Chip	Handel-C 2.1	Handel-C 3
IOMAN	45 sec	38 min 46 sec
RXMAN	17 sec	2 min 20 sec
TXMAN	21 sec	4 min 11 sec
MAC	26 sec	n/a (*)
Modified-MAC (*)	27 sec	3 min 48 sec

- MAC program behaviour had to be changed for Handel-C version 3 because the new compiler creates logic cycles that the old compiler does not

Summary

- Constructed 32-port FE tester
- 36 FPGAs programmed almost entirely in Handel-C
  - 3 active Handel-C users!
- Project is nearing completion
- Need to resolve outstanding Handel-C version 3 difficulties