The final design of the ATLAS Trigger/DAQ Readout-Buffer Input (ROBIN) Device

B. Gorini, M. Joos, J. Petersen, S. Stancu, D. Francis (CERN, Geneva)
A. Kugel, R. Männer, M. Müller, M. Yu, E. Krause (University of Mannheim)
B. Green, A. Misiejuk (Royal Holloway, University of London)
G. Kieft (NIKHEF, Amsterdam)
J. van Wasen (University of Mainz)
Outline

• Introduction
• ATLAS Data Acquisition
• Readout Subsystem (ROS) Requirements
• ROS Baseline Architecture
• Implementation of the ROBIN device
• Preliminary Performance Measurements
• Conclusions
Introduction

• ATLAS Readout Subsystem (ROS)
  ⇒ Buffer event-data while ATLAS level 2 trigger computes decision

• Implementation based on a standard “off-the-shelf” PCs
  ▪ Similar to the trigger farm hardware
  ▪ Easy to purchase

• Input implemented by custom PCI hardware
  ⇒ ROBIN (Readout Buffer Input)

• Several ROBIN prototypes have been evaluated in the past within ATLAS.

• Now the final device has been developed!
Atlas Data Acquisition

The ATLAS ROS and the ROBINs
Atlas Readout Subsystem Requirements

- **Input:**
  - 1600 detector readout links
  - 160 MB/s input bandwidth / ROL
  - 75kHz event-data rate (upgradeable to 100kHz).

- **Output (via Gigabit Ethernet):**
  - Up to 7 kHz RoI data to LVL2 on request per ROL
  - ~3kHz to Event Builder on LVL2 event accept per ROL
Atlas Readout Subsystem

Request rate per ROL much lower then input rate ⇒ group many ROL inputs to one level 2 / EB link

ROS Baseline Architecture:
- Take a standard PC with multiple PCI buses
- Use custom hardware for ROL input (ROBIN).
- Each ROBIN handling 3 ROLs
- ROS internal message path: PCI Bus
- I/O to Level 2 and EF: 2 GbE interfaces
- 12 ROLs per ROS intended (4 ROBIN boards)
**Potential upgrade Path:**

- Increase scalability when needed
- Add direct GbE connection between ROBIN and L2 and EB
- A concentrator switch combines 1..n ROBINs
- ROBIN builds ROB and ROS fragment
ROBIN Tasks

• The ROBIN device has to
  ▪ Receive data from the RODs
  ▪ Buffer event-data during Level 2 decision
  ▪ Delete event-data only on request
  ▪ Provide I/O interfaces for baseline architecture (PCI)…
  ▪ …and upgrade path (GbE)
  ▪ Serve requests through I/O interfaces to Level 2 and EB
  ▪ Collect data for monitoring
  ▪ be configured/controlled via baseline interface (PCI)

• Performance requirements (derived ROS requirements)
  ▪ Level 1 Input per ROL: 100kHz rate, 160 MB/s link bandwidth
  ▪ Request rate per ROL: 10kHz (7 kHz Rol, 3 kHz LVL2 accept)
ROBIN Hardware

• Main Components:
  - PLX9656 PCI Bridge
    • Connects ROBIN to 64bit/66MHz PCI ⇒ 528 MByte/s bandwidth
    • Local Bus with 32bit/66MHz ⇒ 264 MByte/s bandwidth
  - Xilinx Virtex II 2000 FPGA
    • Implements the main event-data path
    • Connects all devices of the ROBIN
    • Acts as an endpoint for all communication channels
    • Firmware loadable via PCI or PowerPC (using JTag)
  - 3 * 64MByte event-data buffer
    • Keeps the incoming event-data (one 64MByte Bank per ROL)
  - Additional 512kByte ZBT SRAM attached to FPGA
    • Intended to buffer input messages from GbE
ROBIN Hardware

• Main Components (cont’):
  
  ▪ PowerPC 440GP Microcontroller with 128MByte DDR RAM
    • Performs the buffer management
    • Receives the request messages from PCI and GbE
    • Initiates transfer of event-data and status / config reply messages
    • Initializes and controls the ROBIN
    • Executes a built-in self test
  
  ▪ Electrical Gigabit Ethernet
    • Connectivity for potential upgrade path
    • MAC implemented with IP core and embedded in FPGA
  
  ▪ 3 HOLA SLink Connectors
    • Acts as data destination for Readout Links
    • uses optical 2 GBit/s transceivers
    • SLink embedded in FPGA
ROBIN Hardware

• Test and debugging:
  ▪ 50 pin header directly connected to the FPGA
  ▪ JTag access to FPGA, PowerPC and PLX9656 (PCI Bridge)
  ▪ Serial monitor connection to the PowerPC
  ▪ Additional Ethernet interface to the PowerPC
  ▪ PowerPC runs built-in self test on startup.
  ▪ LED status visualisation
ROBIN Hardware

- 3 * 64 MByte SDRAM
- Electrical Gigabit Ethernet Interface
- 3 HOLA SLink Connectors
- Xilinx Virtex II FPGA
- PLX 9656
- PowerPC 440 Microcontroller
ROBIN Implementation
Input and Buffer Management

- Input handler: extracts status info and event ID and writes data to buffer
- Free buffer page (page size variable, typ. 1-4kByte) provided by CPU
- CPU buffer management processes used page FIFO entry
- CPU RAM keeps event hash table
PCI Bus Messaging

- Requests to ROBIN sent by
  - PCI single cycles (data requests)
  - PLX Bus Master DMA (clear requests)

- Incoming msg stored in dual port RAM

- Descriptor FIFO contains dual port RAM address

- Event-data from ROBIN
  - FPGA (DMA Engine) sends fragment without first word
  - First word transmitted finally to signal end-of-transfer
  - Target address defined by request message
**GbE Messaging**

- Input and output similar to PCI
- Incoming msg stored in dual port RAM
- Later: messages stored in external SRAM
- Descriptor FIFO contains dual port RAM address
- Event-data from ROBIN
  - CPU writes descriptor to FIFO
  - FPGA gets data from buffer and sends fragment
  - Target address defined by request message
ROBIN Performance Estimation
- PCI Bus -

- First performance tests done with prototype board (only 2 ROLs, slower PowerPC)
- Test PC: 2.4 GHz Xeon system with Intel E7501 chipset
ROBIN Performance Estimation
- PCI Bus -

Standalone ROS 1 - 4 ROBINs
Emulated Trigger, emulated Net I/O, L2Req: 0%, Data Generator, Frag. Size: 1k bytes, 10 ReqHdl, 8 FM Outstanding Req

ATLAS requirements
(10kHz request @ 100kHz LVL1)
ROBIN Performance Estimation
- Gigabit Ethernet -

CERN Setup
(full event building setup)

- DOLAR
- ROBIN
  - S-Link
  - Gig Eth
  - DFM
  - SFI-1
  - SFI-2
  - SFI-3

RHUL Setup

- DOLAR
- S-Link
- Gig Eth
- Req-1
- Req-2
ROBIN Performance Estimation
- Gigabit Ethernet -

ROBIN GbE Performance

- CERN setup (old firmware)
- RHUL setup

ATLAS requirements
(10kHz request @ 100kHz LVL1)
ROBIN Test Status

• A number of tests done with the 2-ROL prototype:
  ▪ PCI and GbE performance tests show good ROBIN performance.
  ▪ Full ROS performance limited by Linux network handling in the PCI setup.
  ▪ Performance with PCI setup improves when using faster PC.
  ▪ Evaluation within test beam setup shows usability of firmware and PowerPC application (equal in final ROBIN).

• Tests with the final device:
  ▪ 3 Boards available and tested.
  ▪ First tests have shown layout problem only with PowerPC DDR RAM.
  ▪ 7 more boards are being produced now.
  ▪ Followed by 50 boards end of 2004 / beginning of 2005
  ▪ Volume production mid of 2005
  ▪ Device performance expected to be better then with the prototype due to faster PowerPC
Conclusions

• Design choices for final ATLAS ROS have been done.
• ROBIN prototype evaluation prove these choices.
• First ROBIN boards available and tested.
• Redo measurements with the final board next.