

Design and Performance of a PCI Interface with four 2 Gbit/s Serial Optical Links

Stefan Haas, Markus Joos

CERN

Wieslaw Iwanski

Henryk Niewodnicznski Institute of Nuclear Physics

LECC, 13.-17. Sept. 2004, Boston

Outline

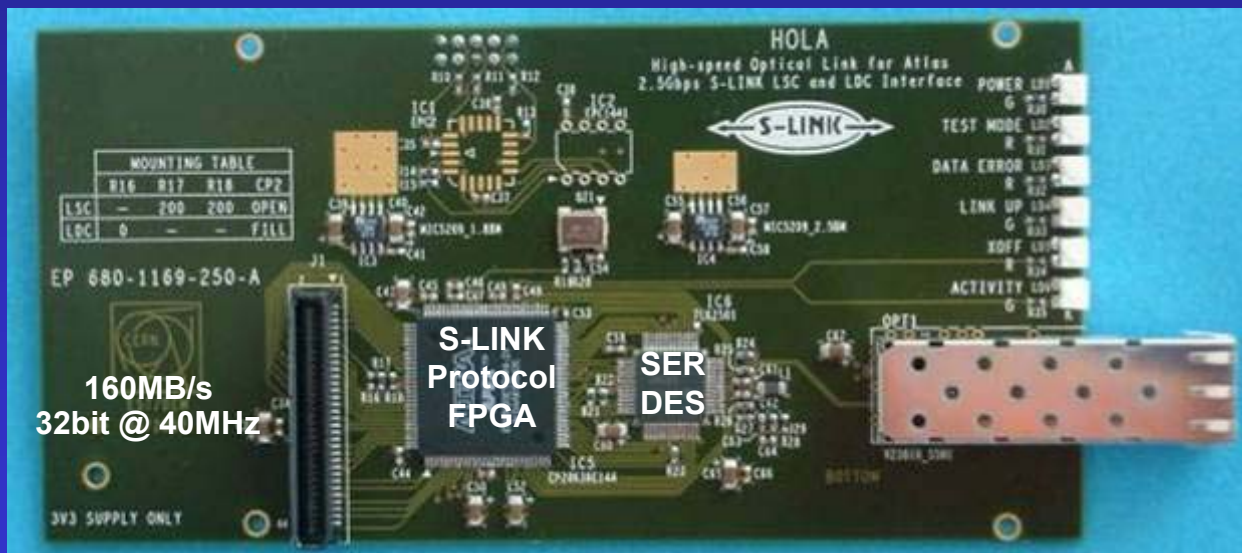
- Introduction
- Interface Card Hardware
- Firmware Description
- Software
- Performance Measurements
- Summary

Introduction

- **DAQ systems for current and future experiments depend on reliable high-speed data transmission**
- **S-LINK specification addresses this type of application:**
 - ▶ Point-to-point data link, bandwidth 160 MB/s (32-bit @ 40 MHz)
 - ▶ Flow control (XON/XOFF)
 - ▶ Error detection (e.g. CRC),
 - ▶ Self-test mode & return line signals
 - ▶ CMC mezzanine card format
- **ATLAS Read-Out Link (ROL)**
 - ▶ ROL implementation is based on S-LINK
 - ▶ Connects front-end electronics interface modules (Read-Out Drivers) to the Read-Out system (ROS)
 - ▶ ROS is based on commodity PCs and custom PCI interface cards (ROBin)
 - ▶ ~1650 ROLs will be used in ATLAS

ROL Source Card

- High-speed Optical Link for ATLAS (HOLA)
- Standard S-LINK mezzanine card
- Industry standard pluggable (SFP) 850nm F/O transceiver
- Serial link speed 2 Gb/s with 8B10B line encoding
- Low-power: ~2W typical



160MB/s
32bit @ 40MHz

S-LINK
Protocol
FPGA

SER
DES

Cage for SFP F/O
Transceiver

CMC mezzanine
Connector

Quad S-LINK PCI Interface (FILAR)

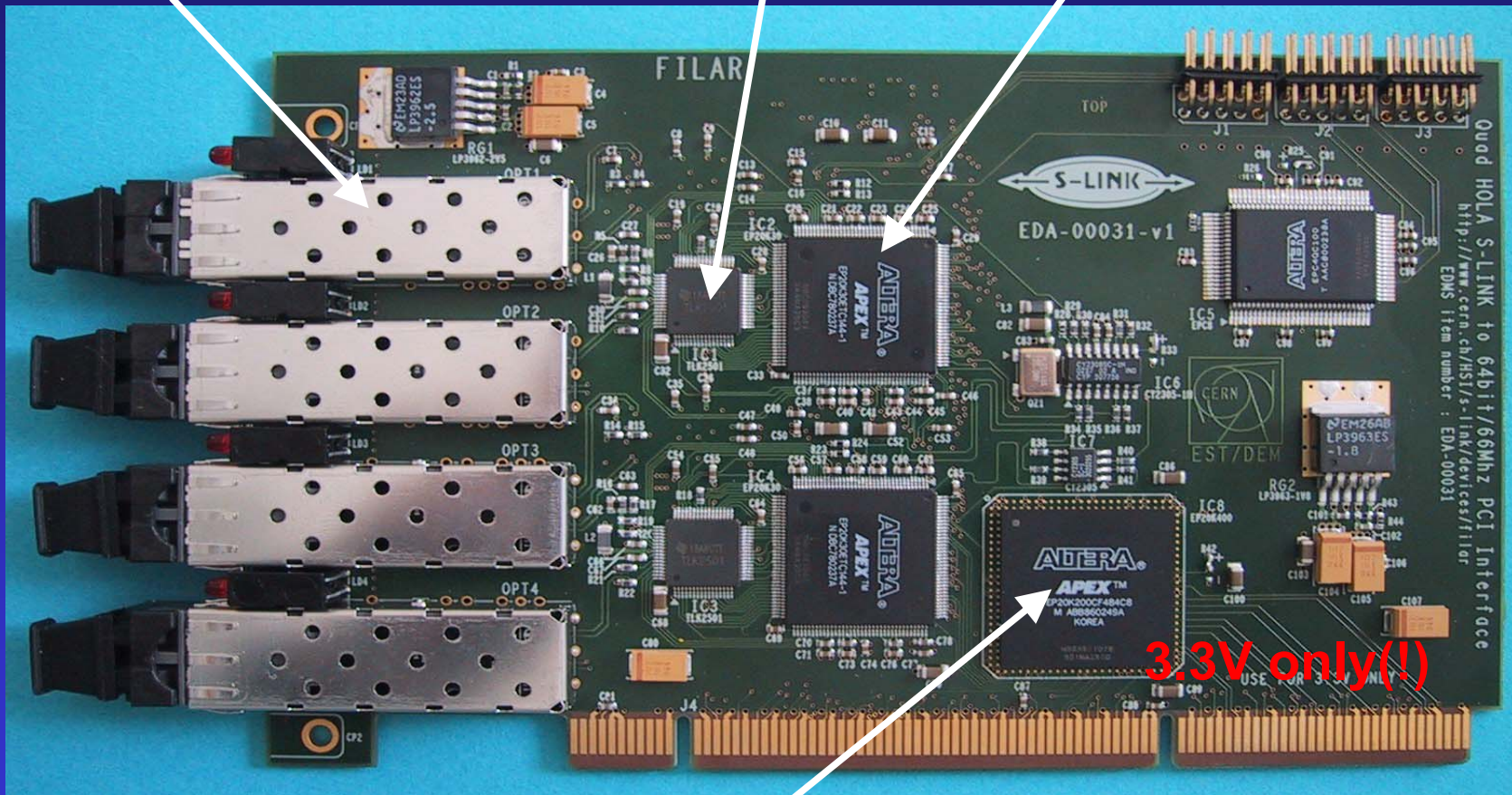
- **FILAR Features:**
 - ▶ Four 2 Gb/s HOLA link channels integrated on-board
 - ▶ 64-bit/66MHz PCI interface (3.3V slots only)
 - ▶ Move data between 4 link interfaces and the host PC memory
 - ▶ Based on S32PCI64 interface design: one slot for S-LINK mezzanine card
- **Applications: small readout systems for lab & test beam**
- **FPGA-based (in-system reconfigurable)**
 - ▶ PCI I/F implemented using a commercial PCI IP core
- **Firmware versions:**
 - ▶ Quad S-LINK receiver (S-LINK to PCI)
 - ▶ Quad S-LINK transmitter (PCI to S-LINK)
 - ▶ Quad S-LINK data source (for performance measurements)

FILAR Hardware

SFP Fiber Optic Transceiver

SERDES

HOLA Interface FPGA



PCI Interface FPGA

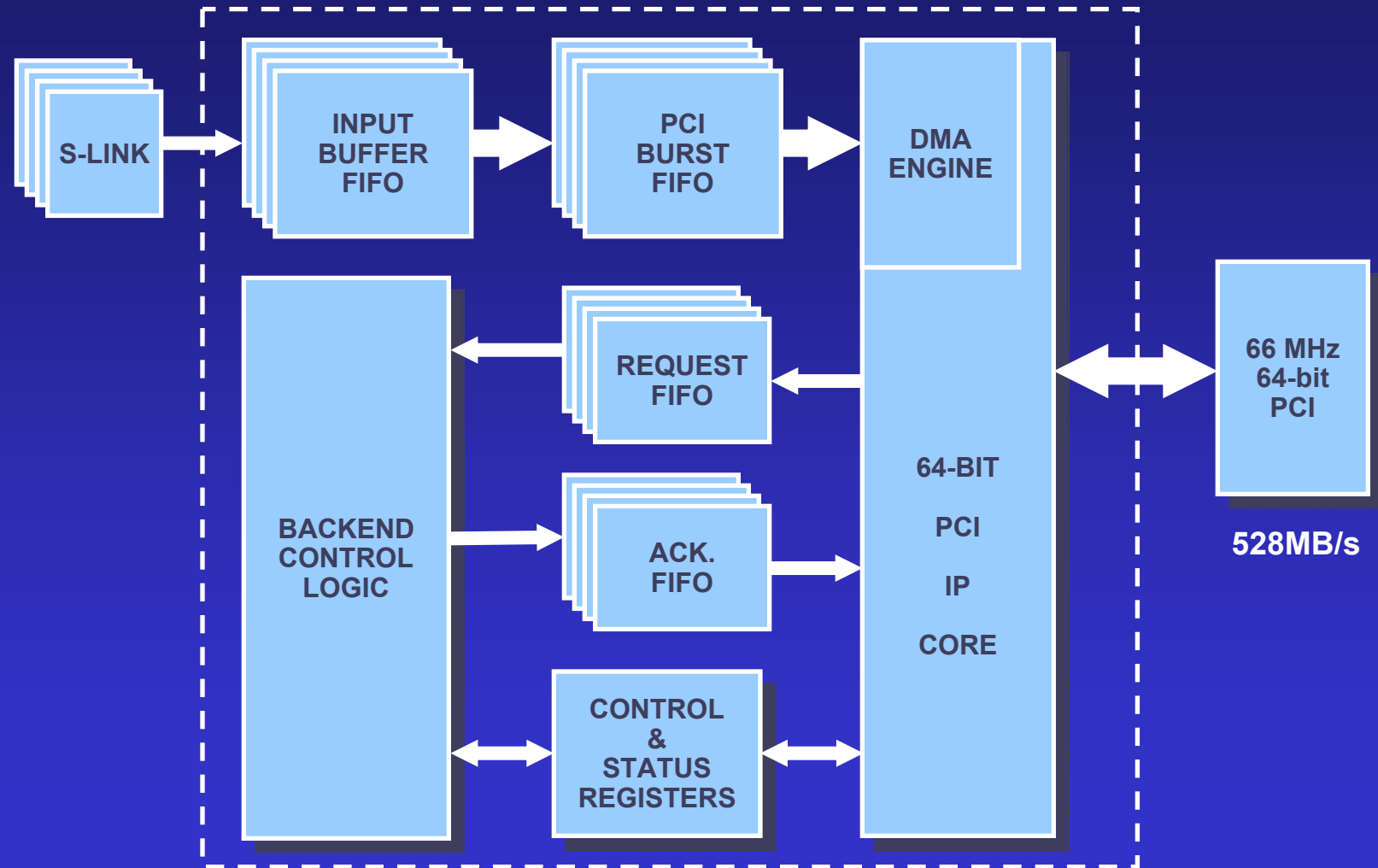
64-bit/66MHz PCI interface

3.3V only (!)

Receiver Firmware Operation

- **Host processor:**
 - ▶ 1) Fills a request FIFO on the interface card with addresses of free memory buffer pages
 - ▶ 5) Reads the results from the acknowledge FIFO and processes the data
 - **Interface card:**
 - ▶ 2) Transfers data fragments from S-LINK to host memory as bus master using PCI bursts of up to 1kB for maximum performance
 - ▶ 3) Stores length, status and control words for received fragments in an acknowledge FIFO
 - ▶ 4) Asserts an interrupt (optional)
-
- ```
graph LR; H1[1) Fills a request FIFO...] --> I2[2) Transfers data fragments...]; I4[4) Asserts an interrupt...] --> H5[5) Reads the results from the acknowledge FIFO...];
```

# Receiver Firmware Block Diagram





# Firmware Optimization

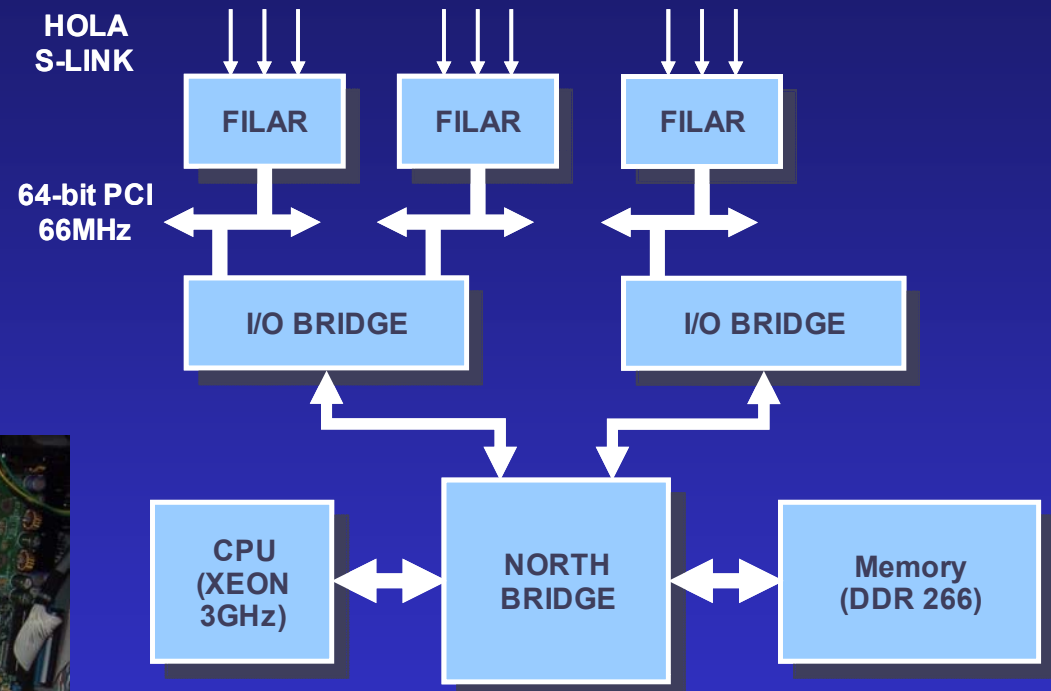
- **Single-cycles do not use the PCI bus efficiently**
- **Performance optimized version receiver firmware was developed (DMA protocol firmware):**
  - ▶ **Interface card transfers request and acknowledge data using DMA**
  - ▶ **CPU prepares a descriptor block with buffer addresses for one or more channels in system memory**
  - ▶ **Firmware fetches the block using DMA and fills the on-board request FIFOs**
  - ▶ **Firmware transfers a block with the length and status information from the acknowledge FIFOs to the system memory using DMA when a threshold is reached**
  - ▶ **Requires additional memory resources in the FPGA, only 3 receive channels can be implemented on the current hardware**
- **Reduces PCI bus overhead and CPU load**

# Software

- **FILAR software package:**
  - ▶ Linux device driver (loadable module)
  - ▶ Library provides easy to use programming API for applications
  - ▶ Test and benchmarking programs
- **Software written in C**
- **Separate drivers for the different receiver firmware versions**
- **Supports multiple channels & PCI cards**
- **Interrupt driven: device driver is called when a predefined number of fragments are available in any channel**
- **Code optimised for maximising throughput**
  - ▶ Manage the card with minimal attention from the application layer
  - ▶ Reduce the number of context switches
- **Fully integrated into the ATLAS DataFlow software**
- **Requires cmem driver/library for allocation of contiguous memory**
- **Similar package available for the transmitter firmware**

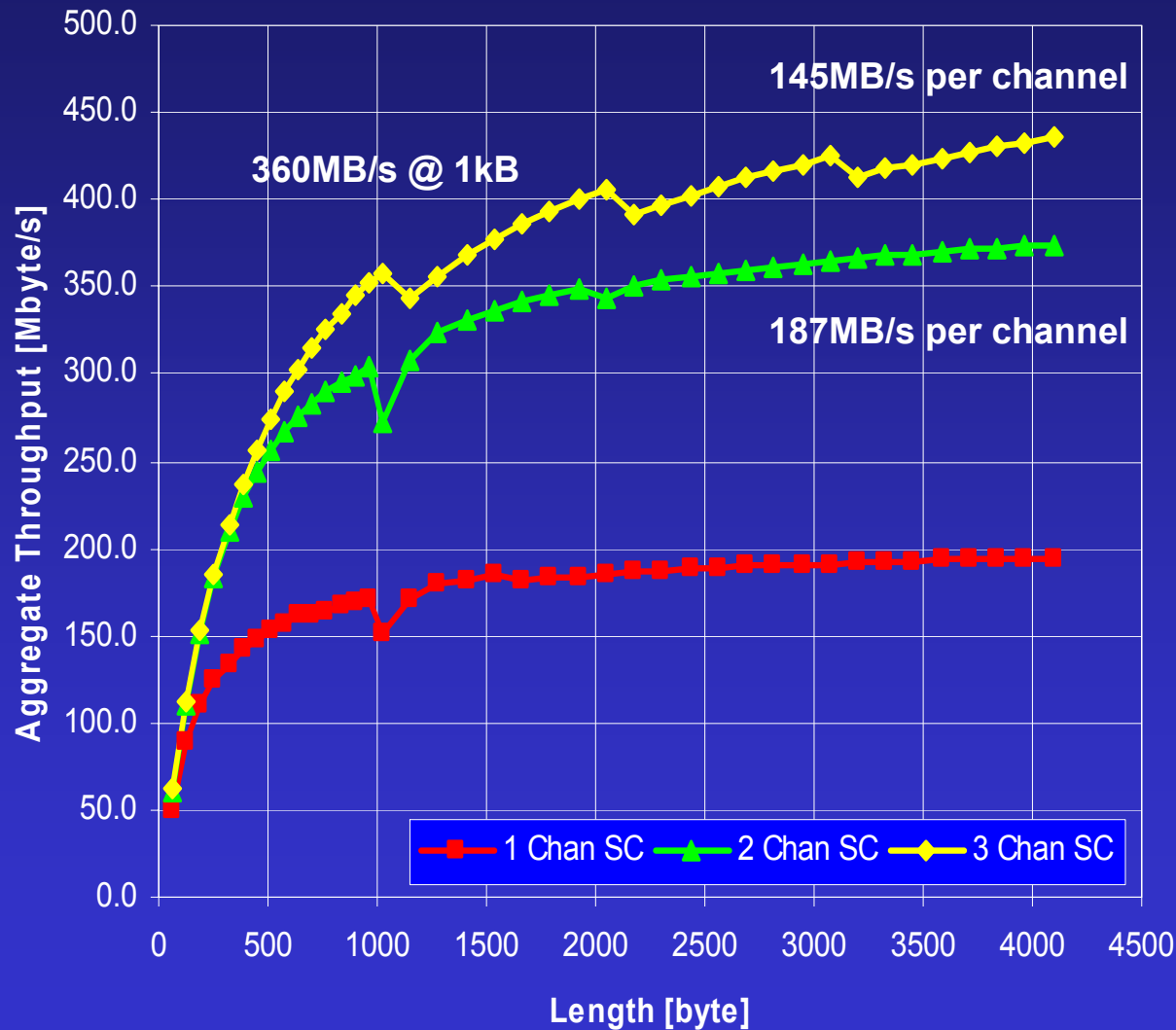
# Measurement Setup

- PC with Supermicro server motherboard (ServerWorks GC-LE chipset)
- 4 independent 64-bit PCI bus segments
- Intel Xeon CPU (3 GHz)
- S-LINK input channels driven by HOLA data sources



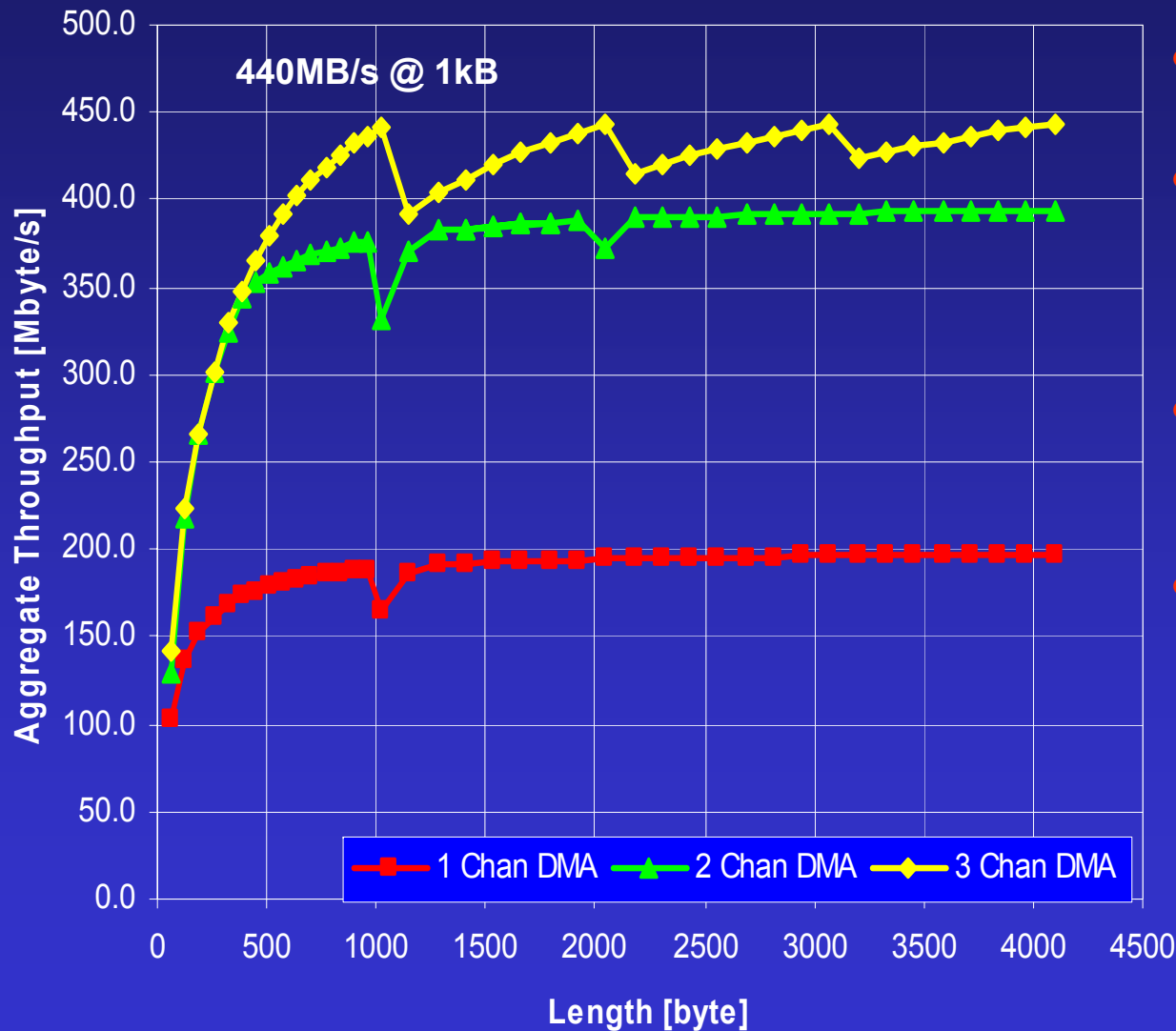
- Chipset architecture is important to obtain the maximum performance

# Performance: Single-Cycle Firmware



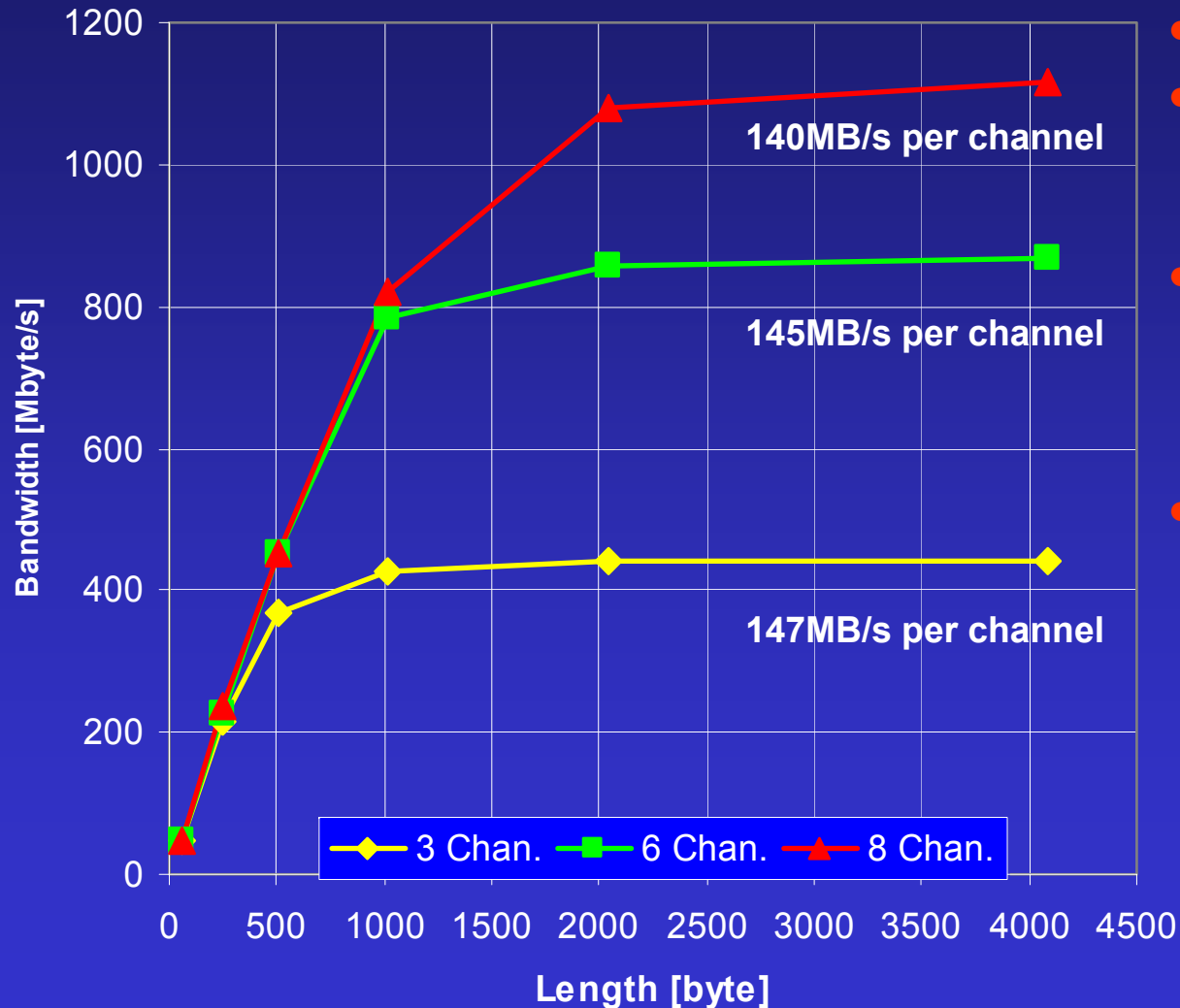
- FILAR receiver with SC firmware
- Sawtooth structure due to overhead for setting up a PCI burst (1kB)
- Performance for one channel is limited by link bandwidth
- Throughput with 3 channels is limited by PCI interface
- Maximum throughput is ~450MB/s

# Performance: DMA Protocol Firmware



- FILAR receiver with DMA firmware
- Better performance than SC firmware, in particular for short fragments
- 25% improvement for 3 channels at 1kB fragment length
- Performance for long fragments is similar for both firmware versions

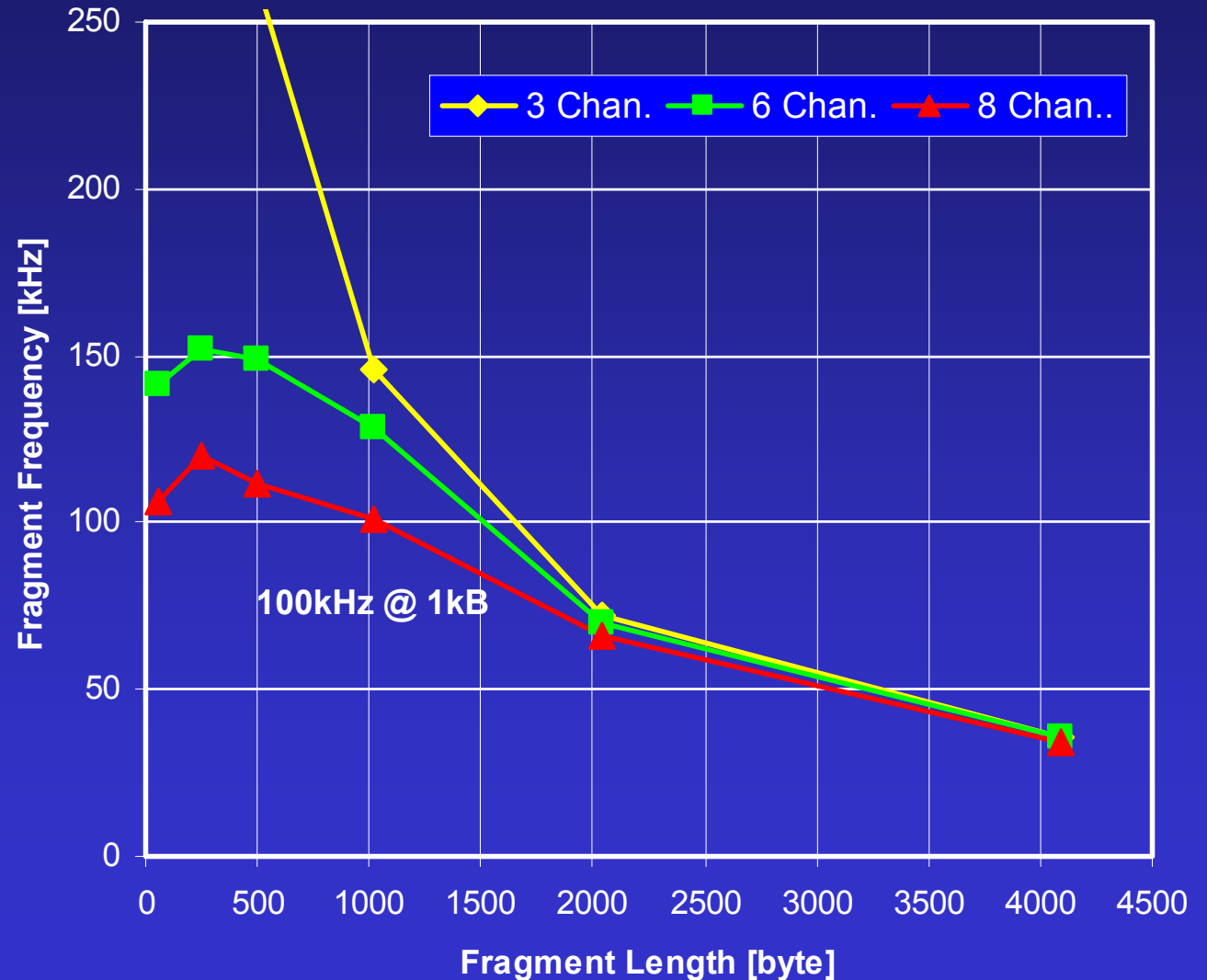
# Throughput: Multiple FILAR cards



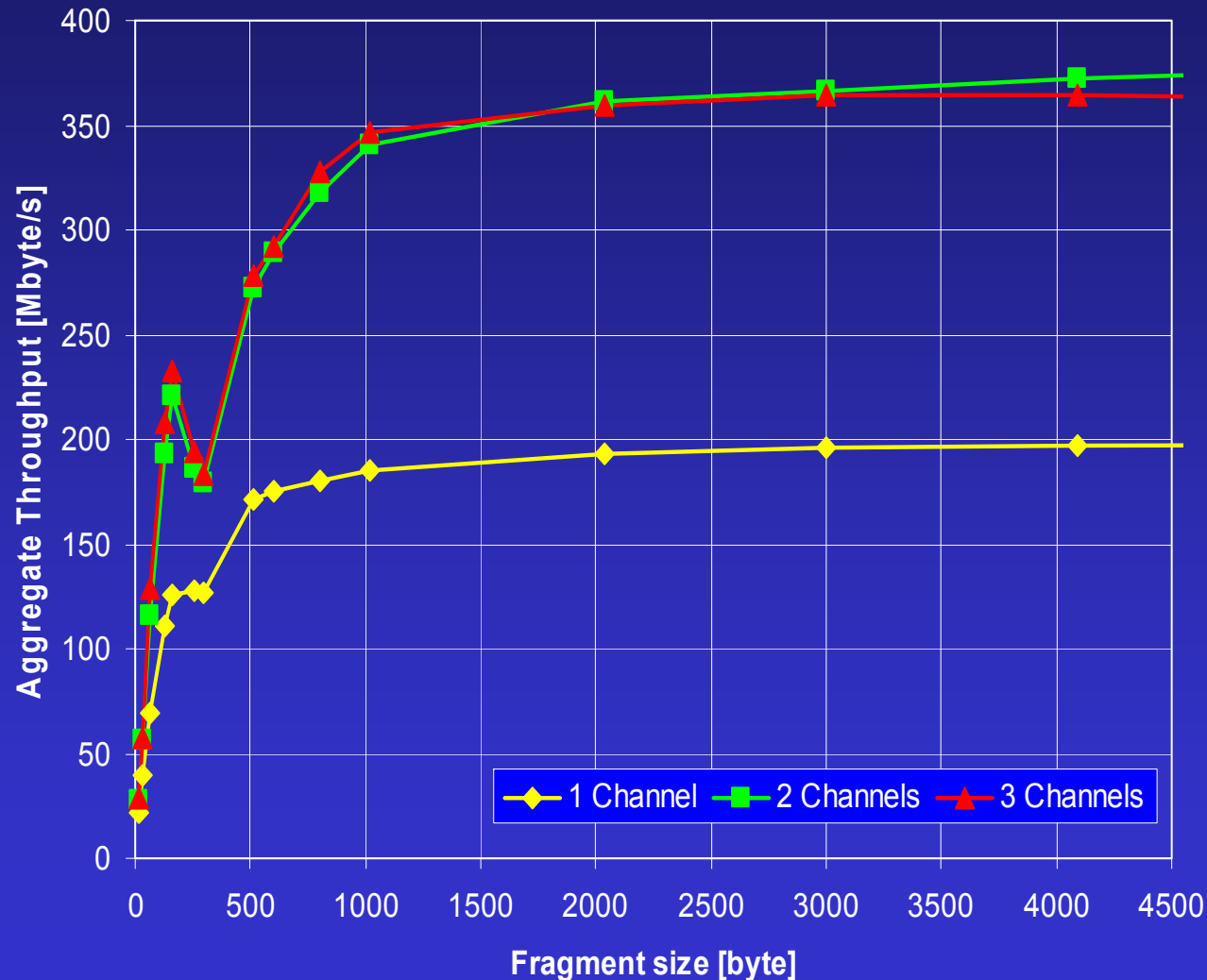
- DMA protocol F/W
- Maximum throughput of 1.1GB/s with 3 receiver cards
- Throughput scales with the number of channels for fragments of 2kB and more
- For fragments of 500B and less the system is rate limited

# Fragment Rate: Multiple FILAR cards

- Received data fragment frequency per channel vs. fragment length
- Fragment rates of 100kHz can be sustained with 3 cards for fragments of less than 1kB



# S-LINK Transmitter Performance



- Transmitter connected to a FILAR receiver in another PC
- PCI interface is saturated with 2 active channels
- Maximum throughput obtained is 360MB/s
- PCI memory read performance is not as good as write



# Summary

- **FILAR high-performance PCI interface card with 4 on-board 2 Gb/s S-LINK channels (HOLA) has been designed**
- **Quad S-LINK receiver, transmitter and data source firmware versions have been developed and optimized**
- **Software package with Linux device driver and API library are integrated in the ATLAS DataFlow software**
- **Maximum throughput for one receiver card ~450MB/s**
- **Aggregate data rate of > 1GB/s to system memory has been measured with 3 receiver cards**
- **Event rates of over 100kHz can be achieved for 1kB fragments**
- **FILAR applications and users:**
  - ▶ **Test readout of front-end electronics interface modules**
  - ▶ **ATLAS subdetector groups (LAr, SCT, TileCal, TRT, Pixel, LVL1 Calo), DAQ & ROBin, MDT chamber tests**
  - ▶ **Readout system for the ATLAS combined test beam**
  - ▶ **Stable design, ~50 cards produced so far**