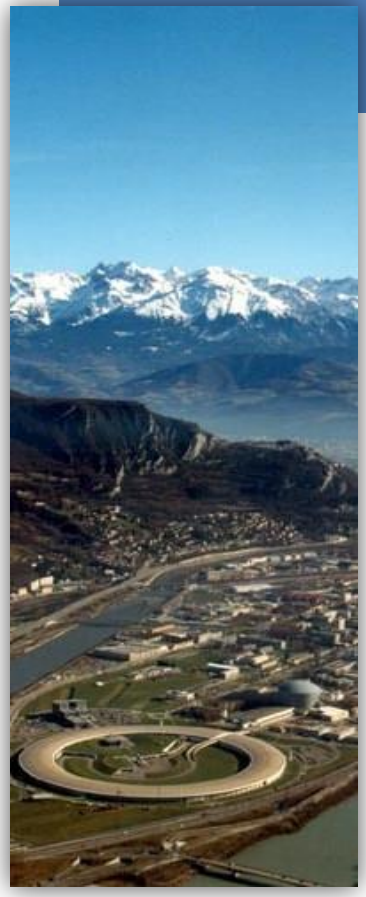


High Throughput Data Readout Architecture

Nicolas Janvier

Wassim Mansour

ESRF | The European Synchrotron

OUTLINE

- **Introduction**
- **DAQ framework Concept**
- **Advanced Demonstrator**
- **Current Activities**
  - **100 GbE**
  - **RASHPA Backend board**
- **Plans and Perspectives**

Introduction

Motivations:

- Current and upcoming advanced detectors produce very high data throughput.
- Industrial DAQ protocols do not handle the required data throughput
- The need for a standard tool
- Sufficiently generic and scalable
- Reusable in a wide range of high performance detectors

pco.edge

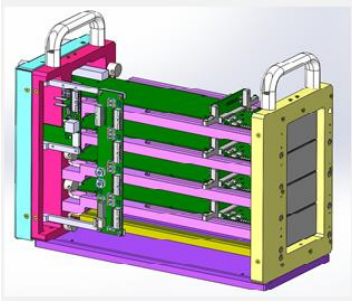
> 1 GByte/s

PSI/Eiger 2M

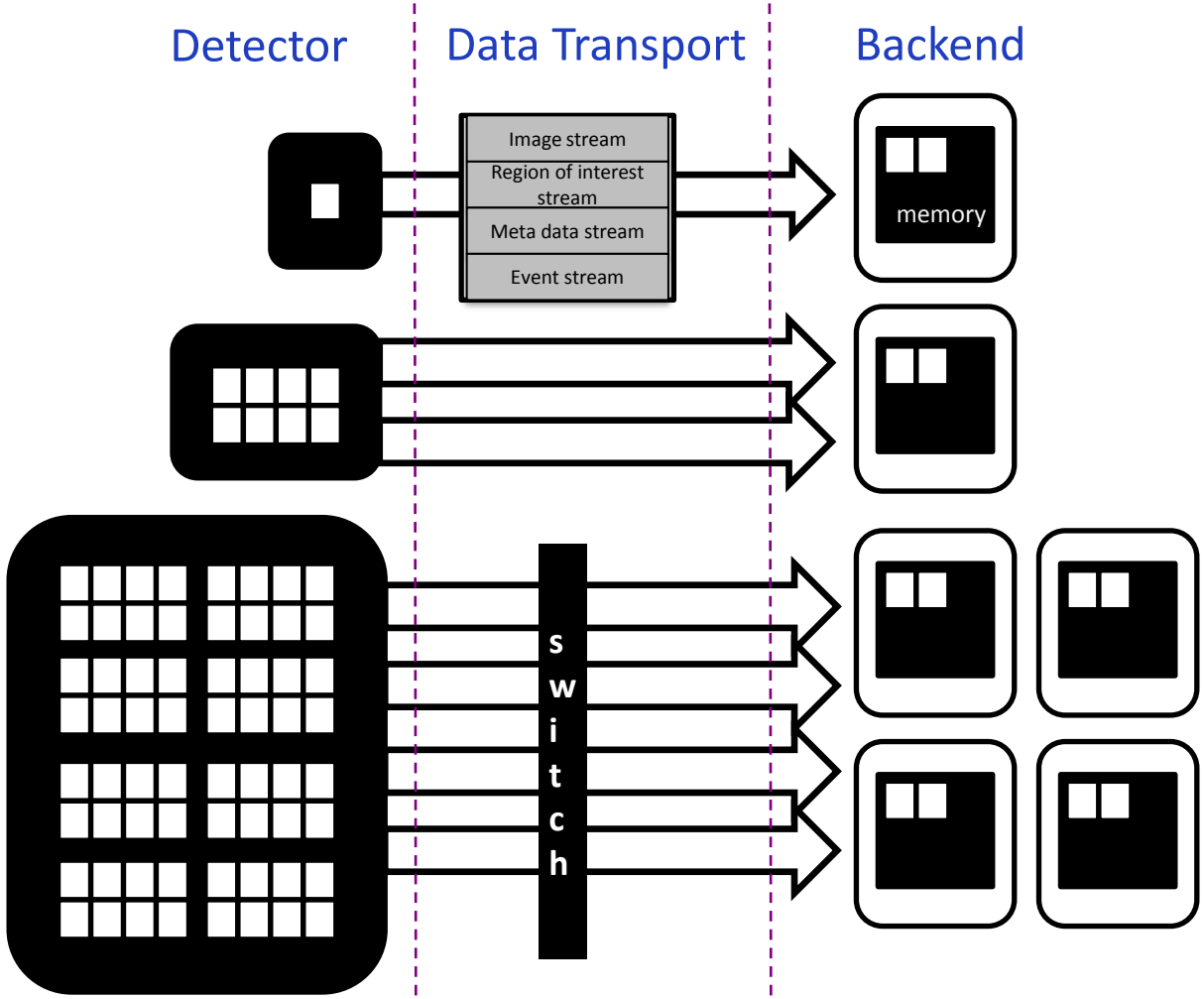
~ 8 GByte/s

Objective:

- Building such platform:
  - RASHPA: RDMA-based Acquisition System for High Performance Applications
  - Initiated in the frame of the European Project: CRISP
  - And continued in the frame of the European Project: Eucall



RASHPA Concept



High Performance

- RDMA to copy data from source to memory buffer
- Multiple and simultaneous data streams

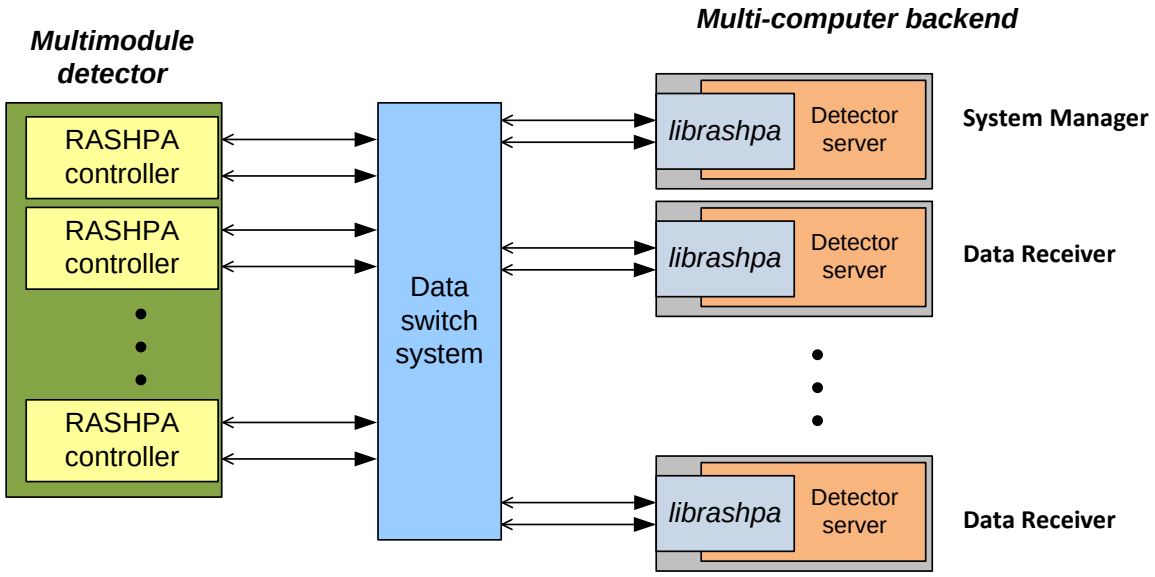
Scalability

- Support single or multiple sensors and workstations
- Adjustable bandwidth

Flexibility

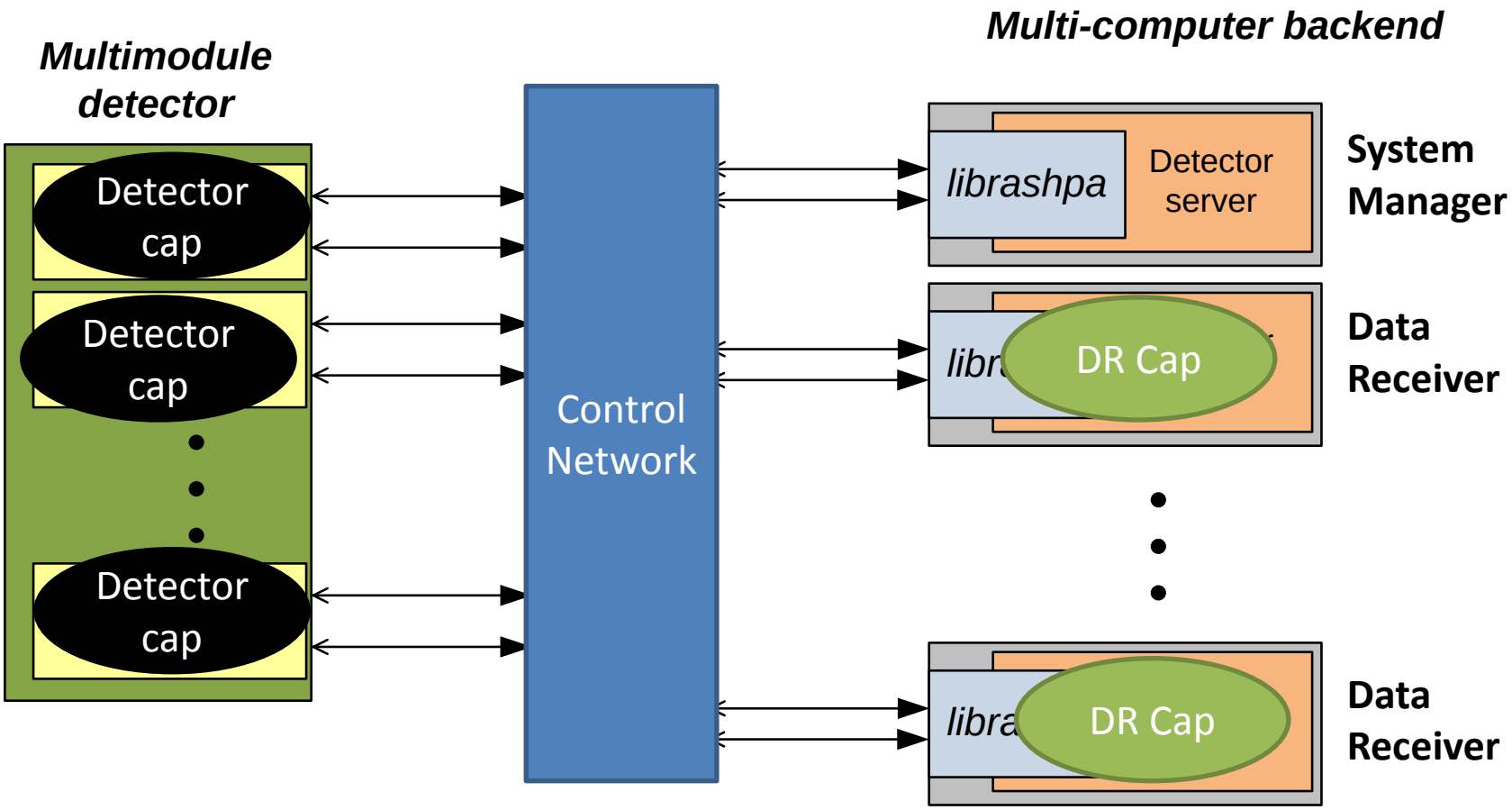
- Adapt to detector geometry (offset, stride, …)
- Configurable data streams (image, ROI, event)
- Data transport layer (PCIe over Cable, GbE, …)

RASHPA Concept

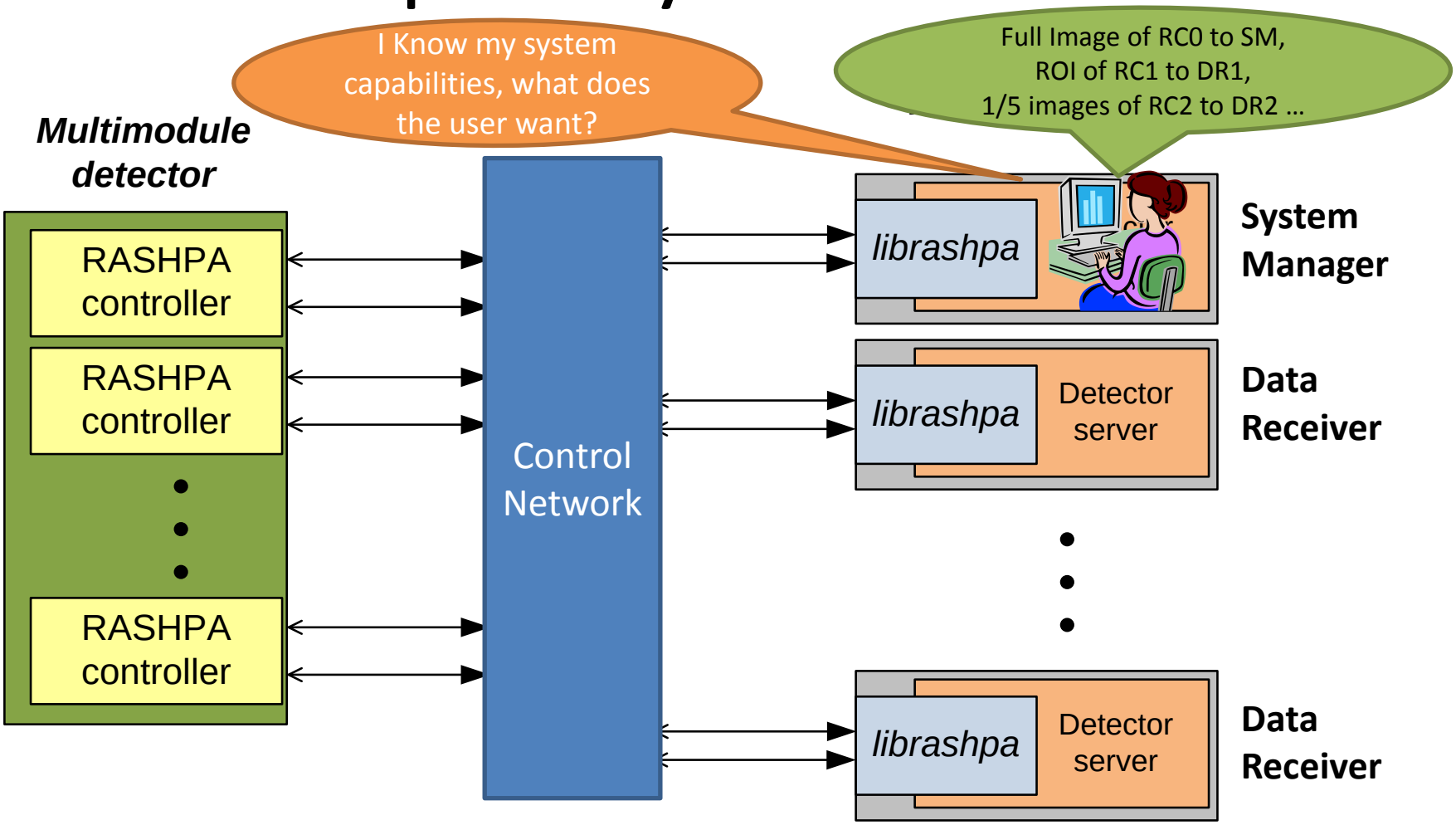


- RASHPA allows detectors to push data into backend computers
- Usual destinations: Memory Buffers
- Other destinations: GPU, Coprocessors, disk controllers …
- BCs receiving data from DM are called data receivers (DR)
- BC that configures and initializes RASHPA is called System Manager (SM).
- Both managed by C-Library: Librashpa

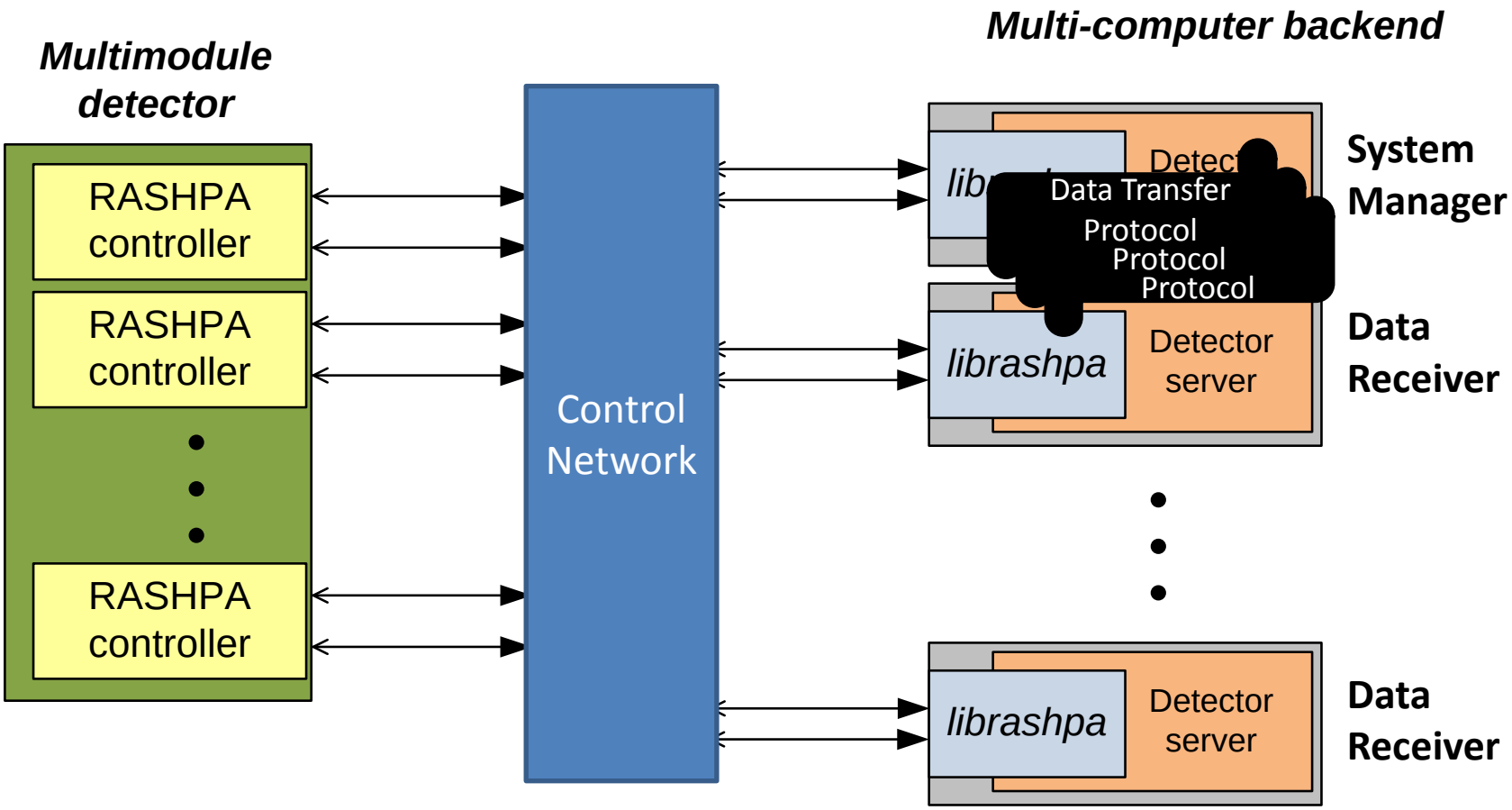
Capability retrieval



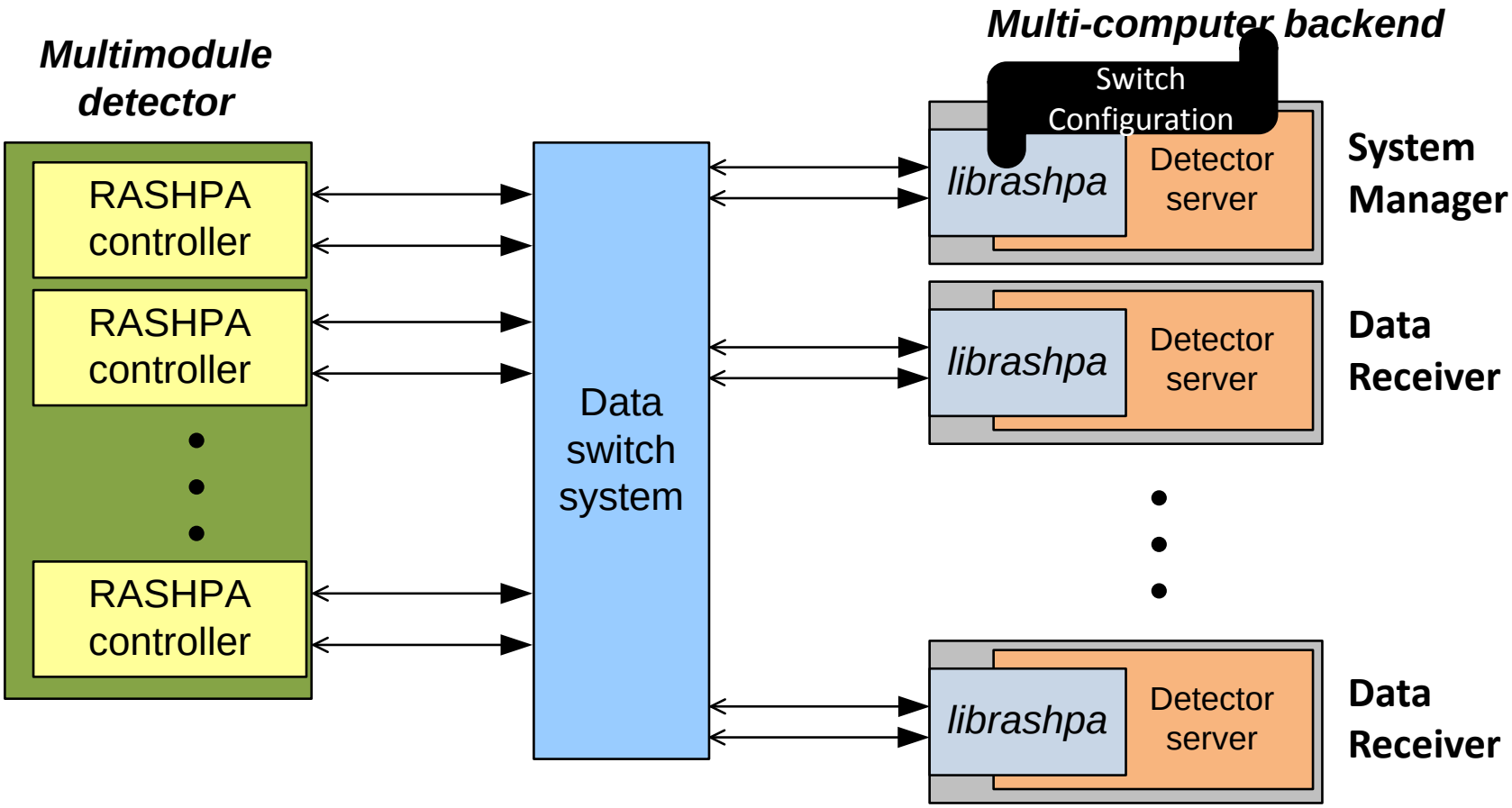
Capability retrieval



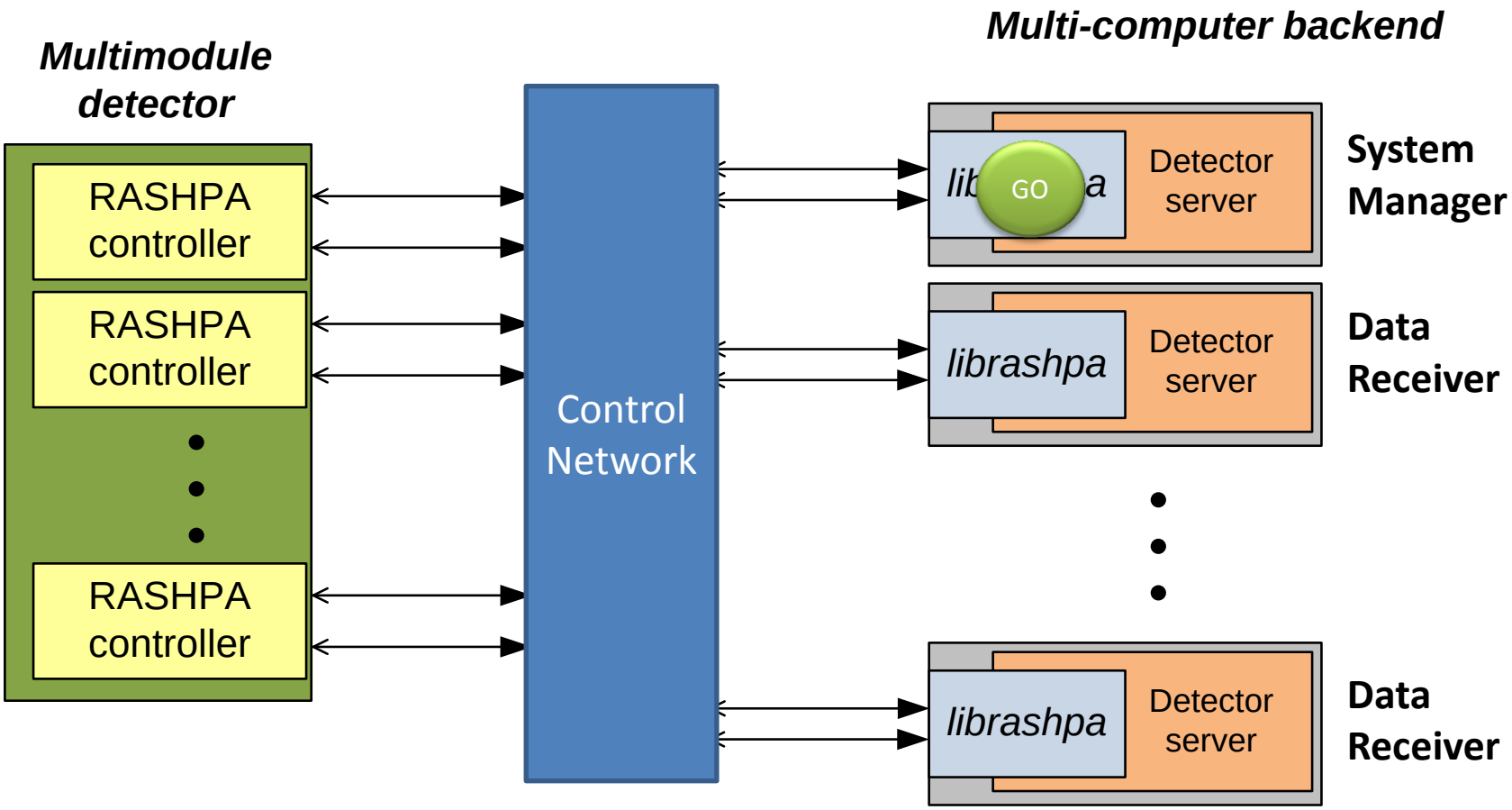
Capability retrieval



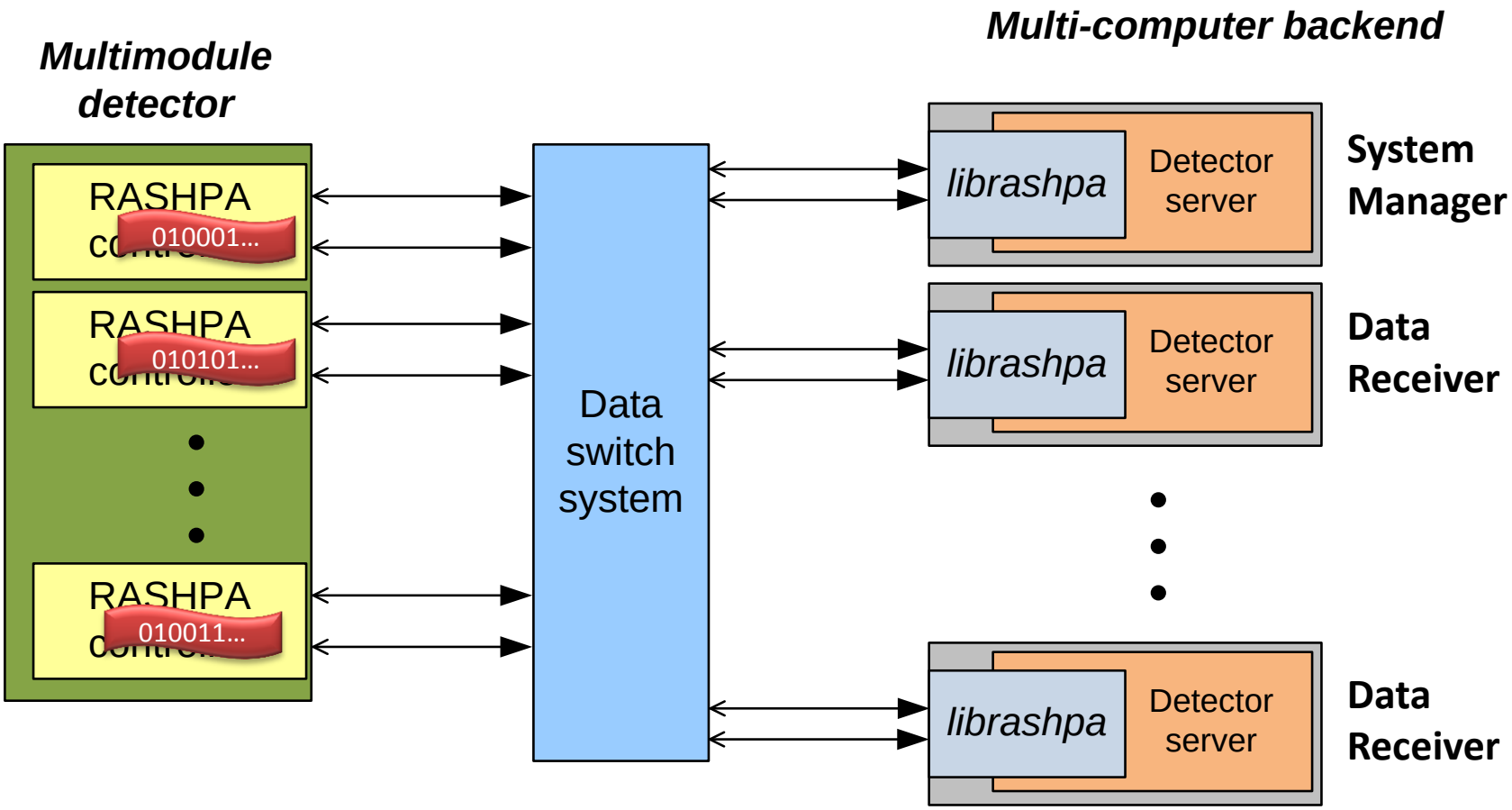
Capability retrieval



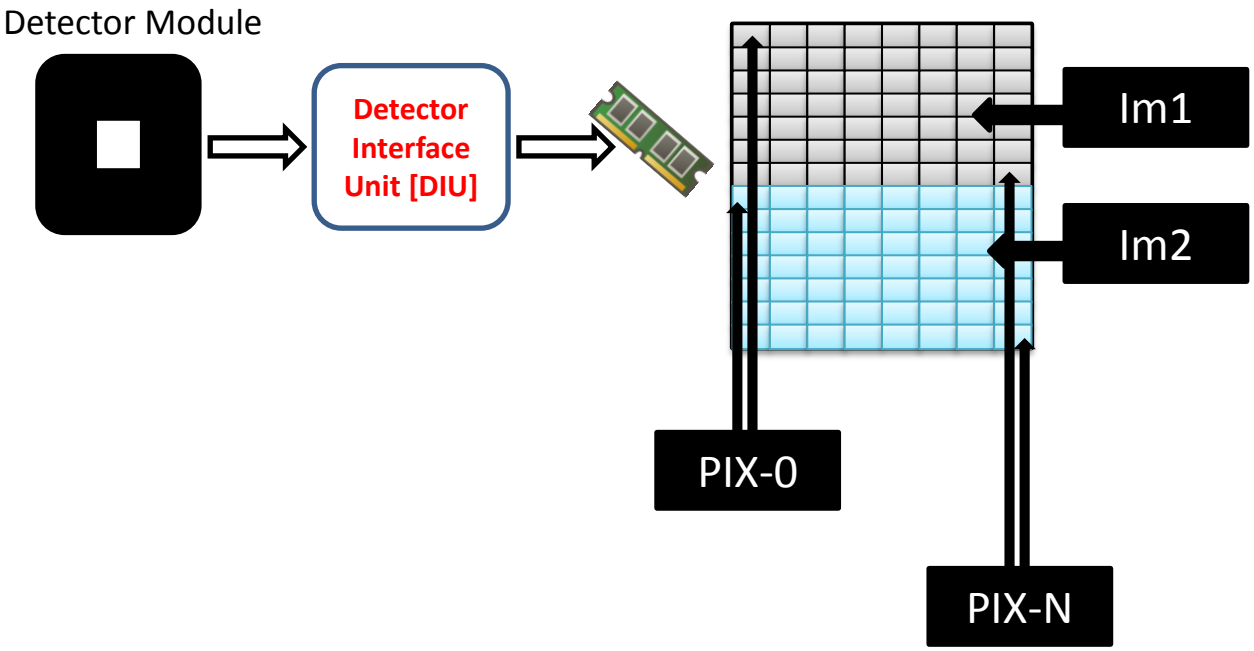
Capability retrieval



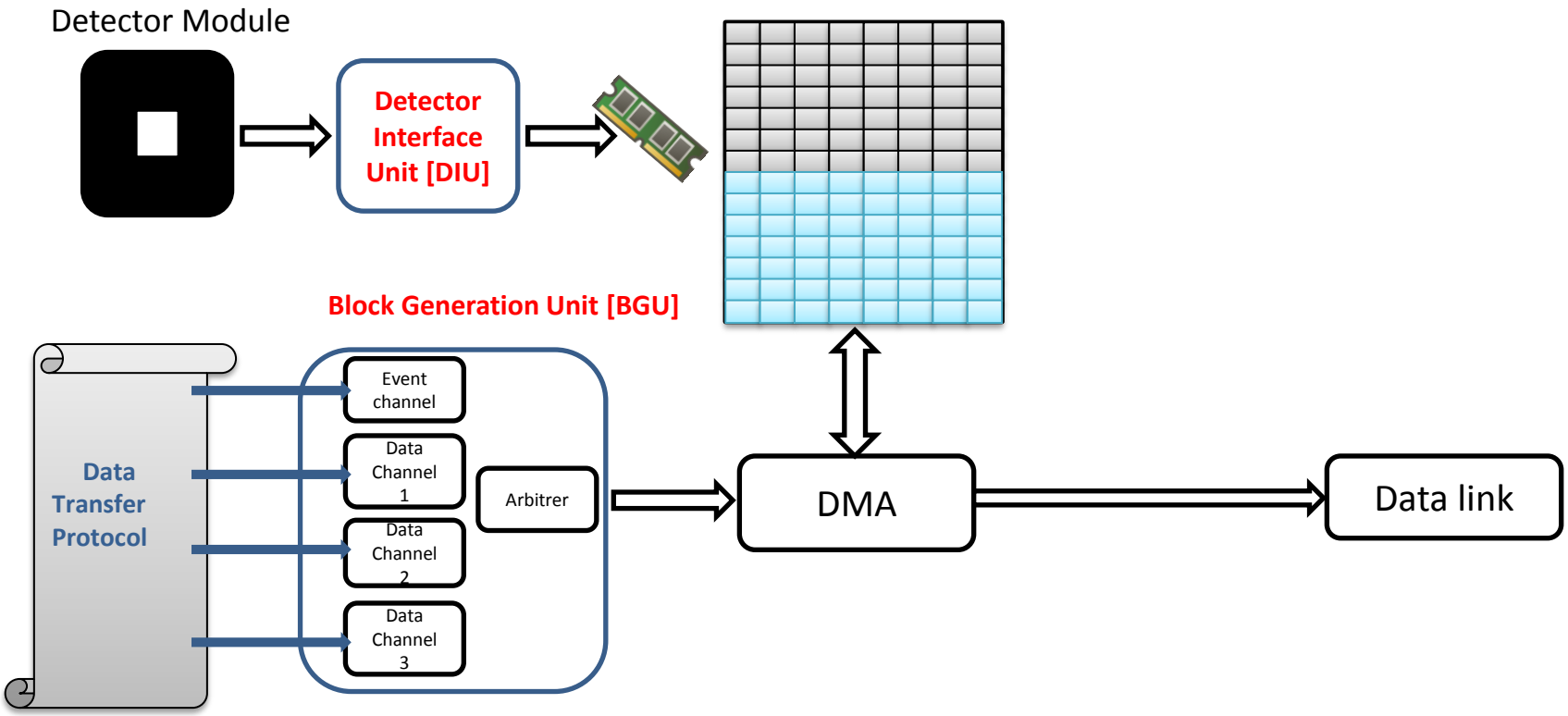
Capability retrieval



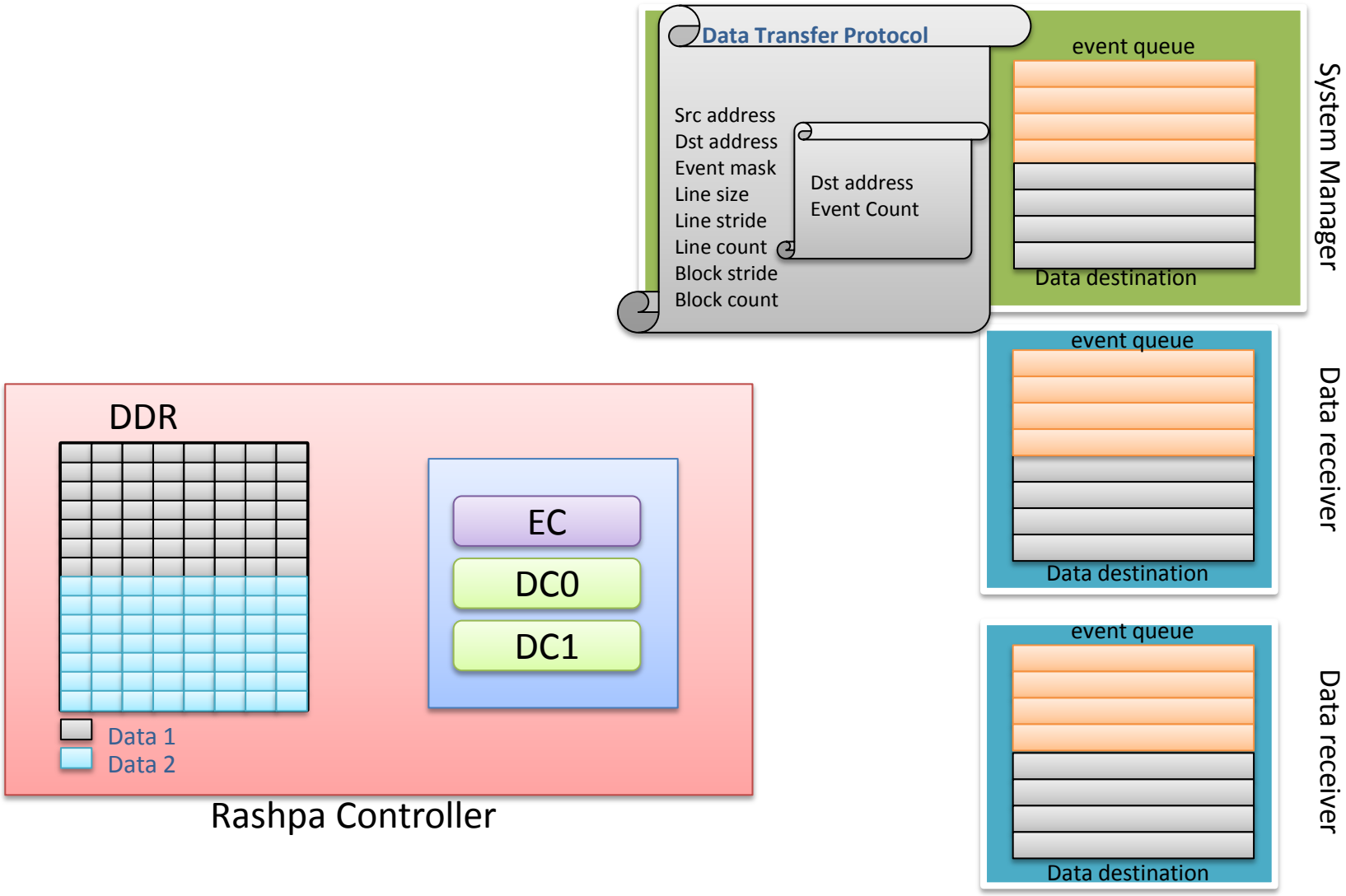
Data Flow


Detector Module
Detector Interface Unit [DIU]
Im1
Im2
PIX-0
PIX-N

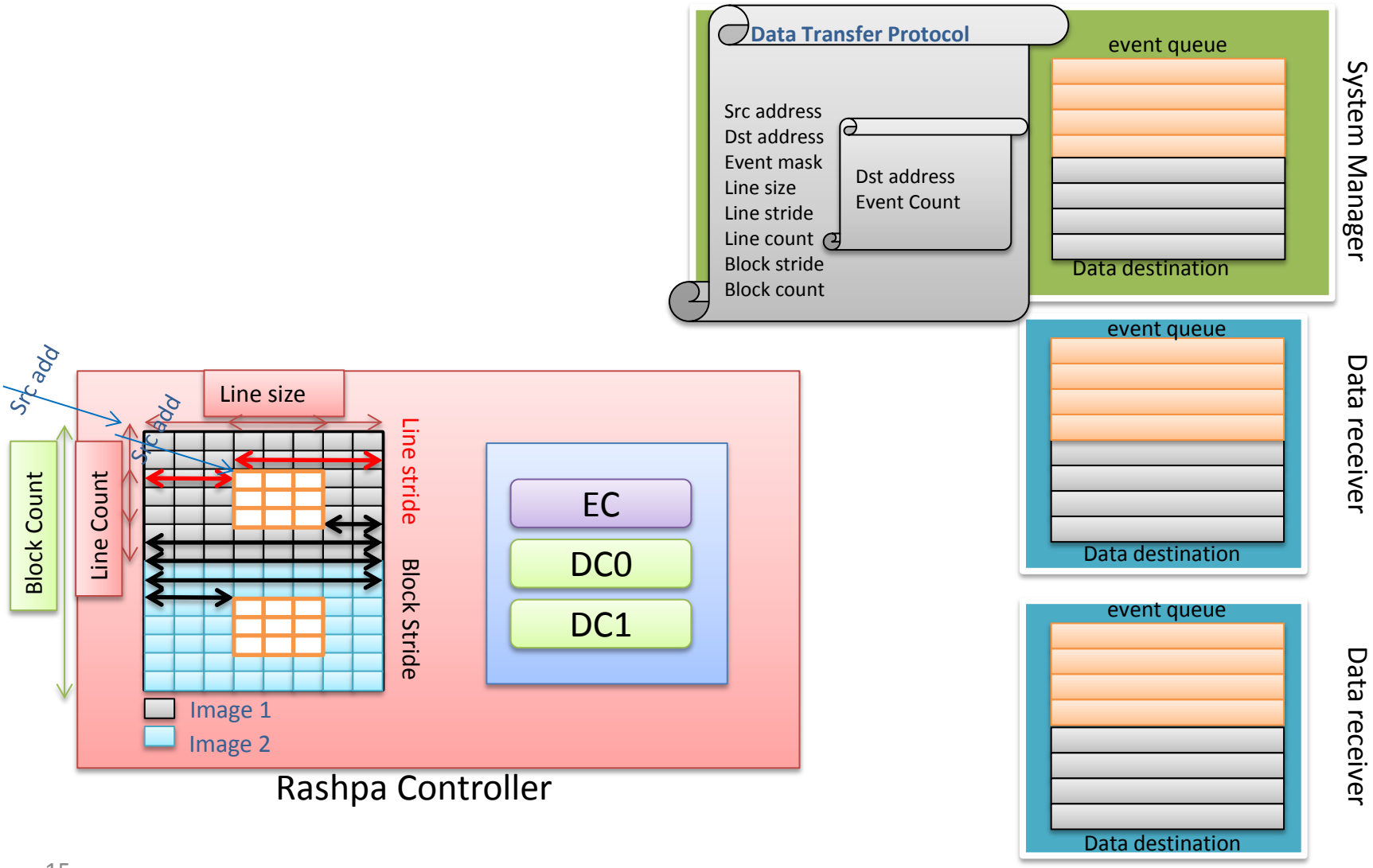
Data Flow


Detector Module
Detector Interface Unit [DIU]
Block Generation Unit [BGU]
Data Transfer Protocol
Event channel
Data Channel 1
Data Channel 2
Data Channel 3
Arbitrer
DMA
Data link

Rashpa Configuration



Rashpa Configuration



Data Link Selection

Link Selection

- PCIe over cable was selected for data transfer
  - Native integration
  - RDMA Capable
  - Limited Capabilities
  - Lack of Standardization

One Stop Systems switch

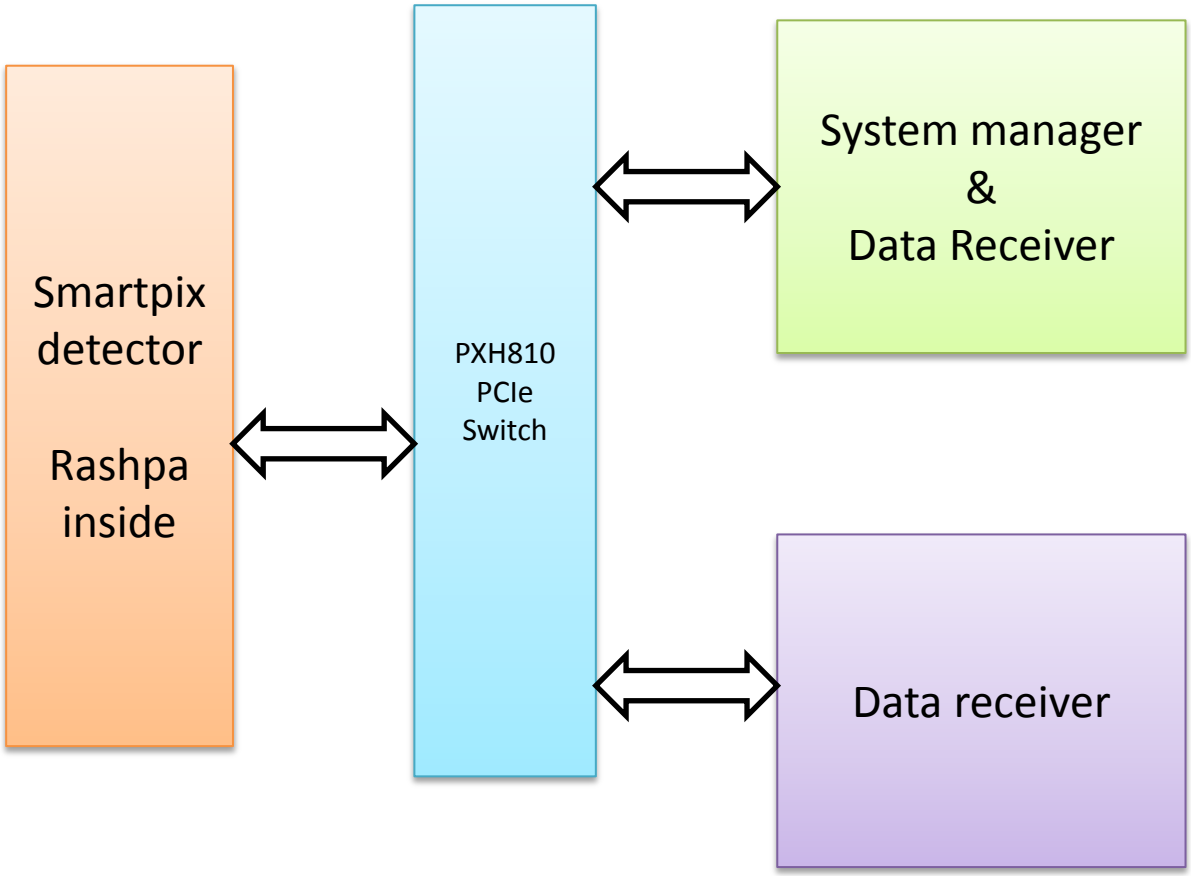
Dolphin IXS600

Supported link

One Stop Systems cable adapter

Dolphin PXH810

RASHPA Advanced Demonstrator

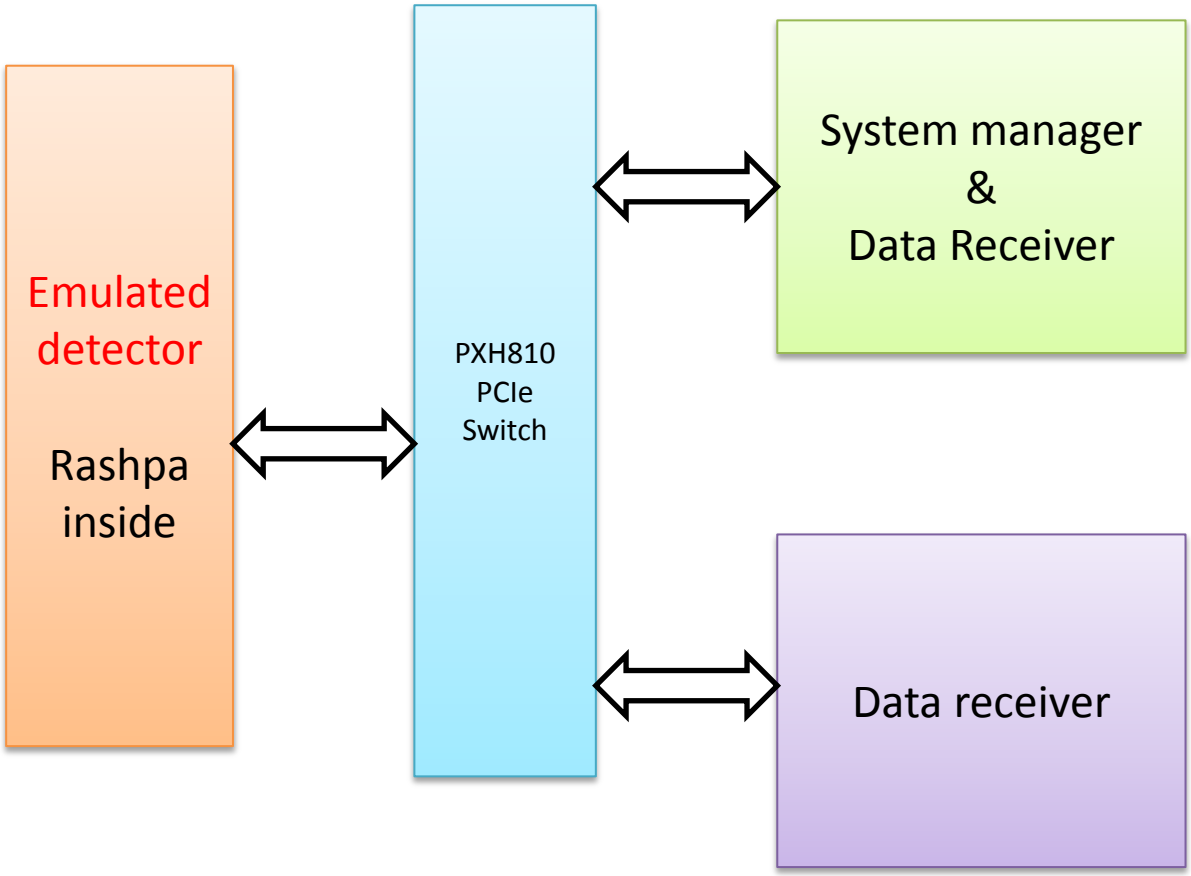


RASHPA implemented on PFPKX7 commercial board from techway and integrated to smartpix detector

The system manager acts as a data receiver and implement the PCIe switching network.

An industrial PC is used as a second data receiver.

RASHPA Advanced Demonstrator



RASHPA implemented on PFPKX7 commercial board from techway and integrated to smartpix detector

The system manager acts as a data receiver and implement the PCIe switching network.

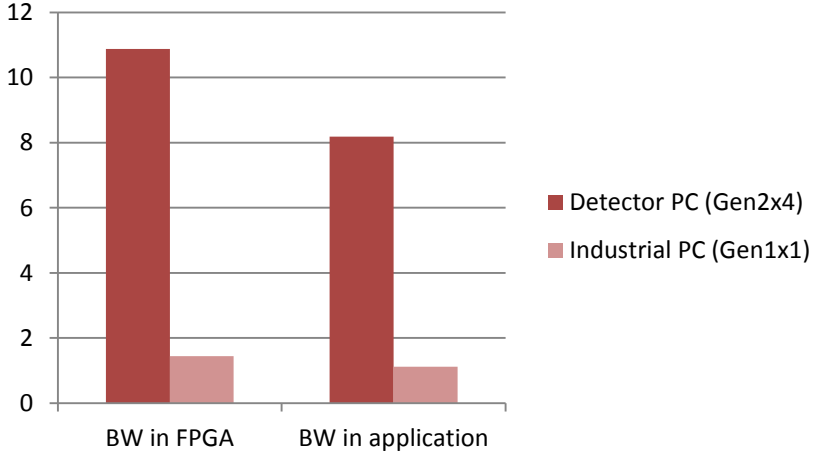
An industrial PC is used as a second data receiver.

Smartpix detector is used as the XRAY detector for testing but replaced by an emulator to get significant data.

Experimental Results

Performance Tests

	Bandwidth within the FPGA	Bandwidth when data available to the application
Detector PC Gen2x4	10.88Gbps 68%	8.18Gbps 51.15%
Industrial PC Gen1x1	1.44Gbps 71.99%	1.12Gbps 56.28%



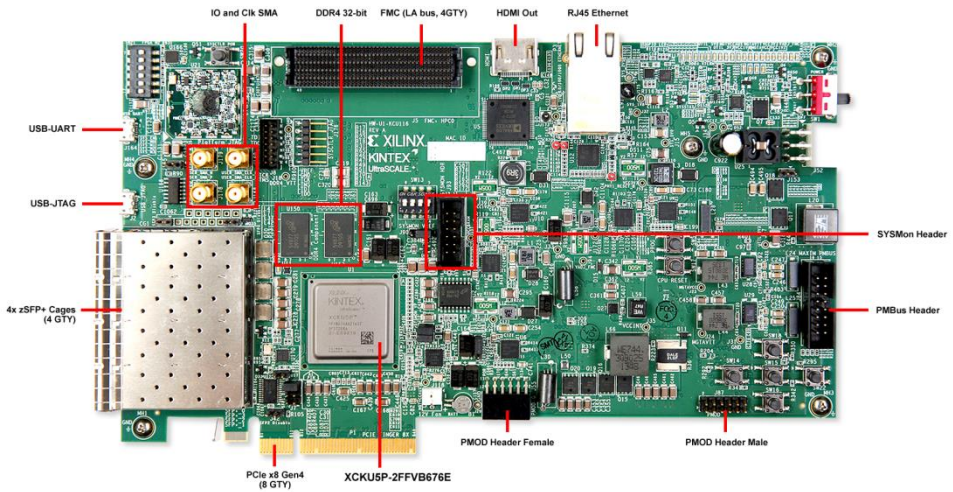
100G Ethernet as Data Link

- **100G Ethernet**
  - Free Hard IP core on ultrascale+ families
  - Ethernet does not have native RDMA Capabilities

- **What RDMA protocols to use?**
  - Candidates could be:
    - RoCEv2: RDMA over Converged Ethernet (UDP based)
    - iWarp: internet Wide-area RDMA Protocol (TCP based and now UDP)
    - An ESRF customized protocol

Mellanox ConnectX-4

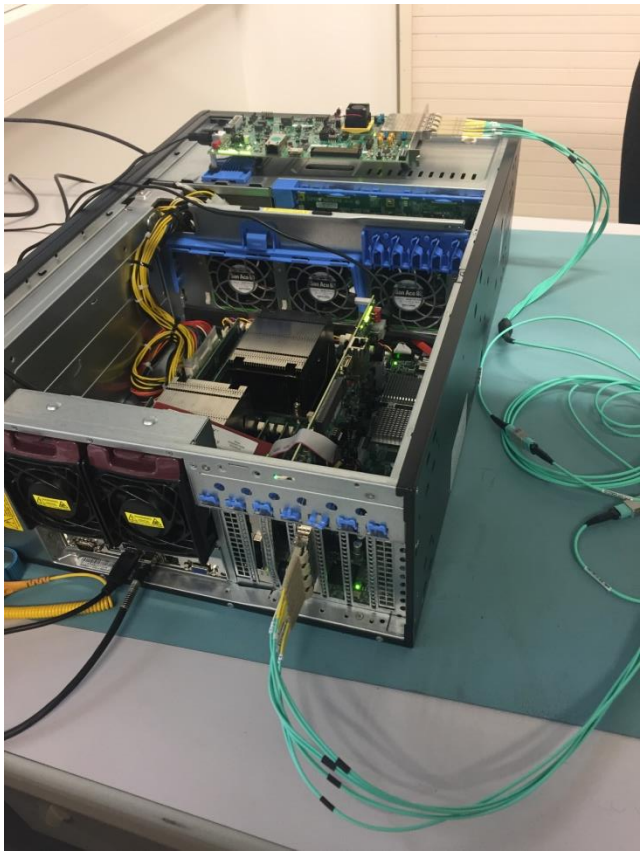
KCU116 Xilinx development kit

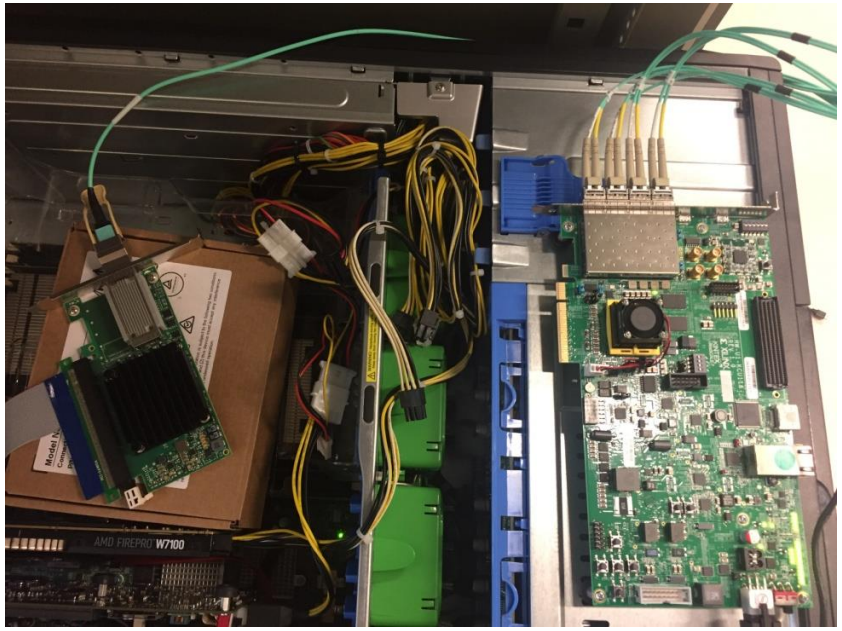


Lab Tests

- **Successful connection schemes**
  - Network Adapter to Network Adapter (Win10, Linux)
  - Board to board (KCU116)
  - Board to Network Adapter (KCU116 to Mellanox)

  No Switch test yet
- **UDP Protocol**
  - C and python applications
  - Wireshark





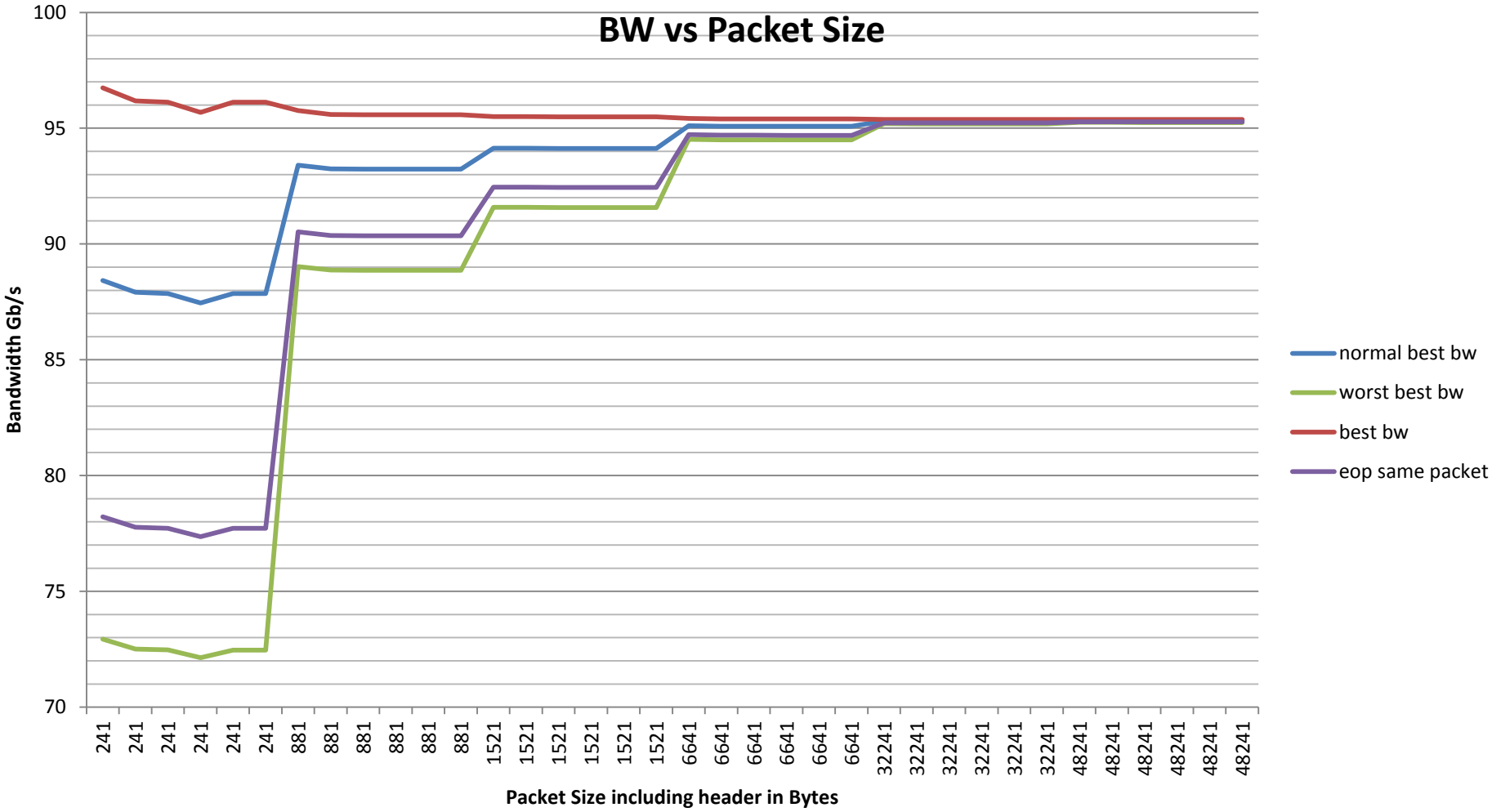
RoCE-V2 has been successfully tested for

- Soft-RoCE (over ESRF network)

- RoCE-V2 (Mellanox/FPGA to Mellanox)

(Thanks to Raphael Ponsard)

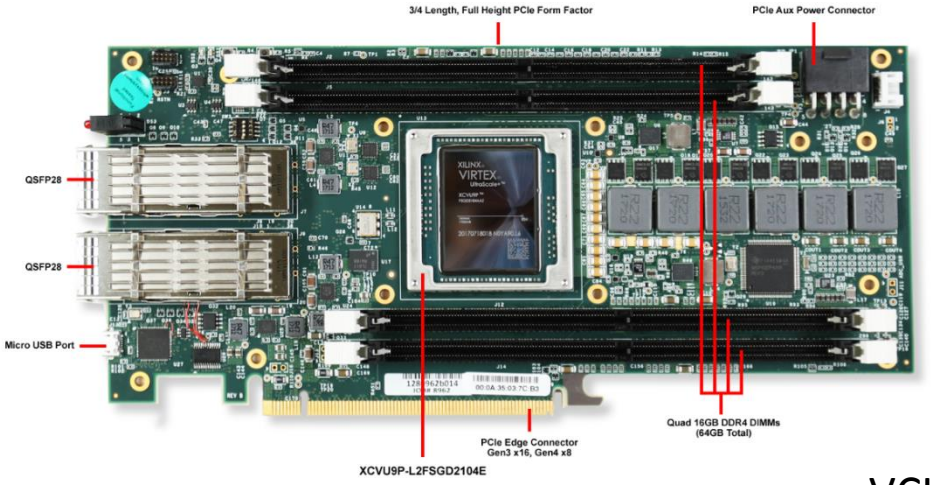
Bandwidth Tests



Rashpa Backend Board

- **Backend board features**
  - Virtex Ultrascale + VU9P
  - High density and parallel memory banks
  - Two QSFP28 100G Interfaces
  - PCIe Gen3x16 or Gen 4x8 edge connector

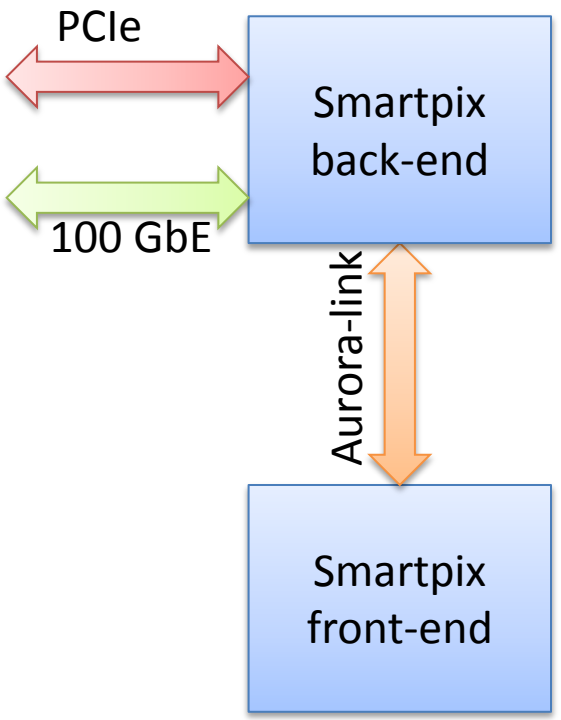
XpressVUP from reflexces



VCU1525 from Xilinx

Integration with Smartpix detector

- **This board can serve as a backend for the ESRF smartpix detector in its current version**
  - Link between smartpix front-end and this board is Aurora
  - Two data link types can be used
    - PCIe Gen3x16
    - 100G ethernet
  - Smartpix and Rashpa control
    - PCIe Gen3x16

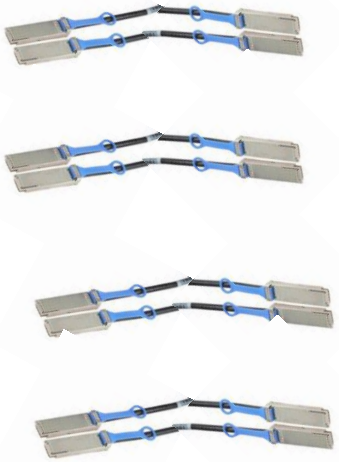


Perspectives

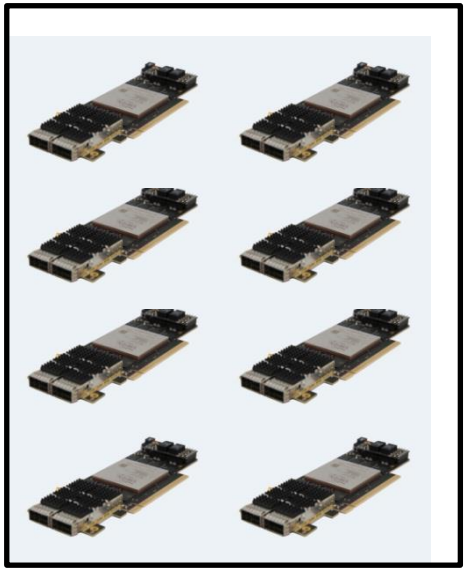
- **This board can also be used racked to serve as a rashpa box for other detectors (such as the EIGER or JUNGFRAU detector)**

Eiger detector

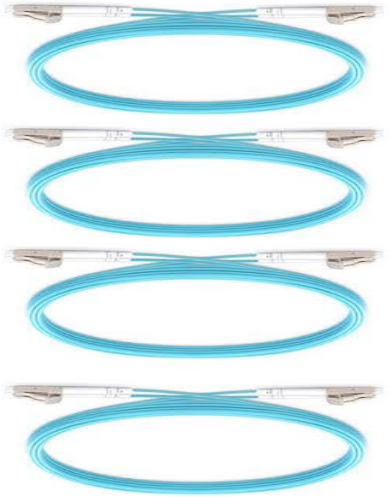
Gb Ethernet



Rashpa Box



100 GbE link

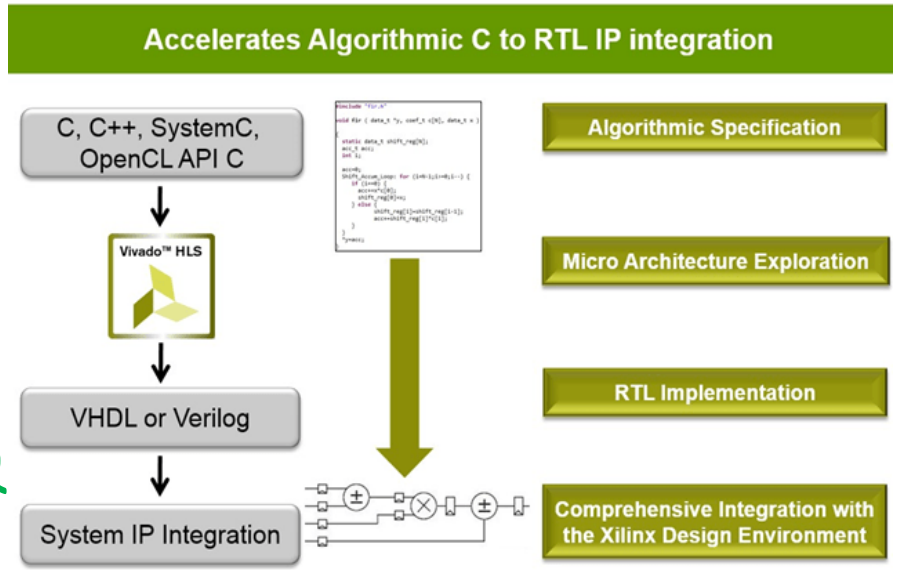


And [maybe] More!!!

Today at ESRF: data is streamed through a LIMA based 'detector server'

All image manipulation is 100% software based

→ Hardware assisted DAQ and image manipulation (RASHPA)



THANK YOU!