

The background of the slide is a complex, abstract network diagram. It features a dense web of thin black lines connecting various nodes, some of which are represented by small circles. A prominent feature is a thick, bold black line that forms a large, irregular loop on the left side of the image. The overall aesthetic is technical and futuristic, suggesting a high-speed or interconnected network.

RapidIO as a multi-purpose interconnect

CHEP, October 2016
Sima Baymani



IDT

- › **Integrated Device Technology**
- › **CERN openlab partner since 2015**
- › **Products target hard real time systems**



CERN openlab

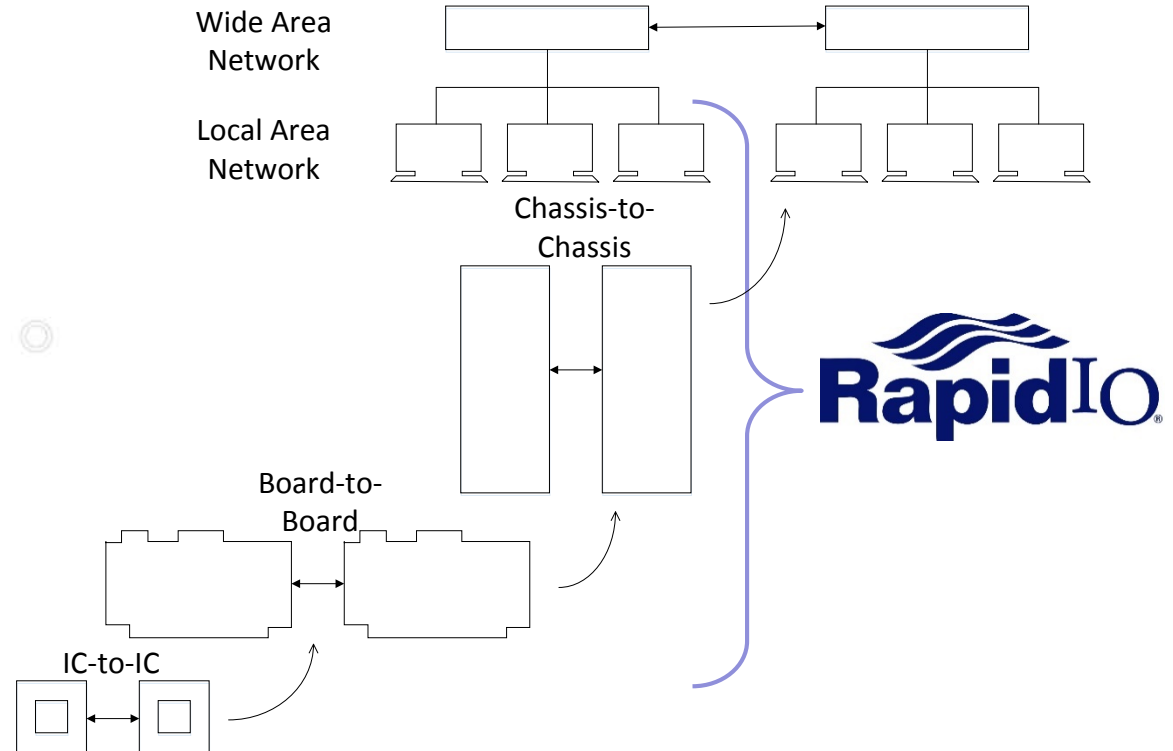
CERN openlab

- › **Public-private partnership**
- › **R&D on new technologies**
- › **More about CERN openlab at poster 318 on Thursday**



- › **Introduced in 1997**
- › **From front side bus to system level interconnect**
- › **Open standard – rapidio.org**
- › **Meets real time needs as well as scalability**

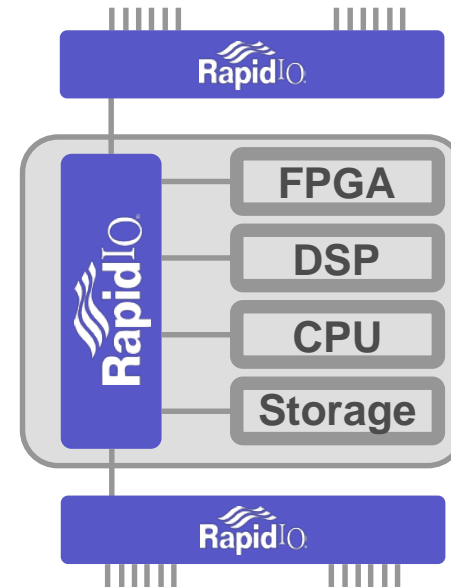
RapidIO Introduction



- › **Combines scalability with low latency**
 - Switch latency ~100 ns
 - Memory to memory latency < 1 μ s
- › **CPU offload**
 - Error recovery in physical layer
 - Protocol stack processed in HW
- › **Operations include**
 - Read/write (remote DMA)
 - Messaging (4KB)

Features

- › **Heterogeneous Systems**



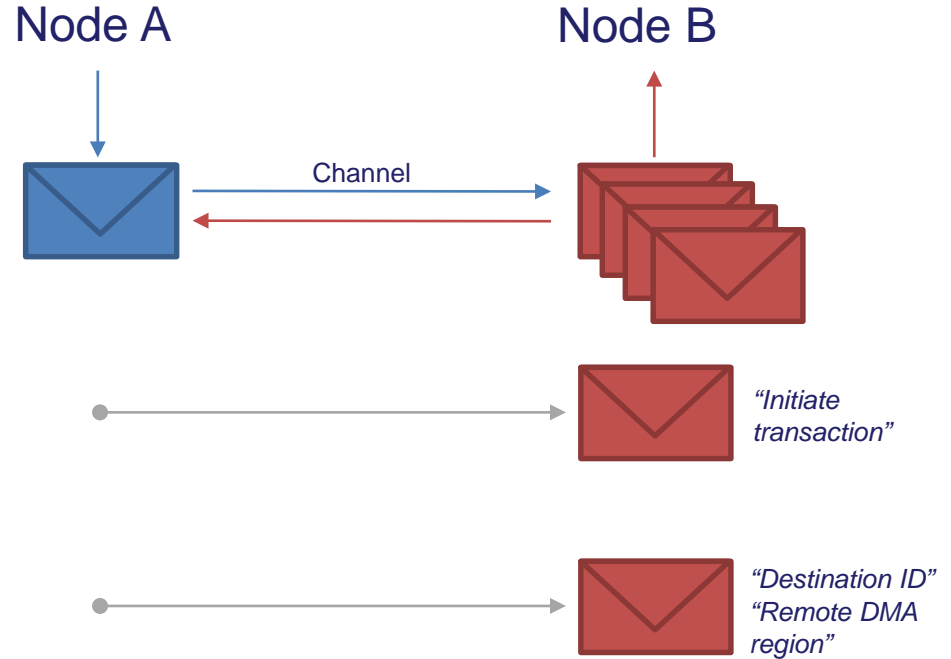
- › **4 server nodes equipped with RapidIO-PCIe bridge cards**
- › **38-port Top of Rack RapidIO switch**
- › **RapidIO drivers for the Linux kernel**
- › **User space libraries for Linux**



Image source: CERN, http://cds.cern.ch/record/2136852/files/DSC_2204.JPG?version=1

- › **Up to 4 KB/message**
- › **Socket-like interface**
- › **Good for orchestration**
 - **Initiate transactions**
 - **Exchange remote DMA information**

Channelized Messages

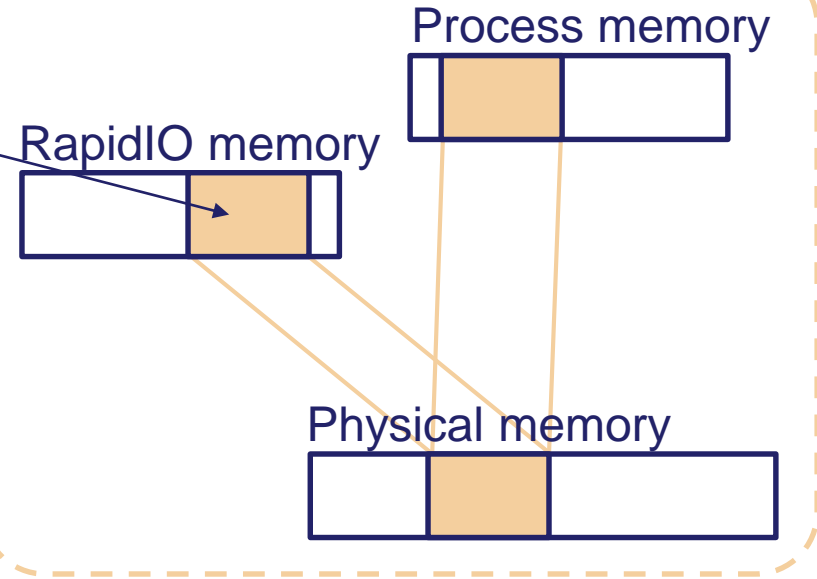


Remote DMA Write

Node A

```
write(0xbadf00d, B)
```

Node B



> **Multicast**

- List of remote endpoints

> **Zero copy**

- Sender: user buffer mapped to kernel space
- Receiver: shared memory

Project Use Cases

**Data
Analytics**

(ROOT, Hadoop)

**Data
Acquisition**

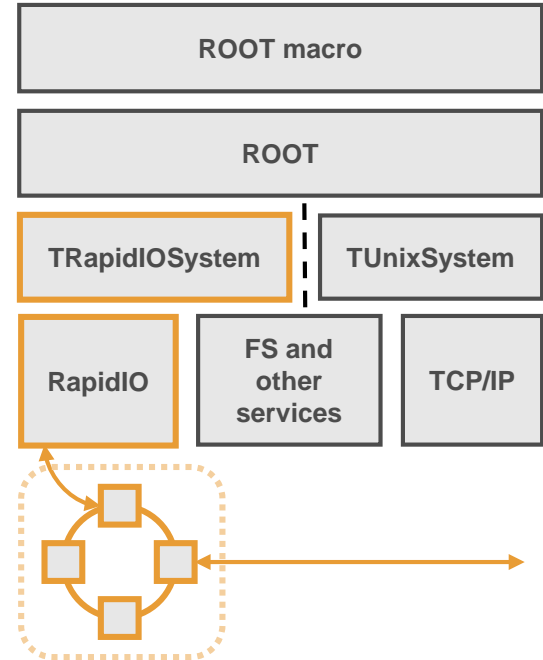
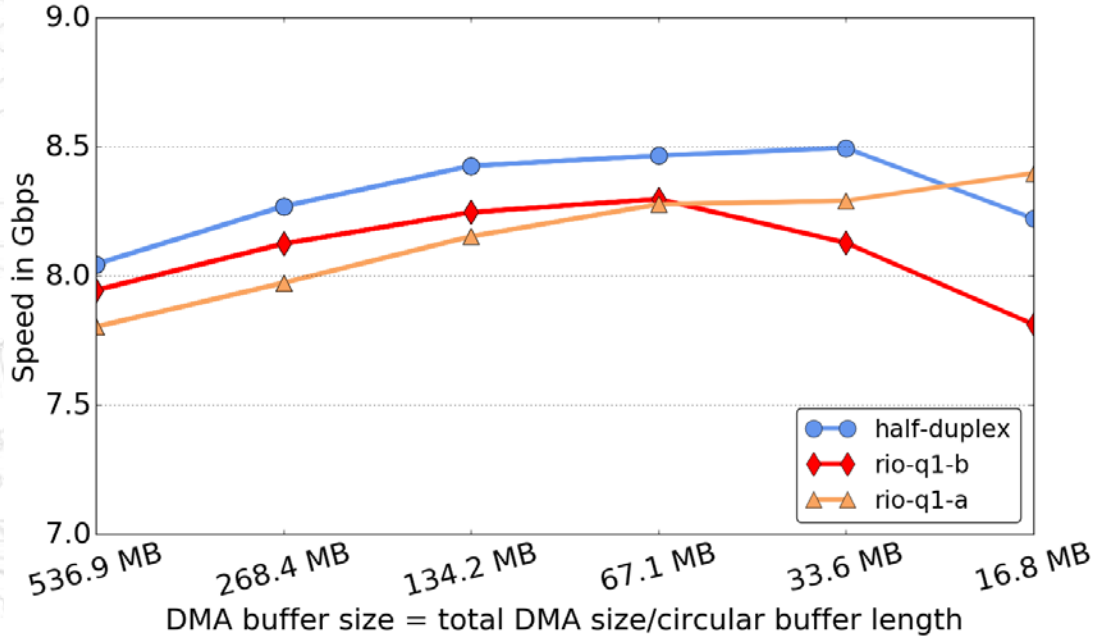
(DAQPIPE)

**Real Time
Trigger
Systems**

- › **Explore RapidIO**
- › **Focus on different areas**
- › **Evaluate suitability**

Use Case 1: ROOT

Half-duplex vs Duplex



Use Case 1: Hadoop

- › **riosocket (drivers)**
 - **Emulate TCP/IP over RapidIO**
 - **Create an Ethernet interface**
- › **No porting work**

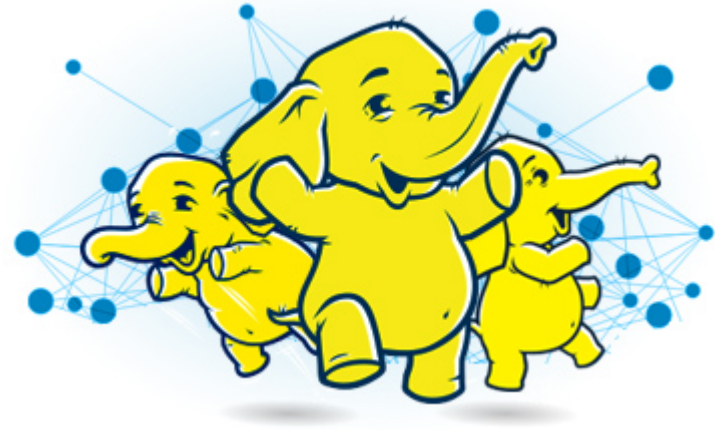
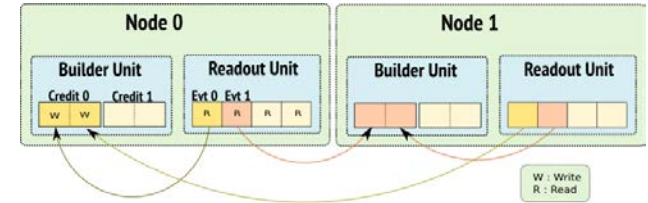
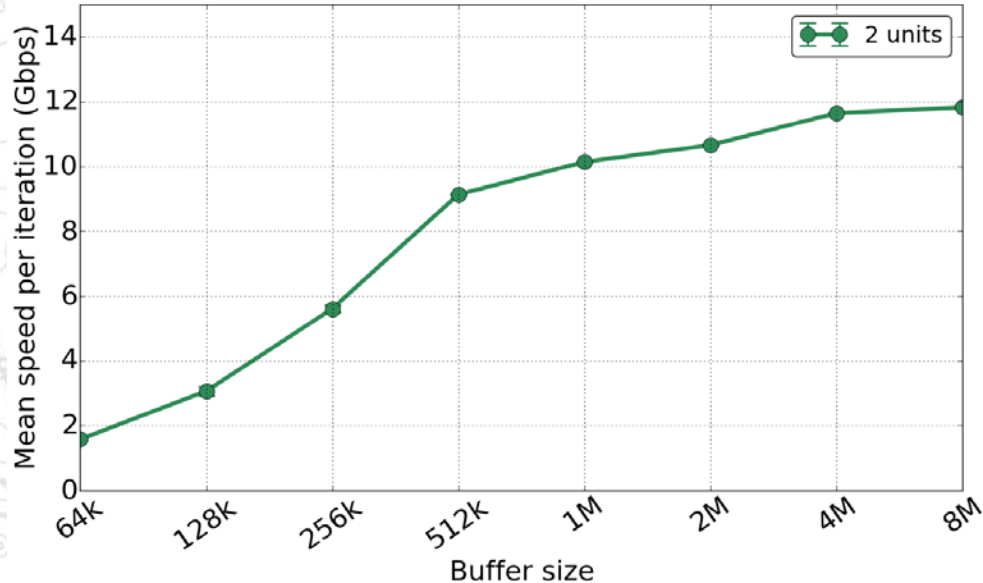


Image source: <http://insidebigdata.com/2015/07/21/hadoop-for-hpc-it-just-makes-sense/>

Use Case 2: LHCb DAQPIPE

Varying buffer sizes (work in progress)



- › LHCb benchmark for DAQ network
- › All-to-all communication
- › Event data spread across network
- › Nodes collect data for one event

Next Steps

- › **Scale up to 16 (32) nodes with current applications**
- › **Use case 3: Triggering systems**
- › **Explore multicast**

Thank you!