

The CBM First-level Event Selector Input Interface

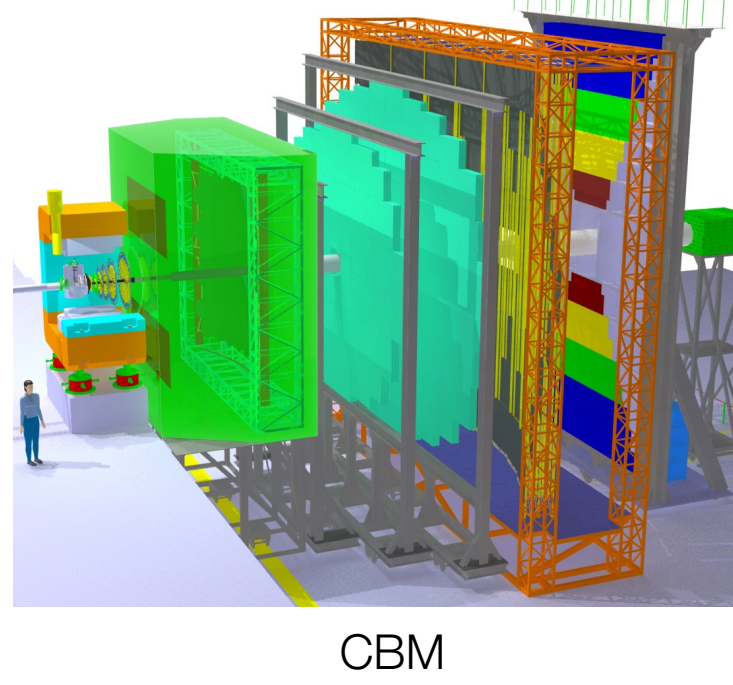
D. Hutter*, J. de Cuveland*, V. Lindenstruth*

* Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe-Universität Frankfurt am Main, Germany

CBM First-level Event Selector: FLES

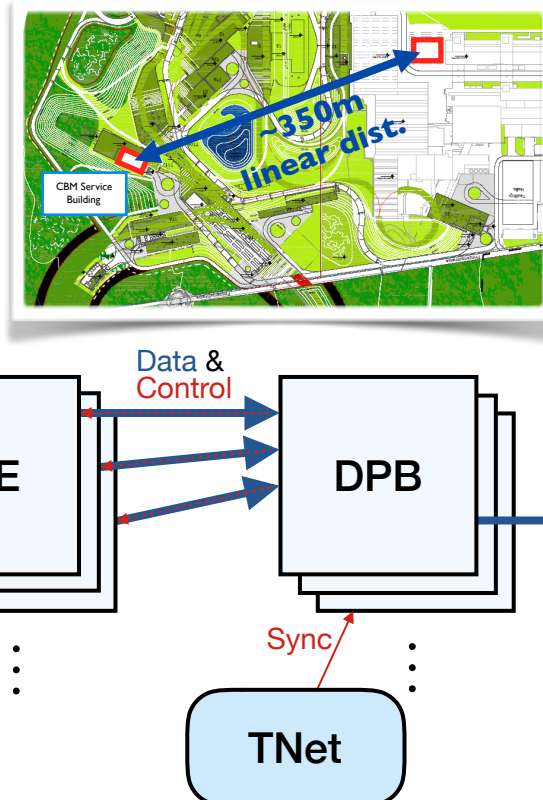
- Central physics selection system for CBM
- HPC processor farm with FPGAs, GPUs and fast interconnect
- Online analysis using fast, vectorized many-core track reconstruction algorithms

- 10⁷ events/s**
- Self-triggering front-end
 - Data push readout
 - Complex global triggers



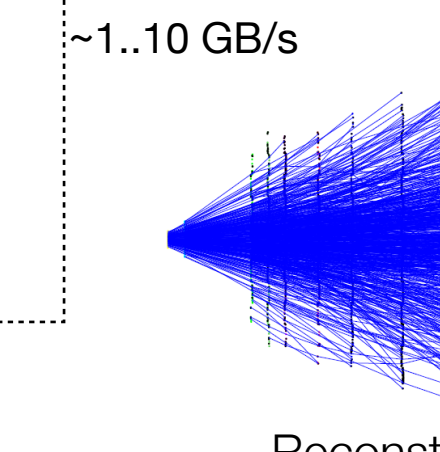
CBM

- Input Interface**
- >1 TByte/s input data rate
 - FPGA-based interface boards
 - Long-distance links to front-end



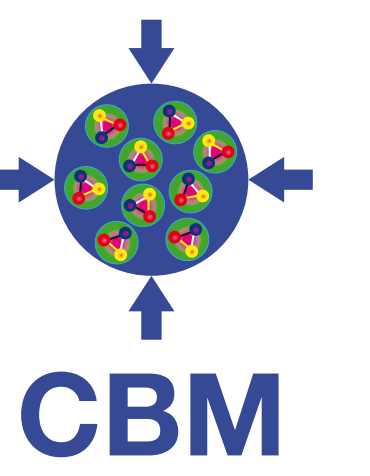
- High-throughput Event Building**
- Event building at full input data rate
 - ~1000 input nodes
 - Delivers fully build timeslices to reco
 - RDMA-based framework, zero-copy

Storage

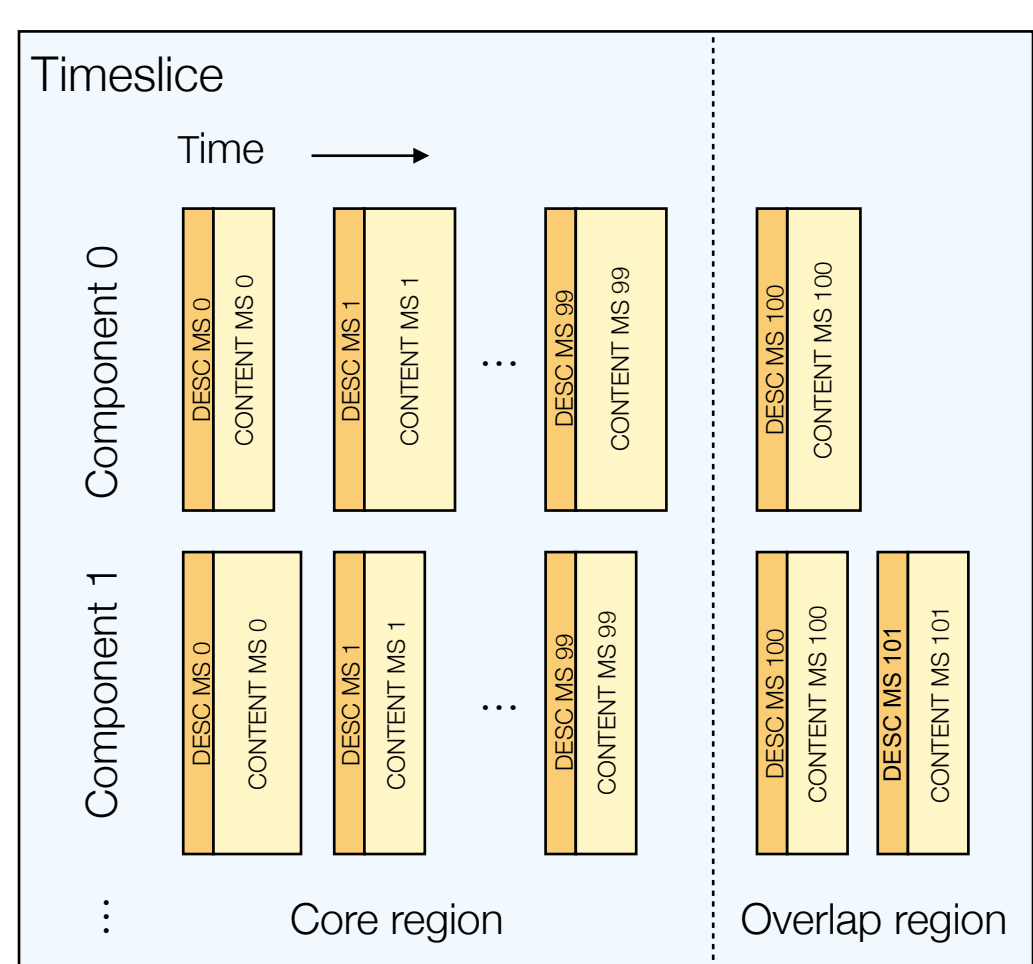


Reconstructed event in STS

- Online Event Selection**
- ~60.000 cores
 - 4-D tracking
 - Identification of leptons and hadrons
 - High-precision vertex reconstruction



Timeslice & Microslice Data Structures



Logical view of timeslice data structure

Microslice

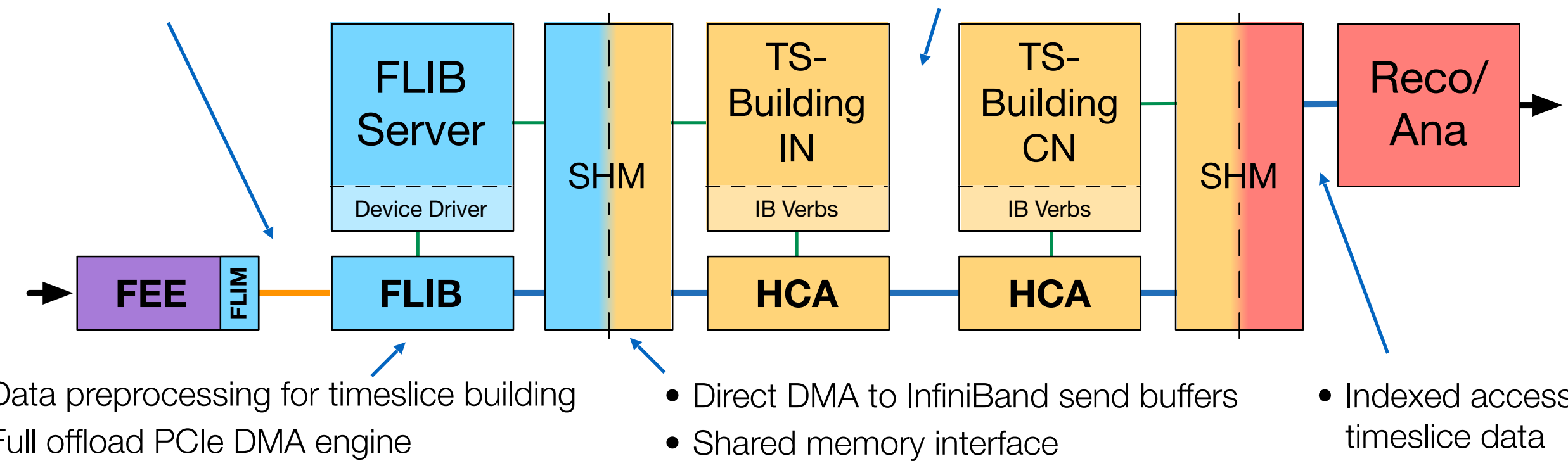
- FLES input data container created in front-end
- Constant in experiment time, variable in data size
- Timeframe covered by one MS is identical for all subsystems
- Descriptor: start time and metadata for timeslice building
- Data content: self-contained subsystem data, detector-specific, unknown to FLES data management

Timeslice

- Two-dimensional collections of microslices
- Processing entity consumed by analysis
- Overlap region allows independent processing

Data Flow

- 10 Gbit/s custom optical link
- Common front-end interface module
- Timeslice building
- InfiniBand RDMA, true zero-copy

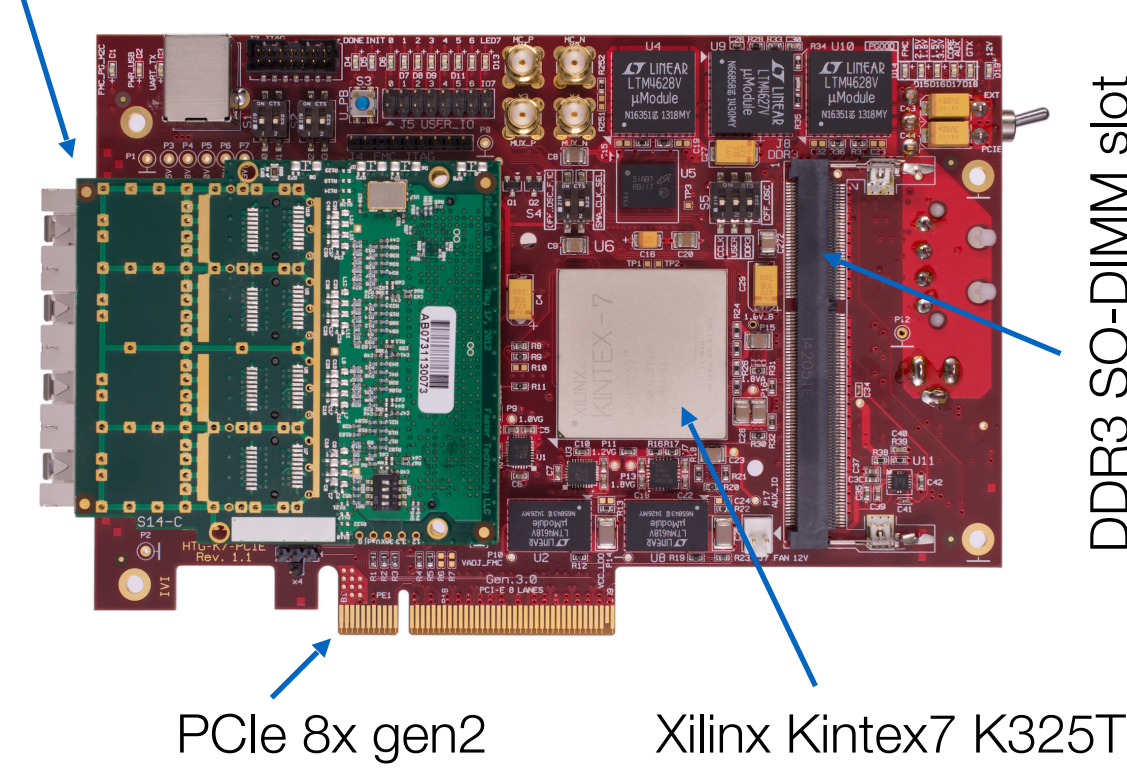


- Data preprocessing for timeslice building
- Full offload PCIe DMA engine
- Direct DMA to InfiniBand send buffers
- Shared memory interface
- Indexed access to timeslice data

FLES Interface Board: FLIB

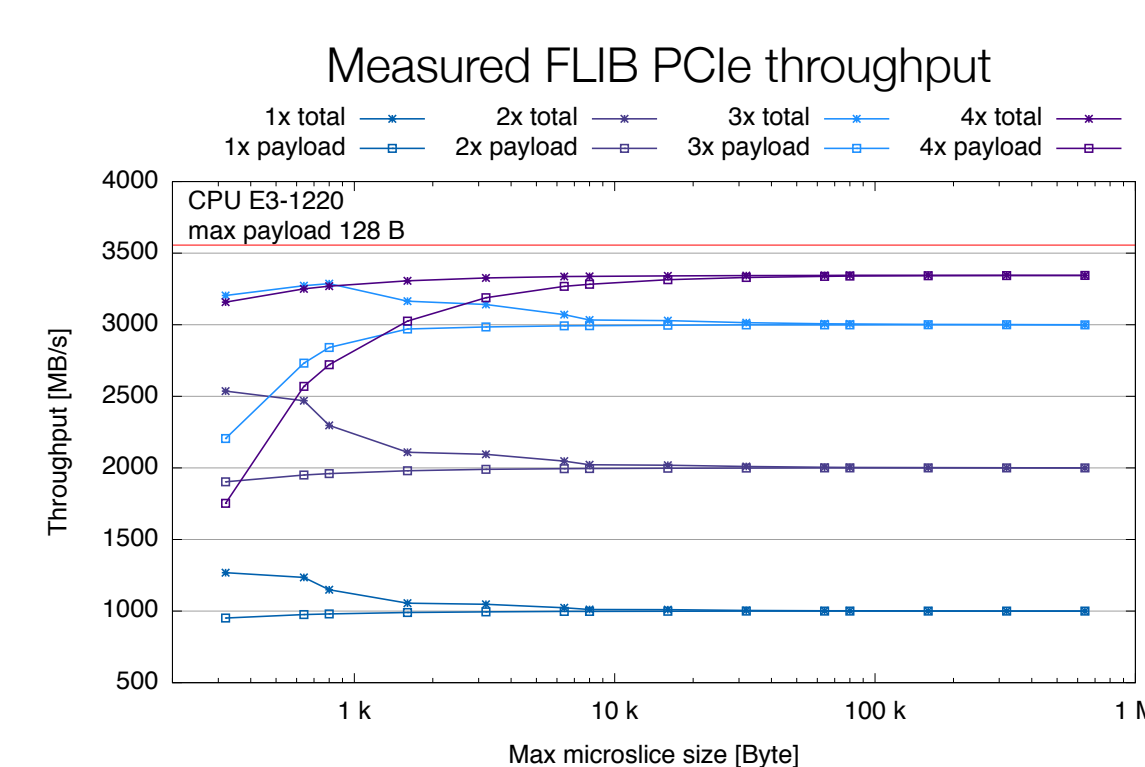
- FPGA-based PCIe card
- Provides interface between FLES and readout tree
- Prepares data for efficient timeslice building
- Current platform: HTG-K7-PCIE board
- Additional purposes for current version:
 - Test platform for FLES hardware and software development
 - Readout device for test beams and lab setups without full infrastructure

Up to eight 10 Gbit/s links via FMC

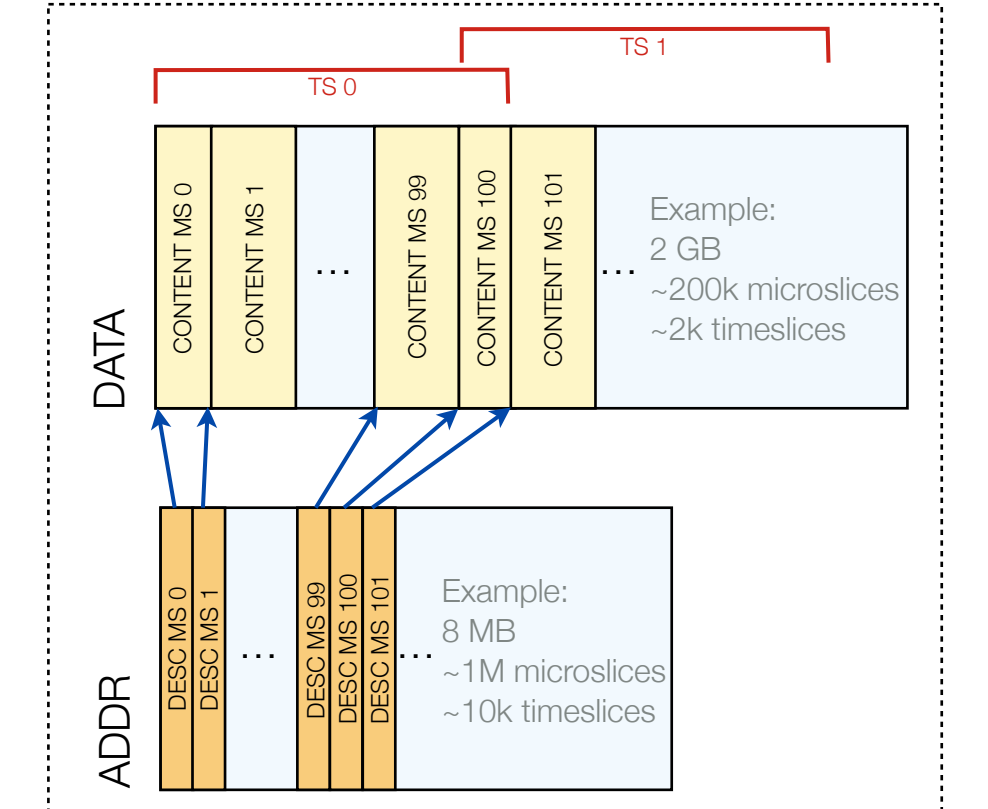


PCIe Host Interface

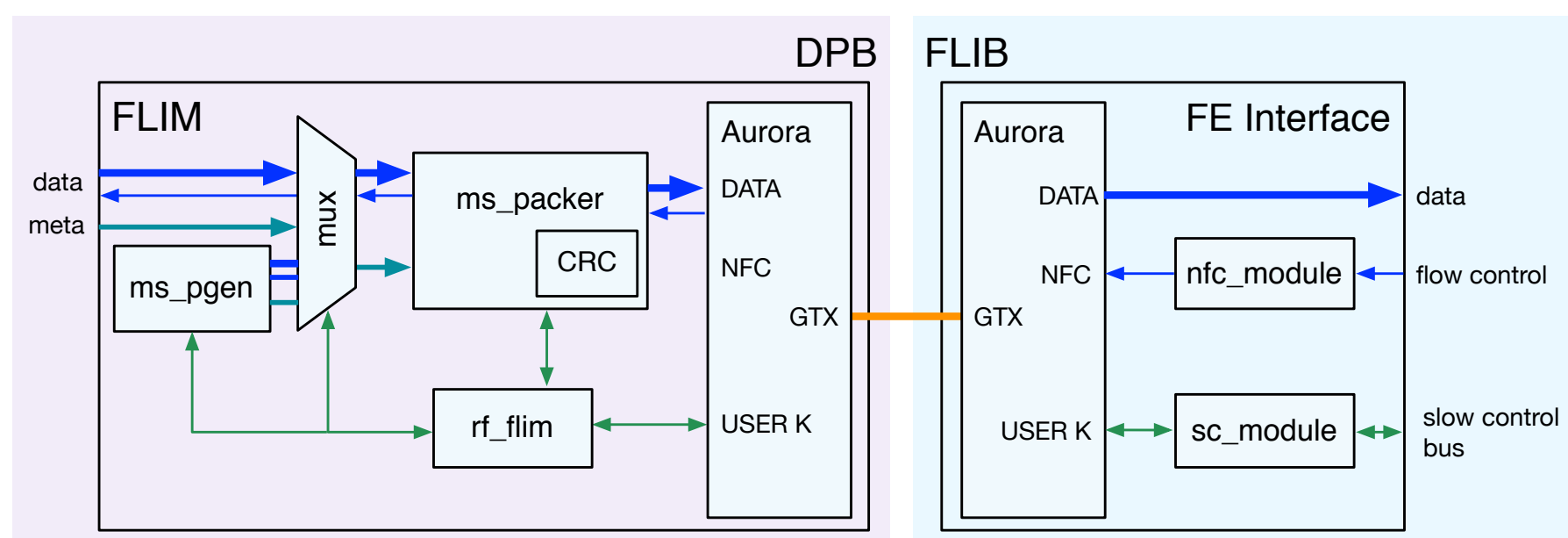
- Full offload DMA engine manages dual ring buffer structure
 - Data buffer for microslice data content
 - Descriptor buffer for index table and microslice meta data
- Random access to data via index tables
- DMA directly to posix shared memory buffer
 - Buffer can be registered in parallel for InfiniBand



Dual Ring Buffer in Shared Memory



Front-End Interface



FLES Interface Module: FLIM

- Common HDL module implementing front-end logic interface and link protocol
- Runtime configuration via FLES link
- Internal pattern generator for test and validation

Optical long-distance links

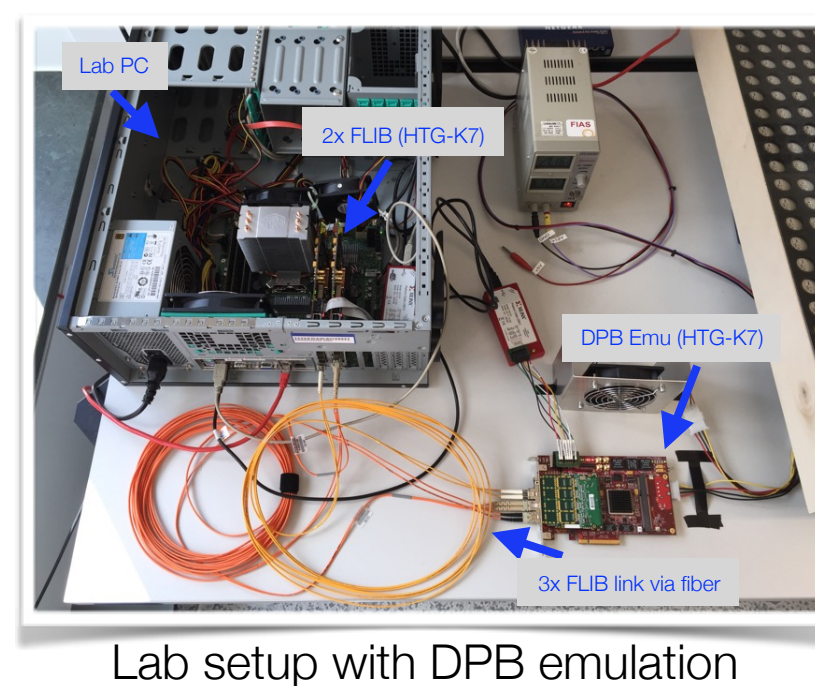
- 10.3125 Gbit/s line rate
- 64B/66B Aurora transport layer
- CRC-32C (Castagnoli) end-to-end checksum on data content

Status

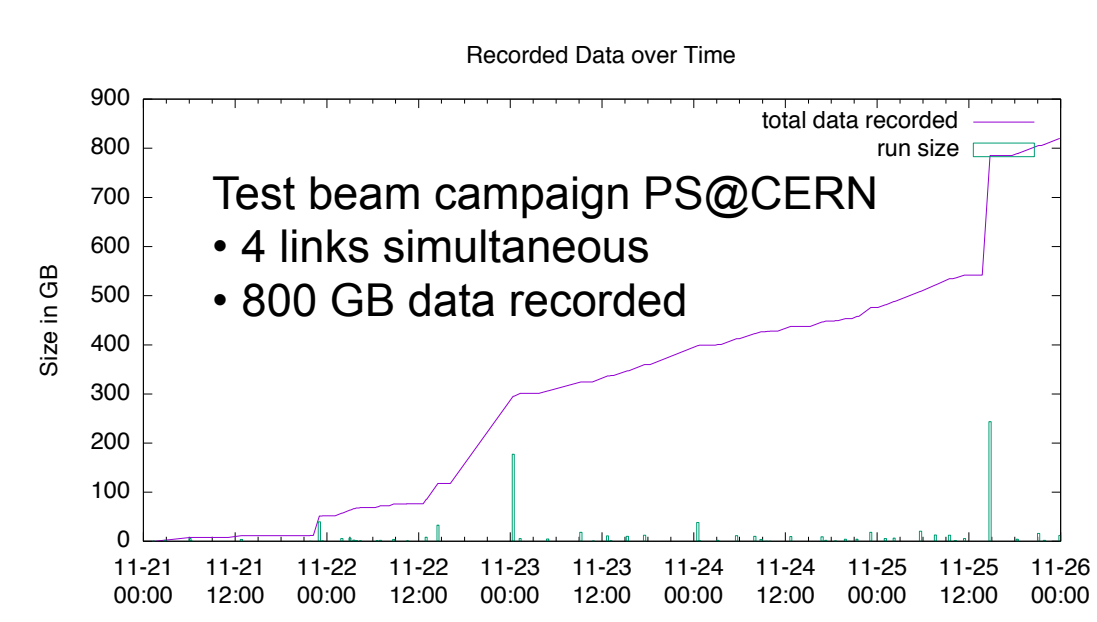
- FLES input interface is successfully used in numerous setups
- System bit error rates <10⁻¹⁶ per link measured in stability tests
- FLIB and FLES framework used to read out test beam campaigns in 2014 and 2015 without DPB
- 16 link setup with DPB layer in preparation for 2016 test beam



Setup with DPB layer in MTCA crate



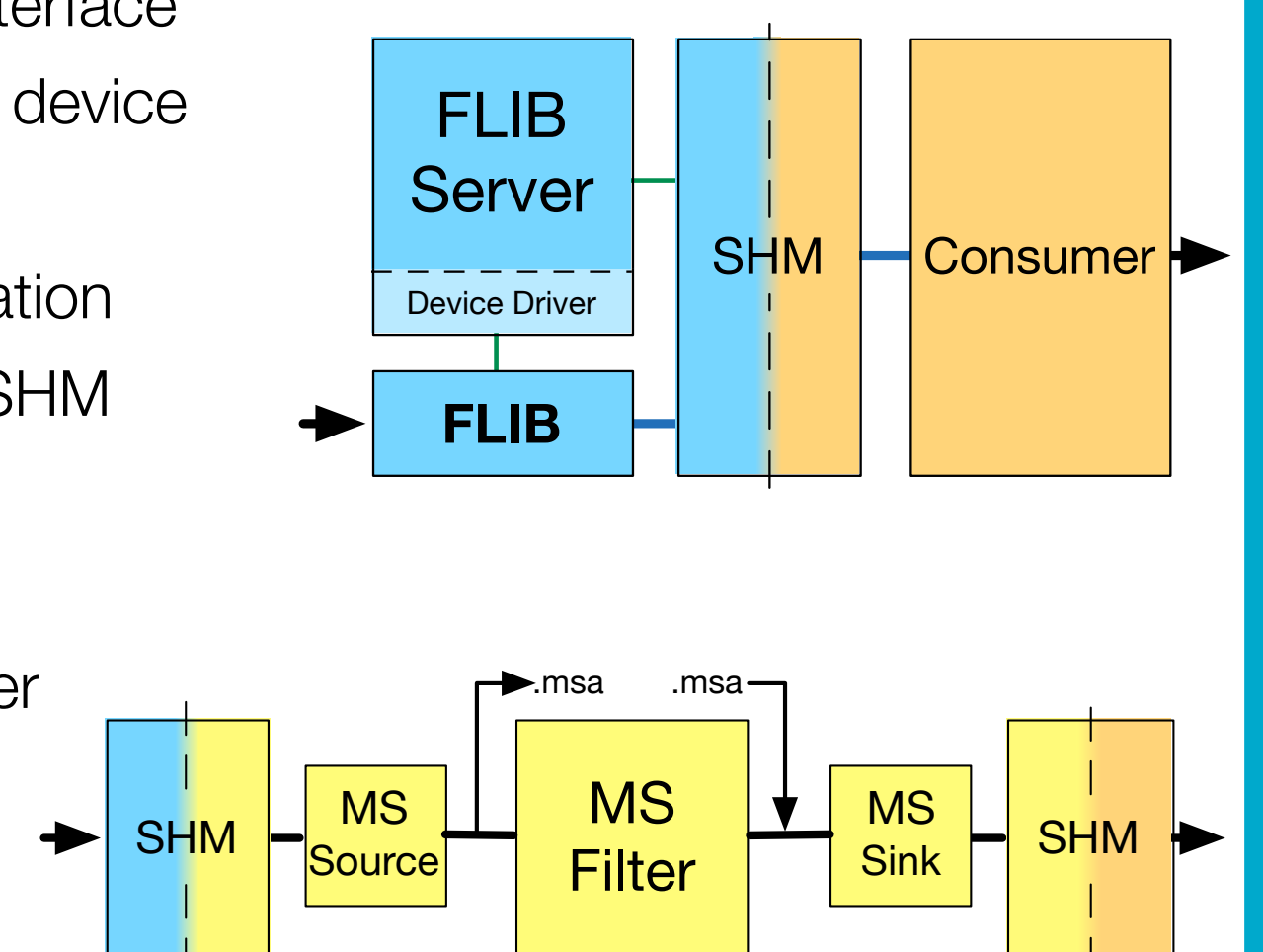
Lab setup with DPB emulation



Test beam campaign PS@CERN
• 4 links simultaneous
• 800 GB data recorded

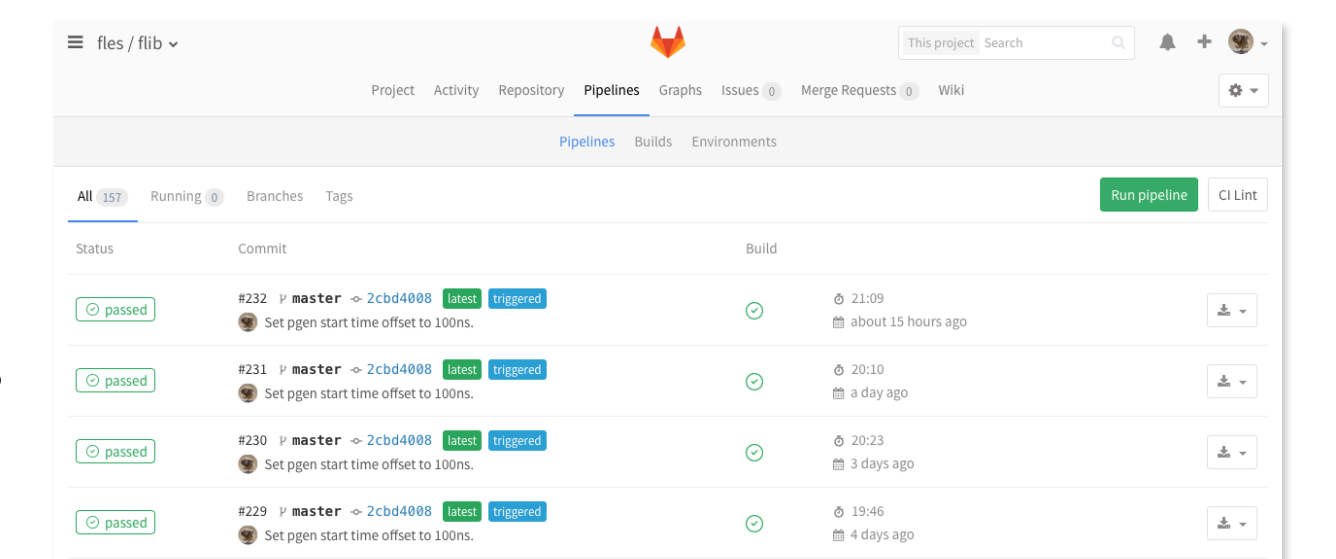
Application Interface

- Consumers access data via generic ring buffer read interface
 - Memory can be filled from any source, not only PCI device
 - Multiple consumers possible
- All device specific actions handled by server application
- Data and needed synchronization structures in same SHM
 - Low-level indexed ring buffer interface
 - Raw access to ring buffers in shared memory
 - Read/writer pointer updates towards FLIB/consumer
- High-level microslice stream interface
 - Allows to implement filters before timeslice building



Firmware Management

- All sources are version controlled
- Fully automated build flow
- Everything is build from source
 - Sophisticated build information included into bitfiles
 - E.g., build date, repository revision, ...
 - Accessible in online system via slow control
- Continuous integration via Gitlab CI



Continuous integration of the FLIB hardware project

References

[1] J de Cuveland et al, 2011 J. Phys.: Conf. Ser. 331 022006