

# Data Acquisition Backbone Core

J.Adamczewski, H.G.Essel, N.Kurz, S.Linev

GSI, Experiment Electronics: Data Processing group

Recall of the CBM FutureDAQ

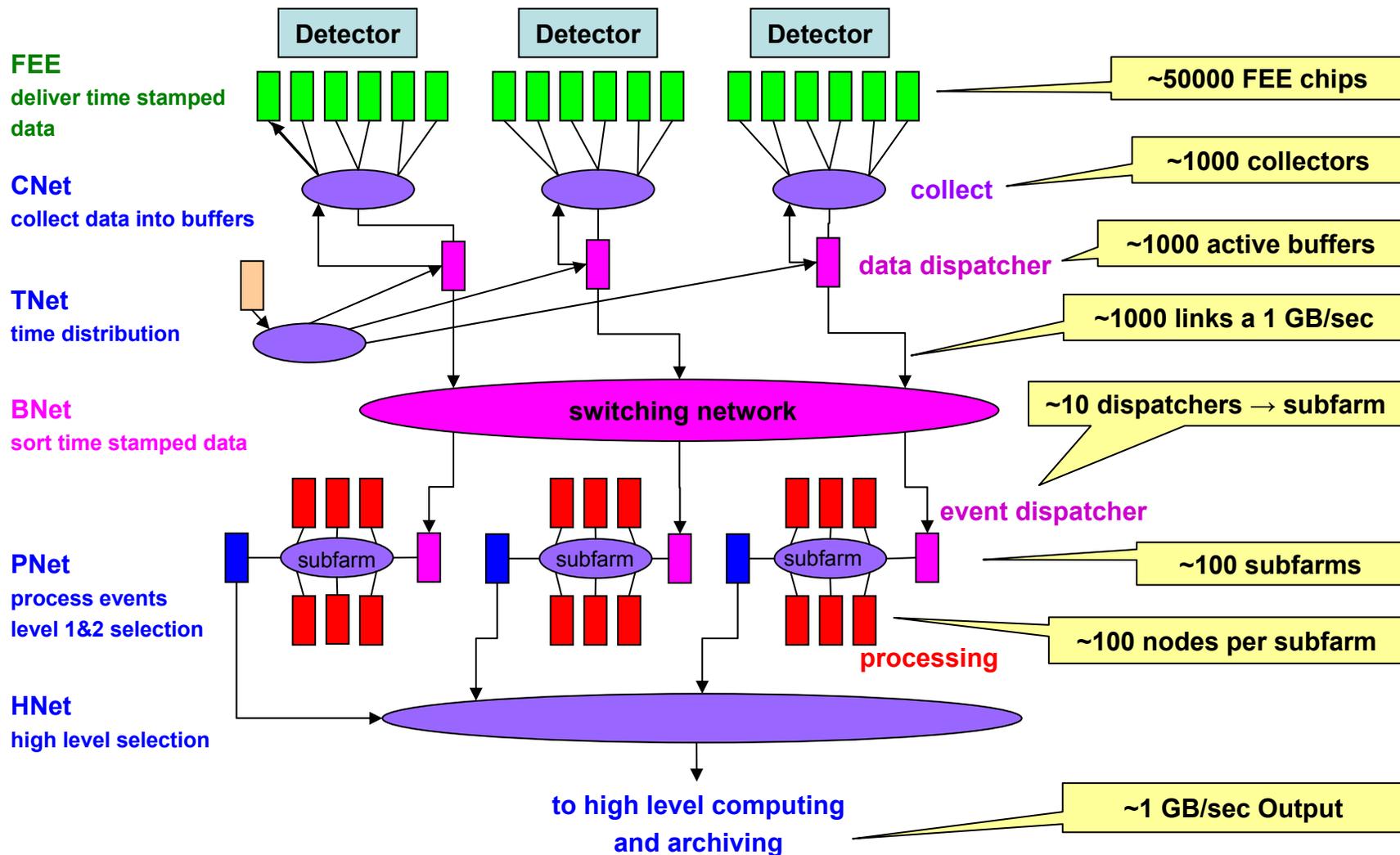
What we have done up to now

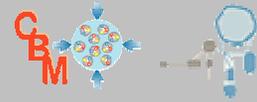
What comes next: **DABC**

Work supported by [EU RP6 project JRA1 FutureDAQ RII3-CT-2004-506078](#)

CBM GSI DABC - J.Adamczewski, H.G.Essel, N.Kurz, S.Linev

W.F.J.Müller, 2004





## New paradigm: switch full data stream into event selector farms

1. A conventional LVL1 trigger would imply full displaced vertex reconstruction within fixed latency.
2. Strongly varying complex event filter decisions needed on almost full event data

➔ **No common trigger! Self triggered channels with time stamps! Event filters**

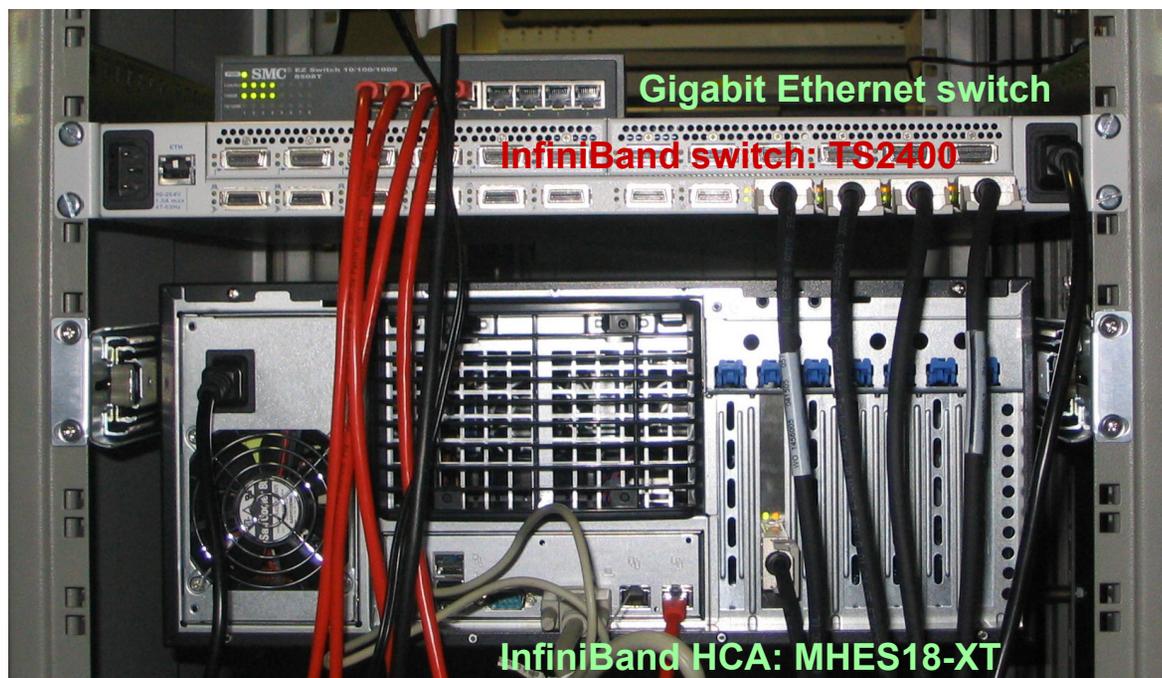
- **10 MHz** interaction rate **expected**
- Time stamps in all data channels, typical resolution **~ns**) **required**
- **1 TByte/s** primary data rate (Panda < 100 GByte/s) **expected**
- **1 GByte/s** maximum archive rate (Panda < 100 MByte/s) **required**
- Event definition (time correlation: multiplicity over time histograms) **required**
- Event filter to **20 KHz** (1 GByte/s archive with compression) **required**
- On-line track & (displaced) vertex reconstruction **required**
- Data flow driven, no problem with latency **expected**
- **Less complex** communication, but **high data rate** to sort



- **Triggerless** data acquisition and transport until filter farm
- FPGA controlled data flows
- Event building on **full data rate ~1TB/s**
- BNet: **~1000 nodes, high-speed interconnections**
- **Linux** may run on most DAQ nodes (even FPGAs)

What has been done?

- 4 nodes:
  - Double Opteron 2.2 GHz, 2GB RAM
  - Mellanox MHES18-XT host adapter (PCIe)
  - 2x Gigabit Ethernet host adapters
  - SuSE Linux 9.3, x64bit version
- Mellanox MTS2400 24X InfiniBand switch



See Talk by S.Linev tomorrow

## Extensive simulations of event building Bnet with SystemC

- Mellanox IB Gold 1.8.0
  - IPoIB: IP over InfiniBand driver
  - uDAPL: User Direct Access Programming Layer
  - MPI: Message Passing Interface
- OpenFabric Enterprise Distribution (OFED) 1.1
  - IB Verbs (similar to uDAPL, plus Multicast)
- MVAPICH2: MPI 2 implementation for IB
- IBAdmin
- OpenSM (subnet manager) (for Multicast)

Very extensive testing of all packages.

Scheduled all to all data transfer (basic pattern for switched event building):  
**500-800 MByte/s (2K-64K buffer) per node.**

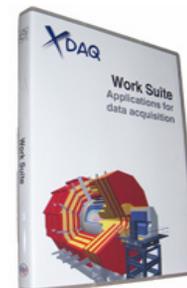
Cooperation with Forschungszentrum Karlsruhe established to scale the tests to 32-64 InfiniBand nodes. **Next week first test with 24 nodes.**

See Talk by J.Adamczewski tomorrow



## Standard DAQ framework for LHC CMS experiment \*

- C++ libraries on Linux, modular packages (SourceForge)
- Distributed xDAQ applications
- Configuration: XML
- Data transport: I<sub>2</sub>O protocol (Intelligent IO)
- Communication: http, cgi; SOAP messages
- InfoSpace: Global parameter access (subscription)
- State machines (sync/async FSM)
- Message logger, error handler
- Monitoring tool
- Hardware access library (HAL)
- Job Control (task handler for node control)
- others: exceptions, threads, data (de)serializers...



\* Orsini, Gutleber <http://xdaqwiki.cern.ch>



## Developments at GSI for DABC (FDAQ)

- DABC application base classes prototypes
- xDAQ peer transport for InfiniBand (uDAPL)
  - uDAPL buffers managed within XDAQ memory pool
  - avoids memcopy and new buffer allocation for each send package:
  - lookup if posted memory reference is known as send buffer
  - user code can write directly into uDAPL send buffer
- multiple threads for sending, releasing, and receiving buffer
- PCI/PCIe drivers\* integrated in Hardware Access Library HAL
- DIM server as interface to xDAQ InfoSpace to control applications
- Test GUI in Labview over DIM
- Test GUI in Java speaking SOAP or DIM\*\* with xDAQ executives

\* Collaboration with Guillermo Marcus, Mannheim

\*\* to be done

## Tools for controls

- State machines, Infospace, message/error loggers, monitoring
- Communication: Webserver, SOAP, DIM
- Connectivity through DIM to: LabView, EPICS, Java, any DIM client/server
- GUI not yet evaluated, Java worth a close look (Eclipse)
- Front-end controls?
- Mix of cooperating control systems

So, what are the next plans?



2004 → EU RP6 project JRA1 FutureDAQ\*

2004 → CBM FutureDAQ for FAIR

2005 → FOPI DAQ upgrade (skipped)

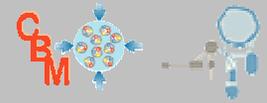
1996 → MBS future  
50 installations at GSI,  
50 external  
<http://daq.gsi.de>

Intermediate  
demonstrator

- Detector tests
- FE equipment tests
- Data transport
- Time distribution
- Switched event building
- Software evaluation
- MBS event builder
- General purpose DAQ

**D**ata  
**A**cquisition  
**B**ackbone  
**C**ore

\* RII3-CT-2004-506078

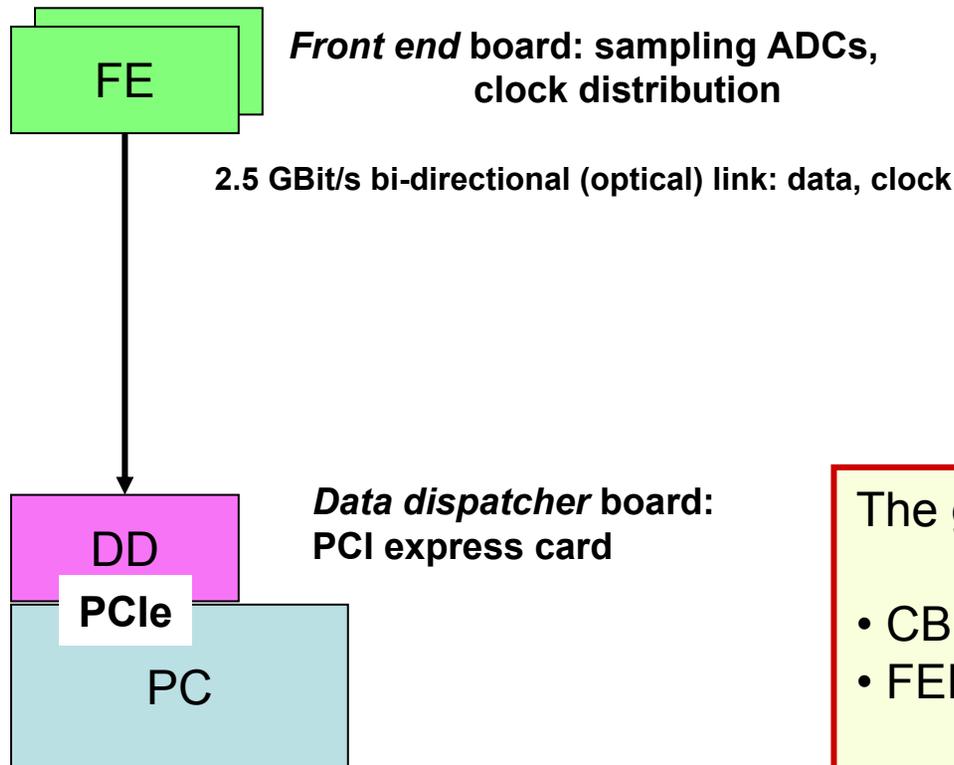
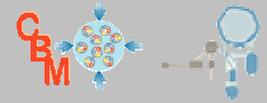


**D**ata  
**A**cquisition  
**B**ackbone  
**C**ore

## Requirements

- build events over fast networks
- handle triggered or triggerless front-ends
- process time stamped data streams
- provide data flow control (to front-ends)
- connect (nearly) any front-ends
- provide interfaces to plug in application codes
- connect MBS readout or collector nodes
- be controllable by several controls frameworks

## Use xDAQ infrastructure

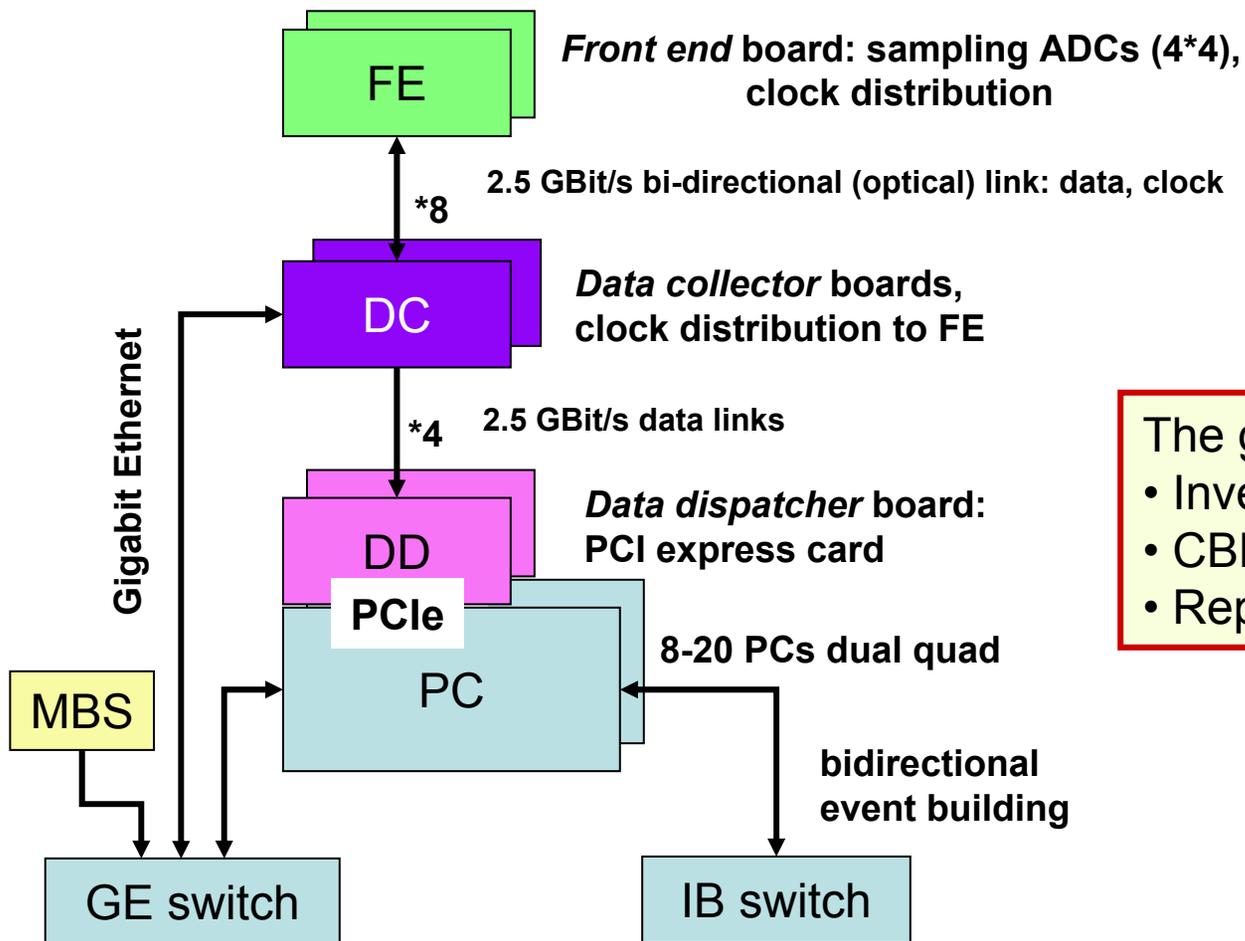


FE: Frontend board  
 DC: Data combiner board  
 DD: Data dispatcher board  
 GE: Gigabit Ethernet  
 IB: InfiniBand

The goal:

- CBM detector tests
- FEE tests

See W.F.J.Müller: N-XYTER  
 U.Brüning: ABB/DCB

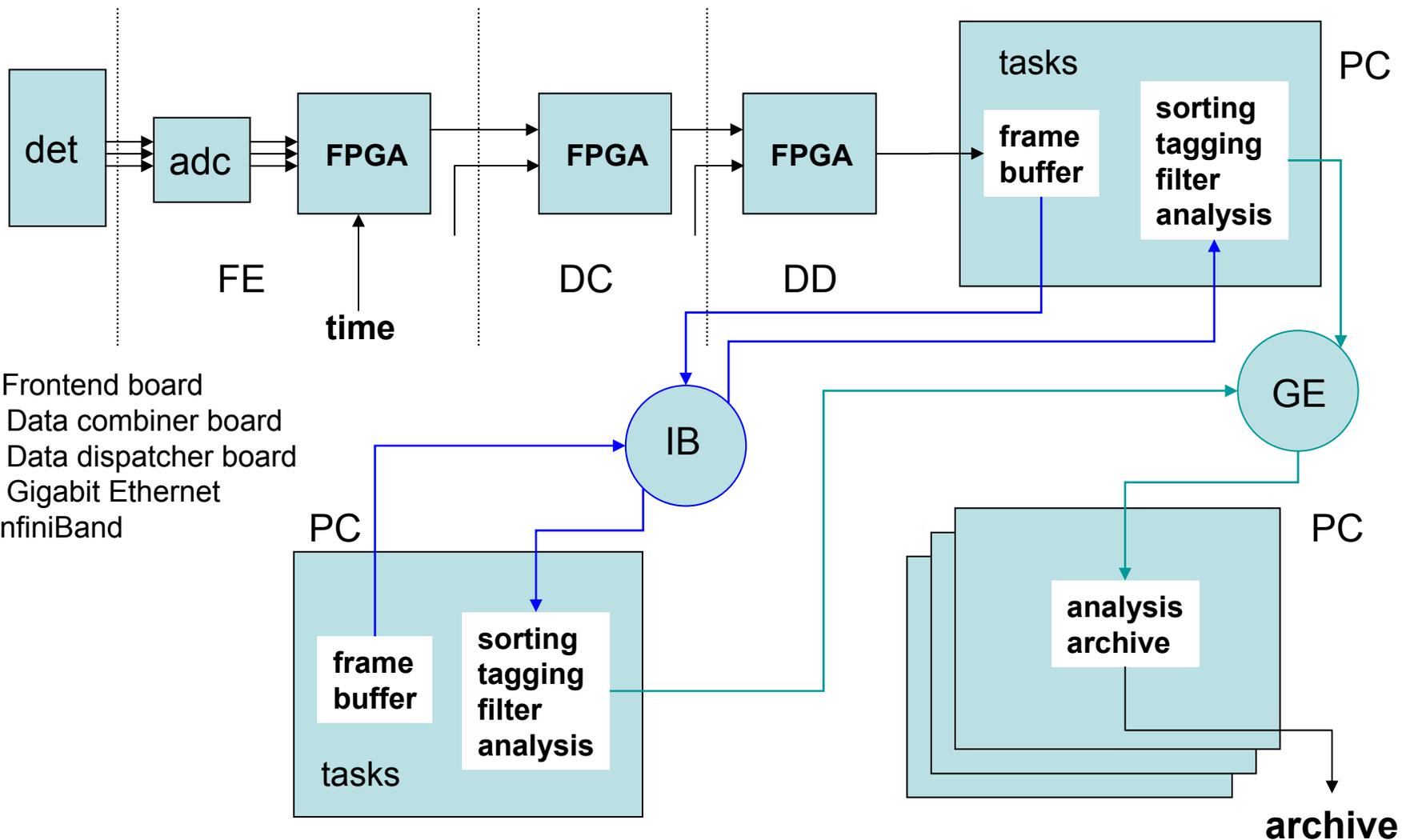


FE: Frontend board  
 DC: Data combiner board  
 DD: Data dispatcher board  
 GE: Gigabit Ethernet  
 IB: InfiniBand

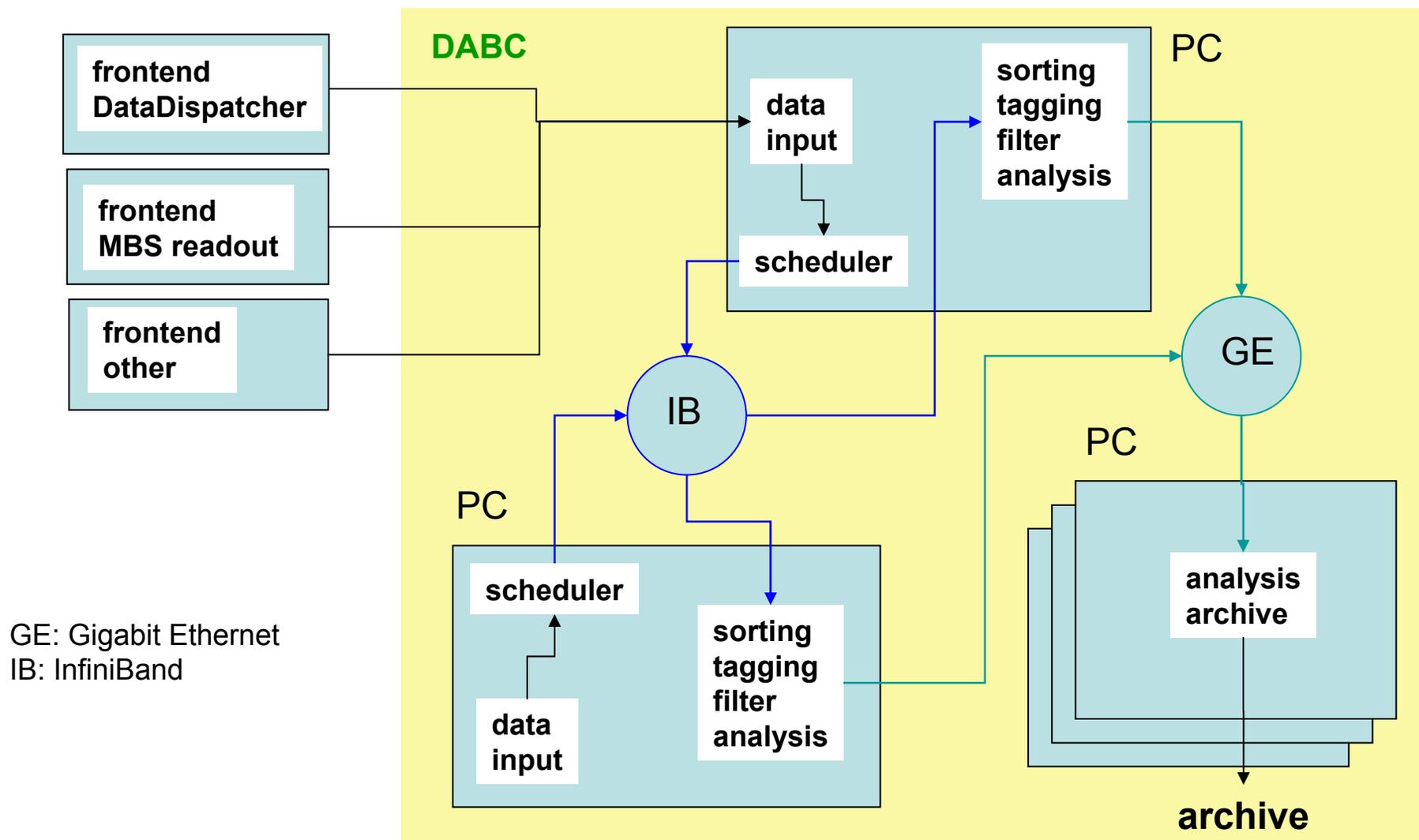
- The goal:
- Investigate critical technology
  - CBM detector tests
  - Replace existing DAQ

Scales up to 10k channels, 160 CPUs

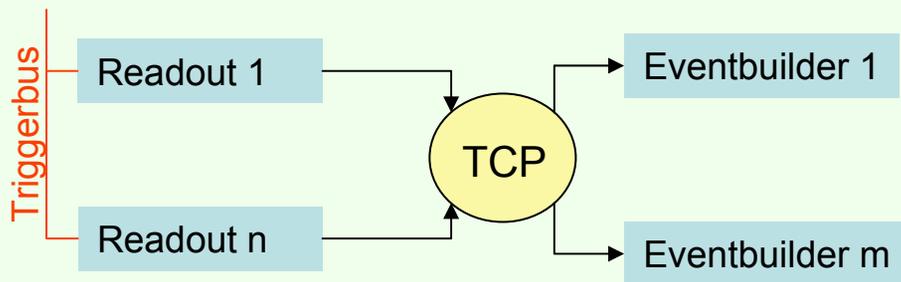
See W.F.J.Müller: N-XYTER  
 U.Brüning: ABB/DCB



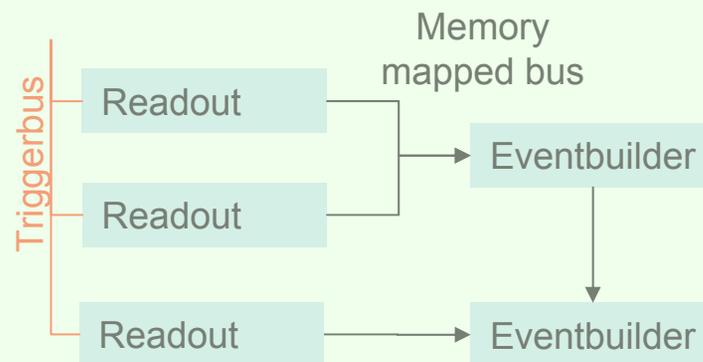
FE: Frontend board  
 DC: Data combiner board  
 DD: Data dispatcher board  
 GE: Gigabit Ethernet  
 IB: InfiniBand



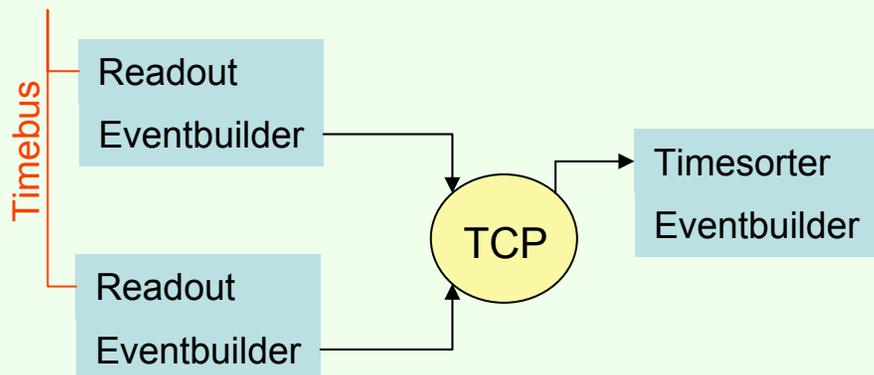
**n x m multiple EBs**



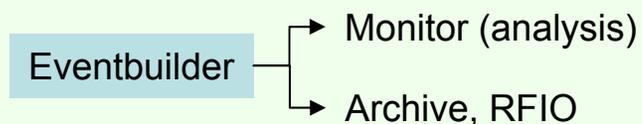
**EB hierarchy**

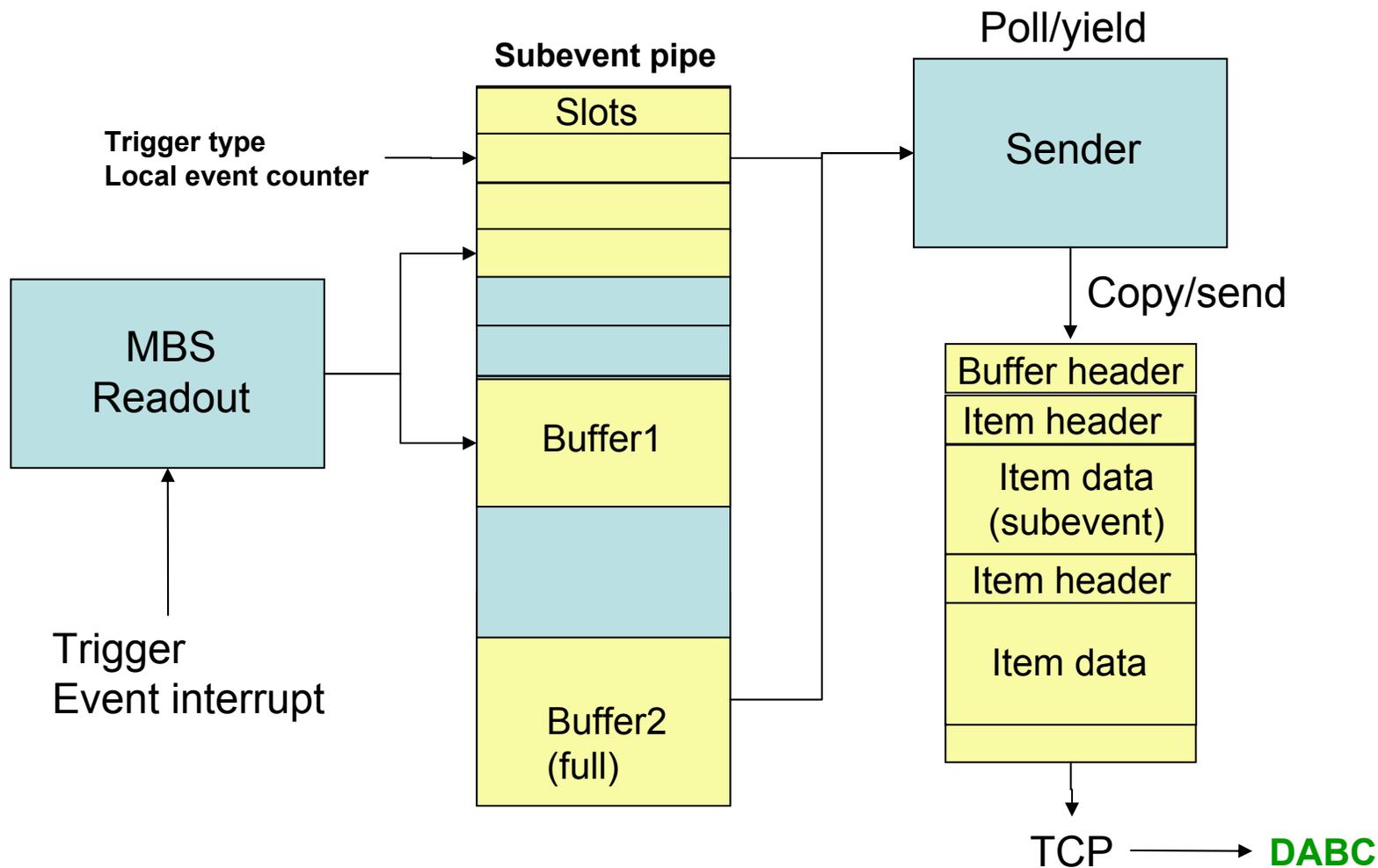


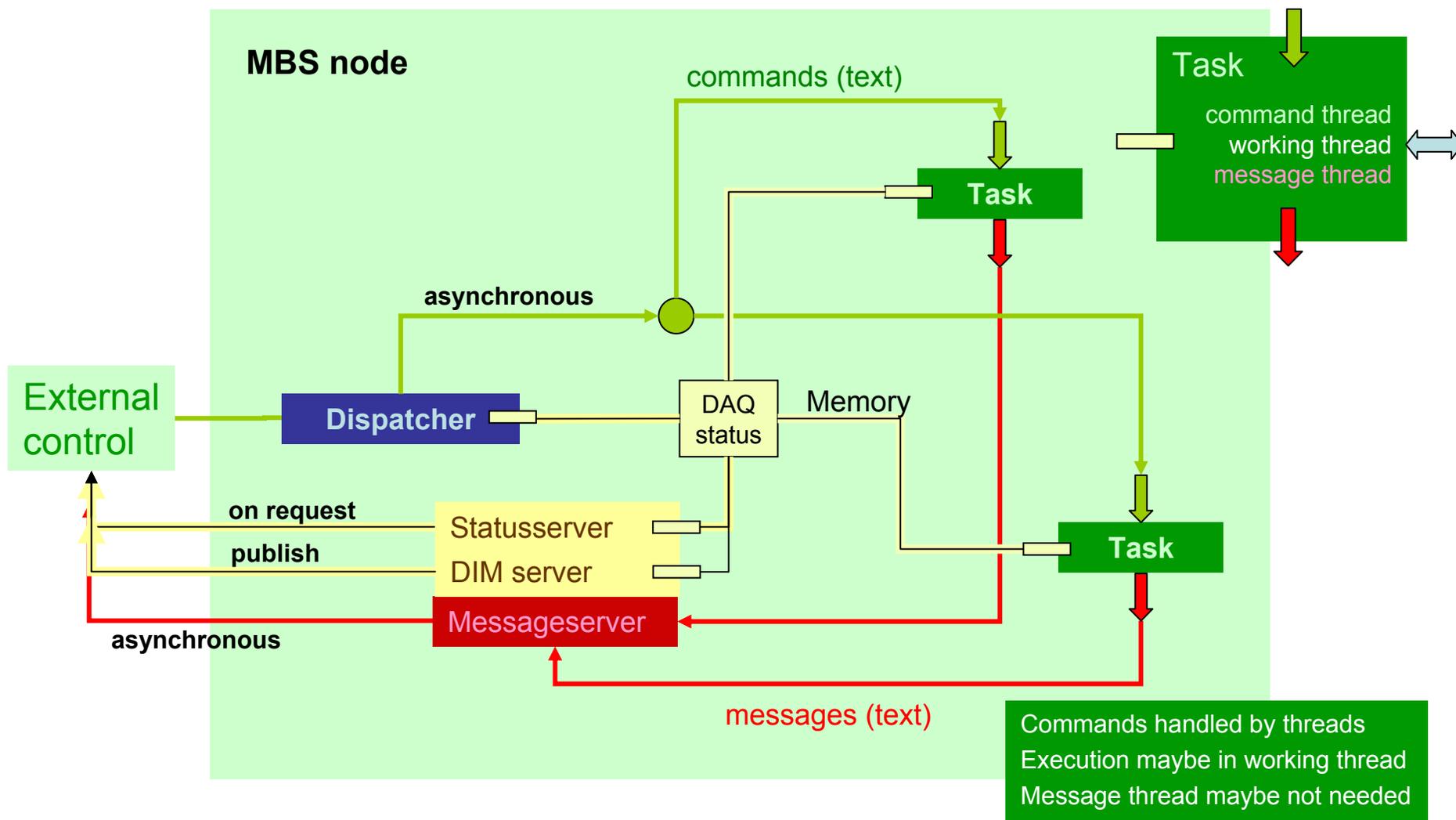
**Time stamped**

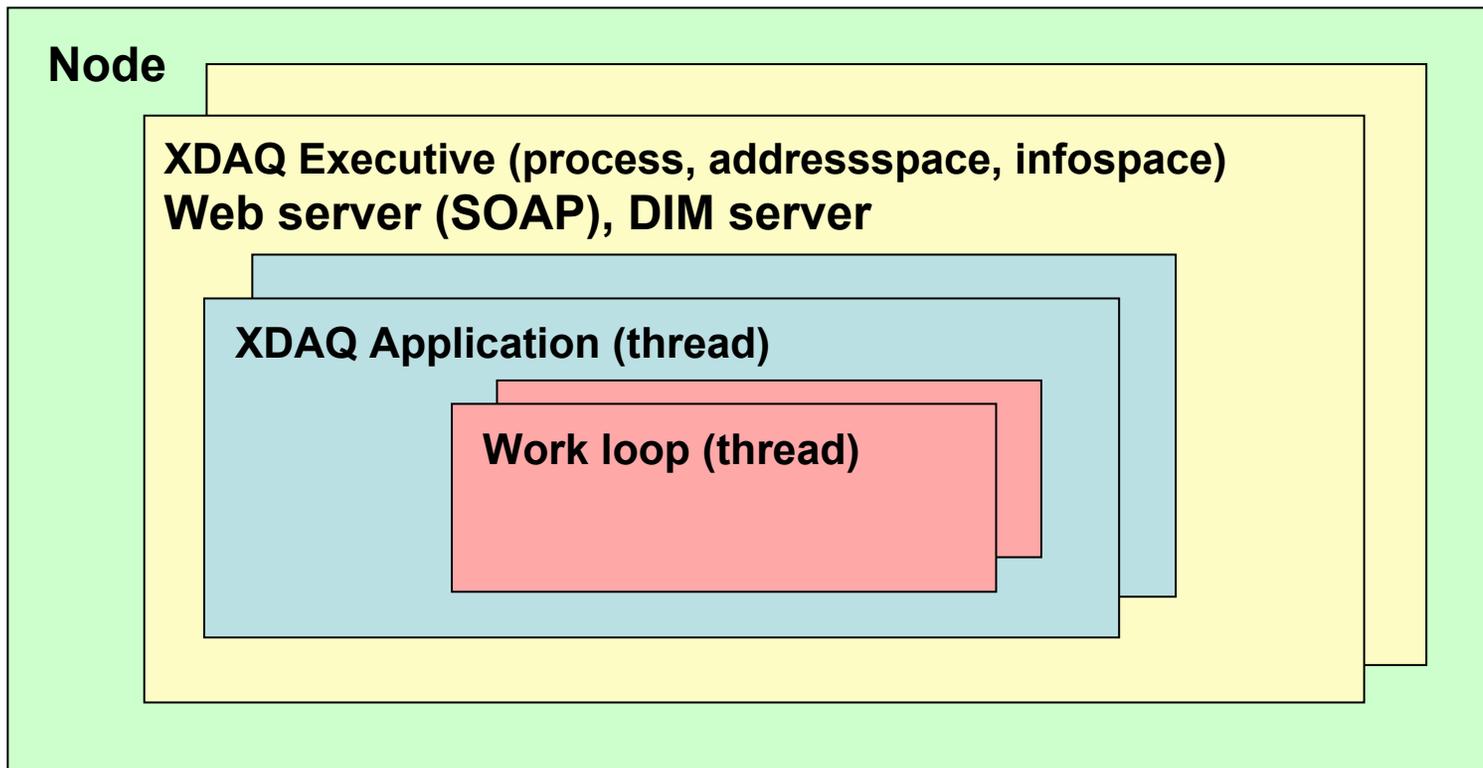


**EB data transport**

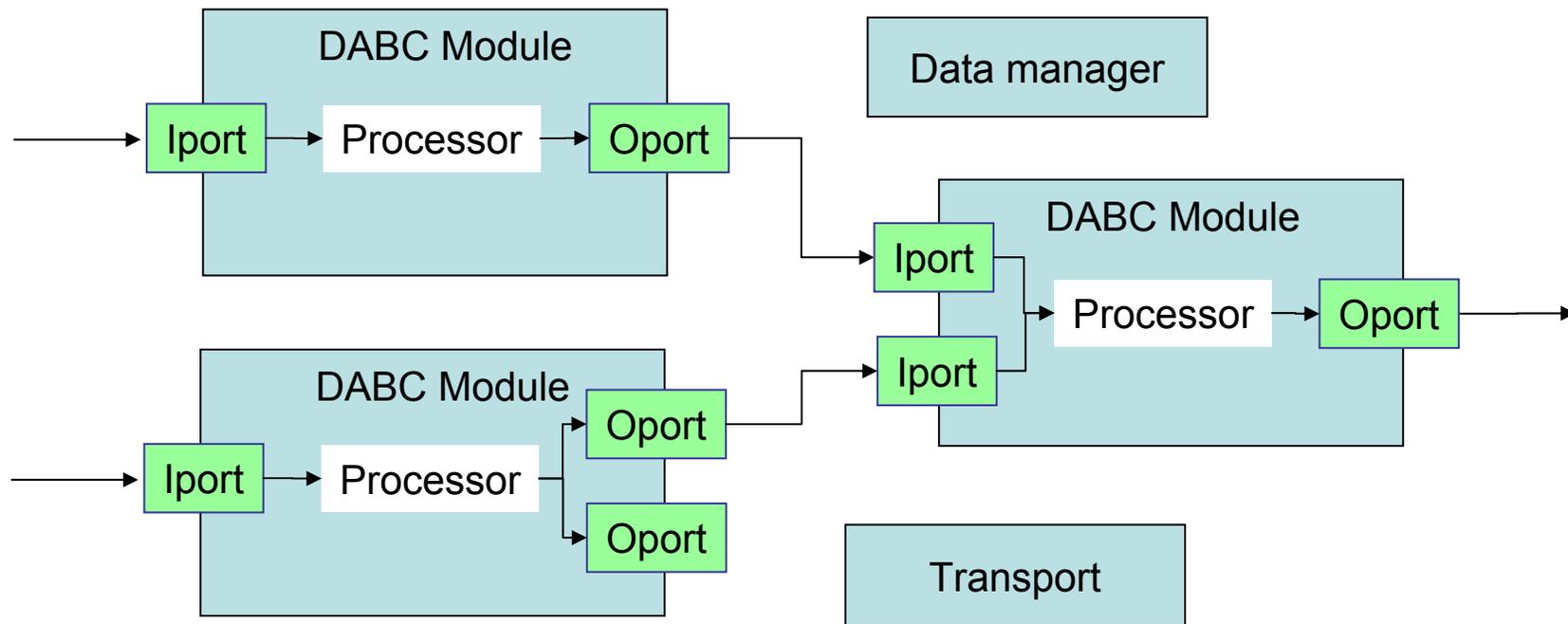








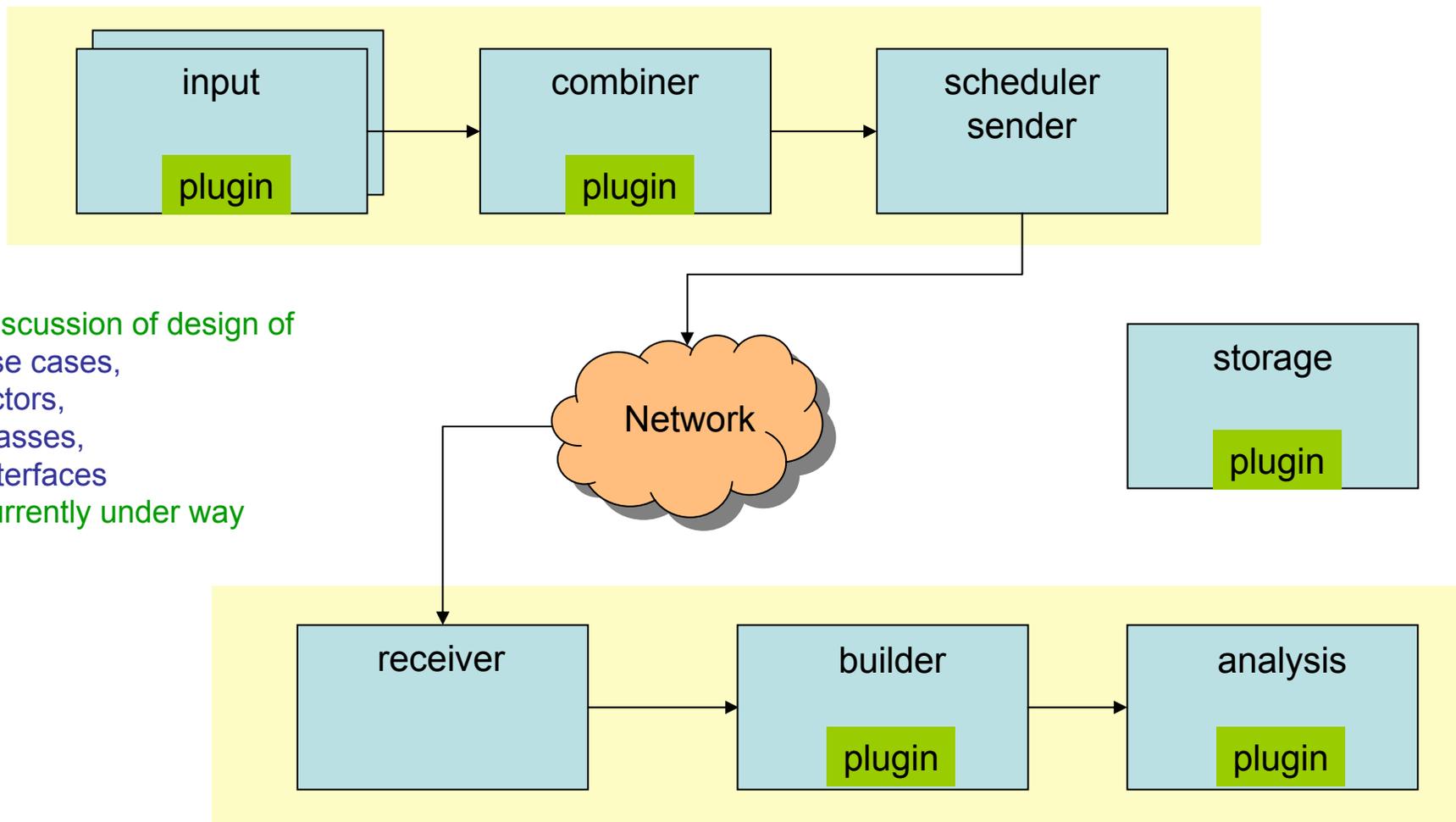
Prototype implementation runs!



Discussion of design of  
use cases,  
actors,  
classes,  
interfaces  
currently under way

Prototype implementation runs!

Discussion of design of  
use cases,  
actors,  
classes,  
interfaces  
currently under way





- Data input channels
  - ABB board: GSI
  - MBS: GSI
  - others
- Combiner/Time sort
  - ABB and MBS: GSI
- Event building
  - ABB and MBS. GSI
- Analysis
  - Framework: GSI
- Storage
  - ROOT IO frame: GSI



- People of data processing group  
H.G.Essel  
J.Adamczewski  
N.Kurz  
S.Linev
- People of controls group  
maybe one FTE
- People from CBM  
hopefully
  
- **CBM requires in 2007 a data taking system**  
**Start with small system, grow on demand**  
**Preliminary controls**

**Tomorrow talks:**

J.Adamczewski: XDAQ developments

S.Linev: Status IB software and event building