Next Generation Data Acquisition Systems for FAIR

based on TrbNet

Jan Michel
Goethe-University, Frankfurt, Germany

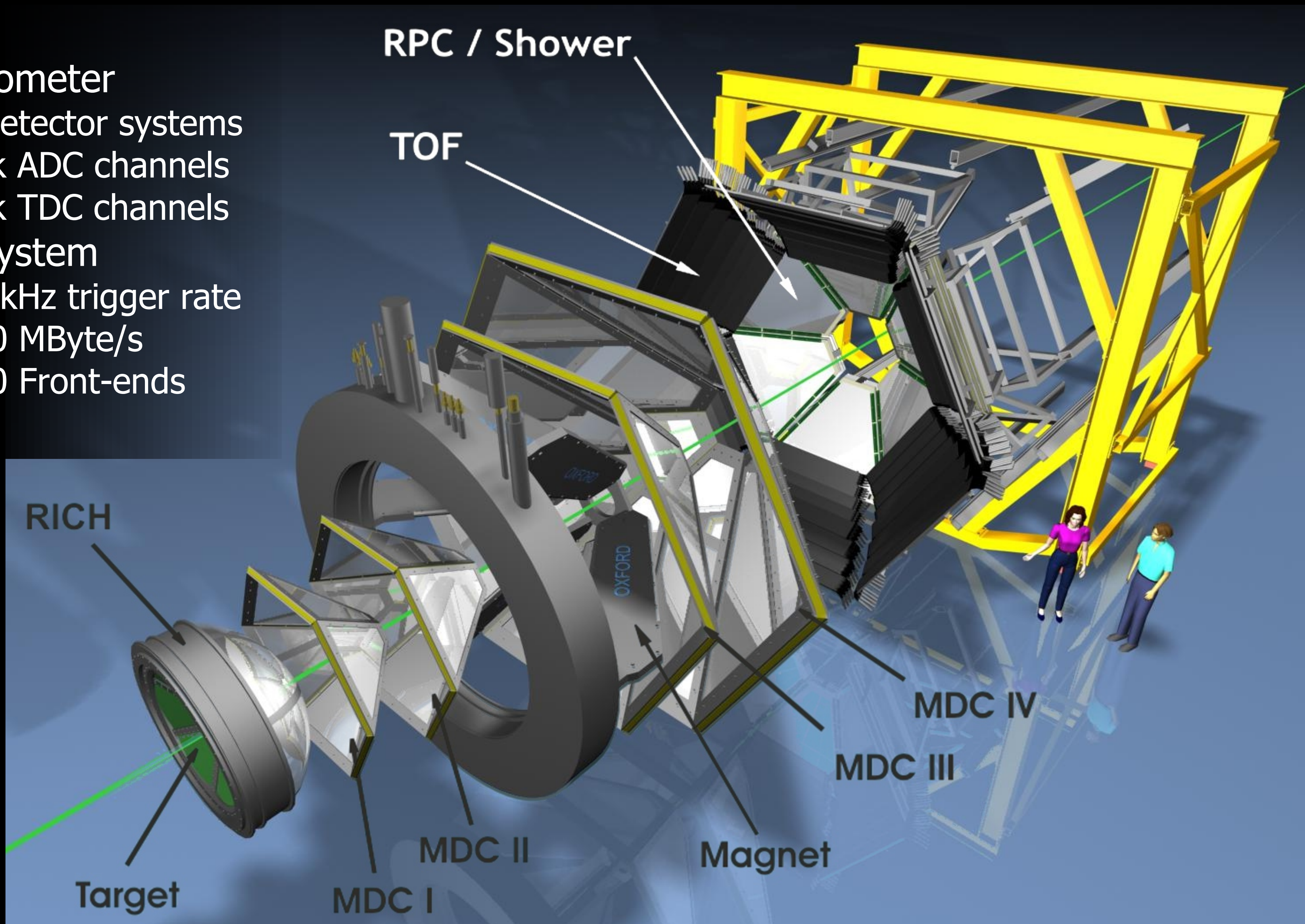
Contact: j.michel@gsi.de

Outline

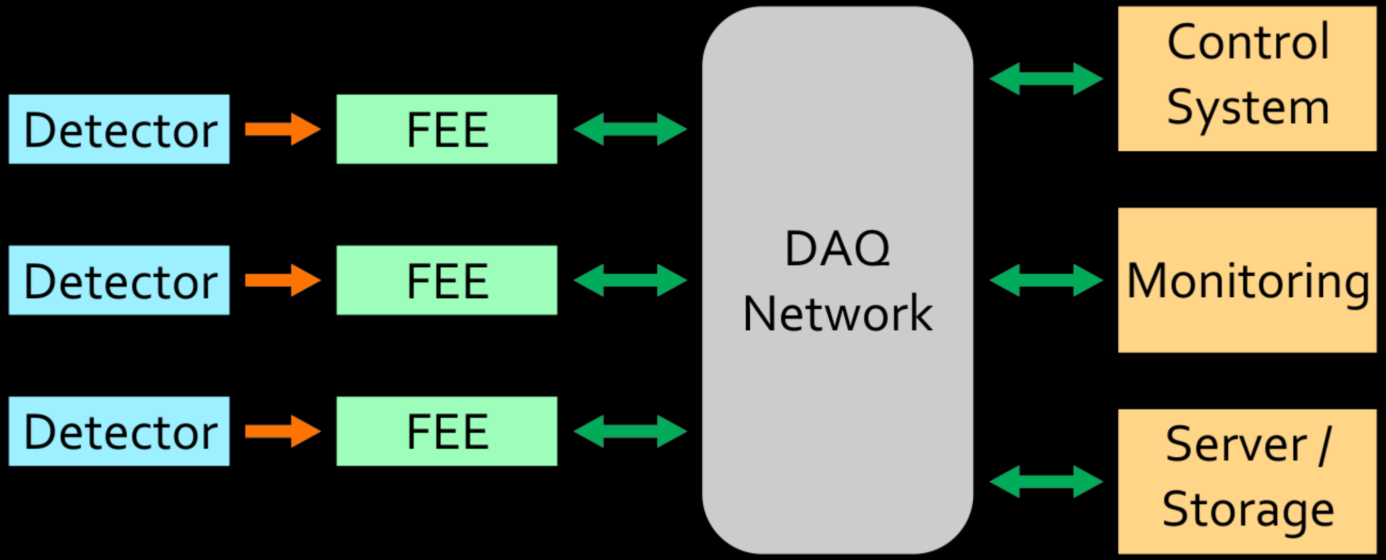
- The HADES Data Acquisition System
  - Set-up and Functionality
  - Succesful Au+Au run (Apr 12)
- Synergy Effects
  - TRB3 as wide-spread platform
  - Example: CBM-RICH test @ Cern, PANDA DISC-Dirc (Nov '12)
- DAQ for FAIR Experiments
  - Time distribution system for PANDA (SODA)
  - CBM-MVD read-out

The HADES Detector @ GSI

- Spectrometer
  - 7 detector systems
  - 50k ADC channels
  - 30k TDC channels
- DAQ System
  - 50 kHz trigger rate
  - 500 MByte/s
  - 500 Front-ends



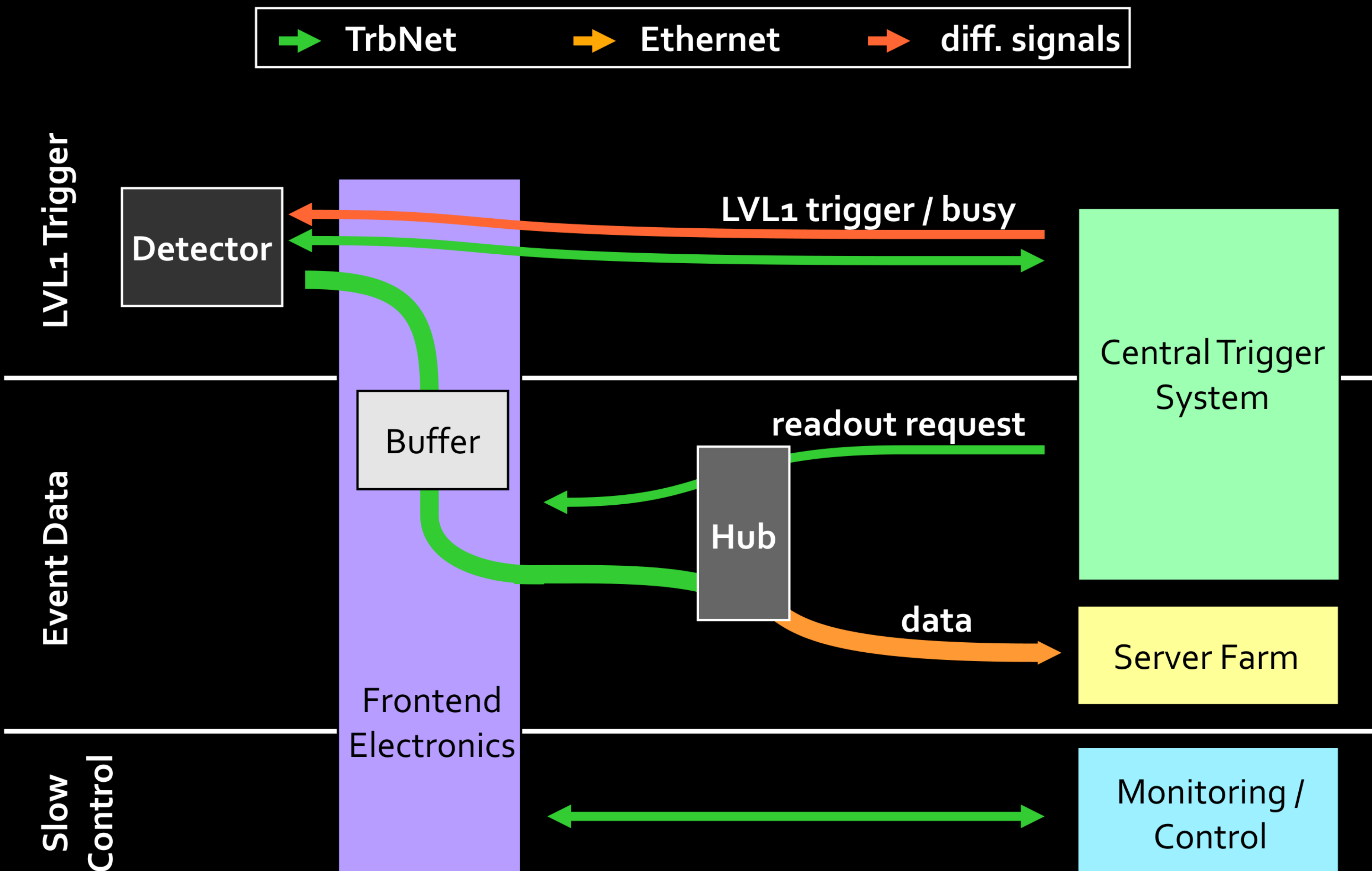
Data Acquisition



HADES DAQ network

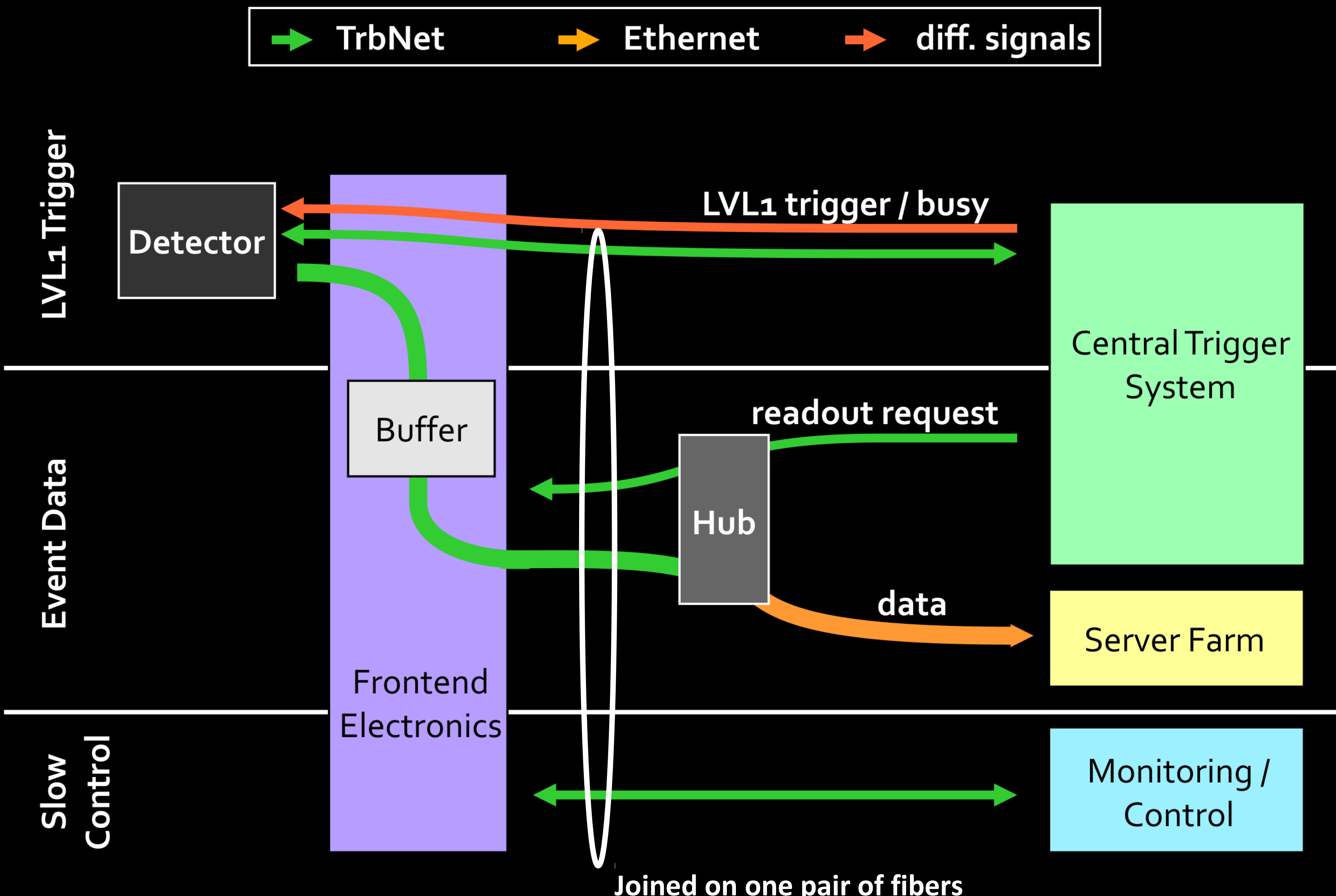
- located inside detector
  - space constraints, optical network
- 50 kHz event rate in trigger-busy-release architecture
  - low latency (5us) triggers, asynchronous read-out
- uniform network for all sub-systems
  - code reuse, identical tools, easy maintenance

DAQ Network



TrbNet
Ethernet
diff. signals
LVL1 Trigger
Event Data
Slow Control
Detector
LVL1 trigger / busy
Central Trigger System
Buffer
readout request
Hub
data
Server Farm
Frontend Electronics
Monitoring / Control

DAQ Network



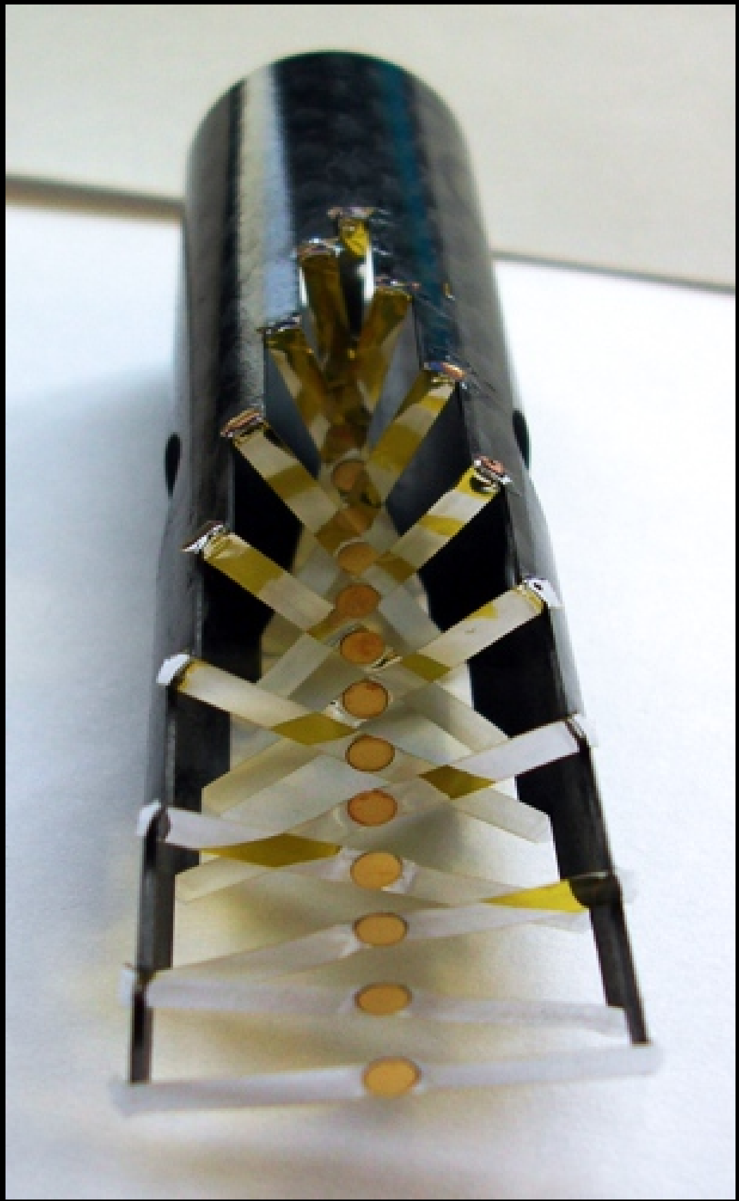
TrbNet
Ethernet
diff. signals
LVL1 Trigger
Event Data
Slow Control
Detector
Buffer
Frontend Electronics
LVL1 trigger / busy
Central Trigger System
readout request
Hub
data
Server Farm
Monitoring / Control
Joined on one pair of fibers

The April 2012 Beamtime

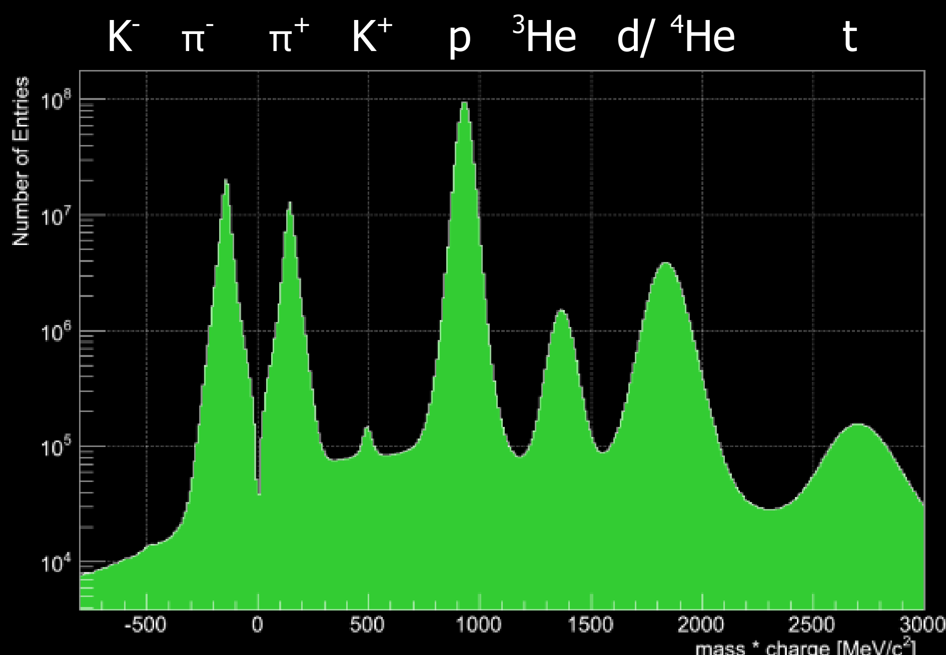
- Au target, 15-fold segmented, 1% interaction
- Au beam, 1.23 AGeV, 2 million ions/second

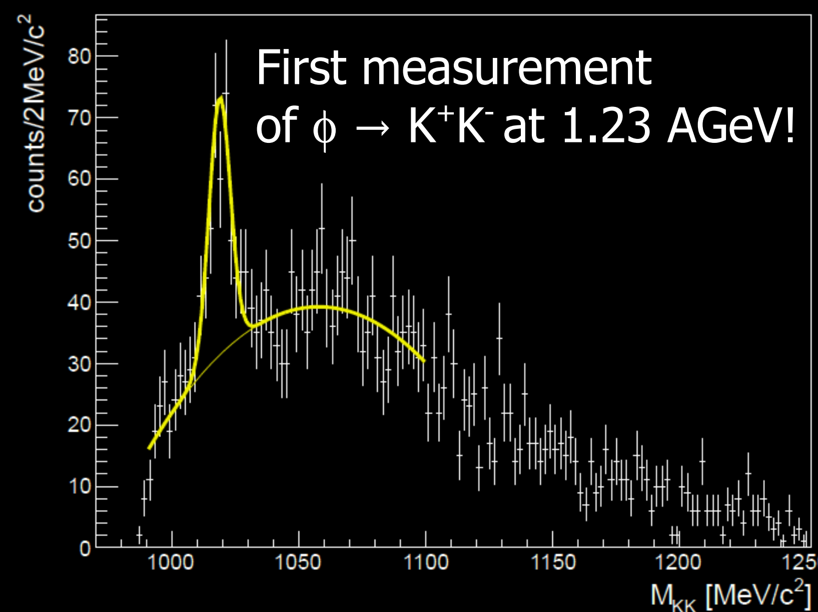
- Typical rates 150 MByte/s, 10 kHz
  - Limited by detectors
  - DAQ test: 500 MByte/s, 65 kHz

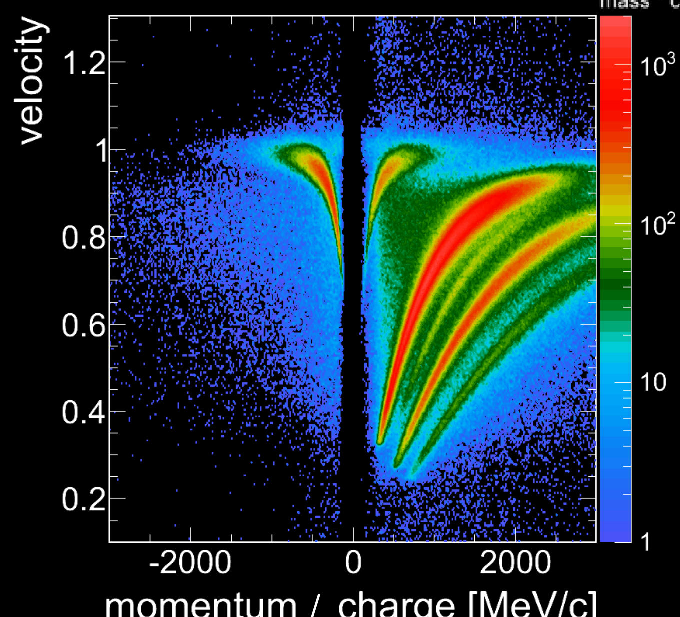
- 5 weeks
- Result: 140 TByte of data, 7.7 billion recorded events



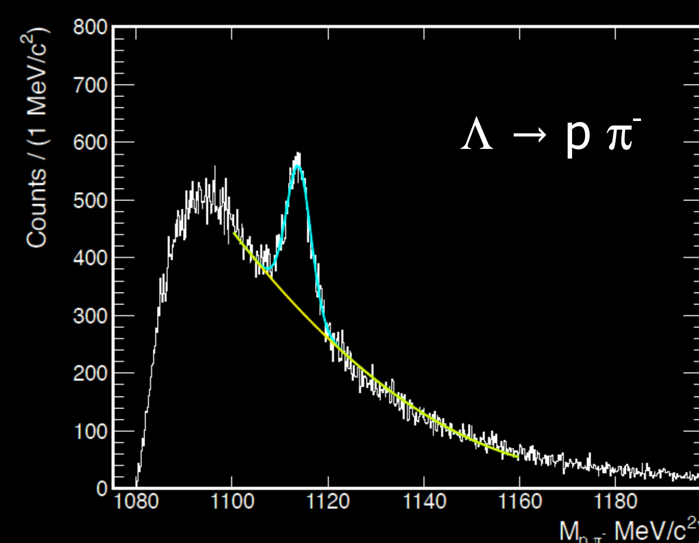
Au + Au (1.23 AGeV) = ?







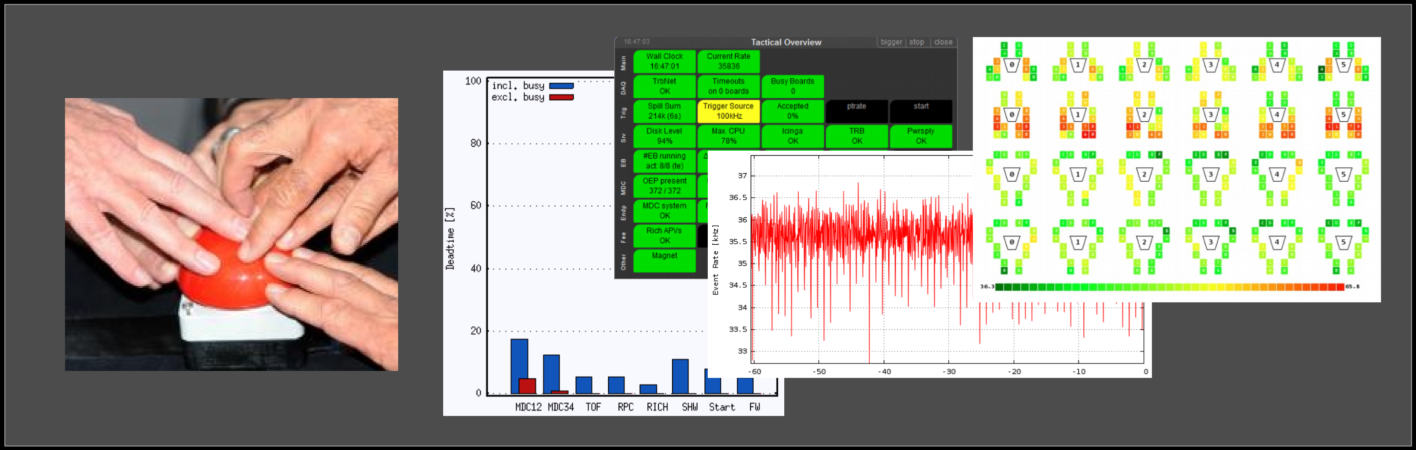
Very preliminary!



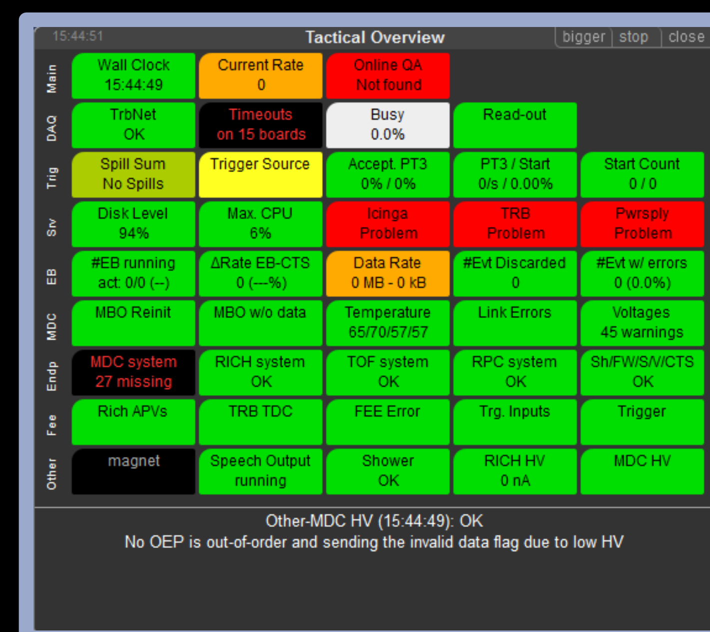
Status Visualization

- All Front-ends are individually adressable and contain a huge number of status and control registers
- Operators for 5-week beam-time are mainly non-experts
  - Intuitive GUI
  - Prepared automatic procedures
  - needed to assure fast handling of all error conditions

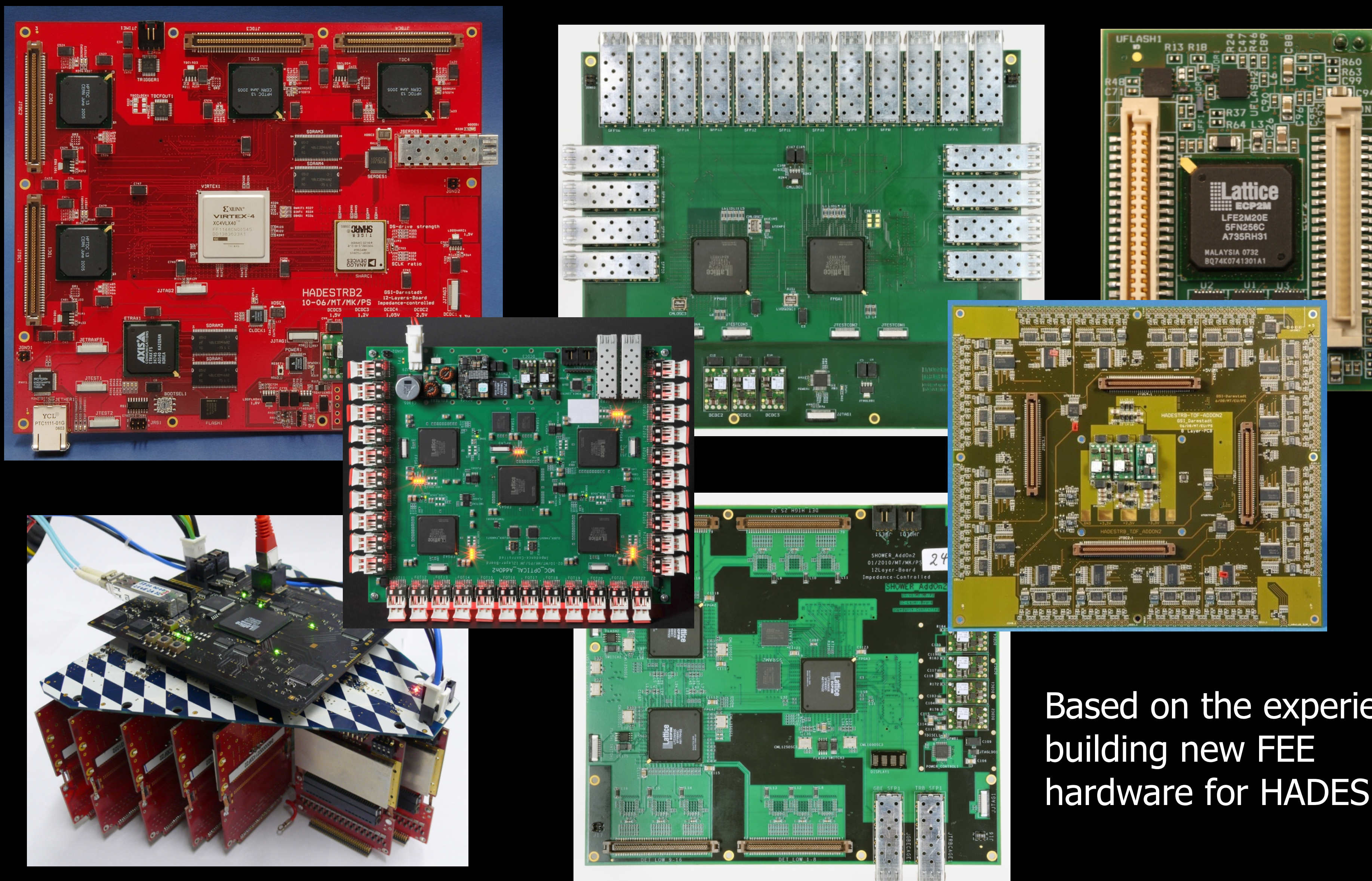
0x00000004	0x00000000	0x00000000	0x00000c00
0x00000100	0x00000000	0x00140020	0x0010de0b
0x00000000	0x00000000	0x00000000	0x00000000
0x00000000	0x00000000	0x00000000	0x00000000
0x0005a860	0x00000000	0x00819518	0x000009a4
0x003bcdfc	0x00408356	0x00000000	0x00000000
0x00000000	0x004d7132	0x00000000	0x03479216
0x00000000	0x00000000	0x00000002	0x00000000
0x8b2bae66	0x0000198f	0x003f4304	0x07ce9695
0x0039b414	0x0000951a	0x00000000	0x00000000
0x00000000	0x00000000	0x00000000	0x00000000
0x00000000	0x00000000	0x00000000	0x00000000



Tactical Overview



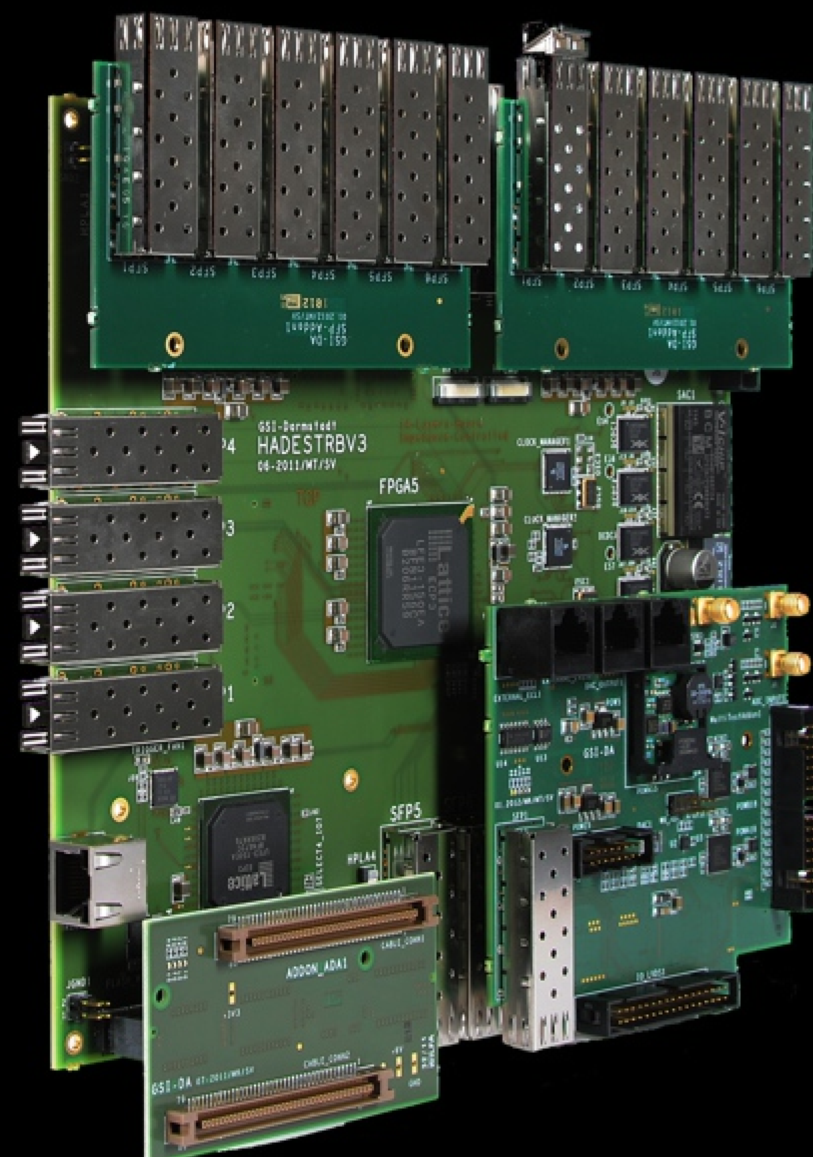
Electronics built for HADES



Based on the experience building new FEE hardware for HADES...

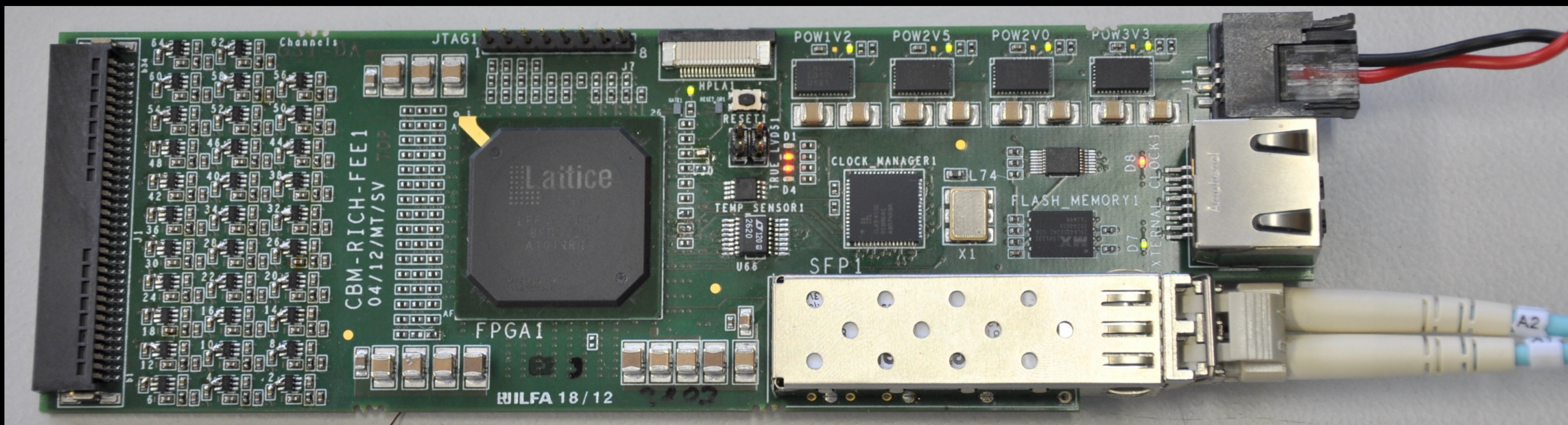
New Versatile Platform: TRB3

- Multi-purpose FPGA platform
- Extension via AddOn-Boards
- 200 I/O per FPGA
- Full control via GbE
  - Stand-alone operation
  - Read-out & Slow-Control
  - Inter-FPGA communication via TrbNet
- Internal trigger system
  - many trigger options
  - connection to other DAQ systems
- Fully scalable to any number of boards



In use / planned to be used / evaluated
by many groups:
HADES (Start, Pion-Tracker)
CBM (RICH, TOF, PSD, MVD)
PANDA (Barrel/Disc-DIRC, STT)
A2 (Crystal Ball/TAPS)
...

CBM-RICH (Nov 12)



Amplifiers

FPGA with
65 TDC ch.

indiv. thresholds

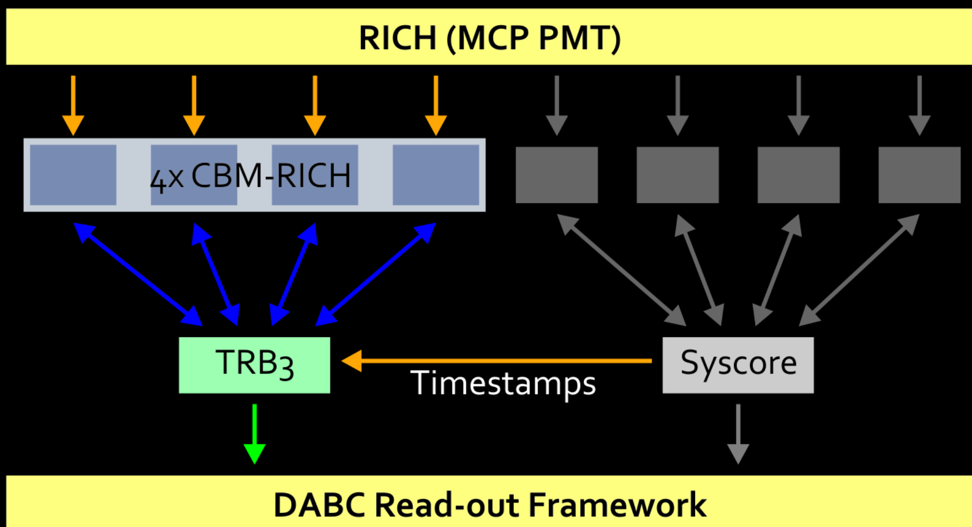
power supply
optical 2 GBit/s link

- CBM-RICH prototype test @ CERN PS
- very short development time:
  6 months from schematics to beam-test
- 64 channel PMT read-out
  - 5x16 cm to be plugged on the back of an MCP-PMT

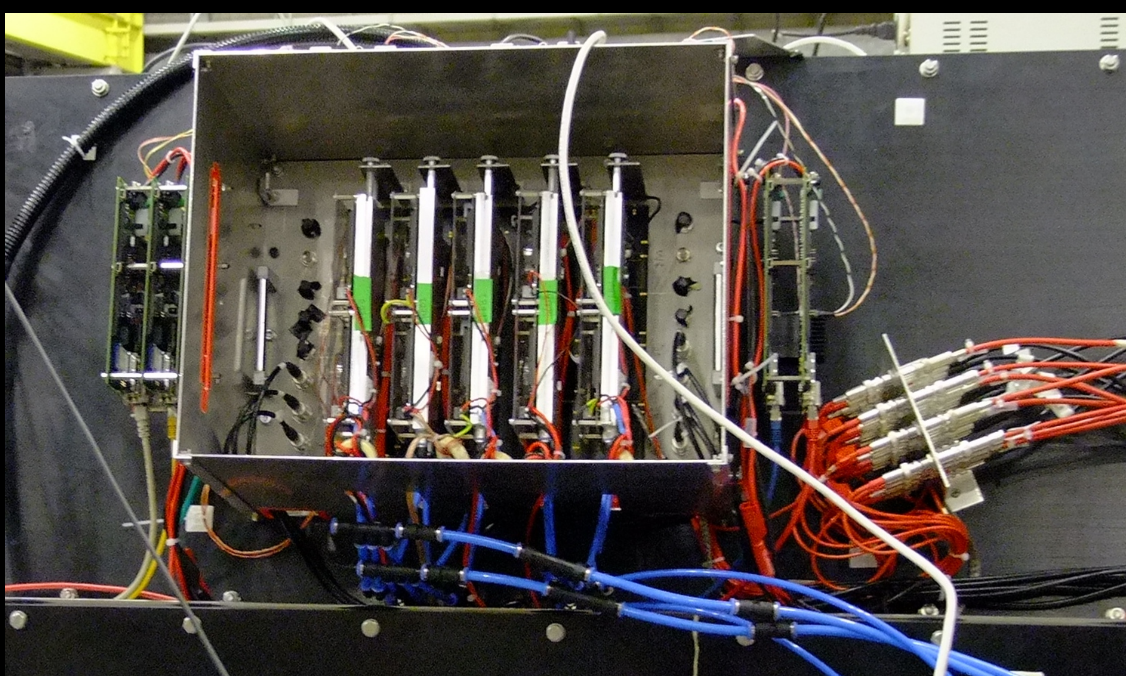
- optical link to TRB3
  - synchronization to DABC read-out of CBM-TRD
  - send data via GbE

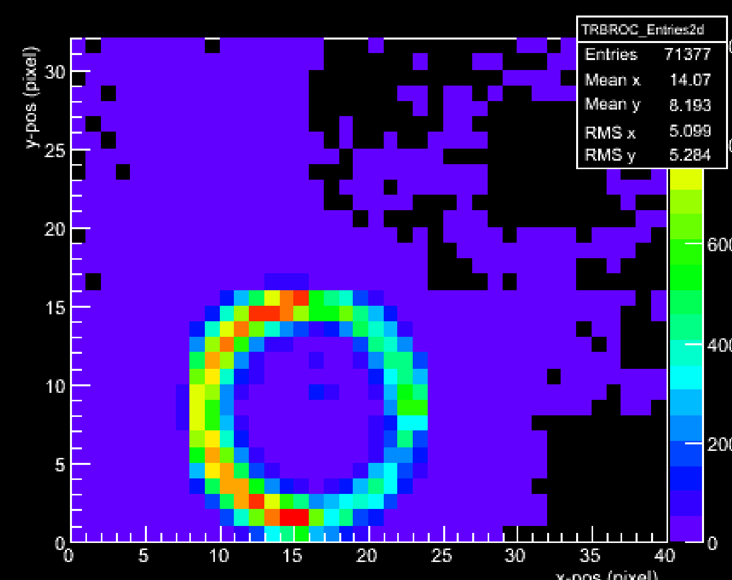
More about hardware: Next talk by Cahit Ugur

CBM-RICH Setup



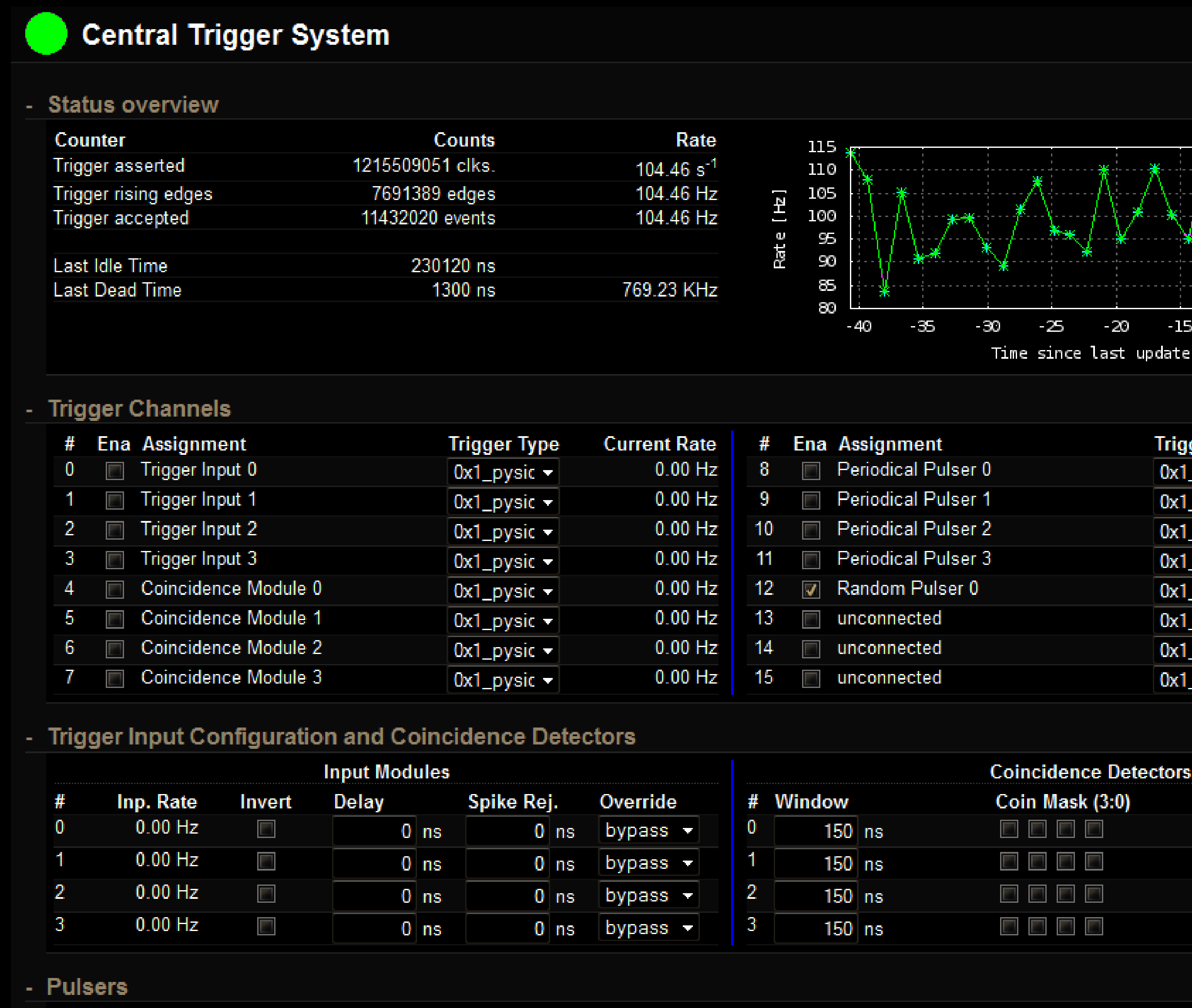
- Combined system to test different front-end hardware
- Full system run synchronized
- Data combined on-line
- TRBnet in quasi-free-running mode
  - triggered for coordinated read-out
  - recording data full time



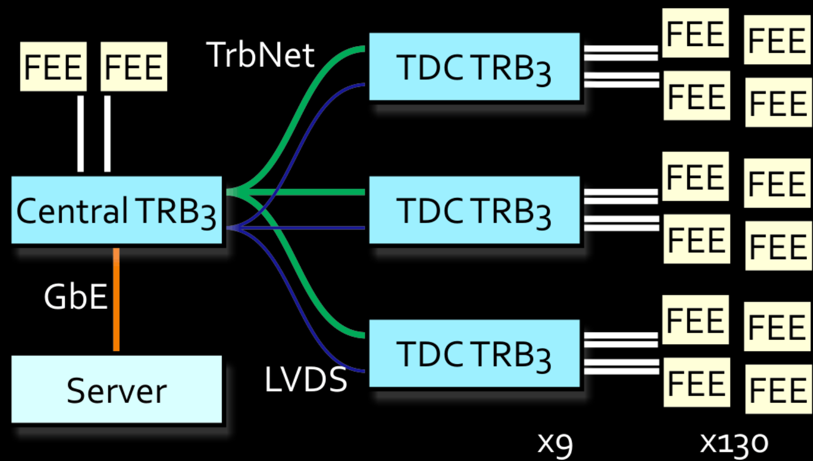


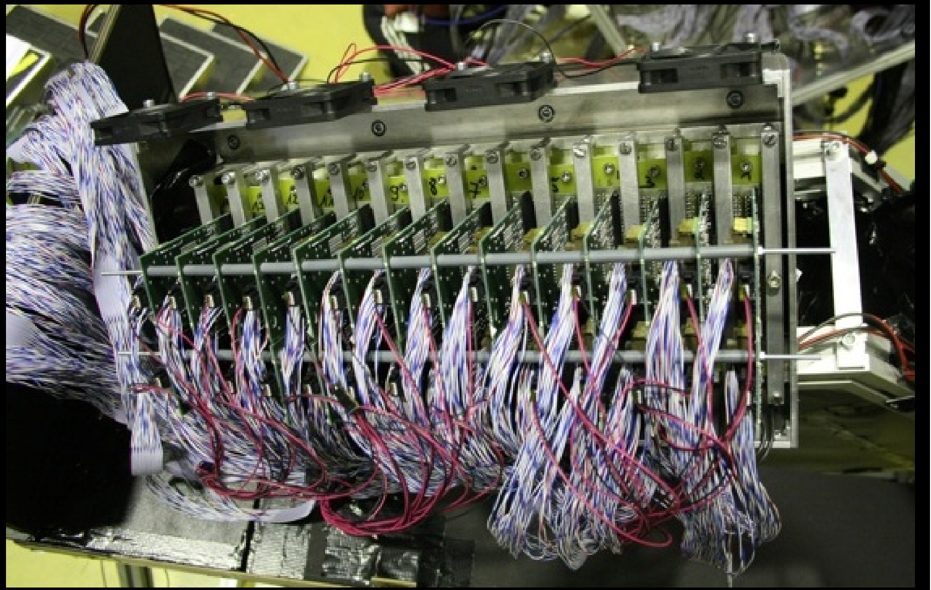
Control System

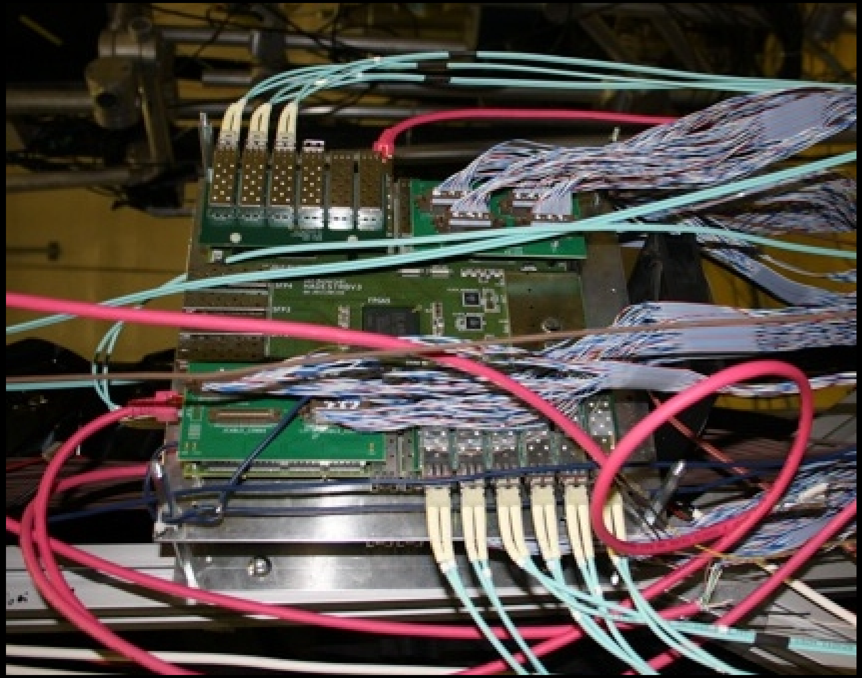
- Control & Monitoring of the TRB3-integrated trigger system
- Full control via web interface



PANDA Disc-DIRC

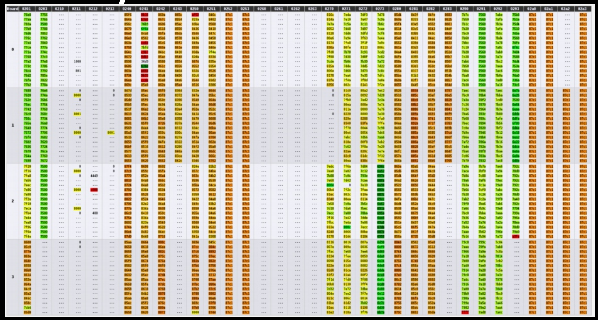






Univ. Erlangen (Adrian Schmidt)
test experiment @ Cosy November 2012

2600 TDC channels + beam hodoscope
Fully stand-alone system



PANDA

- Several detector groups of Panda already use TRB-based hardware
  - in test setups running with TrbNet
- Can parts of the final DAQ be run with TrbNet?

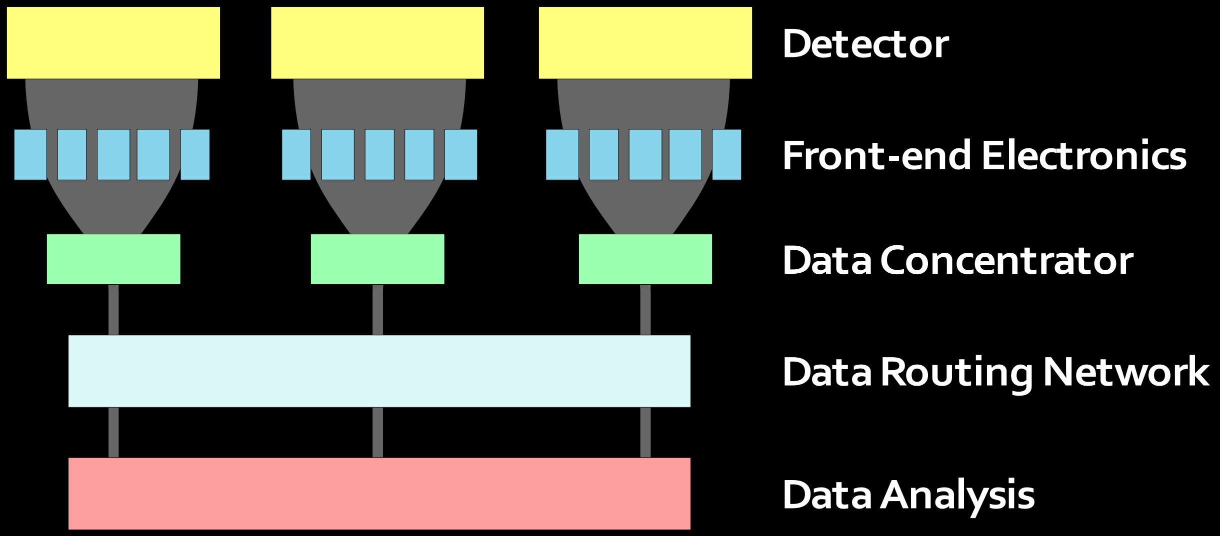
SODAnet for PANDA: Requirements

Synchronous messages

- Reach all read-out controllers with small offset (ideal: < 10 ns)
- High timing accuracy (most systems: jitter < 100 ps)
- Synchronize to beam structure in storage ring (2us beam + 0.4us break)
  - Distribute common burst number
- Send control messages (start, stop, calibration...)

Slow-Control

- Debugging & status reporting
- Online configuration...



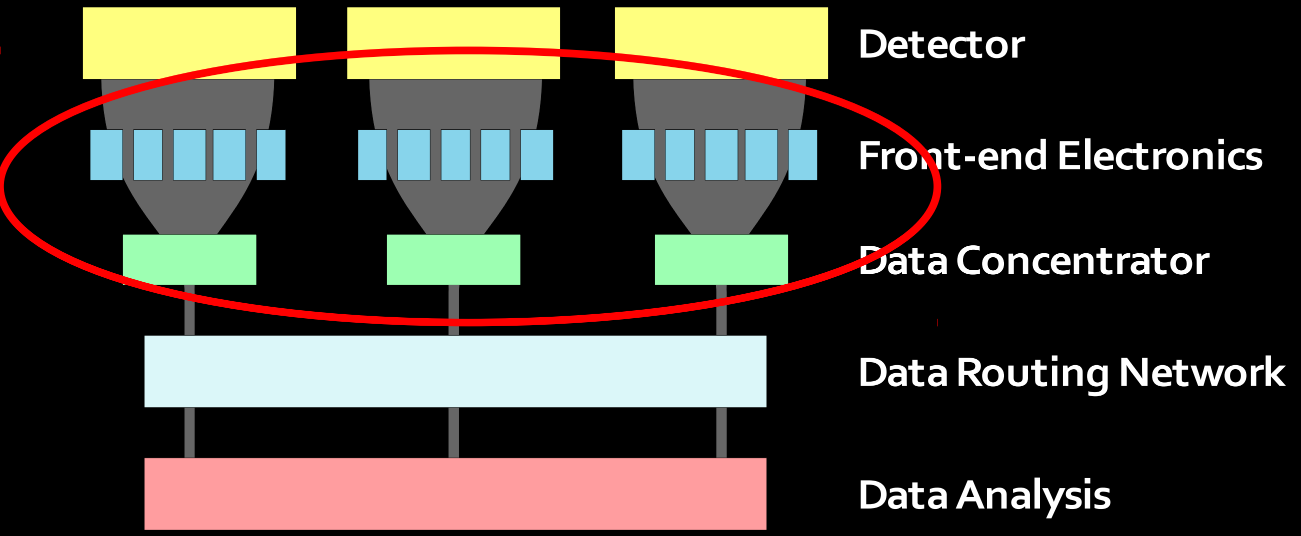
SODAnet for PANDA: Requirements

Synchronous messages

- Reach all read-out controllers with small offset (ideal: < 10 ns)
- High timing accuracy (most systems: jitter < 100 ps)
- Synchronize to beam structure in storage ring (2us beam + 0.4us break)
    - Distribute common burst number
- Send control messages (start, stop, calibration...)

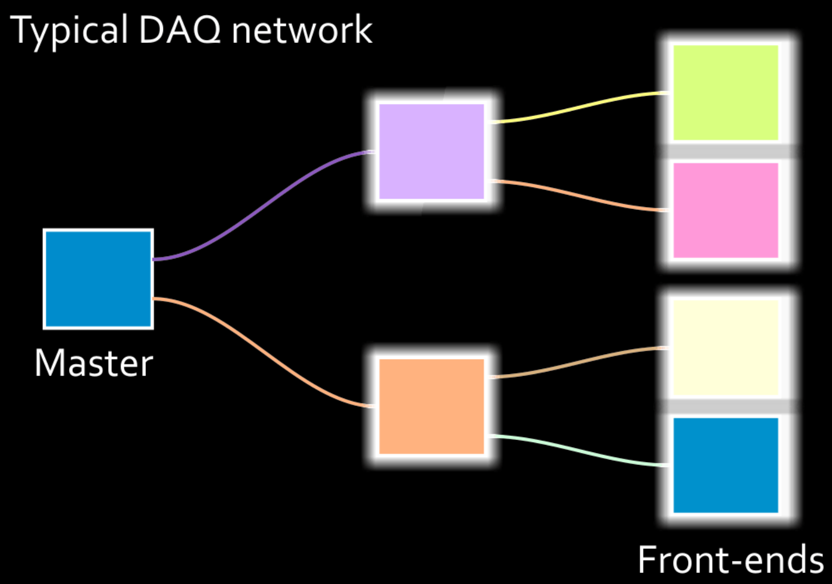
Slow-Control

- All boards accessible for debugging & status reporting
- On-line configuration of read-out, pre-processing modes...

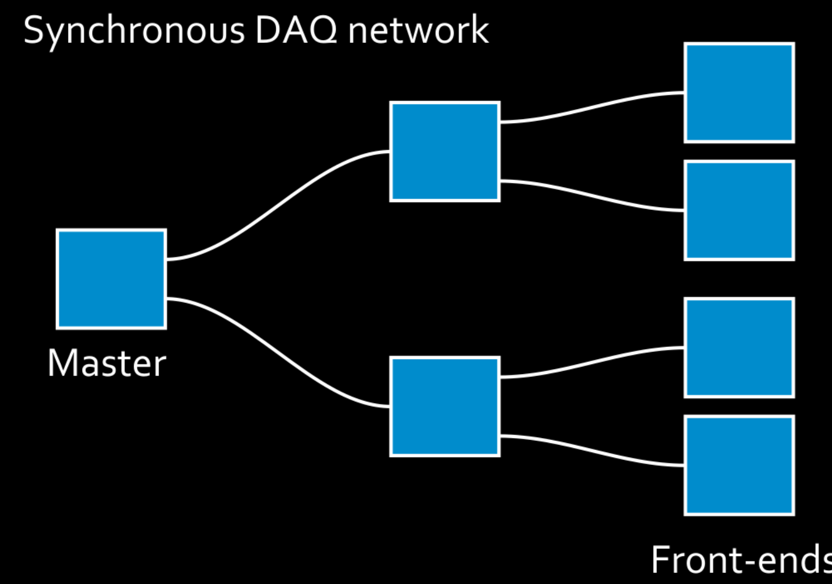


=> Can be based partly on TrbNet

Synchronous Networks



- Usual network (as in HADES)
  - all components run independently
  - buffers within data stream
  - high-level features to "make life easier"
- → delays change over time
  - precision ~ 100 ns
  - external signals for synchronization

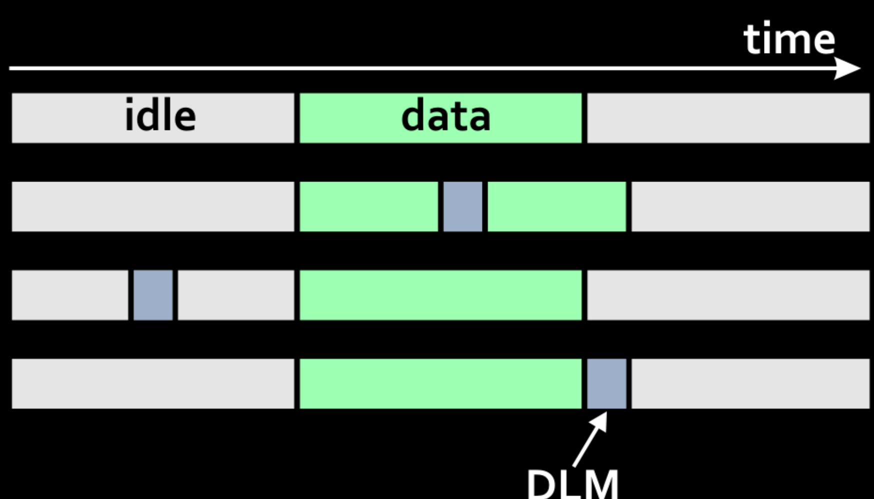


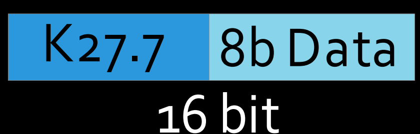
- Deterministic network
  - run all boards with the same clock
  - remove all non-determinism

→ phase & latency between clocks unknown, but constant
  - can be corrected off-line
  - precision (jitter) < 100 ps

Synchronous Messages

- Have a look on the implementation in CBMnet, because:
  - Combination of TrbNet with CBM infrastructure neccessary, e.g. for CBM-MVD
  - It is available and thoroughly tested
- Insert a defined data word at any time into the data stream
  - Special words can not appear anywhere else in the data
  - ”Deterministic Latency Message” (DLM)
  - Restrict to 16 defined messages → possibility to correct single bit errors





- Reuse existing definitions:
  - development work only needed once
  - easier combination of mixed networks

SODA Messages

- Define two types of messages, each with 32 Bit payload
  - "Start of burst" to mark the beginning of each super-burst (16 bursts)
    - number of super-burst used to idenfy data
    - sent in fixed intervals
  - "Control" to select different operation modes, trigger calibration...
    - checksum to prove correctness of packet
    - can be sent at any time

- Link length measurement by returning messages
  - resolution: 4 ns (byte clock) is trivial, 400 ps (bit clock) should be possible
  - used for adjustment of delays in sending DLMs

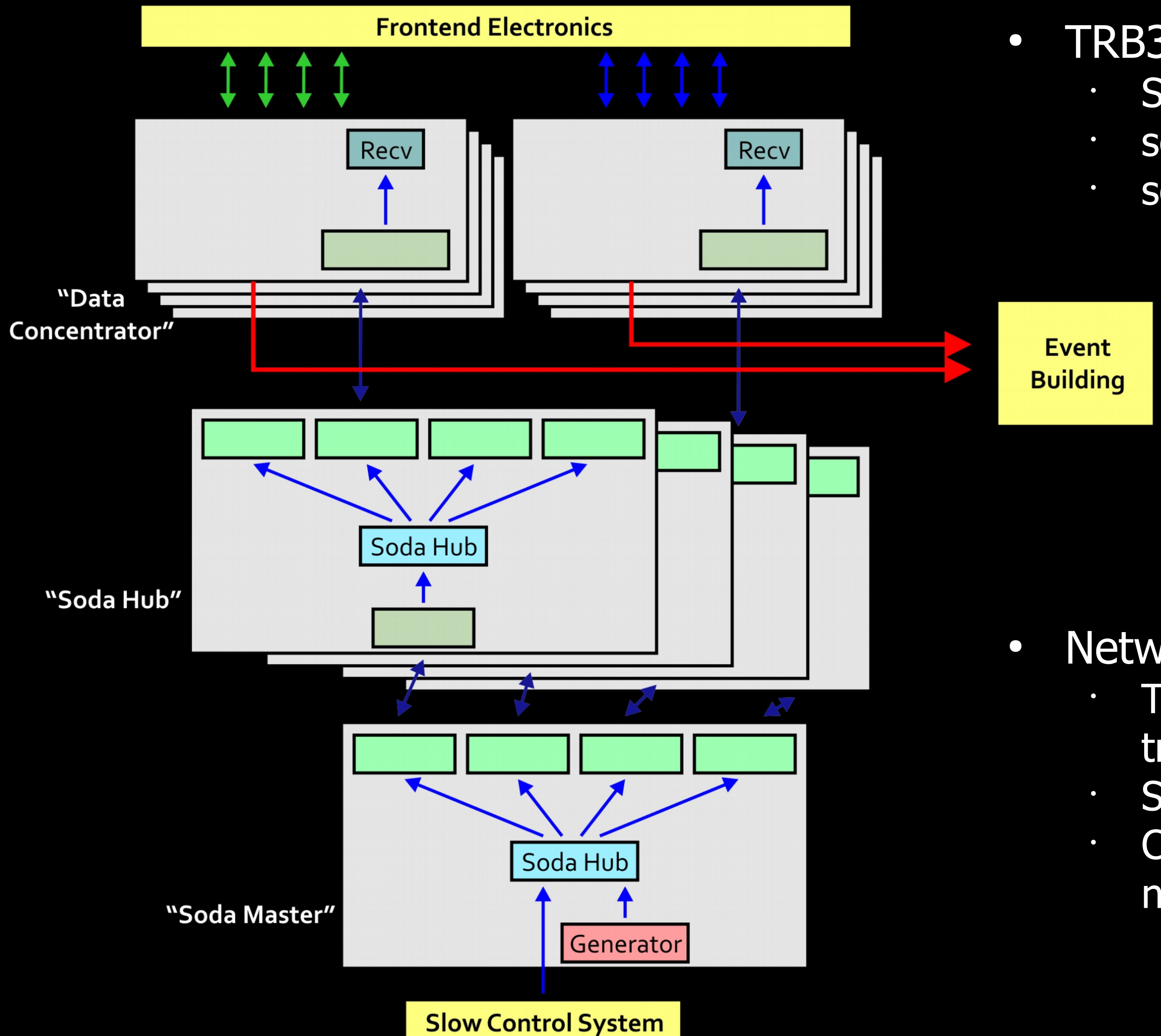
K27.7	cnt 1	K27.7	cnt 2	K27.7	cnt 3	K27.7	cnt 4

"Start of burst": 31 Bit super-burst number

K27.7	ctrl 1	K27.7	ctrl 2	K27.7	ctrl 3	K27.7	check

"Control": 24 Bit for control tasks & checksum

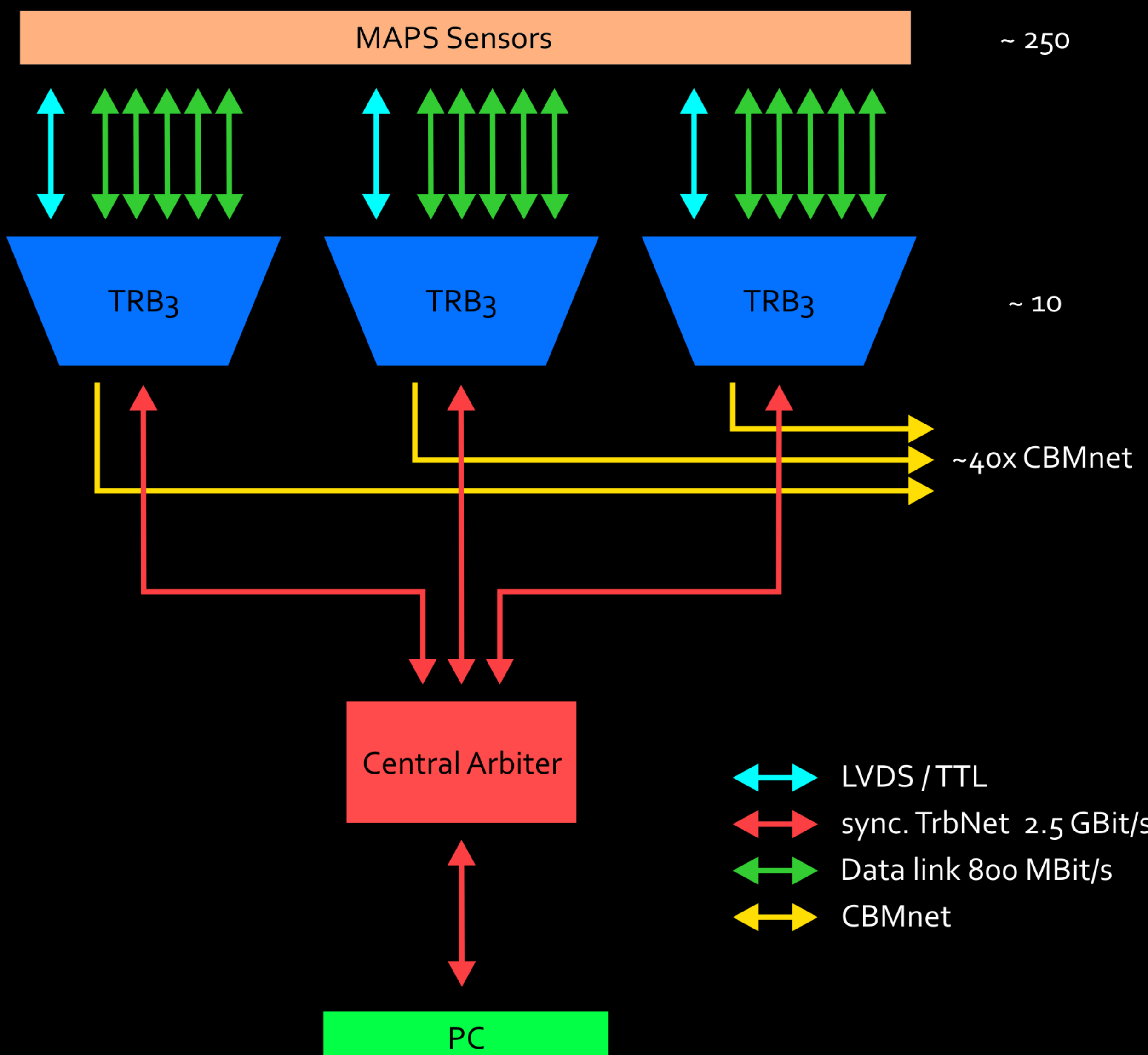
SODAnet for PANDA

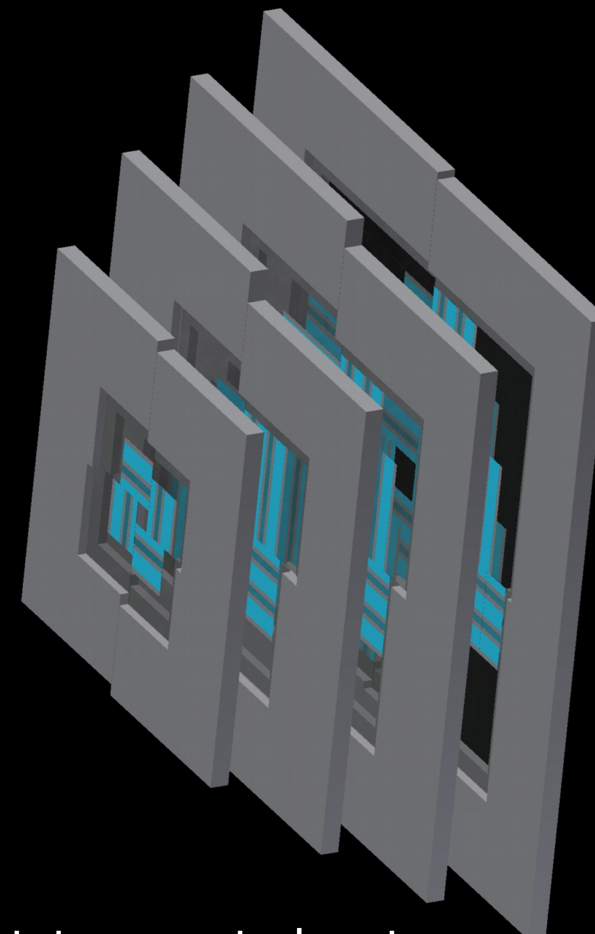


- TRB3 will be used as platform:
  - Soda Hub & Master
  - some Data Concentrators
  - some Frontends

- Network Protocol
  - TrbNet as basis for data transport and slow-control
  - SODA messages added
  - Concentrator → Eventbuilding: not yet decided, GbE?

CBM-MVD





Cur. prototype control system runs with TrbNet,

may be replaced with CBMnet + additional slow-control features later

Timescales

Work based on existing systems gives fast results:

- CBM-RICH
  - First idea: January
  - Board fabrication: April
  - Beamtest: November
- Panda SODA
  - January: DAQ workshop
    - Decision: Use TrbNet as basis for SODA
  - February:
    - Bi-directional communication in synchronous FPGA network
    - Validation of fixed-latency of messages
  - April (planned)
    - Finish first implementation of Soda message master & receiver
    - Test automatic link latency measurements
  - No full-time worker on this project!

The TRB3 "Collaboration"

- Hardware: Michael Traxler (GSI)
- TDC: Cahit Ugur (GSI)
- GbE: Grzegorz Korcyl (Krakow)
- CTS: Manuel Penschuck (Frankfurt)
- Data Network: Jan Michel (Frankfurt)
- Slow-Control Framework: Ludwig Maier (Munich)
- Data Unpacker: Matthias Hoek (Mainz)
- Event-Server: Jörn Adamczewski-Musch (GSI)
- DABC: Sergey Linev (GSI)

- and all the people...
  - … who write there own additional code
  - … who helped designing and implementing TrbNet
  - … who were involved in the HADES DAQ upgrade
  - … I forgot to mention

Conclusion

- The TRB3 platform is in use by many different groups and experiments
  - 70 TRB3 in operation up to now
- Trb3 and TrbNet are an easy-to-run solution to acquire data
  - both for small test setups and huge detector setups
  - compatible to other trigger systems
  - almost plug'n'play – few software parts to install, few cables to connect
- Rather small modifications are needed to adapt the HADES DAQ network to other FAIR experiments:
  - run as synchronous network
  - messages with deterministic latencies
  - free-running read-out
- The DAQ concepts for all FAIR experiments share many similar aspects
- Not everything needs to be as individual as it may seem

Next Generation Data Acquisition Systems for FAIR

based on TrbNet

3rd Hic for FAIR Detector Systems Networking Workshop

Next Generation Data Acquisition Systems for FAIR

based on TrbNet

3rd Hic for FAIR Detector Systems Networking Workshop

"Network networking makes your net work"

HADES DAQ System: Design Criteria

- Highly granular (500 FEE)
  - Placed directly on detector
  - Strong space constraints for electronics and cables
- High trigger rate (50 kHz)
  - Compatible to trigger-busy-release architecture
    - With asynchronous data transfer
  - Low latency transport of trigger information (< 7 us Trigger System - FEE)
- High Bandwidth (500 MByte/s)
- Simple Maintenance
  - High amount of reuseable code
  - Common architecture of all sub-systems

Triggerless Systems

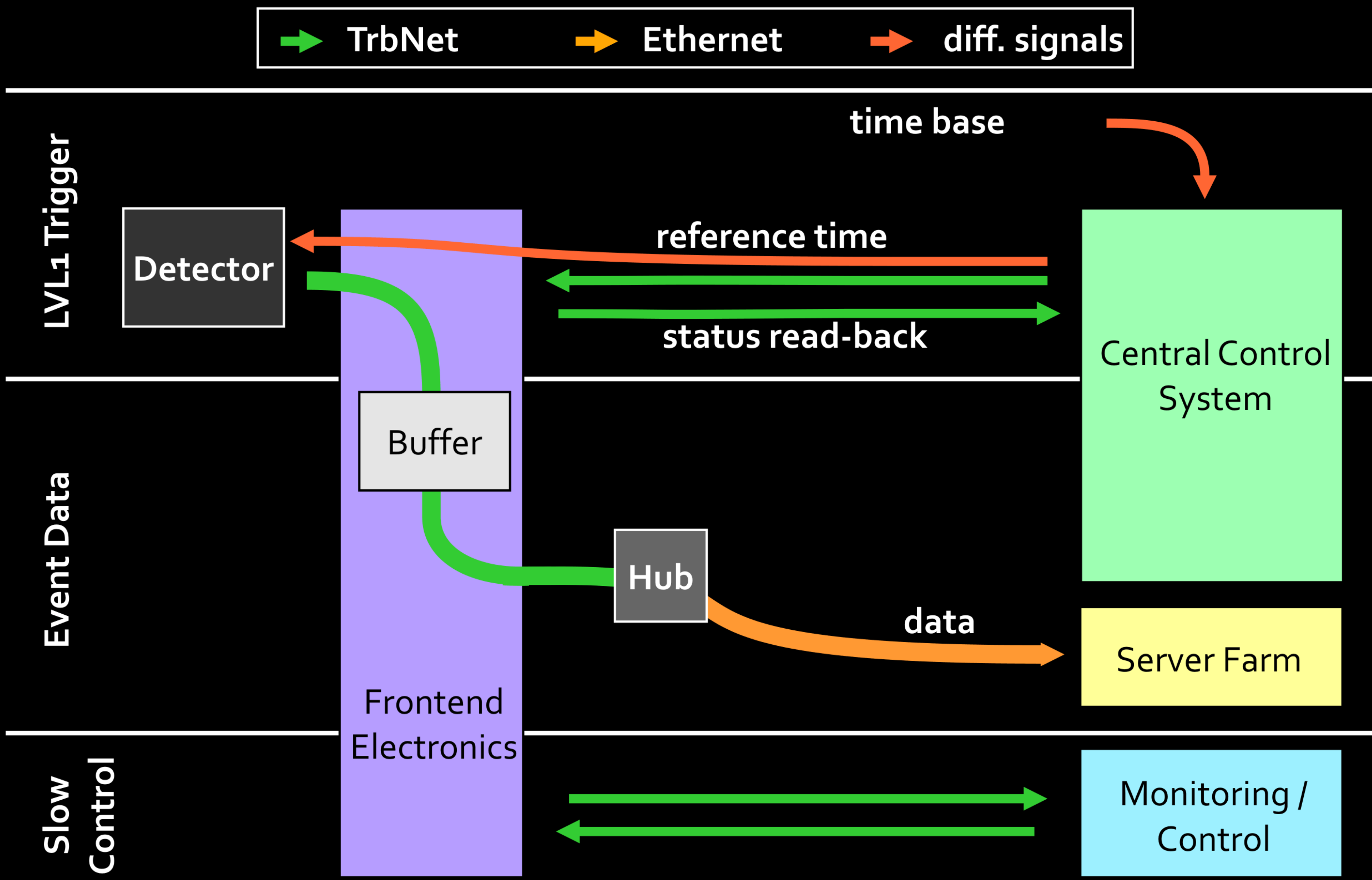
- Conventional DAQ uses triggers
  - e.g. hand-shake between all sub-systems for each recorded event
  - robust
  - easy to handle, discrete data

- Introduces time overhead for each event
- Not usable in high rate experiments (PANDA, CBM ...)

- "Free running" architecture
  - all sub-systems continuously collect data
  - data streams are sent to computing farm

- Need for time stamping

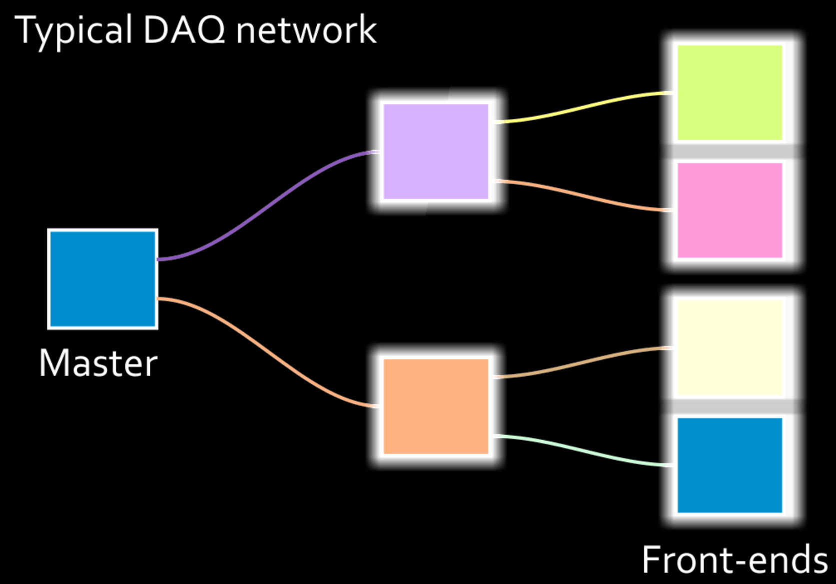
Triggerless System

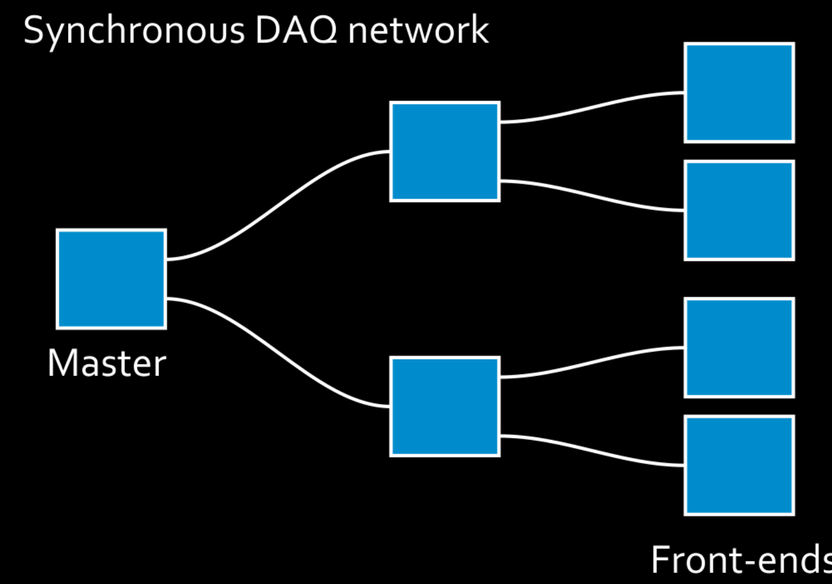


Time Synchronization

- Task: Multiple front-ends need a precise time reference to combine measurements from different parts of a detector
- Classical approach:
  - Put a dedicated cable to each board
- In densely packed detectors, any additional cable should be avoided
  - An optical link for data transport is available

A high-speed serial link is not a cable

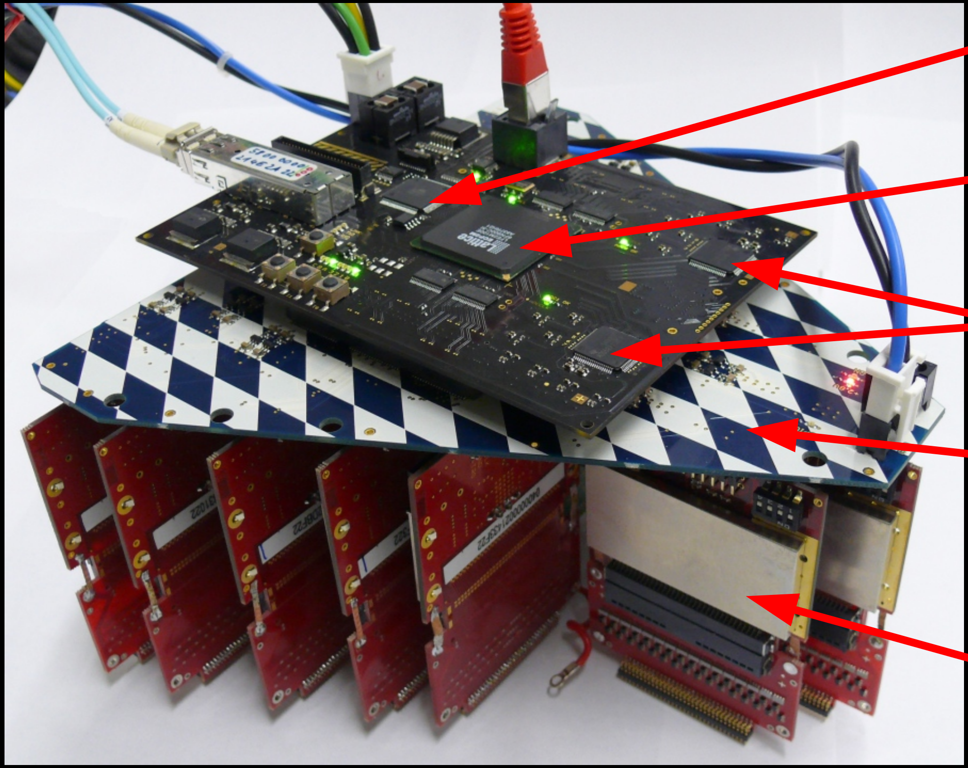




- Normal usage
  - data stream is recovered in each intermediate device
  - clock is extracted, data is parallelized
  - buffers in receivers for data
  - high-level features to ”make life easier”
  - → delays change over time
    - precision ~ 200 ns

- Deterministic network
  - run all boards with the same clock
  - remove all non-determinism
  - → phase & latency between clocks unknown, but constant
    - can be corrected off-line
    - precision (jitter) < 100 ps

Typical Set-up

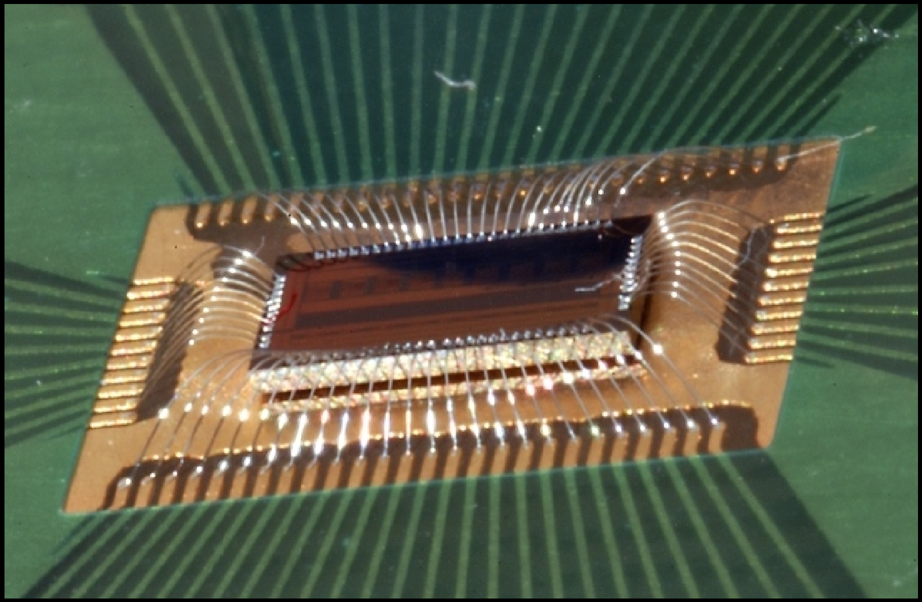


- Controller
- FPGA
- ADC
- Interconnection
- Signal Shaping & Amplification

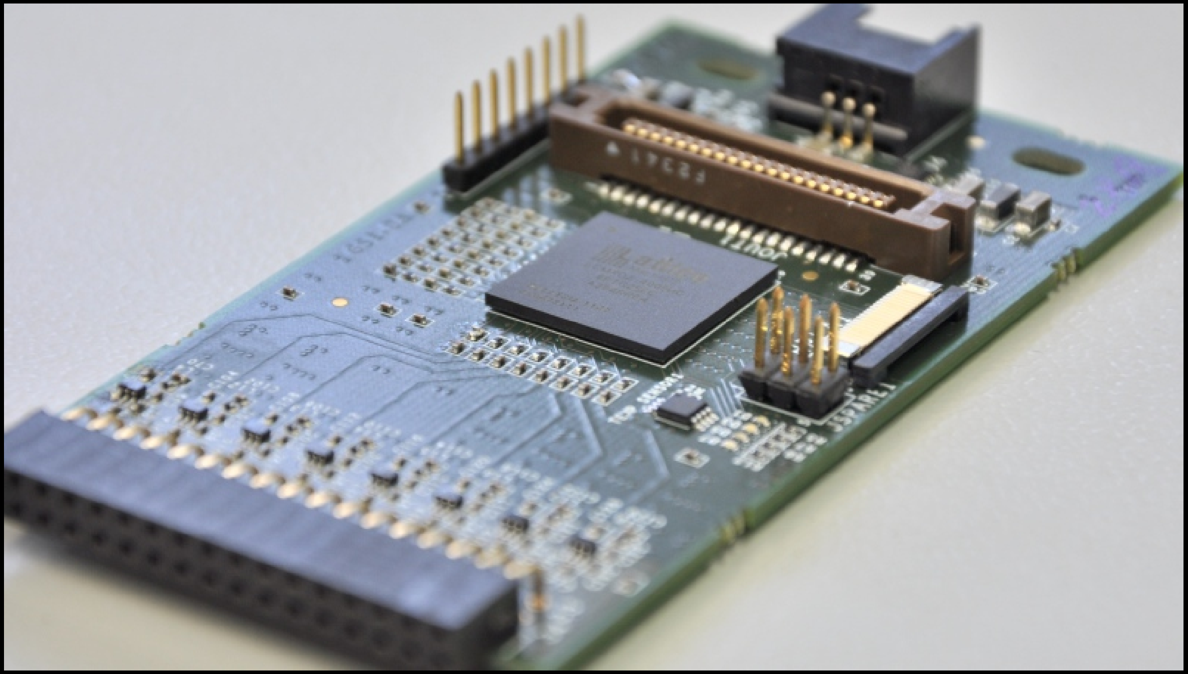
One step further

- Typically, FPGAs serve as digital data processors only
- Analog signals are handled by dedicated chips

- Many customized chips (ASIC) needed for our applications
  - **Very helpful for special applications** (e.g. designed for radiation hardness, tailored to specific constraints...)
  - Common disadvantages:
    - Very long design and production phase
    - Expensive due to low quantities
    - No big vendors, i.e. no guaranteed lifetime
    - Difficult to handle because of non-standard housings
    - If it's versatile, it doesn't fit perfectly



FPGA based signal processing



- 40x amplification
- adjustable thresholds
- digital signal shaping
- control via SPI:
    - temperature, input status, edge count, id, non-volatile memory, in-system programmable

- 16 channels
- component cost: 26 €
- power consumption: 1.2 W
- e.g. analog 500 uV, 6 ns signals → LVDS, 20 ns → $\Delta t \sim 200$ ps

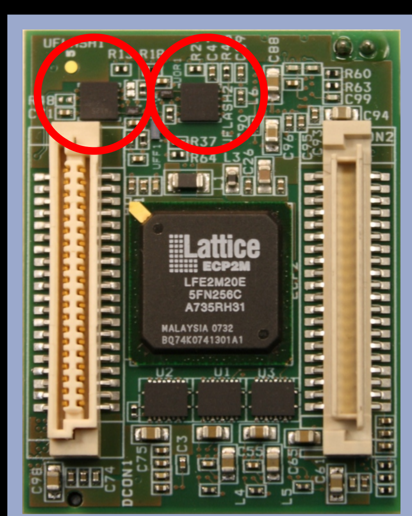
Conclusion

- FPGA are well-suited for common DAQ systems
- The flexible architecture allows to adapt to any front-end electronics

- By using FPGA in non-standard ways, one can
  - measure time
  - discriminate signals
  - measure analog properties with ADC, QDC
- Very flexible designs, easy to adapt to other requirements

Slow Control Features: Updates

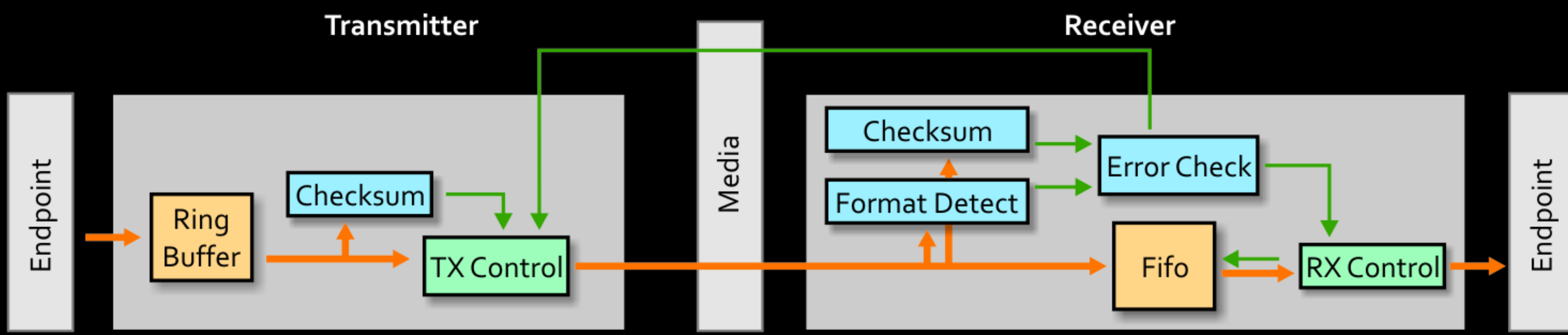
- How to upgrade firmware in ~400 FPGA distributed all over the detector?
  - Connect flash ROMs to logic in FPGA
  - One command in software framework
  - 30s to program + 60s to verify (full system, ~500 MByte)
- Failsafe?
  - Dual-boot with "golden image"




```
> trbflash program 0xfffd bit/mdcoep_20120307a.bit
Found 372 Endpoint(s) of group MDC_OEP_V3
NAME: mdcoep_20120307a.bit
DATE: Wed Mar  7 22:30:59 2012

Start programming ImageFile
'bit/mdcoep_20120307a.bit'
Programming Endpoint(s) @ Address 0xfffd
Symbols:
  E: Erasing
  P: Programming
  @: Success
  .: Skipped

Block: 0 1 2 3 4 5 6 7 8 9 A B C D E F
0      @ @ @ @ @ @ @ @ @ @ @ @ . . . .
1      . . . . . . . . . . . . . . . @

Success

Verifying Endpoint(s) @ Address 0xfffd

Block: 0 1 2 3 4 5 6 7 8 9 A B C D E F
0      @ @ @ @ @ @ @ @ @ @ @ @ . . . .
1      . . . . . . . . . . . . . . . @

Success
```

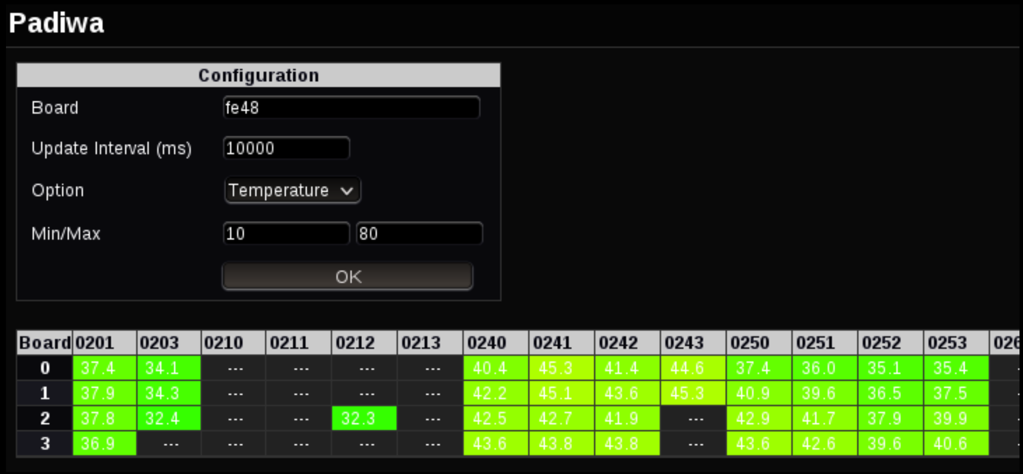
Transmission Error Correction

- Optical transceivers show BER > 10^{-}



- Error topology was even used to track down HV problems in TOF PMTs

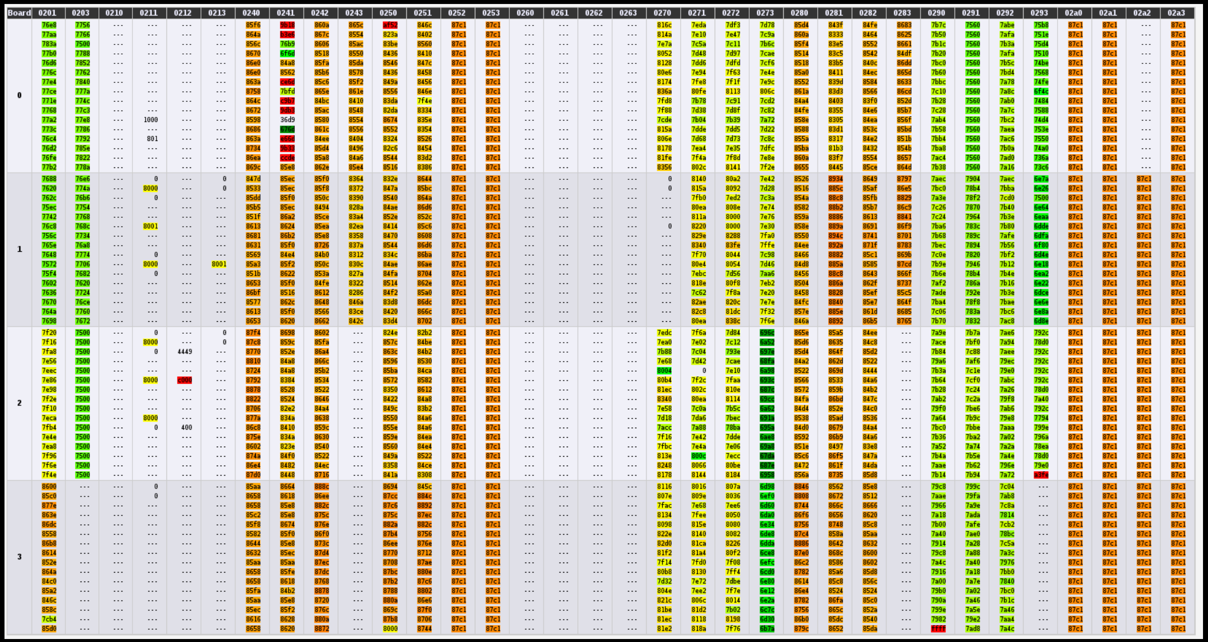
Monitoring for Disc-DIRC



Board	0201	0203	0210	0211	0212	0213	0240	0241	0242	0243	0250	0251	0252	0253	026
0	37.4	34.1	---	---	---	---	40.4	45.3	41.4	44.6	37.4	36.0	35.1	35.4	-
1	37.9	34.3	---	---	---	---	42.2	45.1	43.6	45.3	40.9	39.6	36.5	37.5	-
2	37.8	32.4	---	---	32.3	---	42.5	42.7	41.9	---	42.9	41.7	37.9	39.9	-
3	36.9	---	---	---	---	---	43.6	43.8	43.8	---	43.6	42.6	39.6	40.6	-

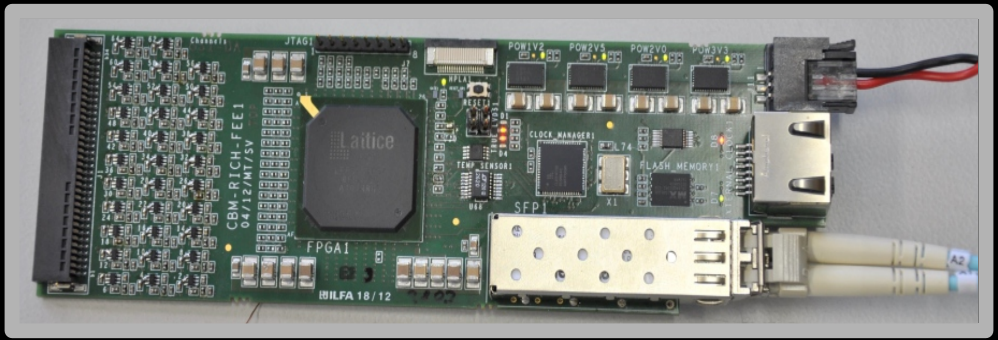
Temperatures of few FEE

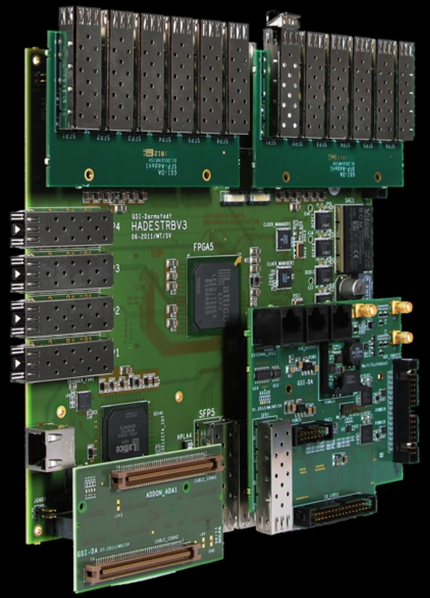
Individual thresholds for 2600 channels

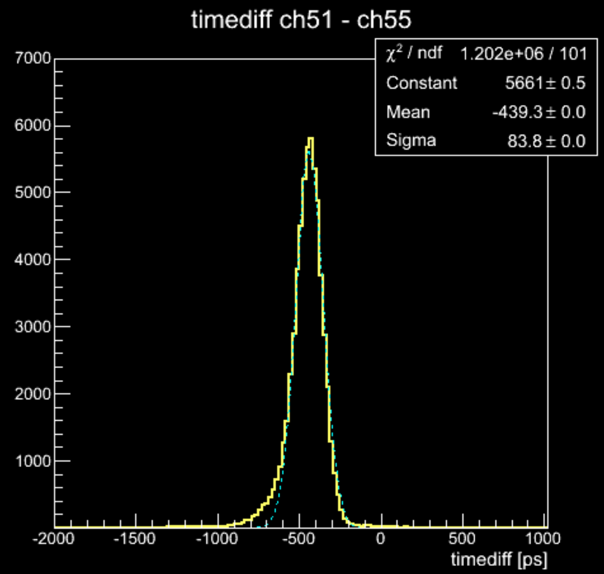


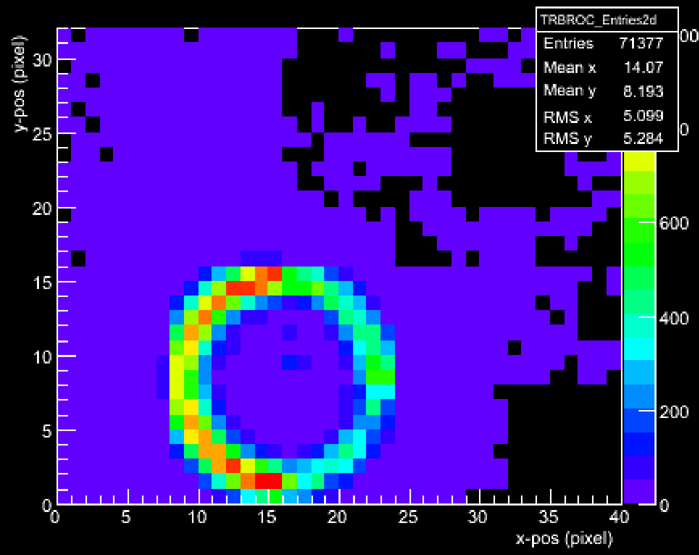
E.g. CBM-RICH

- CBM-RICH test-beam at CERN, Nov 12
- unfortunately only 84 ps time-resoultion due to some layout errors
- Combined with DABC based read-out for other components

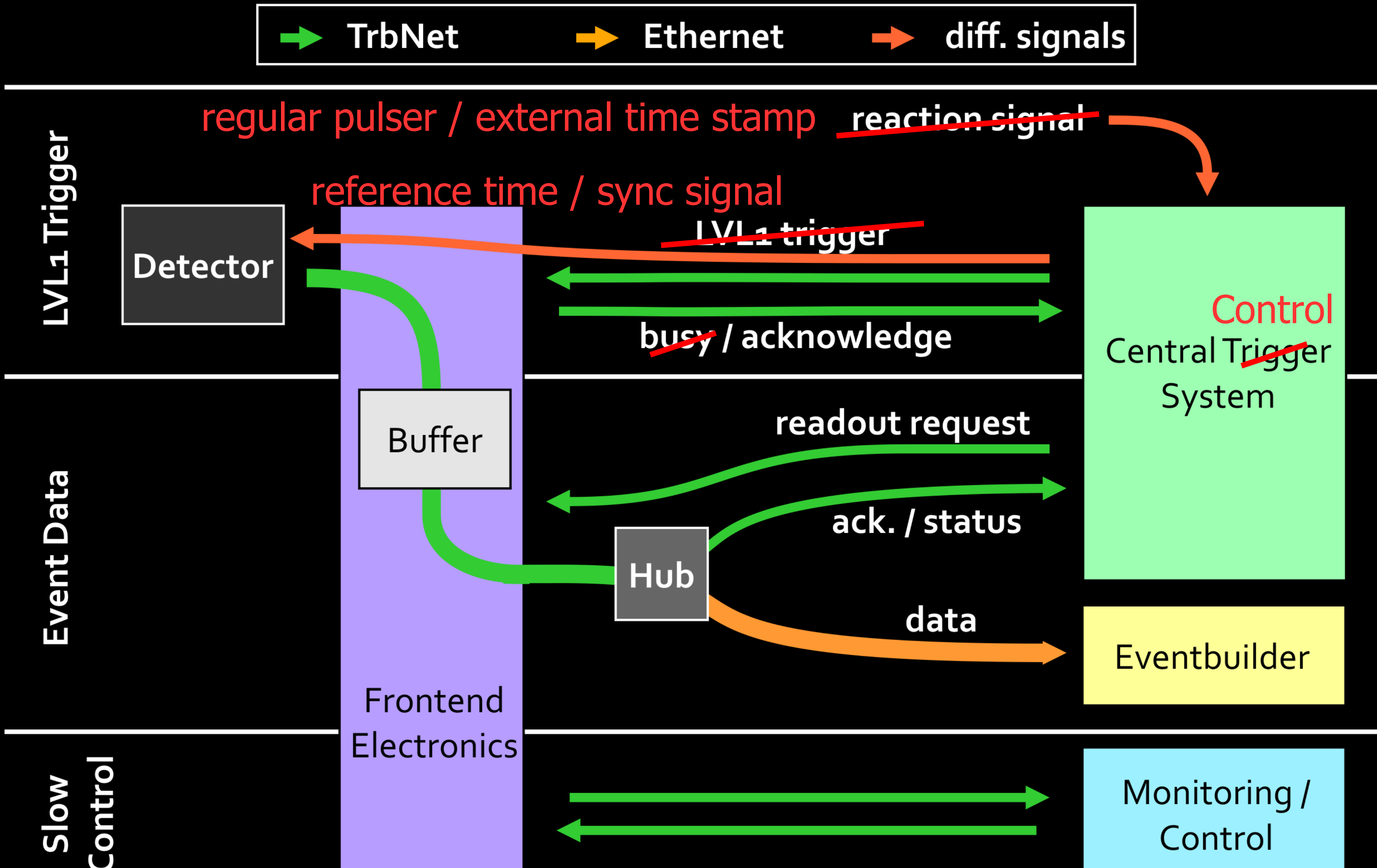






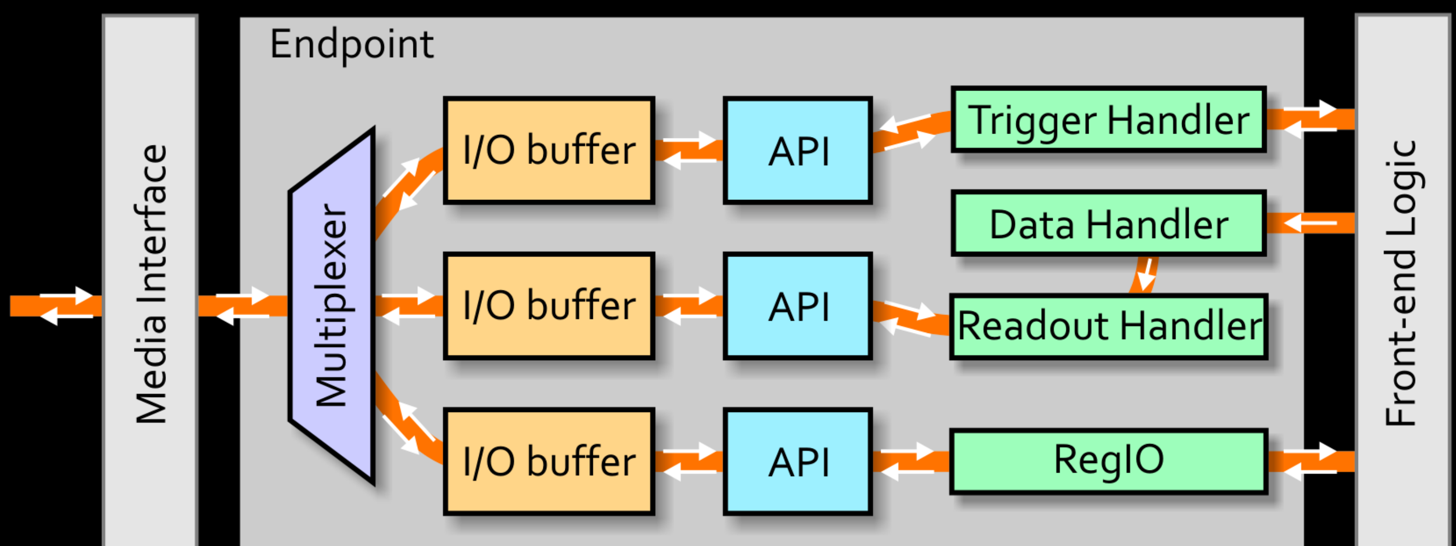


Triggerless?



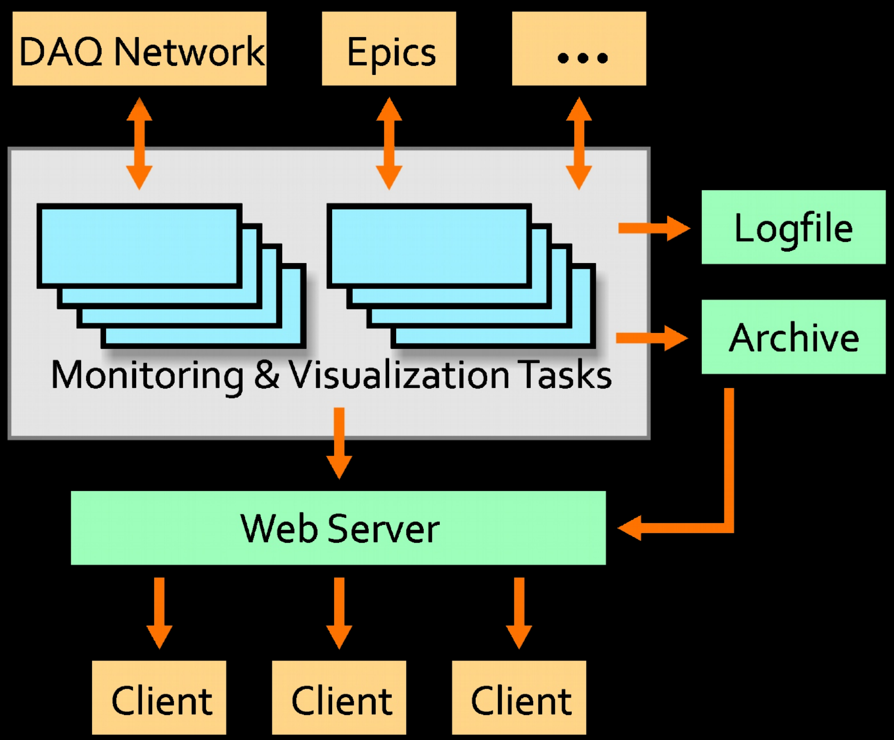
The Network Endpoint

- All network entities are encapsulated
- the "user" only has to deal with few signals
  - Trigger information & status feedback
  - Data input
  - Slow control bus
- All buffering, read-out control, error checking is internal

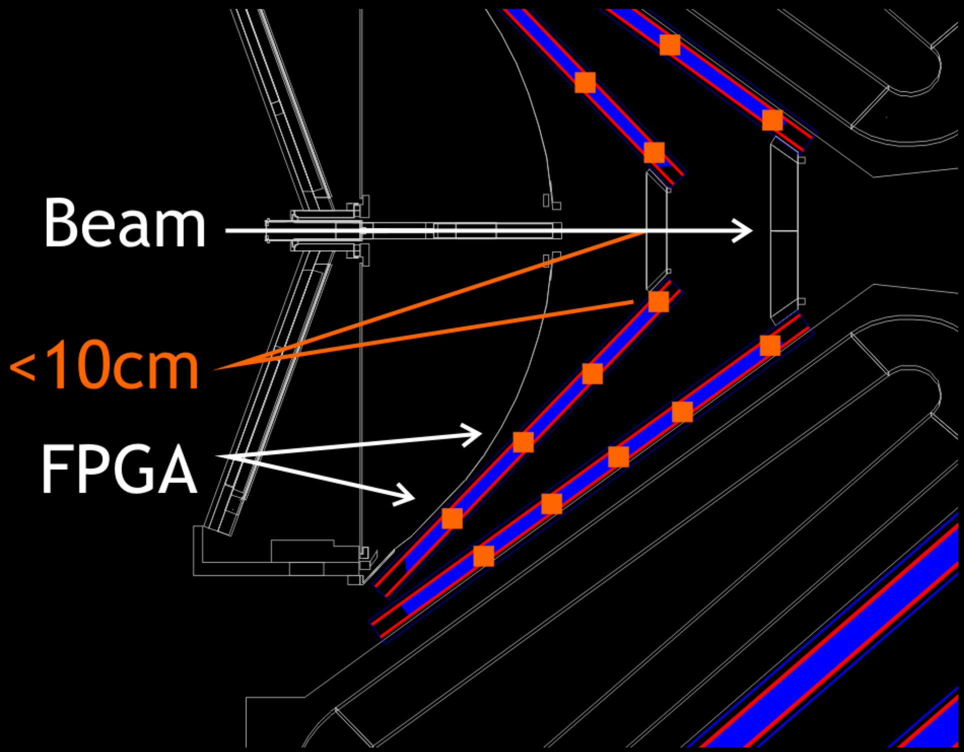


The Hmon Framework

- Standard Tools
  - Perl Environment
  - Apache Webserver
  - Gnuplot
- Interface Libraries
  - DAQ-Network (C-Library)
  - Epics
  - Icinga / Nagios
  - All system tools
- GUI
  - XHTML / JavaScript based
  - Viewable from any device & anywhere
  - Access to data plots from Root

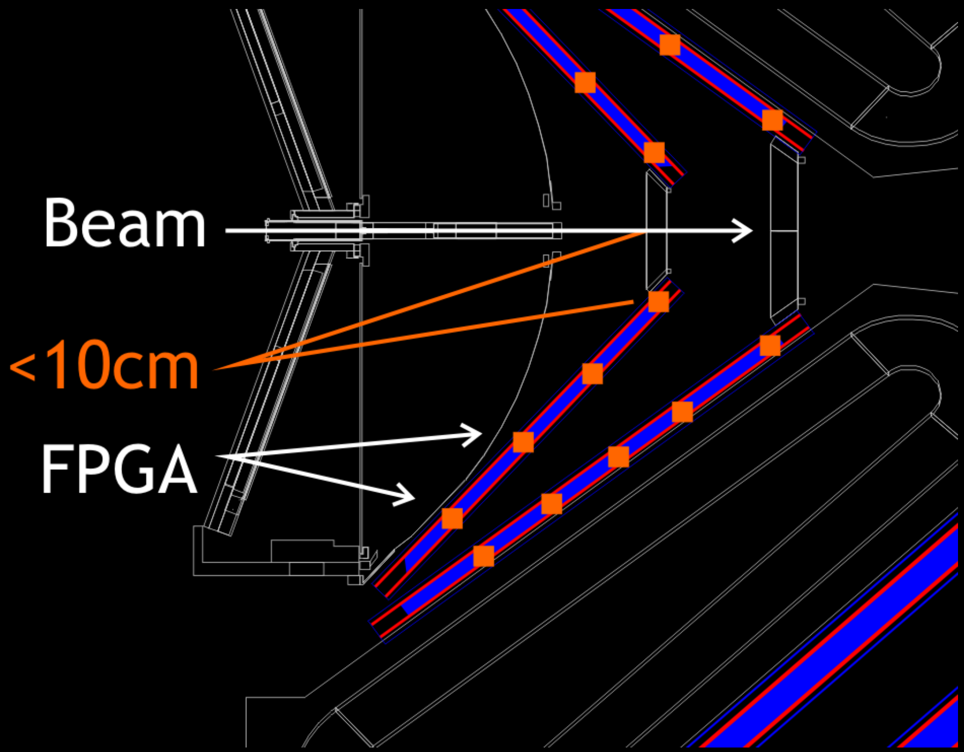


Radiation?



- Theory
  - $\sim$ 50 ions/cm^2/s
  - 3 SEU/h (full system)
- Experiment
  - 2 FPGA/h failing for unknown reasons
  - Topology of errors hints at SEU

Radiation?



Cross-section of HADES

- FPGA are sensitive to ionizing radiation (as any electronics)
- Temporary damage by changing value of storage cells
  - data corruption
  - change of configuration
- HADES: 3 SEU/h expected

- Triple redundancy
  - three instances of the same logic for cross-check of data
  - lowers error rate up to 100x
- detection of configuration changes
  - reload correct configuration

Interface / GUI Selection

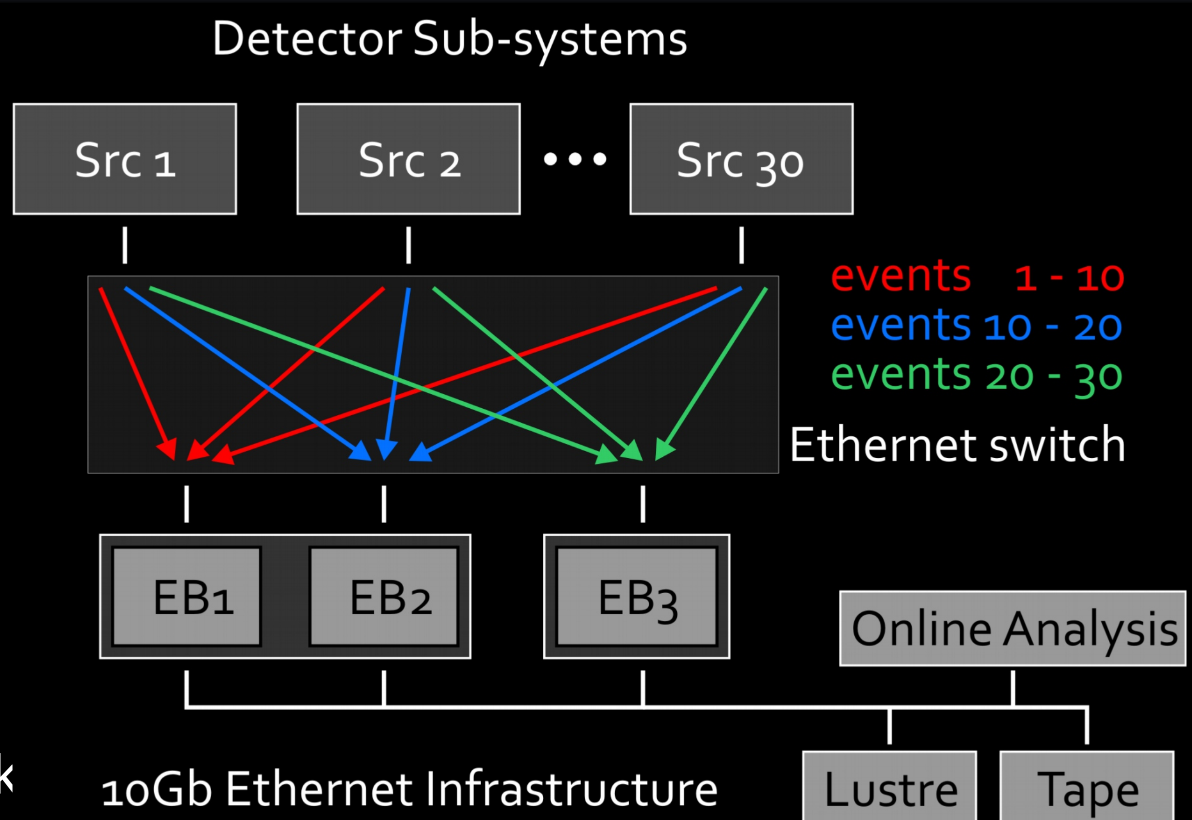
"commercial" Tools,
e.g. EPICS

- Experts are quite rare
- Modern GUI requires lot of ressources
- Many interfaces / modules readily available
- Special software required for each client

Self-made code
e.g. Perl & C-tools

- Very flexible
- Easy to extend by non-experts
- Needs some preparation time
- No I/O modules available
- Many tools ready from development phase

Eventbuilding

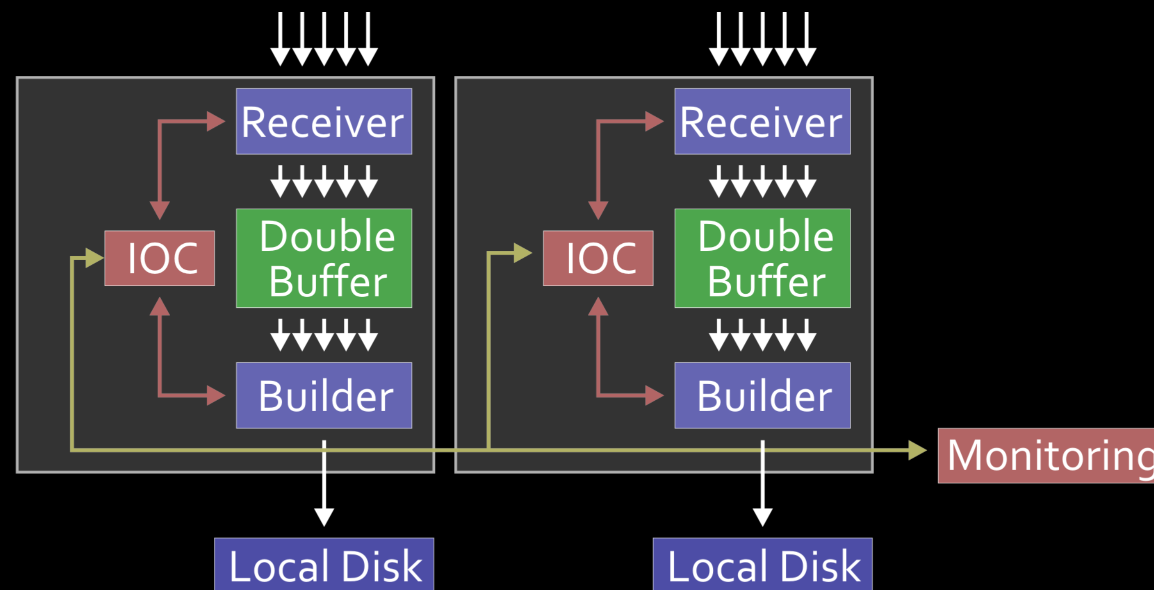


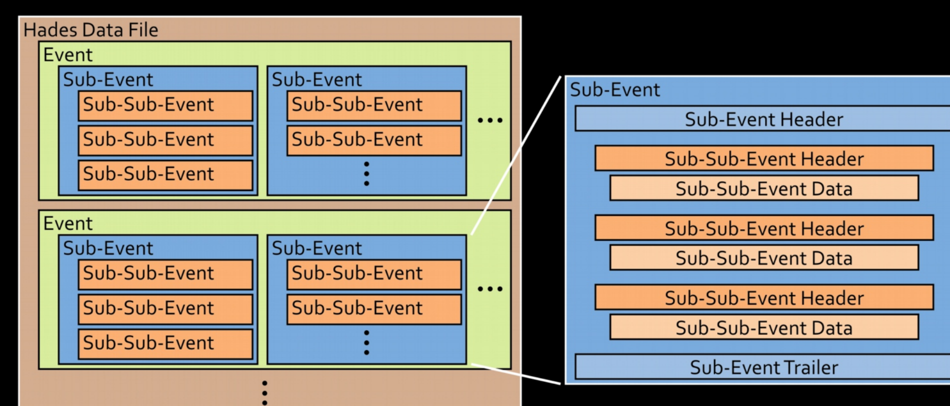
- 30x 1 GbE from detector
- 4 servers / 32 cores / 40 TB disk
  2x 10 GbE
  - Combine data to complete events
  - Temporary local storage
  - Transport to network file system (Lustre) and tape
  - Subset of data to Online Analysis

Eventbuilding

- UDP packets received from >30 sources
  - Stored in buffer memory
- Individual sub-events collected in one event
- Synchronization between 16 servers
  - Data automatically distributed by DAQ network
  - Inter-process communication with EPICS IOC

- Integration in DABC framework available

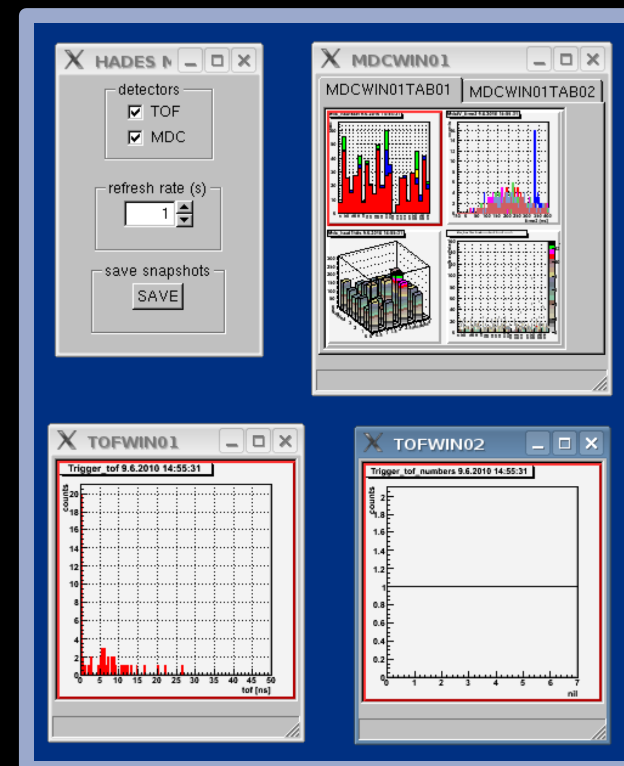




→ J. Adamczewski-Musch

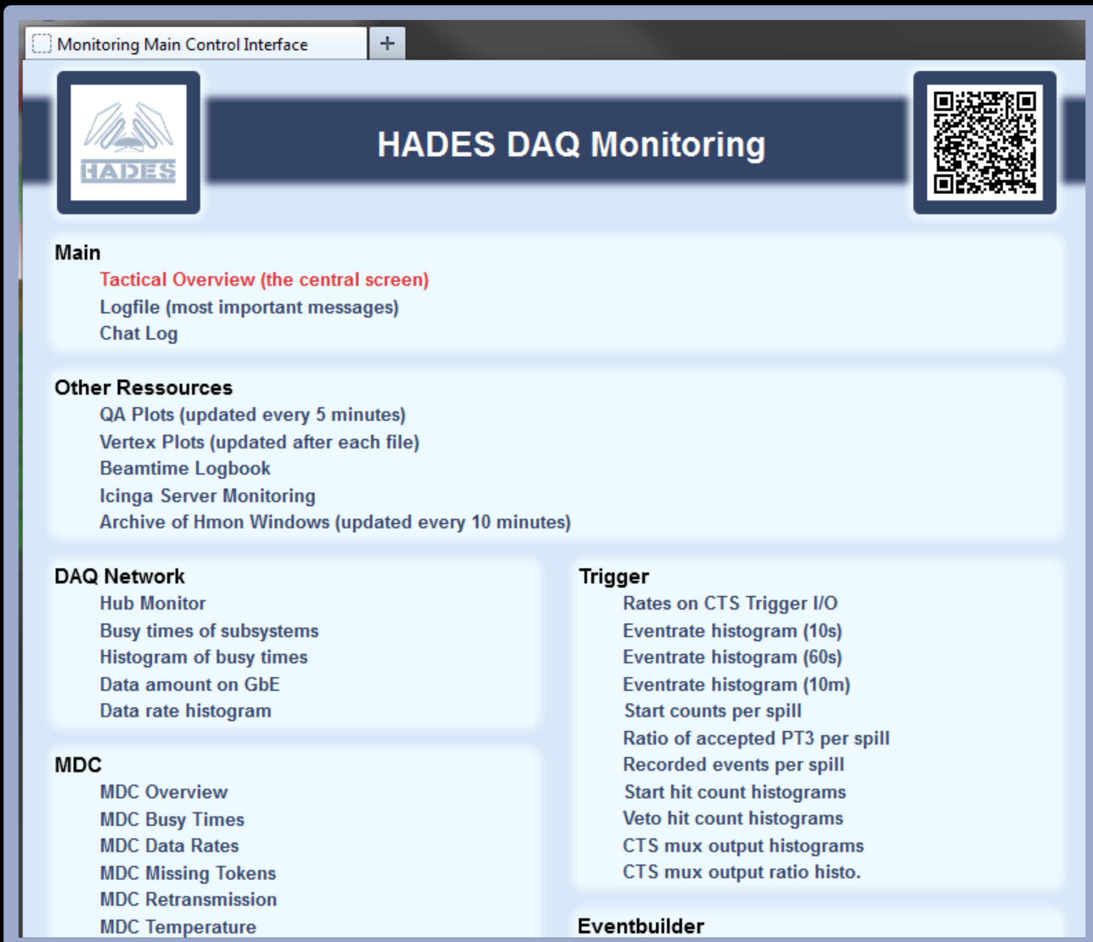
Online Analysis

- Idea: Use only basic Root classes, no external libraries
- Server
  - Reads & analyses HADES data files (hld)
  - Fills all defined histograms
  - Provides histograms over TCP
- Clients
  - Access histograms from server via TCP
  - Encapsulated in one executable
- Configuration
  - All configuration, histogram definitions defined in one XML file

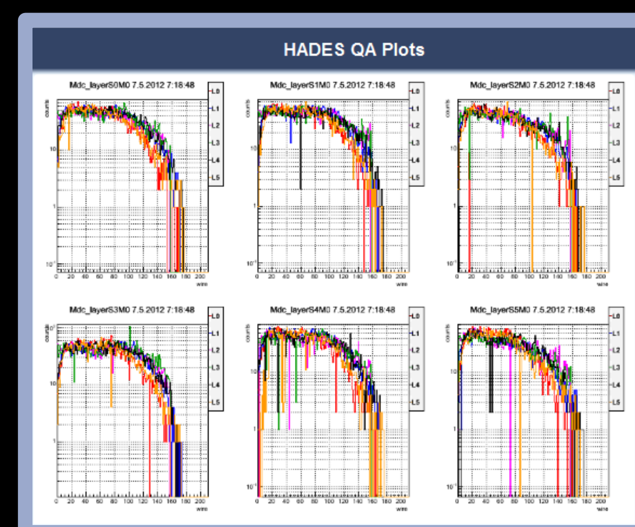


(very small demonstration set-up)
Typical: 10 windows / 10 tabs / 6 histograms ~ 600 histograms

DAQ Monitoring



- Central Webpage gives access to all monitoring features
  - Data QA Plots
  - DAQ Network
  - Operator Logbook
  - Server Status



Conclusions

- The upgraded Hades DAQ System was succesfully used during the 5-week April 2012 beam-time
- The DAQ / Eventbuilder / Monitoring performance was more than sufficient
- All error conditions could be succesfully identified and resolved
  - Most common errors within < 30 seconds
  - Very few calls to experts necessary
- The framework is flexible to be adapted to many other projects