XFEL Data Acquisition Workshop

3rd European XFEL Users' Meeting

28-29th January 2009

DESY, Hamburg

C. Youngman (for WP76)

DAQ workshop

XFEL DAQ and Control for photon beam systems workshop
10-11th March 2008

The workshop was a Pre-XFEL project partially funded by the European Commission under the 7th Framework programme.

- 47 registered participants:
  - ALBA, DESY, Daresbury, ESRF, ITEP, JINR, NIKHEF, PSI, RAL, SLAC, Spring8, Bologna, Heidelberg, Konstanz, Slovakia,
- Agenda and presentations available at:
  - https://indico.desy.de/conferenceDisplay.py?confId=762

Workshop aim and topics

- Aim
  - Review ALL areas of DAQ and control
    - List the requirements, and
    - Look for solutions
- Topics covered (20 talks with 550 slides)
  - Machine parameters and timing
  - Photon beam line instruments and detectors
  - Control systems
  - Archiving and data processing
  - DAQ and control at other Labs
  - Infrastructure requirements
  - Perspectives for data rejection and size reductions
  - 2D pixel detectors DAQ

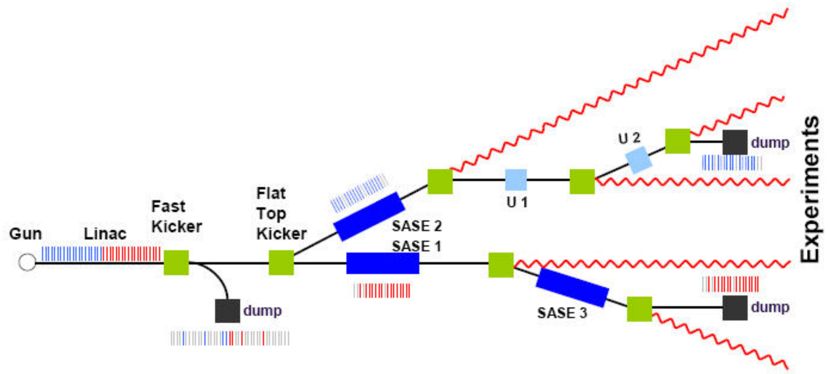
Scope of this talk

- This talk does not review the workshop!
  - Too much information for time available
  - Issues identified have been developed

- Talk about hot issues and what is being done
  - XFEL challenge to DAQ
  - DAQ architecture
  - 2D pixel detector backend systems
  - Data Management (DM) architecture
  - The trigger or data reduction problem

XFEL challenge to DAQ

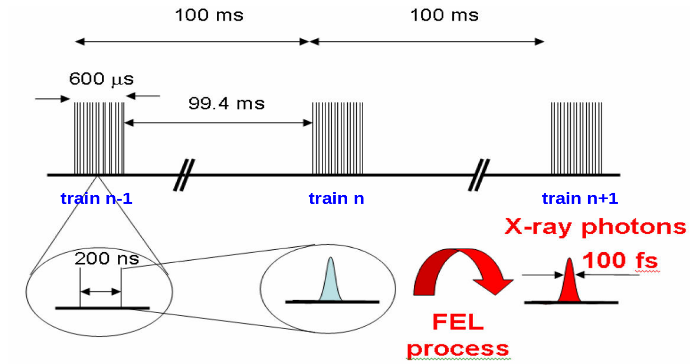
Photon beam lines



≤1500 bunches / train / beam line
(limited by dump, not 3000)

5 concurrent experiments

Beam time structure (nominal)



5MHz bunch rate

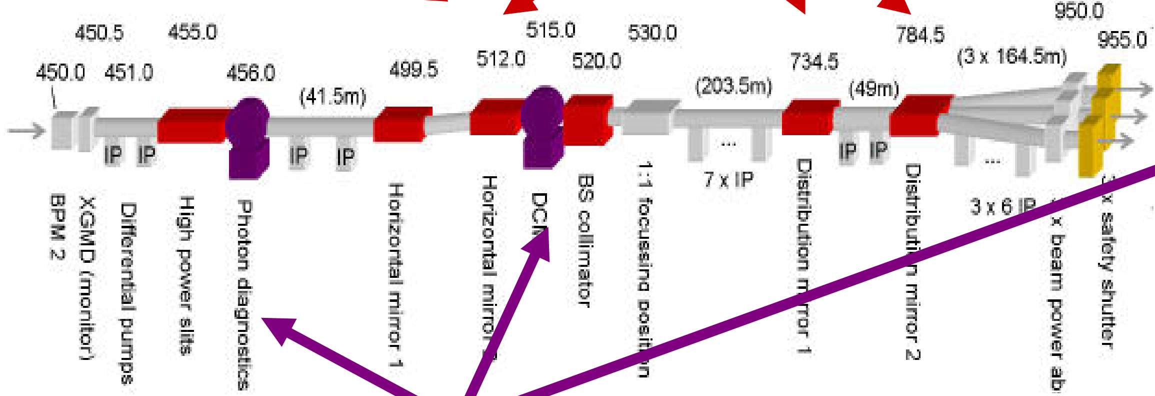
10Hz train rate

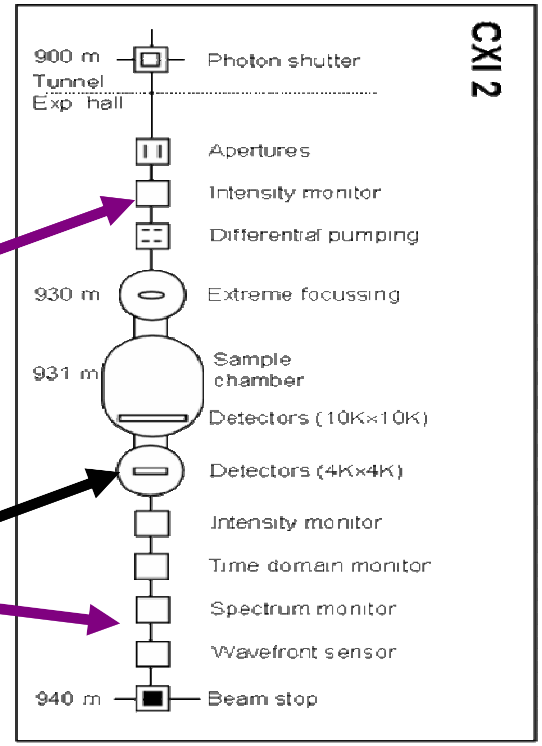
Data volume guidelines:	2MB frame = 1500 x 10 x 2 = 30GB / s 4kB frame = = 60MB / s

Data volumes are very challenging !! Front end bandwidth rates challenging

Today's XFEL DAQ component data sizes

Beam = mirrors, monochramators, shutters, … } ~few channels per component, e.g. motor control





Diagnostic detectors = pulse position, intensity,

~1kB(?) data per pulse, e.g. intensity monitor

Expt. Detectors = cameras, …

~2MB data per pulse per Mpixel of 2D detector

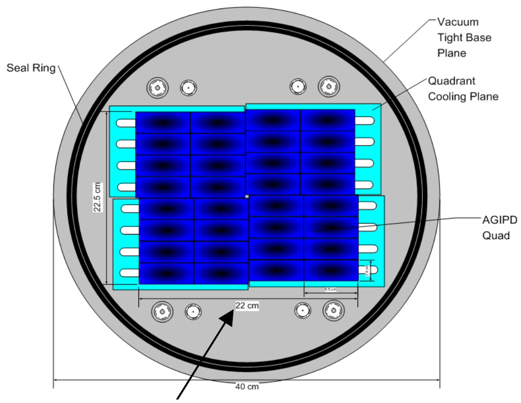
Data size per pulse = 2D Detector : Diagnostic : Beam = MB : kB : ~0 Bytes

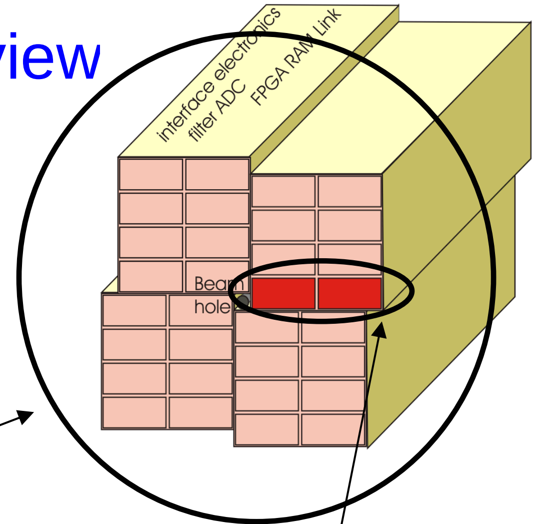
DAQ: satisfying 2D requirements will automatically satisfy others

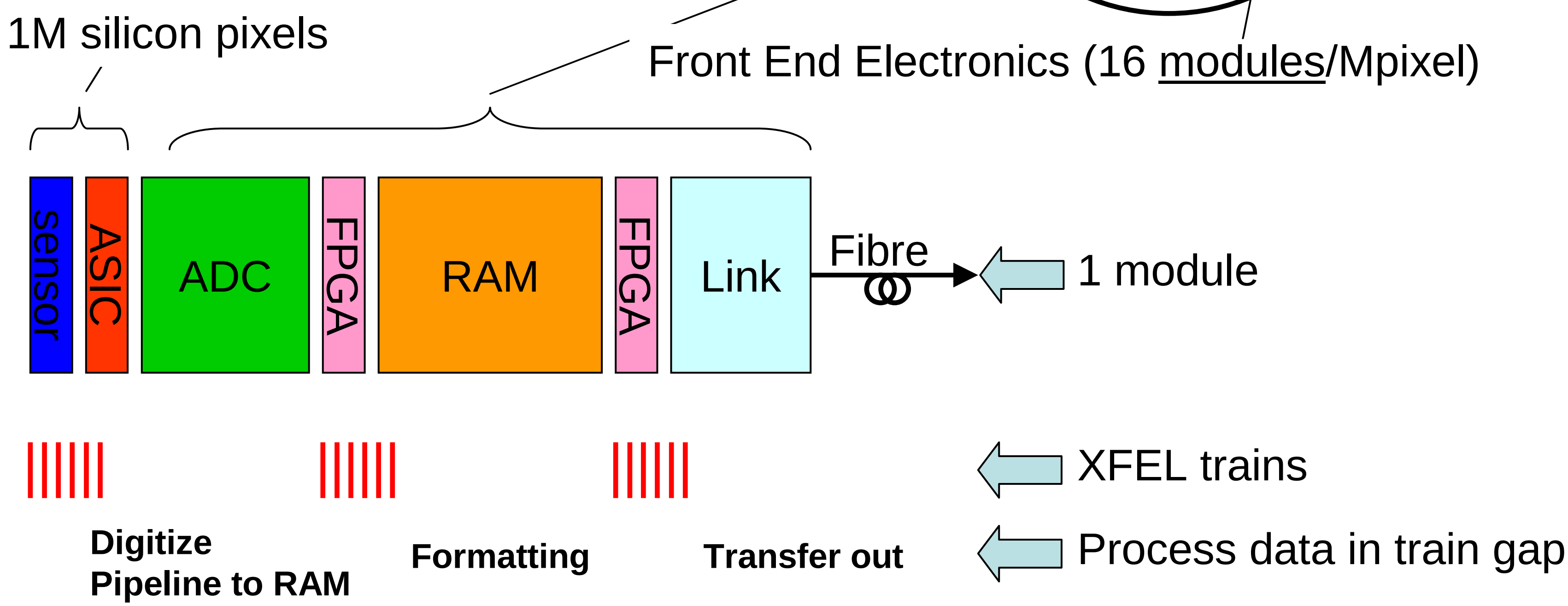
2D pixel detector DAQ developments

- Three 2D silicon pixel detectors being designed for XFEL
  - AGIPD (HPAD)– Adaptive Gain Integrating Pixel Detector
  - LPD – Large Pixel Detector
  - DSSC (DEPFET)– Depmos Sensor Single Compression
- Have similar designs
  - Baseline design for 2013 detector 1k x 1k (2MB/frame)
  - Typically 200µm pixel dimensions
  - Like the Pilatus camera design is modular
    - Allows building whole detector from smaller identical modules
    - Allows large detectors to be build 2k x 2k …
  - Use of capacitor storage pipelines to store pixel charge
    - Use inter-train gap to digitize and readout
- Differences: sensor design, ASIC and signal processing…

Example: AGIPD data processing overview

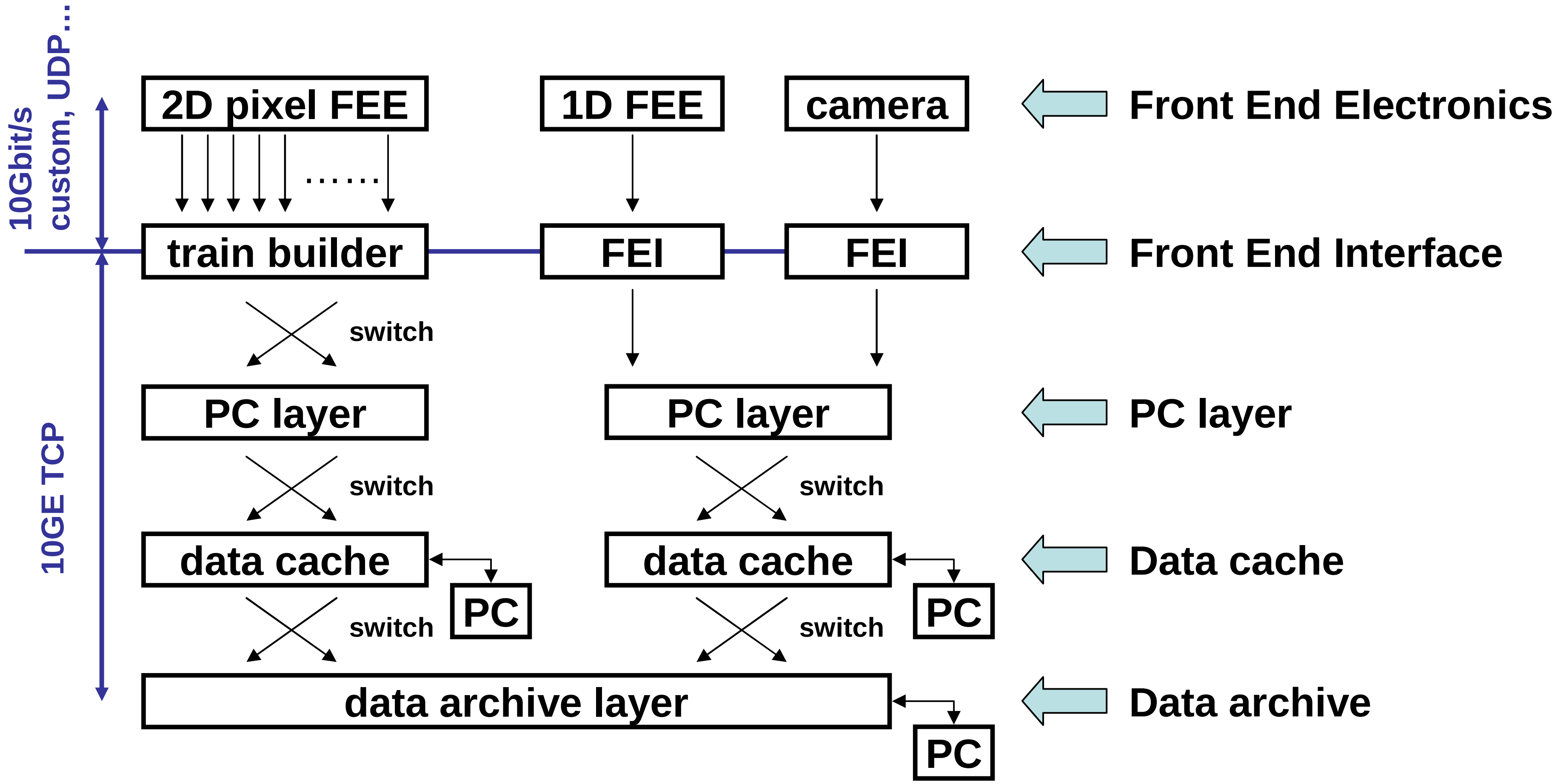






Challenge: 5MHz DAQ rate, otherwise might use commercial camera

XFEL readout architecture (todays version)



Multi layered design should be flexible and scaleable (within limits)
Scale by replicating devices in each layer as required.

XFEL DAQ readout architecture layers

- Five layer system and potential data processing functionality:
  - FEE = experiment specific (ASIC + ADC + data link)
    - data forwarding for non 2D
    - data formatting and forwarding for 2D
  - FEI = interface systems (data readout + control)
    - formatting, train building, pedestal correction, zero suppression, frame rejection tagging…
  - Train builder
    - formatting, frame building, train building, pedestal correction, zero suppression, pixel counting + region of interest finding & tagging…
  - PC layer = interface to data cache
    - monitoring for control purposes, tagged frame and train rejection, additional analysis and further rejection and/or compression, file size optimization…
  - Data cache = temporary storage (processing)
    - Additional processing and data rejection and/or compression
  - Data archive = disk & tape storage (analysis processing)

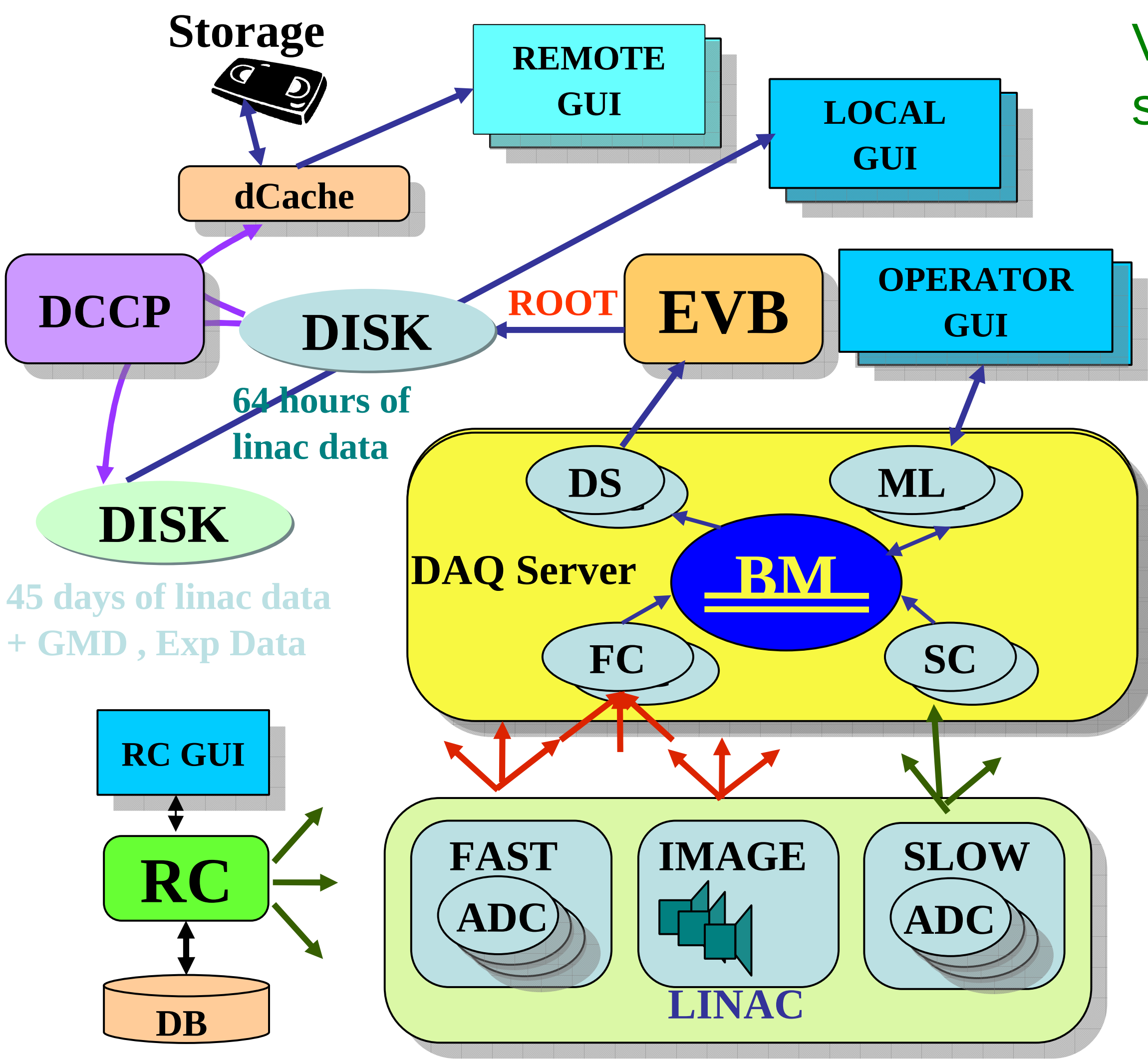
XFEL DAQ readout architecture concepts

- Design concepts
  - Avoid bandwidth bottlenecks (caveat: volume cutoffs, see later)
  - Processing foreseen in all layers
    - Driven by unknown rejection and compression requirements
  - Use 10 Gbit/s link connection as standard (today)
  - Basic unit of data transferred is a train (= train building at FEI layer)
  - Plan for full pixel payload (user compression input needed)

- 2D detectors agreed to use common implementations of
  - FEI layer = Train Builder (TB)
    - Proposed in-kind UK contribution: STFC (J.Coughlan et al)
  - detector control and timing interface (C&C) systems
    - Proposed in-kind UK contribution: UCL (M.Wing et al)
  - 10GE development
    - Led by DESY-FEA (in collaboration with Mannheim+STFC)

Processing = rejection (of frames or trains) + compression (vol. reduction) is an important issue. Its understanding is at an early stage.

Comparison: FLASH readout architecture



Very different approach with single “all data” big buffer.

multicast

Fast data (every micropulse)

Beam relevant info:

ADCs (BPM, BLM, TOR, etc)

CAMERAs

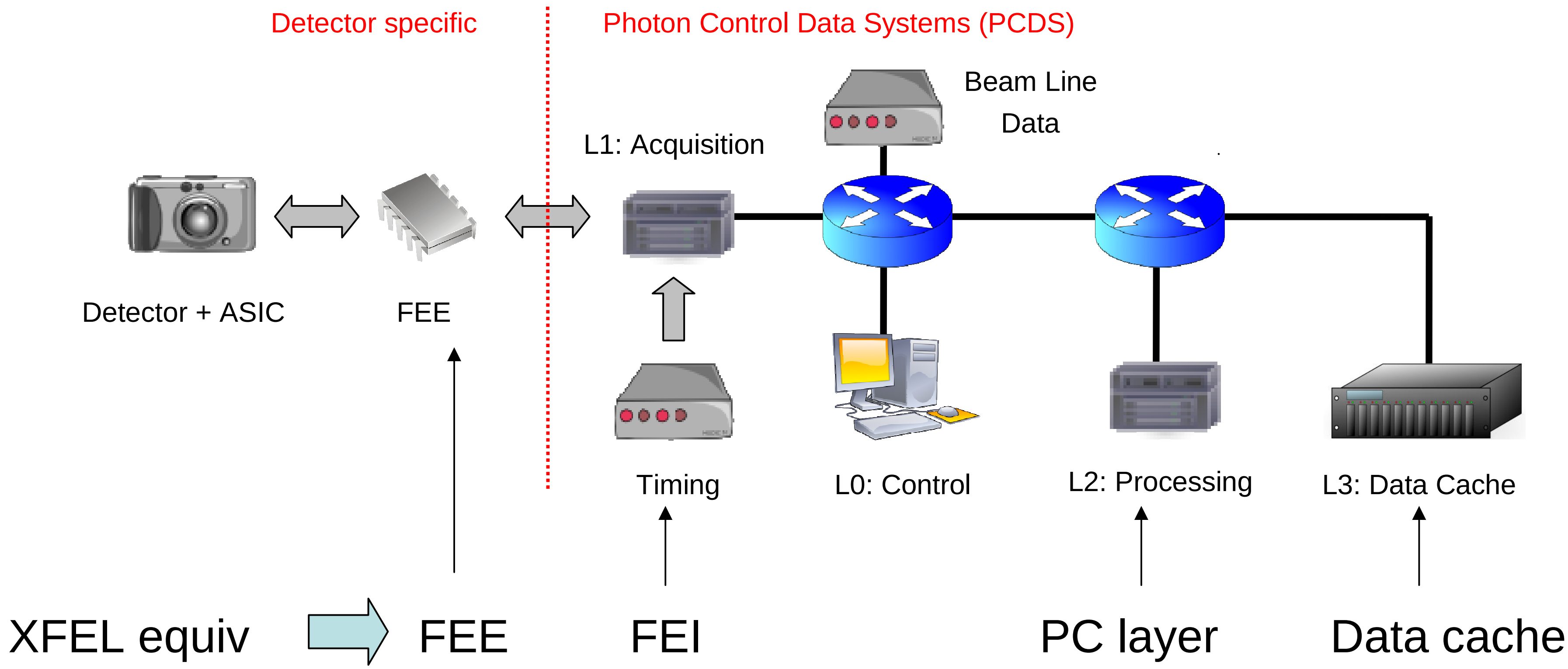
DOOCS (TINE)

Slow data (max 1Hz)

Data from slow ADCs (MAG, V, etc.)

DOOCS channels (Masks, params, etc.)

Comparison: LCLS data system architecture

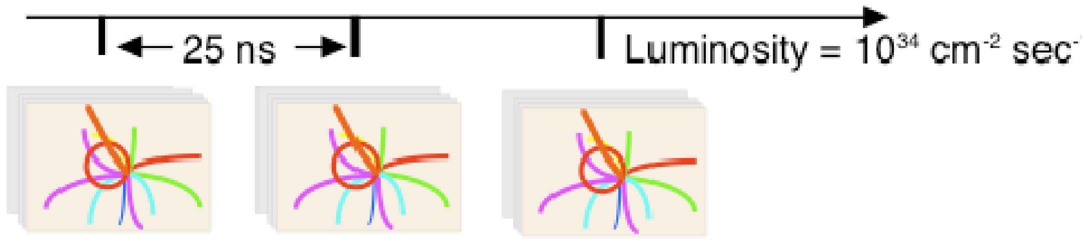


Similar approach at least functionally, hardware implementation differs.

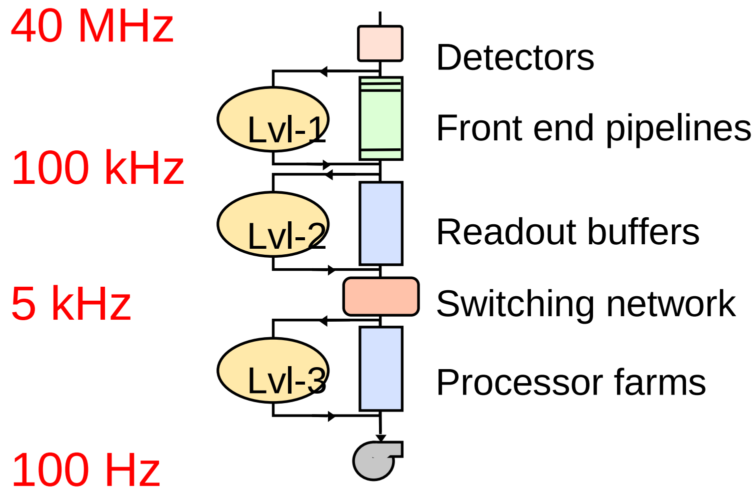
Comparison: HEP DAQ readout architectures

LHC = ≈ 30 Collisions/25ns
(10^9 event/sec)

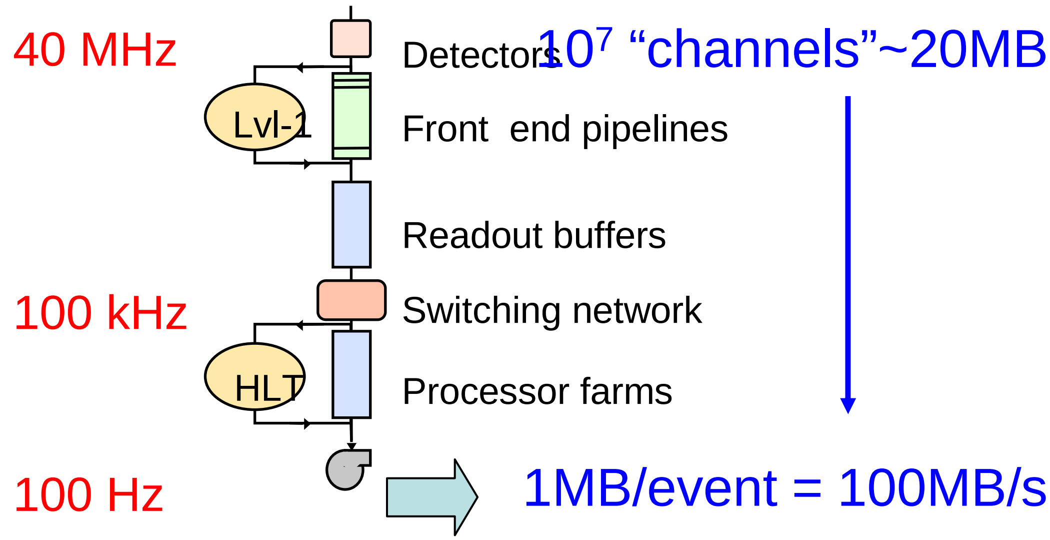
10^7 channels
(10^{16} bit/sec)



ATLAS: 3 trigger levels

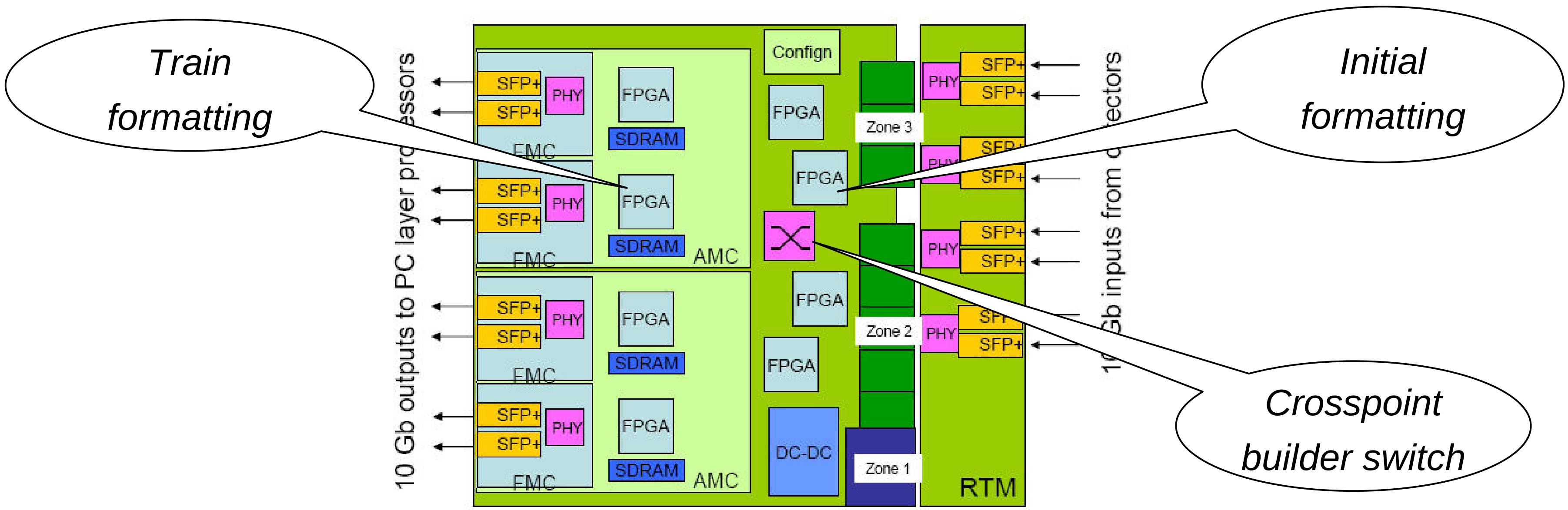


CMS: 2 trigger levels



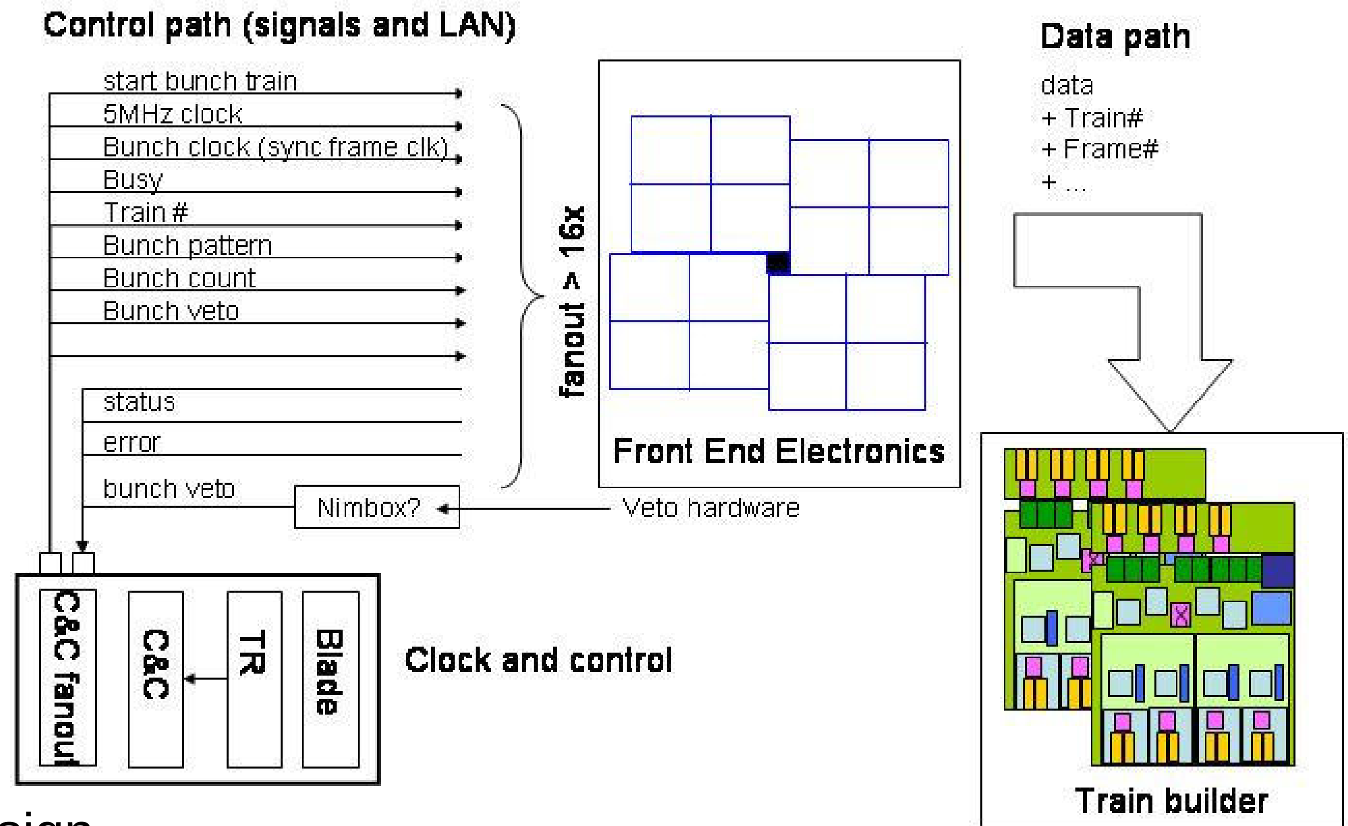
HEP experiments employ successive "trigger" levels to reduce the DAQ rate. The works because "relatively" simply trigger primitives (high transverse momentum leptons, missing energy balance, …) can be cut on. And zero suppression, etc. to reduce size.

Train builder = 2D pixel detector FEI



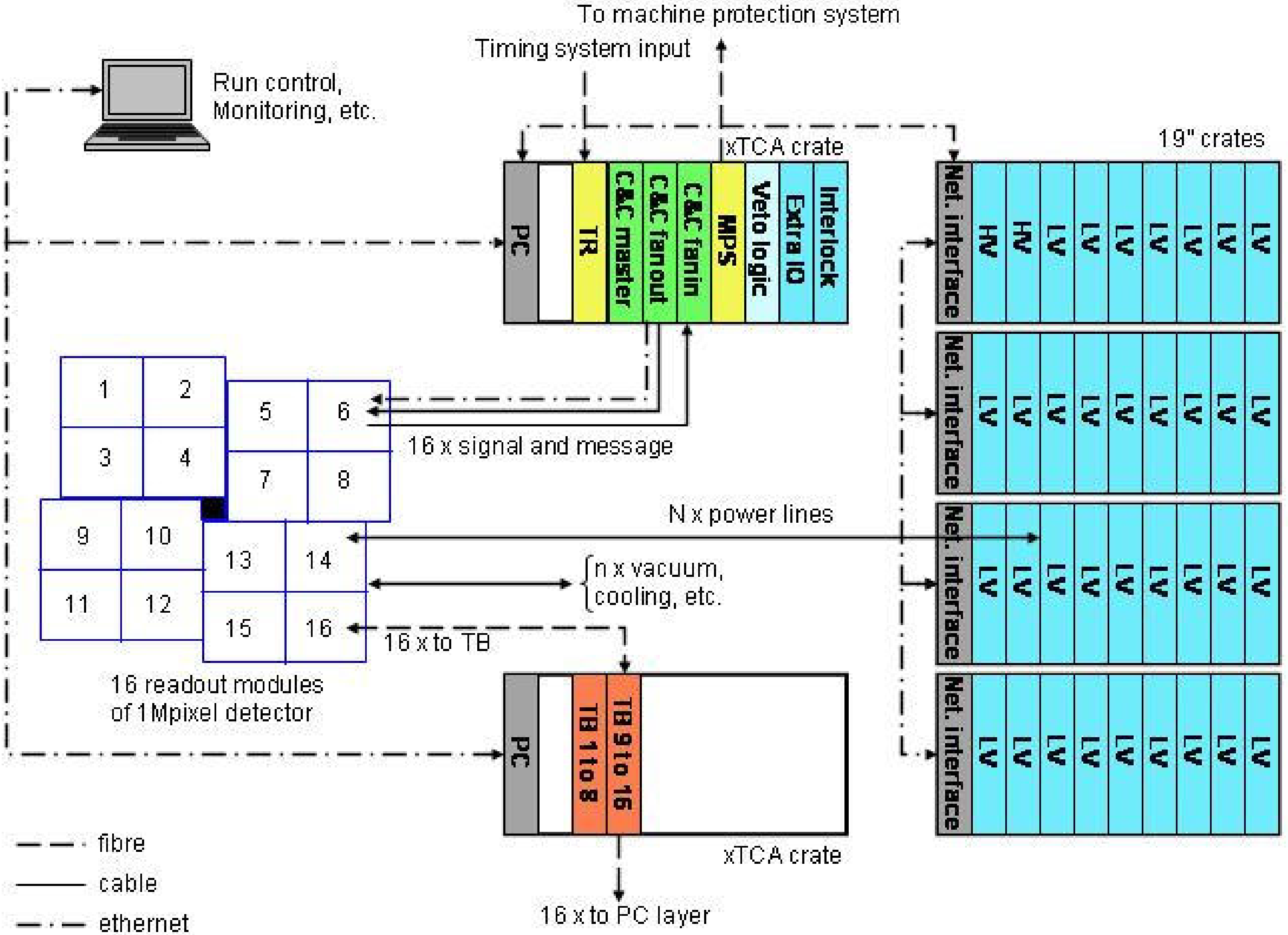
- Train builder
  - Rebuilds and orders frames from modules into single contiguous train unit
  - Protocol conversion = custom hardware-hardware input, TCP output
  - Processing = maybe pedestal correction, zero suppression, trigger sum counting, etc.
  - Implementation = currently double AMC board and FMC mezzanines
  - Currently 8 channels = ½ Mpixel
  - Final design will develop with understanding and technology improvements (e.g. CPUs mezzanines for more processing, 16 channels…)

C&C (Clock and Control) for 2D pixel detectors



- Design
  - Implementation = AMCs boards
  - Master (C&C) interfaces to XFEL timing board (TR)
  - Control distributed to detector FEE (and back) via signal cable and LAN
  - No signal connections to TB, just LAN

Example: Possible DAQ implementation for AGIPD



caveat: DAQ volume cutoffs

- 2D detector data volume limitations:
  - FEE ASIC pipeline lengths: now 100-200 and 2013 ≤ 512
  - TB design baseline 512 x 2MB frames/train (10GB/s)
  - Link speed 10Gbit/s is ~55% saturated by 512 x 2MB frames/train
- Therefore 512 frames/train is max. at 1Mpxl 2D payload now
  - For > 1Mpixel 2D detectors use additional parallel slices
- Keep an eye open for improvements
  - 2010 TB prototyping phase appraisal breakpoint
  - Improved technology (faster links, memory access…)
  - Data reduction and compression feedback
    - 512 x 2MB = 1024 x 1MB = 2048 x ½ MB…
  - New technology solutions

Worst case analysis data volume scenarios

Assume for machine:

220 days of scheduled running per year
70% XFEL beam delivery efficiency during the 220 days.

Assume for the experiments:

3 (of 5) beam lines operating with 2D pixel detectors,
70% detector and DAQ efficiency
90% target efficiency
Frame size per 2D detector 2MB

Three 2D detector running:

Instantaneous bandwidth = 30GB/s
Data volume at full bandwidth at 30GB/s = 105TB/hour continuous working
Data volume per day = 105TB x 24 x 70% x 90% = **1590TB / day**

Three 2D average daily rates to archive:

Average bandwidth = 30GB/s x 70% x 90% = 19GB/s

Do not take to seriously, yet
Data rejection and
Compression ignored

XFEL: 3 detector data volume = 220days x 1590daily x 70%xfel = **245 PB / year**
LHC: 4 expt. data volume = 220days x 0.3daily x 100%lhc = **1PB / year**

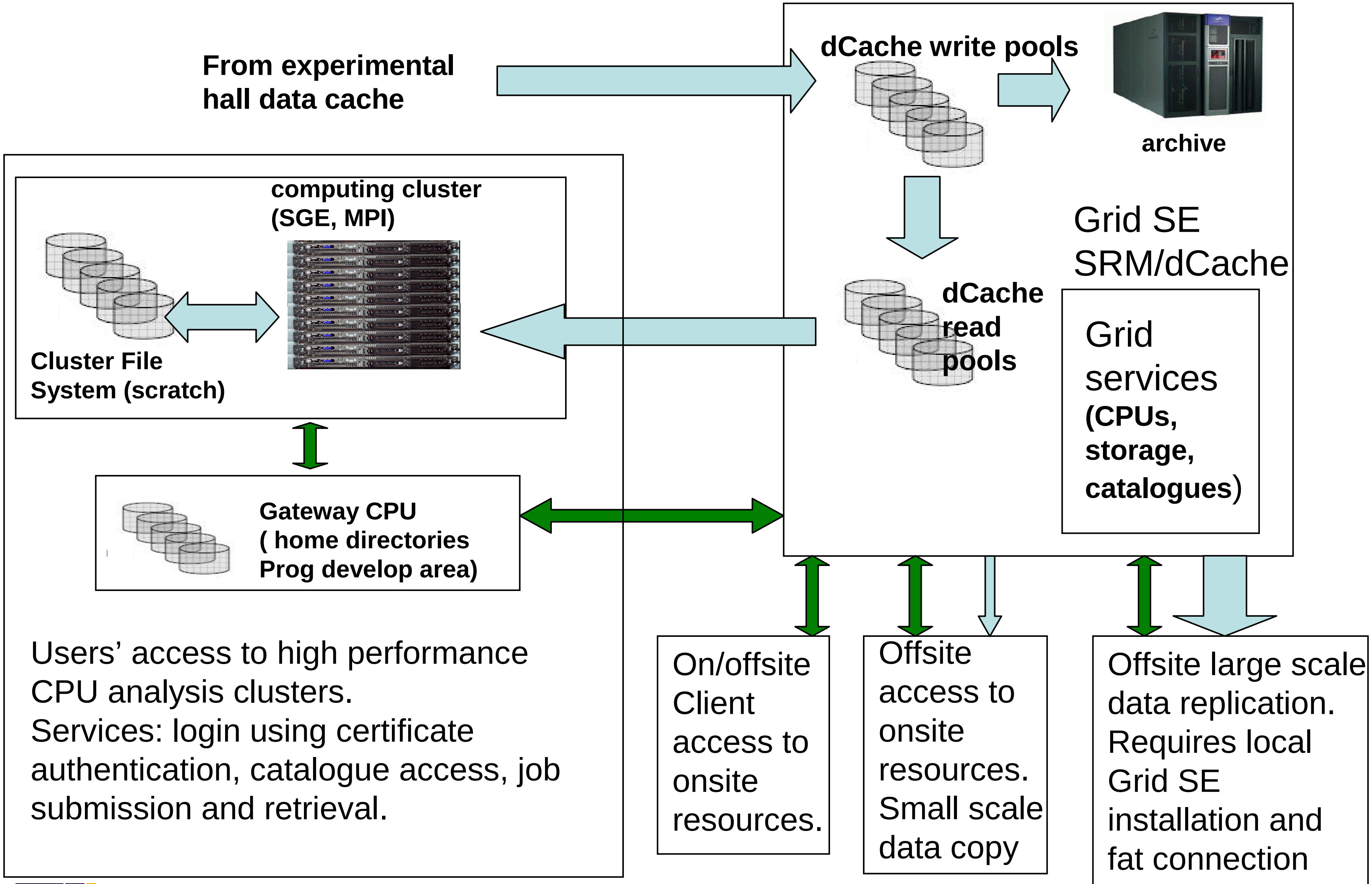
Serious problem = how to store and analyze this amount of data?
Reduce size: reject bad quality data, compression, quick analysis and delete…
Increase size: bigger detectors, more detectors, more frames / train…
245 PB/year storage technologically possible (analysis?) costs money.

Incomplete data management requirements zoo

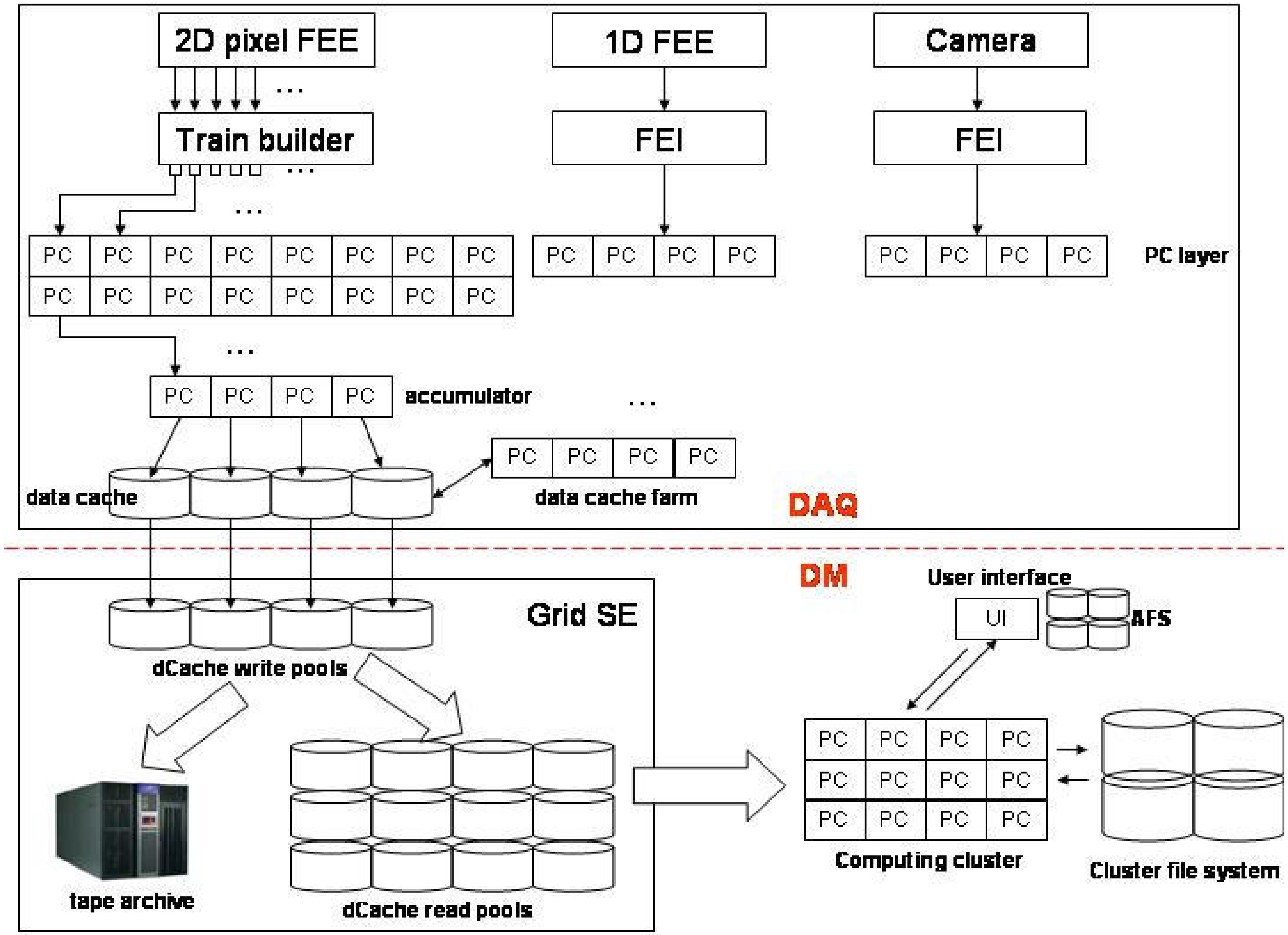
- DAQ to DM connection
  - avoid bandwidth limits
- Centrally organized storage services
  - tapes for secure backup of data
  - disk storage for onsite analysis farms
- Centrally organized computing (CPU) services
  - base load computing service
  - higher performance computing clusters
- User analysis models
  - Onsite users (storage and CPUs)
  - Offsite users (incl. small volume data copying to local CPUs)
  - Offsite for large data volume users
- User authentication (who), authorization (role allowed), data access (copy, delete) and job submission software
  - Looking into using GRID based tools, standard and light (single user no GRID infrastructure).
  - Web service clients, etc.

Ideas not knew: HEP derived, similar developments at ESRF, Diamond…

A candidate DM architecture

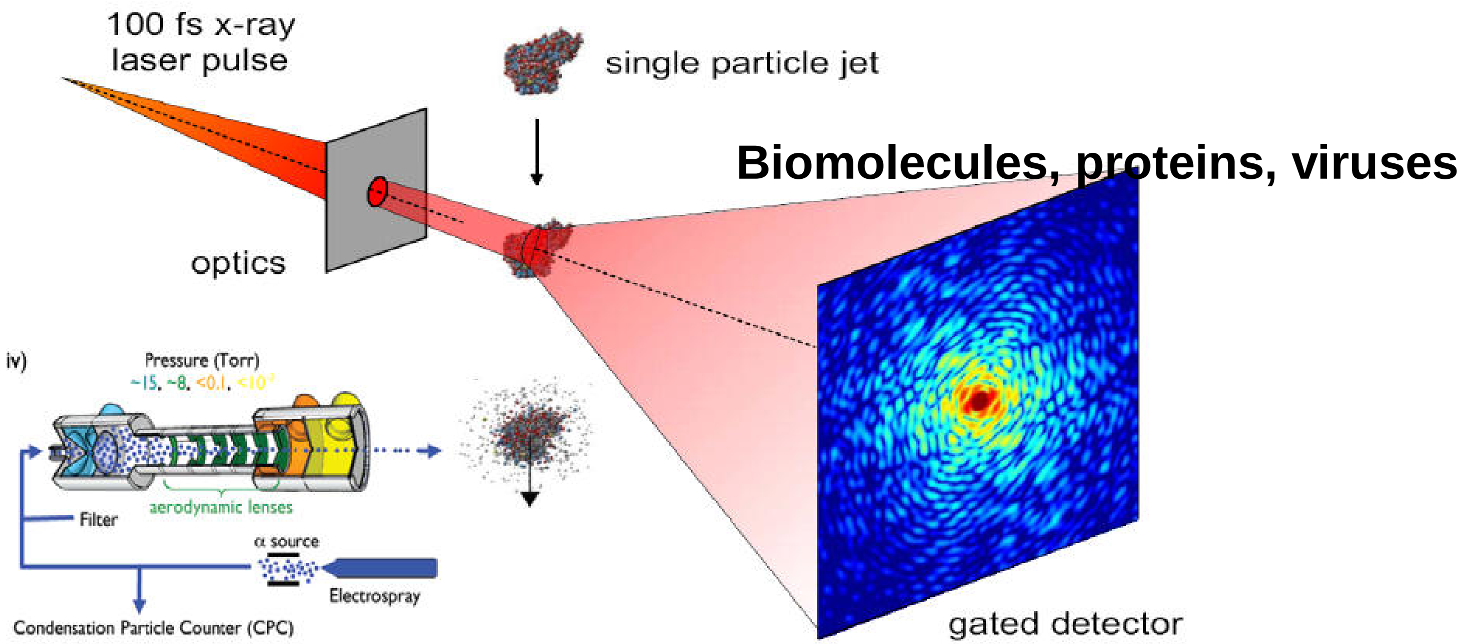


DAQ and DM architecture



Putting DAQ and DM architectures together

The trigger or data reduction and compression problem



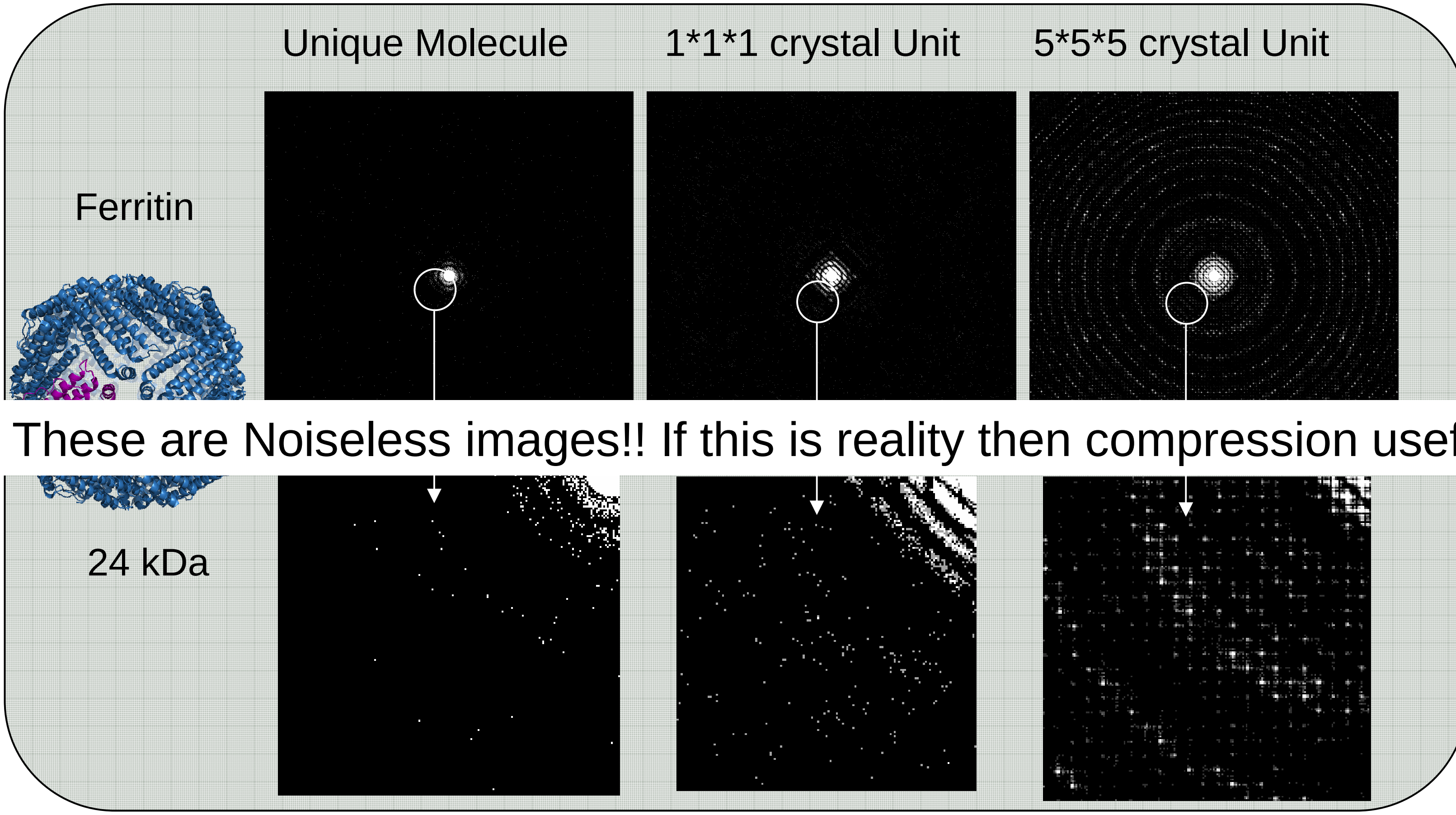
M. Bogan, NanoLetters 8, 310 (2008)

At light sources there are usually no simple primitives = no HEP reject trigger.

PS's traditionally readout all the data and store it – this works at low rate machines (LCLS and SCCS) by 2013 this has to change for 2D DAQ otherwise system will be bandwidth and data volume limited

Data rejection and compresion outlook

- At workshop first results of 2D simulations shown



These are Noiseless images!! If this is reality then compression useful

Source: G.Potdevin

Images courtesy of Pr. Franz Pfeiffer, SLS & Lausanne University

- Worry that noise (detector, beam line optics and vacuum impurities) would be relatively large = no compression

Data rejection and reduction outlook

- Improved simulations suggest that noise is less
  - Detector ~few pixels with noise photons / M pixel
  - vacuum impurities and optics produce lower noise

- Use input from simulations and LCLS detectors
  - Evaluate priority of implementing compression
  - Also use input for rejection algorithms

Rejection and compression algorithms will be very important.

Summary of DAQ and DM challenges

- The data volume problem
  - bigger 2D detectors, more bunches acquired, more beam lines…
  - how long the data is kept for analysis
  - requires rejecting poor quality data as early as possible
  - requires compressing data as early as possible
- User experiment profiles needed
  - Pixel hit per frame
  - Storage size per experiment
  - CPU analysis per experiment
- User access to and analysis of the data will change
  - e.g. GRID type tools will be needed
  - users will analyze their data on and offsite
- Today's design solution is
  - Scalable: but there are limits !
  - Flexible: swap out or insert entire layers if better solution appears
  - but, need to keep an eye open for new ideas/technologies

Items from workshop not discussed during talk

- Photon beam line DAQ and control
- Photon diagnostic DAQ and control
- Control systems software
- Infrastructure requirement

Acknowledgements

- Thanks to the speakers and participants at the workshop
- Thanks to those who contributed to this talk
- Thanks to the other members of WP76:
  - S.Esenov – Software development
  - K.Wrona – Data Management
- Thanks for listening

Spares

Grid services useful for distributed users

The Grid provides a set of generic services

- **Information system**
  - **B**erkley **D**atabase **I**nformation **I**ndex
- **World wide authentication and authorization**
  - **V**irtual **O**rganization **M**anagement **S**ystem
- **File catalogue**
  - **L**CG **F**ile **C**atalogue
- **Storage management**
  - **S**torage **R**esource **M**anager
- **Wide area network file transfer**
  - **F**ile **T**ransfer **S**ervice, lcg-util
- **Site local data access**
  - **G**rid **F**ile **A**ccess **L**ibrary, lcg-util
- **Job management**
  - **W**orkload **M**anagement **S**ystem

Computing TDR covering DAQ and DM at XFEL is in preparation.