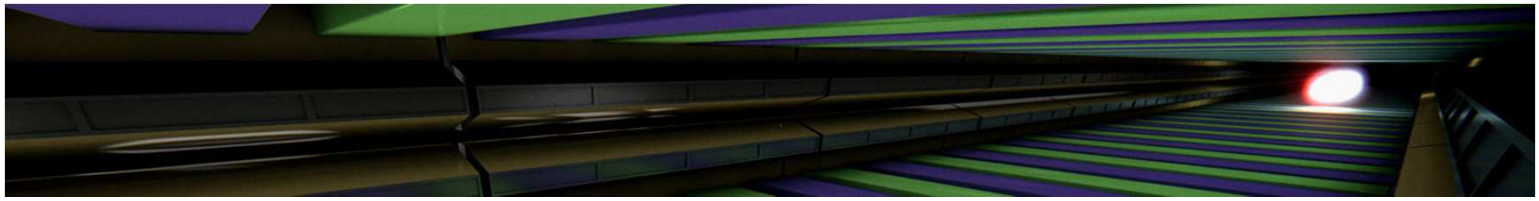


Data acquisition and Controls

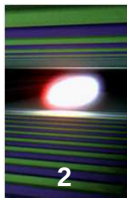
XFEL Users' Meeting

25.Jan.2012

C.Youngman for WP76

European XFEL

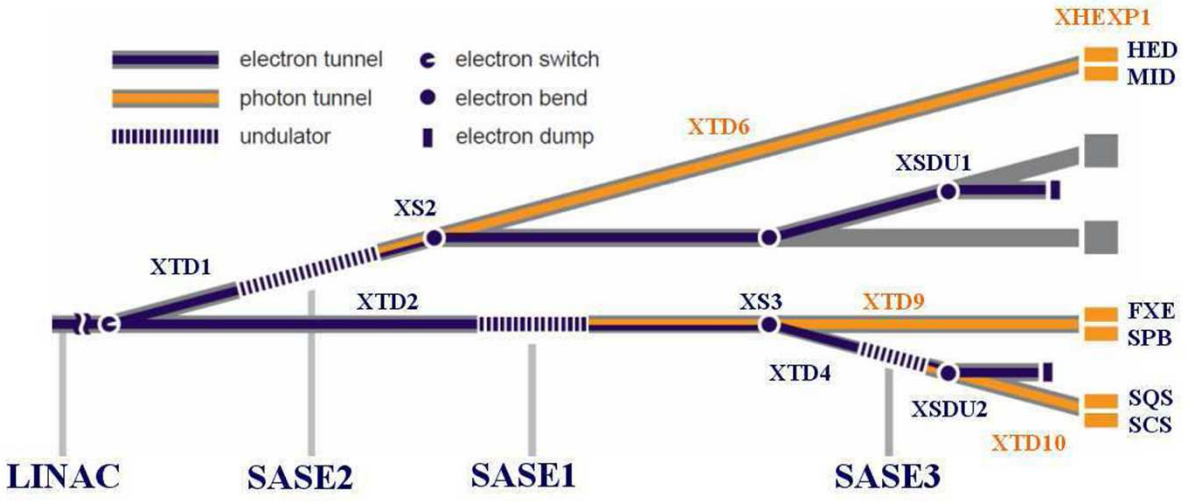
Aim of talk



- This talk describes
  - DAQ and control developments
  - Control of beam line systems
  - Data management
  - Slice test developments
- The following talk from Burkhard Heisen concentrates on software developments for control and scientific computing
- Both talks are concerned with photon beam line systems, they do not address electron machine DAQ and control

Beamline layout

- XFEL first stage consists of 3 SASE beamlines

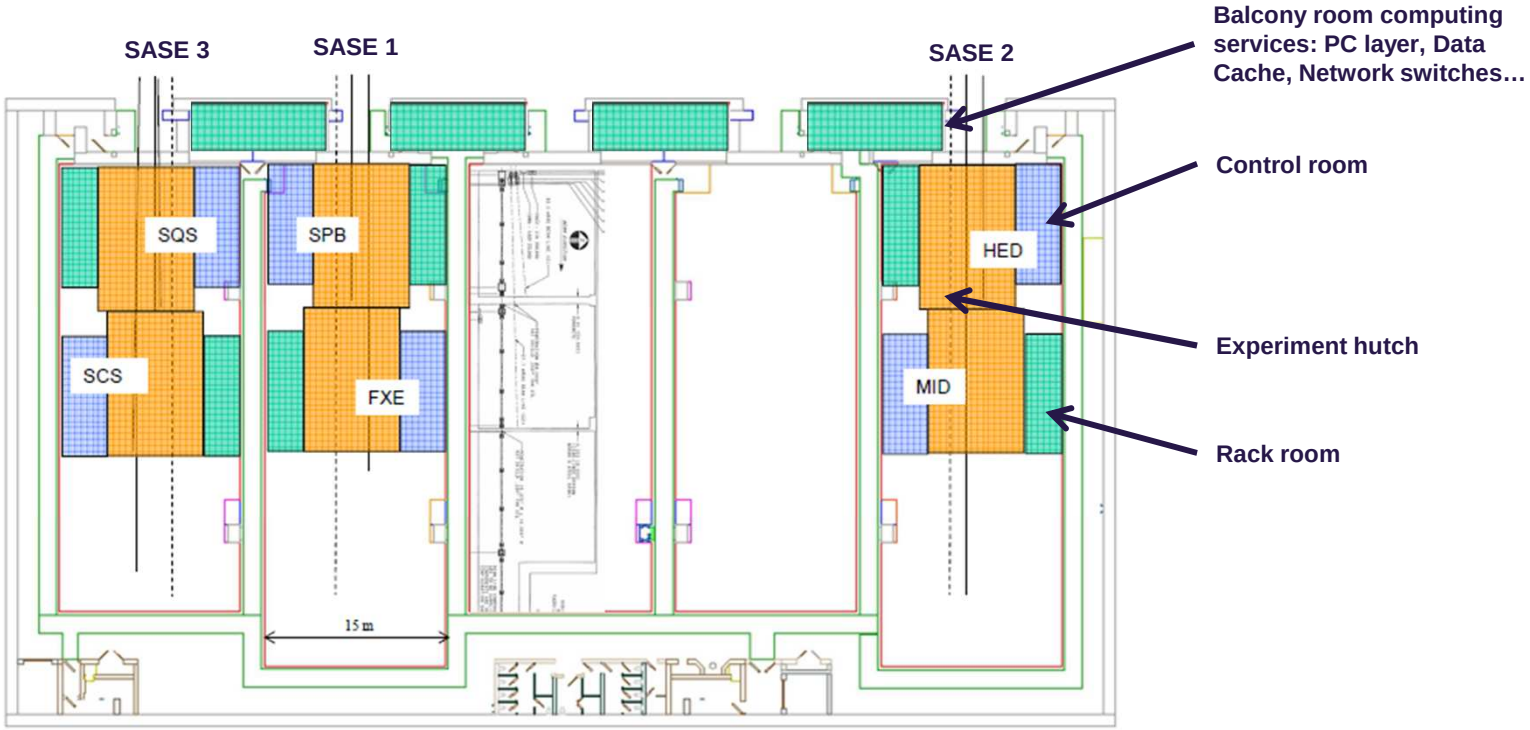


Beamline	Start installation	First operation	experiments
SASE1	2014	2015	HED, MID
SASE2	2014	2015	FXE, SPB
SASE3	2014	2015	SQS, SCS

- DAQ and control required for diagnostic and optics beamline instruments

Experiment layout

- XFEL first stage consists of 6 named experiments in XHEXP1



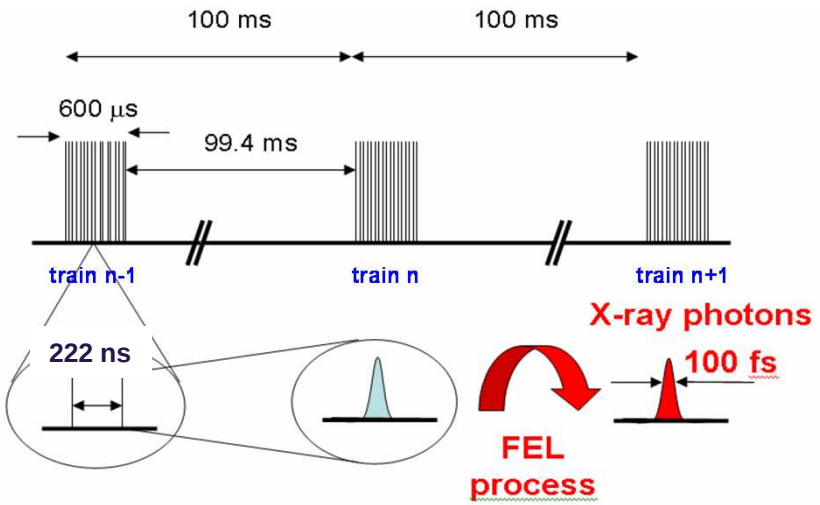
- One experiment taking data, second experiment preparing in beamline
- SPB (+SFX) experiment approximate control size:
  - 3 x 2D-detectors, 2 sample injectors, 1 spectrometer, 2 filter banks, 4 slits, 2 KB lenses, 1 opt. laser, 5 BPMs, 5 screens, 6 chambers, etc.

List of instruments to control

- Optics (WP73)
  - KB mirrors for focusing
  - Refractive lens focusing
  - Monochromatic
  - Collimator
  - Slits
  - Attenuators
  - …
- Sample environment (WP79)
  - Particle injector
  - Cryostat
  - Precision stages
  - …
- Beam diagnostics (WP74)
  - Intensity monitors
  - Beam positioning monitor
  - Photon-electron spectrometers
  - K-monochromator
  - Screens and cameras
  - …
- Measurement instruments (WP8x)
  - e- and ion TOF
  - Spectrometers
  - …
- Laser systems (WP78)
  - Pump laser and diagnostics
  - …
- Vacuum systems (WP73)
  - Turbo pumps
  - Ion pumps
  - …
- 2D detectors (WP75)
  - AGIPD
  - LPD
  - DSSC
  - pnCCD
  - …

Beam time structure drive DAQ

- **Readout rate driven by bunch structure**
  - **10 Hz train of pulses**
  - **4.51 MHz pulses in train**
- **Data volume driven by detector type**



Detector type	Sampling	Data/pulse	Data/train	XFEL/sec	LCLS/sec
1 Mpxl 2D camera	4.5 MHz	~2 MB	~1 GB	~10 GB	~300 MB
1 channel digitizer	5 GS/s	~2 kB	~6 MB	~60 MB	~0.2 MB

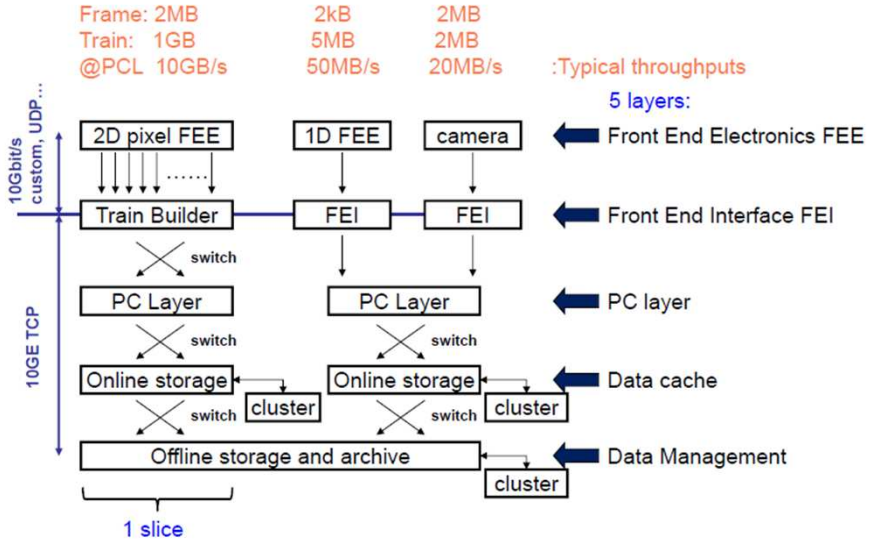
- **Largest data volume are produced by 2D area cameras**
  - ➔ **Solving DAQ for 2D cameras should provide solution for all instrument types**
- **Data volumes are considerably larger than at LCLS**

DAQ and control architecture and standards

DAQ and control Architecture

- Multiple layers with well defined APIs
- Multiple slices for partitioning and scaling
- Allow full speed write through to online storage, but discourage usage
- Enforce data reduction and rejection in all layers

Architecture based on findings of 2009 computing TDR



DAQ and control standards

- Data transfer links are 10 Gbps with SFP+ or compatible plug
- TCP protocol is used for data transfer downstream of FEI
- Control interface to FEI is via 1 or 10Gbps TCP
- Time synchronization is performed using the XFEL timing system (DESY-MCS4)
- Data blocks transferred are complete trains of pulse ordered frames
- Data is tagged with the unique train number (and pulse number)
- Control and data s/w interfaces must be complied to

A new instrument to be integrated must obey the above

Standardizing crates and modules - Overview

- To handle high data volume and rate operations DAQ standardizes on Telecommunications (xTCA) crate and module standards
  - Point-to-point backplane switched connections allow high speed interconnections between FPGAs, host CPUs, etc.

This standardization has been driven by the XFEL e-machine control group MCS4

- The MTCA.4 standard defines extensions required for Physics applications:
  - Timing interface on Backplane
  - High speed module interconnections
  - Double size modules compared to micro-TCA
  - Allows custom Rear Transition Modules (RTMs) use with multi-purpose digital board
- DAQ uses MTCA.4 extensively: 2D Clock and Control, VETO, APD DAQ …

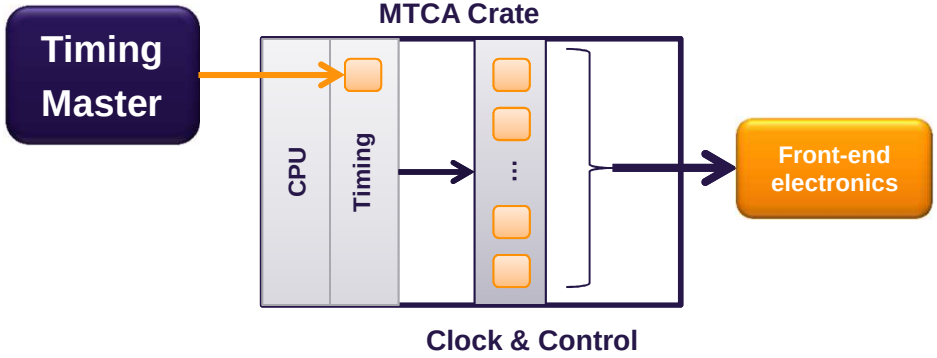
UCL Clock & Control RTM for 2D detectors

DESY DAMC2 (front)

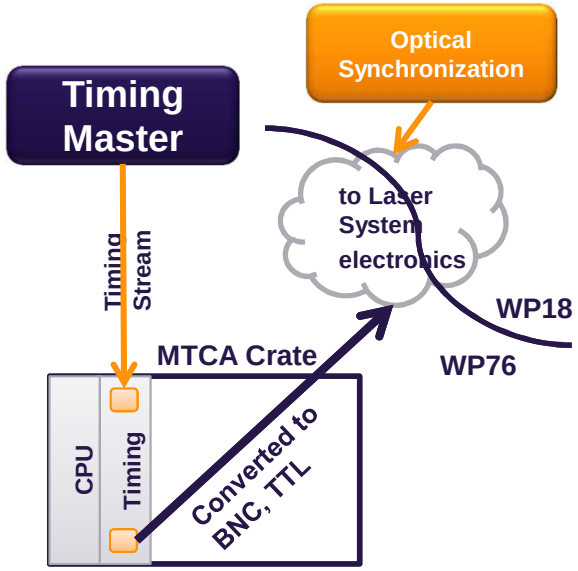
Lab MTCA.4 Schroff crate

Timing and Synchronization Systems for XFEL

	Timing System (WP28)	Optical Synchronization (WP18)
Provides	Clocks, Triggers and Data	Clocks
Stability	Less than 10ps	Less than 50fs
Applications	DAQ and Detector sequencing	Synchronize lasers to beam



- Timing master distributes reference clock with encoded data
- Drift is actively compensated
- Transmits events used fur triggers (START) and bunch clocks
- Transmits bunch pattern and related information

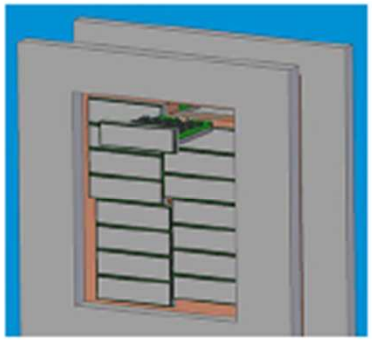


- Optical synchronization is required to be in phase with beam
- Timing is required to select the correct pulses

European XFEL

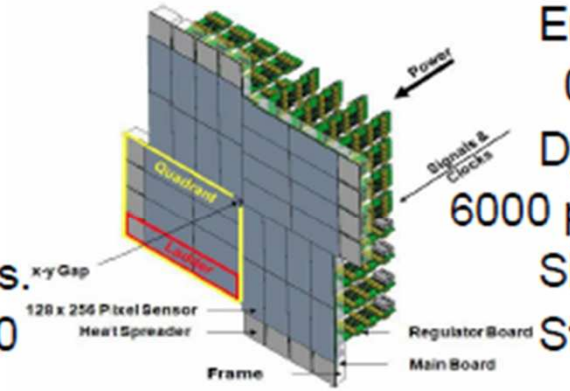
XFEL 2D imaging detectors (M.Kuster WP75)

AGIPD Adaptive Gain Integrating Pixel Detector (AGIPD)



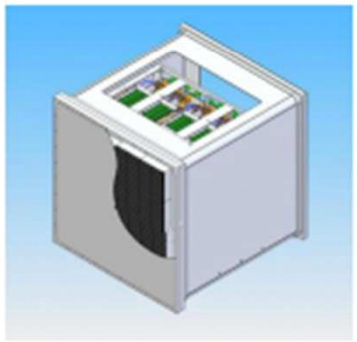
Energy range
3 - 13 keV
Dynamic range
10^4@12 keV
Single Photon Sens.
Storage Cells ≈ 300

DEPFET Sensor with Signal Compression (DSSC)



Energy range
0.5 - 6 keV (25 keV)
Dynamic range
6000 ph/pix/pulse@1 keV
Single Photon Sens.
Storage Cells ≈ 640

Large Pixel Detector (LPD)



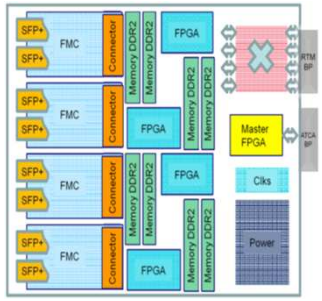
Energy range
5 (1) - 20 keV (25 keV)
Dynamic range
10^5@12 keV
Single Photon Sens.
Storage Cells ≈ 512

Other Detectors

- 0D/1D detectors for high repetition rate applications (e.g. veto, dispersive spectrometers)
- Small areas, low rep. rate, low energy 2D imaging detectors
- Particle detectors (eTOF, iTOF)

Common backend 2D detector development

- **Common detector DAQ and sequencing features**
  - 16 modules = 1 Mpxl
  - One readout 10GE link / module
  - One fast signal sequencing link / module (or quadrant)

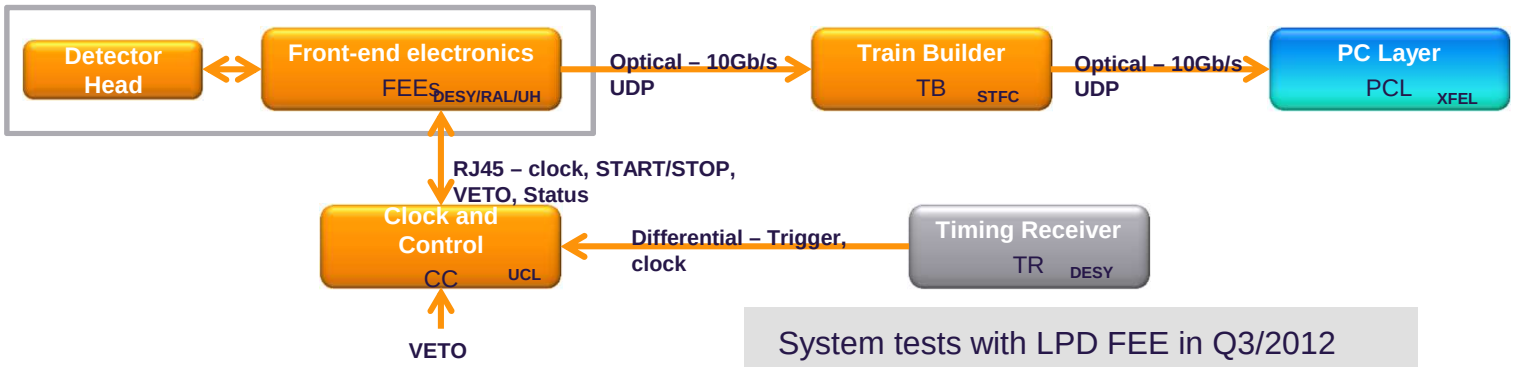


Train builder demonstrator board (STFC)

- **DAQ "train builder" ATCA boards developed**
  - collect image fragments
  - reorganize into complete trains of pulse ordered frames
  - data processing in FPGA – remove empty, no ROI… frames
  - send trains Round-Robin to PC layer

Prototype Clock and Control RTM (UCL)

- **Detector sequencing is controlled by MTCA boards connected to the XFEL timing board in same crate**



System tests with LPD FEE in Q3/2012

DAQ and 2D detector limitations and VETOs

- **Original train builder specification for 1Mpxl detector acquired 512 frames/train**
  - The design can be scaled to larger size, currently 2Mpxl
    - ➔ More depends on memory and internal link speed improvements
    - ➔ A limit will always exist = replicate slices

Detector Mpxl	FEE Links	Nr TB ATCA	RTM & Switch Board	Memory Buffer per FEE MBytes	Data Rate GBytes/sec
1/4	4	1	NO	128	2.5
1/2	8	2	NO	256	5
1	16	4	YES	512	10
2	32	8	YES	1,024	20

- **Detector pipe line limits are**

Detector	Pixel data size	Pipeline depth	Pipeline technology
AGIPD	16 bit	200-400	Capacitor
DSSC	~10 bit	~512	Digital
LPD	16 bit	512	Capacitor

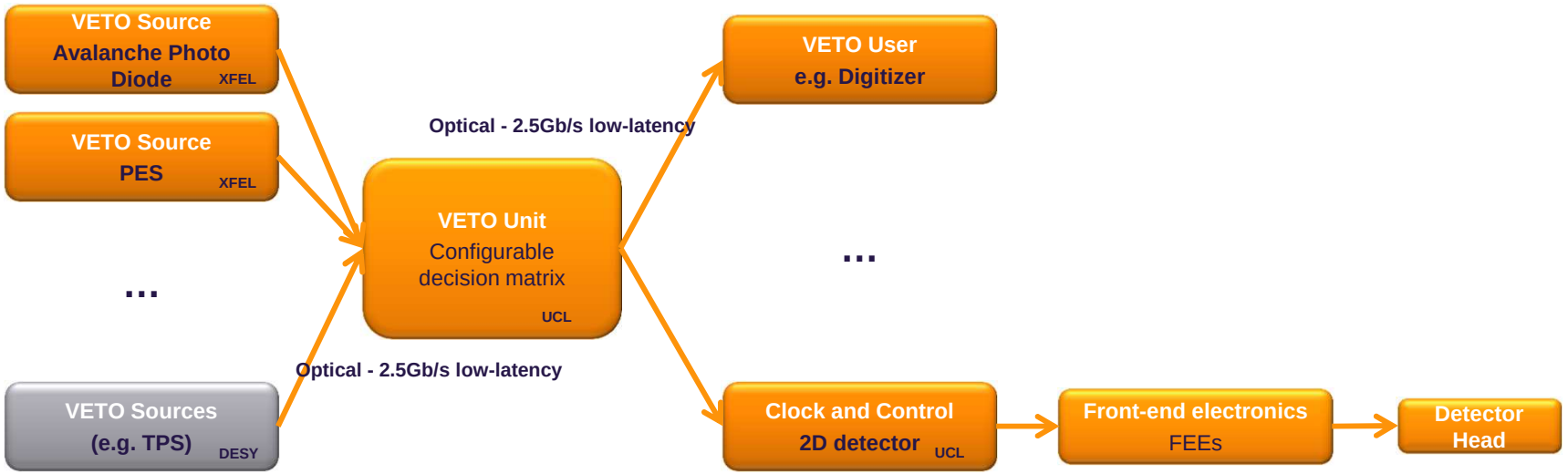
- **Develop VETO system to reject poor quality frames**

Improving image quality – VETO system

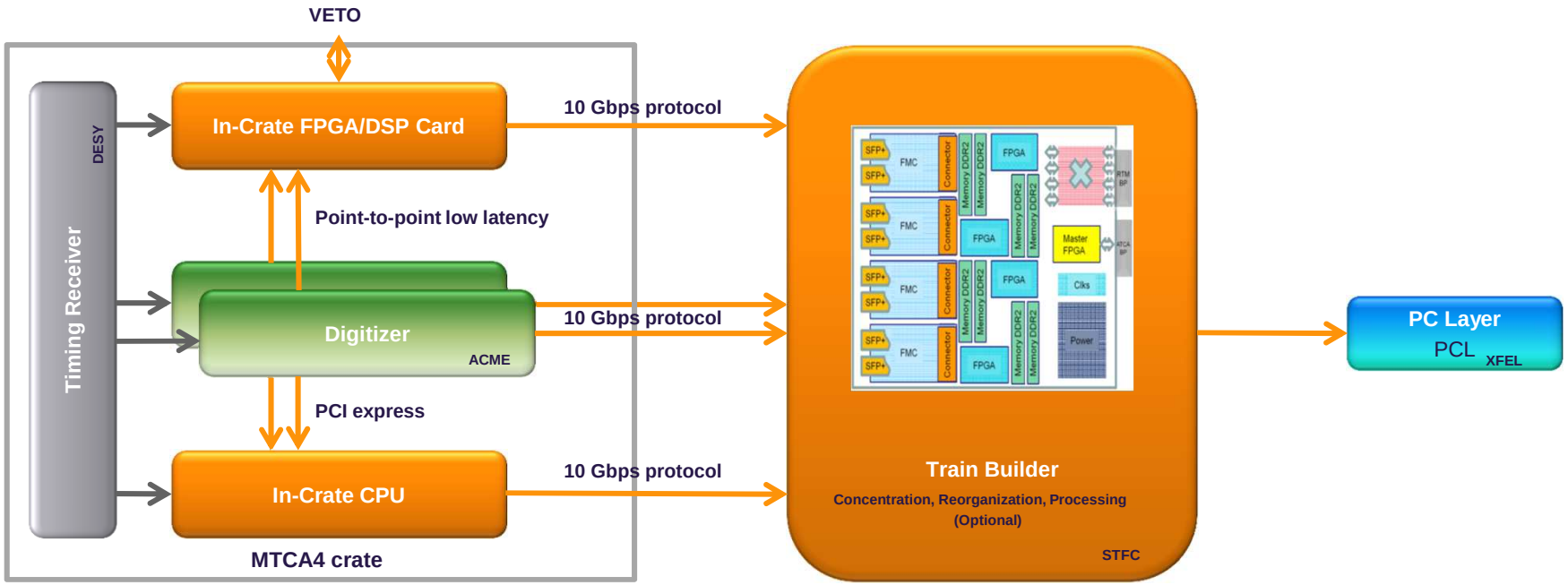
- **A VETO system is being implemented**
  - Clear for reuse storage pipeline cells occupied by poor pulse data
  - Reduce amount of data to transfer or save

- **Centralized VETO unit per experiment**
  - Processes VETO pulse quality measurements from fast diagnostic and measurement devices
  - Trigger decision distributed to VETO users

- **All intelligent FEIs should participate in VETO – specification being consolidated**

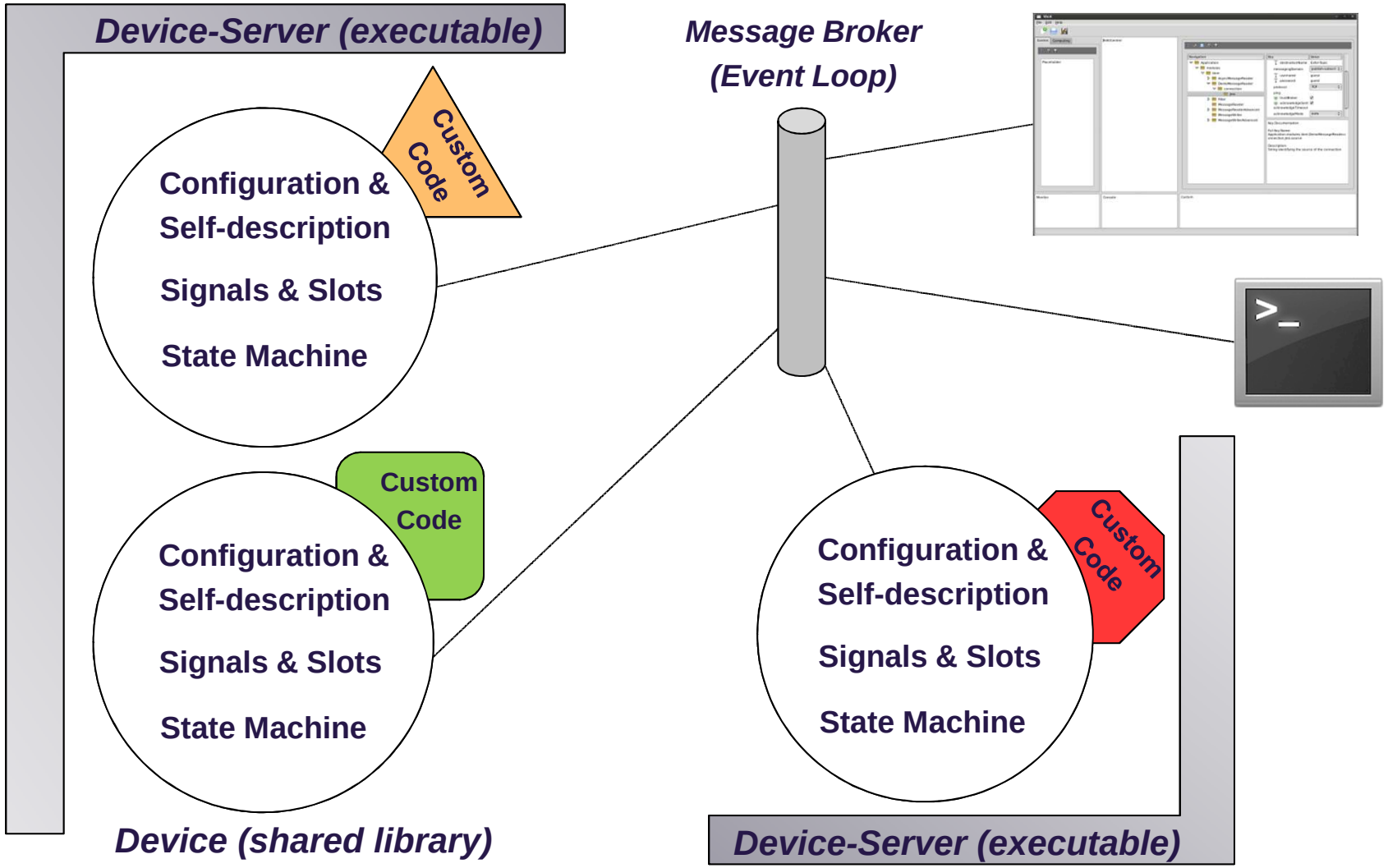


Digitizer integration into DAQ



- **Digitizer selection**
  - 4.5 MHz – 250 MHz using STRUCK SIS8300 (used in single crate APD system)
  - Currently selecting with XFEL+DESY users: 1 – 5 GS/s, 14 – 12 bit resolution digitizers
- **Requirements**
  - Ideally MTCA4 board
  - On-board processing of data
  - Data streaming of all or reduced data to downstream

Control in a nutshell



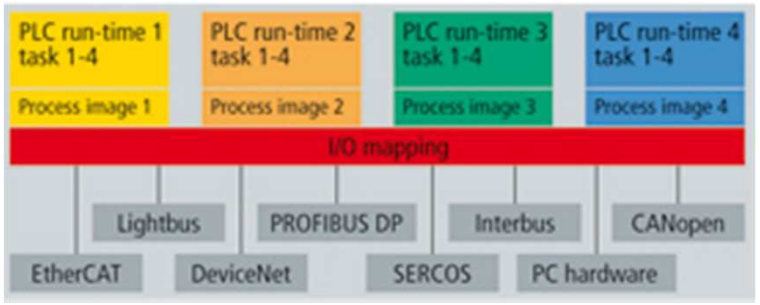
See Burkhard Heisen's talk for details

Beamline control

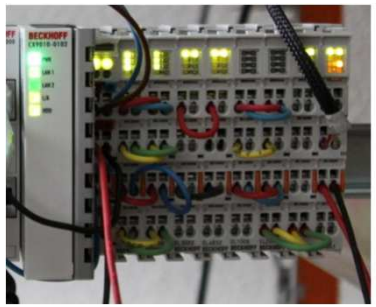
- Beckhoff PLC and terminals are used for beamline control and readout.
  - Already in use with the undulator group (WP71) and at PETRA3
  - Similar implementation ALBA – R&B, ESRF – WRAGO, DIAMOND – ORMRON…
- EtherCAD bus and TwinCAT PLC programming guaranteed real time

Aim is to use only Beckhoff for this type of control work

Programmable Logic Controller (PLC)



Bus terminals



Control

Motors,
Pumps,
Gauges
Sensors,
Actuators,
…

PLC development path similar to ALBA

- PLC development path:
  - Code generated using scripts form csv, later DB, definition of terminals associated with bus, IDs of connected devices, etc.
  - Structures describing properties of devices widened from ALBA implementation to include motors (ALBA use ICEPAP)
  - Firmware downloaded and started with safe configuration, thereafter controlled by a TCP connected device

- Status of development:
  - Software to manage Firmware download and start written
  - Firmware has been developed to configure and control: stepping motors, pump controller(s) , gauges, etc. Most use serial communication.

Important to standardize on a small number of devices

Data management - Policies

- The following policy decisions related to DM are in place:
  - The initial size of the data storage system will be **10PB**, **scalable to 100PB**
    - ➔ Implementation of 10PB storage system is agreed
    - ➔ Expanding the system beyond 10PB requires additional funding
  - Store second copy of data files in archive
  - Data will be archived for at least one year before deletion
  - A reasonable amount of computing power on site will be provided to scientists for data analysis
  - Use DESY IT infrastructure and basic services for implementation of XFEL.EU data management system

Data Management tasks – 1

- Record and maintain data and metadata needed for complete analysis
  - Detectors data, calibration, diagnostic, environmental (conditions), and reduced data
  - Data files stored on disk servers and in the archive (tapes)
  - All stored files registered in the catalogue

- Define logical and physical model for data and metadata
  - IO interfaces (API)
  - Data structures, compression
    - → HDF5 as a data container
    - → database schemas

European XFEL

Data Management tasks – 2

- Provide support and infrastructure for online and offline data reduction and analysis on site
  - Online and offline computing
    - ➔ Computing clusters, CPU and GPU based
  - Storage with highly optimized access
    - ➔ data servers → computing nodes
- Efficient file transfer between different components
  - ➔ PC layer → DAQ data cache
  - ➔ exp. hall → CC
- Small scale data export service
  - For small data producers or reduced data

European XFEL

Data Management tasks – 3

- Coherent user identity management, authentication and authorization scheme for all services
  - User registry, authentication (EU/photon science wide)
  - ACL – role based access to services, data protection
  - Report resource usage
- Provide software framework needed for DAQ, DM and SC
  - Core functionality implementation using C++
  - Python interfaces for non computing experts
  - Main platform is Linux (cross platform GUIs)
  - Pluggable software architecture
  - Software repositories, building and deployment system

European XFEL

Slice test motivation

- Build a stand for testing the complete end-to-end data chain
- Prove the capability of receiving data with max. rate from the train builder, formatting it and sending to the storage devices
- Measure processing capabilities on the PC layer including data compression and rejection
- Test data storage capabilities with different storage hardware variants, find the proper balance between capacity and access speed
- Assess scalability of the selected storage system
- Find and test optimal model for running multiple concurrent experiments, investigate usage of shared resources and data access protection.
- Test and improve data access based on several representative data access patterns

Manpower supported by CRISP, cooperation with other labs

Infrastructure for slice test

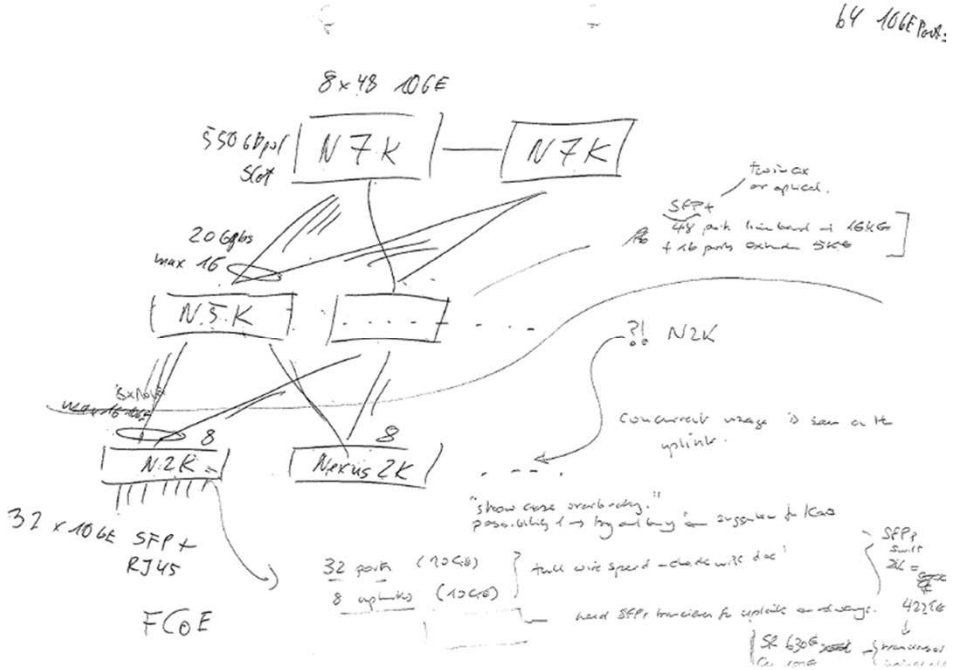
Informal agreement with DESY CC

- Space
- Initially two racks with possibility to extend to four
- Power, UPS, initially ~20kW required, later 40kW
- Cooling (additional power required)
- Network infrastructure for external connectivity
- Access to tape archive
- Personal access to CC

European XFEL

Network setup

Cisco Nexus 5596UP Switch



- Up to 96 x 10Gbps ports
  - This should allow us to transfer data with around 1.1[GB/s]
- Second stage
  - Find the best network topology

European XFEL

PC layer

- Hardware selected
  - 8 x Intel based hosts
    - 2 x 6 core CPU per host
    - 96GB RAM per host
    - 2 x 10Gbps network interface card (NIC)

Functionality	Required software	Software status	Constrains
Receive data from FEI	UDP protocol	Needs final agreement with TB developers, basic UDP implementation exists	1GB/s per stream
Calibration	Calibration framework	Does not exist yet. Needs cooperation with WP75	single image processing
Monitor data quality	Monitoring framework, checksum calculation, histograms, data viewers	Does not exist yet. Some components can be used from existing toolkit	
Data compression	Parallel compression	Need real use cases	Lossless compression, standard algorithms(?)
Format data	HDF5, high level data format, file naming convention	Basic HDF5 implementation exists. High level data format to be defined	1GB/s per stream file buffered in RAM
Send files to storage system	TCP protocol, data shuffling, aggregation	TCP protocol implemented	

European XFEL

Storage

- IBM storage system
  - 14 x 1TB SAS drives per host – 112TB
    - ➔ Raid6 configuration ⇨ **96TB**
  - 96GB RAM per host
  - 2 x SAS/SATA controller [2 x 6Gbps]

Test different configurations for both online and offline storage systems

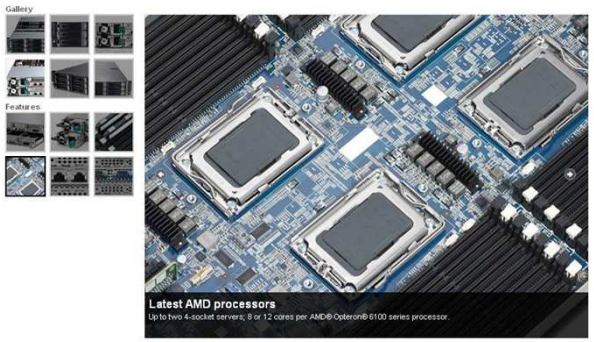
- Cluster file systems (Lustre, GPFS)
- NFS4.1, dCache
- Specialized and optimized setup with data flow controlled by our software – only for online system where the data flow can be well understood and where the requirements are very high in terms of data rates

European XFEL

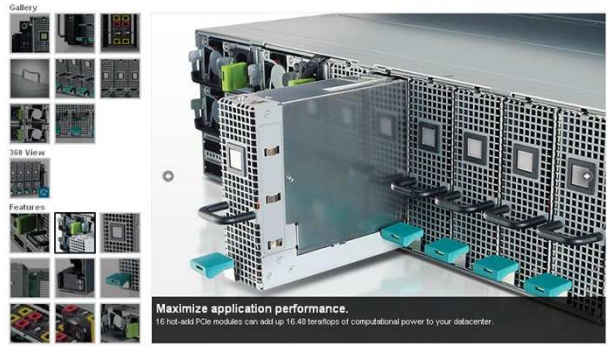
CPU and GPU cluster

- 2 x PowerEdge C6145
- Total 128 CPU cores (8 x16)
- Total RAM: 384GB

PowerEdge C6145 AMD Processor-based 2U Rack Server



PowerEdge C410X PCIe Expansion Chassis



High performance, large memory Fermi-based GPGPU

- 515 GFlops Peak DP
- 6 GB memory size (ECC off)
- 150 GB memory bandwidth (ECC off)
- 448 CUDA cores

Functionality	Required software	Constrains
Data analysis	Scientific computing algorithms ie. single molecule reconstruction	CPU and GPU interoperability
Sharing data between nodes	Pipeline	

Conclusion

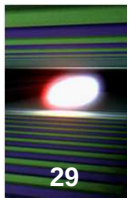
- Progress is being made
  - systems are reaching the test phase which allows design confirmation
  - additional capable manpower available which improves possibilities
  - data reduction and rejection will be a constant companion:
    - ➔ How much can be done on-board in FPGAs and in the compute farms ?

- What was not covered
  - Infrastructure, Network implementation, Data content and format, Person safety interlocks, Machine protection interface, Interface to machine control system, Interface to undulator global control system, Interface to laser systems, Calibration and data correction, QoS, monitoring, etc., Potential of Simulink in FPGA programming, Single crate DAQ system (APD)…

- Thanks to colleagues and suppliers of material !

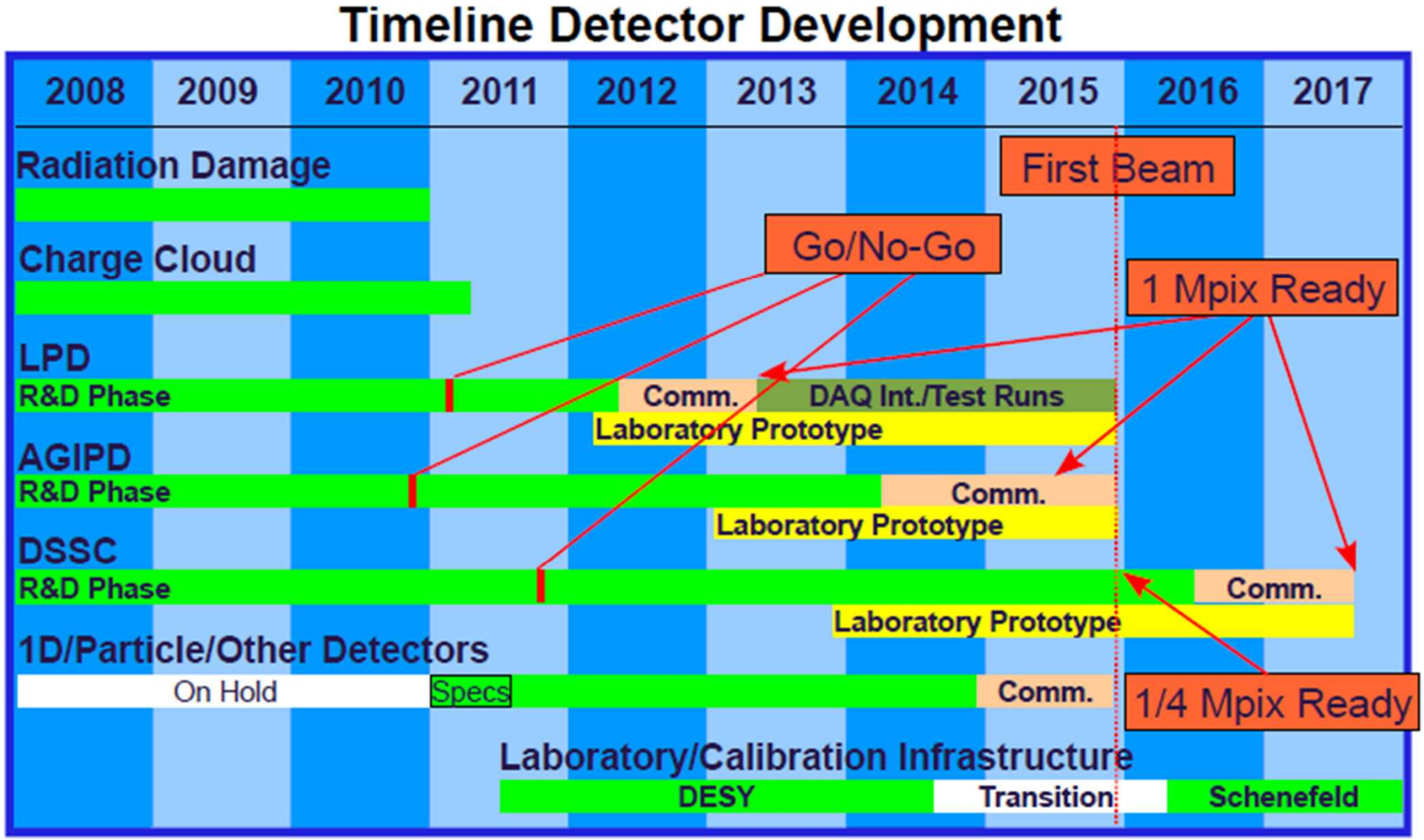
- Poster session contributions:
  - 130 = Overview of DAQ electronics systems for Photon Beamlines and Experiments (XFEL)
  - 131 = The Train Builder Data Acquisition System for the European-XFEL (STFC)
  - 150 = Clock and Control Sequencing System for 2D Detectors at the European XFEL (UCL)
  - 152 = A homogeneous software framework for XFEL.EU

Spare slides

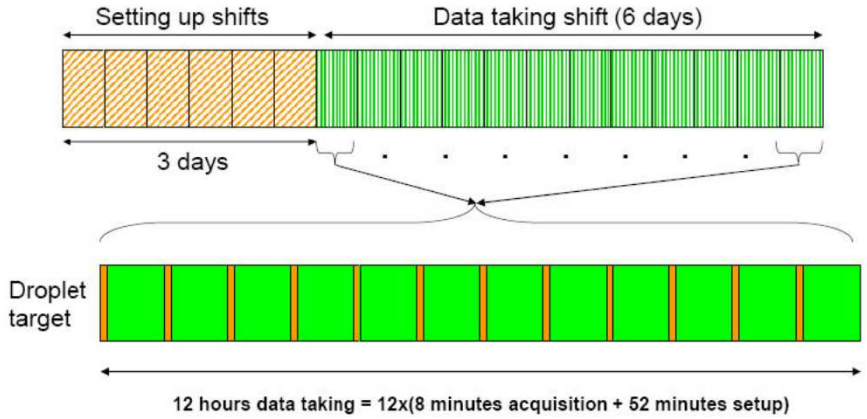


European XFEL

Detector development schedule (M.Kuster WP75)



SPB data volume – revisited



- **Revisited the 2009 Computing TDR's SPB data volume calculation:**
  - **100 kHz sample injection with 2% frame efficiency**
  - **10**5 good frames needed for analysis = ~16 minutes running at 512 frames / train**
  - **10**5 – 10**6 hit pixel multiplicity (not gas target)**
  - **44 minutes target swap**
- **Raw data volume ~200TB/day**
- **Raw data volume after "empty" frame rejection = 4TB/day**

Message is again: SPB and DAQ need to optimize VETO, rejection