Data acquisition in high energy physics

Matthew Wing
(UCL, DESY and UHH)

- Introduction
- Physics rate and collider and detector parameters
- Technology
- Example I: CALICE DAQ
- Example II: CMS tracker upgrade
- Comparisons HEP/PS and summary

Joint Workshop on Detector Development, Hamburg, 16-17 October 2008

Introduction

Introduction - HEP

High energy physics has long been large-scale science:

• Particularly collider-based projects are massive undertakings of great complexity

• Two strands:

  • Search for new (usually rare) phenomena beyond the Standard Model

  • Precision measurement and understanding of known phenomena

• Quantum mechanics is probabilistic, so the enquiry is closely related to statistics

  • Balance between collecting all data and all “useful” or “exotic” data - searching for a needle in a haystack whilst understanding all you can about the haystack

Introduction - HEP DAQ

DAQ systems have had to cater for the needs of HEP:

- Cope with high bandwidth data
- Pick out (“trigger” on) the interesting or spectacular events
- Collect data with 100%(!) efficiency
- Work for a long time without(!) fault
- Integrate into a system of many different components
- Cope with upgrades of accelerators and detectors

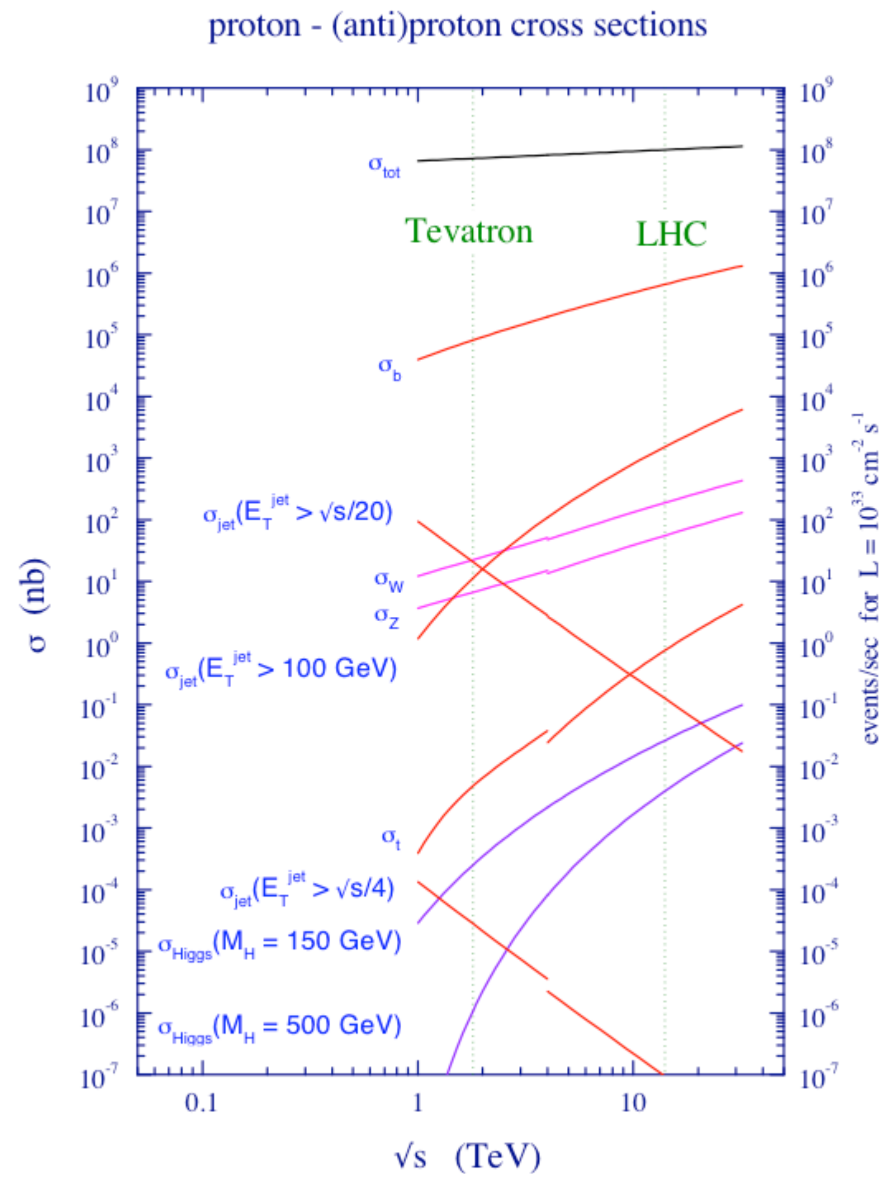
Introduction - HEP DAQ

- HEP has been a driving force for technology - the collection and transfer of data of such high volume and rate was not common at the advent of collider physics
- HEP physicists and engineers had to be clever and design electronics and DAQ systems for their needs
- Each accelerator/experiment/detector had a bespoke system to cater for its needs
- With industrial advances, have the opportunity to use commercial off-the-shelf equipment
- Can reduce costs, development time and risk
- Being considered by new experiments, SuperLHC and ILC, and should apply to any others

Physics rates and collider and detector parameters

Physics rates

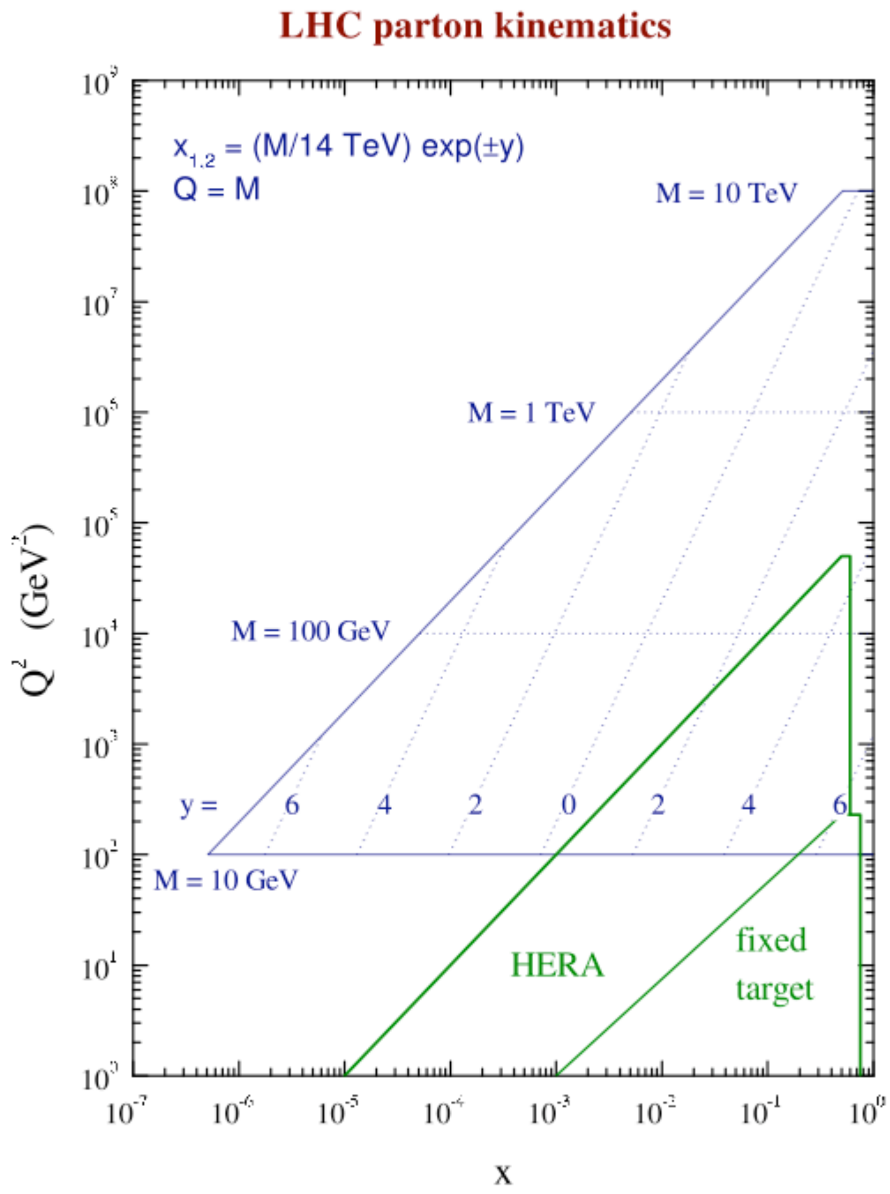
- Highly collider dependent, cf. e^+e^- and pp
- Not all events are as interesting as others
- E.g. keep all events at e^+e^- colliders and not at pp ⇒ “trigger”
- Note “really interesting” events are 10 orders of magnitude less likely
- Note *total* number of events 10^8/sec
- Also have backgrounds, e.g. beam-gas interactions



Campbell, Huston & Stirling, RPP **70** (2007) 89

Physics range

- Have detector geometry and kinematic variables: we don't want “holes” in either.
- Keeping events with high energy will be no issue.
- Have to be clever for lower energies (high rate) to cover all kinematic plane.
- Need calibration samples too.



Campbell, Huston & Stirling, RPP **70** (2007) 89

Collider parameters

- Bunch spacing
- Number of bunches in a train
- Train spacing (repetition rate)
- Luminosity
- Beam energies

These are the starting point for determination of:

- Detector occupancy, physics rate (acquisition rate)
- Triggering
- Detector calibration scheme, when and how much (acquisition rate)
- Readout structure (inter-bunch/inter-train)
- Timing (control) structure - timing in to accelerator and when to pass controls
- Radiation studies

Example colliders - e^+e^- linear colliders

Parameter	ILC	CLIC
E_{CM} (GeV)	500	3000
Bunch separation (ns)	369	0.5
No. bunches/train	2625	312
Repetition rate (Hz)	5	50
Luminosity (10^{34} $cm^{-2}s^{-1}$)	2	5.9
N_{hadron} per crossing	0.12	2.7

Colliders with differing parameters will need different DAQ systems(?)

Good planning needed for a DAQ system to cope with both colliders

ILC Reference Design Report, http://www.linearcollider.org/cms/?pid=1000437
2008 CLIC Parameter note, http://cdsweb.cern.ch/record/1132079?ln=fr

Example colliders - LHC

Parameter	LHC	SuperLHC
E_{CM} (GeV)	7000	7000
Bunch separation (ns)	25	25
No. bunches/train	2808	2808
Luminosity (10^{34} $cm^{-2}s^{-1}$)	1	15.5
Events per crossing	19	294

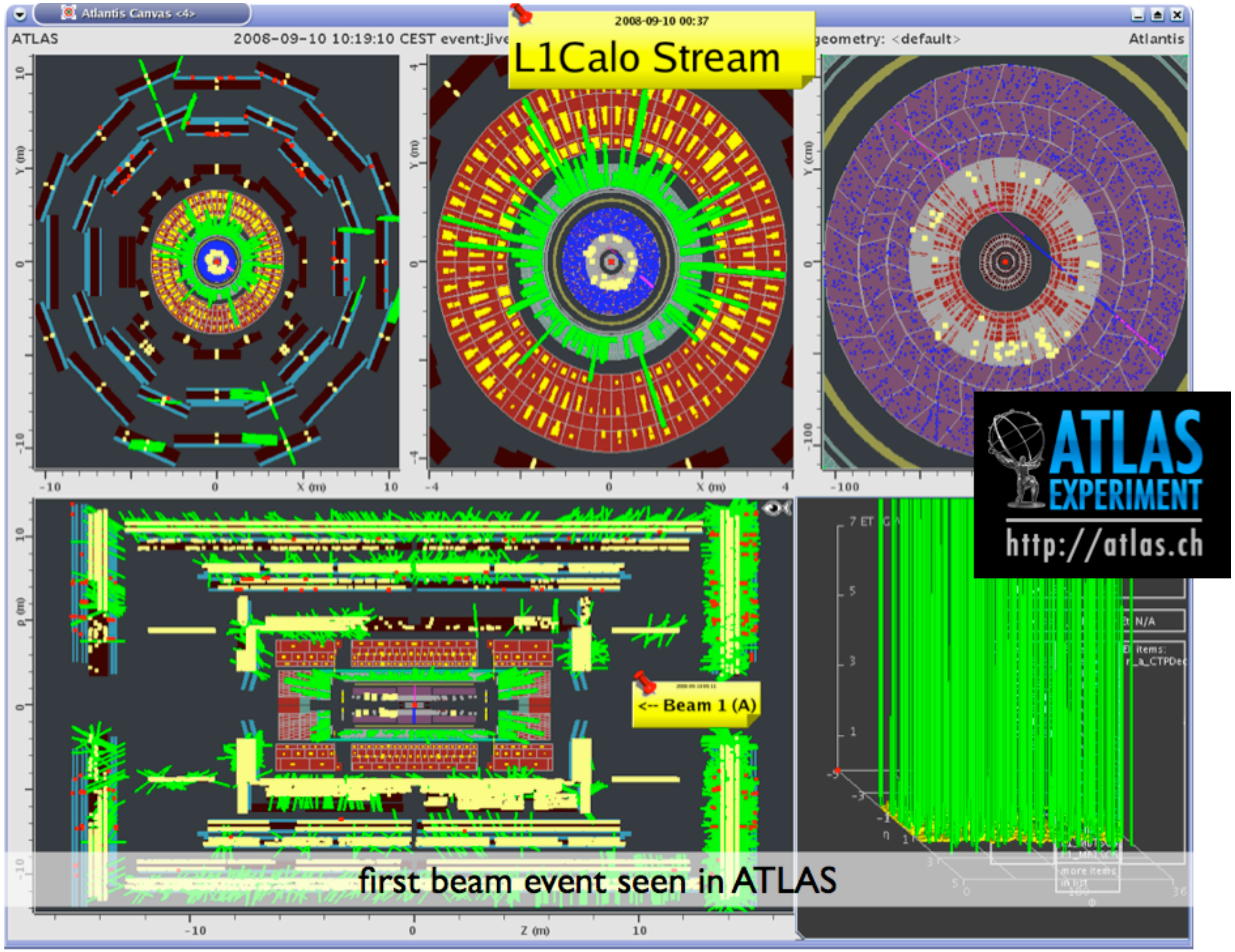
Upgrades will need radically new DAQ systems

O. Brunning, LHC challenges and upgrade options, JoP Conf. Ser. **110** (2008) 112002

A colliding-beam detector

- Chosen as a generic example
- HEP experiments contain similarities
  - o tracker(s)
  - o calorimeter(s)
  - o muon chambers
  - o luminosity monitors
  - o triggers
- Many sub-detectors
- Many technologies
- Different occupancies

- Other experiments can be challenging for DAQ systems, but concentrate on colliders here



Detector parameters

- Dimensions: number of channels, space for connections
- Response time, speed of data transfer on detector
- Number of different sub-detectors and their integration
- Zero suppression, reduction of noise
- Triggering
- Calibration scheme (how often, how much of detector?)
- Amount of data which can be buffered and where
- Readout scheme (all/some data, between bunch crossings/trains?)

Example detector - CALICE ECAL

CALICE propose a high granular calorimeter for detecting (electromagnetic) deposits at the ILC

Assume:

- 0.5 x 0.5 cm^2 channels
- 100 M channels in total
- 6000 detector "slabs"
- Raw data size, 2 Bytes/channel, with 4 Bytes/channel for labelling

Data size/bunch train = (100 x 10^6) x (2625) x (6) = 1575 GBytes

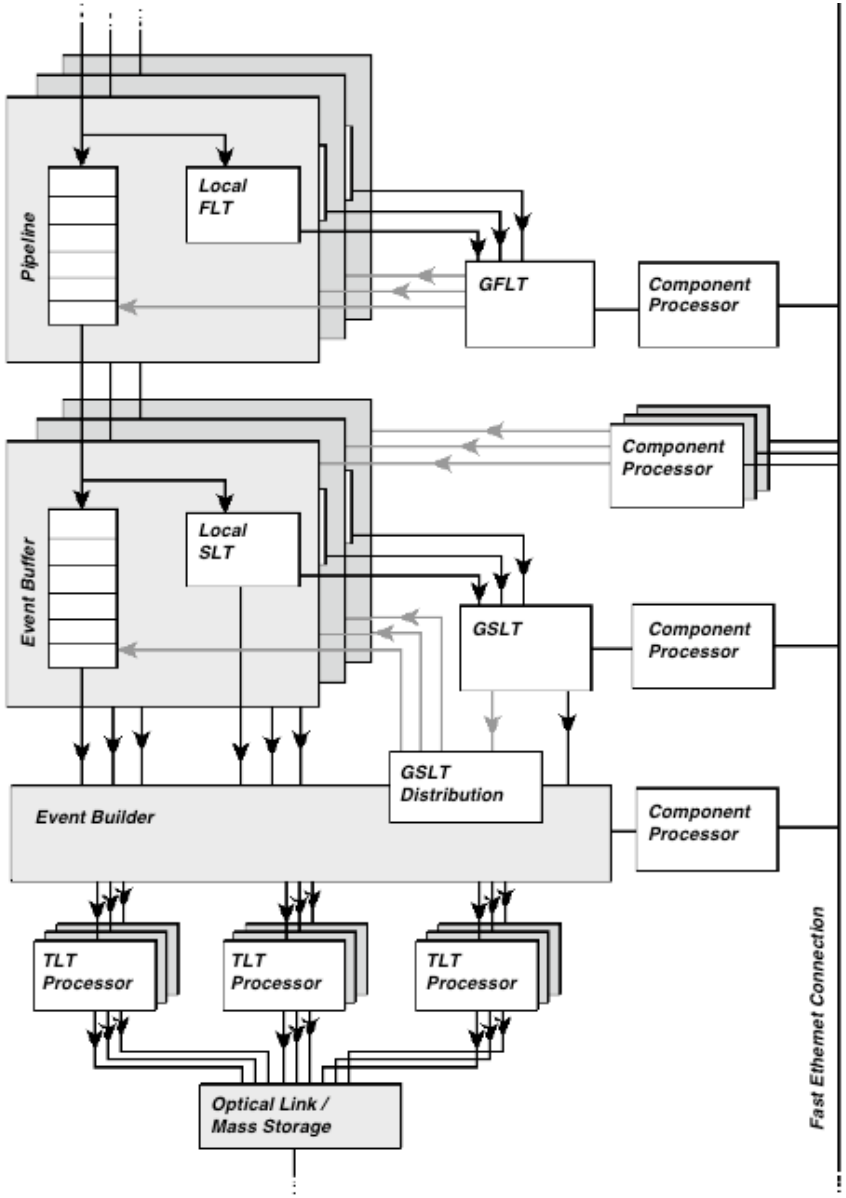
Readout during bunch train = (1575 GB) / (2625 x 369 ns) = 1626 TBytes/s
Or 271 GBytes/s/slab which is clearly very challenging!

Solution:

- Assume pessimistic threshold suppression reduction factor of 100
- Data read-out between bunch trains (200 ms instead of 970 μs)

Readout speed = (1575 GB)/(100)/(0.2 s) = 79 Gbytes/s
Or 0.1 Gbit/s/slab which is clearly far more manageable!

Example detector - the ZEUS trigger system



An example:
- three levels, 1 h/w, 2 s/w
- input rate ~10-100 kHz (mainly beam-gas), output rate <10 Hz (mainly physics)

First Level Trigger
- reduce rate to 1 kHz
- individual components have own FLT: energy, timing, vetos, etc.
- data pipelined

Second Level Trigger
- reduce rate by factor of 10
- individual components have own SLT, simple reconstruction: energy, tracks, kinematic cuts

Third Level Trigger
- final reduction
- fast version of final reconstruction: kinematic cuts, algorithm output

Technology advances

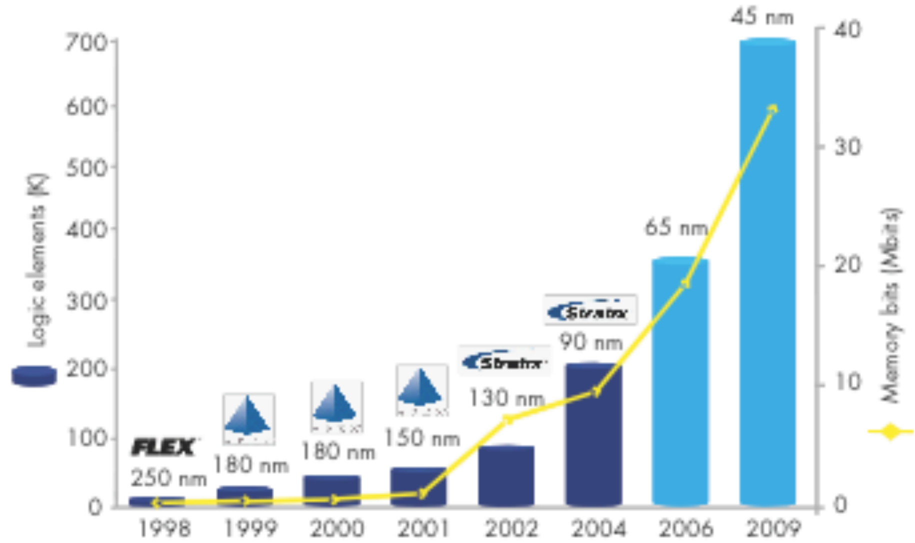
Technology advances

- Globalisation and WWW have created the need in everyday life for cutting-edge technology
- Telecommunications industry has particularly seen many advances.
- Less need now for HEP (or science in general) to develop bespoke equipment
- Academic research needs to embrace the technical expertise and advances of the commercial world.
- Some are:
  - Capability of FPGAs
  - New crate standards
  - Links, networks, switching
- Large-scale science can still contribute though…

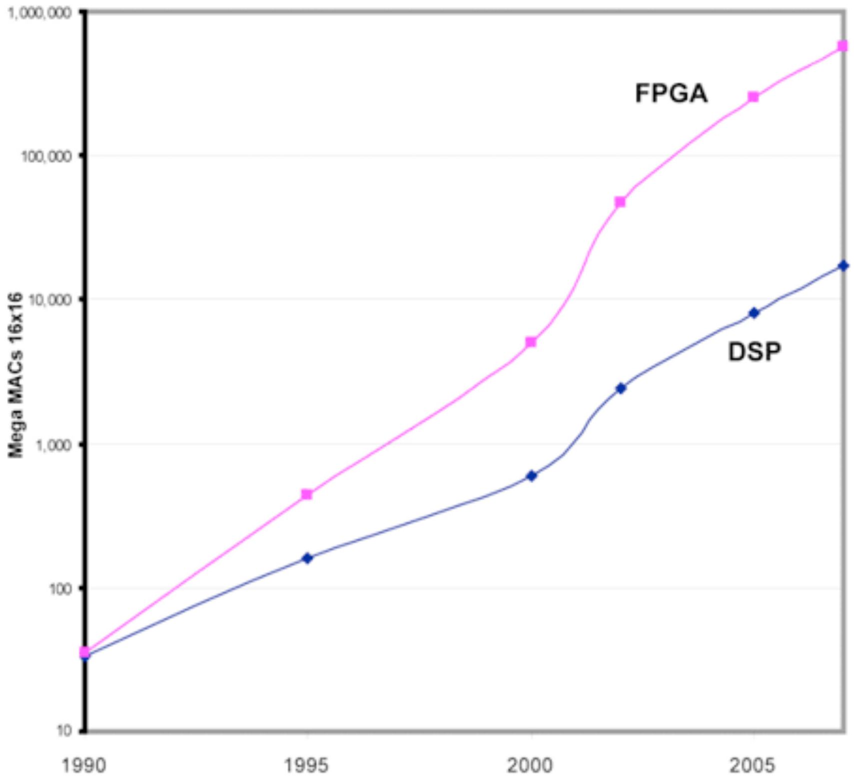
FPGA performance

- Performance increasing dramatically
- Include PowerPC cores
- Low cost

Figure 3. FPGA Memory and Bandwidth Continue to Scale



Accelerating High-Performance Computing with FPGAs, Altera White Paper, http://www.altera.com/literature/wp/wp-01029.pdf



Note: Assume 33% of device used for 16x16 @ 250LEs per MAC

- Similar to increase computing capacity
- Used in wide range of industrial and scientific applications

Paul Ekas, Altera Corporation, http://www.dsp-fpga.com/articles/ekas/

xTCA crate systems

Advanced Telecommunications Computing Architecture (ATCA) is a new standard

- 100 companies
- PMC/AMC mezzanine cards

μTCA (small/dev system)

1	MicroTCA Backplane	6	Power Supply 1 (optional)
2	Cable Tray	7	Fan Tray
3	Air Filter	8	DC Outputs Power Supply 1
4	DC Outputs Power Supply 1	9	ESD Wrist Strap Terminal
5	Grounding Terminals	10	Slot for Power Supply 2

ATCA (the new standard)

- μTCA uses AMC cards

Looks to be the way forward for DAQ systems

Links and networks

Serial links are becoming the norm:

- faster, more reliable and cheaper than parallel architectures
- ethernet, Serial Attached SCSI, PCI Express, …

E.g. PCI Express

- x1 … x32 lanes
- 250-500 MB/s per lane each way, growing linearly with lane
- Total throughput 8-16 GB/s

10 Gigabit ethernet becoming the standard and should be considered for future systems

- Ethernet has been going for over 25 years
- 10 Gb switches becoming perfomant and cheaper

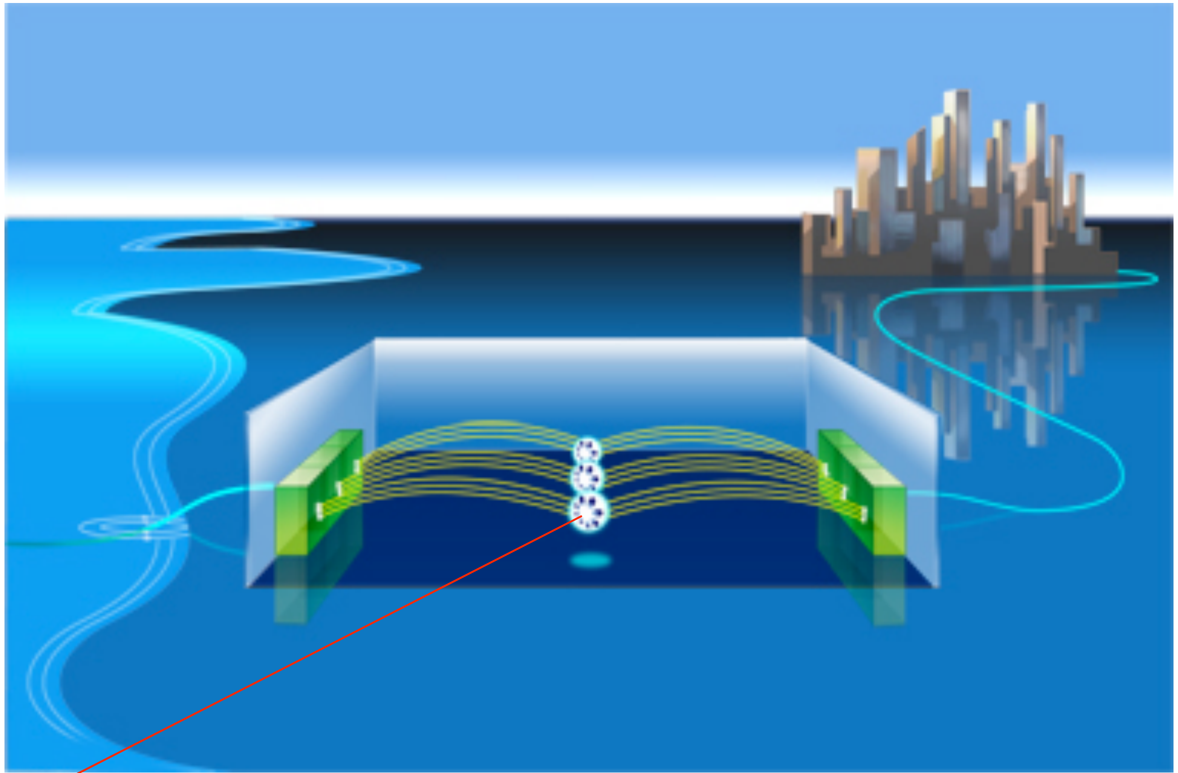
Optical switches

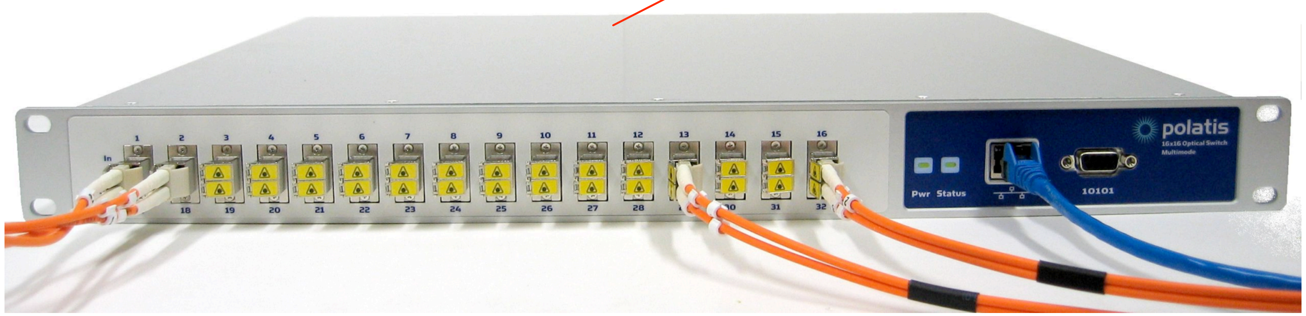
• Becoming a new telecom standard

• New method of managing optical fibres, efficiently and securely

• Used in defence, undersea cabling, …

• Could be used in HEP/science for data transfer from detector





The Grid

• Can make use of the processing and storage capabilities of the grid.

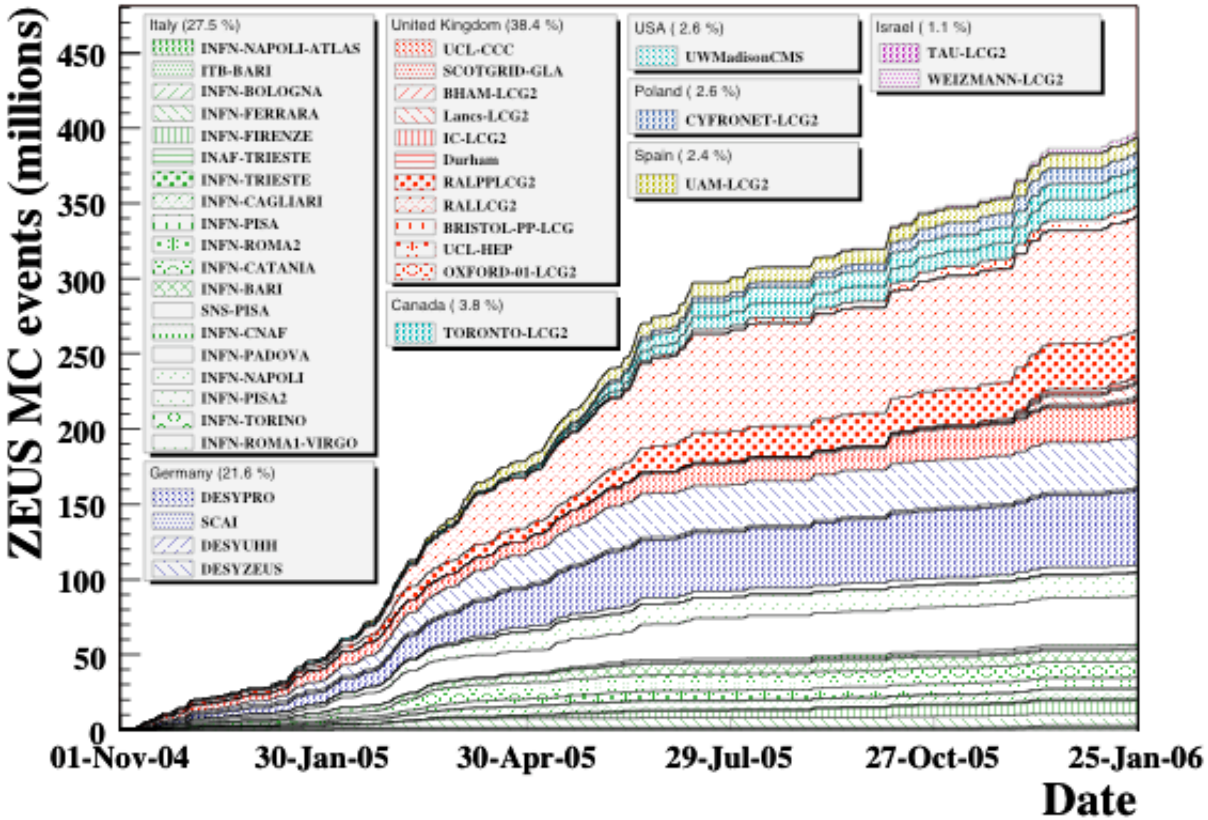
• Set-up for the vast amounts of data to come out of the LHC.

• HERA experiments used grid to increase MC production

• Note number of sites

• CALICE has all test-beam data on the DESY grid

• Data transferred directly from test-beam site (e.g. CERN) to grid for processing and storage



Bespoke versus COTS

Bespoke

- In control of the scientists and engineers designing it
- This may be the only solution
- Can still be “gaps in the market”: WWW, Grid

COTS

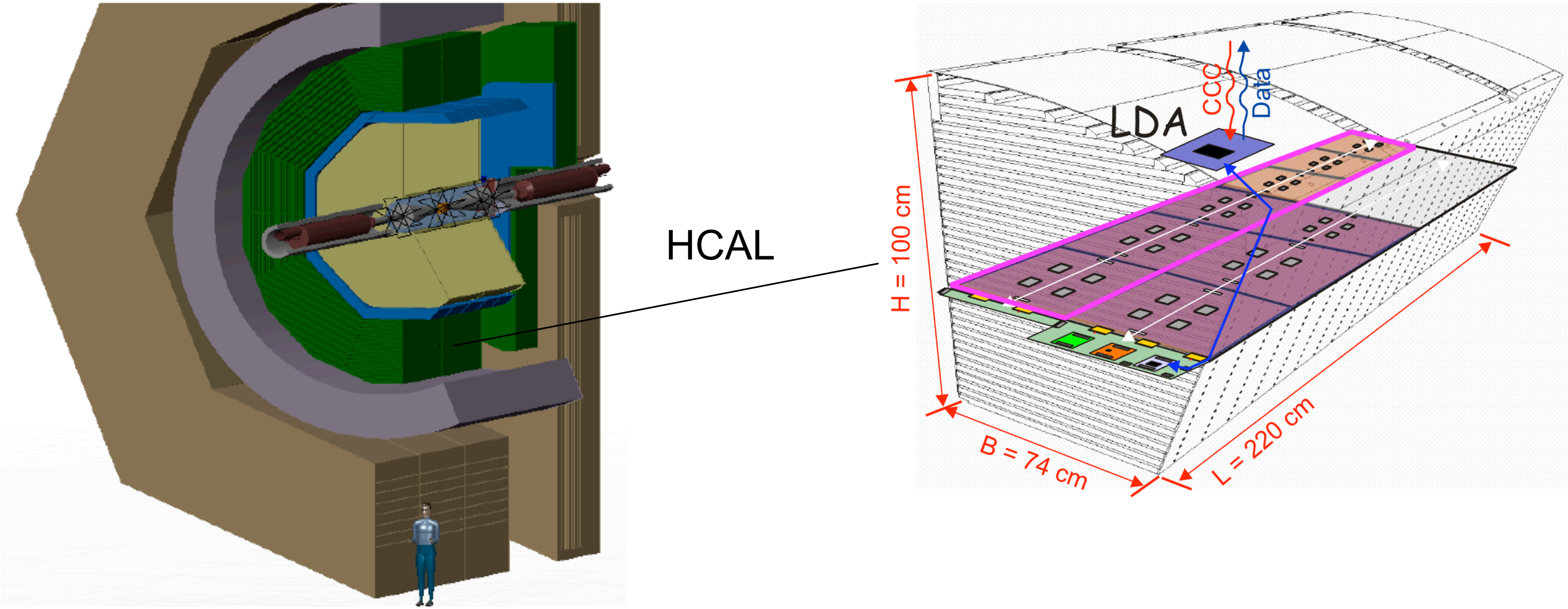
- Save on costs and development time
- Should be flexible, scalable and upgradeable
- Provides a good platform for knowledge exchange with industry and also in research itself

Example DAQ systems

Example I: CALICE DAQ

- Basic R&D into data acquisition systems for itself and for calorimeters at the ILC
- Have a conceptual design of a DAQ system for calorimetry at the ILC (even though far off)
- Develop a system using industrial standards and advances: flexible, high-speed serial links, scalable, using commercial off-the-shelf components
- Deliver working DAQ system for prototype calorimeters
- DAQ system could be applicable for final system or other detector systems

Calorimetry at the ILC



- Building large-scale prototypes for technological tests and test-beam verification
- The DAQ system will cope with several calorimeters: AHCAL, ECAL, DHCAL (+)
- Different beam and/or timing structures - flexible
- Comparable to a conventional HEP experiment

Overall DAQ architecture

Detector Unit: ASICs

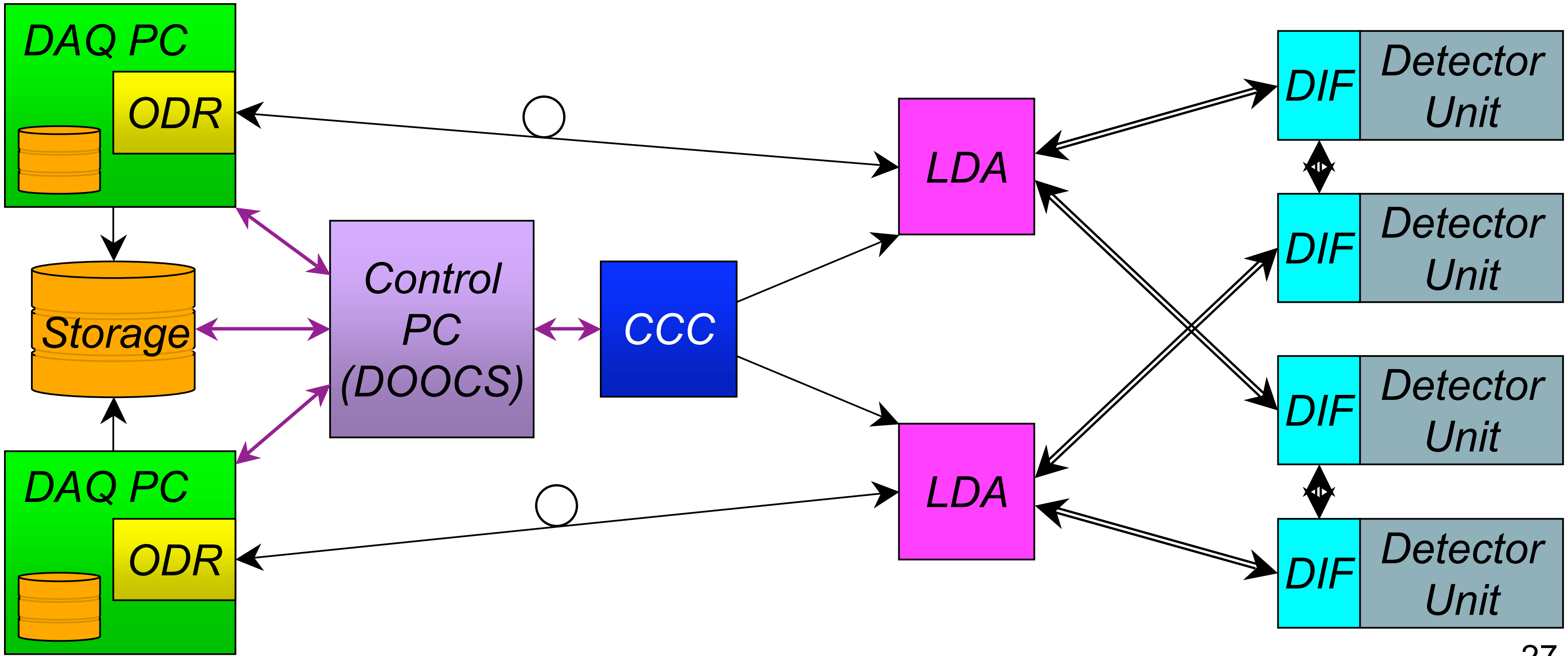
DIF: Detector InterFace connects generic DAQ and services

LDA: Link/Data Aggregator fansout/in DIFs and drives links to ODR

ODR: Off-Detector Receiver is PC interface

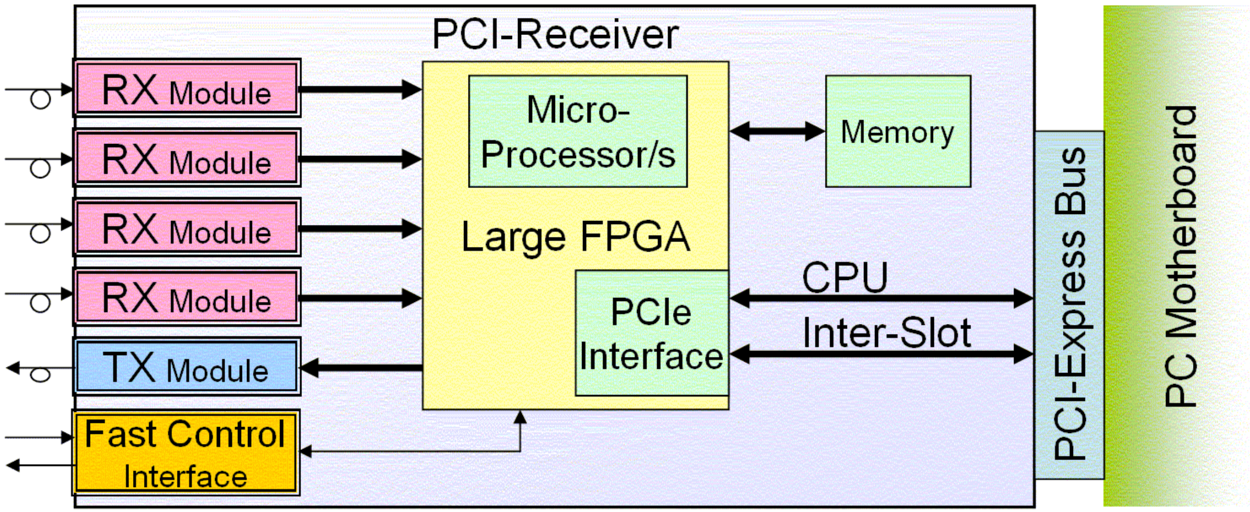
CCC: Clock and Control Card fansout to ODRs (or LDAs)

Control PC: Using DOOCS

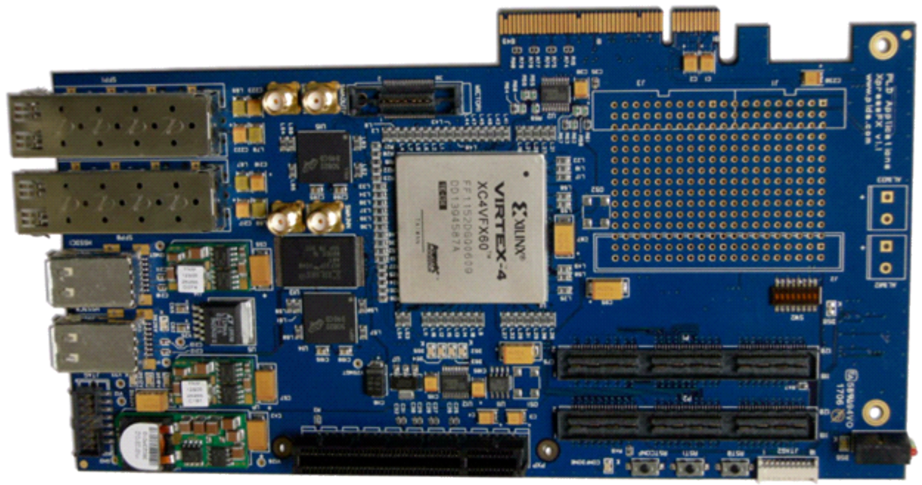


Off-detector receiver

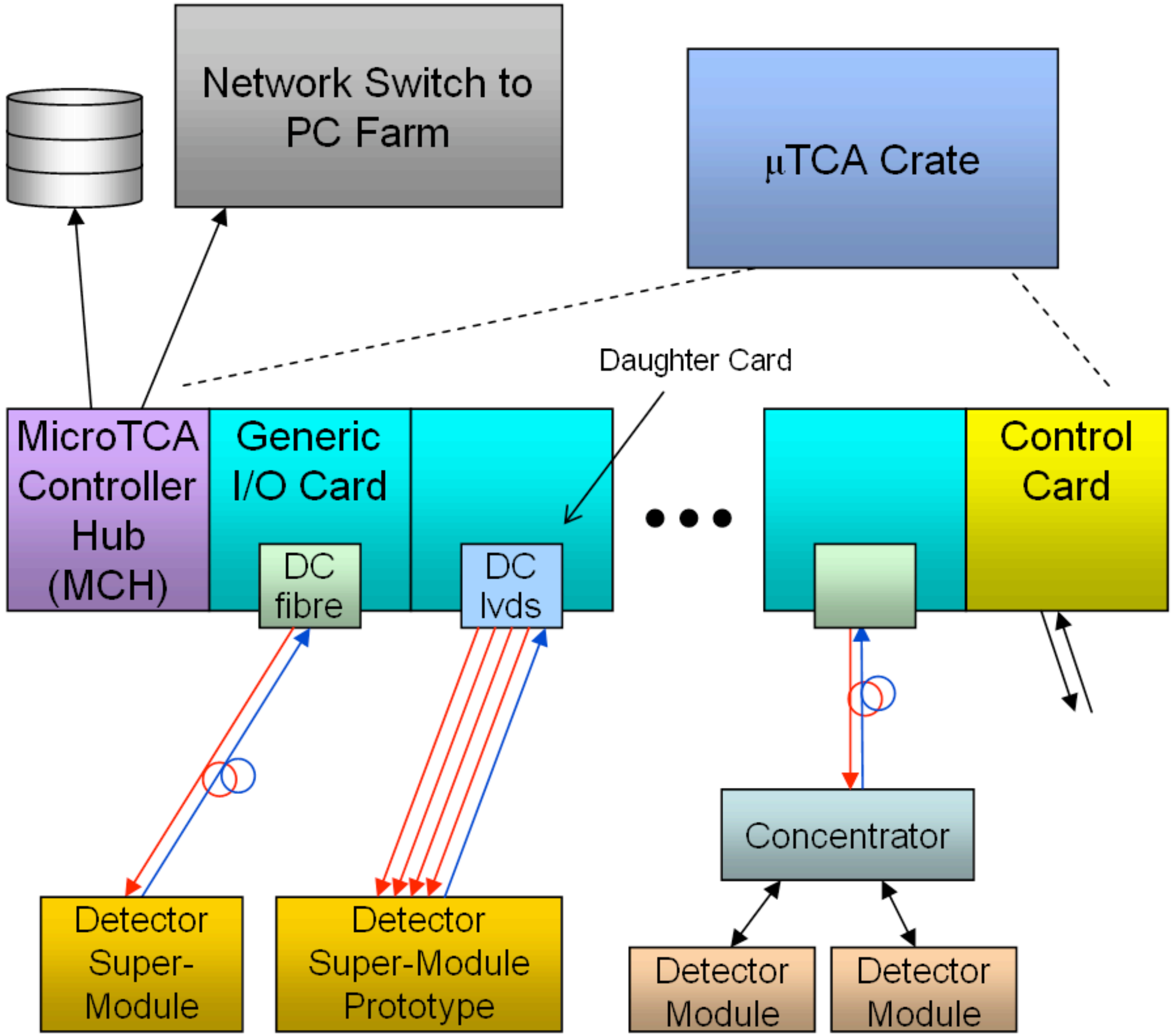
- Based around a PCI card housed in PC
- PCI Express bus
- Large FPGA
- Rx/Tx modules
- xTCA crate systems were appearing when originally thought about this



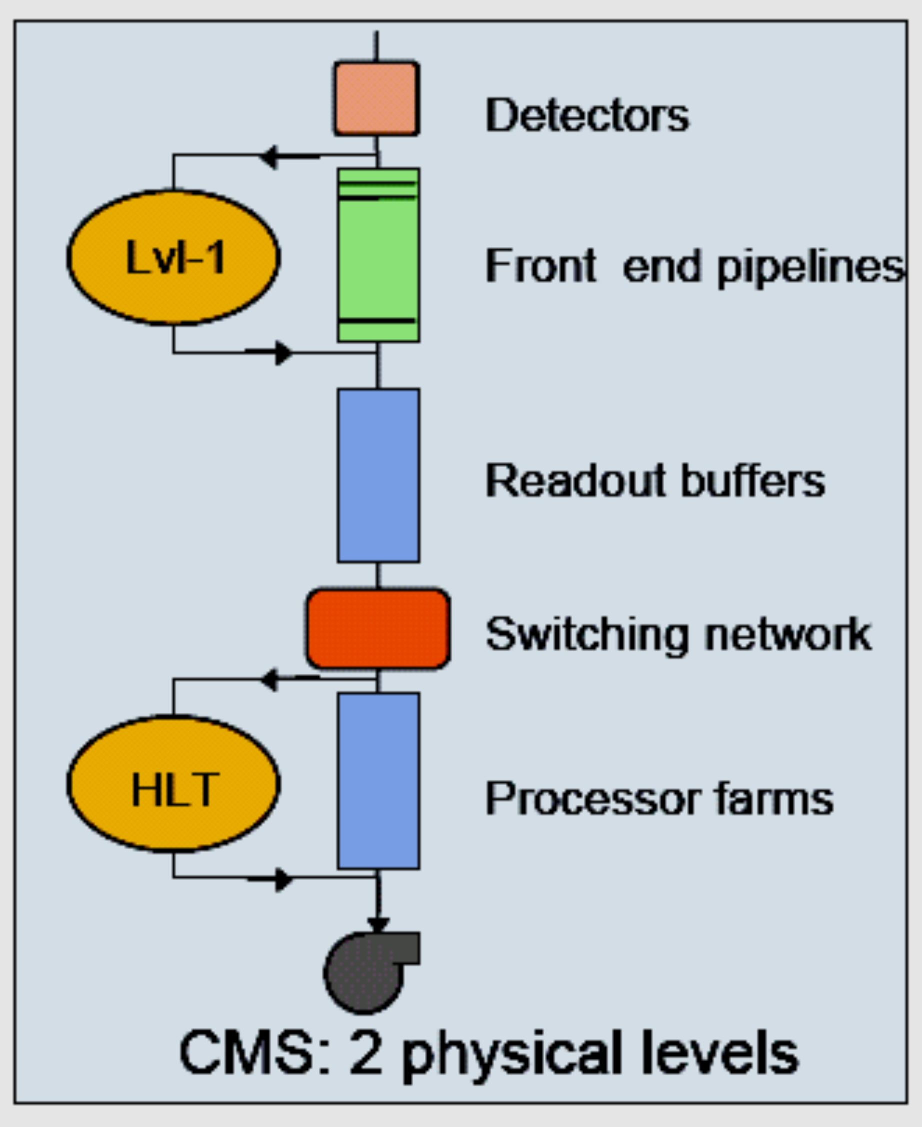
- Originally thought to design card, but COTS card existed from PLDA
- Will use this DAQ system and evaluate its performance
- Comparison to xTCA systems will be valuable
- Much of this should be transportable to xTCA crates and cards



Layout of a μTCA system



Example II: CMS trigger - LHC and SLHC



- 40 MHz input rate
- Data stored at L1 for 3.2 μs
- 100 kHz output L1 rate
- 100 Hz written at output
- Extremely demanding environment

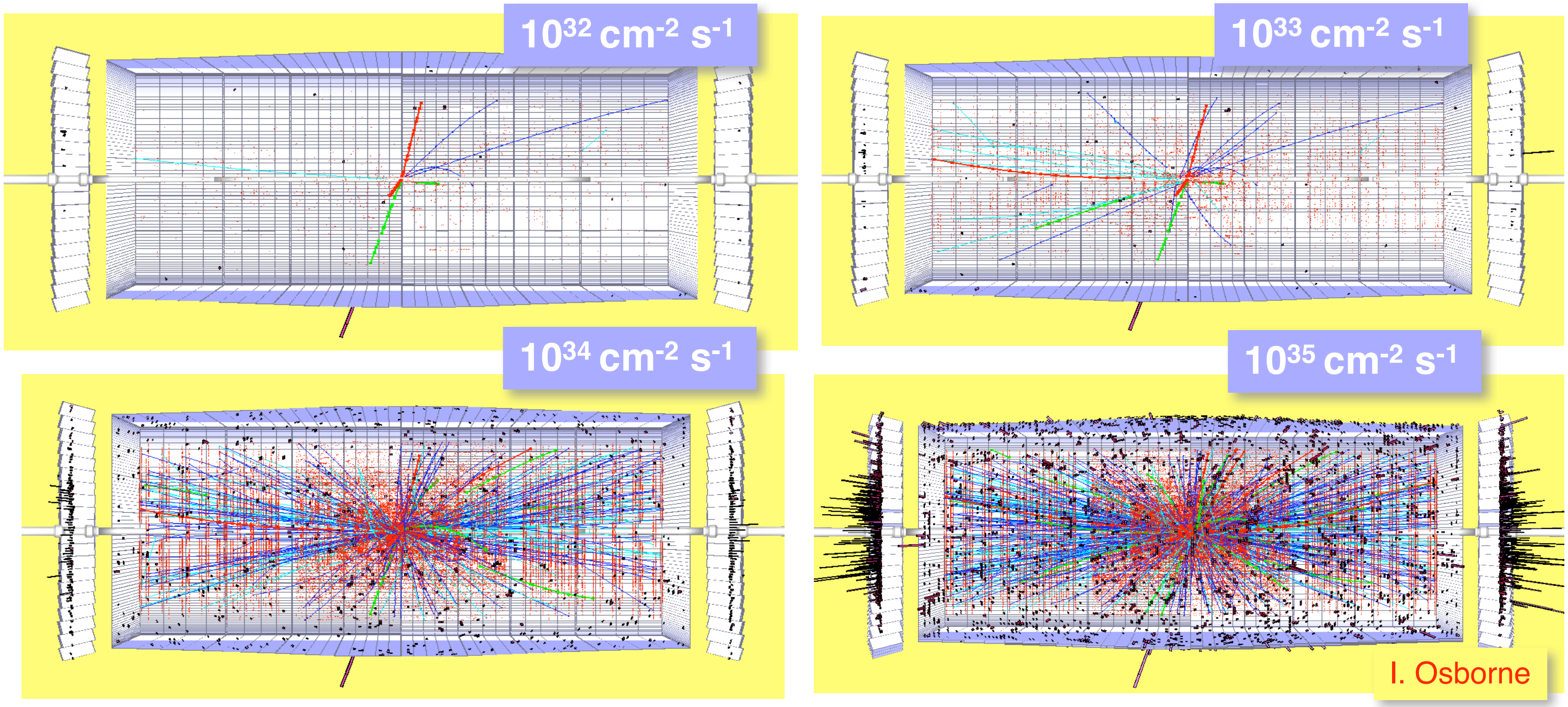
- L1 uses custom hardware processor
- HLT is a PC farm with software filters

- L1 is based on calorimeter and muon detectors:
    - High E_T jets/e±/γ
    - High p_T muons

- This will all suffice for the LHC, but not SLHC.

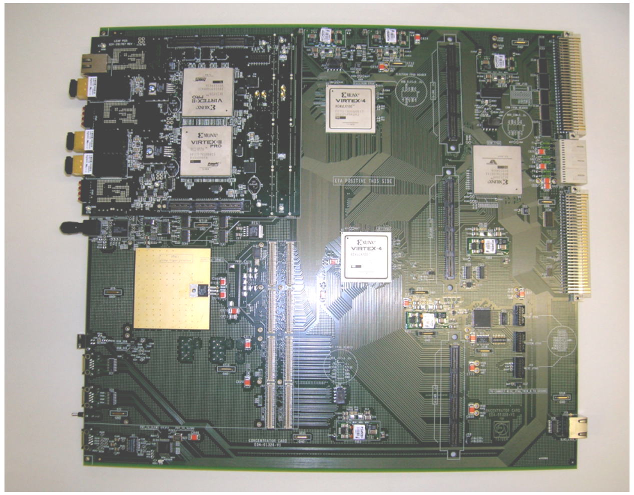
C. Foudas et al., The CMS trigger at LHC and SLHC, ICHEP08, Philadelphia

The SLHC challenge



- Raising thresholds might not work and might not be desirable
- Need to use a tracking trigger for significant reduction
- Need fast and powerful electronics for algorithms

The hardware solution



Custom hardware processor → μTCA crate

- Designed to accept data from different detectors
- Can it become a CMS wide standard?
- Could have significant cost savings

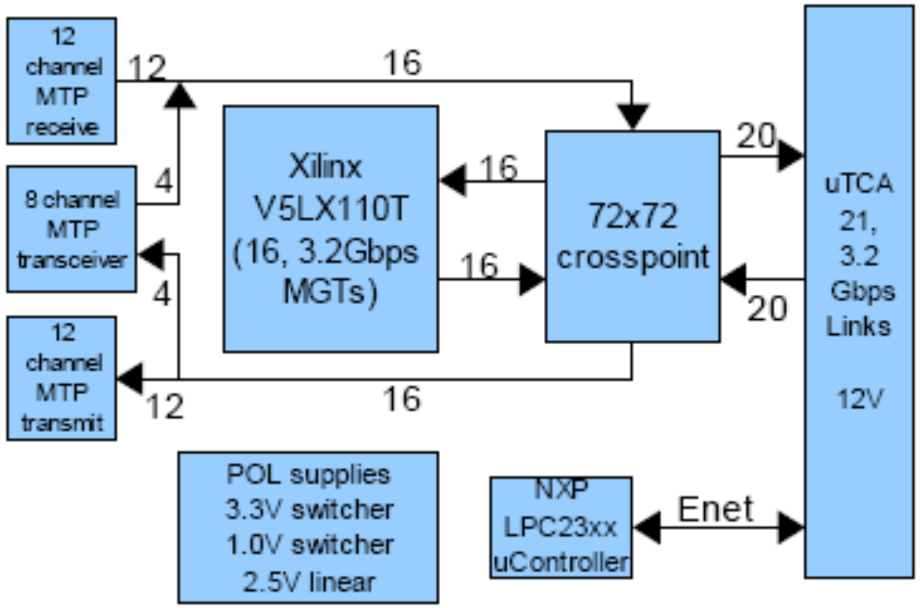


Figure 2: Processing module block diagram

Discussion and summary

Comparisons: HEP/PS

- Due to the facilities required and nature of enquiry, HEP has (traditionally) pioneered and pushed back boundaries in detector development and specifically DAQ
- Photon science is now using large-scale facilities with many coming online in the near future (more than HEP)
- Industrial advances are changing the landscape
- Strong parallels between the two groups, HEP and PS, cf. XFEL and ILC
- What is similar and what is different?

Contrasts

- HEP has generally had dedicated DAQ systems, PS detectors will be used at different light sources
    - Need flexible system from the start for PS
    - Building generic (flexible) systems in HEP is a way to go
- Accelerator structures are similar
- PS detectors more simple(?) but also has large detectors: issues of noise, threshold suppression
- HEP you let run and don't change (single user), PS will have many different users
- Will PS have to think of triggers?
- Overall, I think that HEP and PS are thinking along similar lines about data acquisition systems

Summary

- The use of large-scale apparatus is unifying the DAQ needs
- Technology advances are driving commonality
- Due to the timing of large experiments, PS has the opportunity to push this forward. HEP will need it on a similar time-scale for SLHC
- Overall, I think that HEP and PS are thinking along similar lines about data acquisition systems and we should share our experiences more closely and in more detail