UK input to European Strategy: perspective from Trigger & DAQ



## Veronique Boisvert

With inputs from: D Newbold, F Pastore, S George, P Teixeira-Dias, B Green, A Tapper, N Konstantinidis, M Wing, J Brooke, P Dauncey, D Sankey



ROYAL HOLLOWAY UNIVERSITY

# Approach of this talk



- Quick overview of TDAQ for HL-LHC (not part of strategy!) as an intro to other future projects:
  - ILC, CLIC (FCC-ee)
  - FCC-hh (HE-LHC, chinese colliders)
  - DUNE (Hyper-K)
  - Not covering: g-2, MICE, CTA, SKA, XFEL detectors, smaller experiments, etc.
- Direction of technologies relevant to TDAQ
- Answers to questions
- Discussion

2

### LHC Experiments in the middle of Upgrades!





## **TDAQ** Requirements



### Three major TDAQ challenges:

- Search for rare physics:
  - high rejection or large data collection
- ➡ Face High Luminosity:
  - high frequency to resolve individual bunch crossing fast electronics
  - Iarge detectors with fine granularity to avoid pile-up in the same detector element is high data volume
- Be radiation resistant
- ATLAS/CMS: p-p collisions @70 mb
  - + full Luminosity, high rejection
- + LHCb: p-p collisions
  - reduced Luminosity for rare topologies
- ALICE: heavy-ion collisions ~2000 mb
  - + high energy density



### Pushing the limits







# ATLAS & CMS: complementary approaches

Detector Front-Ends (FE) UXC Trigger Processors Global Trigger Trigger and detector data. ~ 50,000 x 1-10 Gbps GBT links DTH DTH DTH TTC/TTS USC 120 ATCA crates TCDS/EVM 1111 Data to Surface 200m fibers 560x100 Gbs data links 40 32x100 Gbs switches Data to Surface routers Ы пп 500 IO servers  $\bowtie$ ..... 1111 111 SCX ~ 500x200 Gbs switch 100 Tbs bisection bandwidth  $\bowtie$ Event networks 64 32x100 HLT switches Ш 111 ~ 5 PB local storage HLT PC farms/clouds SMTS CDR/T0 Storage 60 GBs access ~ 9.2 MHS06

6

ROYAL HOLLOWAY



	LHC	HL-LHC Phase-2		
CMS detector	Run-2			
Peak $\langle PU \rangle$	60	140	200	
L1 accept rate (maximum)	100 kHz	500 kHz	750 kHz	
Event Size	2.0 MB <sup><i>a</i></sup>	5.7 MB <sup>b</sup>	7.4 MB	
Event Network throughput	1.6 Tb/s	23 Tb/s	44 Tb/s	
Event Network buffer (60 seconds)	12 TB	171 TB	333 TB	
HLT accept rate	1 kHz	5 kHz	7.5 kHz	
HLT computing power <sup>c</sup>	0.5 MHS06	4.5 MHS06	9.2 MHS06	
Storage throughput	2.5 GB/s	31 GB/s	61 GB/s	
Storage capacity needed (1 day)	0.2 PB	2.7 PB	5.3 PB	

 CMS: allow large data flow bandwidth and invest in scalable commercial network and processing systems

CMS-TDR-018

# ATLAS & CMS: complementary approaches

 ATLAS: minimize data flow bandwidth by using multiple trigger levels and regional readout (RoI)



ROYAL HOLLOWAY

ATLAS TDAQ Phase II TDR (publicly out soon!)





### LHCb Run 3: No low-level Trigger!





→ all at reasonable cost: R&D ongoing on network, versatile links





## ILC/CLIC: different beam timing structure





ILC/CLIC DAQ is triggerless, needs to perform zero suppression and undergoes power pulsing

### ILC/CLIC vs LHC/HL-LHC : some comparisons



	ILC	CLIC (380 GeV, 1.4, 3TeV)	LHC (design)	HL-LHC
number of bunches	1312 or 2625	354, 312, 312	2808	2748
bunch spacing	366 ns or 344 ns	0.5 ns	25 ns	25 ns
bunch train length	1 ms	156 ns	N/A	N/A
time between bunch train	199 ms	20 ms	N/A	N/A
bunch train repetition rate	5 Hz	50 Hz	N/A	N/A
collision rate	13 kHz (ave) ~ MHz (peak)	50 Hz	40 MHz	40 MHz
event building rate	13 kHz	50 Hz	100 kHz	1 MHz
detector readout channels	2-5x10^9	3-4x10^9	10^8	7x10^8
max data throughput	~500 Gb/s	~2.4 Tb/s	3 Tb/s	20-40 Tb/s

FCC-ee: no time structure like ILC/CLIC, but similar requirements in terms of detector readout channels, etc.

### SiD and ILD DAQ



Figure II-9.1 Simplified blockdiagram of the SiD detector control and readout chain using the ATCA RCE and CIM modules (defined later in this chapter).



#### ILC TDR Volume 4

### Beam Telescope from AIDA(-2020)

- EUDET-style telescope:
  - Mimosa26 (MAPS)
  - NI FlexRIO system:
    - LVDS front-end
    - FPGA card (Virtex 5)
    - PXIe crate
  - Trigger Logic Unit (TLU)
- Triggerless readout and improvements
  - custom FPGA card to replace NI
  - AIDA-2020 TLU
- Caribou:
  - Xilinx ZC-706 (1/10 Gbit ethernet), FMC, interface board, chip boards, etc.







CLICdp DAQ





Fig. 10.4: Overview of the DAQ scheme.

#### CLIC CDR

### Future very high energy colliders: eg FCC-hh



	LHC (design)	HL-LHC	FCC-hh
Energy	14 TeV	14 TeV	100 TeV
Circumference	26 km	26 km	100 km
Dipole field	8.33 T	8.33 T	16 T
number of bunches	2808	2748	10600 (25 ns) 53000 (5 ns)
bunch spacing	25 ns	25 ns	25 (5) ns
Max Luminosity	3 x 10^34	7.5 x 10^34	1-5 x 10^34
collision rate	40 MHz	40 MHz	40 MHz (200 MHz)
event building rate	100 kHz	1 MHz	what can we achieve?
detector readout channels	10^8	7x10^8	?
max data throughput	3 Tb/s	20-40 Tb/s	need 10k Tb/s?
Peak Pile up	27 (hahaha!)	200	171 (34)

### FCC-hh: 100 TeV simulations



### Using CMS simulation



- Collecting EWK & Higgs physics via singleobject triggers is going to be challenging
  - Improvements to E/G algorithms and muon resolution will be needed



Bologna, Brooke, Newbold, Sphicas, FCC week 2018 Amsterdam

### Now for something a bit different... DUNE (Hyper-K)



## DUNE

- Extremely varied physics program
  - Neutrino beam -> external trigger possible
  - Supernova explosion -> very late trigger
  - Proton decay, atmospheric & solar neutrino measurements -> local and rare signature
- Challenge for the Trigger and DAQ system:
  - Fit very different requirements
- TPC sampled at 2 MHz continuous readout, photon detectors sampled at 150 MHz (local triggering)
  - Signal for a particle forming over msecs
  - Downstream TDAQ elements decide when anything interesting happened inside the active volume
  - Combination over time windows of thresholds, tracking, distributed activity signatures, ...



EP-DT Detector Technologies

Enrico Gamberini

ISOTDAQ 2018, Vienna 22/02/2018

### DUNE (Hyper-K)



- For 10kT, plan on 150 Anode Plane Assemblies (APAs) -> 9 Tbps over 12k links
  - > 10 PB/year in first year Ο
- All data CAN be streamed out of the detector... so why not do it?!



### **Possible DUNE TDAQ**

- Readout with very large buffer to account for long LO/L1 latency (tens
- integrated into readout (or carried

ISOTDAQ 2018, Vienna 22/02/2018

#### How to buffer ~10 Tb/s for 10 s !?

### Trying things out this Summer...



## **ProtoDUNE SP TDAQ environment**

- 6 Anode Plane Assemblies (APA)
  - TPC ~ 430 Gb/s (continuous readout; 15360 ch @ 2MHz)
  - Photon Detectors ~ 1 Gb/s (locally triggered)
- SPS super cycle structure: 2 x 4.8 s bursts in 48 s
  - Full readout -> ~85 Gb/s
  - Too much for DAQ as well as for storage and offline!
- Introduction of a simple global trigger to mitigate data flow
  - Retain full readout off detector
  - Cannot rely on triggering on TPC signatures, because there is too much activity from cosmic rays.
- Lossless data compression to reduce event size
- 5 APAs will be readout via ATCA boards (12800 ch), 1 APA (2560 ch) via FELIX
  - 2 firmware variants in front-end electronics
  - API for transparently treating data at offline software level

# Summary of the future experiments tour



- LHC experiments TDAQ performed very well!
- Started building Phase I and Phase II TDAQ Upgrades (PU=200!)
  - similar philosophies to current LHC
  - Physics needs require same (or lower) trigger pT threshold compared to today:
    - high trigger rates controlled by use of hardware tracker trigger
  - LHCb pioneering full readout for Run 3 (for their small event size...)
- ILC/CLIC (FCC-ee)
  - At face value very feasible compared to LHC, but high peak rates, large number of channels and power pulsing might prove to be tricky
- FCC-hh
  - large rates!! large data throughput and 5 ns operation sounds tricky (porting LHC or HL-LHC methods to FCC-hh implies very large pT threshold, ok with that?)
  - Reminder: for discovery (not precision) ok with large pT thresholds and prescaled triggers
- Dune (Hyper-K)
  - challenging parameters and need for versatile system

### Technology trends to help us accomplish this

- Trigger & DAQ components:
  - Readout links/buffers
  - Timing
  - Processors
  - Protocols
  - Switching networks



A DAQ system



## Frontend readout

Pixel readout: RD53 collaboration



CERN-RD53-PUB-17-001

Version 3.21, February 7, 2018

#### The RD53A Integrated Circuit

ABSTRACT: Implementation details for the RD53A pixel readout integrated circuit designed by the RD53 Collaboration. This is a companion to the specifications document and will eventually become a reference for chip users. RD53A is not intended to be a final production IC for use in an experiment, and contains design variations for testing purposes, making the pixel matrix non-uniform. The chip size is 20.0 mm by 11.8 mm.





RD-53 will develop the tools and designs needed to produce the next generation of pixel readout chips needed by <u>ATLAS</u> and <u>CMS</u> at the <u>HL-LHC</u>. There is also interest and participation by <u>CLIC</u>. More details can be found in the <u>collaboration proposal</u>.



Simulation WG pages RD53 Wiki CDS Internal Submit document to CDS (instructions)

### Frontend readout: Optical links

### Example: Versatile Link

- 3.2 Gbit/s user bandwidth; in uplink 4.48 Gbits/s
- Optional FEC
- SFP-like form factor
- Deterministic latency in both directions
- Radiation hard qualified for:
  - 1 MRad total dose
  - 5x1014 neq/cm<sup>2</sup>
- FE interface: 10 to 40 E-links: SLVS based with 320, 160 or 80 Mbit/s
- "Low"-power: <1.5W, 2.2W Max
  - 500mW version under design



Courtesy: Paulo Moreira and Versatile link team https://espace.cern.ch/project-versatile-link/public/default.aspx

### GBT architecture

Experiment control (SC/DCS/ECS)

Modest bandwidth (bidirectional link)



https://espace.cern.ch/GBT-Project/default.aspx

#### P Durante ISOTDAQ 2018



### Frontend readout: Optical links



# **Trends For Next Generation**

- Higher speed using advanced modulation formats
  - PAM4 for 56G electrical and 110G optical
  - Matches FPGA, Ethernet switches, and CPU evolution
- Power consumption goes up
  - More equalization electronics, CDR, PAM4 circuitry
  - Effort under way to bring it back down to around a few pJ/W for close to chip on-board
- BER goes up
  - Standards have very loose BER (10e-5 at 28G), requires strong FEC
  - Currently BER < 10e-12 at 28G, BER at 110G still unknown</li>
- Silicon Photonics integration
  - Higher speeds, single mode, but higher power consumption







56GBaud Optical Out



#### P Durante ISOTDAQ 2018





## **Buffering?**

- How to buffer ~10 Tb/s for 10 s !?
- Development of a KeyValue storage system based on new Intel<sup>®</sup> memory technology:

https://indico.cern.ch/event/669648/contributions/2802031/attachments/1581153/2499892/fogKV.pdf

- Decouple real time data acquisition from asynchronous event selection:
  - Large, temporary storage of O(100) PB
  - High throughput of O(10) TB/s
- Fits DUNE long term needs:
  - O(100) TB storage
  - O(10) TB/s throughput



E Gamberini ISOTDAQ 2018

## Timing systems



# Solutions for HL-LHC



#### • New custom ASIC: submission in March

- First proto tape out: Q3 2018
- Lower Power 500mW/750mW (5.12/10.24Gbps)
- Higher radiation hardness TID 200 Mrad
- Lower jitter <5ps rms
- Higher upstream bandwidth (10.24Gbps)
- ...and much more in the specs!

#### https://espace.cern.ch/GBT-Project/LpGBT/Specifications/LpGbtxSpecifications.pdf

#### White Rabbit (Backbone for LS3?)

#### 

#### Inovative concept

- Self synchronous...but not to the Bunch Clock!
- Standard Ethernet network
- Future part of PTP standard
- IEEE1588-2018 (High Accuracy)
- High accuracy synchronisation to the GPS time
  - Precise GPS distribution
- Precise round trip measurement & compensation
  - Wander ~0, even over 10km
- Bounded and low-latency Control Data

... not enough to distribute the Bunch Clock! => An additional layer is needed



#### S Baron ISOTDAQ 2018









worlds by analysing which strengths of FPGA, GPU and CPU best fit the different demands of the application.





0+0

Evolution in programming paradigms, tools and libraries

5

10

15

20

concurrency

25

30

Exploiting HW is more complicated (vectors, memory sharing...)

#### F Pastore ISOTDAQ 2018

thread memory

35

40

45



- Use Case example: Pattern recognition (tracks) in hardware
  - GPU: ALICE uses Cellular Automaton and Kalman filtering for their TPC tracking: 10 times faster than CPU
  - FPGA: LHCb studying the Retina approach

#### **Retina prototype**

- LHC-b moving to a trigger less design
  - Event processing at 40 MHz
  - FPGA based tracker before Event Builder can help to make online tracking affortable



Associative Memories

						Mito Ista Arto Ista Arto		
	Vers.	Design	Tech.	Area	Patterns	Package	]	
	1	Full custom	700 nm		128	QFP	SVT @CFD	
	2	FPGA	350 nm		128	QFP		
Insued	3	Std cells	180 nm	100 mm <sup>2</sup>	5 k	QFP	SVT upgrade	
	4	Std cells $+$	65 nm	$14 \text{ mm}^2$	8 k	OEP		
XILINX* //2046* 1154/F5025 115974		Full custom	00 1111	14 11111	UK	- UKI I		
	mini-5	Std cells +		$4 \text{ mm}^2$	0.5 k	QFP		
		Full custom	65 nm					
	5	+ IP blocks		12 mm²	3 k	BGA		
	6	Std cells +						
		Full custom	65 nm	168 mm <sup>2</sup>	128 k	BGA	FIK@Atlas	
		+ IP blocks						
	7	Std cells +	28 nm	10 mm <sup>2</sup>	16 k	BGA,		
		Full custom	20 1111	10 1111	10 %	SiP		

#### **AM evolution**

### A Negri ISOTDAQ 2018

## Processors: GPUs





Theoretical Peak Floating Point Operations per Watt, Single Precision

G Lamanna ISOTDAQ 2018

FPGAs



## Major Manufacturers

- Xilinx
  - First company to produce FPGAs in 1985
  - About 55% market share, today
  - SRAM based CMOS devices
- Intel FPGA (formerly Altera)
  - About 35% market share
  - SRAM based CMOS devices
- Microsemi (Actel)
  - Anti-fuse FPGAs
  - Flash based FPGAs
  - Mixed Signal
- Lattice Semiconductor
  - SRAM based with integrated Flash PROM
  - low power











# Ever-decreasing feature size





## System-On-a-Chip (SoC) FPGAs





## FPGAs in Server Processors and the Cloud

- Since 2016: Intel Xeon Server Processor with FPGA in socket
  - Intel acquired Altera in 2015



- FPGAs in the cloud
  - Amazon Elastic Cloud F1 instances
    - 8 CPUs / 1 Xlinix UltraScale+ FPGA
    - 64 CPUs / 8 Xlinix UltraScale+ FPGA



## PCIe example: ATLAS FELIX

- 2016
- ≤ 48 duplex optical links
- XilinX Ultrascale FPGA
- 2x DDR4 SO-DIMM
- PCle 3.0 x16
- Wupper DMA (Open Source!)



14/02/2017

ISOTDAQ 2018 - Introduction to PCIe

P Durante ISOTDAQ 2018





# Example: Gen3 x8, 256 Bytes MPS • $\rho = 64 \times 0.98 \times \frac{256}{256+24} = 62.7 \times 0.91 = 57 \text{ Gb/s}$

P Durante ISOTDAQ 2018



# PCle Gen4 – On Silicon

### **Mellanox ConnectX®-5**

<section-header>

LnkCap: Port #0, Speed 16GT/s, Width x16, ASPM L0s L1

### **IBM Power AC922 (2018?)**

- 2 POWER9 Processors
- 190, 250W modules
- 4-6 NVidia "Volta" GPU's
- 300W, SXM2 Form Factor, NVLink 2.0
- 6 GPU configuration, water cooled
- 4 GPU configuration, air or water cooled
- 2 Gen4 x16 HHHL PCIe, CAPI enabled
- 1 Gen4 x4 HHHL PCIe
- 1 Gen4 Shared x8 PCIe adapter
- 16 IS DIMM' s
- 8, 16, 32, 64, 128GB DIMMs
- 2 SATA SFF HDD / SSD
- 2 2200W power supplies
- 200 VAC, 277VAC, 400VDC input
- N+1 Redundant
- Second generation BMC Support Structure
- Pluggable NVMe storage adapter option

# Switching networks



CMS

100

80

70

60

30

20

10

0

Others

## Run 1: 100 GB/s network

### Myrinet widely used when DAQ-1 was designed

- high throughput, low overhead
- direct access to OS
- flow control included
- new generation can suppost ⁵
   10GBE ₄₀

## Run2: 200 GB/s network

- ⇒ 2MB/event
- Technology allows single EB network (56 Gbps FDR Infiniband)
- Myrinet —>10/40 Gbps Ethernet



Myrinet

SP Switch

# Switching networks



# Going beyond TCP/IP

#### High Performance Computing

- HPC technologies: very high throughput and very low latency within data centers
- Standard: Infiniband, implemented mostly by Mellanox and Intel
- Replacement technologies for layers 1 to 4 at least
  - TCP is not suitable for intra-data center communications (timeout too long)
  - IP is often not needed
  - Ethernet has too much overhead
- RDMA, remote direct memory access:
  - Network packets are written directly into host memory
  - Minimal latency, no OS overhead

### F Le Goff ISOTDAQ 2018

## Questions from Input committee



- I. What are potential developments in this field?
  - see previous slides
  - reminder from A Tapper: Machine Learning in Trigger systems (NN in μs!)
  - reminder from M Wing: UK very active in DAQ for smaller experiments, developments and synergies happening there as well
- 2. What consensus / conflicts (on what should be done in longer term european particle physics) are there in this area?
  - Commercial vs custom components
  - Firmware done by engineers vs physicists/PhDs... (issues of design, maintenance, etc.)
  - 2 main future strategies:
    - Process data on-detector and move all of it without trigger to offline processing
    - Implement sophisticated multi-layer trigger algorithms using fast hardware components
- 3. What are experimental possibilities to do that? Are different scenarios already envisaged?
  - As shown in previous slides, some options currently being studied and looked into
  - Remember that detectors including TDAQ systems need a lot of R&D and long lead time

## Questions from Input committee



- 4. What are the choices for the strategy? What can the UK agree to input?
  - Given the future experiments TDAQ challenges:
    - I. more collaboration between projects is needed!
      - European national labs could host week-long TDAQ specific conferences
      - UK-centric: organize IOP-like UK TDAQ workshops
    - 2. more collaboration with industry is needed!
      - European national labs should help university groups with industry contacts
      - UK-centric: make use of ISCF
    - 3. more training of PhD students in this area is needed
      - UK-centric: need a CDT in detector/TDAQ technologies!
    - 4. CERN RD Collaboration useful? Extension of OpenLab?
    - Others:
- Discussion:



# Back-ups



## ILC machine parameters from TDR



			Baseline 500 GeV Machine		1st Stage L Upgrade		$E_{\rm CM}$ (	$E_{CM}$ Upgrade	
								A	В
Centre-of-mass energy	$E_{\rm CM}$	GeV	250	350	500	250	500	1000	1000
Collision rate	$f_{\rm rep}$	Hz	5	5	5	5	5	4	4
Electron linac rate	flinac	Hz	10	5	5	10	5	4	4
Number of bunches	nb		1312	1312	1312	1312	2625	2450	2450
Bunch population	N	×10 <sup>10</sup>	2.0	2.0	2.0	2.0	2.0	1.74	1.74
Bunch separation	$\Delta t_{\rm b}$	ns	554	554	554	554	366	366	366
Pulse current	$I_{\rm beam}$	mA	5.8	5.8	5.8	5.8	8.8	7.6	7.6
Main linac average gradient	$G_{*}$	MV m <sup>-1</sup>	14.7	21.4	31.5	31.5	31.5	38.2	39.2
Average total beam power	Pheam	MW	5.9	7.3	10.5	5.9	21.0	27.2	27.2
Estimated AC power	$P_{\rm AC}$	MW	122	121	163	129	204	300	300
RMS bunch length	$\sigma_{2}$	mm	0.3	0.3	0.3	0.3	0.3	0.250	0.225
Electron RMS energy spread	$\Delta p/p$	%	0.190	0.158	0.124	0.190	0.124	0.083	0.085
Positron RMS energy spread	$\frac{\Delta p}{p}$	%	0.152	0.100	0.070	0.152	0.070	0.043	0.047
Electron polarisation	$\overline{P}_{-}^{r/r}$	%	80	80	80	80	80	80	80
Positron polarisation	$P_+$	%	30	30	30	30	30	20	20
Horizontal emittance	262	um	10	10	10	10	10	10	10
Vertical emittance	Yey	nm	35	35	35	35	35	30	30
IP horizontal beta function	8*	mm	13.0	16.0	11.0	13.0	11.0	22.6	11.0
IP vertical beta function	$\beta_{y}^{*}$	mm	0.41	0.34	0.48	0.41	0.48	0.25	0.23
IP RMS horizontal beam size	<i>a</i> *	nm	729.0	683.5	474	729	474	481	335
IP RMS veritcal beam size	$\sigma_y^*$	nm	7.7	5.9	5.9	7.7	5.9	2.8	2.7
Luminosity	L	$\times 10^{34} \mathrm{cm}^{-2} \mathrm{s}^{-1}$	0.75	1.0	1.8	0.75	3.6	3.6	4.9
Fraction of luminosity in top 1%	$L_{0.01}/L$		87.1%	77.4%	58.3%	87.1%	58.3%	59.2%	44.5%
Average energy loss	δns		0.97%	1.9%	4.5%	0.97%	4.5%	5.6%	10.5%
Number of pairs per bunch crossing	Nonire	×10 <sup>3</sup>	62.4	93.6	139.0	62.4	139.0	200.5	382.6
Total pair energy per bunch crossing	Enairs	TeV	46.5	115.0	344.1	46.5	344.1	1338.0	3441.0



## CALICE DAQ: Architecture





Detector Unit: 1 layer of a Calo module (30-50 layers) 1 LDA = 10 DIFs 1 ODR = 4 LDAs



## CALICE DAQ: Architecture





Cambridge

Detector Unit: Sensors & ASICs DIF: Detector InterFace -connects generic DAQ and services



Link/Data Aggregator – fanout/in DIFs & drive link to ODR

Manchester



Clock & Control Card: Fanout to ODRs (or LDAs)

UCL



**ODR:** Off Detector Receiver – PC interface for system

## CALICE DAQ: Performance



- DIF-LDA link:
  - theoretical limit: 40 Mbits/s measured: 28 Mbits/s (40% higher than worst-case scenario of detectors)
  - in practice: 20 Mbits/s: ASICS organised in 4 parallel daisy-chains, each running at a 5 MHz clock
- LDA-ODR link:
  - data rate of 28 x 10 = 280 Mbits/s << link speed of 1 Gbit/s</p>
  - 4 x 200 Mbits/s = 800 Mbits/s received by each ODR
- ODR writing to memory or RAID:
  - 2 ODRs in 1 DAQ PC = 200 MBytes/s
  - measured ODR writing to RAM =
    - I PCIExpress lane: 310 MBytes/s (constant with data size)
    - 2: doubles, 3: increase by 100MBytes 4: no gain → max transfer rate: 700 MBytes/s
  - measured ODR writing to disk using scatter-gather: 280 MBytes/s > 200 MBytes/s



## CALICE DAQ: Lessons

ROYAL HOLLOWAY UNIVERSITY Or LONGON

- PCs are cheap but unreliable: use TCA crates
- since used FPGAs and PCIExpress → easily port to TCA
- using commercial components is good but not commercial boards
  - don't necessarily contain all the functionality
  - experienced different performance from advertised
  - introduces a middle-man



## ODR (RHUL)





- Receives module data from LDA
  - PCI-Express card, hosted in PC.
  - 4 links/card, 1-3 cards/PC
  - Buffers and transfers to store as fast as possible
- Fibre optic link to detector via SFP modules (std networking hw)
  - Currently GigE (1.25Gb), but could be higher and use different proto.
- Sends controls and config to LDA for distribution to DIFs
- Interfaces to CCC for synchro running
  - Goal to send clock and prompt controls over optic link too
  - Reset and reprog FPGAs



#### Hardware:

- Using commercial FPGA dev-board:
  - PLDA XPressFX100
  - Xilinx Virtex 4, 8xPCIe, 2x SFP (or 3 with expansion board)
  - Early cards are faulty, investigation with supplier ongoing
- Our own firmware and Linux driver software

# CCC (UCL)





- CCC unit provides machine clock and fast signals to 8x ODR/LDA.
- Logic control (FPGA, connected via USB)
  - Command encoders
  - Remote signal enable, clock selection
  - But capable of stand-alone, dumb mode
- Provision for async scintillator type signals (VFast)
- LDA provides next stage fanout to DIFs
  Eg CCC unit -> 8 LDAs -> 10 DIFs = 80 DUs.
- Signalling over same HDMI type cabling
- Facility to generate optical link clock (~125-250MHz from ~50MHz machine clock)
- Commercial systems are not ideal here.
  - Looking at custom protocol on fibre optic link
  - Prompt signals and low jitter clock recovery needs further investigation

## **Current activities: AIDA**





Advanced European Infrastructures for Detectors at Accelerators

- EU FP 7 AIDA: Advanced Infrastructures for Detectors and Accelerators
  - http://aida.web.cern.ch/aida/index.html
  - started Feb 1st 2011 for 4 years
  - 80 institutes and labs from 23 EU countries
  - 8m € from EU and 26m € in total

■ It aims to upgrade, improve and integrate key European research infrastructures and develop advanced detector technologies for future particle accelerators (LHC upgrade, Linear Colliders, Neutrino facilities and Super-B factories) in line with the <u>European</u> <u>Strategy for Particle Physics</u>.

#### coordinated by CERN

52

 UK institutes: QMUL, RHUL, STFC, UCAM, UNIGLA, UNILIV, UNIBRIS, UOXF, USFD

### Switching networks



## The OSI Model

The ISO's (International Organization for Standardization) project OSI (Open Systems Interconnection) has defined a **conceptual model** (ISO/IEC 7498-1) that provides a common basis for coordination of standards development for the purpose of **systems interconnection**.

- Defines **7 layers** that splits responsibilities and functionalities of networking communication
- Layer interfaces allow actors of the "industry" to develop functionalities independently
- It's a framework not an actual implementation nor a strict guide
- Most network technologies reflect this layered structure





# SILICON PHOTONICS

- Traditional VCSEL (Vertical Cavity Surface Emitting Lasers)
  - Maximum NRZ rate is about 28Gb/s
  - can achieve 56Gb/s using PAM4 (28 GBaud)
  - Limited distance (<100m, multimode fiber) and BER
  - But: lowest power consumption, lowest packaging cost
- Silicon Photonics
  - Integrated optical components on a Silicon Wafer, using silicon manufacturing technology
  - Faster modulators: 56Gb/s and 112Gb/s (56 Gbaud)
  - Longer distance (500m to 2 km), due to single mode fiber
  - Higher power consumption, higher packaging cost
- SiPho integration gives a path to 100Gb/s and much beyond
  - Use the integration and WDM (Wavelength Division Multiplexing)
  - We are working to reduce power consumption in line (~2.5W at 800 Gb/s)
    - Assumes on-board optics and close proximity to the FPGA



Courtesy LETI





Courtesy Macom