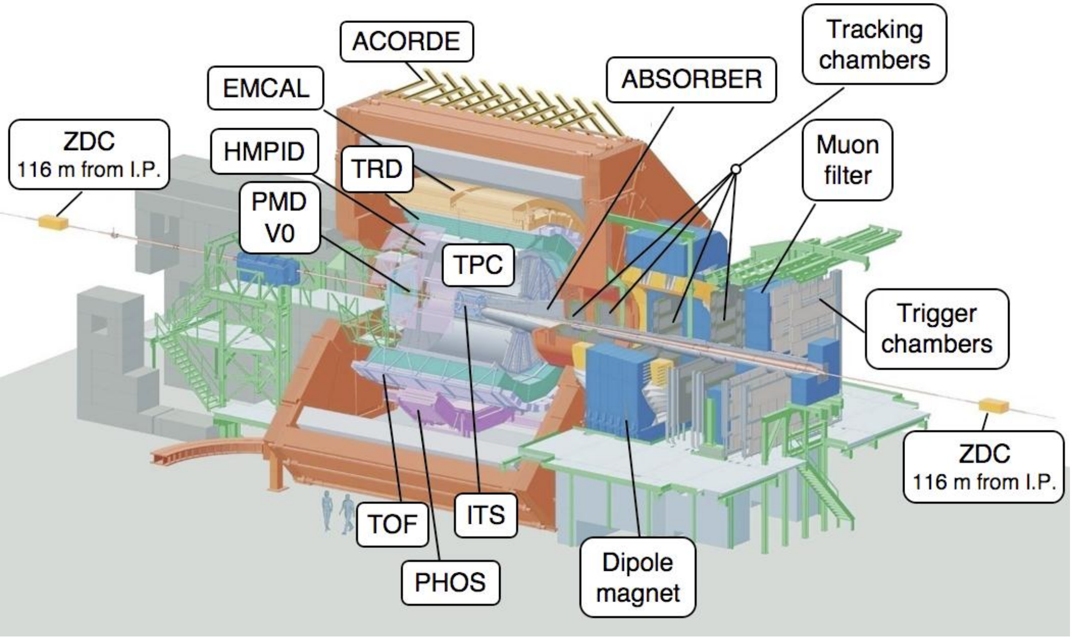
A Large Ion Collider Experiment

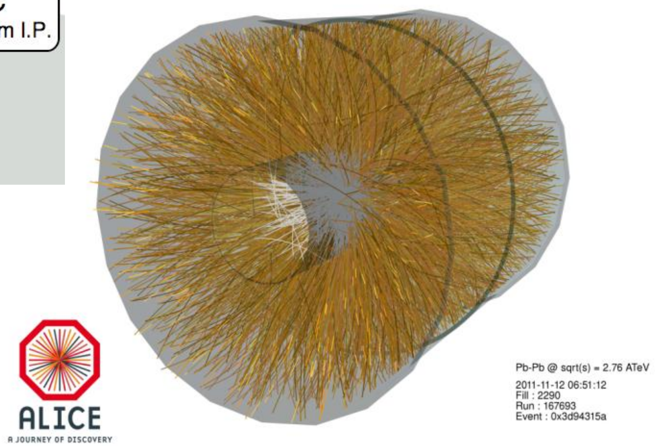
DAQ@LHC Workshop

ALICE Upgrade Overview

Thorsten Kollegger
for the ALICE Collaboration

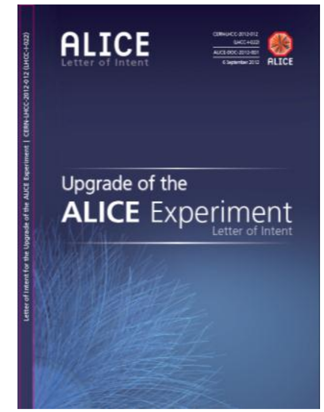
A Large Ion Collider Experiment





ALICE Upgrade Overview

Planned for 2018 (LHC 2nd Long Shutdown)
(“Upgrade of the ALICE Experiment”, LoI, CERN-LHCC-2012-12)



Physics goals

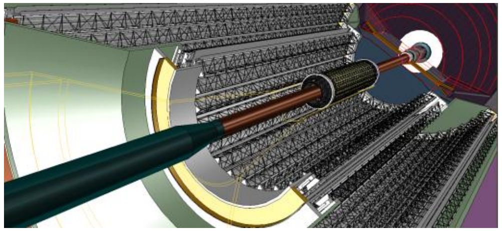
High precision measurements of rare probes at low $p_{T,}$
which cannot be selected with a trigger,
require a large sample of events recorded on tape

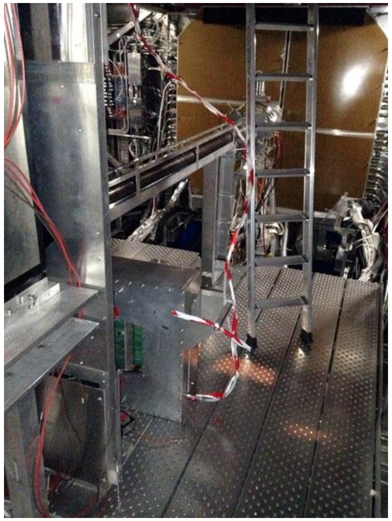
Target

- Pb-Pb recorded luminosity $\geq 10\ nb^{-1}$ ➡ 8×10^{10} events
- pp (@5.5 Tev) recorded luminosity $\geq 6\ pb^{-1}$ ➡ 1.4×10^{11} events

…and significant improvement of vertexing and tracking capabilities

ALICE Upgrade Overview





The upgrade plans entails building

- New, high-resolution, low-material ITS
- Upgrade of TPC with replacement of MWPCs with GEMs and new pipelined readout electronics
- Upgrade of readout electronics of: TRD, TOF, PHOS and Muon Spectrometer
- Upgrade of the forward trigger detectors and ZDC
- **Upgrade of the online systems**
- **Upgrade of the offline reconstruction and analysis framework and code**

Requirements

Sample full 50kHz Pb-Pb interaction rate
(current limit at ~500Hz, factor 100 increase)

Typical event size of PbPb collisions@5.5TeV: 22 MByte
⇨ ~1.1 TByte/s detector readout
⇨ ~500 PByte/HI period (1 month)

However:

- storage bandwidth limited to ~20 GByte/s
- many physics probes have low S/B:
  classical trigger/event filter approach not efficient

… and all this data has to be reconstructed

Strategy

Data reduction by online reconstruction

Store only reconstruction results, discard raw data
- Demonstrated with TPC data since Pb-Pb 2011
- Optimized data structures for lossless compression
- Algorithms designed to allow for “offline” reconstruction passes with improved calibrations

⇨ Implies much tighter coupling between online and offline computing systems

ALICE

Data Bandwidth

Detector	Input to Online System (GByte/s)	Peak Output to Local Data Storage (GByte/s)	Avg. Output to Computing Center (GByte/s)
TPC	1000	50.0	8.0
TRD	81.5	10.0	1.6
ITS	40	10.0	1.6
Others	25	12.5	2.0
Total	**1146.5**	**82.5**	**13.2**

LHC luminosity variation during fill and efficiency taken into account for average output to computing center.

TPC Data Reduction

	Data Format	Data Reduction Factor	Event Size (MByte)
	Raw Data	1	700
FEE	Zero Suppression	35	20
HLT	Clustering & Compression	5-7	~3
	Remove clusters not associated to relevant tracks	2	1.5
	Data format optimization	2-3	<1

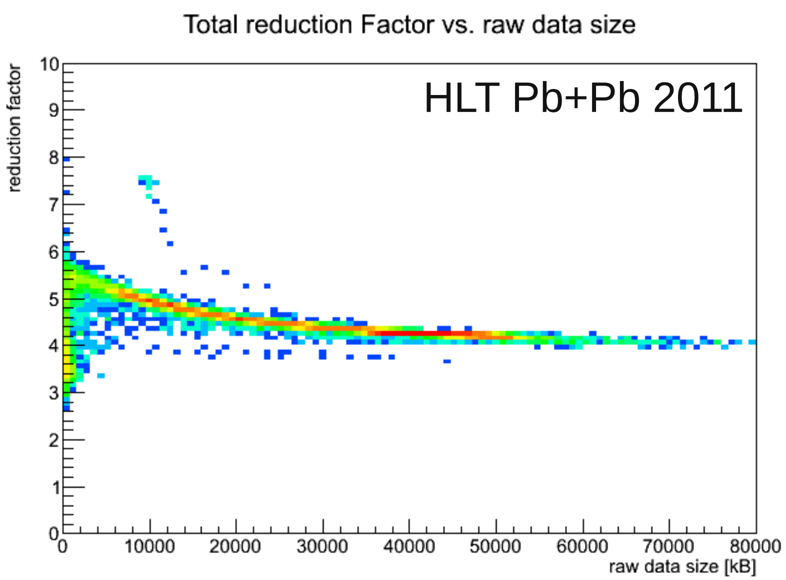
First steps up to clustering on the FPGA of the detector link receiver
Further steps require full event reconstruction, pattern recognition requires only coarse online calibration

TPC Data Reduction

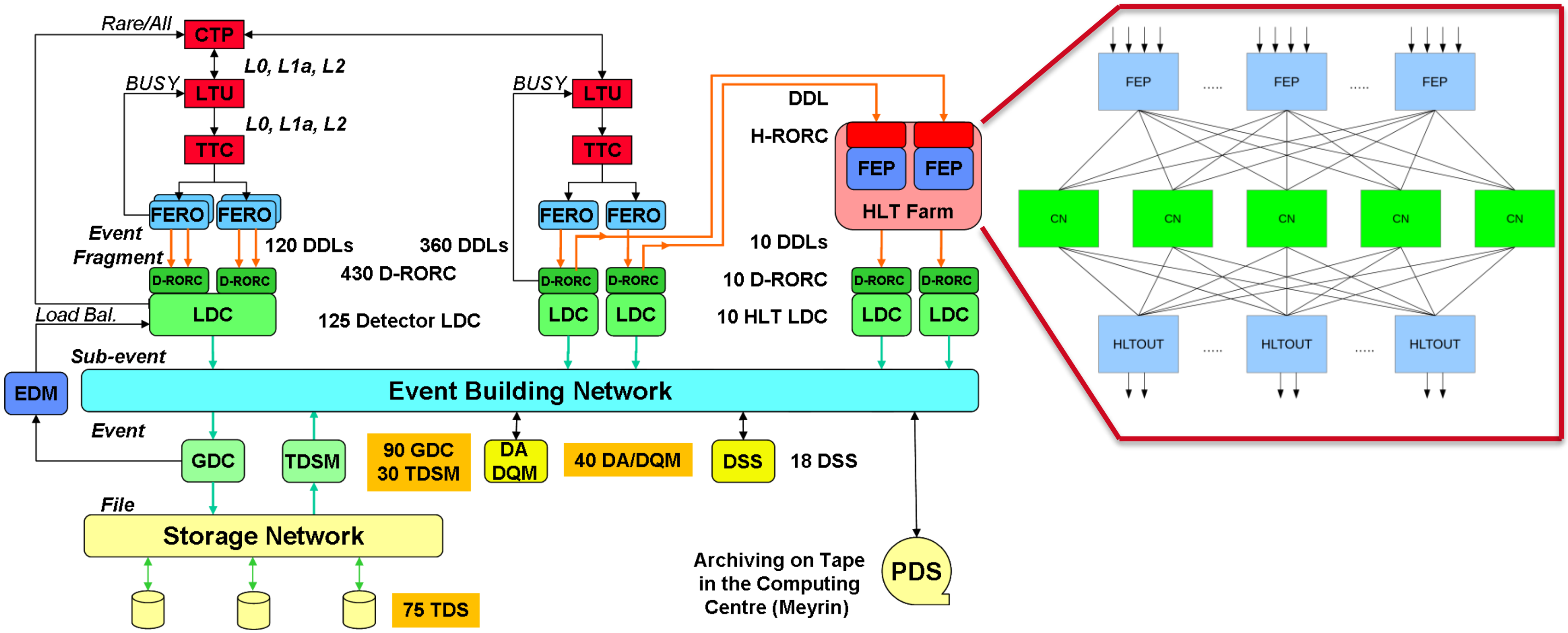
First compression steps used in production starting with the 2011 Pb+Pb run

Online found TPC clusters are basis for offline reconstruction

Currently R&D towards using online found TPC tracks to complement offline seed finding and online calibration

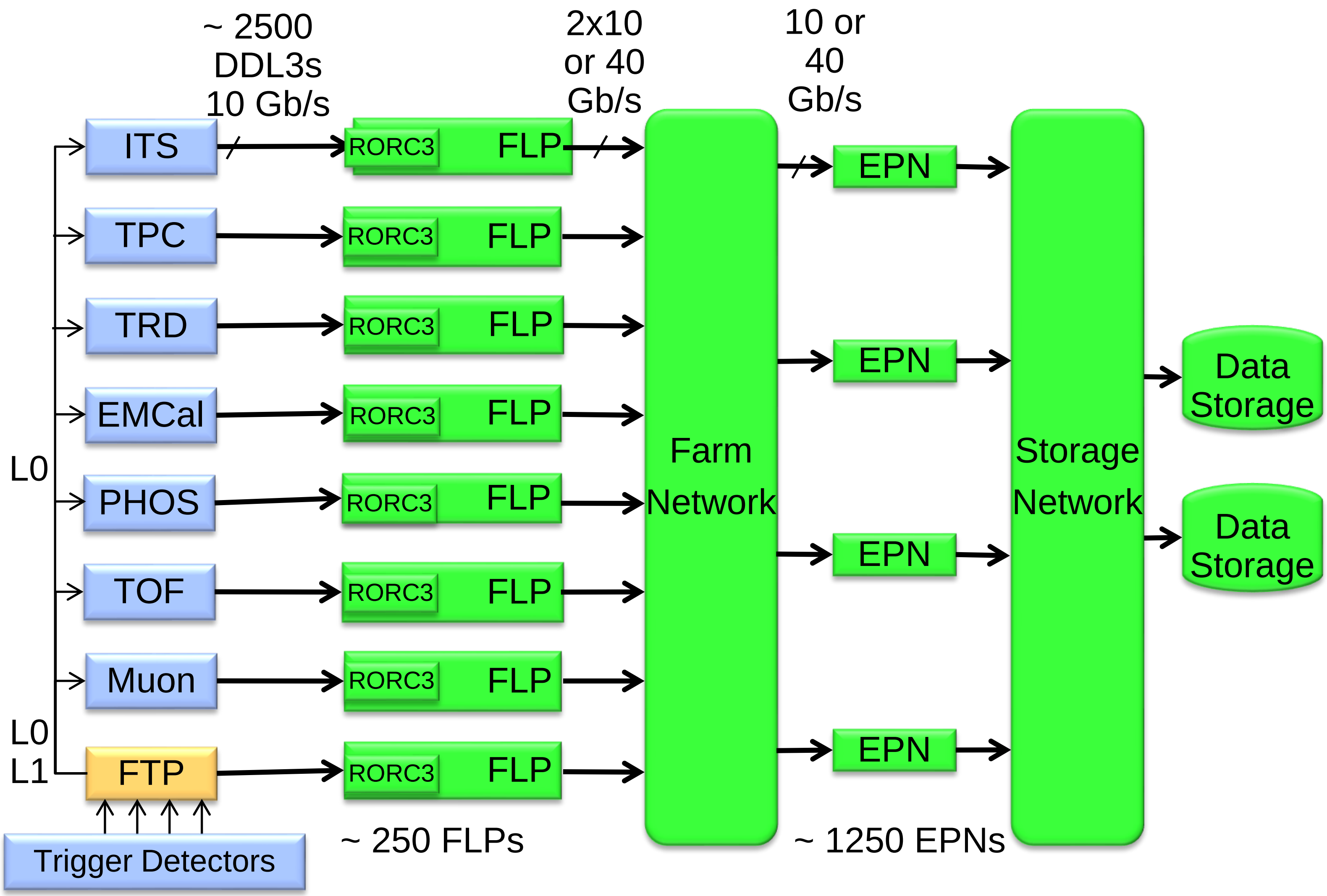


Current Online Systems



Different technologies/techniques used in DAQ/HLT
e.g. Ethernet <-> Infiniband

Combined DAQ/HLT System



Detector Readout

Combination of continuous and triggered readout

Continuous readout for TPC and ITS

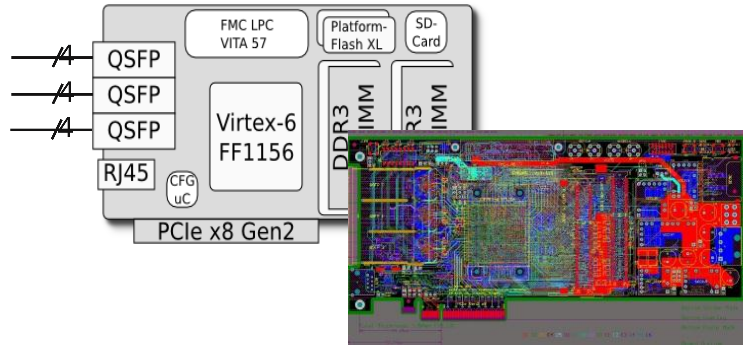
- At 50 kHz, ~5 events in TPC during drift time of 92 µs
  Continuous readout minimizes needed bandwidth
- Implies event building only after partial reconstruction

Fast Trigger Processor (FTP) complementing CTP

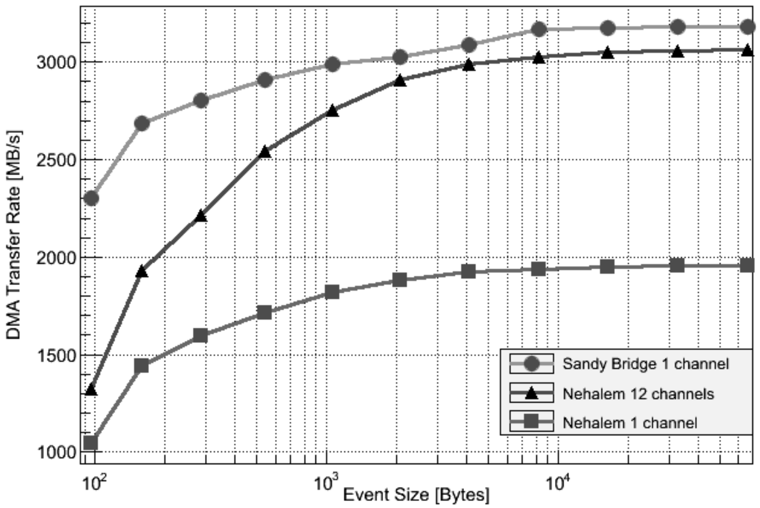
- Provides clock/L0/L1 to triggered detectors and TPC/ITS for data tagging and test purposes

DDL/RORC Development

Data Link
DDL1 (now): 2Gbit/s
DDL2 (LS1): 6 Gbit/s
DDL3 (LS2): 10 Gbit/s



Receiver Card (FPGA)
RORC1 (now)
- 2 DDL1, PCI-X&PCIe Gen1x4
RORC2 (LS1 HLT)
-12 DDL2, PCIe Gen2x8
RORC3 (LS2)
- 10-12 DDL3, PCIe Gen3



ALICE

Network

Requirements

Total number of nodes: ~1500
FLP Node Output: up to 12 Gbit/s
EPN Node Input: up to 7.2 Gbit/s
EPN Output: up to 0.5 Gbit/s

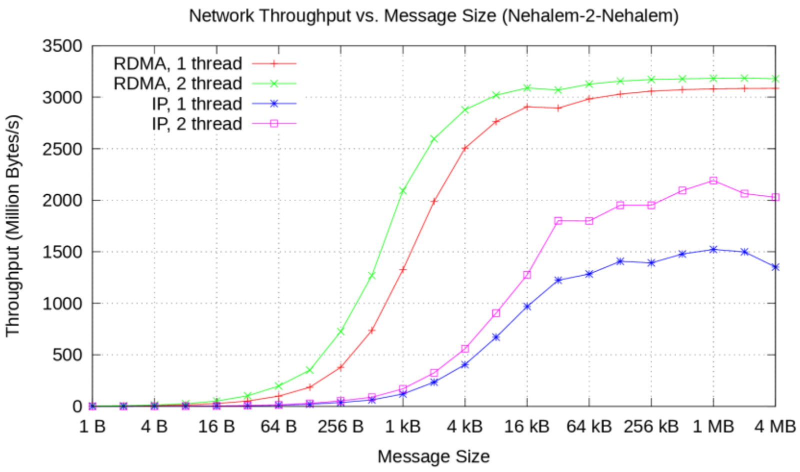
Two technologies available

- 10/100Gbit Ethernet (currently used in DAQ)
- QDR/FDR Infiniband (40/52Gbit, used in HLT)

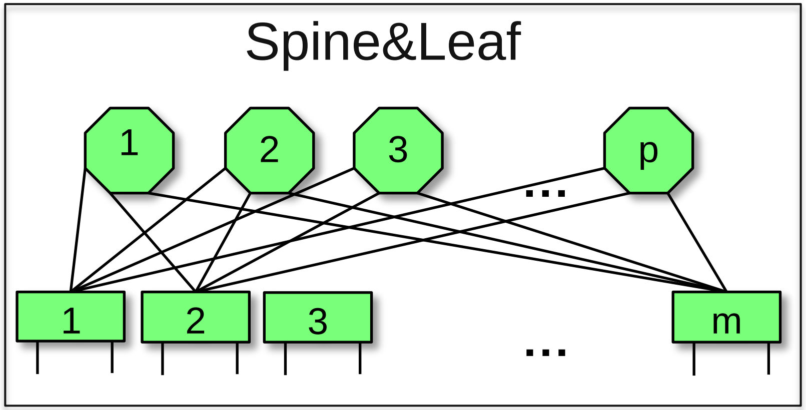
Both would allow to construct a network satisfying the requirements even today

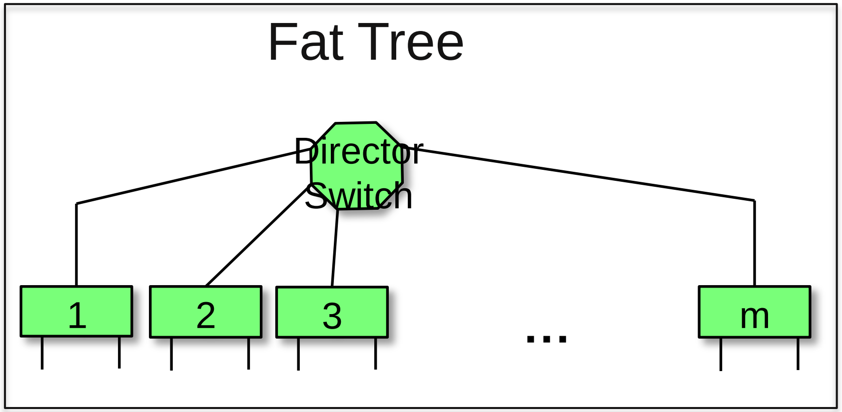
Network

Throughput tests



Different topologies under study to
- minimize cost
- optimize failure tolerance
- cabling





Processing Power

Estimate for online systems based on current HLT processing power

- ~2500 cores distributed over 200 nodes
- 108 FPGAs on H-RORCs for cluster finding
  1 FPGA equivalent to ~80 CPU cores
- 64 GPGPUs for tracking (NVIDIA GTX480 + GTX580)

Scaling to 50 kHz rate to estimate requirements

- ~ 250.000 cores
- additional processing power by FPGAs + GPGPUs

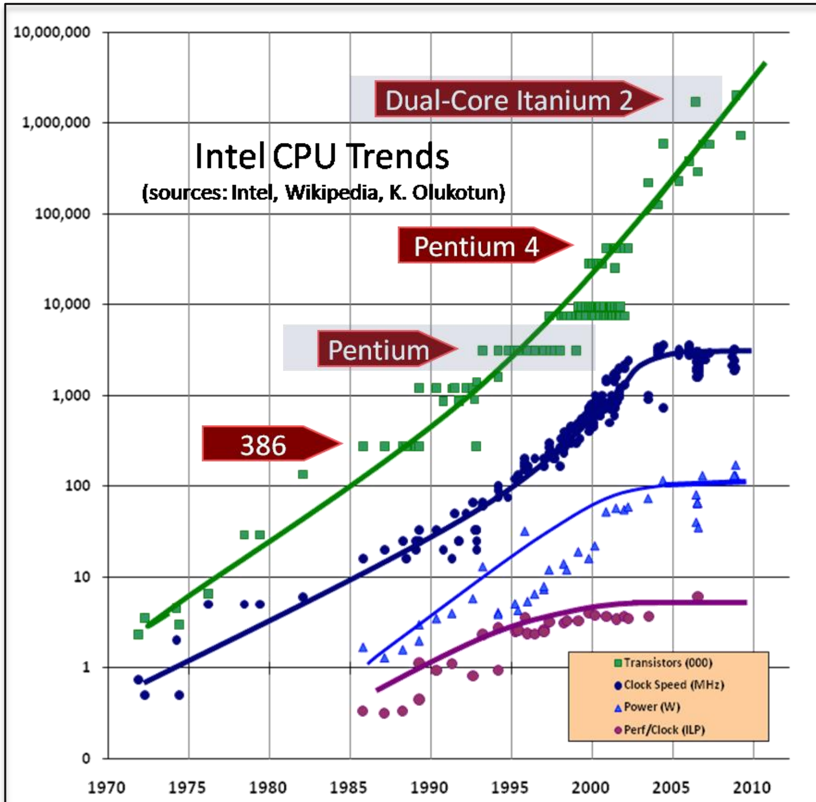
⇨1250-1500 nodes in 2018 with multicores

Processing Power

Estimate of processing power based on scaling by Moore's law

However: no increase in single core clock speed, instead multi/many-core

Reconstruction software needs to adapt to full use resources

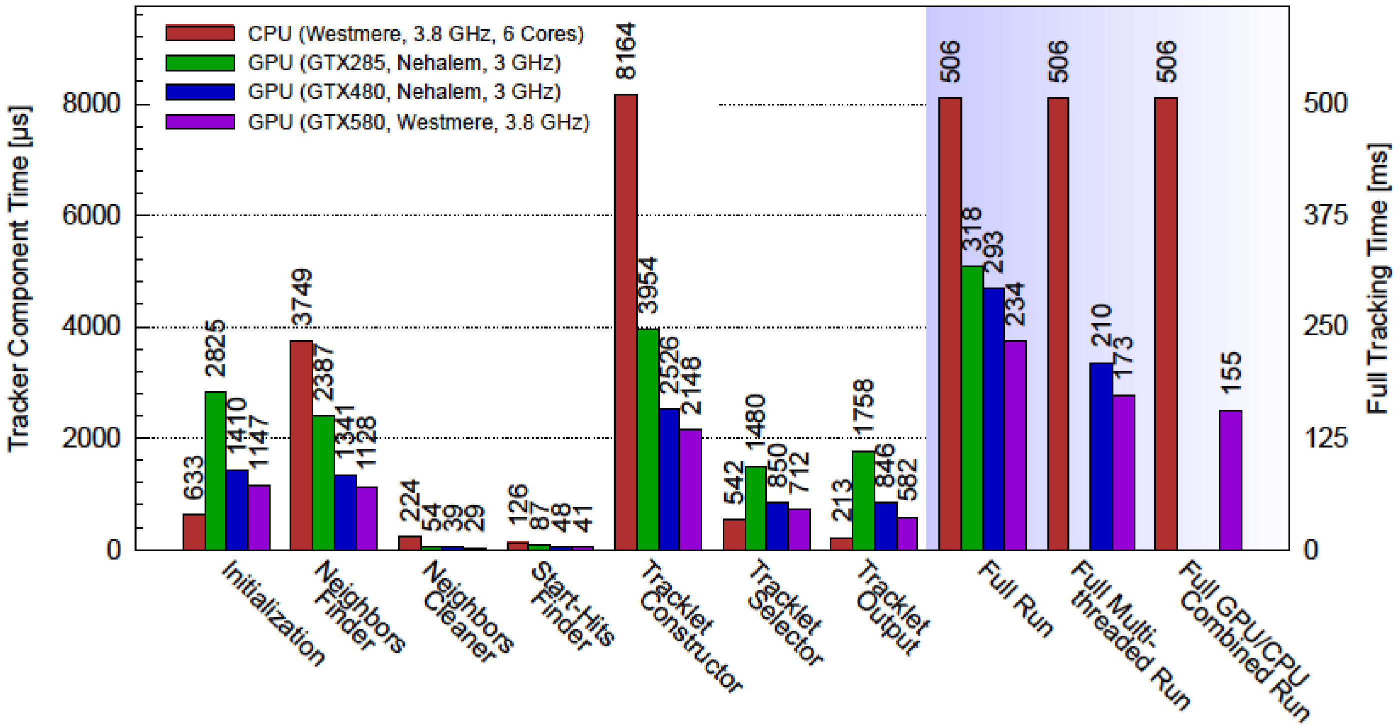


Picture from Herb Sutte: The Free Lunch Is Over
A Fundamental Turn Toward Concurrency in Software
Dr. Dobb's Journal, 30(3), March 2005 (updated)

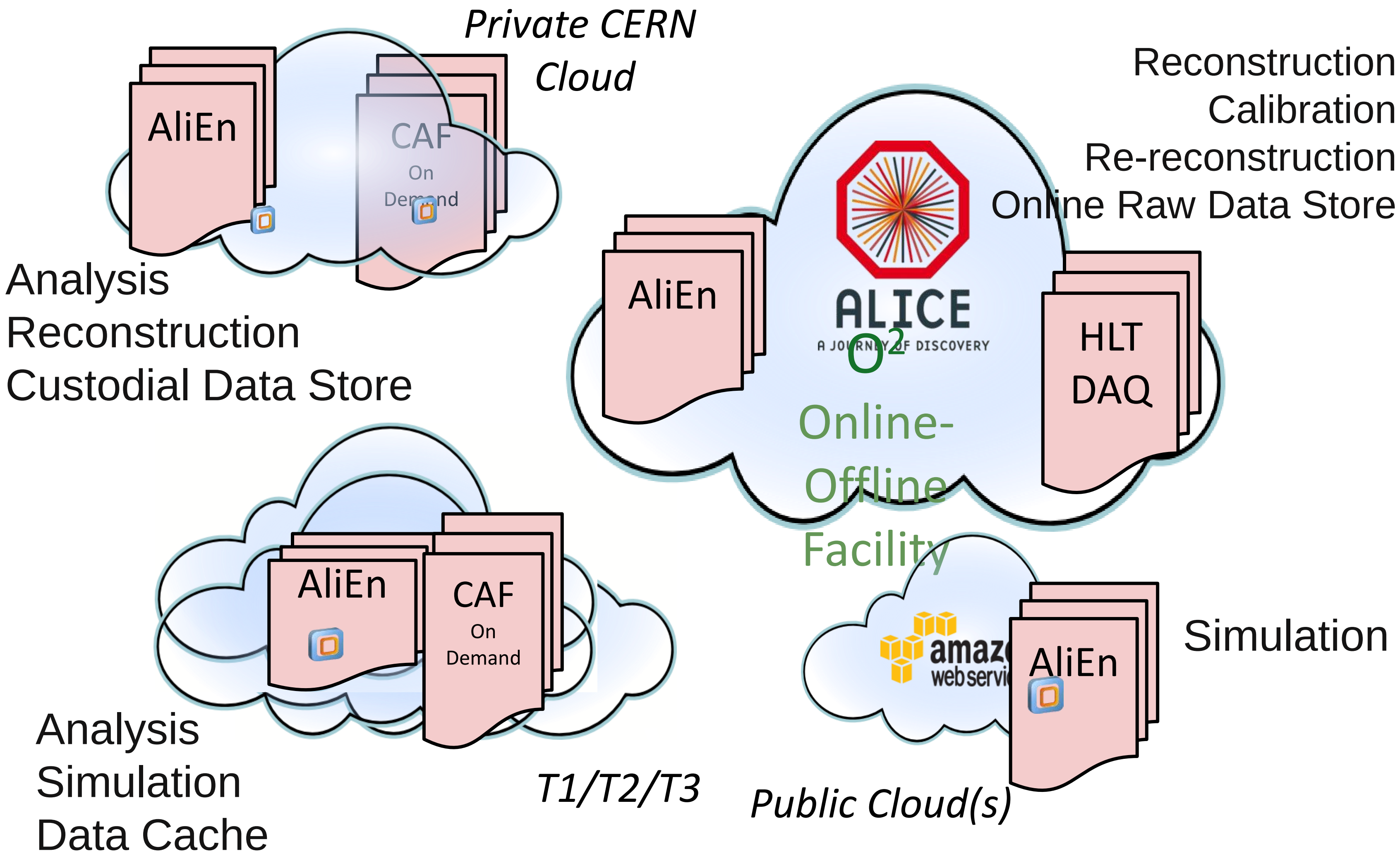
ALICE

Parallel Reconstruction

Tracking most time-consuming step in ALICE
Developed multi-threaded tracking algorithm for the HLT
Also adopted to GPUs (NVIDIA Fermi, CUDA)



Computing Model



ALICE

Summary

ALICE physics program requires handling
of 50kHz minimum-bias Pb-Pb collisions (1TByte/s)
from the online and offline systems

Strategy to handle the load is an ambitious data volume reduction by a first pass online reconstruction & discarding of raw data on a combined DAQ/HLT/offline farm (O2)

R&D towards the Online&Computing TDR at end of 2014

An interesting future ahead…

ALICE

Backup - Event Size

Detector	Event Size (MByte)	
	After Zero Suppression	After Data Compression
TPC	20.0	1.0
TRD	1.6	0.2
ITS	0.8	0.2
Others	0.5	0.25
Total	**22.9**	**1.65**

Expected data sizes for minimum bias
Pb-Pb collisions at full LHC energy

Why triggering does not work? The open charm case

A Large Ion Collider Experiment

Estimated signal statistics and trigger rate for minimum-bias Pb-Pb collisions
Hadronic Interaction rate of 50 kHz

Particle	Eff	S/ev	S/B	B'/ev	trigger rate (Hz)	S/nb^{-1}
D^0	0.02	$1.6 \cdot 10^{-3}$	0.03	0.21	$11 \cdot 10^3$	$1.3 \cdot 10^7$
D_s^+	0.01	$4.6 \cdot 10^{-4}$	*0.01*	0.18	$9 \cdot 10^3$	$3.7 \cdot 10^6$
Λ_c	0.01	$1.4 \cdot 10^{-4}$	$5 \cdot 10^{-5}$	11	$5 \cdot 10^4$	$1.1 \cdot 10^6$
Λ_c ($p_t > 2$ GeV/c)	0.01	$0.8 \cdot 10^{-4}$	0.001	0.33	$1.6 \cdot 10^4$	$0.6 \cdot 10^6$
$B \to D^0(\to K^-\pi^+)$	0.02	$0.8 \cdot 10^{-4}$	0.03	$11 \cdot 10^{-3}$	$5 \cdot 10^2$	$0.6 \cdot 10^6$
$B \to J/\psi(\to e^+e^-)$	0.1	$1.3 \cdot 10^{-5}$	*0.01*	$5 \cdot 10^{-3}$	$3 \cdot 10^2$	$1 \cdot 10^5$
$B^+ \to J/\psi K^+$	*0.01*	$0.5 \cdot 10^{-7}$	*0.01*	$2 \cdot 10^{-5}$	1	$4 \cdot 10^2$
$B^+ \to \overline{D}^0\pi^+$	*0.01*	$1.9 \cdot 10^{-7}$	*0.01*	$8 \cdot 10^{-5}$	4	$1.5 \cdot 10^3$
$B_s^0 \to J/\psi\phi$	*0.01*	$1.1 \cdot 10^{-8}$	*0.01*	$4.4 \cdot 10^{-6}$	$2 \cdot 10^{-1}$	$9 \cdot 10^1$
$\Lambda_b(\to \Lambda_c + e^-)$	*0.01*	$0.7 \cdot 10^{-6}$	*0.01*	$2.8 \cdot 10^{-4}$	14	$5 \cdot 10^3$
$\Lambda_b(\to \Lambda_c + h^-)$	*0.01*	$0.7 \cdot 10^{-5}$	*0.01*	$2.8 \cdot 10^{-3}$	$1.4 \cdot 10^2$	$5 \cdot 10^4$

B' is the background in the broad invariant mass range (±12σ)

Triggering on D^0, D_s and Λ_c (p_T>2 Gev/c) ➡ ~ 36 kHz