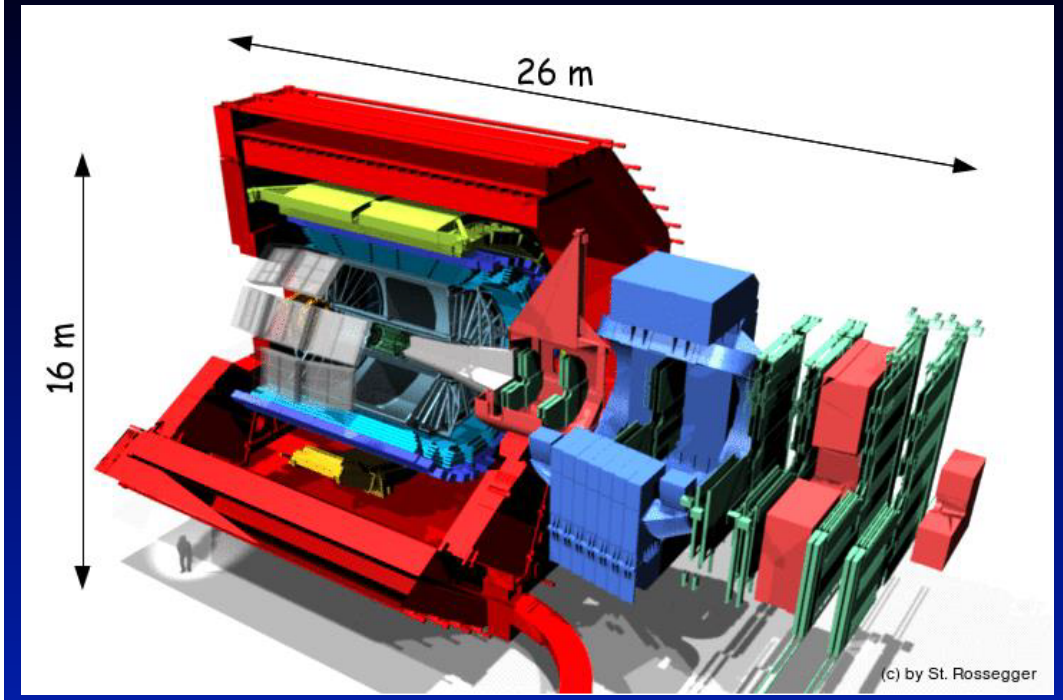
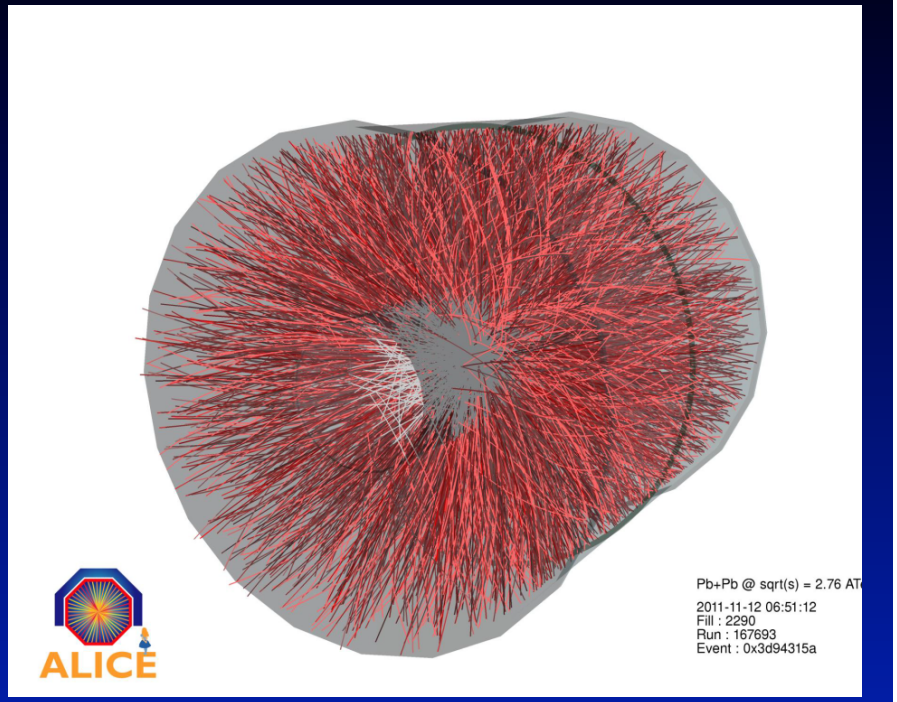




Evolution of ALICE Computing

3.12.2013

P.Malzacher@gsi.de

7th Annual Workshop of the Helmholtz Alliance "Physics at the Terascale"

ALICE

Evolution of ALICE Computing

Components of the ALICE Computing Model

Evolution of ALICE

Run 2

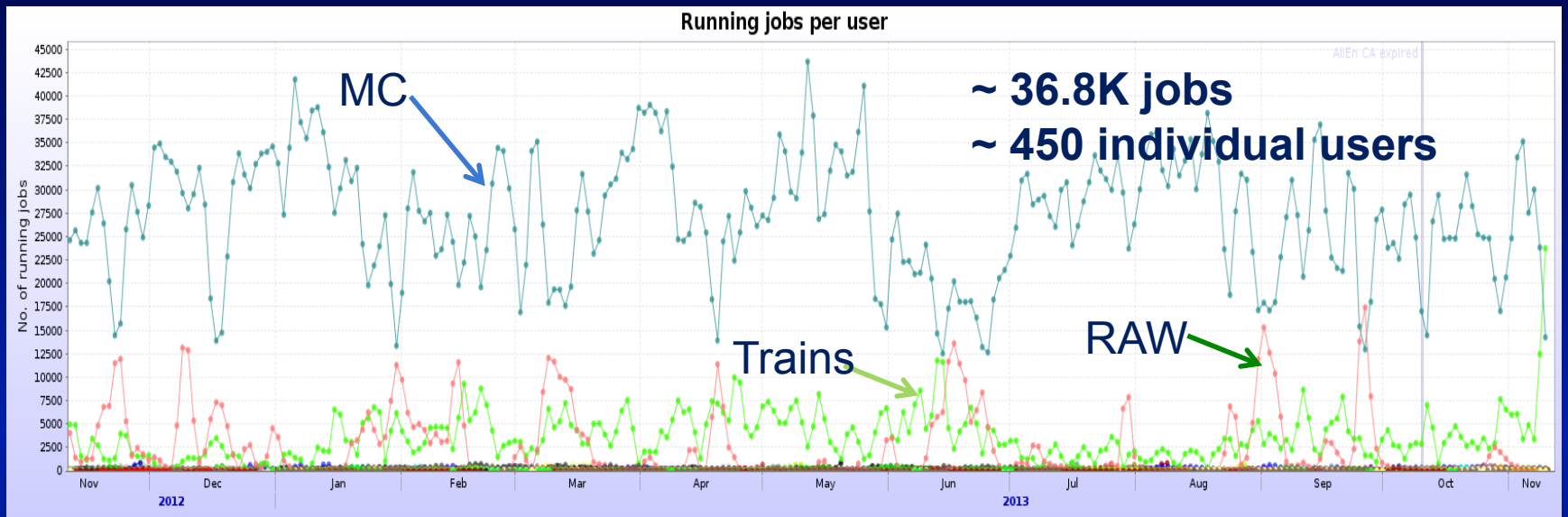
Run 3 / O2

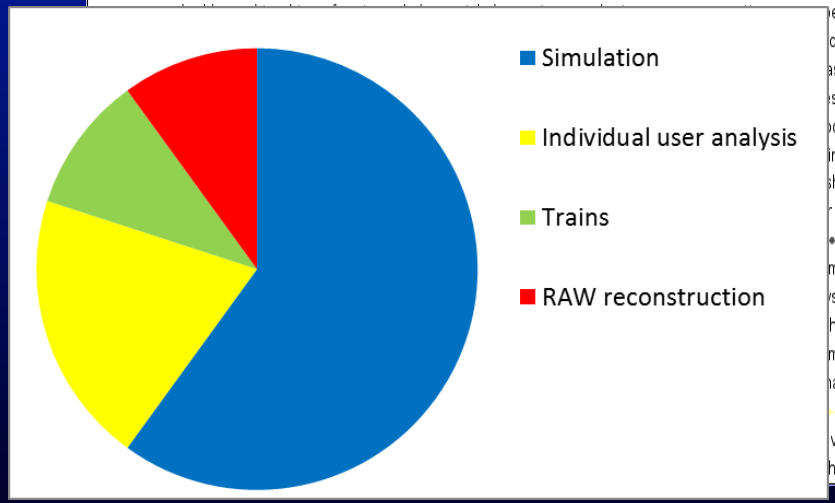
FAIR Computing

Summary

Run 1			LS 1		Run 2			LS 2		Run 3	
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021

Components of ALICE Computing: ROOT, XRootD, AliRoot, AliEn & MonALISA

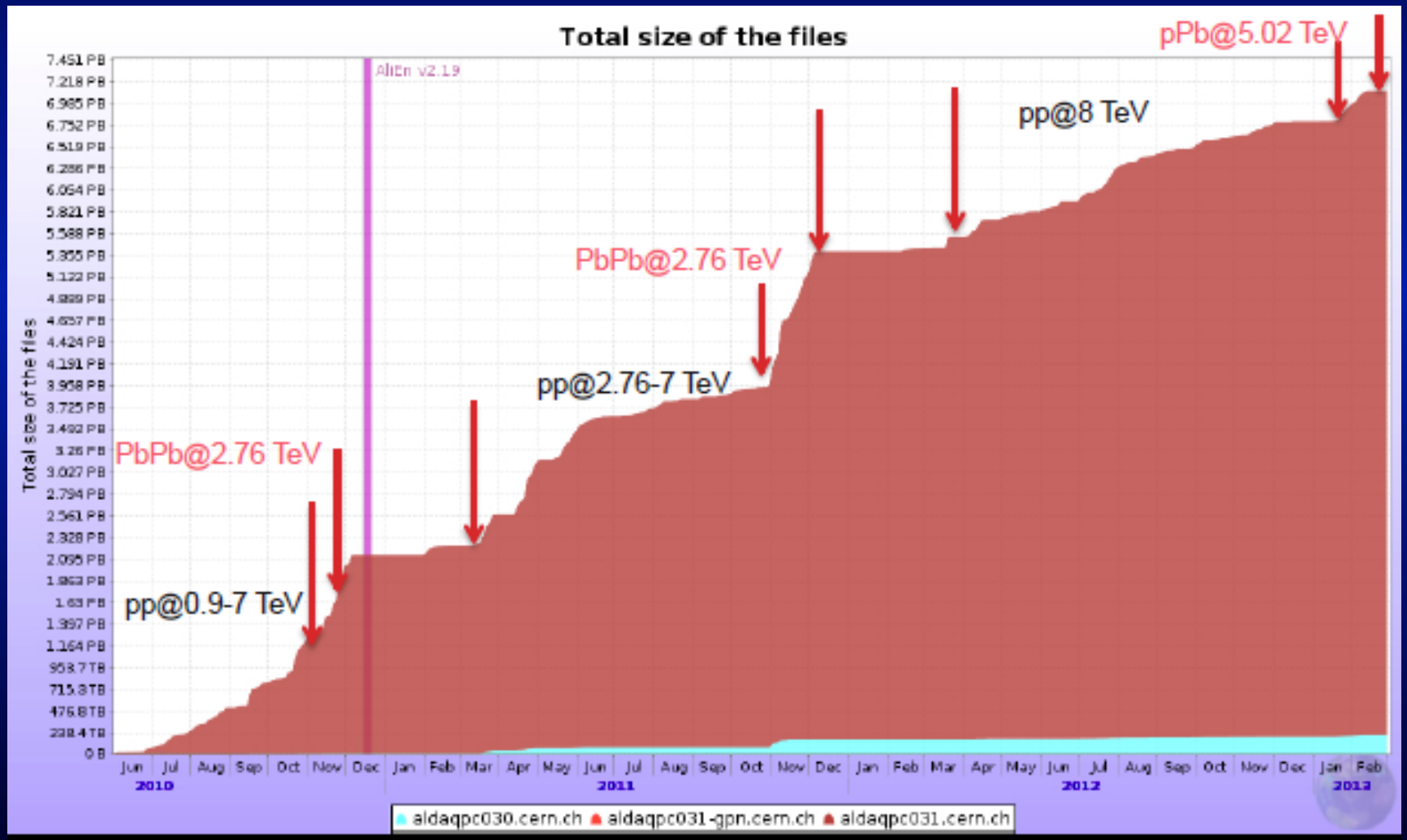




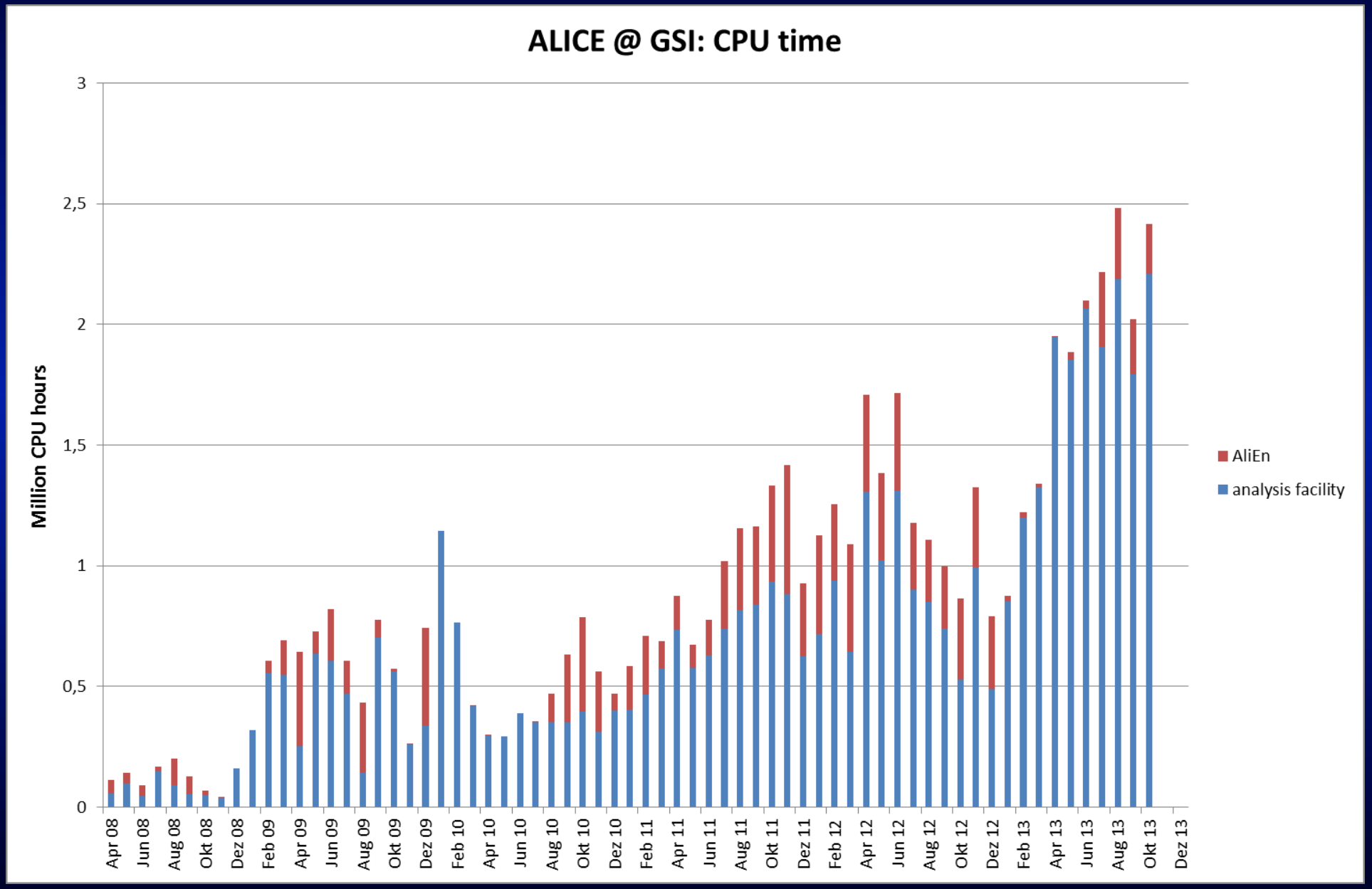
Stable Operation
~ 80 % efficiency
Problems:

- **user analysis**
- **widly distributed datasets**
- **lack of manpower**

7.3 PB of raw data collected during RUN1
16 PB of derived data on disk (MC,ESD, AOD)



The ALICE T2 / AF ~ 20% of ALICE T2 requests



Analysis

Trains = organized analysis

- 10-20 analysis tasks (wagons) running together on the same data sample
- 57 active trains (from all Physics Working Groups)
- Running twice weekly
- Average train duration – 21 hours
- 10% of all resources in the past year

The Evolution of ALICE: From Run 1 to Run 2

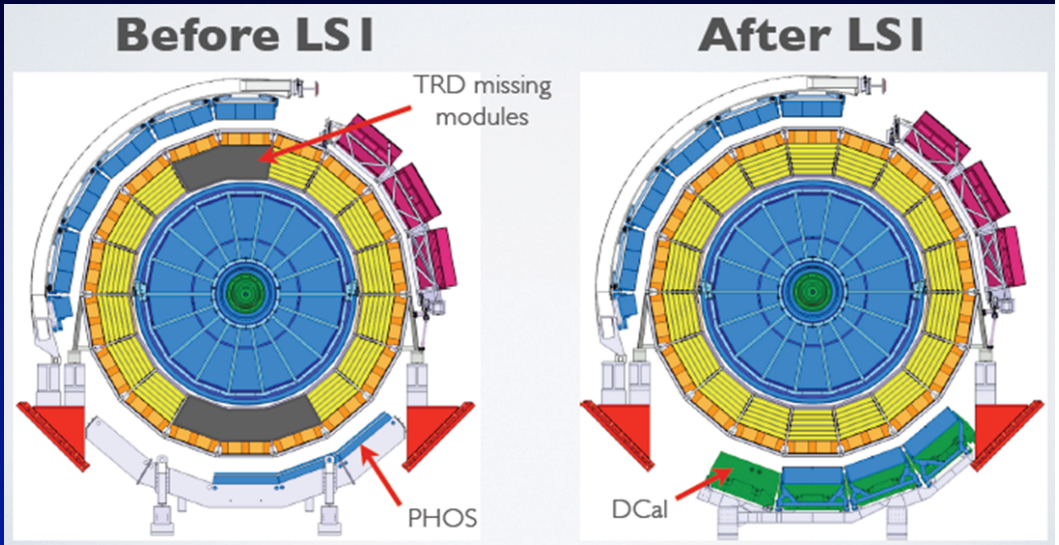
instant luminosity Pb-Pb * 4

additional detector coverage

upgrade DAQ & HLT

upgrade of readout electronics

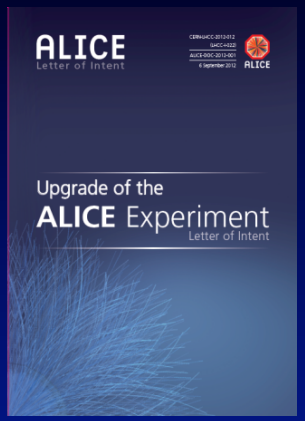
readout rate * 2



Year	System	Instant Lumi. (cm^{-2} s^{-1})	Interaction Rate (kHz)	Running Time (s)	# Events ($*10^9$)
2015	p-p unbiased	$2*10^{29}$	20	$3.1*10^6$	1.5
	Pb-Pb	10^{27}	8	$0.7*10^6$	0.35
2016	p-p rare triggers	$5*10^{30}$	500	$5.2*10^6$	2.6
	Pb-Pb	10^{27}	8	$0.7*10^6$	0.35
2017	p-p rare triggers	$5*10^{30}$	500	$7.1*10^6$	3.4
	p-Pb	10^{28} - 10^{29}	20 - 200	$0.7*10^6$	0.35

Run 1			LS 1		Run 2			LS 2		Run 3	
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021

Evolution of ALICE: Upgrade during LS 2



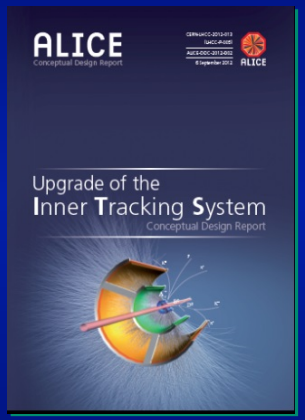
Inner Tracking System (ITS)

New, high-resolution, low-material ITS

Time Project Chamber (TPC)

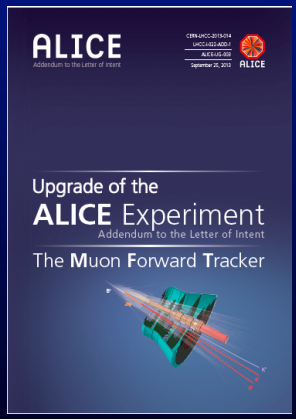
Upgrade of TPC with replacement of MWPCs with GEMs

New pipelined continuous readout electronics



New 5-plane silicon telescope in front of the Muon Spectrometer

New and common computing system for online and offline computing



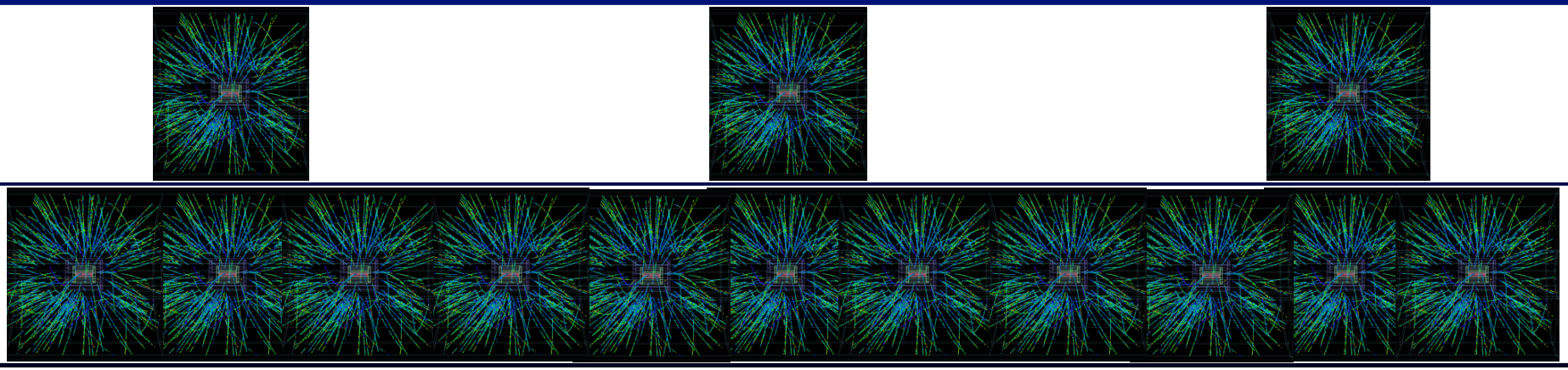
Requirements: Event Rate

Rate increase: from 500 Hz to 50 kHz

- Physics topics require measurements characterized by very small signal-over-background ratio → large statistics
- Large background → traditional triggering or filtering techniques very inefficient for most physics channels.
- Strategy: read out all particle interactions 50 kHz (anticipated Pb-Pb interaction rate)

TPC intrinsic rate << 50 kHz

- In average 5 events overlapping in the detector
- Continuous read-out



(R)evolution of ALICE Computing

Evolution: AliRoot 5.x
Based on Root 5.x ?
Improved algorithms
Implements needs of the PWGs

Similar vision for distributed computing

Revolution: AliRoot 6.x
Based on Root 6.x, C++11
Optimized for I/O
FPGA, GPU, MIC, ...
Steered by O2 CWGs

Run 1			LS 1		Run 2			LS 2		Run 3	
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021

The foreseen improvements of the software for Run 2 will focus on:

Moving one calibration iteration into the Online environment

Using HLT track seeds to speed up the Offline reconstruction

Replacing Geant 3 with Geant 4

Improving the performance of Geant 4 in ALICE

Fast and parameterized simulation to reduce the CPU needs

Collaboration with other experiments to make use of opportunistic resources for simulation

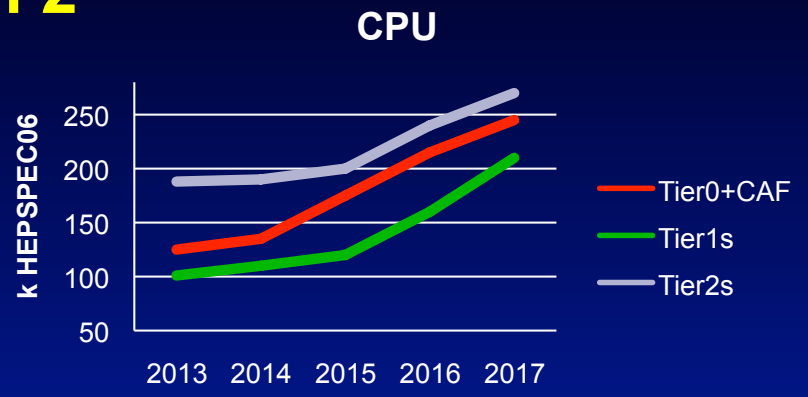
Adapting to cloud environments, use HLT for Offline processing

Reducing turn-around time of analysis trains

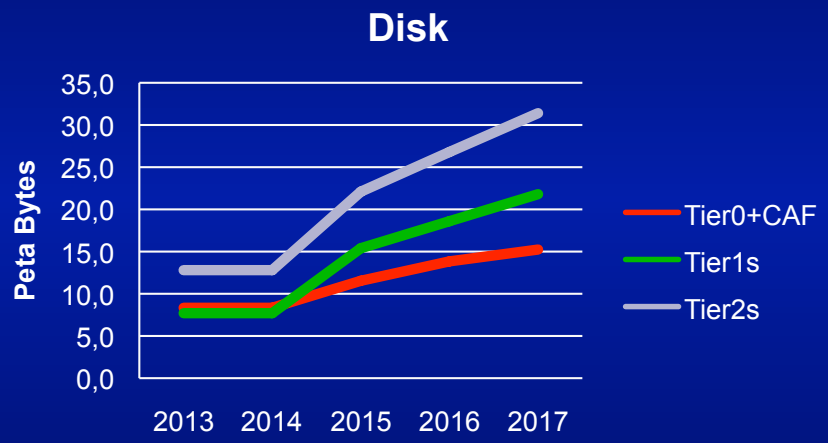
Consolidating popular datasets on fewer sites

Resource Requirements for Run 2

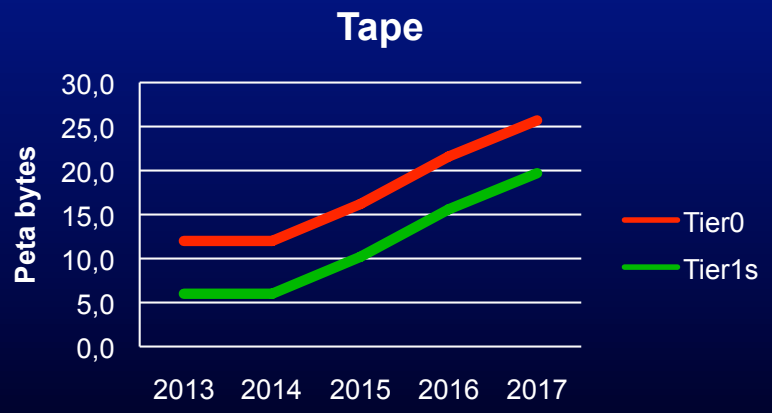
	CPU (kHEPSPEC06)			
	Tier0	CAF	Tier1s	Tier2s
2015	130	45.0	120	200
2016	170	45.0	160	240
2017	200	45.0	210	270



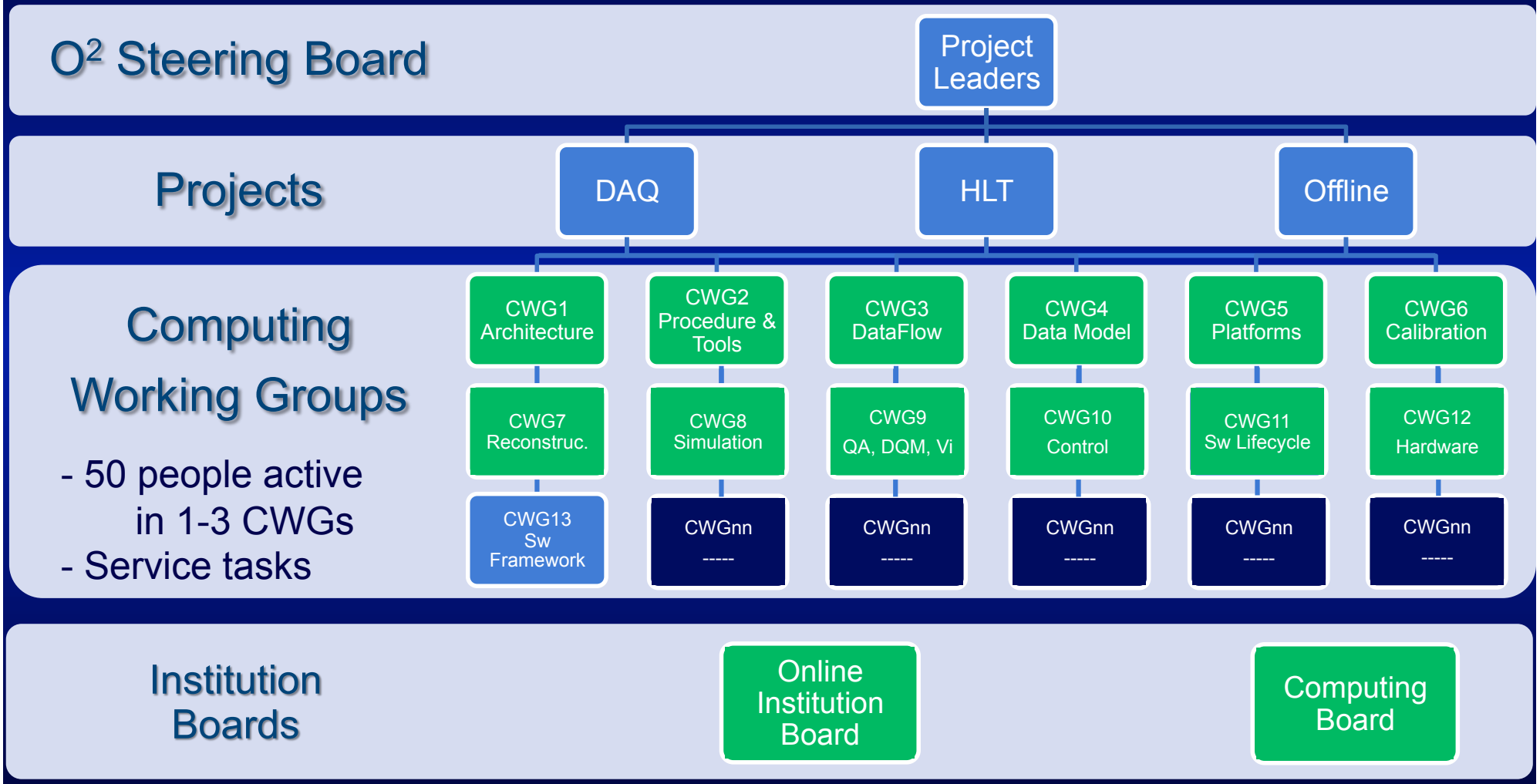
	Disk (PB)			
	Tier0	CAF	Tier1s[1)]	Tier2s
2015	11.2	0.34	15.4	22.1
2016	13.4	0.44	18.6	26.8
2017	14.7	0.54	21.8	31.4



	Tape (PB)	
	Tier0	Tier1
2015	16.2	10.2
2016	21.6	15.6
2017	25.7	19.7



Preparation for Run 3: O^2 Project



Sep 2012 ALICE Upgrade LoI

Jan 2013 Report of the DAQ-HLT-Offline software panel on “ALICE Computer software framework for LS2 upgrade”

Mar 2013 O^2 Computing Working Groups

Sep 2014 O^2 Technical Design Report

Intensive period of R&D :

Collect the requirements: ITS and TPC TDRs

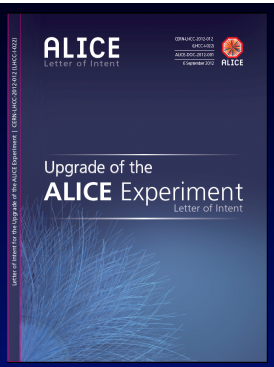
System modeling

Prototyping and benchmarking

Technology and time are working with us

New options

Massive usage of commercial equipment very a




EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

ALICE Computing software framework for LS2 Upgrade version 1.0

O2 Computing Working Groups

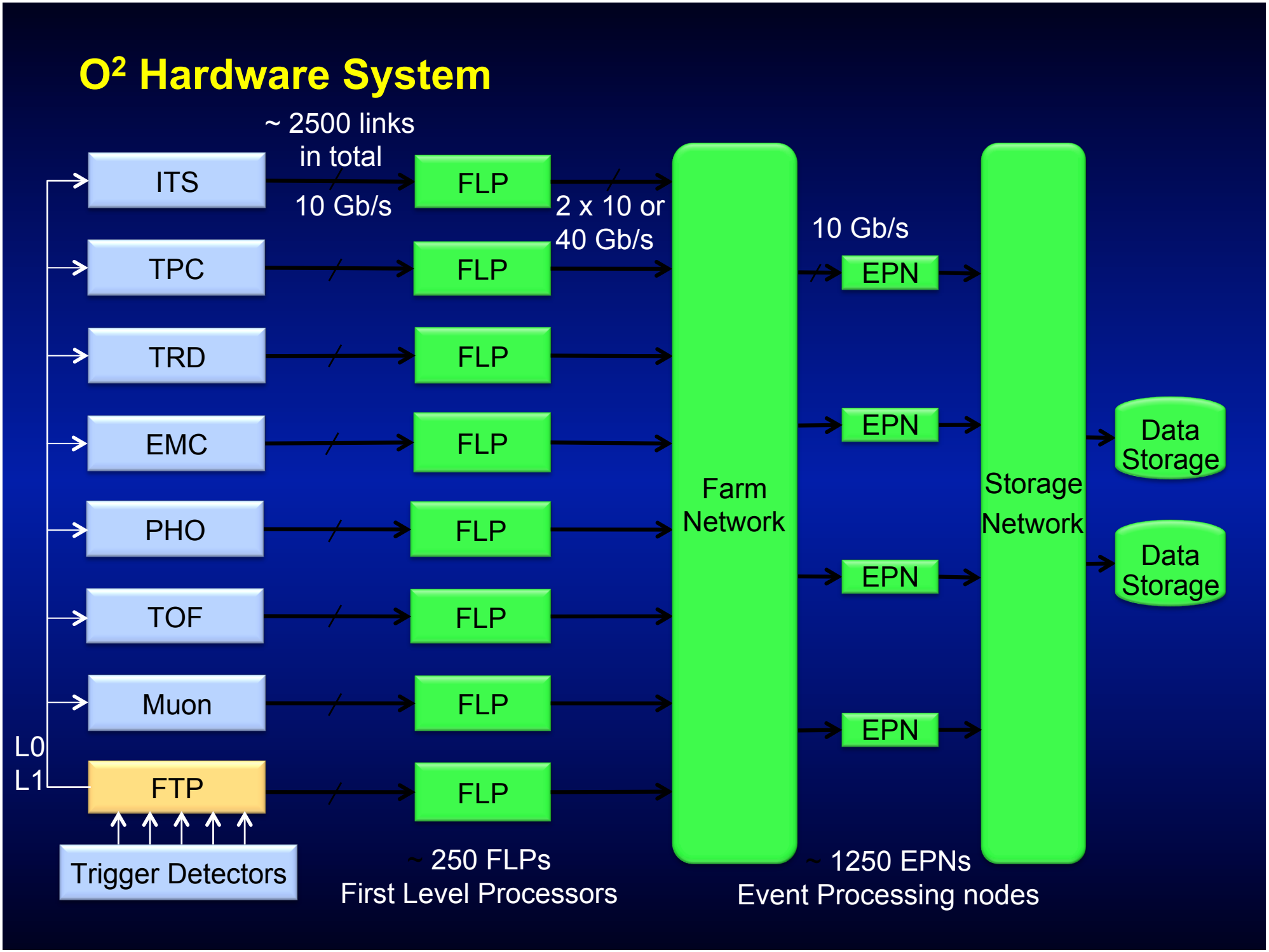
O^2 Technical Design Report

Requirements for Run 3: Data Volume

Detector	Event Size After Zero Suppression (MByte)	Bandwidth @50 kHz Pb-Pb (GByte/s)
TPC	20.0	1000
TRD	1.6	81.5
ITS	0.8	40
Others	0.5	25
Total	**22.9**	**1146.5**

Massive data volume reduction needed

Only option is by online processing


O2 Hardware System
~ 2500 links in total
10 Gb/s
2 x 10 or 40 Gb/s
10 Gb/s
ITS
TPC
TRD
EMC
PHO
TOF
Muon
FTP
L0
L1
Trigger Detectors
FLP
FLP
FLP
FLP
FLP
FLP
FLP
FLP
~ 250 FLPs
First Level Processors
Farm Network
EPN
EPN
EPN
EPN
~ 1250 EPNs
Event Processing nodes
Storage Network
Data Storage
Data Storage

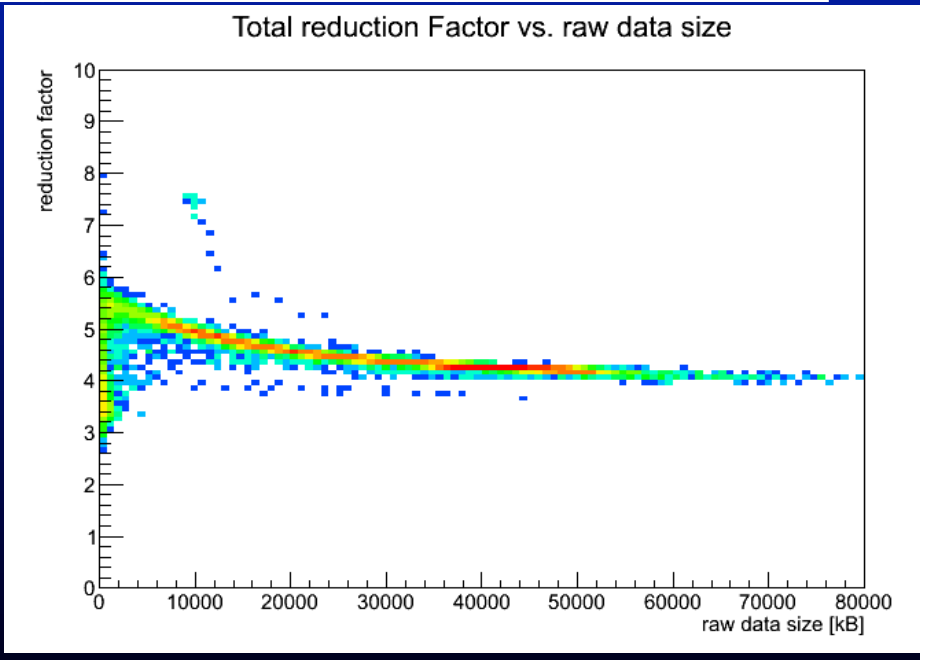
TPC Data Volume Reduction

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

TPC data volume reduction by online event reconstruction

Discarding original raw data

In production from the 2011 Pb-Pb run

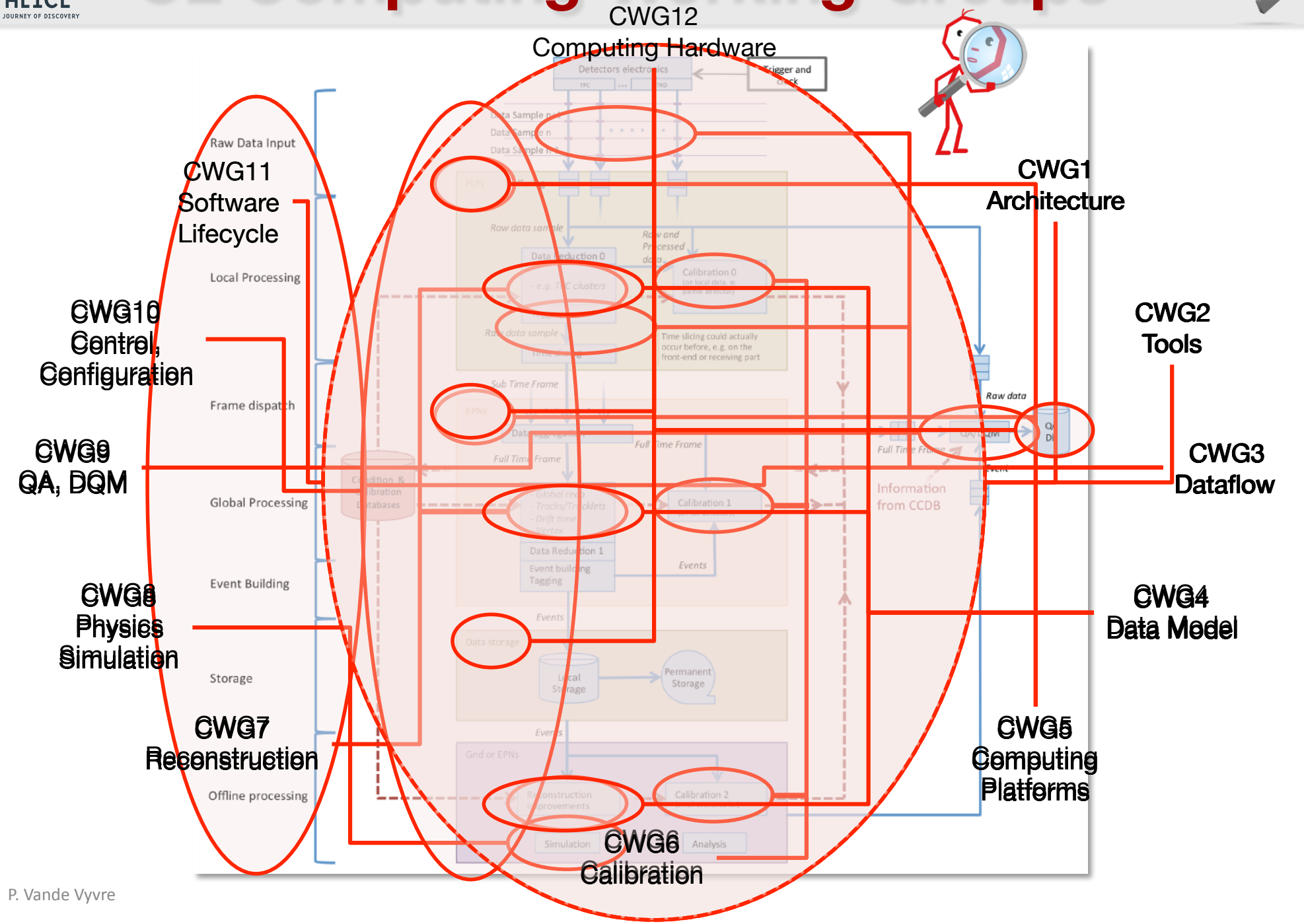


Total Data Volume

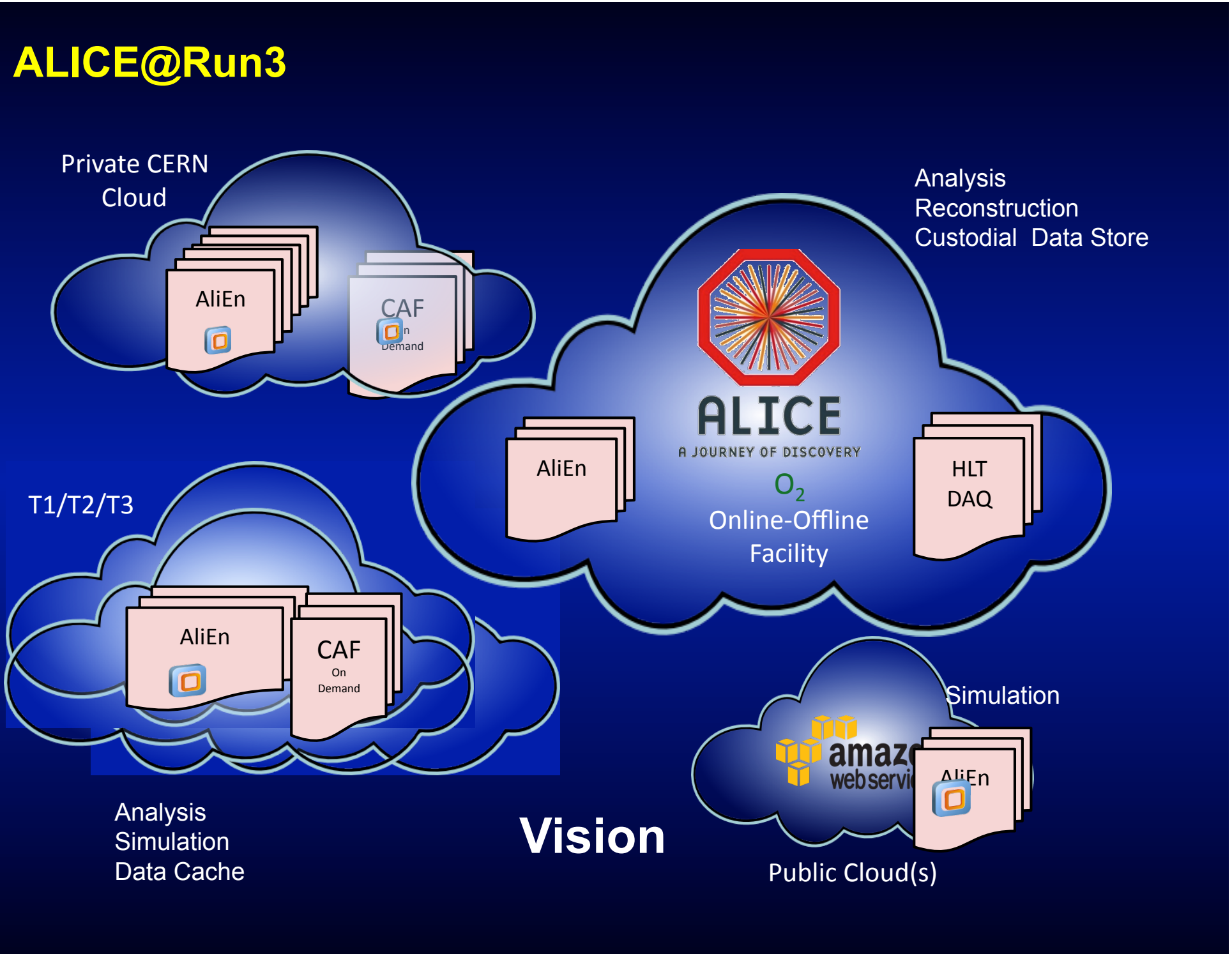
Detector	Input to Online System (GByte/s)	Peak Output to Local Data Storage (GByte/s)	Avg. Output to Computing Center (GByte/s)
TPC	1000	50.0	8.0
TRD	81.5	10.0	1.6
ITS	40	10.0	1.6
Others	25	12.5	2.0
Total	**1146.5**	**82.5**	**13.2**

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

O2 Computing Working Groups



ALICE@Run3


Private CERN
Cloud
AliEn
CAF
On
Demand
Analysis
Reconstruction
Custodial Data Store
ALICE
A JOURNEY OF DISCOVERY
AliEn
O_2
Online-Offline
Facility
HLT
DAQ
T1/T2/T3
AliEn
CAF
On
Demand
Analysis
Simulation
Data Cache
Vision
Simulation
amaz
web servi
AliEn
Public Cloud(s)

FAIR: Facility for Ion and Antiproton Research: First Beam ~2018, explore commonalities, CTDR due 2015

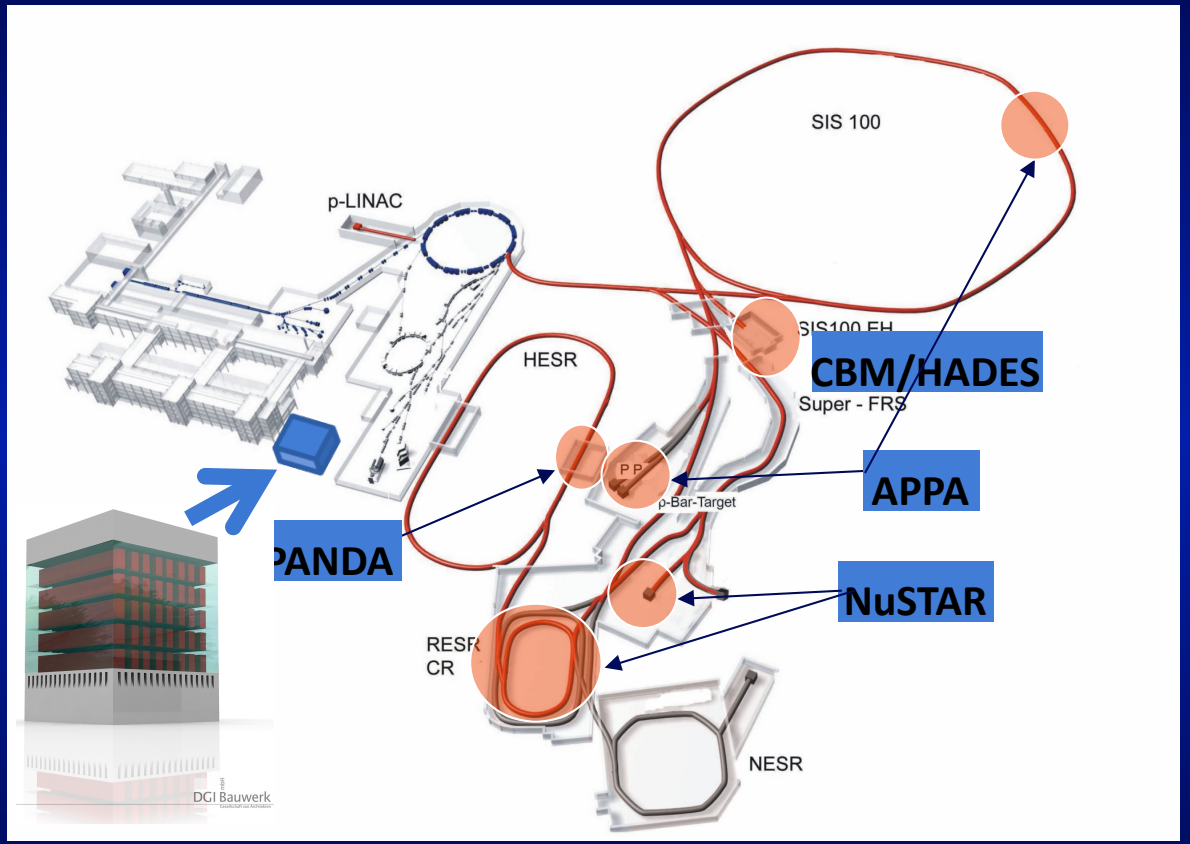
GSI computing 2013

- ALICE T2/AF
- HADES, Theory, FAIR
- ~ 14000 cores,
- ~ 5 PB lustre

FAIR computing 2020

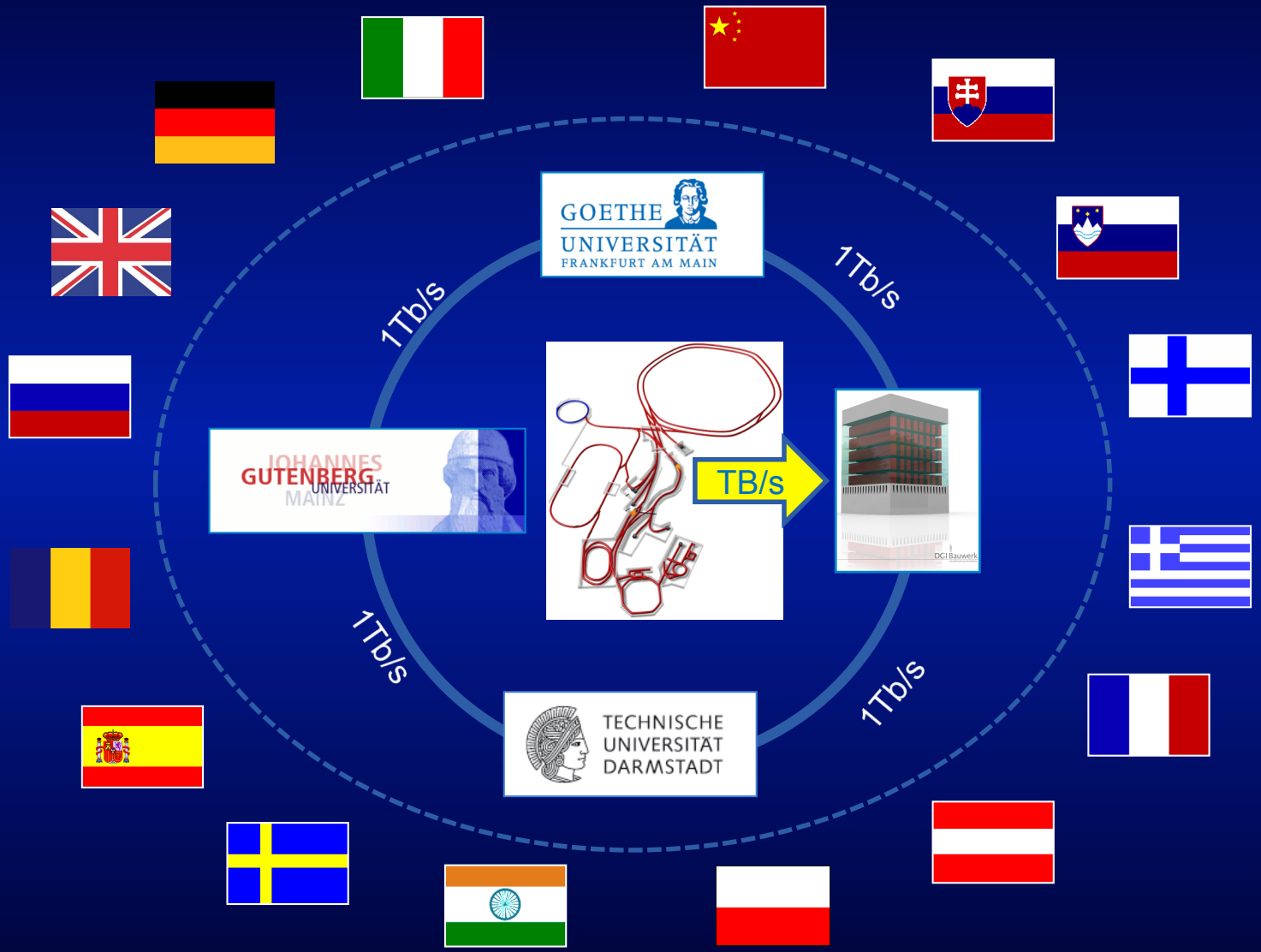
- CBM
- PANDA
- NuSTAR
- APPA
- LQCD
- 300000 cores
- 40 PB disk
- 40 PB archive

& FAIR MAN
& FAIR Grid/Cloud



open source and community software
budget commodity hardware
support different communities
scarce manpower

FAIR T0/T1 Metropolitan Area Network integrated in an international Grid/Cloud



In discussion: Joining effort from O2 and FairRoot

Data Acquisition

Analysis

Simulation

ØMQ

high-performance asynchronous messaging library; patterns: request-reply, publish-subscribe, push-pull, exclusive pair

Cluster Finding

Analysis

Tracking

ESD, AOD

Summary: (R)evolution of ALICE Computing

Three phases, each jumping one order of magnitude in statistics and progressively improving the detectors, e.g.:

Run1 Pb-Pb: 0.1 nb^{-1} Run2: 1 nb^{-1} Run3&4: 10 nb^{-1}

ROOT, XRootD, AliRoot, AliEn & MonALISA the key components for the successful analysis of Run 1.

Computing for Run 2 is an evolution of the systems from Run 1 by the three projects DAQ, HLT and Offline.

For Run 3 the O2 project develops a complete redesign of:

- one common new online and offline computing system
- a common computing farm and software framework.

The main risk is scarce manpower $\longrightarrow$ team up with other experiments and projects from industry and open source.