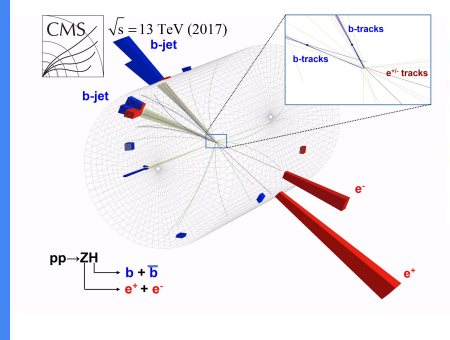


CMS towards HL-LHC - a small update

WLCG OB Nov 30th 2018

Tommaso Boccali (INFN Pisa/CERN)

Markus Klute (MIT)



Observation of Higgs boson decay to bottom quarks

CMS Collaboration

(Submitted on 24 Aug 2018)

- ▶ WLCG Strategy document will undergo a review process at the start of 2019
- ▶ In the expected activity items, CMS is advanced in a few critical ones
- ▶ **Storage:** the “physics” data tier shrunk by a factor **3000x** between 2010 and 2018
 - ▶ **NanoAOD** are **not** considered “commissioned” for Physics today, but a very important step was done on Aug 24th:
 - ▶ **Observation of Higgs boson decay to bottom quarks:** first published analysis using NanoAOD data/MC; but in general not to be considered physics-commissioned yet!
- ▶ While there is no intention / possibility to move all the analyses to NanoAOD, a 50% utilization as projected allows to safely reduce the # of MiniAOD copies on Disk
 - ▶ Starting to be used for the Phasell modelling

Data Tier	Size (kB)
RAW	1000
GEN	< 50
SIM	1000
DIGI	3000
RECO(SIM)	3000
AOD(SIM)	400 (8x reduction)
MINIAOD(SIM)	50 (8x reduction)
NANOAOD(SIM)	1 (50x reduction)

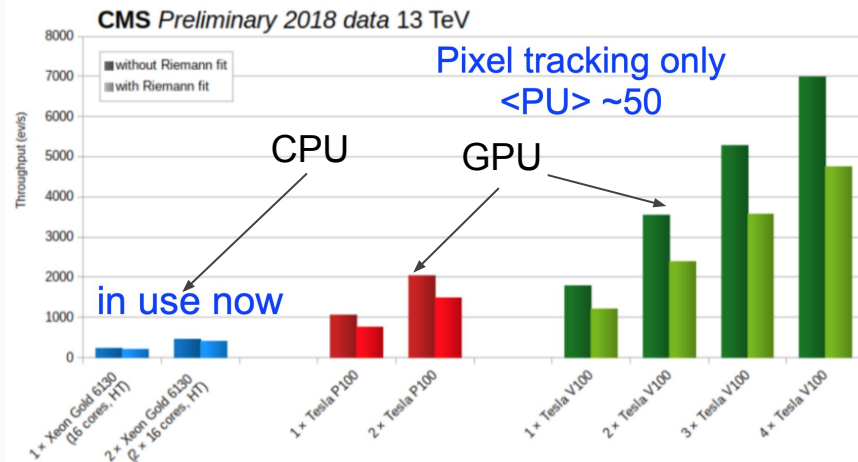
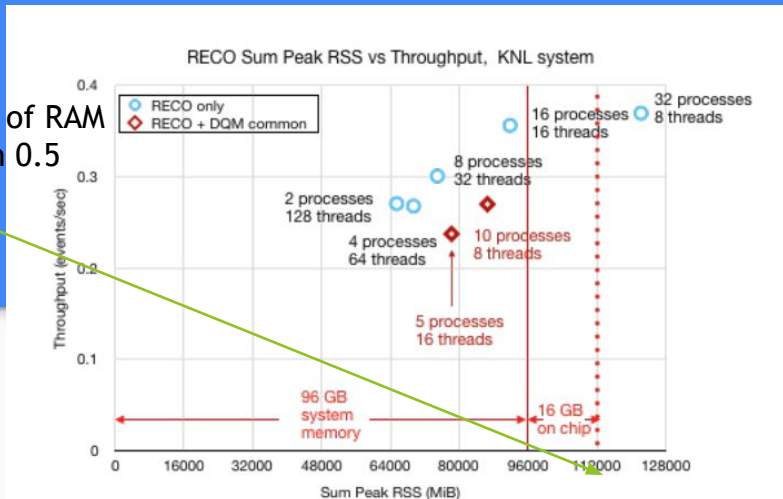
Analysis data formats

@ typical RunII PU

Multithreading and more

KNL only provides 96 GB of RAM over 272 cores, less than 0.5 GB/core

- ▶ Real multithreading (not fork!) enabled @ Tier0 AND HLT since 2015, and for distributed computing since Summer 2016
 - ▶ Using Intel's TBB as workhorse
 - ▶ # of threads (cores used) and streams (events processed in parallel) independently tunable parameters, allow to choose for better efficiency or less RAM
- ▶ Next CMSSW release (10_5) will our first "pre-production ready" Heterogeneous Framework deployed
 - ▶ Abstract interface accommodates GPU, FPGA,
 - ▶ Pixel Tracking for GPU first guinea pig



Software - RunIII used as a RunIV dress rehearsal

- ▶ Moving to Community projects and reduce CMS specific codes
 - ▶ Impact on sustainability and expected reduced support
- ▶ Data Management:
 - ▶ Review in late July of the solutions for RunIII+HL-LHC; Rucio (ATLAS, AMS, Xenon1T, ...) selected
 - ▶ Deployment plan being prepared. Idea is to test in a 2020 via a RunIII dress rehearsal
- ▶ Information system:
 - ▶ Moving to CRIC, WLCG supported. First CMS service (SiteDB) already substituted
- ▶ Geometry Database
 - ▶ Moving from DDD (CMS, 2002) to DD4HEP (AIDA, LC, ...)
- ▶ Testing GeantV transport engine
 - ▶ CMS is the most advanced LHC experiment in testing GeantV transport engine, with alpha version now; testing with beta expected during LS2 and decision for transition during RunIII (a community-wide decision?)
 - ▶ In any case, already using today GeantV features, like a backported VecGeom

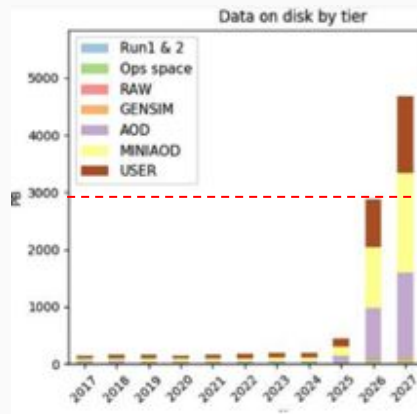
Configuration	Relative CPU usage	
	MinBias	ttbar
No optimizations	1.00	1.00
Static library	0.95	0.93
Production cuts	0.93	0.97
Tracking cut	0.69	0.88
Time cut	0.95	0.97
Shower library	0.60	0.74
Russian roulette	0.75	0.71
FTFP_BERT_EMM	0.87	0.83
All optimizations	0.21	0.29

Improvements in utilization of Geant4 in CMS in the last 10 years
A factor 3-5 achieved!

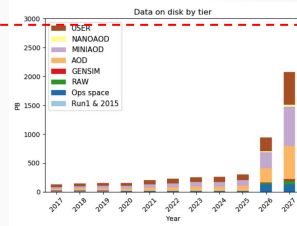
Updated HL-LHC projections (Sep'18)

- ▶ **Sept'18: large difference wrt 2017 projections on the storage aspect, the most critical for our FAs**
 - ▶ CPU can also come from Commercial, HPC, Opportunistic
 - ▶ There is currently no modelling to have not owned disk
- ▶ Main change comes gaining confidence that NanoAOD would be usable during 2027 (after commissioning in 2026).

- ▶ This allows to reduce the AOD/MiniAOD residing on Disk
- ▶ the effect is close to a factor 2x in disk storage
- ▶ also has an effect on expected CPU for analysis
- ▶ Many other lower impact re-tunings
- ▶ Changes
- ▶ **Tape ~ constant**
- ▶ **CPU -15% (less user level analysis needed)**
- ▶ **Disk -50%**



When on the same scale



Projected 2027 needs (Sept17 → Sept18):

- CPU 50 → 44 MHS06
- Disk 5 → 2.2 EB
- Tape 3 → 3 EB

Evolution of Computing Model 2X - ECoM2X

- As reported at LHCCs, **ECoM2X** is a task force with all the CMS coordination areas (Physics, Computing, Trigger, PPD)
- Activities going on on many fronts, **aiming to a first assessment of results by the end of 2018**
 - a. Modelling of resource needs
 - b. Understanding of Physics inputs (rates, MC/DT, ...)
 - c. Technology tracking - HW and heterogenous SW solutions
 - d. External funding opportunities
- A notable fact: together with ATLAS O+C Management, we are applying for a EU funded Training Network on SW R&D for Phase II
 - Endorsed by CERN, experiments, IRIS-HEP (“twin project” in US - already granted 25M\$)
 - To be submitted by Jan 2019
 - CERN, ES, FR, DE, UK, SK, NO, IT
 - **PI: Borut Kersevan (Ljubljana)**