CMS simulation long-term timeline

V. Daniel Elvira

CMS Simulation: next 4+ years

Full simulation only schedule – CMS also has program to use Fast Sim techniques in its Geant4 full simulation application (machine learning, parametrizations), and to execute full simulation in a hybrid cloud environment which includes HPC systems. (The GeantV transport engine design utilizes SIMD vectorization for fine-grained parallelism.)

R&D timeline: (targets HL-LHC)

GeantV alpha tag within Toy/actual CMSSW: tests of multithread models, track-level basketization, user interfaces

Development of Computing Performance (CP) infrastructure for Geant4 versus GeantV CP tests within CMSSW

GeantV beta tag within CMSSW: CP tests to evaluate speedup, iteration with developers

Community decision. If YES, iteration with developers: features, speed, integration to Geant4 toolkit

?

Commissioning 202X-2025/2026: computing performance, robustness, physics validation: full support from Geant4 Collaboration

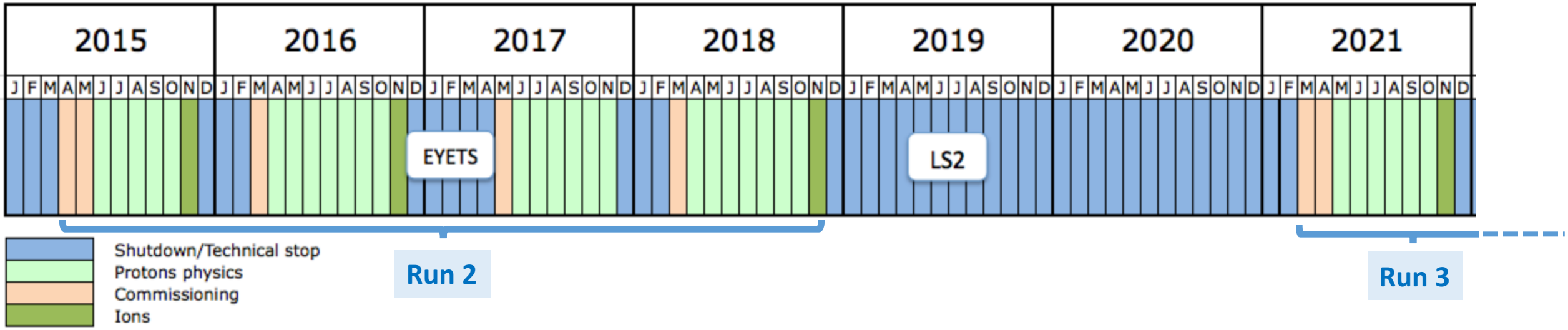
?

?

Production timeline:

Upgraded shower libraries, G4 DormandPrince stepper, smart tracks, RN handling for full reproducibility

DD4HEP migration, improved pileup infrastructure, potential G4 upgrades (new releases, improved navigator, predictions with systematic uncertainties – VMP, other modular packages from R&D)



Interesting background information about CMS and ATLAS Detector Simulation CP

Detector simulation CP performance

HEP Software Foundation Community White Paper Working Group - Detector Simulation
https://arxiv.org/abs/1803.04165

ATLAS-CMS CP comparison exercise (benchmark configuration):

- Machine: olhswep16.cern.ch (CERN's OpenLab), one thread runs
- Compiler: gcc 6.3
- Geant4: version 10.2, FTFP_BERT physics list
- Pythia [17] events: 13 TeV Pythia minimum bias (300 events) and ttbar (300 events), pseudo rapidity cut $|\eta| < 5.5$
- Particle gun: 50 GeV e's, muons, pions with a flat η, ϕ distribution in $\eta=[-0.8,0.8]$, $\eta=[2,2.7]$, and $\phi=[0,2\pi]$
- Geometry: 2015 or 2016

(Default configuration what the experiments actually run in production.)

For the default configuration, CMS was 7-10 times faster than ATLAS (as of July 2017), half of it comes from approximations, and half from a more complex detector.

Numbers correspond to Geant4 module only. No Gen, no digi, no reco

CMS:	Minbias (sec*/evt)	ttbar (sec*/evt)	(* 1 sec = 11 HS06 for machine used here)
default (As described in first row of Table 2)	4.3	24.6	
Benchmark (As described in second row of Table 1)	4.9	29.6	
All approximations removed (As described in last row of Table 2)	23.2	100.9	

ATLAS:	Minbias (sec*/evt)	ttbar (sec*/evt)	(* 1 sec = 11 HS06 for machine used here)
default (As described first first row of Table 4, absolute values taken from lines 18,19 of the attached spreadsheets)	42.2	177.6	
Benchmark (Computed from first row in Table 4, and multiplying by factor for no sim above eta=5.5, using no change when switch to FTFP_BERT)	39.7	168.7	
All approximations removed (As described in last row of Table 2)	Cumulative numbers not provided, but approximations other than shower libraries are small		

Absolute numbers in seconds or HS06 not available in CWP simulation note, as per decision of the ATLAS management

In summary:

For the default configurations, the ATLAS/CMS ratios are	9.8 (minbias)	7.2 (ttbar)
For the benchmark configurations, the ATLAS/CMS ratios are	8.1	5.7
For the "All approximations removed" configuration we do not have numbers. The ratios here are computed assuming the individual factors in Table 4 are multiplicative (1.6 for min bias and 1.3 for ttbar)	5.1	4.4

Note that ATLAS does not use Russian Roulette approximation. About half of the ATLAS/CMS ratio comes from different detector complexity and the other half from simulation approximations present in CMS and not present in ATLAS.

Also note that while in CMS there is a significant penalty on time performance as we move to LH-LHC detectors, in ATLAS the increase in CPU time is minimal.

Detector simulation CP performance

CMS

	Minimum Bias		Top Pair	
	CPU Time	Memory (RSS)	CPU Time	Memory (RSS)
2015	1.00	0.75 GB	1.00	0.79 GB
2016	1.01	0.75 GB	1.01	0.75 GB
2017	1.21	0.70 GB	1.13	0.75 GB
2023D4	1.28	0.87 GB	1.65	0.83 GB
2023D17	1.36	0.78 GB	1.57	0.78 GB

Table 1: Time and memory performance of the Geant4 module of the CMS Full Simulation application for minimum bias and ttbar events in current and future detector configurations. The Geant4 version used in the test is 10.2p02. With respect to the 2015 and 2016 configurations, which are very similar, 2017 adds an upgraded pixel detector and modifies the forward calorimeter layer configuration, 2023D4 and 2023D17 are different versions of the HL-LHC CMS detector with an upgraded tracker and a High Granularity end cap calorimeter (HGCAL), the latter sub-detector involving a significant increase in the number of Geant4 volumes.

Upgrade configuration performance compared to the 2015 benchmark

In addition, CMS may need to use a more CPU expensive G4 physics list, for accurate physics modeling in the HGCAL (~25-65% increase)

Detector simulation CP performance

CMS CPU ratios: individual measurements

Configuration	MinBias	ttbar
CMS default (FTFP_BERT_EMM)	0.87	0.83
FTFP_BERT	1.00	1.00
no Russian roulette	1.33	1.41
no HF shower library	1.68	1.36
no tracking cuts	1.45	1.13
no time cuts	1.05	1.03
no cuts per region	1.07	1.03
no static build	1.05	1.07

Table 2: Relative time performance of the Geant4 module of the CMS Full Simulation application for minimum bias and ttbar production. Measurements of the individual effect of several options to reduce CPU time consumption are listed. The Geant4 version used in the test is 10.2.p02. Values of range cuts and tracking cuts are specified in the text.

CMS CPU ratios: cummulative measurements

Configuration	MinBias	ttbar
CMS default	1.0	1.0
no Russian roulette	1.28	1.37
no Russian roulette + no tracking and time cuts	1.84	1.63
no Russian roulette + no tracking and time cuts + FTFP_BERT	2.25	2.04
no Russian roulette + no tracking and time cuts + FTFP_BERT + no HF shower library	4.43	3.22
no Russian roulette + no tracking and time cuts + FTFP_BERT + no HF shower library + no cuts per region	4.73	3.41

Table 3: Relative time performance of the Geant4 module of the CMS Full Simulation application for minimum bias and ttbar production. Measurements of the cumulative effect of several options to reduce CPU time consumption are listed. The Geant4 version used in the test is 10.2p02. Values of range cuts and tracking cuts are specified in the text.

Detector simulation CP performance

ATLAS CPU ratios: individual measurements

Configuration	MinBias	ttbar
Nominal production configuration: shower libraries in the forward calorimeter, nominal range cuts, NystromRK4 stepper, FTFP_BERT_ATL physics list, 250ns neutron time cut, simulation of primary particles with pseudo-rapidity below 6.0	1.0	1.0
No shower libraries	1.5	1.3
ClassicalRK4 stepper instead of NystromRK4	1.09	1.07
No neutron time cut	1.02	1.01
FTFP_BERT instead of FTFP_BERT_ATL physics list	No change	No change
No simulation of primaries with pseudo-rapidity above 5.5	0.94	0.95
All range cuts set to 1mm	0.92	0.90

Table 4: Performance of various configurations of the ATLAS simulation for minimum bias and ttbar production events. The Geant4 version used for this test was G4 10.2p03. No significant performance improvements were introduced in patch 03 with respect to patch 02.

ATLAS CPU ratios: cumulative measurements

Configuration	MinBias	ttbar
Nominal production configuration:	1.0	1.0
No shower libraries	1.5	1.3
No shower libraries + use ClassicalRK4 stepper	1.6	1.4
No shower libraries + use ClassicalRK4 stepper + No neutron time cut	1.6	1.4
No shower libraries + use ClassicalRK4 stepper + No neutron time cut + use FTFP_BERT physics list	1.6	1.4
No shower libraries + use ClassicalRK4 stepper + No neutron time cut + use FTFP_BERT physics list + All range cuts set to 1mm	1.5	1.3
No shower libraries + use ClassicalRK4 stepper + No neutron time cut + use FTFP_BERT physics list + All range cuts set to 1mm + No simulation of primaries with pseudo-rapidity above 5.5	1.4	1.2

Table 5: Performance of various configurations of the ATLAS simulation for minimum bias and ttbar production events. Measurements of the cumulative effect of several options to reduce CPU time consumption are listed. The Geant4 version used for this test was G4 10.2p03. No significant performance improvements were introduced in patch 03 with respect to patch 02.

CMS simulation R&D strategy

CMS is following a co-development model with the simulation community

- Tested and adopted VecGeom with G4 10.4 for production < 2 months after release (G4 10.4 is the latest release, available as of early 2018)
  - VecGeom (scalar mode) gave 7-13% speedup
- Successfully integrated GeantV alpha tag in toy framework (CmsToyGV)
  - Goal is verify compatibility of multithread model, test user interfaces, provide feedback

ATLAS is following a much less aggressive strategy

- Running G4 10.2.pXX
- Putting all eggs in the basket of fast simulation techniques (ML in the future) and HPC

What ATLAS does or does not do is relevant because some decisions will be “community decisions”

- I am increasing communication with the ATLAS simulation conveners/experts on R&D matters to try to bring the to experiments closer on future strategies involving R&D products