



CMS Status Report

Maria Girone, CERN
David Lange, LLNL



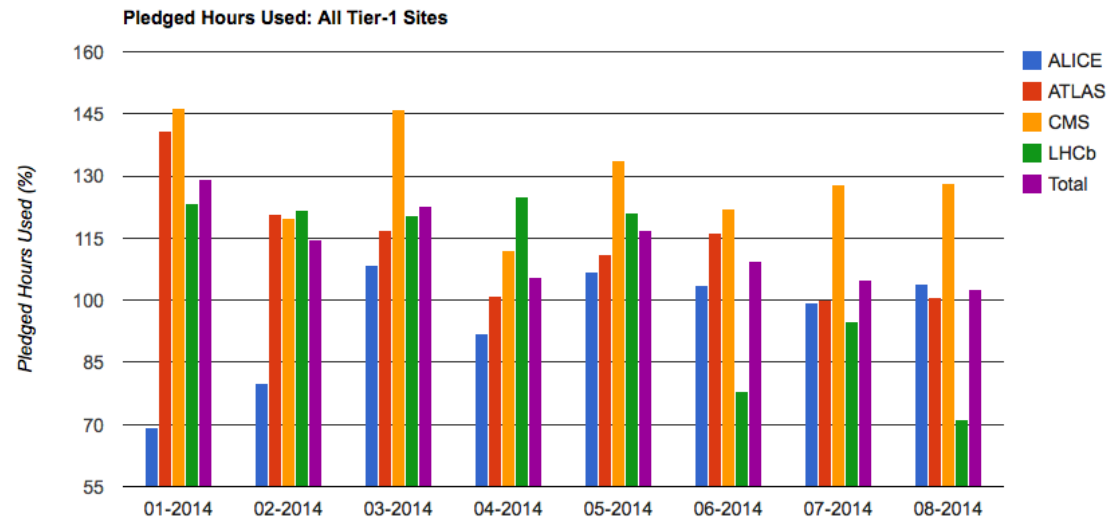
Outline

- Resource usage
- Preparing for Run2: Computing milestones and CSA14
- Software development and production schedule
- Resource requests for 2015
- Planning for 2020+

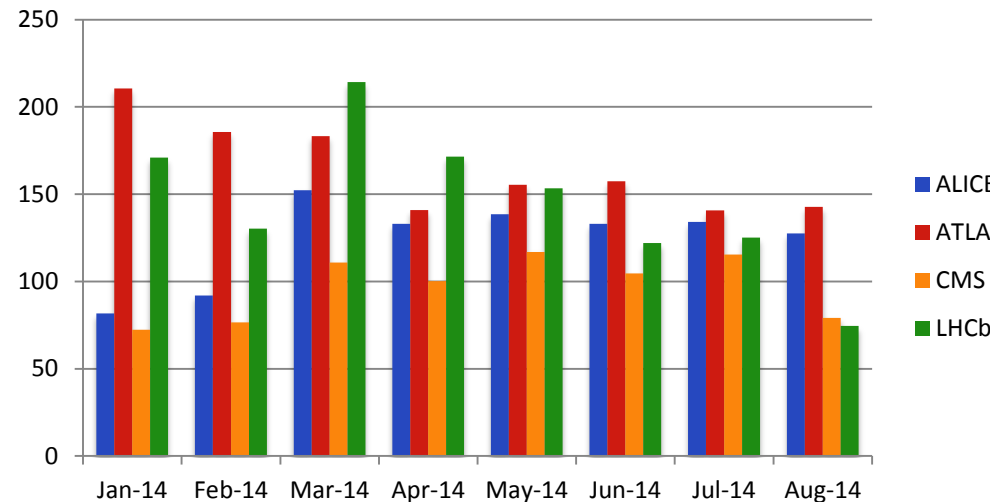


Pledge Usage

- CMS has used the highest percentage of pledged time at Tier-1
 - Many high priority activities are on-going
 - processing of simulation for Run1 analysis
 - production for the CSA14 commissioning challenge
 - preparation of upgrade samples
- Tier-2s have fluctuated around 100% with continued analysis
 - Run2 simulation will ramp up soon



Tier-2 Percentage of Pledge Used





2014 Computing Milestones

An intense testing & commissioning program throughout 2014 for validating changes of the CMS Computing model for Run2

✓ **Data Management milestone: 30 April 2014**

Done!

▪ **More transparent access to the data for users and production**

- Disk and tape separated to manage Tier-1 disk resources and control tape access
- Data federation and data access (AAA)
- Developing Dynamic Data Placement for handling centrally managed disk space

✓ **Analysis Milestone: 30 June 2014**

• **Demonstrate the full scale of the new CRAB3 distributed analysis tool**

Done!

- Main changes are reducing the job failures in handling of user data products, improved job tracking and automatic resubmission

▪ **Organized Production Milestone: 30 November 2014**

▪ **Exercise the full system for organized production**

- Utilize the Agile Infrastructure (IT-CC and Wigner) at scale for the Tier-0
- Run with multi-core at Tier-0 and Tier-1 for data reconstruction
- Demonstrate shared workflows to improve latency and speed



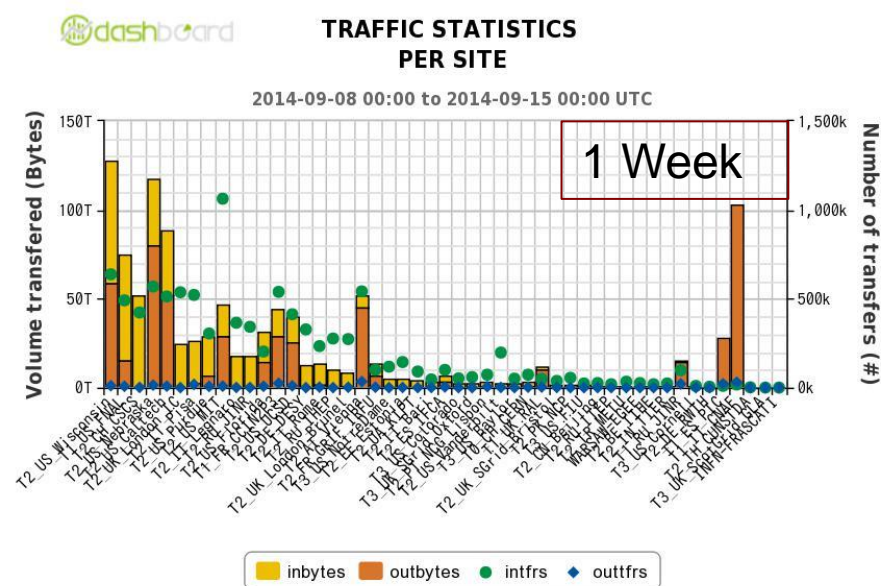
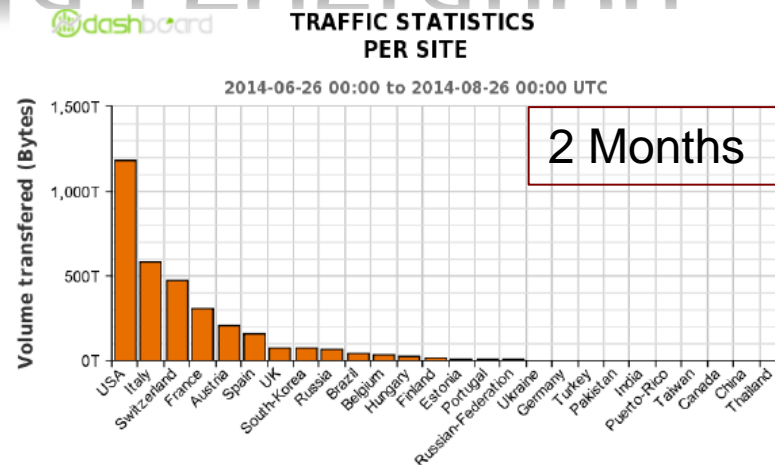
Computing and CSA14

- Summer exercise (Computing, Software and Analysis challenge CSA14) for physics users and developer community: >150 users involved, started 2nd week of July, wrapping up now
 - **For Computing and Offline this involved**
 - Commissioning of a new mini-AOD format (nearly a factor of 10 smaller than the Run 1 AOD while targeting 80% of analysis users)
 - Commissioning of the data federation (AAA). The target is 20% of analysis access served over the wide area
 - Commissioning of the production workflows (SIM, DIGI, RECO)
 - Commissioning of the new analysis submission tool (CRAB3)
 - Production of the samples for the challenge
 - Production and reconstruction of ~2 billion 13TeV events under various pile-up scenarios for 25ns and 50ns running conditions
- Good opportunity to identify remaining weaknesses of the production system under high-load, which are currently being addressed



Data Federation

- During CSA14, AAA has been used to access data over the WAN via CRAB3
- Target: 60k files/day, O(100TB)/day
- Goal of the data federation **scaling tests** during the last few weeks of CSA14 has been to exceed 20% of analysis access
 - Phase 1: USA only (see plot)
 - Average incoming rate is 30TB/week for the 7 US Tier2 sites (40MB/s)
 - Phase 2: add Europe – starting now
 - AAA deployment in Europe started later and we are expecting to have to invest more in validation





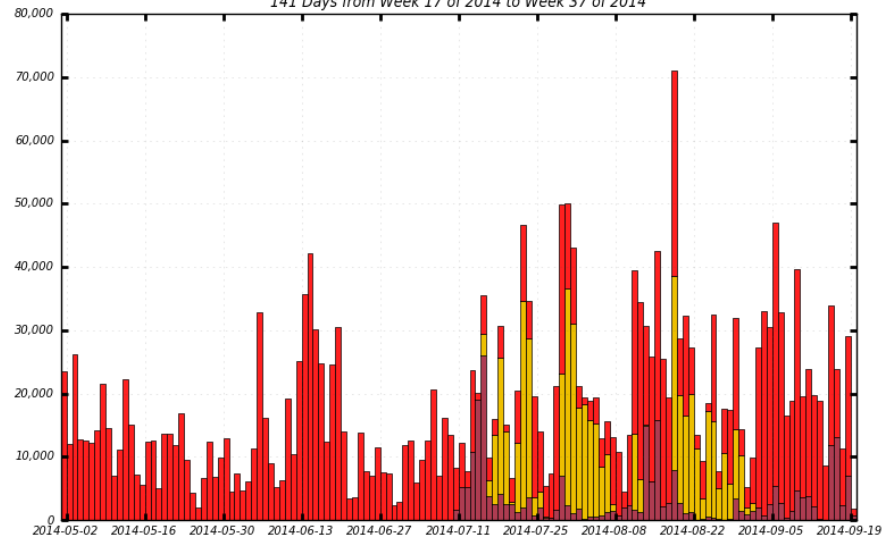
CRAB3

- CRAB3 usage remains a small fraction of the total user submission
 - CSA14 users were asked to use CRAB3
 - We further test operations at scale with Hammer Cloud
- User-driven MiniAOD production was done with CRAB3
 - Initial feedback from users is VERY positive!
 - User production was able to keep pace with central MiniAOD production
- We are developing a CRAB3 adoption plan before the beginning of Run2

dashboards

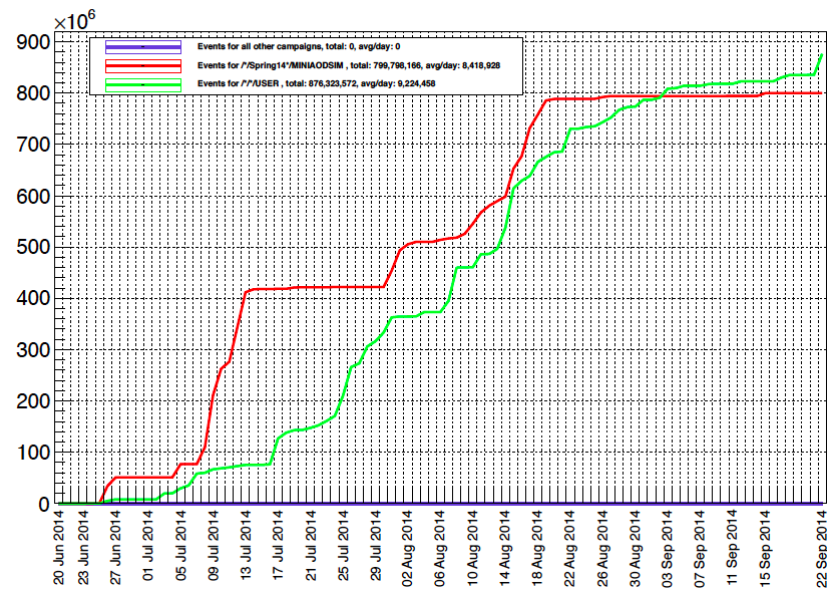
Submitted jobs

141 Days from Week 17 of 2014 to Week 37 of 2014



ar

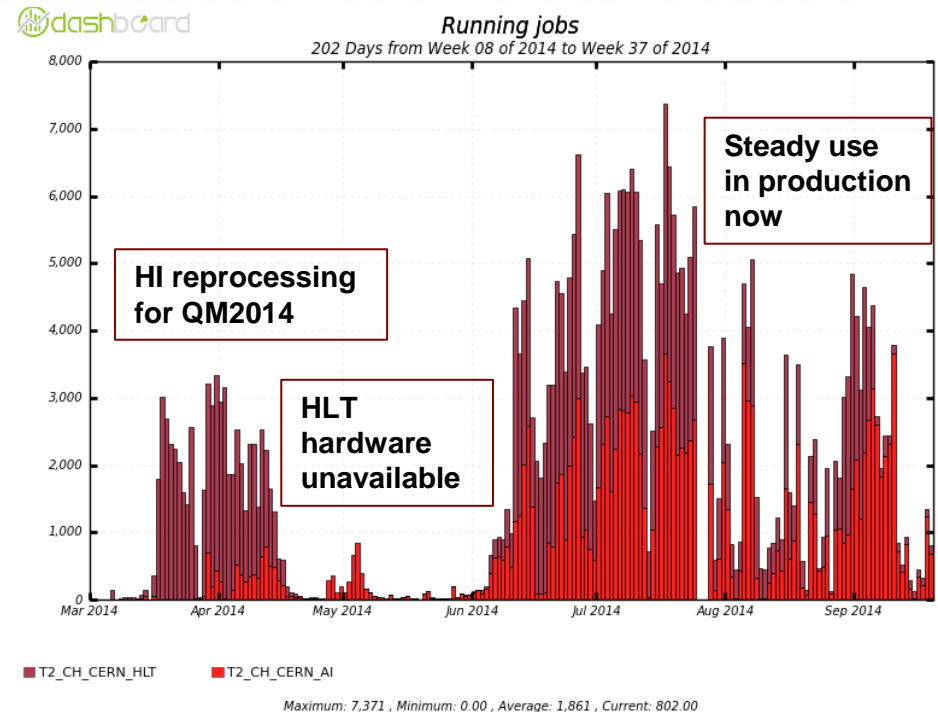
Overview for DBS-status VALID+PRODUCTION, total: 1,676,121,738, avg/day: 17,643,386





Cloud and AI Testing

- CMS has been using the HLT for production activity for all of 2014
 - Began also using the CERN AI for production and Tier-0
 - We have demonstrated up to 5k simultaneous jobs



- Large scale tests are scheduled for later this fall pending that sufficient AI resources are available
 - Short full-scale tests need to be early enough to understand and solve problems identified before the run starts
 - Discussing with IT on how to achieve this goal



Software

We are ready to begin production of Monte Carlo samples for 2015 startup:

- All changes to CMS detectors during shutdown reflected in geometry
- An improved material model for tracker implemented
- The detector simulation is updated to Geant 4.10
- **Technical improvements over past year have resulted in a factor of 2 speed improvement**

CSA14 exercise was first large scale test of recent CMSSW developments

- Simulation, digitization and reconstruction applications proved to be robust in processing of ~1.5 billion events
- Successfully integrated and deployed a new analysis data format: “MiniAOD”
 - **~10x reduction in size relative to AOD achieved through a number of approaches**
 - **Plan to produce the MiniAOD for 2015 in two ways:**
 1. Together with the AOD format in the reconstruction workflow
 2. In a dedicated conversion workflow from the AOD to also include all available physics object recipes needed for analysis



Software

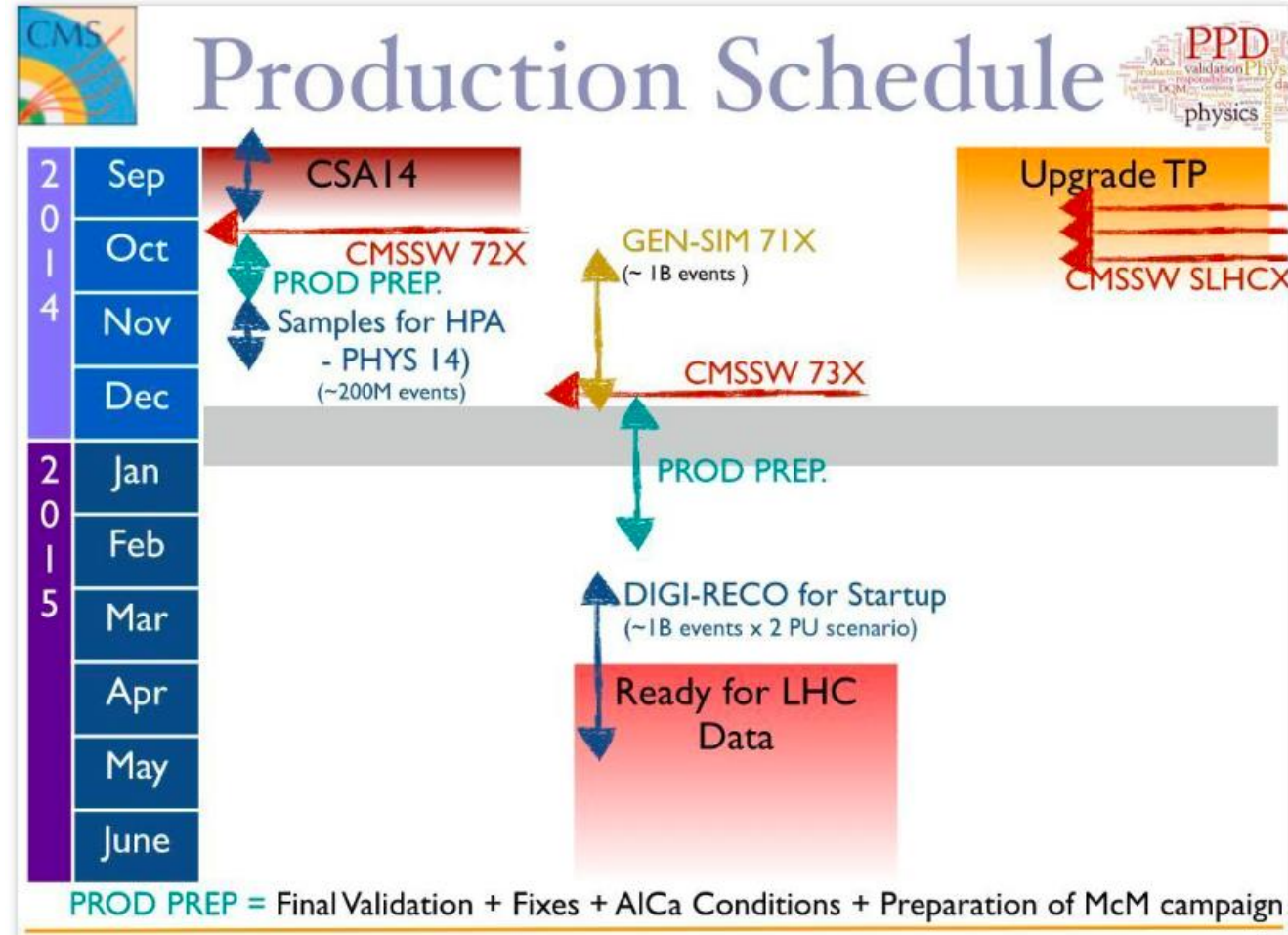
- The CMSSW release for startup reconstruction is on track to be complete in December (“CMSSW_7_3_0”)
 - We are now finishing an intermediate release of CMSSW that includes the first set of algorithmic changes for 25ns bunch spacing operations (for the “PHYS14” exercise)
 - Excellent multi-threaded performance of the reconstruction has been achieved over the summer
- We continue to support releases for Phase-I and Phase-II upgrade development work in dedicated releases
 - Even very high pileup workflows run within usual GRID resource constraints
 - Successful demonstration of recent developments in our MinBias event overlay algorithm and the flexibility of our tracking algorithms
 - Time/event is a challenge at 140-200 PU, but our tools have allowed us to keep up with rapid detector and reconstruction algorithm changes
 - We have started the process of integrating the code developed for Technical Proposal software developments into our mainstream CMSSW releases



Preparation Schedule

Current CMS schedule calls for 2B events prepared in advance for 2 PU scenarios (one for 50 ns and one for 25 ns)

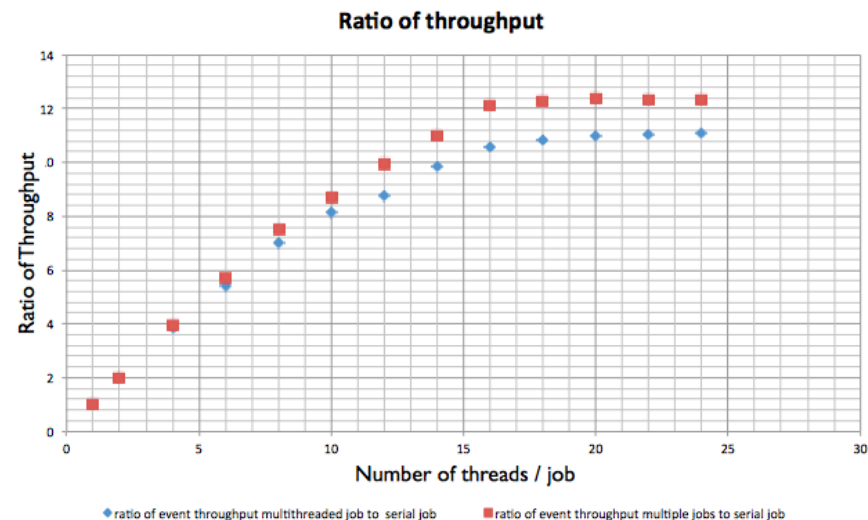
- 1B Gen-Sim will start soon!
- PHYS14 exercise being planned
 - Prepare high priority analysis for speedy results on first 1-5 fb⁻¹
- Run2 production release CMSSW_7_3 by December





Multi-threaded FWK Deployment

- CMS plans to switch to multi-core processing for reconstruction by the beginning of Run2
 - Better I/O behavior, less merging, better memory footprint, fewer processes to track, and generally the direction for the future
 - Multi-core queues now exist at all the Tier-1 sites



[Details presented at ACAT](#)

- CMS has already achieved >99% parallel safe code and has excellent efficiency up to 8 cores
 - Computing is expected to receive CMSSW_7_2 as a production release in October



Resource Requests

- CMS is currently in the interactions with the Computing Resource Scrutiny group
 - **The request endorsed for 2015 has very little contingency**
 - We are continuing to optimize the Computing system to fit within the resource envelope
 - **Setting priorities and making hard choices!**
 - Reduced the number of reprocessing passes
 - Reduced the ratio of Monte Carlo events to the number of real data events (ratio of 1.3 AND the bulk can be reprocessed only once)
 - No contingency for mistakes
 - **This interaction with C-RSG is a first look at 2016**
 - Small changes expected in the instantaneous luminosity and a big increase in the live seconds of running expected wrt 2015



Preliminary Resource Request for 2016

- The largest requested increase is in Tier-2 CPU
 - The total volume of data to analyze drives the increase

- All quantities except the Tier-0 are predicted to grow faster than technology evolution
 - Some optimization is hopefully still possible with experience from the first year of Run2

	2014	Increase from 2013	2015	Increase from 2014	2016	Increase from 2015	2017	Increase from 2016
Tier-0 CPU (kHS06)	121	0%	256 (256)	111%	302	18%	350	18%
Tier-0 Disk (TB)	7000	0%	3250 (3000)	Reallocated to CAF	3250	0%	3250	0%
Tier-0 Tape (TB)	26000	0%	31000 (31000)	31%	38000	23%	50000	31%
CAF CPU (kHS06)	0	0%	15 (15)	-	15	17%	17	21%
CAF Disk (TB)	0	0%	12100 (12000)	-	13100	8%	14000	7%
CAF Tape (TB)	0	0%	4000 (4000)	-	6000	50%	8000	33%
T1 CPU (kHS06)	175	0%	300 (300)	71%	400	33%	525	31%
T1 Disk (TB)	26000	0%	27000 (26000)	4%	35000	30%	45000	28%
T1 Tape (TB)	55000	11%	73500 (74000)	34%	100000	36%	135000	35%
T2 CPU (kHS06)	390	14%	500 (500)	25%	700	40%	800	14%
T2 Disk (TB)	27000	4%	31400 (29000)	16%	40000	27%	48000	20%



Looking forward to Upgrades

- CMS is facing a huge increase in the **scale of the expected computing** needed for Run4

Phase	Pile-Up	HLT Output	Reconstruction time ratio to Run2	AOD Size ratio to Run2	Total Weighted Average Increase above Run2
Phase I (2019)	50	1kHz	4	1.4	3
Phase II (2024)	140	5kHz	20	3.7	65
Phase II (2024)	200	7.5kHz	45	5.4	200

- The WLCG Computing Model Evolution document predicts 25% processing capacity and 20% storage increase per year
 - Factor of 8 in processing and 6 in storage between now and Run4
 - Even assuming a factor of 2 code improvements the deficit is 4-13 in processing and 3-7 in storage



Targets for improvements

- It is unlikely we will get a factor of 5 more money, nor will the experiment be willing to take a factor of 5 less data
 - **Big improvements are needed**
- Roughly 40% of the CMS processing capacity is devoted to task identified as reconstruction
 - **Prompt reconstruction, re-reconstruction, data and simulation reco**
 - Looking at lower cost and lower power massively parallel systems like ARM and high performance processors like GPUs (Both can lower the average cost per processing)
- ~20% of the offline computing capacity is in areas identified as selection and reduction
 - **Analysis selection, skimming, production of reduced user formats**
 - Looking at dedicated data reduction solutions like event catalogs and big data tools like Map Reduce
- The remaining 40% is a mix
 - **Lot of different activities with no single area to concentrate optimization effort**
 - Simulation already has a strong ongoing optimization effort (in Geant4 and CMS)
 - User analysis activities developed by many people
 - Smaller scale calibration and monitoring activities



Organizational Changes

- CMS decided to merge the Offline and Computing organization over the next year (complete by Sept. 2015)
 - Does not put an artificial distinction between these closely related activities
 - Aligns CMS with the other LHC experiments
 - Makes communication lines clearer from outside
- CMS appointed Maria Girone and David Lange to co-lead the new joint project
 - The sub-projects will evolve in the new structure over the next year
 - We hope that the enlarged scope will strengthen the joint project and create new opportunities



Outlook

- CMS is using the long shutdown to improve the efficiency of our software and computing system
 - Data management and data access improvements are in commissioning. Production improvements will continue through the fall. Analysis improvements are in user testing
 - We are pushing forward the adoption of CRAB3 and pushing the testing the organized processing environment, including AI, to a reasonable scale
 - Streamlining software and computing workflows as we expect at best constant level of effort
- In addition to trying to handle day-to-day operations and prepare for Run2, we are assembling projects to do the innovative development needed to close the future resource gap in Run4
 - Offering leading-edge techniques may result in the injection of new people
 - An excellent area for collaborative work