Science and Technology Facilities Council

CMS

Katy Ellis, GridPP52, 29 August 2024

CMS status

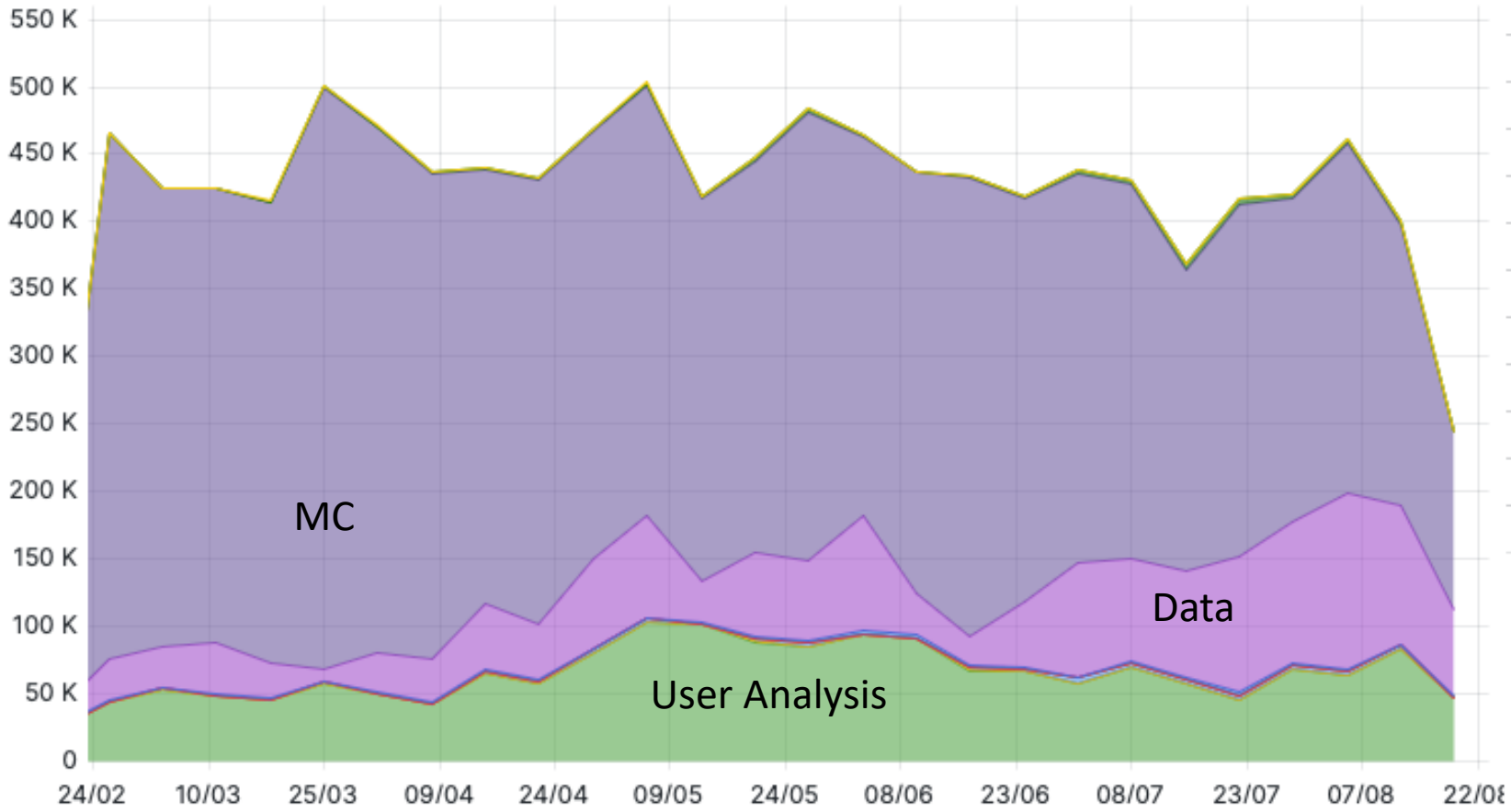
Global view

Job mix in last 6 months (all sites)

Total completed jobs

	total	percentage
Analysis	88862895	50.6%
Production	61564259	35.1%
Merge	6958176	4.0%
Cleanup	6758376	3.9%
Processing	4431810	2.5%

Running cores



Name	Mean
Analysis	65.2 K
Cleanup	73.0
Express	1.47 K
Harvesting	4.44
LogCollect	33.6
Merge	1.30 K
Processing	56.2 K
Production	297 K
Repack	1.70 K
Unknown	160
TOTAL	423 K

Most Analysis jobs are single core; most central production jobs are multicore

Token status

- Most sites are passing IAM-authorized SAM tests on their storage
  - Some sites waiting on new EOS version
- However, token transfers did not remain in production since DC24
  - CMS are currently testing a new version of IAM
  - Hopefully this can be quickly rolled out again to many sites
- Many sites are also passing IAM-authorized SAM tests on their CEs
- Several internal parts of the system are already token-enabled
- Lots more information here: https://twiki.cern.ch/twiki/bin/viewauth/CMS/IAMTokens (CMS members only)

Shoveler - Monitoring for XRootD

- Katy decided to take on a project to test and validate for CMS
  - Making some progress
  - Last 3 weeks would have been better if the operator at CERN hadn’t been on holiday ;-) Most issues are related to Grafana and OpenSearch
  - Basic tests are positive - thanks to Alex R for help with vector reads!
- Shoveler has been monitoring AAA proxy machines at RAL T1 for a couple of years
  - Recent progress -> Starting to monitor the WNs (Echo reads by XRootD)
  - Thanks to Jyothish and Tom Birkett for all the help!
- Several presentations already lined up
  - Last week’s CMS O&C weekly
  - XRootD workshop
  - CHEP

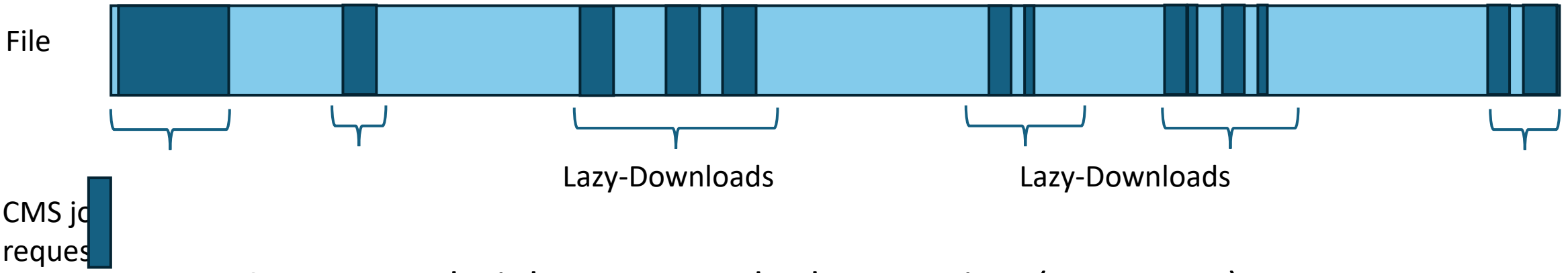
CMS@UK sites

UK T2 site status

- RALPP: Currently preparing to make part of the GPU smallholding accessible to the Grid
  - Three nodes each with two NVidia A100 cards
- Brunel: Migration from retiring storage software, DPM, to Ceph was troublesome and time-consuming
  - Required a lot of consultation with CMS support and UK admins
  - Much of the batch farm unused during this time
- Bristol: Currently out of action for complete storage re-design
  - Person-power issues are a major blocker

RAL T1 – removal of Lazy-Download

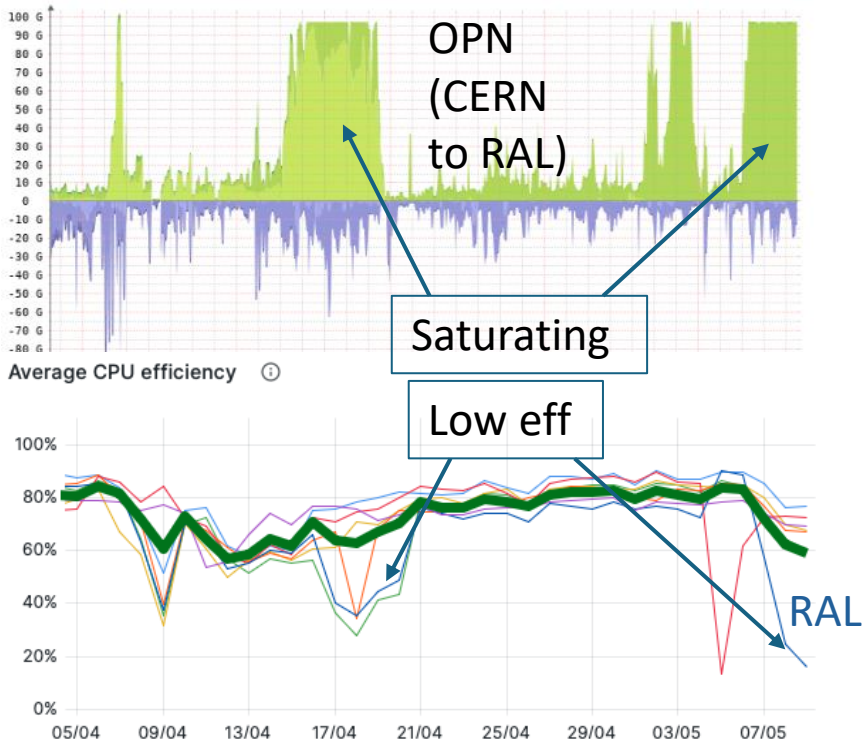
- CMS jobs stream the part of the data they need direct from storage, whether locally or remotely (e.g. via AAA)
  - Lazy-Download copies larger chunks of data containing the needed parts
  - Was required at RAL Tier 1 (and other sites) in the past where the storage could not handle lots of tiny reads

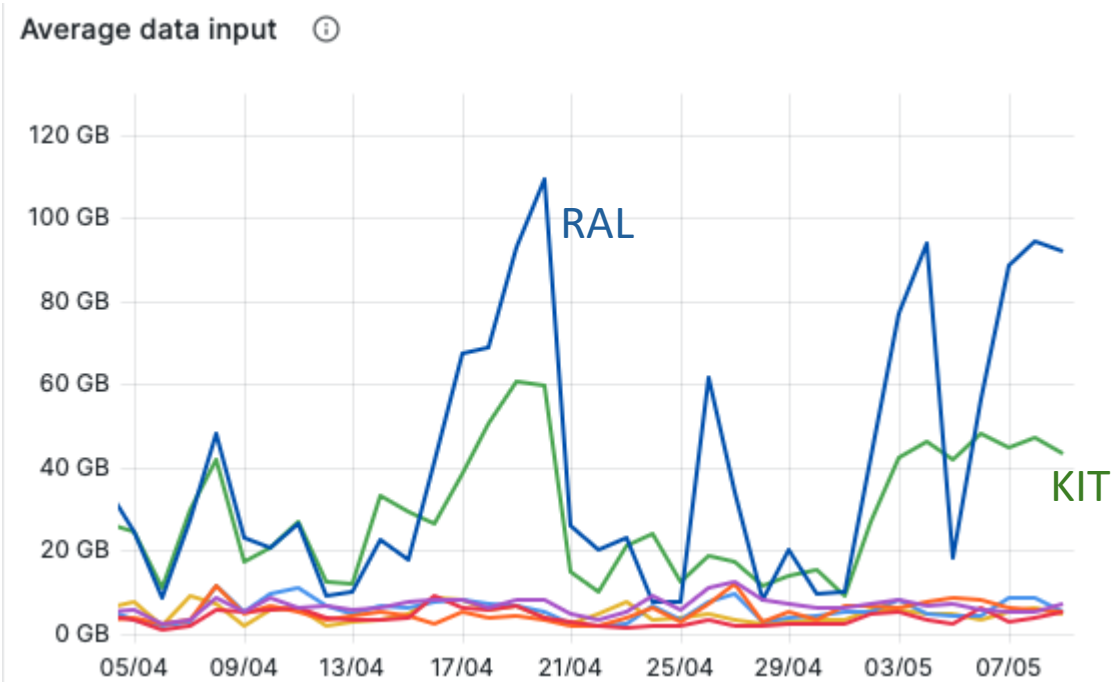


- Consequently, jobs at RAL and other L-D sites (e.g. KIT T1) can, at times, read significantly higher data volume

RAL T1 – removal of Lazy-Download

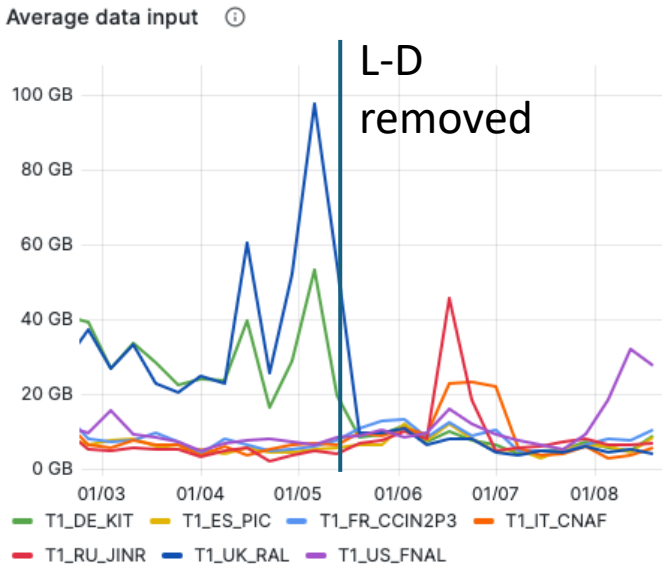
- L-D configured so that if you need it for local reads, you get it for remote reads too
  - CMS jobs regularly blamed (rightly) for filling the network from CERN to RAL batch farm. Starts to have effect on job efficiency, especially when core usage is high

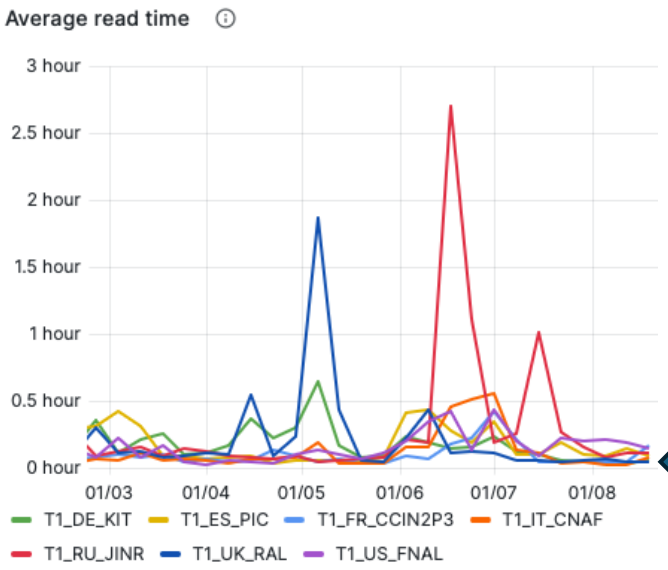




RAL T1 – removal of Lazy-Download

- Recent performance at T1 has been fantastic





RAL

	CPU Eff	Failure rate
T1_US_FNAL	71.9%	19.89%
T1_UK_RAL	80.4%	6.11%
T1_RU_JINR	74.7%	8.05%
T1_IT_CNAF	76.2%	9.05%
T1_FR_CCIN2	81.1%	23.47%
T1_ES_PIC	67.7%	41.36%
T1_DE_KIT	77.8%	9.34%

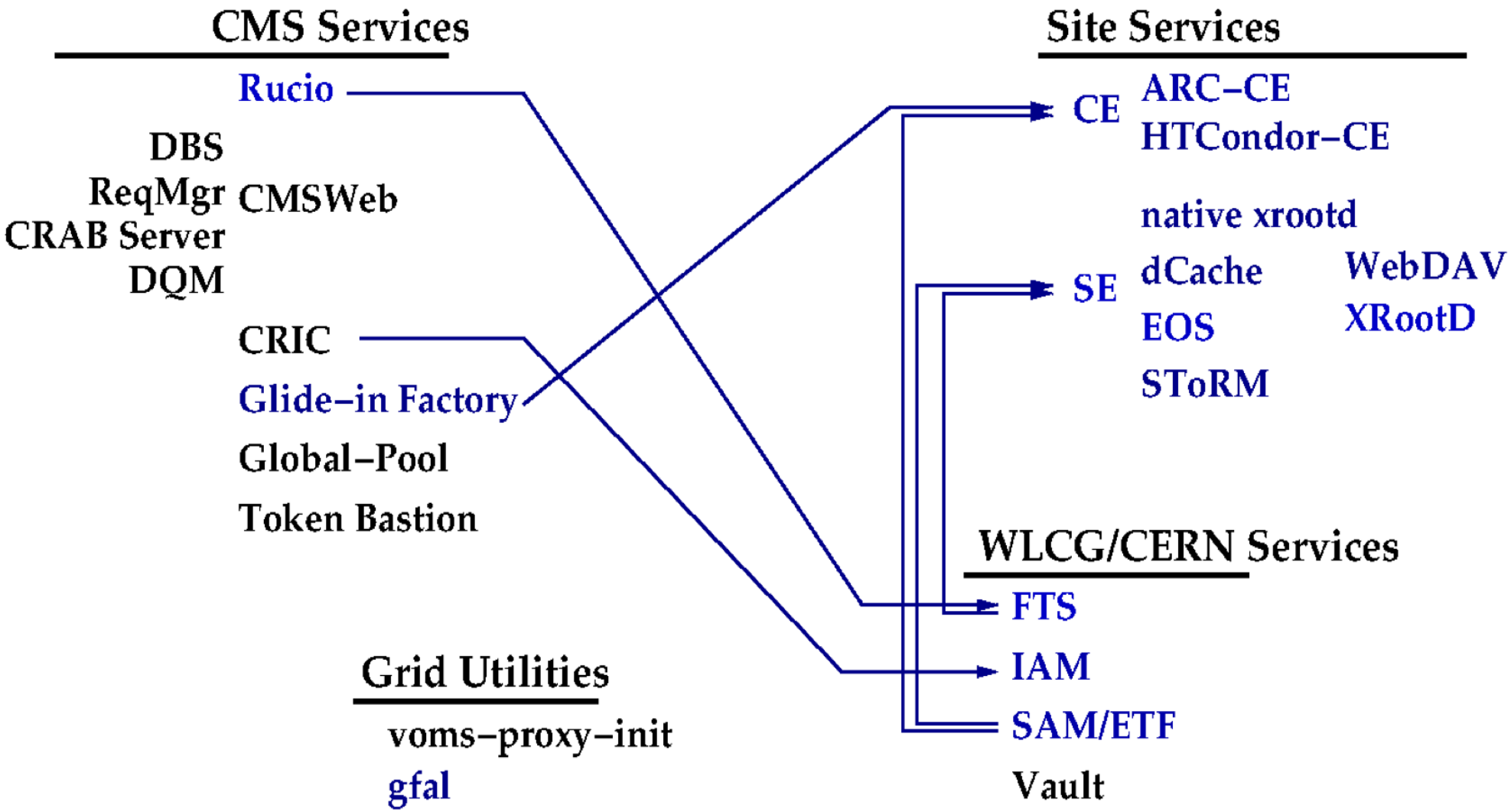
CMS computing insights

Stolen from the Offline & Computing F2F Management Meeting

Token roadmap

- Stating how CMS will use tokens all over our infrastructure
  - Both internal to the system, and User-facing
  - See next slide
- How we will protect tokens of different value
- How we will maintain our User information and group membership
- The planned schedule of the roll-out
- https://twiki.cern.ch/twiki/bin/viewauth/CMS/IAMTokens

Token interactions



Review of review of Workload Management

- Last formal review took place in 2018-2019
  - Involving a committee and input from our experts
  - Already looking forward to HL-LHC
  - Scope included CRAB software (User analysis)
  - Made a number of recommendations

Review of review of Workload Management

- Last formal review took place in 2018-2019
  - Involving a committee and input from our experts
  - Already looking forward to HL-LHC
  - Scope included CRAB software (User analysis)
  - Made a number of recommendations
- Attempting to answer several questions, such as:
  - How will our software scale for HL-LHC?
  - How can we develop and maintain the software most efficiently?
  - Possibility for joining ‘community projects’? (akin to Rucio)

Review of review of Workload Management

- Last formal review took place in 2018-2019
  - Involving a committee and input from our experts
  - Already looking forward to HL-LHC
  - Scope included CRAB software (User analysis)
  - Made a number of recommendations
- Asking several questions, such as:
  - How will our software scale for HL-LHC?
  - How can we develop and maintain the software most efficiently?
  - Possibility for joining ‘community projects’? (akin to Rucio)
- What was achieved following the review?

Review of review of Workload Management

- Last formal review took place in 2018-2019
  - Involving a committee and input from our experts
  - Already looking forward to HL-LHC
  - Scope included CRAB software (User analysis)
  - Made a number of recommendations
- Asking several questions, such as:
  - How will our software scale for HL-LHC?
  - How can we develop and maintain the software most efficiently?
  - Possibility for joining ‘community projects’? (akin to Rucio)
- What was achieved following the review?
- Discussion of scope for a new review

Possibility for a CSA?

- What is a CSA?
  - Computing, Software and Analysis challenge!
  - Hopefully involving groups outside of O&C – e.g. trigger and physics analysis
- Why do we need one?
  - To exercise our computing model in advance of HL-LHC
    - Many things will change in our software
- When?
  - 2027 and/or 2028
  - Will major changes be in place by then?

Examples of major new components

- New CMSSW software components
  - Detector upgrades:
    - All new tracker raw/trigger/reco/sim?
    - All new tracking?
    - All new forward calo/PF raw/trigger/sim/reco?
    - All new L1 trigger emulation?
  - Alpaka-based plugins?
  - RNtuple and fully defined data formats for commissioning and analysis?
  - Modernized G4 components?
- Production
  - GPU aware MC production ?
  - Tier-0 workflow and datasets?
  - Processing Units?
- Calibration workflows
  - New detector workflows?
  - Increased integration w/ Tier-0?
- Data handling automation
  - ML driven certification?
- User Analysis
  - Analysis Facilities?
  - Workflows for on-tape tiers?

Other topics

- RNTuple migration
- Data format efficiency
- Calculating how much MC is needed for HL-LHC
- The Future of FastSim (with respect to FullSim and FlashSim!)

Summary

- CMS computing is busy processing (real) data
- Thinking carefully about the challenges of HL-LHC within constraints
- Another performance jump at Tier 1 for computing jobs!
- Business-as-usual at the Tier 2s