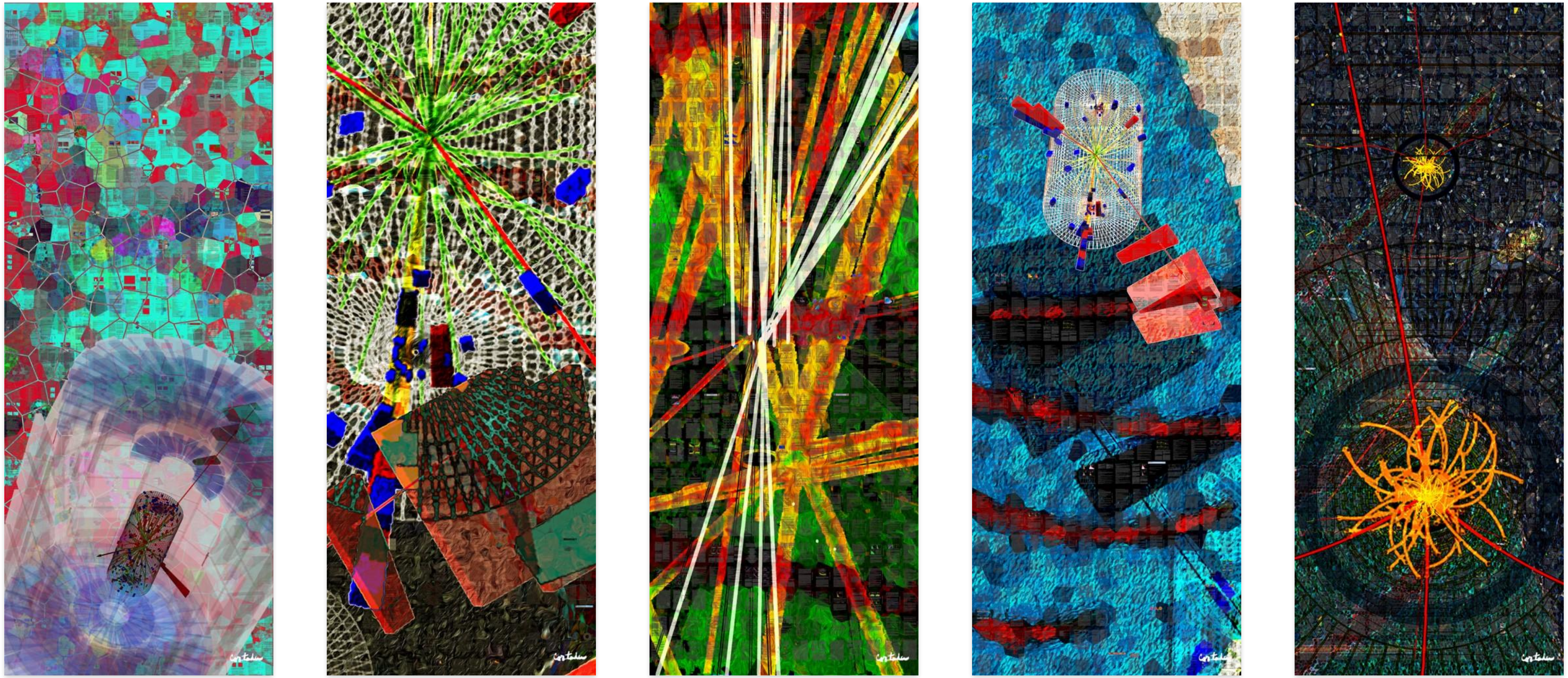
U.S. CMS
Operations
Program

U.S. CMS R&D activities and IRIS-HEP

Ken Bloom & Oliver Gutsche
U.S. CMS Software & Computing Operations Program



Introductory comments

- The HL-LHC data volume presents challenges to CMS computing that must be addressed by a broad array of R&D efforts to ensure that physicists can succeed in extracting scientific results from this multi-MCHF project at reasonable computing cost
- Driving problems and areas of exploration (from our perspective):
  - CPU budget will be driven by event reconstruction needs
    - Need to improve performance on all architectures, expanding beyond x86
  - Storage budget will be driven by the trigger rate
    - Need to revise model of disk/tape/network usage to optimize costs
  - Analysis will be performed with less physicist effort
    - Need to identify technologies that will reduce effort needed for analysis
- In addition, expect that funding agencies are planning on providing a sizeable fraction of HL-LHC computing via HPC centers -- won't look like the grid, each system is different

Introductory comments

- To address these challenges, U.S. CMS Software and Computing Operations supports a number of R&D efforts
  - In some cases, partner with other funded efforts to jointly support projects of mutual interest
    - Many of our partners are sitting in this room!
  - In other cases, provide seed funding to start a project until it develops enough to be able to attract external funding
  - In general, supporting fractions of people working on each individual project
- U.S. CMS R&D efforts cannot be too “blue sky” -- need to have clear potential for payoff to CMS
  - Projects descriptions to follow are targeting multiple timescales, not just HL-LHC
- We expect that specific projects will evolve over time as products get launched or efforts get funded by someone else or priorities evolve
- See extra slides for many more details about the projects

Near-term and ongoing efforts

- ROOT development
  - The language of HEP, CMSSW depends on it, THE analysis tool of CMS
  - Goals: Improve ROOT I/O on HPC machines, modularize to support external analysis toolkits, elevate Python to first-class language
  - Support leadership for I/O improvements, pyroot in the ROOT team
- Visualization support/development
  - Event display is U.S. CMS responsibility, community contribution to maintain EvE in ROOT
  - In longer term, move to client-server setup based on community solutions, evolve EvE
- HL-LHC resource estimates
  - Important to understand expected resource needs as they will drive the R&D program
  - Difficult multi-dimensional modeling problem, agencies always ask us about this
- HEPCloud
  - New capabilities for CMS via elastic expansion of resources and a homogeneous interface to heterogeneous resources including commercial clouds and HPC centers
  - With Fermilab SCD, supporting integration of HEPCloud into CMS submission infrastructure (SI) and workflow management (WM) systems, and integration of different resources into HEPCloud

Evolutionary more than revolutionary

- Scaling SI and WM systems for HL-LHC
  - No matter what new resources are available, scales expected for HL-LHC will challenge WM and SI
  - We will always be HTC, and need to make sure that we can use all future resources effectively
  - Continue already substantial U.S. CMS efforts on these systems
- Storage infrastructure for HL-LHC
  - Data handling will be a major challenge, need to explore ways that we can improve existing systems to improve efficiency and reduce costs
  - “Data lake” as a natural evolution of data federations
  - Large-scale managed caches to increase availability of popular datasets without over-replication
  - Support a variety of activities in this area, including transition to Rucio and exploration of alternative storage technologies (e.g. Ceph)

Potentially revolutionary explorations

- Future architectures
  - This could help us reduce CPU requirements and take full advantage of HPC centers
  - These will require massive multi-threading and vectorization
  - Within CMS, strengthening effort to bridge between CMSSW framework and physics algorithms
  - Also supporting external R&D efforts, especially on tracking, through integration with CMSSW, scaling and profiling, consulting and testing
  - Other related work: efficient access to ML frameworks, improving pre-mixing for HL-LHC detectors, prototype integration of vectorized transport engine of GeantV
- Analysis tools and machine learning
  - Alternative analysis concepts could potentially reduce resource needs significantly, and allow large datasets to be analyzed quasi-interactively to speed up analysis
  - Supporting R&D for integrating community Python tools and ML into CMSSW ecosystem, enabling alternative analysis tools in root
    - Apache Spark, uproot, machine learning benchmarks….
  - This R&D allows us to develop concepts for a future analysis facility

Comments for IRIS-HEP

- We are very interested in SSL approaches, also looking forward to learn and collaborate
- We have a long history of working with OSG in both operational and development areas; we believe it has been successful and look forward to continuing
- We are supportive of finding common solutions wherever possible, but we recognize it's not always possible
    - Different detectors have different designs and will require different optimization of algorithms
- IRIS-HEP is paying attention to interaction with DOE entities, but we should be thinking about international partners
    - Because these experiments are international projects!
    - HSF? WLCG? CERN R&D plans?
- We are looking forward to many future collaborations with IRIS-HEP!

Extra slides with many more details

Future Architectures

- Strategy
  - Future architectures will require massive multi-threading and vectorization
  - Strengthening of Core Software effort (CMSSW) aimed at bridges between physics algorithm development and core framework development
  - Program supports external R&D efforts through integration into CMSSW, scaling and profiling, consulting and testing
  - Short term goal: continue support of external tracking reconstruction projects
- Justification
  - Progress in vectorizing cpu-intensive reconstruction algorithms could reduce CPU requirements for HL-LHC significantly, enable to use new HPC centers with specialized architectures in the first place
- Support for
  - Integration of new architecture software improvements into CMSSW (generic CMSSW interface to GPUs, FPGAs, ...)
  - Efficient access to machine learning frameworks for training and inference
  - Adding, refining and optimizing pre-mixing for HL-LHC detector components
  - Prototype integration of vectorized transport engine of GeantV
  - Integration of vectorized tracking improvements into CMSSW

Analysis Tools & Machine Learning

- Strategy
  - HL-LHC analysis challenge: data volumes significantly larger compared to LHC
  - Machine learning: industry techniques very attractive for HEP analysis
  - Supporting R&D by integrating community python tools and machine learning tools into CMSSW ecosystem, enabling the CMS community to use alternative analysis tools to ROOT and use machine learning tools
  - Short term goal: study analysis facility components based on industry tools
  - Medium and long term goals: support development of analysis facility concepts and large scale machine learning applications
- Justification
  - Alternative analysis concepts could potentially reduce the needed resources for HL-LHC analysis significantly, as well as speed up analysis so that large datasets can be analyzed quasi-interactively
- Support for
  - Alternative Analysis Tools like Apache Spark in collaboration with DIANA, CERN openlab and FNAL LDRD
  - Machine Learning integration in collaboration with DIANA
  - Machine Learning Architecture Benchmark in the context of the CMS Machine Learning Forum
  - Machine Learning and Analysis Tools integration into CMSSW ecosystem

ROOT

- Strategy
  - ROOT is a major dependency of CMSSW, and THE analysis tool of the CMS collaboration
  - Short term goal: improve ROOT I/O to work on HPC machines
  - Medium to long term goal: Modularize ROOT to support external analysis toolkits (bridges and ferries) and elevate python to first class language
- Justification
  - ROOT is our community analysis and I/O toolkit, investment in ROOT is needed to maintain and evolve its functionality for CMS
  - Especially improvements in the I/O stack are needed for high thread count execution, for example on HPCs
- Support for
  - Leadership of I/O improvement effort in the ROOT team
  - Python elevation in CY19 (pyroot) in the ROOT team

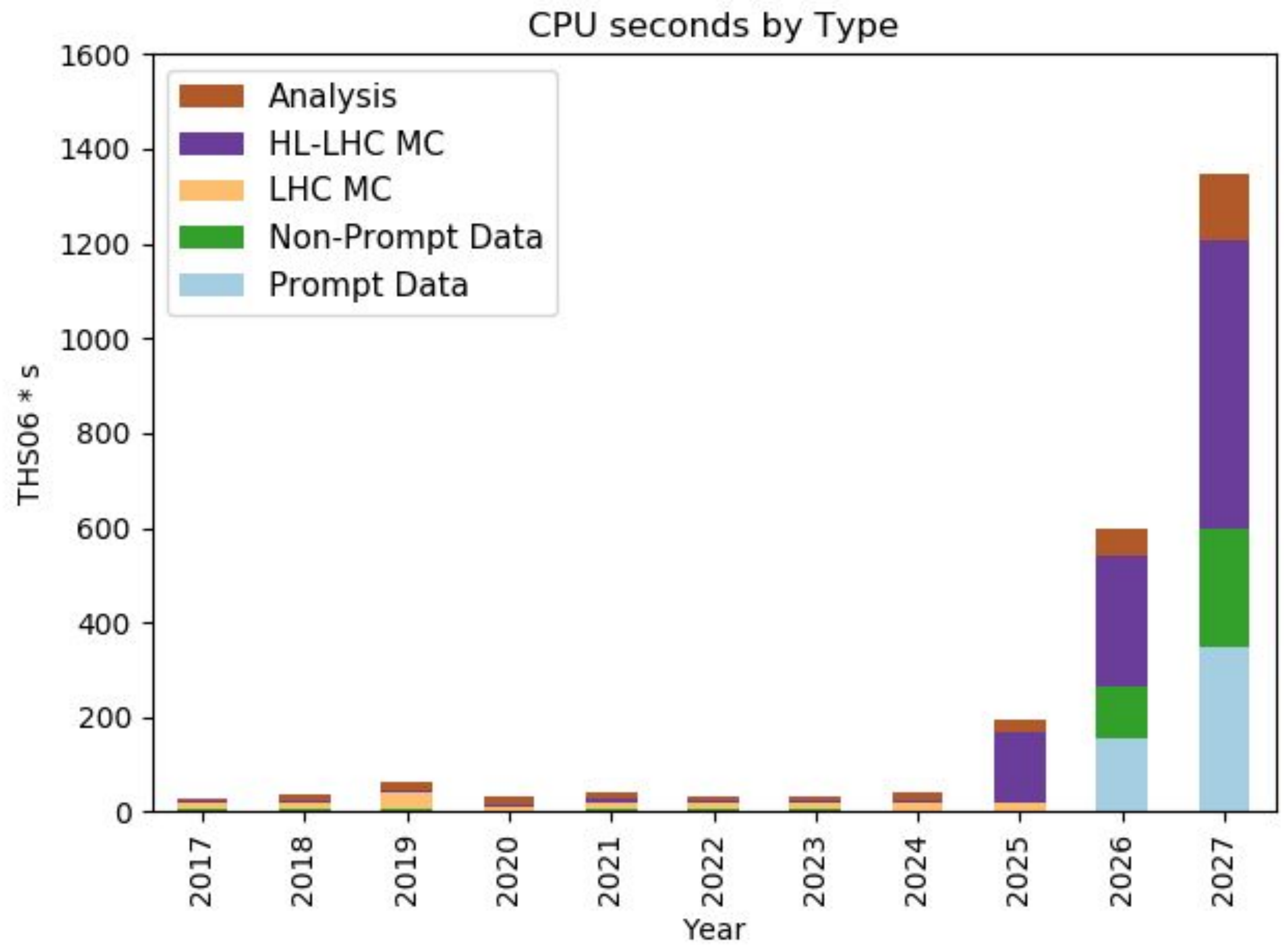
Visualization

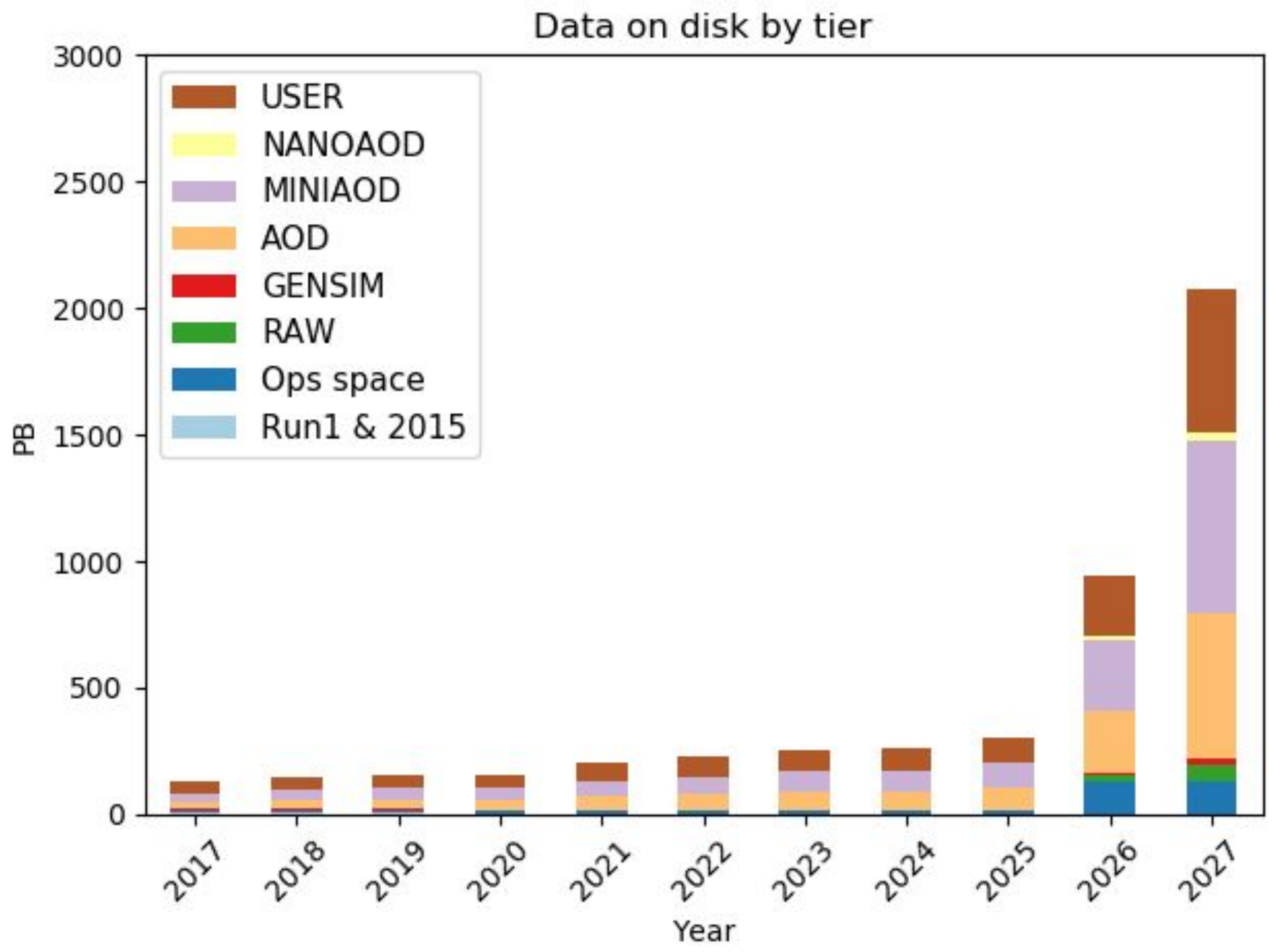
- Strategy
  - U.S. CMS responsibility: Maintain and evolve the Fireworks event display solution
  - Includes community contribution to maintain EvE, the event visualization framework of ROOT
  - Short term goal: support Fireworks for data taking and analysis
  - Medium to long term goal: move to client-server setup based on community solutions, evolve EvE in the light of ROOT7 graphics changing to web-based solutions
- Justification
  - U.S. CMS committed to maintain and evolve Fireworks as the event display solution for CMS
  - CMS relies on the solution not only for user analysis, but also relies on it during data taking at P5 and CMS centers
- Support for
  - Prototype of client-server setup
  - EvE7 studies and prototypes to support EvE7 in ROOT

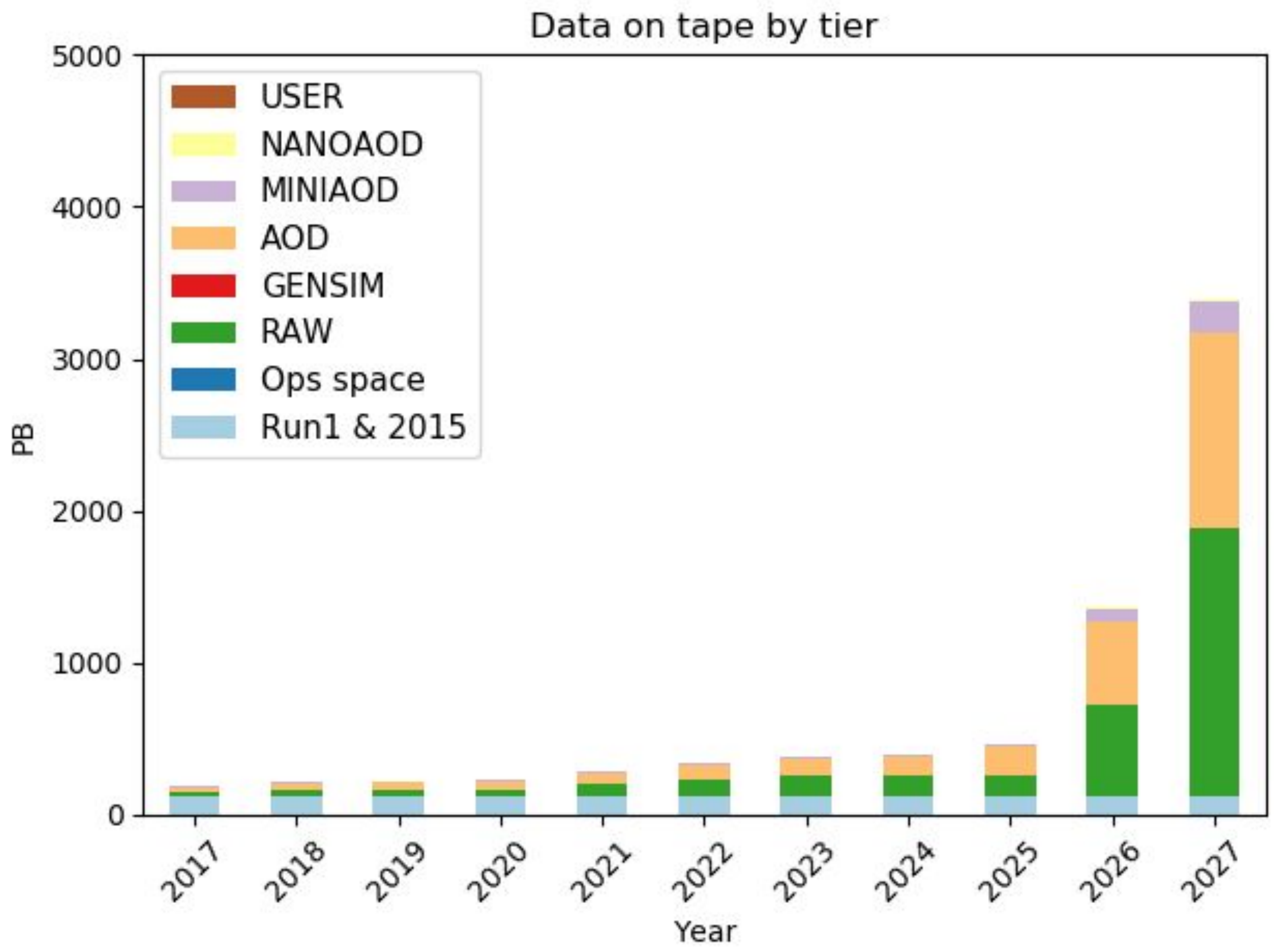
HEPCloud

- Strategy
  - Provide new capabilities to CMS
  - Elastic scaling to commercial clouds and HPC allocations
  - Provide decision engine capabilities to steer workflows depending on their requirements
- Justification
  - Elasticity could relief pressure on CMS release/validation/processing process
  - New resource forms like commercial clouds and HPC centers are a necessity to integrate into the resource mix for HL-LHC
- Support for
  - Integration of HEPCloud into the CMS submission infrastructure and workflow management systems together with the FNAL Scientific Computing Division
  - Integration of resource forms into HEPCloud together with the FNAL Scientific Computing Division
    - Commercial Cloud Providers
    - Access to HPC centers in the US

Updated HL-LHC resource need estimates







- Strategy
  - Resource needs for HL-LHC: modeling a difficult multi-dimensional problem
  - Need to prepare R&D efforts and accompany these efforts to plan and measure their effectiveness and impact
- Justification
  - Modeling of resource needs for HL-LHC critical tool to decide on R&D efforts, clear requirement from funding agencies
- Support for
  - CMS Resource Office to prepare resource requests for LHCC/RRB/C-RSG
  - Long-term resource modeling framework and models

Scale WM and SI for HL-LHC

- Strategy
  - Even aside from HPC resources, the scale expected for the HL-LHC era will challenge the existing workflow management (WM) and submission infrastructure (SI) systems.
  - Allow WMAgent to fit jobs more efficiently into available resources and safely increase the average job length (benefits for opportunistic resources, HPC allocations, …)
  - Scale up global pool (HTCondor and glideinWMS) to support base load but also elastic boosts
- Justification
  - CMS will be HTC and will stay HTC, need to make sure that CMS can use the future resource landscape in a multi-nationally distributed and all encompassing way
- Support for
  - Continued work on StepChain: single HTCondor job runs all steps through MINIAOD.
  - WMAgent partial job completion: allow WMAgent to accept the results of successful jobs which process only a portion of their assigned work.
    - Includes resubmission of incomplete work and improving resubmissions to decrease the need for ACDC workflows.
  - Submission Infrastructure: Combination of HTCondor (job management) and GlideinWMS (fabric management - creates HTCondor pool for CMS)
  - Continue work in scalability and improvements to the global pool.

Storage Infrastructure for HL-LHC

- Strategy
  - As we start preparations for the HL-LHC era, we are beginning to investigate new storage infrastructure possibilities:
    - “Data lake” - more tightly coupled storage systems within the US, allowing us to continue reduction of on-disk replication levels. Thought as a natural evolution of data federations.
  - Large-scale, managed caches - increase availability of popular datasets without over-replication. Improve latency-hiding capabilities.
- Justification
  - Handling data will be one of the biggest challenges of HL-LHC
- Support for
  - Infrastructure migrations include migration from Globus for bulk transfer and authorization layer
  - Data federation scale up and data lakes
  - Xrootd caching and integration into the CMS ecosystem
  - Rucio transition
  - Alternate storage technologies (such as Ceph).