Norfolk, Virginia, USA • May 8-13, 2023
CHEP 2023
Computing in High Energy & Nuclear Physics

HPC resources for CMS offline computing: an integration and scalability challenge for the Submission Infrastructure

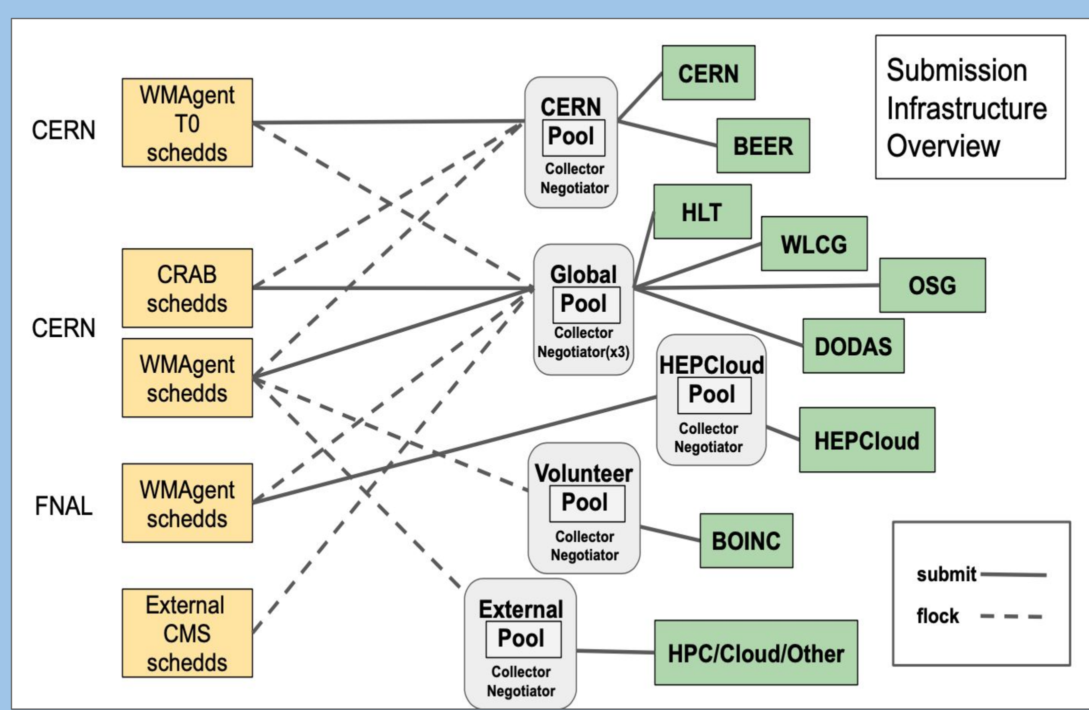
A. Pérez-Calero Yzquierdo[1], Edita Kizinevic[2], Farrukh Aftab Khan[3], Hyunwoo Kim[3], Marco Mascheroni[4], Maria Acosta Flechas[3], Dirk Hufnagel, Nikos Tsipinakis[2] and Saqib Haleem[5], on behalf of the CMS collaboration
1.CIEMAT and PIC (ES), 2. CERN, 3. Fermi National Accelerator Lab. (US), 4. Univ. of California San Diego (US), 5. National Center for Physics (PK)

The CMS Submission Infrastructure

The CMS Submission Infrastructure team in CMS Offline and Computing is in **charge** of operating a set of federated HTCondor pools which aggregates **resources from 70 Grid sites, plus non-Grid resources, where reconstruction, simulation, and analysis of physics data takes place**

The challenges:

- Operate our infrastructure managing an ever **growing** collection of computing resources
- Connecting new and more **diverse resource types** (including non-x86 CPU architectures and GPUs) and **resource providers** (WLCG and OSG, HPC, Cloud, volunteer)
- **Use all of our resources efficiently**, maximizing data processing throughput
- Enforce **task priorities** according to CMS research programme



HPC growing contribution to CMS computing

HPCs can help supporting the computing needs of CMS:

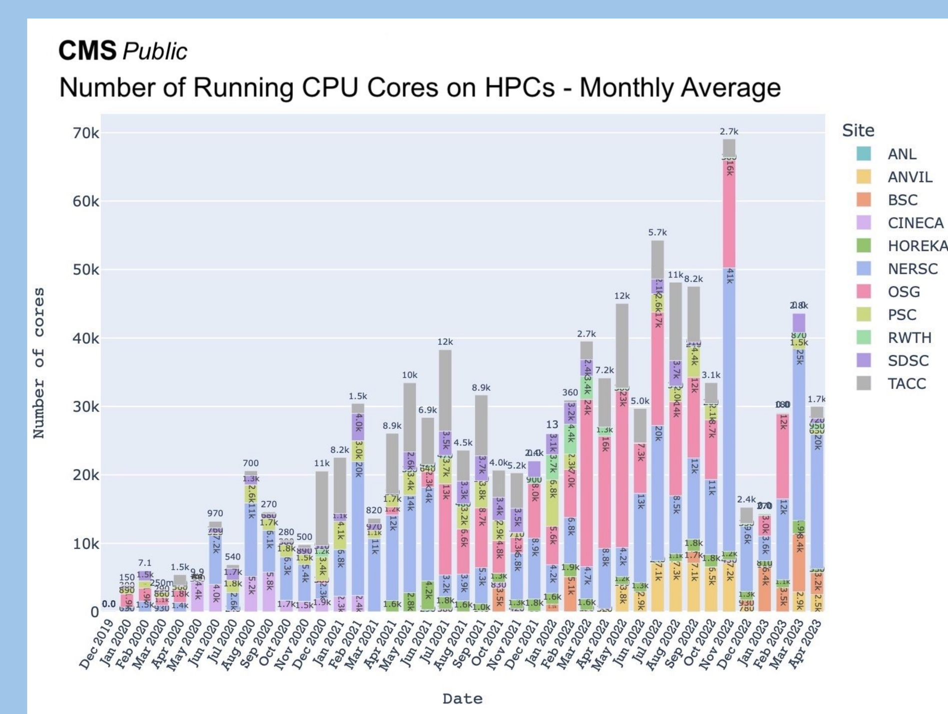
- Substantial national and supranational investments made and planned
- HPCs are part of the scientific computing infrastructure, and there to stay
- Several HPC integration efforts in CMS

Two main models of HPC integration to CMS computing:

- **HEPCloud** (*), based on dedicated sites, for HPC resources in the U.S.
- **Transparent WLCG site-extension**, employed for the EU HPC facilities.

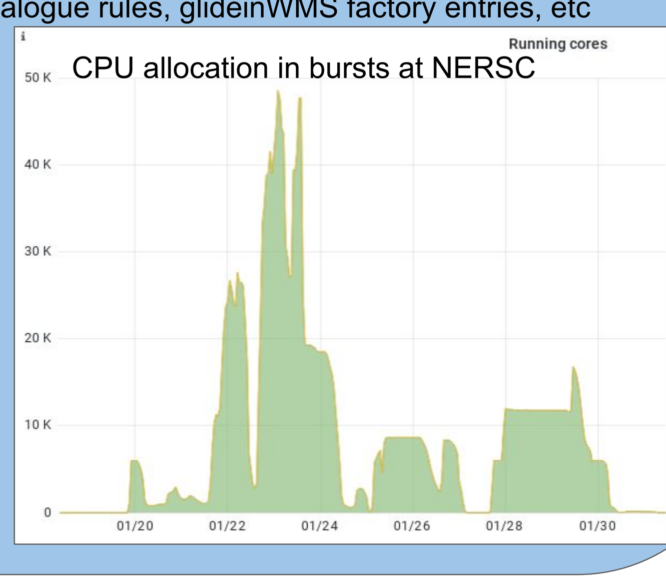
It appears that the **compute capacity used by CMS at HPC facilities** could be **maintained** at the current level, **if not higher**, in the future years.

(*) HEPCloud: arXiv:1710.00100



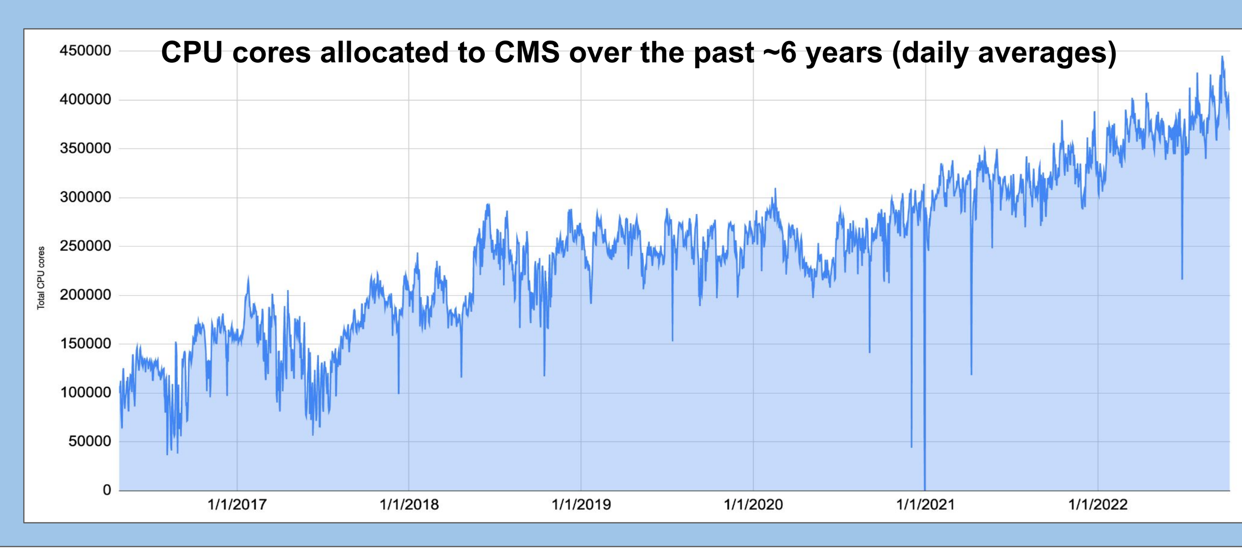
Challenges towards use of HPC resources by CMS

- Each HPC is different
  - Diversity in technical and policy constraints (networking, security, edge services, etc)
- Negotiate deployment of services case by case
  - Singularity, CVMFS
  - Data cache services (for application software, conditions data, event data)
  - Job gateway (CE)
  - http/gridFTP/xrootd doors for input/output data files
- Configuration in our WM/DM systems
  - Processing and Storage site names, site config files, trivial file catalogue rules, glideinWMS factory entries, etc
- **Scalability:**
  - **Sufficient WM and SI capacity to manage and supply workloads to fill increasingly large pools of resources**
  - **HPC vs WLCG provisioning: WM and SI capacity should be dimensioned to profit from HPC resource peaks, in addition to WLCG resource baseline**
- Stability and flexibility of the WM infrastructure and services
  - To ensure effectively using enlarged compute capacity with HPCs



Scalability in the CMS SI

- **Operate away from any scalability limiting factor:** critical aspect for a system that is designed to perform in a dynamic environment, adapting itself to growing **resource demands** by CMS, **resource availability** in the WLCG and the **mix of workloads** it has to manage:
- **Proactively find those limits**, in every direction, and evolve the infrastructure to push them further away:
  - Total computing power our HTCondor pools can harness and use efficiently
    - Collector capacity to process the stream of slot updates and keep resource status fresh
    - Negotiator matchmaking cycle time under control
  - Total number of workflows we can manage and jobs we can run simultaneously with our pool of schedds



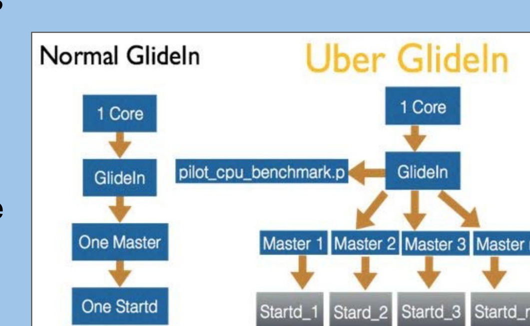
Pushing the limits of the CMS Global Pool

- **Since the start of the Run 2, we have been continuously detecting and solving bottlenecks to our Global Pool,** with the support of the **HTCondor** and **GlideinWMS** developers teams over the years
- Multiple **customized settings** compared to a "standard HTCondor configuration":
  - Using a **CCB** running on a separate host to the CM, with enlarged pool of connection sockets
  - Use of multiple **negotiator** daemons running in multithreaded mode
  - 32-bit binaries for the **shadow processes** running in the schedds (presently disabled)
  - Hierarchy of **secondary collectors** connected to the main top collector process
  - Optimized **slot update conditions** (filter on update triggers, use UDP instead of TCP, enlarged UDP buffer)
  - **Classify queries** reaching the collector from the negotiator as high-prio, as opposed to those from the GlideinWMS FE and CMS WM and monitoring services
  - **Redirect non-high prio queries** to the backup secondary collector (HA infrastructure at FNAL)
- Most of these actions directed at **avoiding the saturation of the top collector**
- **Essential to build an exhaustive and reliable monitoring infrastructure**
- **Previous rounds of our CMS SI scalability tests reported at CHEP:**
  - Exploring GlideinWMS and HTCondor scalability frontiers for an expanding CMS Global Pool, CHEP 2018, https://doi.org/10.1051/epjconf/201921403002
  - Evolution of the CMS Global Submission Infrastructure for the HL-LHC Era, CHEP 2019. https://doi.org/10.1051/epjconf/202024503016
  - Reaching new peaks for the future of the CMS HTCondor Global Pool, vCHEP 2021, https://doi.org/10.1051/epjconf/202125102055

The SI 2023 scale tests (I)

Our main goal for the 2023 tests is to assess the scalability of our Global Pool, considering the following updates since our latest tests (2021):

- Evolution in HTCondor software (tested version 10.0.1)
- A new physical host for the central manager (AMD EPYC 7302 at 3 GHz)
- Adoption of token-based authentication for HTCondor services in our SI
- Incrementally improved configuration of our infrastructure



As in previous tests

- Über-glideins used to generate enlarged copies of our Global Pool
- Secondary collector service in the pool as a source of monitoring data, to reduce stress from non-essential queries on the main collector

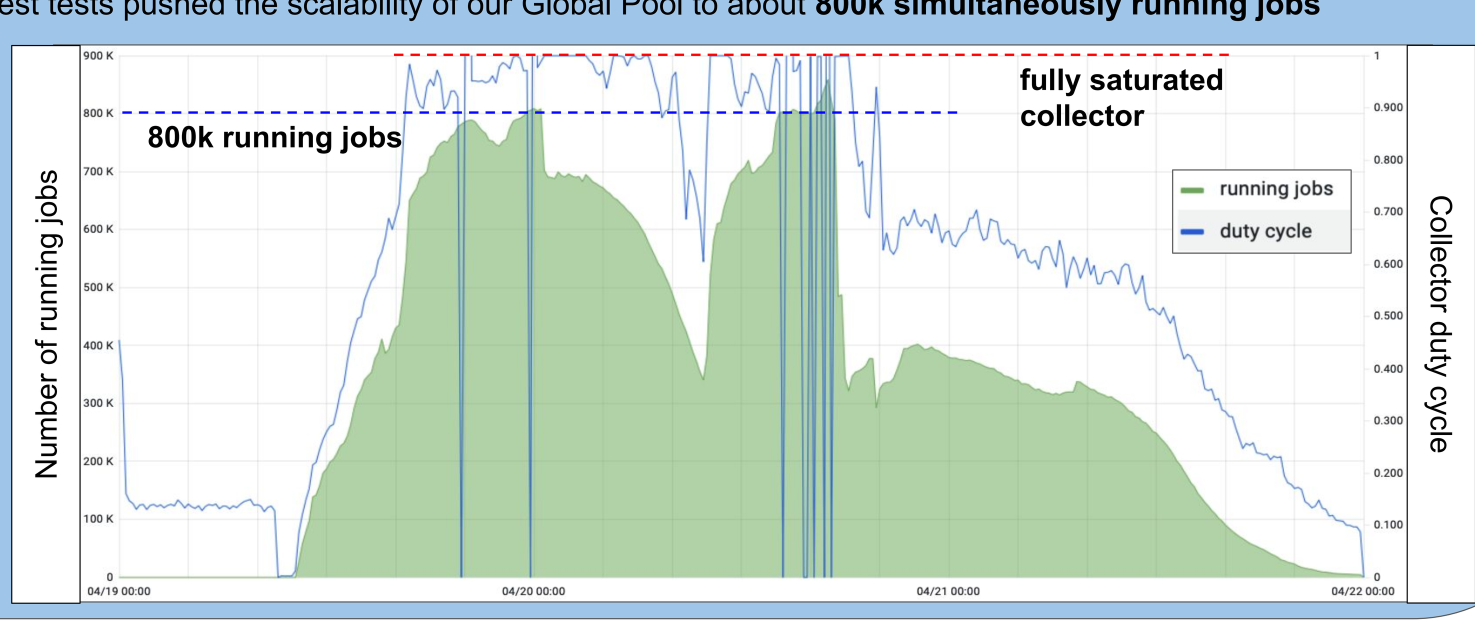
Initial results and mitigations:

- Found **first bottleneck** in the total capacity of our **pool of schedds** to supply jobs to the resource pool:
  - At 1 MB of RAM per running job, our 10 schedds with 50 GB of memory saturated at 50k running jobs each
  - Saturation at 0.5M running jobs, comparable to our 2021 result
  - Schedd pool capacity subsequently enlarged to, in principle, be able to support nearly 1M running jobs
- **Second bottleneck** in the HTCondor Connection Brokering (**CCB) service**, hosted in a VM, limiting the maximum number of TCP connections
  - Moved the CCB service to a physical node

The SI 2023 scale tests (II)

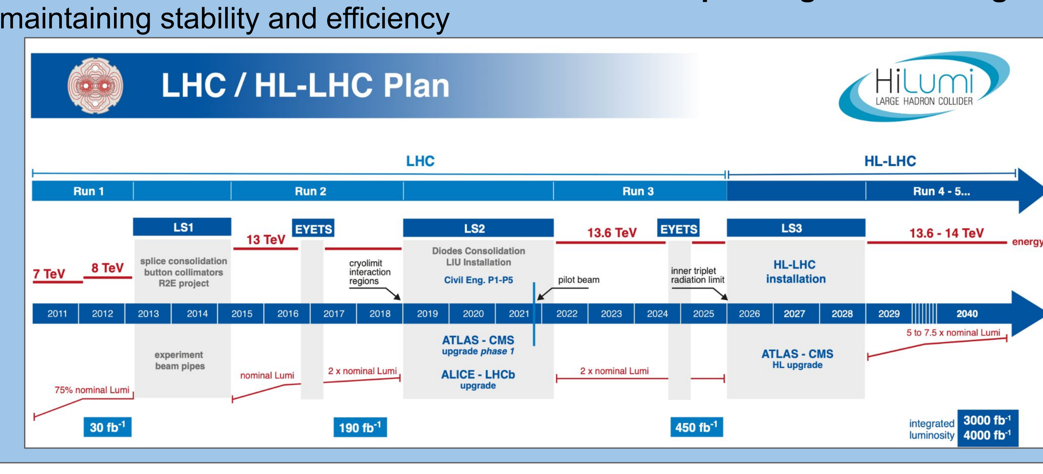
Our latest results:

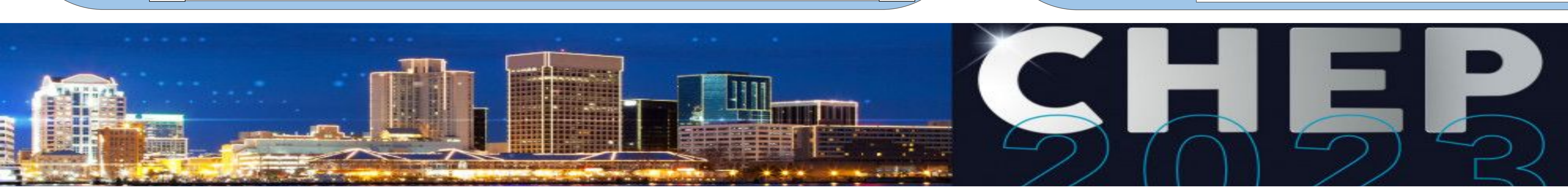
- The main Global Pool collector, which processes slot status updates, becomes increasingly stressed as the size of the pool and the rate of updates keeps growing.
- When the collector becomes fully saturated (collector duty cycle = 1), it can't provide further updated slots status information to the negotiator, thus job to slot matchmaking becomes slow and inefficient
  - This effectively limits the total number of running jobs that a pool can manage.
- Our latest tests pushed the scalability of our Global Pool to about **800k simultaneously running jobs**



Conclusions and Future work

- **HPC resources contribution to CMS Computing** has **increased** in the recent past and is expected to **continue growing** in the coming years
- The **integration** of these resources presents a **number of challenges** to our infrastructure, including that of **scalability**, both in terms of an increment to the CMS baseline compute resources, and specially due to CPU allocation **bursts** at HPCs
- The CMS SI team, in close collaboration with the HTCondor and GlideinWMS developers, **periodically evaluates the scalability** of our infrastructure, in order to **anticipate** and **remedy** future scaling and stability problems and stay off the bleeding edge of limitations
- Our 2023 scalability tests showed the **capacity** of our SI to support up to **0.8M simultaneous running jobs**
- The LHC program extends well into the future, so we need to **continue pushing the SI for higher scales**, as required by CMS needs, while maintaining stability and efficiency





Projects FPA2016-80994-C2-1-R, PID2019-110942RB-C21, BES-2017-082665 and PID2020-113807RA-I00 funded by: