

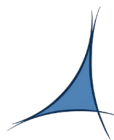


Phasing Out GSI Authentication in the CMS Submission Infrastructure

Nikos Tsipinakis, Marco Mascheroni, Antonio Pérez-Calero
Yzquierdo, Saqib Haleem, Edita Kizinevic

On Behalf of the CMS Submission Infrastructure team

HTCondor Workshop Autumn 2022



PIC
port d'informació
científica

UC San Diego

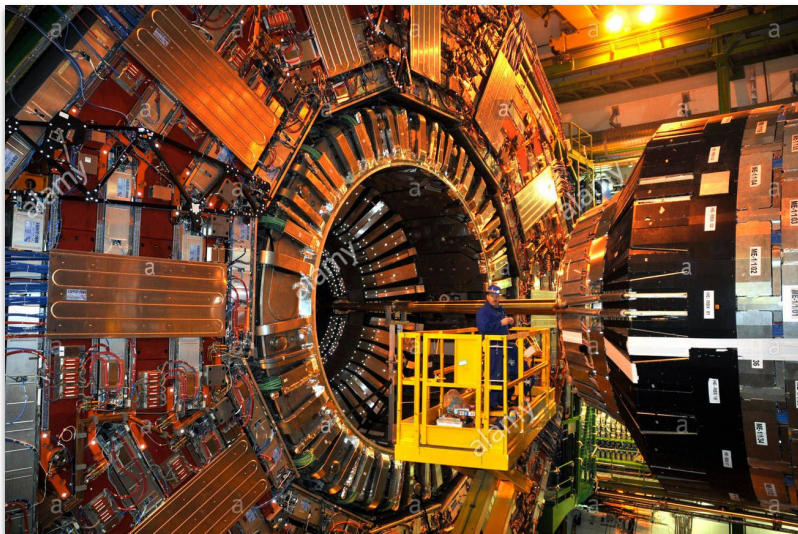


Outline

- Introduction
- The CMS Submission Infrastructure
- IDTokens
- SciTokens
- Migration status
- Conclusion
- Future goals

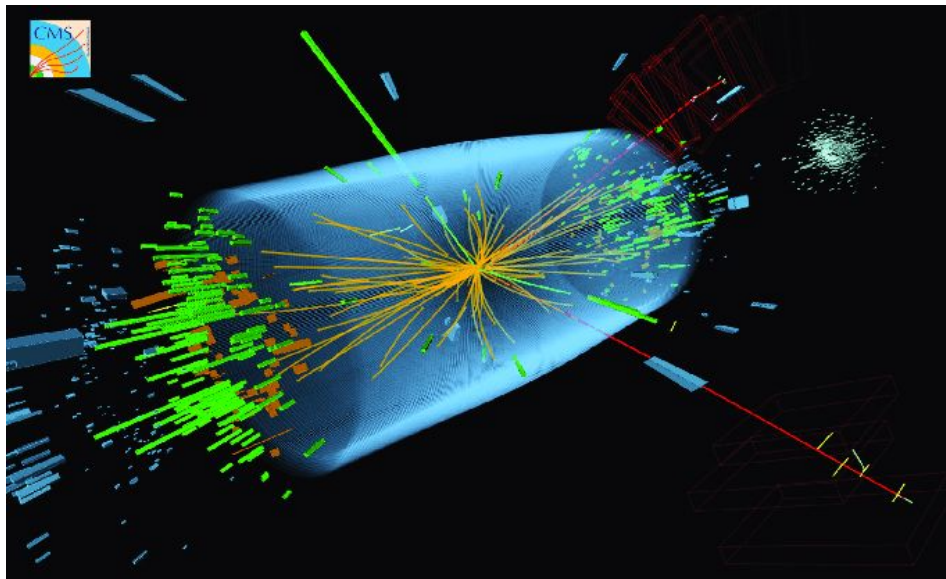


The CMS experiment at CERN



- Experimental data is stored, distributed, reconstructed, and analyzed, comparing to simulated data (Monte-Carlo)
 - **Hundreds of PBs per year**

- **High Energy Physics general-purpose experiment** recording proton-proton collisions at the LHC at CERN





The computing landscape - the WLCG

- Data traditionally analyzed using **WLCG Grid resources**
 - Global collaboration of around 170 computing centers
 - Access based on dedicated resources (**pledges**)

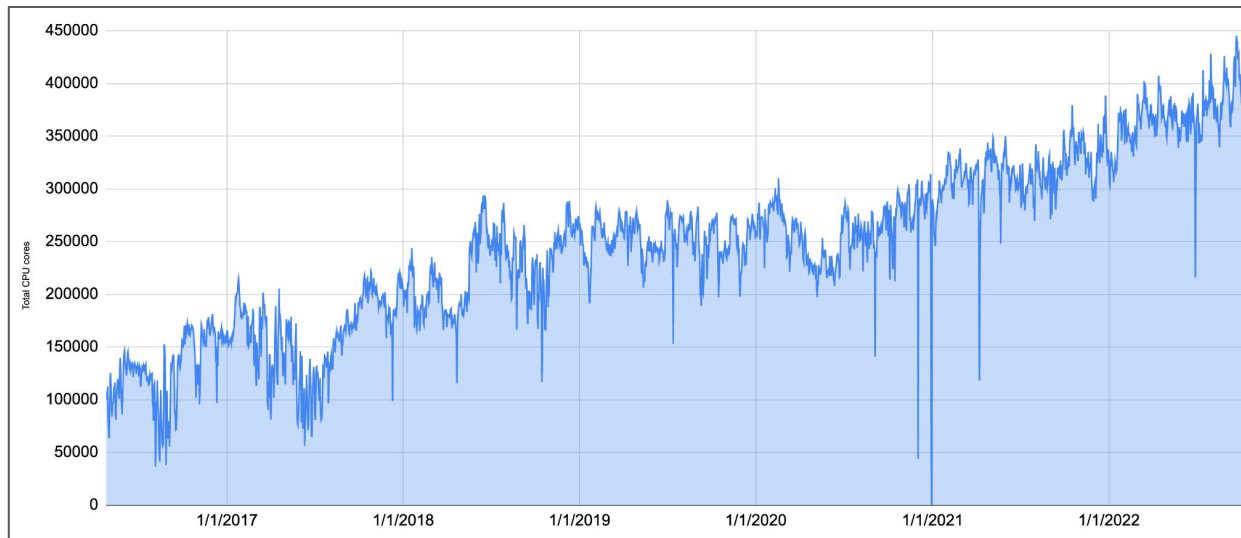




The CMS Submission Infrastructure Group

- Part of CMS Offline and Computing in **charge** of:
 - Organizing HTCondor and GlideinWMS operations in CMS, in particular of the **Global Pool**, an infrastructure where reconstruction, simulation, and analysis of physics data takes place
 - Communicate CMS **priorities to the development teams** of glideinWMS and HTCondor
- In practice:
 - We operate a set of federated HTCondor pools **distributed over 70 Grid sites, plus non-Grid resources**
 - We hold fortnightly **meetings with** HTCondor and glideinWMS **developers** where we discuss
 - Current operational limitations
 - Feature requests
 - Future scale requirements

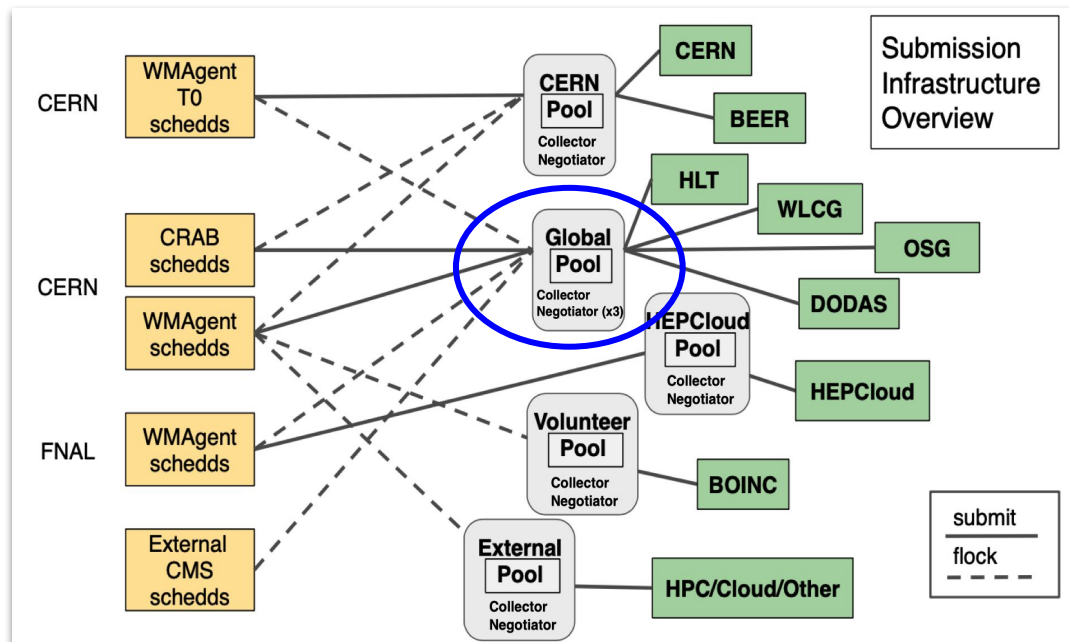
**CPU cores allocated to CMS
over the past ~6 years (daily
averages)**





A federated infrastructure

- The CMS SI model has evolved to running multiple federated pools, with extensive use of flocking
- Multiple sets of specialized workflow managers (CRAB & WMAgent) attached to schedds
- The main Global Pool:
 - Peaks at ~450k CPU cores
 - Up to 210k running jobs
 - 50+ schedds
- Redundant infrastructure for HA





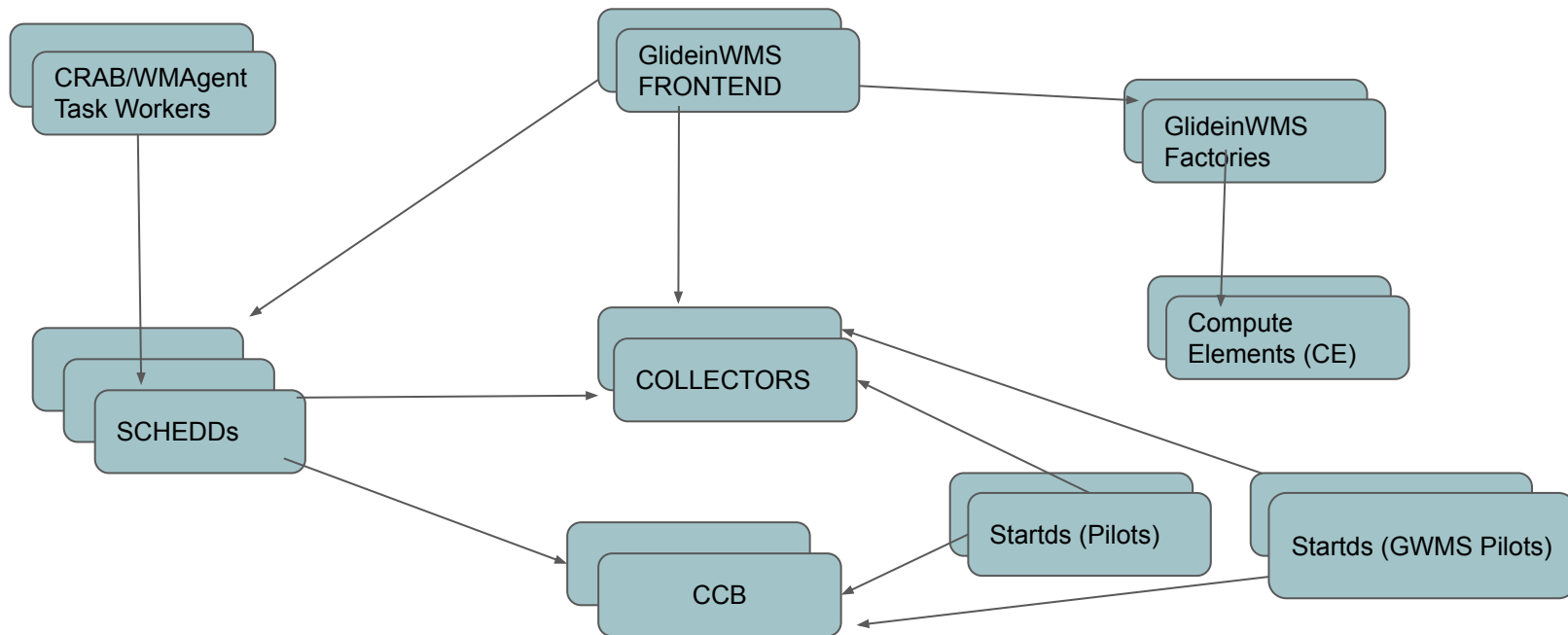
Moving to tokens in the CMS Submission Infrastructure

- Motivation
 - Towards **industry** standards : Capabilities based authorization for distributed services (new)
 - Globus Toolkit **retirement**
 - Practical example: *multiple tokens with different capabilities instead of a single identity i.e. powerful pilot (GSI) proxy*
- Timeline
 - **March 2023**: HTCondor GSI End Of Life



Components of the CMS Global Pool

- Authentication between SI Internal Components (IDTOKENS)
- Authentication between Factories <-> Sites (SciToken)



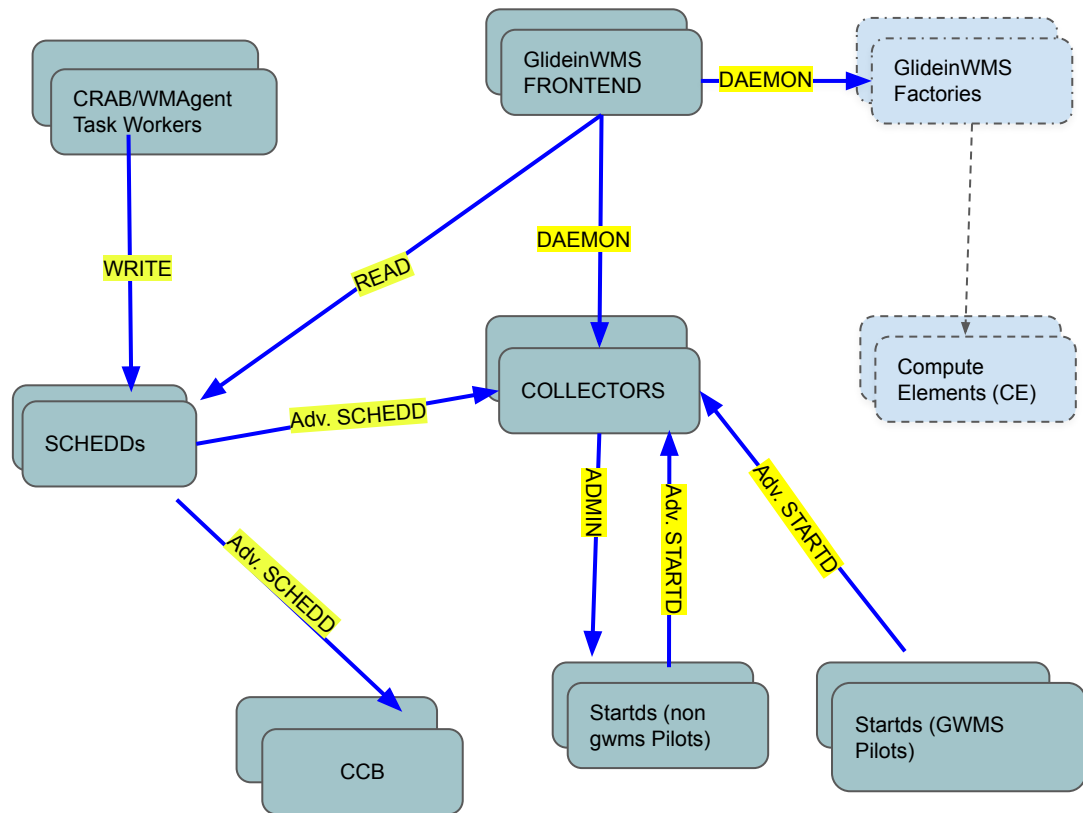


1.- IDTOKENS



IDTOKEN Implementation

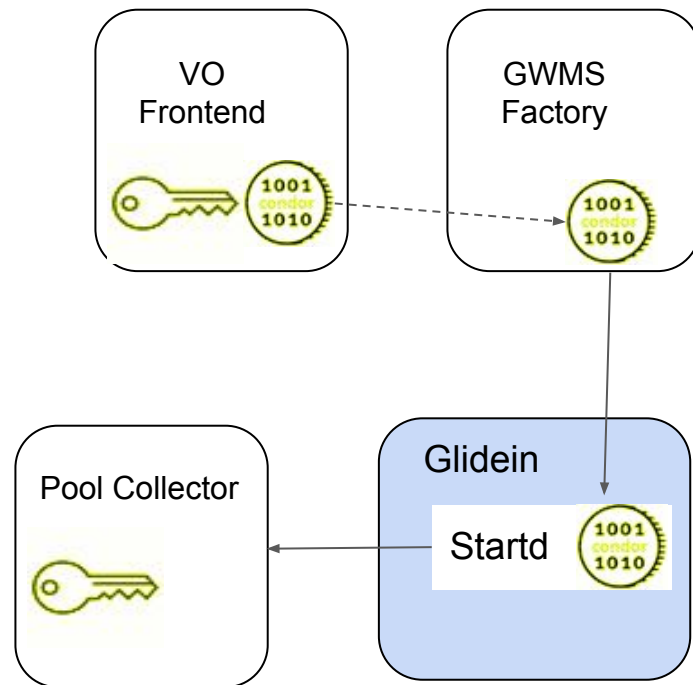
- Whole Infrastructure is **already using** IDToken as preferred authentication method between components.
- **96%** startds using IDTokens, 4% falling back to GSI due to token expiration.
- IDTOKENS created with need based authorization capabilities and limited lifetime.





IDTOKEN authentication for GlideinWMS pilots

- **Same signing key** placed on both **Frontend** and **Collector**
- Frontend generates an IDTOKEN for each site.
- **IDTOKEN is transferred** to the factory, the CE, Batch System, **WN** (as pilot proxy before)
- **Startd is then authenticated and authorized** by the collector

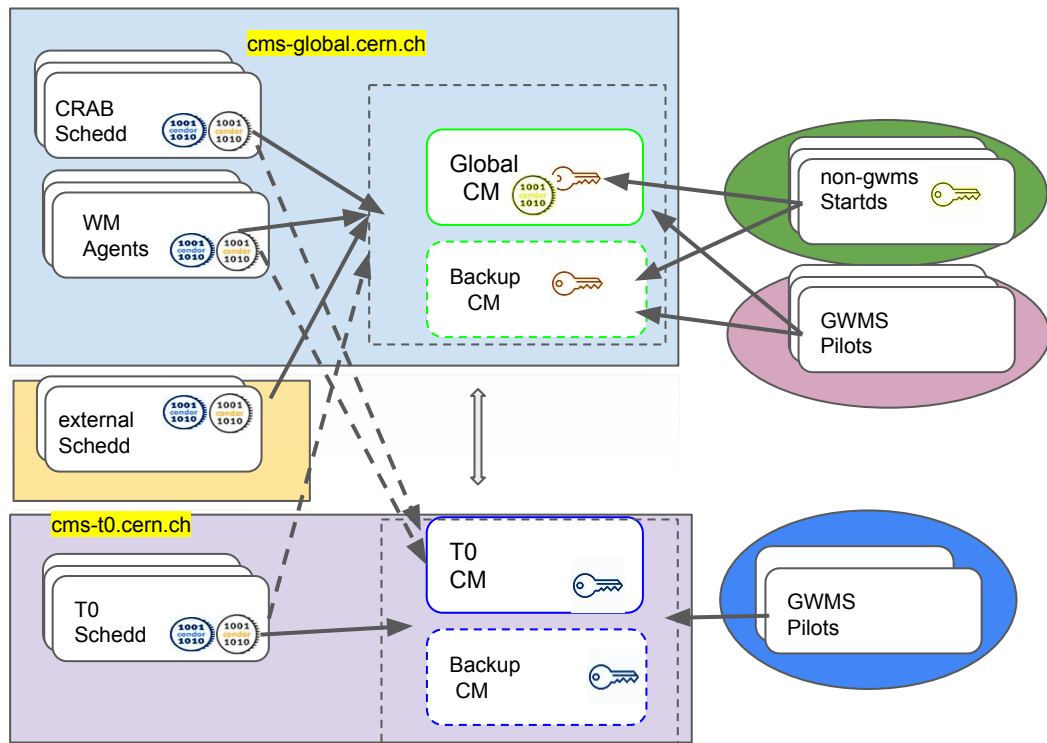




IDTOKEN Implementation

TRUST DOMAINS:

- Separate trust domains for:
 - Global and CERN pools
 - External schedds
 - Pilots and non-gwms startds.
- Each trust domain has its own signing key.
- External startds issues IDToken to CM which is required for admin operations: e.g **condor_drain**



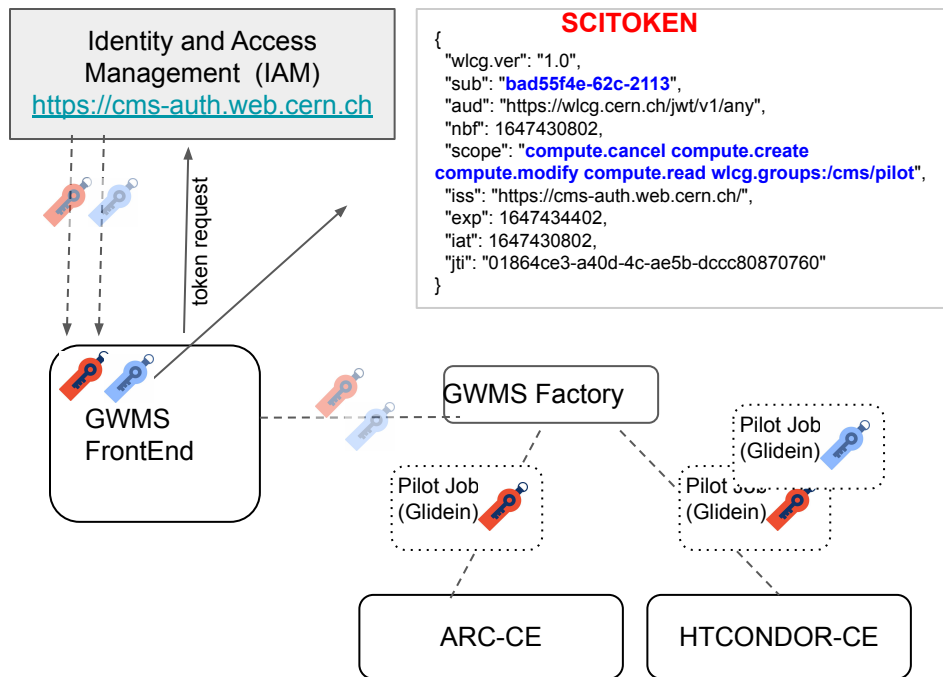


2.- SCITOKENS



SciToken Implementation

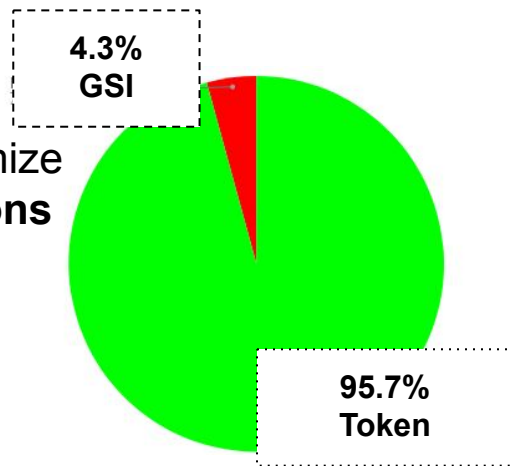
- SciToken authentication used for pilot submission between **Factory** -> **Compute Element (CE)**
 - HTCondor-CE (newer version)
 - ARC -CE (REST API)
- SciToken with different scopes/subjects are issued for different categories of pilots. e.g. local vs generic pilots.
- **CronJob**: Registered clients with CMS IAM fetches fresh token after every 10 minutes, and put it on FE, which is then used by factory for pilot submission.





SciToken Implementation (HTCondor CEs): Almost there!

- Nearly **96%** of the HTCondor CEs we interact with have been upgraded to a recent enough HTCondor version and support SciTokens.
- CMS is working in a systematic way with each grid site to minimize disruption during transition (= **transparent from CMS Operations point of view**)
 - Separate glideinWMS FE group "main-token" created for submitting jobs with SciToken credentials.
 - Individual CEs are moved to token group after successful condor_ping test.
- Site admins perform mapping of different jobs based on scitoken's subject in htcondor-CEs e.g:



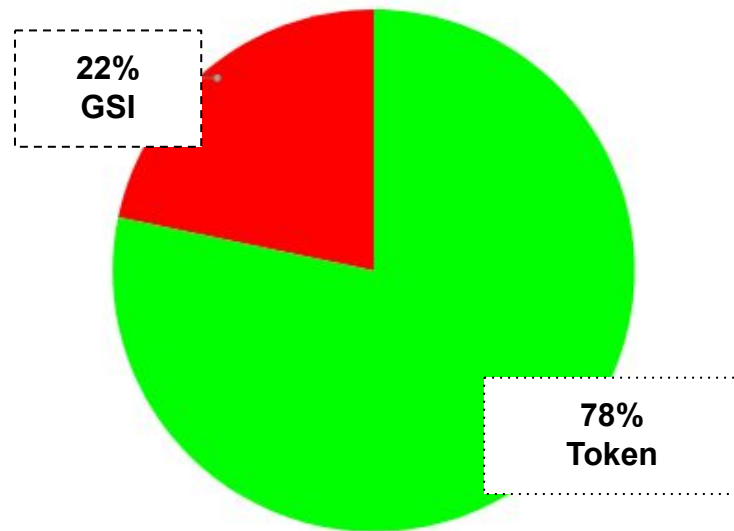
scitoken enabled
CEs (HTCondor)

```
# CMS ITB generic pilots:  
SCITOKENS /^https://cms-auth.web.cern.ch/V,07f75a9a-bb78-4735-938b-7e61b2b6d5c$/ cmspilot  
# CMS ITB local pilots:  
SCITOKENS /^https://cms-auth.web.cern.ch/V,efbed8c1-f9a7-4063-92f7-f89c04c04a3$/ cmslocal
```



SciToken Implementation (ARC CEs)

- In the case of ARC CEs (about $\frac{1}{3}$ of our total sites and CEs use this technology), our strategy so far has been to test that we can interact with them via x509 proxies but with the new REST interface
 - About **78%** of all ARC CEs we use already OK



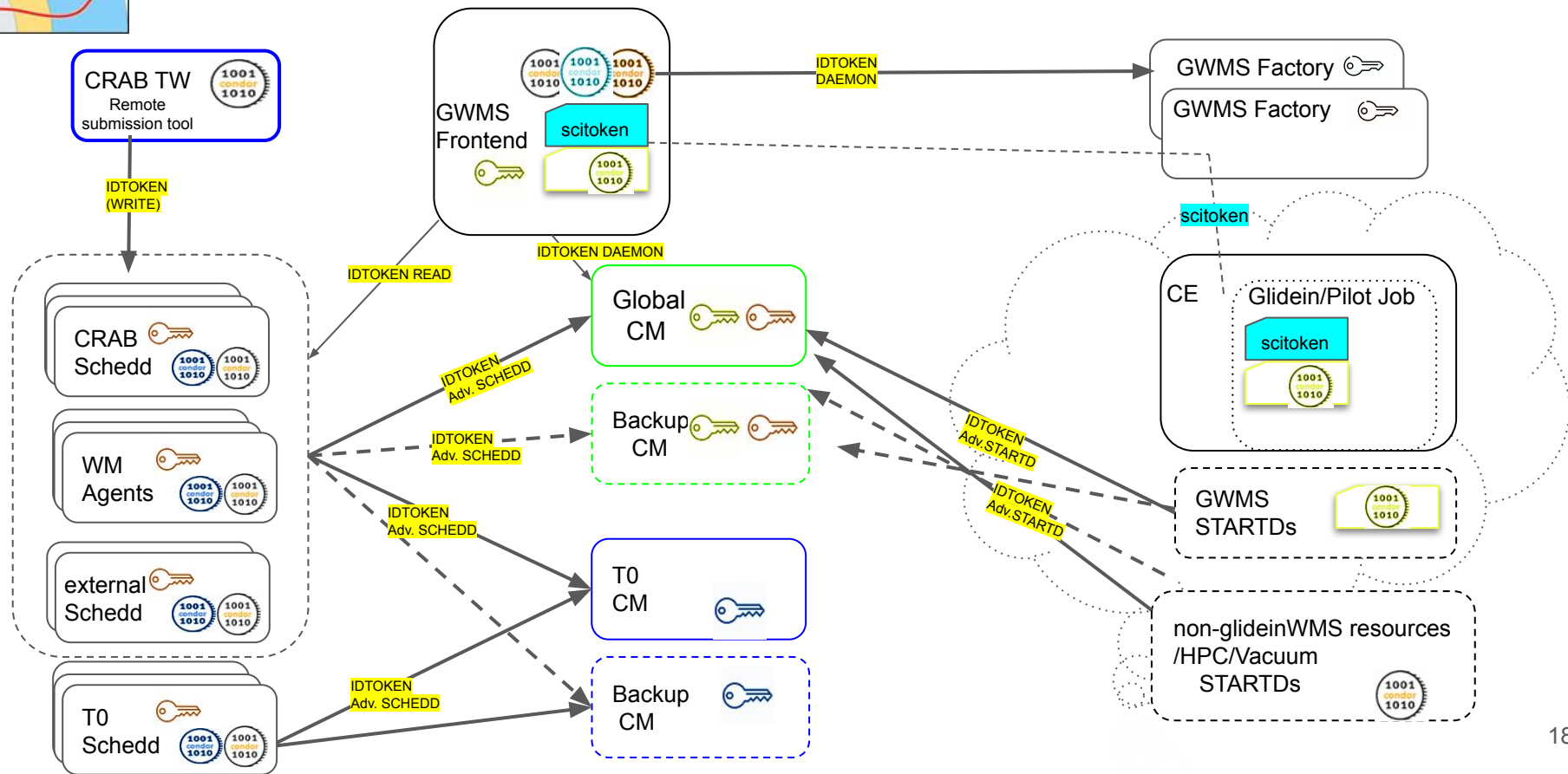
ARC-CEs REST
interface status



Summary



Summary





New challenge: Secret management

- GSI secrets needed minimal maintenance as they used one certificate per-host
- IDTokens need to be securely generated, tracked and distributed to their intended user
 - Need for an audit log
 - Need for a signing key revocation/rotation procedure
 - **Work on these has already started!**



CONCLUSION AND FUTURE PROSPECTS



Conclusions and Future Prospects

A round of applause broke out in the CERN Control Centre on 5 July at 4.47 p.m. CEST when the Large Hadron Collider (LHC) detectors started recording high-energy collisions at the unprecedented energy of 13.6 TeV

5 JULY, 2022



Celebrations at the CERN control centre (CCC) to mark the start of LHC Run 3 (Image: CERN)

Additional motivation as the LHC has restarted operations (Run3, 2022-25)

...at **unprecedented collision energy** (13.6 TeV) providing new territory for Physics exploration

Time to accumulate loads of new data with CMS!

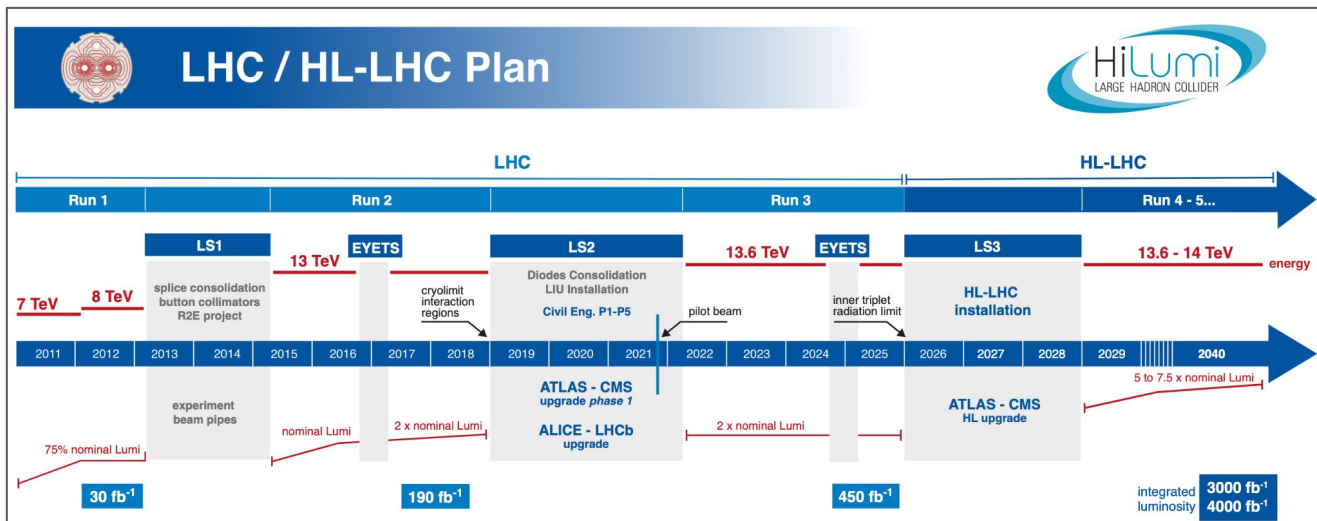
[CERN news](#)



Conclusions and Future Prospects

...and the LHC program extends well into the future so we need to:

- continue pushing for **higher scales**, as required by CMS needs...
- ...while maintaining **stability** and **efficiency**
- remain relevant by **adapting** to tools/technology changes





**Thanks to the
HTCondor team for the
great support over
many years!**

Backup Slides

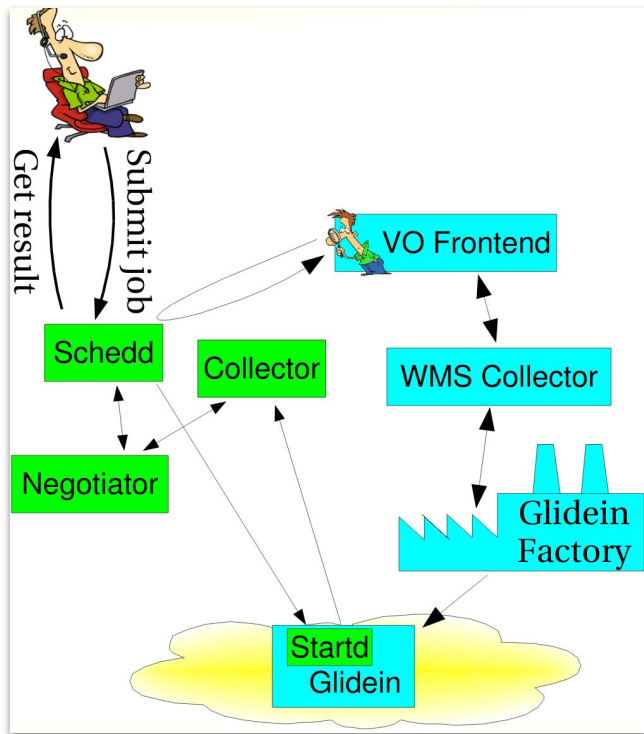
Abstract

The CMS Computing Submission Infrastructure group manages and exploits a set of HTCondor pools to satisfy the experiment computing needs. The biggest of those pools, the so-called CMS Global Pool, currently aggregates nearly 400k CPU cores dynamically from pledged and opportunistic WLCG resources.

Historically, authentication among the diverse components of the infrastructure used the Grid Security Infrastructure (GSI), based on identities and X509 certificates. However, more modern authentication standards based on capabilities and tokens have emerged over the years. The CMS Submission Infrastructure group is in the process of phasing out the GSI part of its authentication layers in favor of IDTokens and Scitokens. In this contribution we will report on the current status of this migration, and our plans for the final GSI phase out.



Building dynamic HTCondor pools with GlideinWMS



- CMS computing pool is build using two components:
 - **GlideinWMS** : Resource provisioning overlay batch system which grows and shrink based on Job pressure.
 - **HTCondor**: Batch system for Job scheduling.