CMS Token Transition

Strategy and Status

Brian Bockelman, on behalf of the co-authors and the CMS Collaboration

MORGRIDGE
INSTITUTE FOR RESEARCH

Many thanks to the co-authors

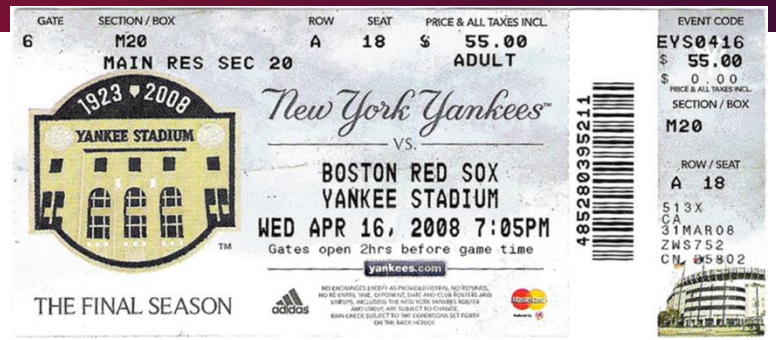
- Moving an experiment's authorization model forward is not a light lift! I'd like to start by acknowledging the co-authors.
  - Alan Malta Rodrigues (University of Notre Dame (US))
  - Brian Paul Bockelman (University of Wisconsin Madison (US))
  - Chan-Anun Rungphitakchai (University of Wisconsin Madison (US))
  - Dave Dykstra (Fermi National Accelerator Lab. (US))
  - Diego Ciangottini (INFN, Perugia (IT))
  - Edita Kizinevic (CERN)
  - Eric Vaandering (Fermi National Accelerator Lab. (US))
  - Marco Mascheroni (Univ. of California San Diego (US))
  - Panos Paparrigopoulos (CERN)
  - Rahul Chauhan (CERN)
  - Sarun Nuntaviriyakul (Chulalongkorn University (TH))
  - Stephan Lammel (Fermi National Accelerator Lab. (US))
  - Vaiva Zokaite (Vilnius University (LT))

Authorization, authentication, and credentials: A recap

Three important terms for this presentation:

- **Authorization**: Deciding whether an entity is permitted to perform an action.
- **Authentication**: Mapping an entity to an identifier.
  - Note: *Authentication* is often part of an authorization scheme.
- **Credential**: knowledge that establishes a fact (e.g., identity).
  - Not too far off from the ‘credentials’ the university provides: a diploma establishes the bearer has particular knowledge.
  - Classic example: a username/password is used as a credential to perform authentication.

Moving from identity mapping to capabilities



- Authorization on WLCG was always based on identity mapping*:
  - A request was authenticated to a global identity.
  - The global identity was mapped to a local identity.
  - The request was authorized if the local identity was authorized to perform the action.

Scheme	Credentials	Authentication	Authorization
Gmail login	Password, 2FA	Username	Access to your inbox
Building access	ID card	Identity in HR database	Elevators
International Travel	Passport	Identity according to US Government	Enter Switzerland
Baseball Game	Ticket	**NONE!**	Sit in section 4, seat 34B
Workshop	Zoom URL	**NONE!**	Attend this wonderful talk!

A transition in Two Ways -> Credentials

X.509

JWT


```
{
  "aud": "ANY",
  "sub": "brian.bockelman@ligo.org",
  "uid": "brian.bockelman",
  "ver": "scitoken:2.0",
  "nbf": 1728392225,
  "scope": "read:/frames read:/shared gwdatafind.read
gracedb.read read:/staging read:/kagra
write:/staging/brian.bockelman read:/ligo dqsegdb.read
read:/virgo",
  "iss": "https://cilogon.org/igwn",
  "exp": 1728403030,
  "iat": 1728392230,
  "jti":
"https://cilogon.org/oauth2/4add24b32504f17477b7392ca7a
ab0d5?
type=accessToken&ts=1728392230685&version=v2.0&lifetime
=10800000"
}
```

A transition in Two Ways -> Philosophy

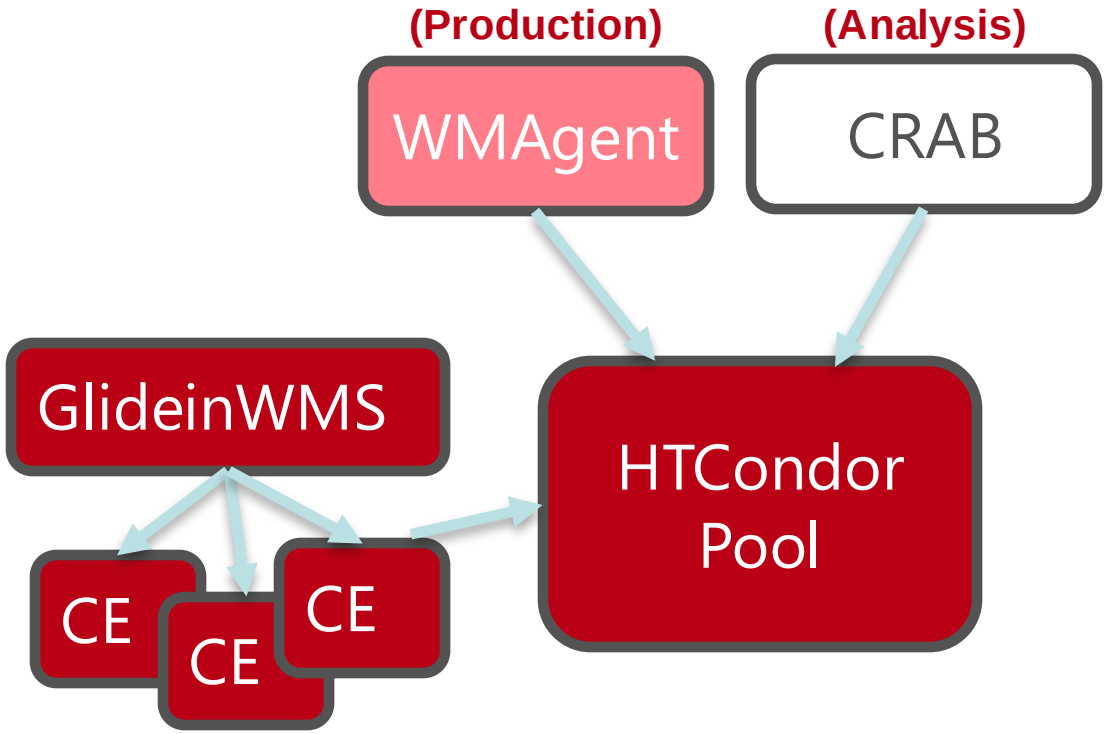
- Perhaps more important than credential format is a change in **authorization philosophy**: capability-based authorization instead of identity-mapping.
- Capabilities state **what you are allowed to do**.
  - Identity mapping states **who you are** and assumes the service can map that to a list of permissions.
  - Identity mapping requires out-of-band coordination across all the distributed endpoints to ensure identities are mapped correctly (meaning changes are hard to implement).
  - Identity mapping schemes are hard to attenuate (remove or limit permissions from the credential).
- Example capabilities for the WLCG:
  - **compute.create**: submit a job to a CE
  - **storage.create:/store/user/bbockelm/foo.txt**: create a specific file within the CMS space at a SE.

CMS Strategy for the Token Transition

- Outside-in: **start with the distributed services**, verify they are working, and then convert CMS services.
  - Begin with services that are expert-centric (e.g., production system) and expand out to the user base (analysis tools, standalone laptop environments).
  - When possible, pick:
    - Standard protocols (OAuth2 for token acquisition)
    - Shared/external services (FTS, ETF)
    - Shared solutions (Rucio, Vault, HTCondor)
    - Minimize the CMS-specific technologies!
- Technology choices:
  - **Token profile**: WLCG Common JWT Profiles (10.5281/zenodo.3460257)
  - **Token issuer**: IAM instance operated by CERN IT.
  - **Token provider**: HTVault (a specially-configured, patched version of Vault)

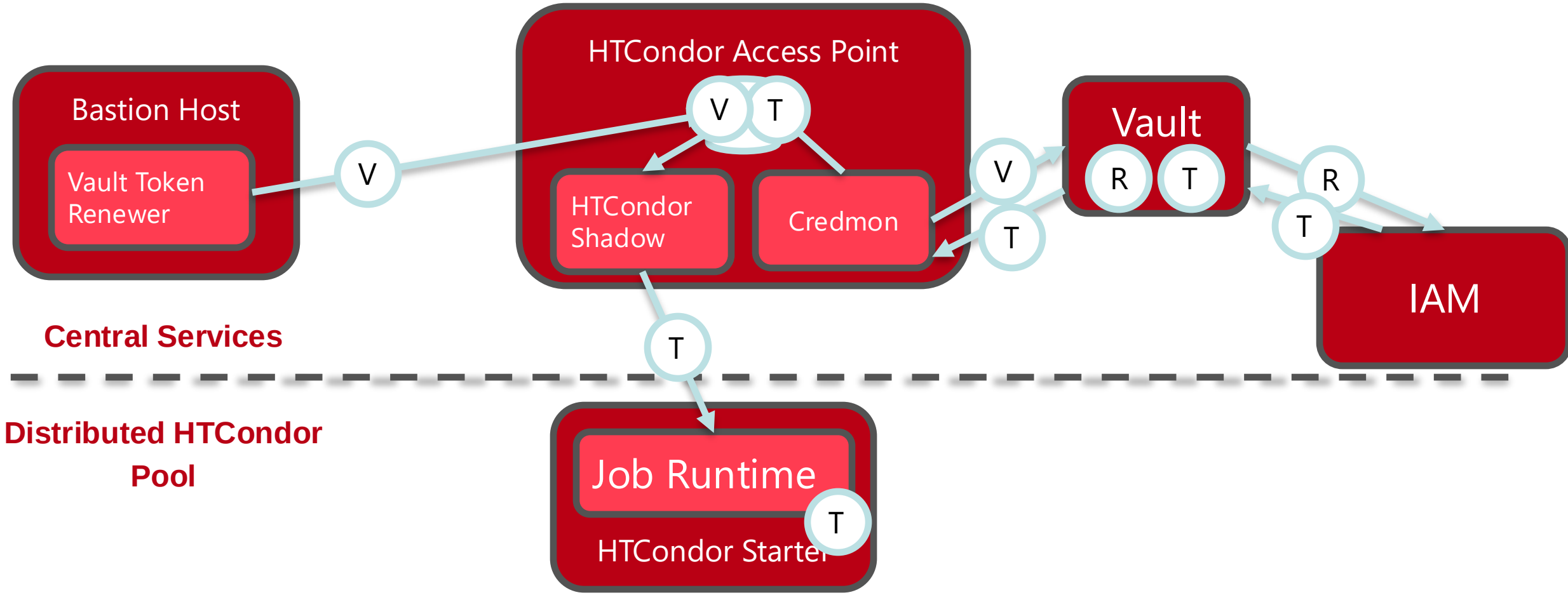
Compute Ecosystem

- CMS uses the GlideinWMS software to interact with the Compute Entrypoints (CEs) at sites to build a HTCondor pool.
  - **Status**: GlideinWMS and HTCondor have been converted to token-based submission and internal authorization since 2022.
  - **Status**: Every CE but 3 ARC-CEs use tokens for pilot job submission. 159 CEs using tokens.
- Once the HTCondor pool is constructed, the WMAgent component submits production jobs to the system and CRAB submits analysis jobs.
  - **Status (WMAgent)**: Latest runtime will use tokens if present for reading data and stageout. Rollout of token-enabled version expected in the next quarter. Some central WMAgent services (e.g., deleting temporary files) not converted over.
  - **Status (CRAB)**: Transition largely not started.



Tokens to the Job

- WMAgent's runtime will use a token if present -- but how does a token get to the job?

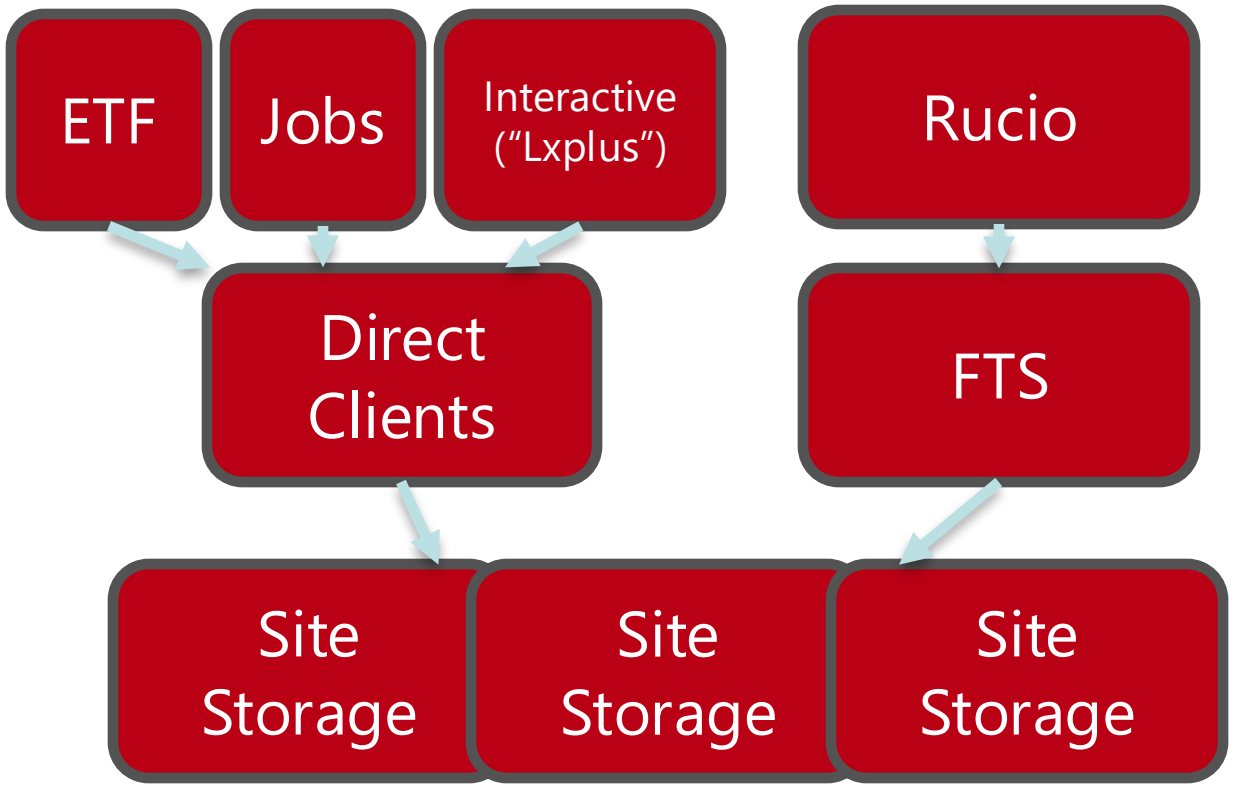


The Previous Slide, But in Words

- The HTCondor **shadow** process is responsible for maintaining the remote job's runtime; it sends a steady stream of new access tokens according to the job description. **Only the needed access token goes to the job**.
  - The shadow pulls the access token from an on-disk database maintained by the **credmon**.
  - The credmon periodically receives the token from the **HTVault** service, authenticating with a Vault token.
  - The HTVault service handles the OAuth2 flows, including exchanging a refresh token for a new access token from the **IAM** service.
  - The bastion host is managed by the CMS operator and periodically pushes new vault tokens to the HTCondor AP.
- **Status**:
  - Setup is in production for Fermilab experiments.
  - CMS was able to reproduce the setup in test instances. Rolling out this calendar year.
  - Finalizing the HTVault instance at CERN with IT.

Storage Ecosystem

- The storage transition has been the most active one in 2024, focusing on bulk data transfers between sites (the “Rucio / FTS / SE” stack).
- **Rucio** uses the client credentials flow to acquire a token per dataset from IAM.
  - The token is passed to **FTS** which performs the OAuth2 token exchange flow with IAM to get a refresh token as well.
  - FTS uses the token to interact with site storage and drive the transfers.
  - Design and development needed for tape transfers.
- **Status**: Tested heavily in DC24; tweaks and adjustment through the summer; put into production in September 2024.



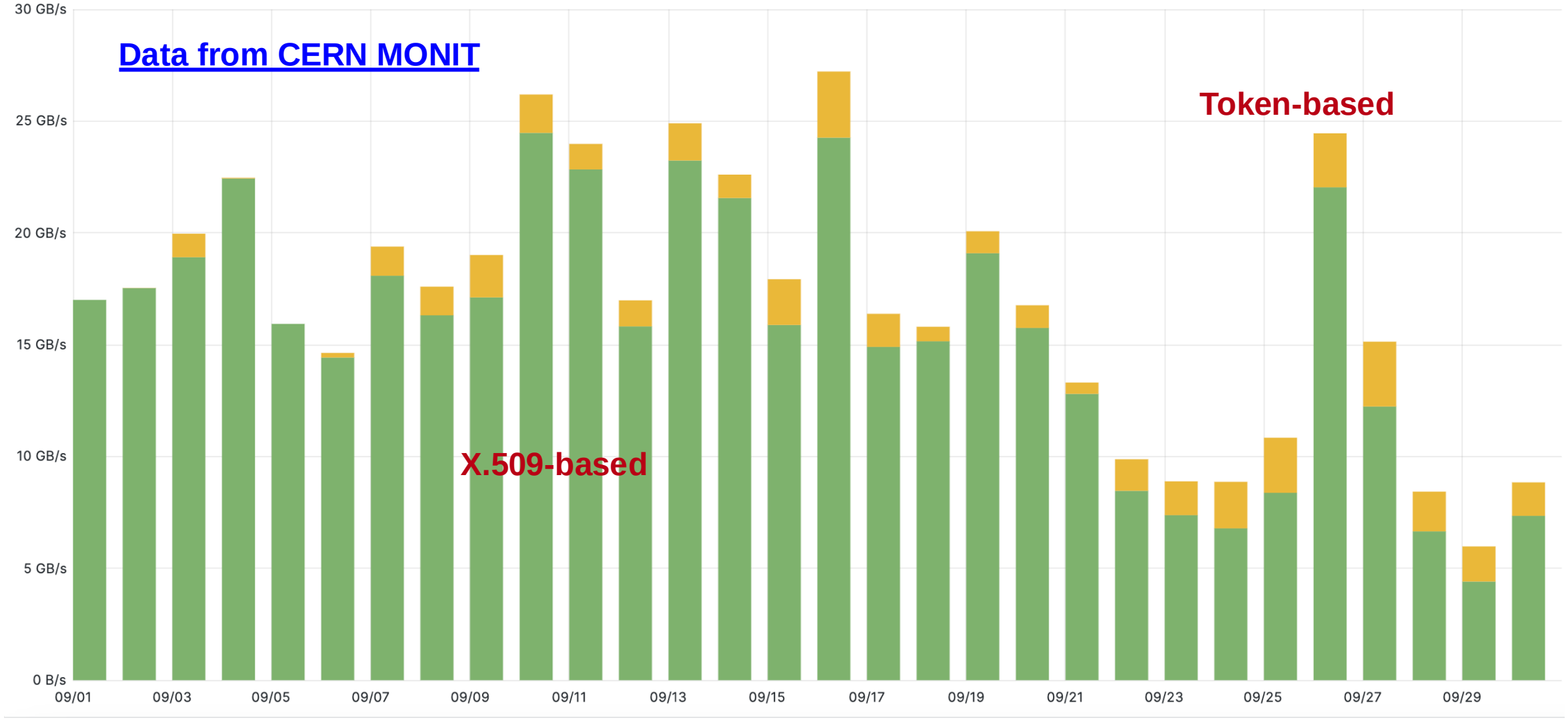
Storage Status

- We continuously monitor token readiness at CMS storage endpoints using ETF and observe:

Technology	Site Count	Ready	Percent	Notes
XRootD (native)	15	15	100%	✅
SToRM	3	3	100%	Critical bug preventing FTS transfers
dCache	32	26	81%	
EOS	9	0	0%	Waiting on new release
DPM	2	0	0%	To be retired, not token-enabled

- Over 30 sites are transferring with tokens in production!

Rucio-based transfers, September 2024



Direct client access to storage

How else is storage accessed remotely?

- **Upload/download files via Rucio**: Implementation is in progress in Rucio; expected in winter 2024.
- **Through the CMSSW physics application**:
  - XRootD client libraries auto-discover token in environment; no code change needed.
  - AAA remote data access infrastructure mostly share endpoints with bulk data transfers. Most pieces are expected to "just work" – but **a rollout campaign is needed**.
    - XCache works with tokens over HTTP – but not xrootd protocol. Development needed.
  - Once HTVault is deployed at CERN, users will be able to acquire tokens via "htgettoken" in their interactive environments.
    - Planning to add a few creature comforts to automatically invoke this in environment setup or when starting a CMS-based container environment.
- **ETF**: Currently working; basis for our continuous testing!

CMS Web Services

- The CMSWeb suite of services provide access to CMS dataset metadata (DAS, DBS) and central production services infrastructure (ReqMgr).
  - The frontend service handles authorization, passes request to backend.
    - Authentication/authorization tightly integrated with X.509/GSI, SL7-based.
  - Significant recent investment to modernize/rework the frontend. Allows CMS to migrate of SL7 and opens the door for token auth. **Status**: New version being deployed.
- The WLCG profile defines various common storage & compute scopes: what scopes should be used for accessing various CMS features?
  - **Status**: We are in the process of defining the CMS-specific scopes for these services.
  - **Status**: Significant work to update client tooling ahead of us in 2025!

IAM Statistics

- How busy is the CMS IAM instance?

6,829	528	>425k
Registered Users	OAuth2 Clients	Tokens in DB

- By far, the driver of the token usage is Rucio. We generate a token per dataset, as needed to transfer. We estimate half of the datasets transferred use tokens, meaning **we need to scale by only 2x** to hit the full expected scale.
  - By optimizing the issuing of refresh tokens, we could reduce the usage by 2x if needed.

Outlook

- Re-envisioning the authorization model for a mature ecosystem like CMS's is a huge undertaking!
  - Early R&D done in 2017; WLCG profile standardized in 2019; first production uses in 2021;
  - CE transition done in 2022; SE in production in 2024.
- Distributed services are in great shape, with few remaining endpoints missing support.
- CMS services and strategy came into sharp relief during mid-2024; we are executing on the plans now.
- By end-2025, likely all services will be in dual X.509 / token mode.

Will be ready to begin retiring X.509 in LS3!

Questions?

This presentation is supported by the National Science Foundation under Cooperative Agreement PHY-2323298. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

FEARLESS SCIENCE

MORGRIDGE INSTITUTE FOR RESEARCH