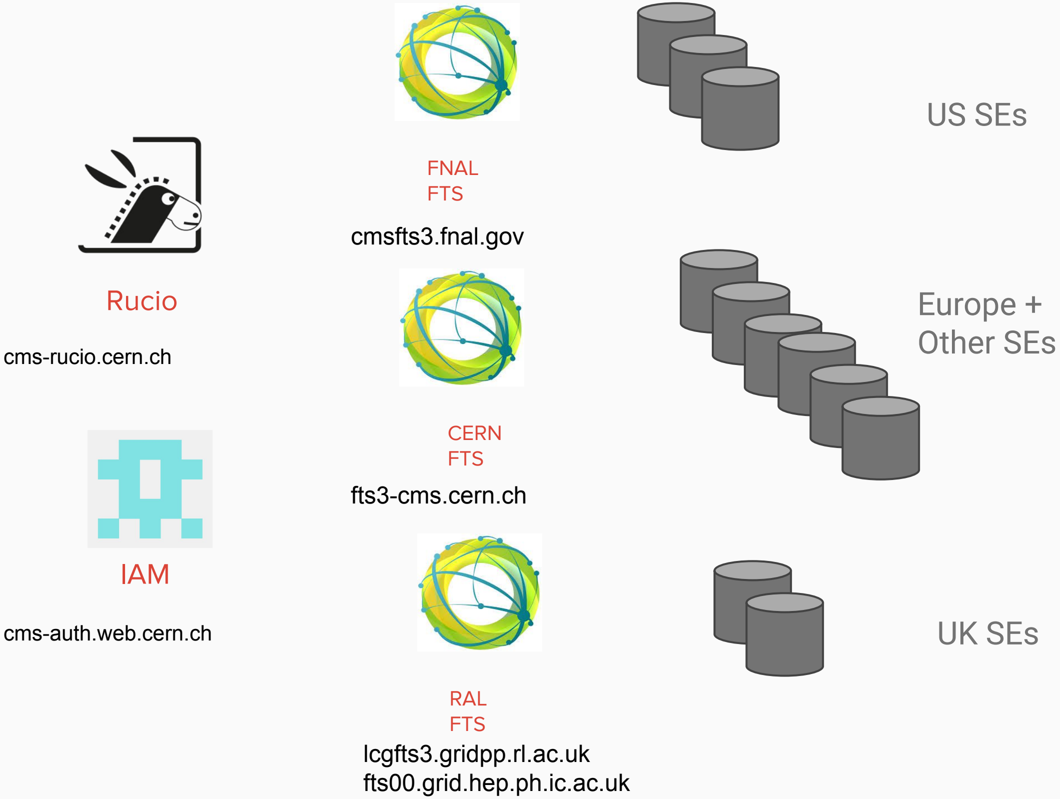
# CMS Token Workflows

Katy Ellis (STFC), Rahul Chauhan (CERN), Stephan Lammel (FNAL)

# Intro



Rucio

cms-rucio.cern.ch

IAM

cms-auth.web.cern.ch

FNAL
FTS
cmsfts3.fnal.gov

CERN
FTS
fts3-cms.cern.ch

RAL
FTS
lcgfts3.gridpp.rl.ac.uk
fts00.grid.hep.ph.ic.ac.uk

US SEs

Europe +
Other SEs

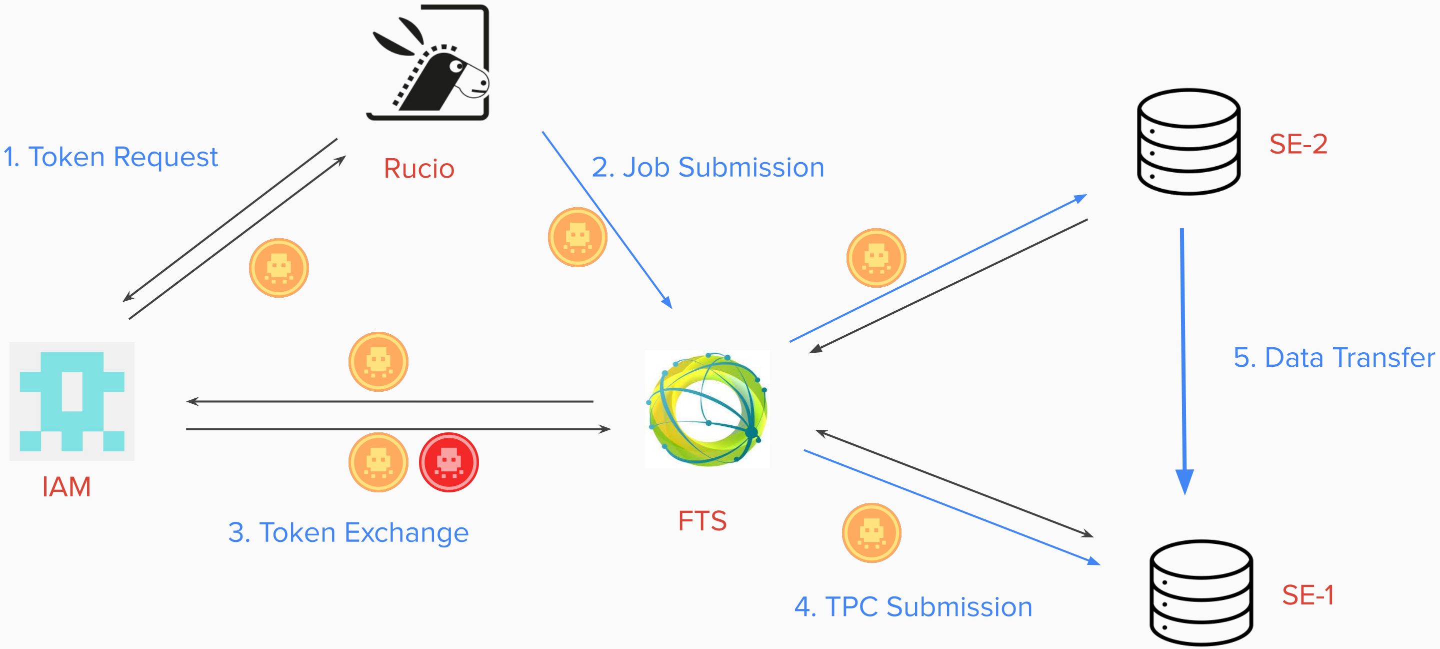
UK SEs

Single Prod RSE per site

**End Goal:**

- Dataset scoped tokens
- About 10k datasets per day
- Spread across ~5 SEs

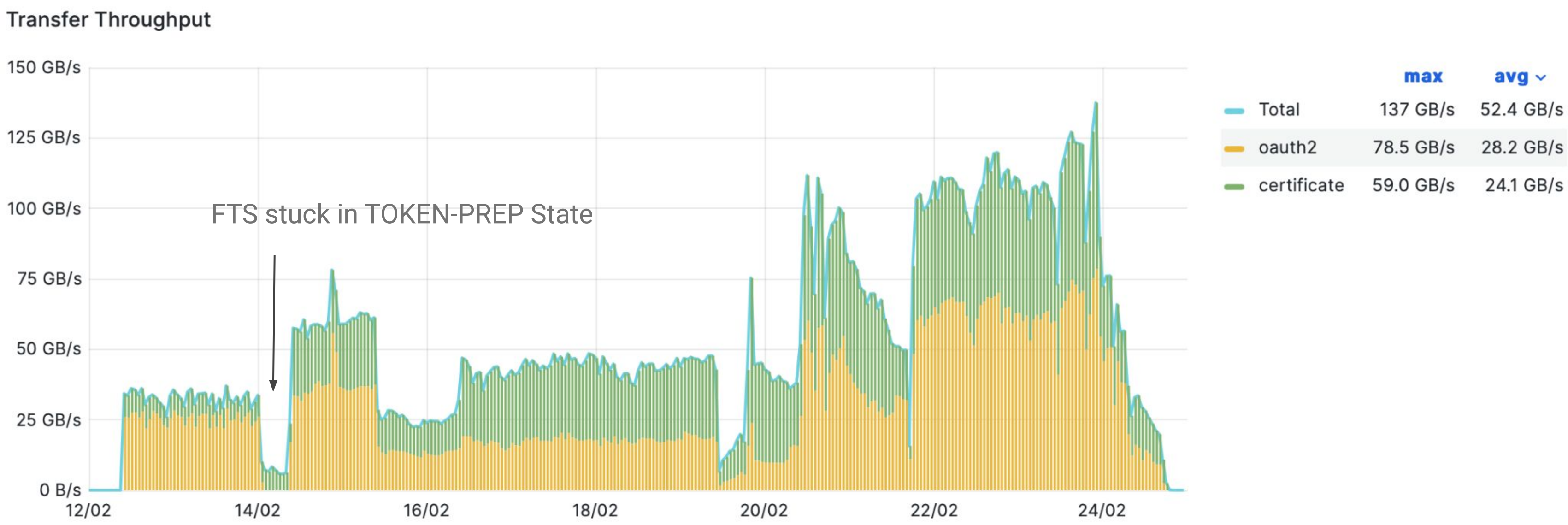
# Rucio - FTS - SE : **Token Exchange flow**



# Tokens in Data Challenge

First step towards tokens:

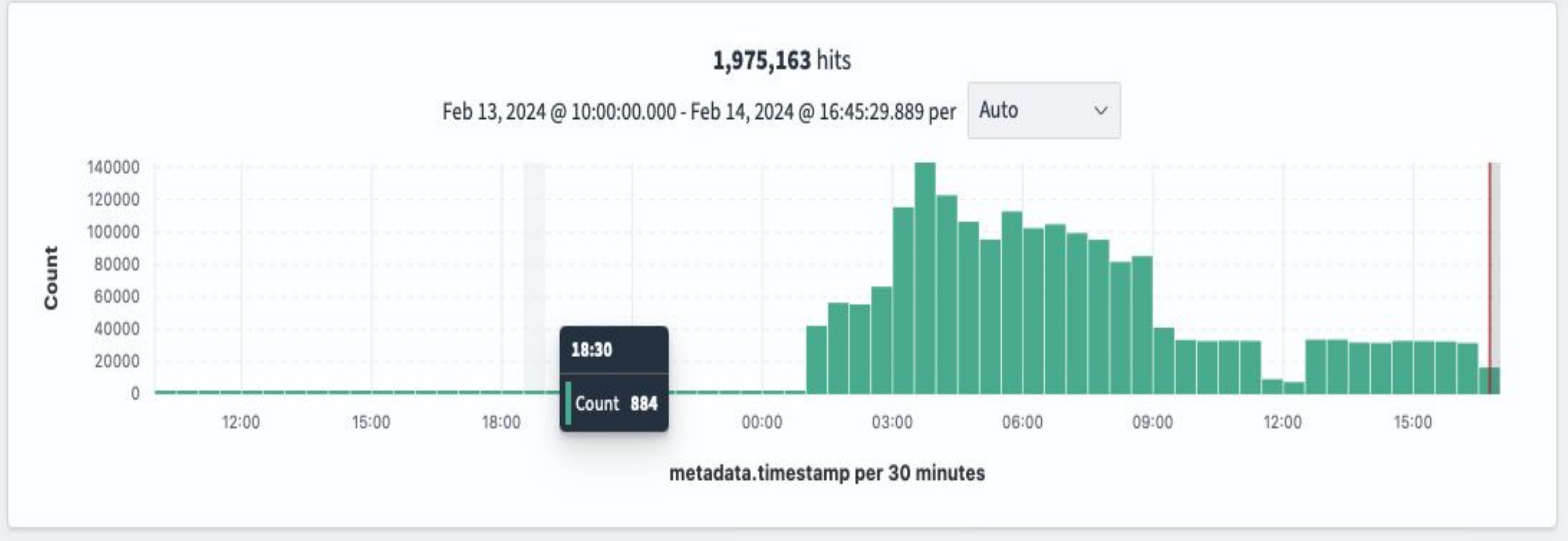
- **Scope** (context in which that token may act): *storage.read/modify: /*
- **Audience** (identifies the recipients that the JWT is intended for): ***wlcg/any***



| | max | avg |
|---|---|---|
| Total | 137 GB/s | 52.4 GB/s |
| oauth2 | 78.5 GB/s | 28.2 GB/s |
| certificate | 59.0 GB/s | 24.1 GB/s |

# Successful DoS attack on IAM

- Tokens not cached in rucio
- Requesting token for every transfer: 60  to 70 Hz



**1,975,163** hits

Feb 13, 2024 @ 10:00:00.000 - Feb 14, 2024 @ 16:45:29.889 per Auto

# New Tests

CMS has "*_Test*" Rucio SEs

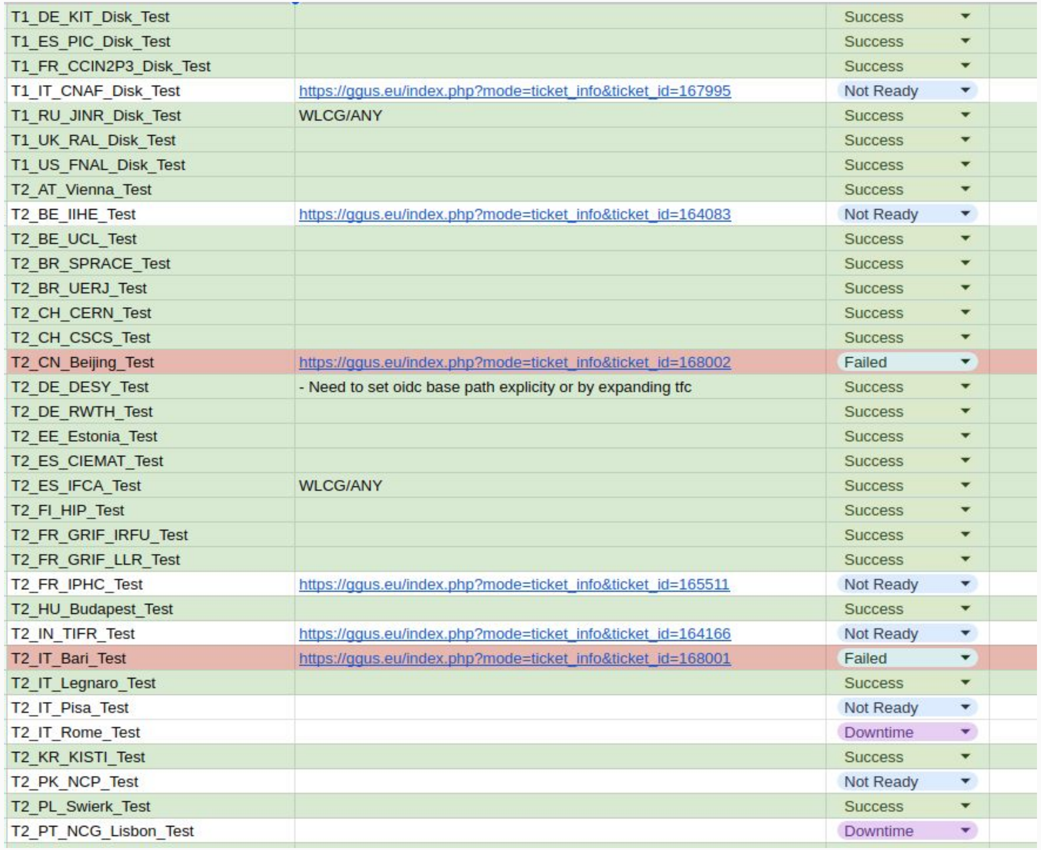
- Same hosts as "*main*" SEs
- Subpath at */store/test/rucio*

60% T1s and T2s were ready

- No Tape
- No T0

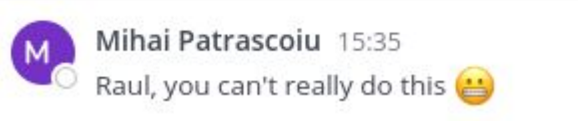
First test:

- 1 Dataset
  - Total files : 1294
  - Total size : 2.279 TB
- Single source
- 38 Destinations



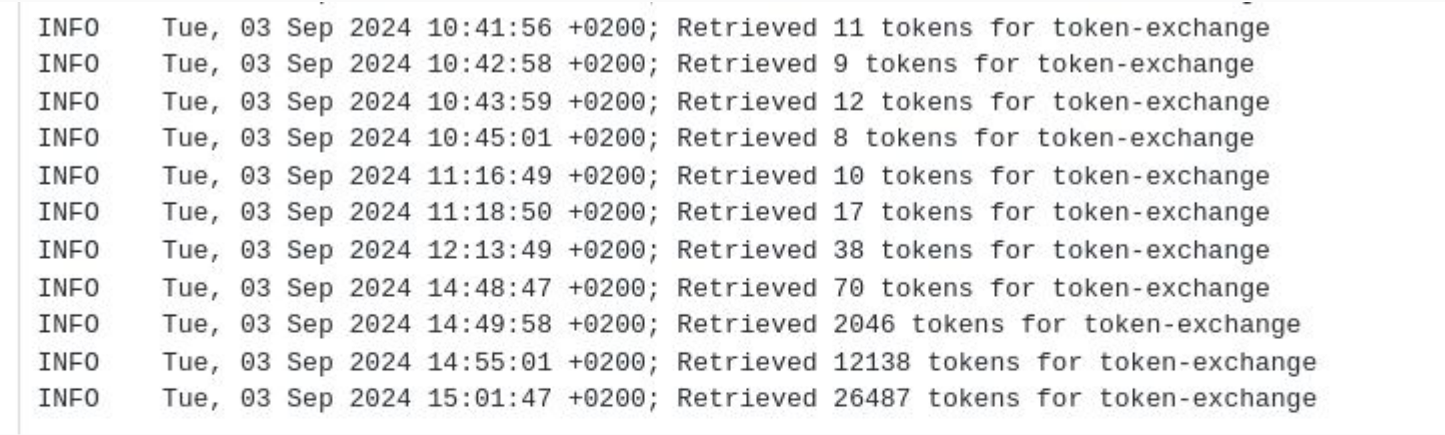
| | | |
|---|---|---|
| T1_DE_KIT_Disk_Test | | Success |
| T1_ES_PIC_Disk_Test | | Success |
| T1_FR_CCIN2P3_Disk_Test | | Success |
| T1_IT_CNAF_Disk_Test | https://ggus.eu/index.php?mode=ticket_info&ticket_id=167995 | Not Ready |
| T1_RU_JINR_Disk_Test | WLCG/ANY | Success |
| T1_UK_RAL_Disk_Test | | Success |
| T1_US_FNAL_Disk_Test | | Success |
| T2_AT_Vienna_Test | | Success |
| T2_BE_IIHE_Test | https://ggus.eu/index.php?mode=ticket_info&ticket_id=164083 | Not Ready |
| T2_BE_UCL_Test | | Success |
| T2_BR_SPRACE_Test | | Success |
| T2_BR_UERJ_Test | | Success |
| T2_CH_CERN_Test | | Success |
| T2_CH_CSCS_Test | | Success |
| T2_CN_Beijing_Test | https://ggus.eu/index.php?mode=ticket_info&ticket_id=168002 | Failed |
| T2_DE_DESY_Test | - Need to set oidc base path explicity or by expanding tfc | Success |
| T2_DE_RWTH_Test | | Success |
| T2_EE_Estonia_Test | | Success |
| T2_ES_CIEMAT_Test | | Success |
| T2_ES_IFCA_Test | WLCG/ANY | Success |
| T2_FI_HIP_Test | | Success |
| T2_FR_GRIF_IRFU_Test | | Success |
| T2_FR_GRIF_LLR_Test | | Success |
| T2_FR_IPHC_Test | https://ggus.eu/index.php?mode=ticket_info&ticket_id=165511 | Not Ready |
| T2_HU_Budapest_Test | | Success |
| T2_IN_TIFR_Test | https://ggus.eu/index.php?mode=ticket_info&ticket_id=164166 | Not Ready |
| T2_IT_Bari_Test | https://ggus.eu/index.php?mode=ticket_info&ticket_id=168001 | Failed |
| T2_IT_Legnaro_Test | | Success |
| T2_IT_Pisa_Test | | Not Ready |
| T2_IT_Rome_Test | | Downtime |
| T2_KR_KISTI_Test | | Success |
| T2_PK_NCP_Test | | Not Ready |
| T2_PL_Swierk_Test | | Success |
| T2_PT_NCG_Lisbon_Test | | Downtime |

# **New Tests:** New token for every job



Mihai Patrascoiu 15:35

Raul, you can't really do this 😬



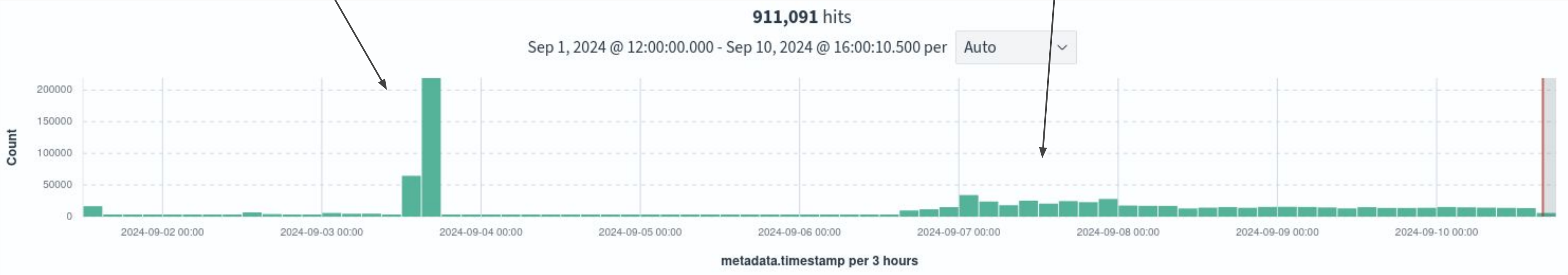
```
INFO    Tue, 03 Sep 2024 10:41:56 +0200; Retrieved 11 tokens for token-exchange
INFO    Tue, 03 Sep 2024 10:42:58 +0200; Retrieved 9 tokens for token-exchange
INFO    Tue, 03 Sep 2024 10:43:59 +0200; Retrieved 12 tokens for token-exchange
INFO    Tue, 03 Sep 2024 10:45:01 +0200; Retrieved 8 tokens for token-exchange
INFO    Tue, 03 Sep 2024 11:16:49 +0200; Retrieved 10 tokens for token-exchange
INFO    Tue, 03 Sep 2024 11:18:50 +0200; Retrieved 17 tokens for token-exchange
INFO    Tue, 03 Sep 2024 12:13:49 +0200; Retrieved 38 tokens for token-exchange
INFO    Tue, 03 Sep 2024 14:48:47 +0200; Retrieved 70 tokens for token-exchange
INFO    Tue, 03 Sep 2024 14:49:58 +0200; Retrieved 2046 tokens for token-exchange
INFO    Tue, 03 Sep 2024 14:55:01 +0200; Retrieved 12138 tokens for token-exchange
INFO    Tue, 03 Sep 2024 15:01:47 +0200; Retrieved 26487 tokens for token-exchange
```

- Bypassing token cache in rucio
- FTS refreshing at about 12 HZ

# **New Tests:** Token Cache enabled

- Test without Rucio cache: 100 Hz for 0.5 hr *then* 30 Hz for 1.5 hrs

Test (In prod) with Rucio Cache: 3 Hz



**911,091** hits

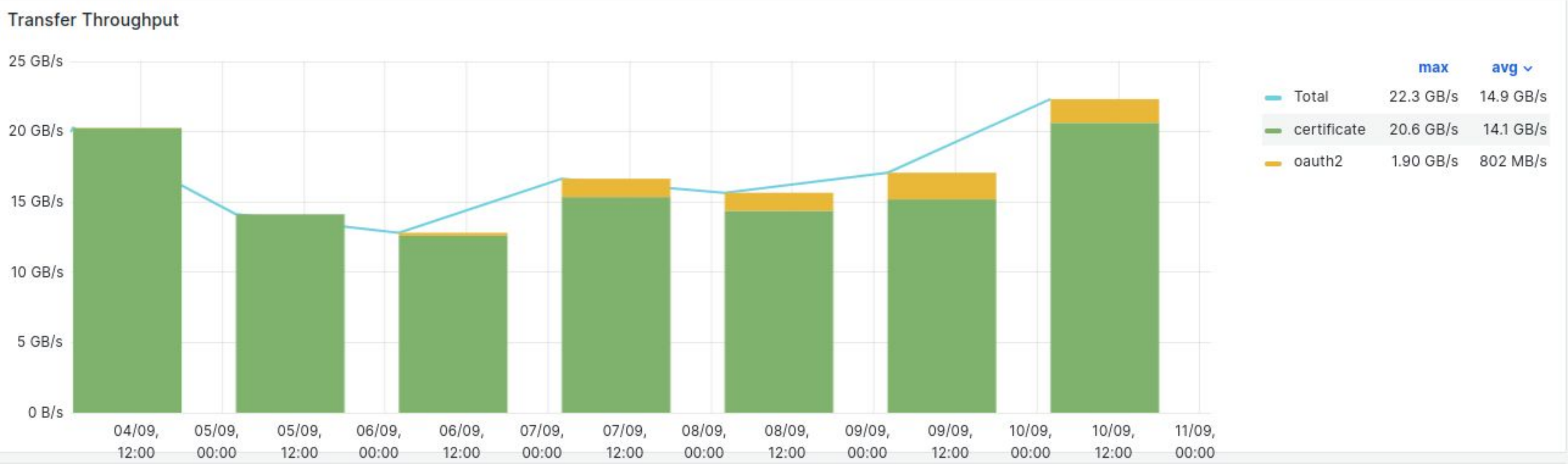
Sep 1, 2024 @ 12:00:00.000 - Sep 10, 2024 @ 16:00:10.500 per Auto

- FTS understanding the token profile and caching based on that would help further reduce duplicate tokens and storage and renewal requirements

- Same goes for IAM

# New Test Results

About 10% of transfers use oauth token authorisation



| | max | avg ⌄ |
|---|---|---|
| Total | 22.3 GB/s | 14.9 GB/s |
| certificate | 20.6 GB/s | 14.1 GB/s |
| oauth2 | 1.90 GB/s | 802 MB/s |

# Tests by IAM Team and Future Plans

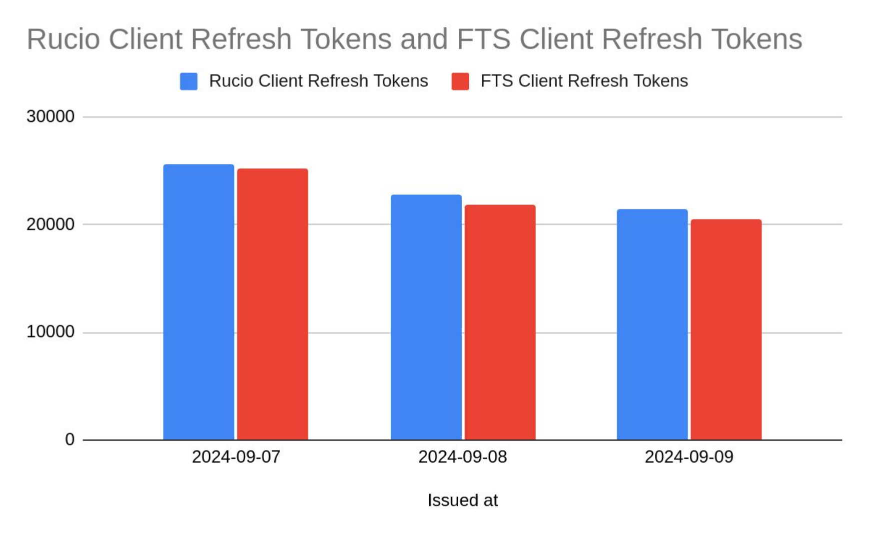
| Pods | Rate (RPS) | Average Response Time (ms) |
|---|---|---|
| **1 Pod** | 350 RPS | 250 ms |
| | 180 RPS | 60 ms |
| **2 Pods** | 600 RPS | 150 ms |
| | 180 RPS | 60 ms |

- Courtesy of Berk Balci

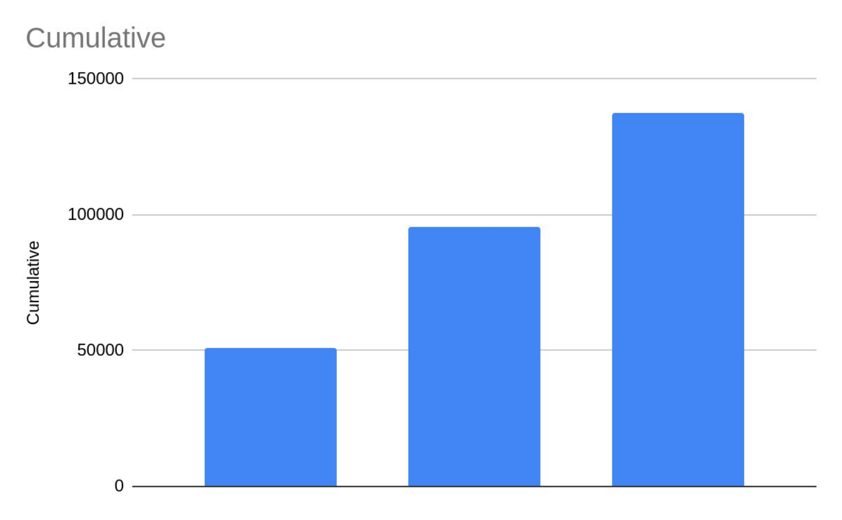
- Currently on production there is a single pod on Openshift.

- “ Our plan is to use 3 pods in HA mode when we migrate to k8s but currently, in our testing environment there are 2 clusters, we are going to deploy the 3rd dev cluster in this week and then I will do tests with 3 clusters “

# What about the DB?



- Can be halved by not issuing refresh tokens to Rucio
- offline_access in scope

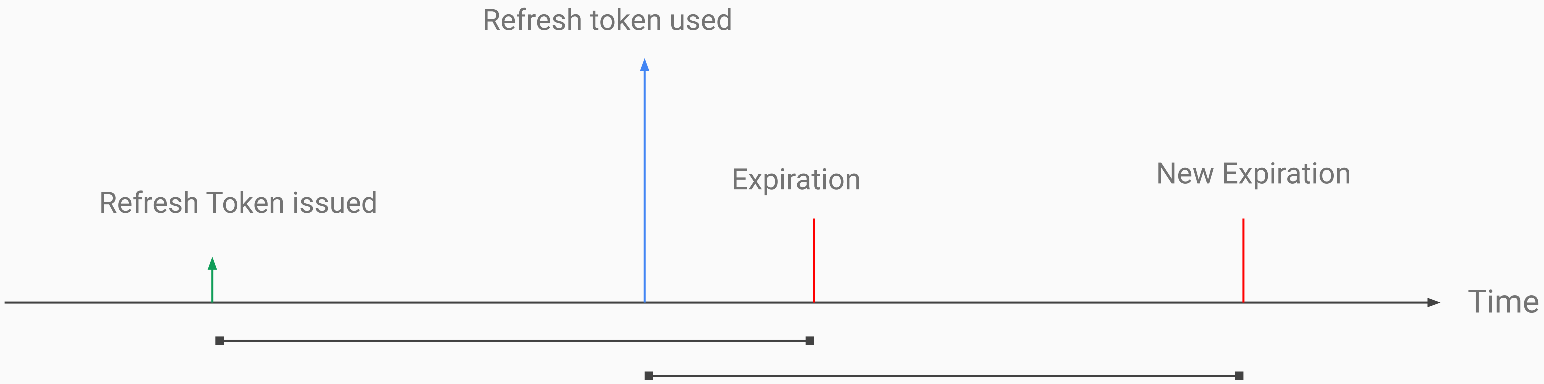


- Expecting saturation at 300k in 4 days

# Miscellaneous

Surprisingly got the spelling right in the first try.
I swear! please believe me!

- /revoke when stale?
- Sliding window expiration for refresh token?



# Optimal OIDC configs?

| Service | Access Token Lifetime | Refresh Token Lifetime |
|---|---|---|
| Rucio | 3 hours | 7 days |
| FTS | 3 hours | 7 days |

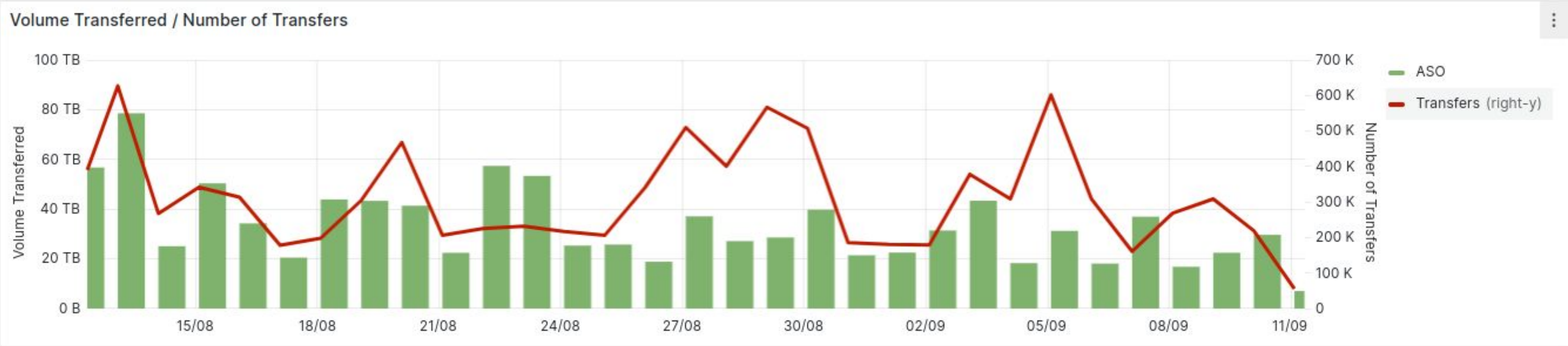
- Short access token lifetimes to reduce chances of *stale tokens* in current FTS refresh procedure

- Do not need 30 days refresh token for Disk transfers

- We envision 6 hr access token lifetimes

# Non-Rucio CMS FTS Clients - CRAB (ASO)

- Stage out of user analysis
  - Low volume and high cardinality
- Rucio to manage the stage out to user area as well
- Tokens will make possible that users still own the files, but rucio can write them
- Impersonation as opposed to delegation[1]



[1] https://datatracker.ietf.org/doc/html/rfc8693#name-delegation-vs-impersonation

# CMS Jobs - Authorisation Code Flow - Vault as a dynamic client

Hashicorp Vault stores long-lifetime credentials and its secret engines allow to acquire (and cache) short-lifetime credentials.

- It is a static-registered client of IAM with confidential credentials.

Users authorize vault, via authorization code flow or device code flow to get a 1 month refresh token.

- In return the user gets a vault token with a lifetime matching the Kerberos ticket-granting ticket (in case of an HTCondor schedd a 1 month vault token) and an access token with a 6 hour lifetime.

# Min Requirements for CMS

- FTS “Just in Time” Token Refresh and subsequent stable release for Fermilab and RAL FTS
- IAM migration to K8s
- Rucio release with policy based scope configuration
- Tape Token Discussions :)

# Thanks!