



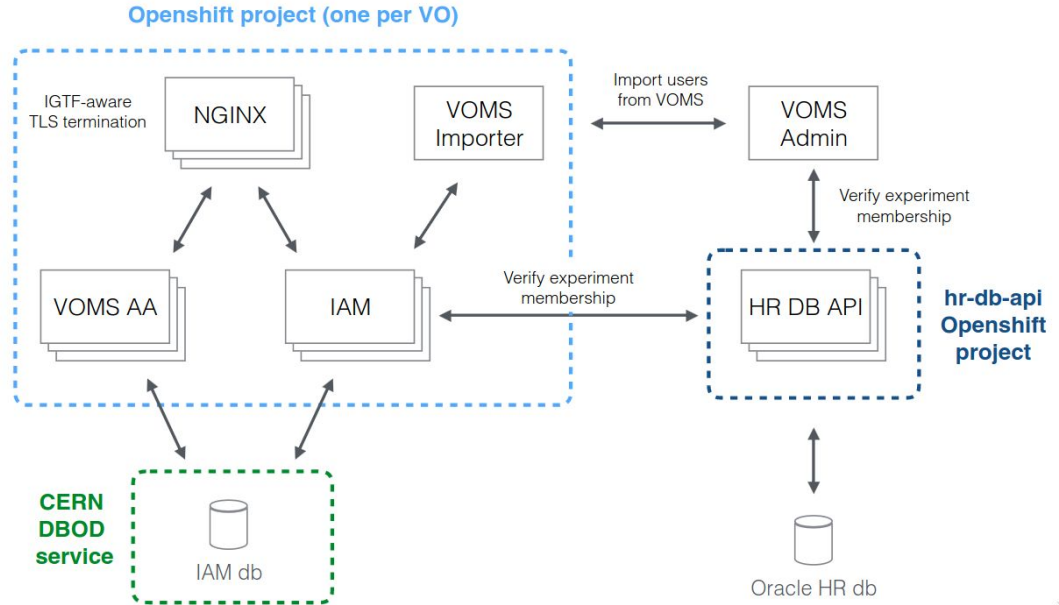
# IAM @ Data Challenge 24

WLCG/HSF Workshop 2024  
14.05.2024

Berk Balci, Maarten Litmaath, Antonio Nappi

# Introduction - IAM

- Token based (OAuth2.0) authentication and authorization service for WLCG
- Uses JSON Web Tokens to authenticate
- Backward compatible for X.509 Certificate authentication



# Preparation for DC24

- Conducted performance tests and stress tests.
- Implemented configuration changes aimed at enhancing performance.
- Increased resource quotas of the instance to optimize performance.
- INDIGO IAM developers (INFN-CNAF) released v1.8.3, incorporating several performance enhancements.
- Implemented additional monitoring tools to track performance metrics effectively.

# IAM Performance

Experiment	Issued tokens	Max. number of tokens in DB	Peak token request rate	Typical token request rates
ATLAS	2.6 M	1.03 M	5 Hz	3 Hz (12 days)
CMS	2.7 M	0.97 M	200 Hz	60 Hz (6 hours), 20 Hz (10 hours), 1 Hz (11 days)
LHCb	3.4 M	1.65 M	120 Hz	25 Hz (2 days), 1 Hz (10 days)

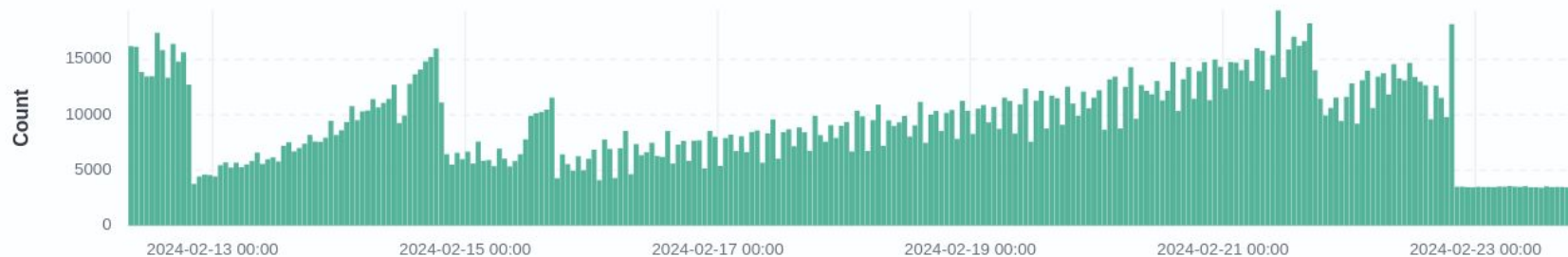
- LHCb tested the “1 token per file transfer” configuration for 2 days which increased their token request rate.
- CMS had high token request rates for ~16h
- During these peak token requests rates on CMS and LHCb, IAM slowed down on issuing tokens
- ALICE instance isn't included in the summaries, as it was not used for any data management operations.

# Token Requests - ATLAS

2,577,632 hits

Feb 12, 2024 @ 08:00:00.000 - Feb 23, 2024 @ 19:00:00.000 per

Hour



metadata.timestamp per hour

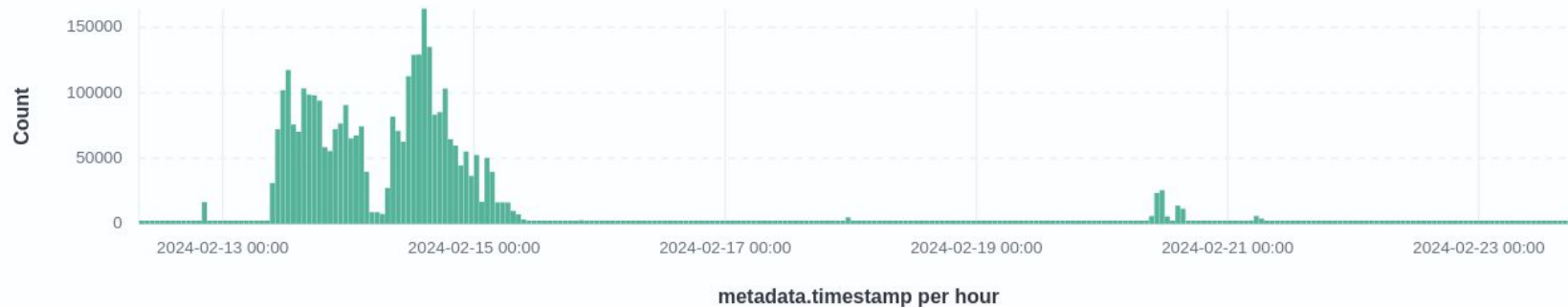


# Token Requests - LHCb

3,376,314 hits

Feb 12, 2024 @ 08:00:00.000 - Feb 23, 2024 @ 19:00:00.000 per

Hour



# IAM Performance - Challenges Faced

- **Database Overload:**
  - Increased token request rates led to database overload, impacting response times.
  - Token lifetimes of up to 30 days delayed cleanup processes during DC24 which led to the database being filled with tokens.
  - The database cleanup algorithm was running slowly and filled up the database connection pool.
- **Token Management:**
  - Suboptimal token usage patterns, especially concerning refresh tokens.



# IAM Performance - Measures Taken

- Performed a comprehensive cleanup of the access token table.
- Increased connection pool size to expedite database cleanup processes.
- Addition of index on refresh token tables to enhance database operations.

# Lessons Learned:

- **Token Lifecycle Management:**
  - Implement shorter token lifetimes to facilitate quicker cleanup processes during peak usage periods.
- **Token Management Enhancements:**
  - Stop storing access tokens in the DB to improve the performance. This needs a modification of token management engine (MitreID). IAM developers are working on this.
- **Collaborative Discussions:**
  - Foreseen discussions between Rucio, DIRAC, FTS, and IAM experts to explore more efficient token orchestration methods for large-scale data transfers.
- **Performance Testing:**
  - Enhance IAM performance tests to make them closer to the real use-cases and include closer examination of latency issues.

# Future plans - Migration to Kubernetes

- It was already planned since a year to move the IAM services from OpenShift to Kubernetes to enhance control, flexibility, HA, GitOps integration, and monitoring capabilities.
- The Kubernetes instances have already been successfully deployed.
  - Existing Openshift instances will be maintained for a transitional period to allow enough time for VOs and Sites to update URIs for both token and VOMS endpoints.
- We already have improvements in load balancing, logging, monitoring, and GitLab integration compared to our OpenShift infrastructure.
- Exploration of high availability (HA) options is underway, including the implementation of multiple clusters across multiple availability zones with transparent load balancing to users. This shows great promise.

## Future plans - Other

- IAM Developers (INFN-CNAF) are working on an update to have access tokens no longer stored in the database.
- Other developments have priority at this time, because of the transition from VOMS-Admin to IAM. They are working on urgent task for this transition:  
<https://github.com/orgs/indigo-iam/projects/8>
- Developing performance tests that incorporate more realistic scenarios involving refresh tokens, with more focus on latency analysis.
- INDIGO IAM Technical Hackathon will take place on the 29-30 May at CNAF

# Summary

- In DC24, the IAM service token endpoints experienced heavy usage from Rucio, DIRAC, and FTS across ATLAS, CMS, and LHCb experiments.
- Despite occasional instances of slow response times and database overload, the service remained available throughout DC24.
- Lessons taken on various topics, we are already working on performance improvements, HA but our main focus is on VOMS-Admin Decommissioning until the end of June.
- Collaborative discussions needed to explore more efficient token orchestration (e.g. token lifetimes, token scopes)



# Thank you for listening!

Questions?