Contribution ID: **76** Type: **Oral Presentation**

Remote S3 storage access using CEPH Rados Gateway (Remote Presentation)

Thursday, 28 March 2024 17:00 (30 minutes)

INFN-CNAF is one of the Worldwide LHC Computing Grid (WLCG) Tier-1 data centers, providing support in terms of computing, networking, storage resources and services also to a wide variety of scientific collaborations, ranging from physics to bioinformatics and industrial engineering.
Recently, several collaborations working with our data center have developed computing and data management workflows that require access to object storage services and the integration with POSIX capabilities.
To accomplish this requirement in distributed environments, where computing and storage resources are located at geographically distant physical sites, the possibility to locally mount a file system from a remote site to directly perform operations on files and directories becomes crucial.
Nevertheless, accessing data must be regulated by standard, federated authentication and authorization mechanisms, such as OpenID Connect (OIDC), which is already adopted within WLCG and the European Open Science Cloud (EOSC).
Starting from such principles, we have implemented a solution that provides fine-grained data access by integrating JSON Web Token (JWT) authentication, provided by INDIGO-IAM as Identity Provider (IdP), Open Policy Agent (OPA) , CEPH Rados Gateway supporting the S3 compatible API and the Security Token Service (STS) for cross-account operations and sts-wire, a Rclone wrap-up to mount cloud storage as a disk.
CEPH RADOS Gateway allows access via OIDC Identity Provider and cross-account operations by offering Security Token Service (STS). In addition, the integration with OPA allows an authorization policy reinforcement: while CEPH administrator creates a role to define both location and type of storage resources available for the identity provider's users, the fine-grained policies are handled by OPA to manage buckets and objects, and to perform S3 operations. This design has allowed us to decouple authorization enforcement from the storage service. Hence, modifying user's access can be executed independently ensuring the availability of the service. Moreover, the policy engine offers scalability that can accommodate increasing resource demanding on geographically distributed infrastructures. The S3 storage made available with the present solution is also accessible via a wide range of client applications such as SDK (Boto3), CLI (s3cmd and sts-wire) and the in-house designed Web Application, developed upon ReactJS, able to exploit the official AWS SDK for JavaScript. The WebApp allows easy access to RADOS Gateway resources, giving the user the ability to list, upload and download file objects, providing authentication with both plain credentials and OpenID connect.
In this work, the design and integration process of the above mentioned solution is presented, together with some examples and related advancements.

Primary authors: ALKANSA, Ahmad (INFN CNAF); COSTANTINI, Alessandro (INFN-CNAF); SPIGA, Daniele; CIANGOTTINI, Diego (INFN Perugia); Dr MICHELOTTO, Diego (INFN); FORNARI, Federico (INFN-CNAF); MALATESTA, Giada; GASPARETTO, Jacopo (INFN CNAF); SGARAVATTO, Massimo; STALIO, Stefano (INFN LNGS)

Presenters: ALKANSA, Ahmad (INFN CNAF); COSTANTINI, Alessandro (INFN-CNAF)

Session Classification: Data Management & Big Data

Track Classification: Track 6: Data Management & Big Data