WP2 Fortnightly Meeting - CERN XCache

Riccardo Di Maria

CERN

October 9th, 2019

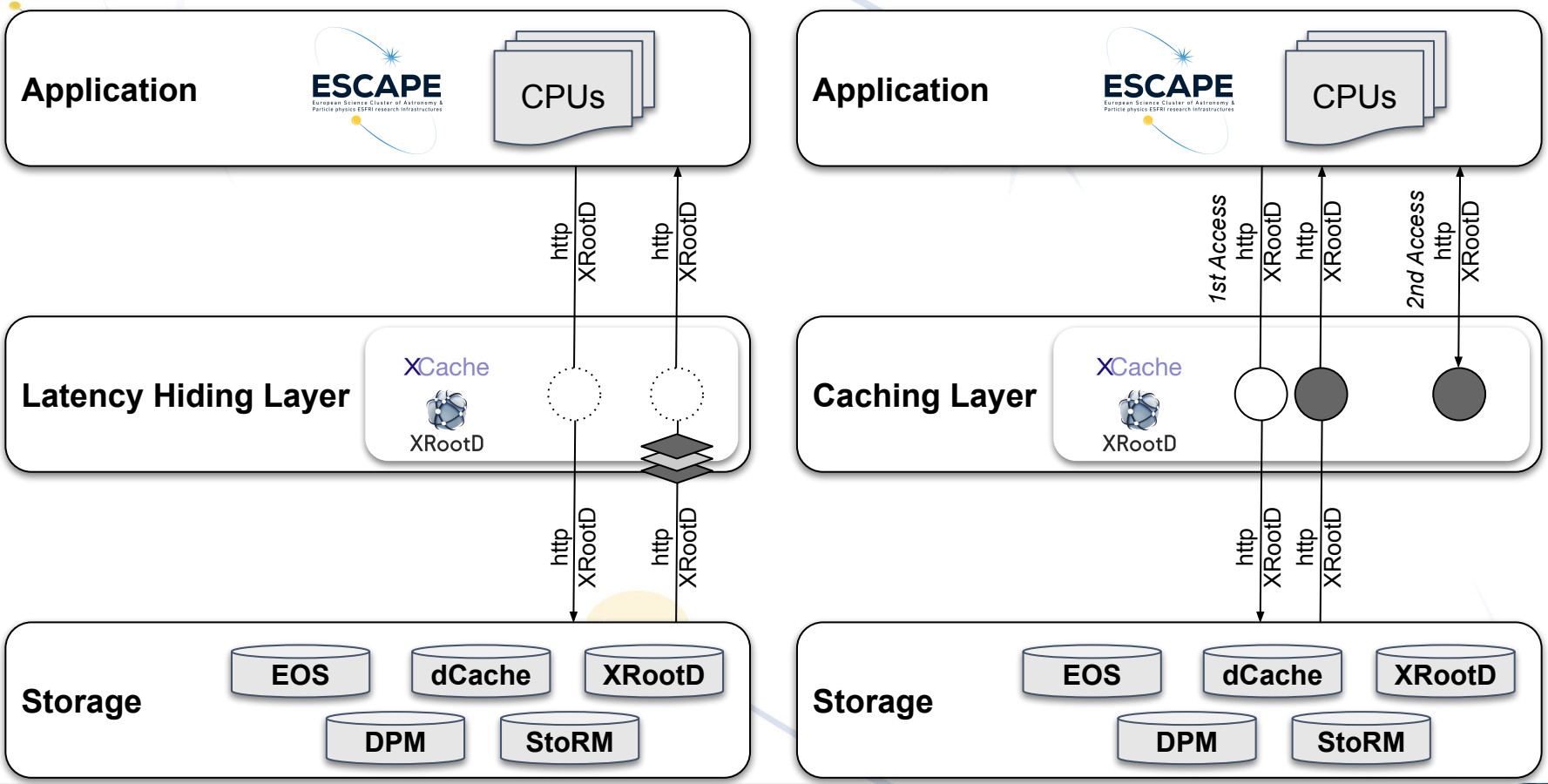
ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Overview

- XCache
- CMS XCache - INFN Experience
- CERN Vanilla XCache - Disk Caching Proxy (DCP) Cluster
- Next Steps - To Do List

Funded by the European Union's Horizon 2020 - Grant N° 824064

XCache



Riccardo.Di.Maria@cern.ch

Funded by the European Union's Horizon 2020 - Grant N° 824064

CMS XCache - Caching On-Demand Effort

- The goal was to acquire knowledge on XCache by reproducing what INFN is currently deploying for CMS.
  - Simple setup at CERN (VM-based); smooth installation for a non-expert.
- The XCache origin point was set as xrootd-cms.infn.it, i.e. the CMS Italian global redirector.
  - xrdcp -f -v xroot://XCache.cern.ch[//xroot://ORIGIN.cern.ch]//path/file.root /dev/null
- Simple Vanilla implementation easily achieved (cf. with ATLAS/CMS “container” approach).
  - Moving towards a Disk Caching Proxy (DCP) cluster.
- The goal is to eventually deploy a caching layer flexible enough to serve both LHC-based experiments and ESCAPE partner experiments.

Funded by the European Union’s Horizon 2020 - Grant N° 824064

CERN Vanilla XCache - DCP Cluster

- Simple VM-based cluster setup (ie. 2-8 GB RAM and 10-40 GB disk).
- First host established as primary server.
  - Once file cached, future calls always redirected to the same cache.
    - PRO: global national primary server - redundancy needed.
    - CONS: work always carried out by the primary server if files are not yet cached.
  - A test to check the behaviour for overloading should be performed.
- If the origin point of the primary server does not host the file, the client request is routed to other caches until the request is fulfilled.

Funded by the European Union's Horizon 2020 - Grant N° 824064

CERN Vanilla XCache - To Do List

- Implement monitoring.
- Integration with other storages within ESCAPE DataLake.
  - Investigate horizontal scaling with several stages, load balancing, etc.
- Investigate other protocols such as http.
- XCache stress test using HammerCloud (stability, reliability, etc.):
  - setup HC jobs from existing analysis functional tests;
  - create specific HC test to stress the storage.
- This is a different operational model wrt ATLAS and CMS.
  - investigate the integration with Rucio as a first step;
  - possibly federated solution with multiple caching layers.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Backup

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Data Infrastructure for Open Science (DIOS)

- Data Lake Infrastructure and Federation Services - Xavier Espinal, CERN
- Data Lake Orchestration Service - Patrick Fuhrmann, DESY
- Integration with Compute Services - Yan Grange, ASTRON-NWO
- Networking - Rosie Bolton, SKAO
- Authentication and Authorization - Andrea Ceccanti, INFN

Simone Campana, CERN as WP leader and Rosie Bolton, SKAO as deputy

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Goals

- Implementing Science Analysis Platforms for EOSC researchers to stage data collections, analyse them, access ESFRIs' software tools, bring their own custom workflows.
- Contributing to the EOSC global resources federation through a Data-Lake concept implementation to manage extremely large data volumes at the multi-Exabyte level.
- Supporting "scientific software" as a major component of ESFRI data to be preserved and exposed in EOSC through dedicated catalogues.
- Implementing a community foundation approach for continuous software shared development and training new generation researchers.
- Extending the Virtual Observatory standards and methods according to FAIR principles to a larger scientific context; demonstrating EOSC capacity to include existing frameworks.
- Further involving SMEs and society in knowledge discovery.

Funded by the European Union's Horizon 2020 - Grant N° 824064