Keeping up With the Evolution of Science

Tigran Mkrtchyan for dCache Team
CHEP2019, Adelaide

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 777367
About dCache

- A distributed petabyte-scale storage system for scientific data
- Supports standard and HEP specific access protocols and authentication mechanisms
- Developed for HERA and Tevatron, used for LHC and others
  - Belle II, LOFAR, CTA, IceCUBE, EU-XFEL, Petra3, DUNE and many more ...
Scientific Data Challenges

- Volume
- Fast ingest
- Chaotic Access
- Sharing
- Access Control
- Persistence & Long term archival
- Immutability
- Data integrity and protection
Keeping up With the Evolution of Science | Tigran Mkrtchyan | 4
User Workflow Shift

- More non HEP tools and POSIX access
  - ROOT $\Rightarrow$ Jupyter Notebook
  - Apache Spark
  - HDF5
- Grow of interactive analysis
  - Analysis Facilities
- Industry standard AuthN
  - OpenID Connect
  - OAuth2
- Hybrid Clouds
- New 3rd-party transfers protocols
- Integration with HPC clusters
XROOTD
- Source/destination support
- GSI authN and delegation
- Interoperability with SLAC xrootd client & server

HTTP
- Source/destination support
- 3rd vendor HTTP server as destination
- X509, Macaroon and SciToken support

dCache 5.2.x is the LTS version with all required changes
- recommended version by DOMA-TPC WG
Caching/Cloud Bursting

- Advanced caching deployment
  - Dataset/protocol based data migration
    - Data for active analysis replicated (NFS, xrootd)
    - Cold data accessed at remote site
  - Cache warm-up
Zone: Geo-location

- Geo-location aware unit
- Dynamically groups services together
- Available in replication rules
- Network topology aware internals communication
  - Always prefer local resources
  -Disconnected operation

```
set storage unit data:resilient@osm -required=2 -onlyOneCopyPer=zone
create pgroup caching-pools -dynamic -tags=zone=A
```
In-transit Encryption

- HTTPS on redirect (upload/download)
  - Like NFS with krb5i and krb5p
- HTTPS on internal copy
  - Pool-to-pool over WAN
  - Zone awareness
Authentication requirements

- OAuth2 and Co.
  - SciTokens
  - OpenID Connect
- Federated IDPs
  - ESCAPE
  - XDC
- Sharing with Macaroons
  - “Adapting ATLAS@Home to trusted and semi-trusted resources” by David Cameron, 15:30 T3
HPC Friendly Enhancements

- Better POSIX (NFS) compatibility
- Scalable byte-range locks
- Listing of large directories
- Squeezing the most out of internal communication
  - “Efficient Message Encoding For Inter-Service Communication” by me, 14:15 T5
UNIX <=> Windows mapping

- Host running Samba configured to use LDAP
  - no user login allowed!
- Samba as domain member
- Custom script for mapping
  - provides UID/GID <=> SID
Nextcloud Instance @ DESY

No single point of failure
dCache as a Storage Backend

- PB-scale storage system
- HA – downtime free maintenance
- No changes in Nextcloud required
- Unique functionality
  - Tape integration
  - File ownership preservation
  - NFS export to selected users
  - Storage events
  - Data visible by all protocols and security flavors
Storage Events

- Trigger actions on user activity
  - Stop polling, Please!
- Storage system becomes a workflow engine
- Producer-consumer model
- For infrastructure
  - Apache Kafka
- For individuals
  - Server Sent Events
Workflow control

Event: new data

store derived file

create derived data

update catalog

extract metadata

by Michael Schuh
Thank You!