

dCache - running a fault-tolerant storage over public networks

Thursday, 22 March 2018 11:00 (30 minutes)

As a robust and scalable storage system, dCache has always allowed the number of storage nodes and user accessible endpoints to be scaled horizontally, providing several levels of fault tolerance and high throughput. Core management services like the POSIX name space and central load balancing components however are merely vertically scalable. This greatly limits the scalability of the core services as well as provides single points of failures. Such single points of failures are not just a concern for fault tolerance, but also prevent zero downtime rolling upgrades. For large sites, redundant and horizontally scalable services translate to higher uptime, easier up-grades, and higher maximum request rates. This becomes more important with growing demand in multi-site distributed deployments.

Since version 2.16 dCache team have made a big effort to move towards redundant services in dCache. The low level UDP based service discovery is replaced with widely-adopted Apache-ZooKeeper, which is a redundant, persistent, hierarchical directory service with strong ordering guarantees. As ZooKeeper itself run in a fault-tolerant cluster mode with strong consistency guarantee, it become a natural place to keep shared state or take a role of service coordination.

Many dCache internal services are updated and can run in replicated mode by providing truly fault-tolerant deployment. However, as any distributed service, dCache is affected by a network partitioning. In terms of so-called CAP theorem, dCache will prefer consistency over availability and return an error or timeout if data consistency can not be guaranteed.

Yet another aspect of distributed deployments, is the security. The different components must be authenticated and communication must be secure. The latest dCache versions provide a mechanism to use standard PKI infrastructure to achieve secure network communication as well as inter-component authentication.

With this presentation we will show how to deploy distributed, fault-tolerant dCache to provide reliable storage. We will discuss some technical details, share experience and lessons learned.

Primary author: Mr MKRTCHYAN, Tigran (DESY)

Co-authors: ROSSI, Albert (FNAL); Dr LITVINTSEV, Dmitry (FNAL); Dr BEHRMANN, Gerd (NDGF); STAREK, Juergen (DESY); Dr SAHAKYAN, Marina (DESY); Dr ADEYEMI, Olufemi (DESY); Dr FUHRMANN, Patrick (DESY/dCache.org); Dr MILLAR, Paul (DESY)

Presenter: Dr FUHRMANN, Patrick (DESY/dCache.org)

Session Classification: Data Management & Big Data Session

Track Classification: Big Data & Data Management