

dCache

Development and Testing on Openstack.



Stefan Bujack, Yves Kemp, Paul Millar, Tigran Mkrtchyan, Marina Sahakyan, Jürgen Starek, Christian Voß, Marus Wengert

dCache is a system for storing and retrieving huge amounts of data, distributed among a large number of heterogeneous server nodes, under a single virtual filesystem tree with a variety of standard access methods. dCache is a joined effort between Deutsches Elektronen-Synchrotron (DESY), Fermi National Accelerator Laboratory and Nordic E-Infrastructure Collaboration.



Integration of the dCache Development and Testing Setup into the DESY IT-Infrastructure

Although dCache is under collaborative development the building, testing and release infrastructure is located at DESY. At the moment it is a rather independent part of the DESY computing centre. The dCache infrastructure consists of several high performance servers. Unfortunately these resources remain unused outside the usual release cycles.

Independent of the dCache developers DESY operates several dCache installations providing about 20 Peta-Bytes of disk space for ATLAS, CMS, ILC, Belle II and the photon science experiments run at the European XFEL and PETRA III. It is also used for the local storage cloud.

These are managed using a combination of Foreman and Puppet for centralised installation and configuration.

Currently the dCache developers and the DESY computing centre are investigating a closer relationship between their independent infrastructures.

As a first step the functionality tests are migrated to the DESY Openstack compute cloud. Rather than on a bare metal machine a test cluster can be spawned using different orchestration methods. Right now, several virtual machines are created with the help of the HEAT module of Openstack.

Using a single template describing the dCache cluster or the client suite the hosts are spawned, receive their DNS entries and are registered with Foreman. As for any machine in the DESY computing centre a role is assigned in Foreman and Puppet manages the installation and configuration.

After the setup is complete the tests are run on the machines in Openstack and the results are reported back to Jenkins for inspection of the developers.

After all tests are completed the spawned machines are deleted and thus the resources are again available for other users.

GitHub

- dCache repository hosted on GitHub
- dCache Project split in several packages

Jenkins

- Pulls code from GitHub
- Builds dCache
- Generates and manages a suite of different tests

Pull Repositories

Run and report test results

Use the Openstack HEAT api to start a small dCache cluster



- Use Openstack infrastructure at DESY
- HEAT Module allows easy orchestration for whole clusters of virtual hosts



- Register hosts with Foreman
- Start puppet run



FOREMAN
puppet

- Foreman manages the host groups for the machines in the test cluster
- Puppet applies configuration according to the host groups provided by Foreman



- Start central services
- Initialise a small storage pool
- Start doors to allow access from outside

Advantages

- Resources are only allocated when actually needed
- The dCache test installation reflects closely the productive environment of sites deploying dCache
- Cloud infrastructure provides great flexibility in testing a wide variety of clients and server setups
- Reliance on infrastructure provided by DESY such as storing and deployment of GRID certificates

Disadvantages

- Additional complexity based on cloud infrastructure
- Latency issues due to spawning and configuration with Openstack and Puppet
- Changes to dCache infrastructure at DESY affects dCache development and testing