Dynamic storage provisioning for elastic cloud services with dCache

International Symposium on Grids & Clouds (ISGC) 2021

Michael Schuh, Patrick Fuhrmann, Tigran Mkrtchyan, Paul Millar, Johannes Reppin, Christian Voss, Tim Wetzel
March 26 2021
Integrate with research infrastructure

5. DNS
6. Network Operations
   • LBaaS
   • Dynamic Certificates
7. Scientific Data, storage
   • dCache
   • High performance storage
8. Event streaming platforms
   • Data Acquisition streams
   • FaaS
9. Scale container registry
   • HPC / HTC
10. Software Repository
    • CVMFS

Source: trailmap.cncf.io
The EGI Federated Cloud

DESY provides resources to the EGI Federated Cloud

Cloud Compute

Run virtual machines on-demand with complete control over computing resources

Syncronised services

- Accounting
- Service discovery
- VM images
- AAI
- DNS (*.desy.fedcloud.eu)

Source: wiki.egi.eu/wiki/Federated_Cloud_user_support
egi.eu/federation/egi-federated-cloud
The ESCAPE Data Lake

Hiding complexity and providing transparent access to data

Heterogenous federated storage and operations model

Source: Slide by Xavier Espinal – PaN ESCAPE Data Management Workshop, 12 January 2012

ESCAPE has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 824064.
PaNOSC - Photon And Neutron Open Science Cloud

EOSC - European Open Science Cloud

EOSC
• FAIR data, effective Open Science

Source: panosc.eu eosc-portal.eu
PaNOSC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852.
The PaNOSC/ExPaNDS (and others) use case

Interactive data analysis environments with Jupyter Notebooks

Science portals
• Find data
• Access data
• Interoperable environments
• Reproducible data analysis

Source: github.com/panosc-portal

PaNOSC has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 823852.
Software as a Service
Containerized applications
Deployments as code

Orchestration
Rancher managed Kubernetes
Helm Package Manager

Containerization
Cloud Native CI/CD
Docker Registry

Infrastructure
Compute Cloud
Storage Systems
Primary Field of Application in Cloud Computing: Container Orchestration with Kubernetes

Openstack
- Used as a virtualization platform

Kubernetes
- Clusters of virtual machines
- For containerized applications, automated deployments and scaling

Source: openstack.org/software
Storage in the Cloud

**CEPH**
- Block Storage for Openstack Cinder (RBD)
  - Disk storage attached to a virtual machine
  - Accessible from attached VM only
- Object Storage for Apps and Openstack Swift (S3)
  - MinIO S3 Proxy: Accessible from anywhere

**dCache**
- Shared file system
  - Disk storage attached to a virtual machine
  - Accessible from many VMs in parallel
- NFS, SMB
- Scale dcache-demo.desy.de to > 1PB
- **Storage for scientific data** (immutable)
Software as a Service
Containerized applications
Deployments as code

Orchestration
Rancher managed Kubernetes
Helm Package Manager

Containerization
Cloud Native CI/CD
Docker Registry

Infrastructure
Compute Cloud
Storage Systems

Software as a Service
Containerized applications
Deployments as code

Orchestration
Rancher managed Kubernetes
Helm Package Manager

Containerization
Cloud Native CI/CD
Docker Registry

Infrastructure
Compute Cloud
Storage Systems

Prometheus
Spark
Grafana
GitLab CI/CD for Container and Cloud Applications

Git as a single source of truth for declarative infrastructure and application

- DevOps Platform
- Auto-scaling CI/CD
- Container Registry

Source: about.gitlab.com
Integrated Docker Registry in GitLab

- Host public and private Container Images
  - Docker (Container Registry)
  - Singularity
    - As Docker Image
    - Singularity Images as build artifacts
Software as a Service
Containerized applications
Deployments as code

Orchestration
Rancher managed Kubernetes
Helm Package Manager

Containerization
Cloud Native CI/CD
Docker Registry

Infrastructure
Compute Cloud
Storage Systems
Rancher Node Templates

- Openstack VMs as k8s nodes
- Node pools
  - Workers
  - Control Plane

```
{
  "name": "apitemplate-test3",
  "driver": "openstack",
  "engineRegistryMirror": [  
    "https://eos-pan-dhub.desy.de:5000"
  ],
  "engineStorageDriver": "overlay2",
  "openstackConfig": {
    "activeTimeout": "200",
    "authUrl": "https://keystone-tank.desy.de:5000/v3/",
    "availabilityZone": "nova",
    "configDrive": "false",
    "applicationCredentialId": "APPLICATION_ID",
    "applicationCredentialSecret": "APPLICATION_SECRET",
    "domainId": "3d1f9e6b4744ac9937c8727163ad560",
    "endpointType": "publicURL",
    "flavorName": "m1.large",
    "imageName": "ubuntu-20-focal",
    "insecure": "false",
    "ipVersion": "4",
    "netId": "eaa0b545b-b1e0-49a7-be18-1a5501ad1758",
    "novaNetwork": "false",
    "region": "RegionOne",
    "secGroups": "ssh,web,container",
    "sshPort": "22",
    "sshUser": "ubuntu",
    "userDotaFile": null
  }
```
Additional software components

“bare” Kubernetes is not enough

**Nginx Ingress Controller**
- Direct traffic to pods

**MetalLB Loadbalancer**
- Level2 Loadbalancer for Kubernetes

**Cinder Storage Class**
- Automatically Provision Volumes in Ceph

**Cert Manager**
- Provides Let’s Encrypt Certificates
- Watches the Kubernetes API for *Ingress* Objects

**dCache**
- Shared Filesystem on cluster nodes
Integrate Kubernetes with Gitlab

Deliver Kubernetes as a Service for GitLab users
- Instance cluster
- Group clusters
- Project clusters
- Users deploy environments for review and production
Managing Apps with Helm

Chart Repository

<table>
<thead>
<tr>
<th>Application</th>
<th>Repository URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>dest-it-helm3</td>
<td><a href="https://charts.desy.de/desy-it">https://charts.desy.de/desy-it</a></td>
</tr>
<tr>
<td>elastic</td>
<td><a href="https://helm.elastic.co">https://helm.elastic.co</a></td>
</tr>
<tr>
<td>gitlab</td>
<td><a href="https://charts.gitlab.io">https://charts.gitlab.io</a></td>
</tr>
<tr>
<td>gitlab3</td>
<td><a href="https://charts.gitlab.io">https://charts.gitlab.io</a></td>
</tr>
<tr>
<td>grafana</td>
<td><a href="https://grafana.github.io/helm-charts">https://grafana.github.io/helm-charts</a></td>
</tr>
</tbody>
</table>

charts.desy.de

- Templated k8s definition files
- Repository for Helm Chart Tarballs
  - Add as Rancher Catalog
  - Install charts as Rancher Apps

---

DESY. Dynamic storage provisioning for elastic cloud services with dCache | Mar 26 2021
Software as a Service
Containerized applications
Deployments as code

Orchestration
Rancher managed Kubernetes
Helm Package Manager

Containerization
Cloud Native CI/CD
Docker Registry

Infrastructure
Compute Cloud
Storage Systems
Managing Apps with Helm

Deployment

- Add DNS alias to Loadbalancer IP
- Add Helm Chart repository
- Customize Values.yaml
- Install to k8s (helm install)

Map Role Based Group Memberships on OIDC Proxy to local accounts and UID/GIDs

- Merge group memberships from user federation in Keycloak
- Export UID/GIDs as non-standard attribute in OIDC Token
- Run Jupyter Servers with UID/GIDs

User Federation
- LDAP (DESY)
- OIDC (EGI Check-in)
- OIDC (Helmholtz AAI)

PetaByte Storage
- NFS Mounts
NFS in elastic cloud environments

• Storage system can’t trust to OS level authentication/mapping
  • Users build and select images VM and Container images
• Storage system can’t trust client’s IP address
  • Use of public networks
  • After disposal VM’s IP returns to shared pool

• NFSv3 based on trusted hosts
  • Server exports based on IP address
  • OS is responsible for proper mapping

• NFSv4.0+
  • Strong authentication is enforced
    • Krb5 + LDAP/AD
  • No kerberos infrastructure provided by public clouds
• Backward compatibility is agreed for migration period

• **On the field, most of sites run NFSv4.0+ in NFSv3 security mode**
Map VM by IP or subnet to a dCache user

```
# /etc/export
/data 10.1.0.0/24(rw,all_squash,anonuid=1001,anongid=1001)
/data 10.2.0.0/24(rw,all_squash,anonuid=1002,anongid=1002)
```
dCache REST interface

- Compatible with OpenStack Manila
- Simple API to manage the export table

Get defined exports

GET https://dcache-demo/v1/exports

Create share ‘/data’

POST https://dcache-demo/v1/exports/data
Summary and outlook

• dCache developers work on better cloud integration
  • Manage shared storage with exports REST API
  • Geo-aware zones
• NFS community works to address cloud challenges
  • RPC-over-TLS
  • 3rd party copy
• Jupyter Hub extensions
  • for ESCAPE Data Lake
  • for Remote Desktops

The 15th International dCache workshop 2021 will take place from 2021-06-01 to 2021-06-02 as a virtual event. [indico.desy.de/event/29564](indico.desy.de/event/29564)
## Contact

**DESY.**
Deutsches Elektronen-Synchrotron  
www.desy.de

Michael Schuh  
Research and Innovation  
michael.schuh@desy.de

Tigran Mkrtchyan  
dcache.org  
tigran.mkrtchyan@desy.de