Gaia Spark analysis platform

Storage technologies

D Morris 1st Feb 2021

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

Hadoop HDFS on Openstack Cinder

Standard deployment well documented in Hadoop manuals

- Original GDAF system – working example
- Cloudera deployment - lots of documentation
- Standard deployment - lots of blogs and howtos

Initial system up and running quickly

- Working on Openstack cloud in 2019
- Automated with Ansible in 2020

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

Hadoop HDFS on Openstack Cinder

Issues with cloud deployment

Hadoop partitions the data into shards spreading the shards across available machines

In a cloud based system, the discs are not local. Cinder storage is often provided by SAN network volumes, not local discs.

A common pattern for Openstack is to use Ceph to provide the backing storage for Cinder volumes.

Ceph will in turn, partition the data into shards and spread the shards across available machines on the storage system.

Two layers of replication, which don't know about each other.

Data ends up being partitioned, sharded and replicated twice.

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

Amazon S3 on Openstack Swift

Amazon S3 is the de facto standard for BigData storage

Very simple web service API

- buckets and objects
- opaque BLOB objects
- flat structure
- no directories

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

Amazon S3 on Openstack Swift

Everyone provides S3 storage

- Amazon
- Azure
- Google
- Digital Ocean
- ….

Everyone can access S3 storage

- command line
- Python
- Java
- C++
- Go
- ….

…. apart from Openstack !?

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

Amazon S3 on Openstack Swift

Openstack object store is Swift

Swift supports a version of S3

- Standard Amazon style - bucket name in hostname
- Swift S3 style - bucket name in URL path
- Still valid, but less common
- Less well documented

S3 command line tools work well

- Uploaded a complete copy of DR2 using command line client

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

Amazon S3 on Openstack Swift

Issues with Hadoop AWS Java library

- Java client library fails with large datasets
- Works fine up to 1/4 of DR2
- More than 1/4 of DR2 and client hangs

- Probably specific to Hadoop-AWS and Openstack Swift
- Need to test Hadoop-AWS with alternative S3 service
- Plan to test with commercial clound providers
- Testing with STFC/IRIS echo service in progress

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

CephFS via Openstack Manila

Storage for Cambridge Openstack is implemented using Ceph

Direct access via CephFS service is available to end users

- Data access is distributed across multiple Ceph nodes
- Single layer of partitioning and sharding

Openstack Manila provides an API to manage access to the Ceph system

We started working with Manila in Kubernetes in August 2020, just at the point where Openstack and Kubernetes were transitioning from platform specific implementations to inter-operable CSI plugins.

Migrated to CSI plugins in September
Identified and reported an issue with the Manila CSI plugin
Another issue pending diagnosis

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

CephFS via Openstack Manila

Static shares for catalog data:
- Gaia
- WISE
- Pan-STARRS
- 2MASS

Dynamic shares for temp data:
- Created on demand
- Lifetime of the Pod that requested them
- Managed by Kubernetes

Static shares isolate data from code:
- Single set of shares provides data for multiple deployments
- Hadoop and Kubernetes deployments both share the same data
- Create update and delete deployments without disturbing the data

Managed storage cluster:
- Backups and replication are SEP

D.Morris
Institute for Astronomy,
Edinburgh University

Storage technologies

Hadoop HDFS on Openstack Cinder

- Standard deploy for Hadoop in cloud
- Well documented
- Performance issues
- “not recommended for production”

Amazon S3 object store

- De facto standard for BigData
- Issues with Openstack Swift and Hadoop-AWS
- Data size limit is a blocking issue

CephFS with Manila

- Good performance
- Isolates data and code
- Managed storage with backups and replication
- Best option so far

D.Morris
Institute for Astronomy,
Edinburgh University