EVALUATION OF THE HUAWEI UDS CLOUD STORAGE SYSTEMS FOR HEP APPLICATIONS

Maria Arsuaga-Rios, Seppo Heikkila and Dirk Duellmann

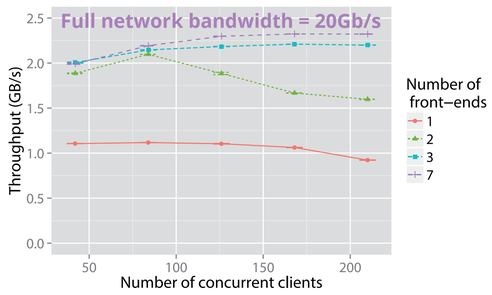
We evaluated the recent UDS version V100R002C00 focusing on scalability in realistic HEP applications such as ROOT analysis and software distribution via CvmFS/S3. Results show that both Huawei storage systems fill the 20 gigabit network bandwidth by obtaining a successful scalability regarding throughput and metadata performance measurements. Moreover, we used the ROOT framework to simulate end user analysis access, which is often characterized by sparse, random access. During this evaluation, we discovered that Amazon S3 and Ceph do not support the multi-range HTTP requests that are commonly used in HEP analysis, when both UDS generations support them and even reach the full network bandwidth in all cases. Both Huawei cloud storage systems have been demonstrated to function as expected as back-end for a large scale software repository hosting nightly builds of the LHCb experiment software.

1ST UDS GENERATION
- 384 disks
- 768 TB
- 7 front-ends
- S3 multipart uploads not supported

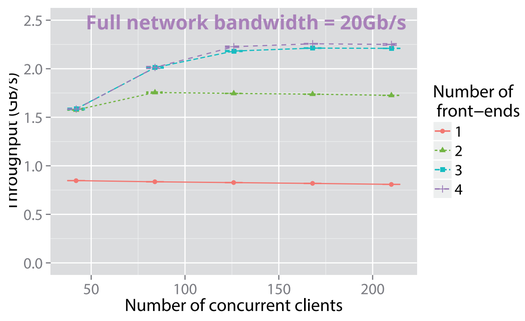
2ND UDS GENERATION
- 300 disks
- 1.2 PB
- 4 front-ends
- Supports S3 multipart uploads
- Less rack space
- More compact storage nodes

1 RAW DATA PERFORMANCE COMPARISON

Each additional front-end node is able to download around 1000 MB per second

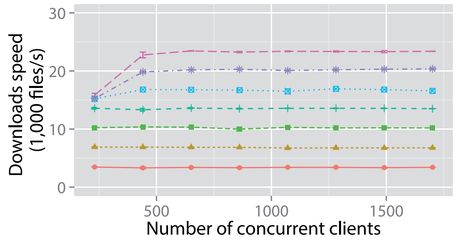


Throughput performance for downloads in the first UDS generation



Throughput performance for downloads in the second UDS generation

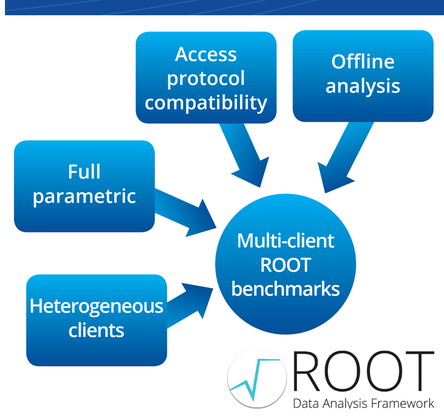
Each front-end node adds linearly around 3500 files per second to the total download rate for both UDS generations



Metadata download speed in the first UDS generation

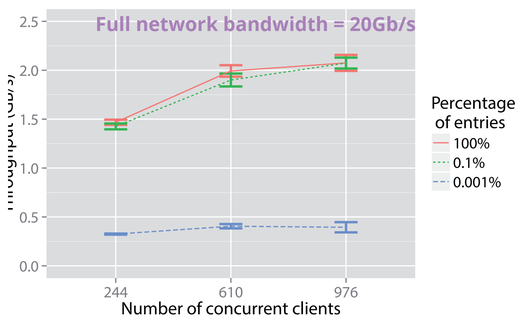
Metadata download speed in the second UDS generation

2 S3 DATA ACCESS PATTERNS WITH ROOT DATA ANALYSIS FRAMEWORK

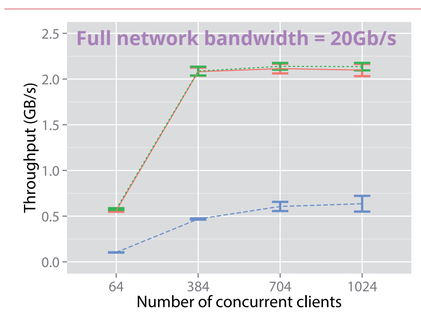


- Concurrent client accesses to real ATLAS ROOT files
- Sparseness, entries accessed
  - 100%, 0,1%, 0,001%
- Two use cases:
  - 1 bucket
  - 64 buckets
- Huawei supports multi-range requests:
  - Ceph and Amazon S3 do not support multi-range requests

Scalability of sparse entry accesses through one bucket in first UDS generation



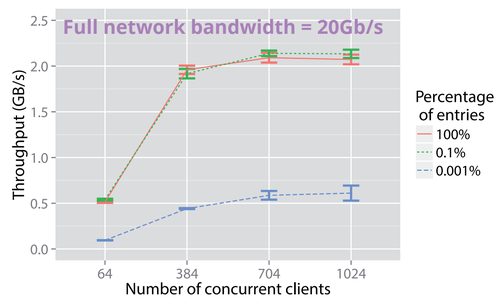
The first UDS generation successfully scales until it reaches the full network bandwidth when reading 100% and 0.1% of entries. In case of 1% of entries, the bucket is over-stressed due to too many clients are accessing to the same small amount of data.



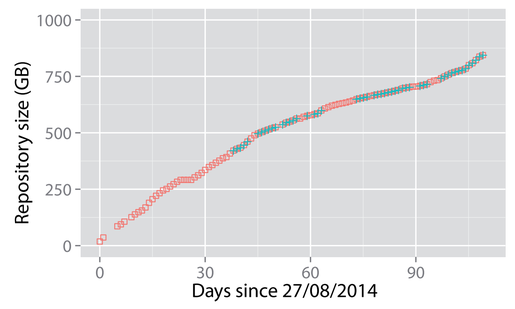
Scalability of sparse entry accesses through 64 buckets in first UDS generation

Scalability of sparse entry accesses through 64 buckets in second UDS generation

Both UDS generations successfully scale until they reach the full network bandwidth.



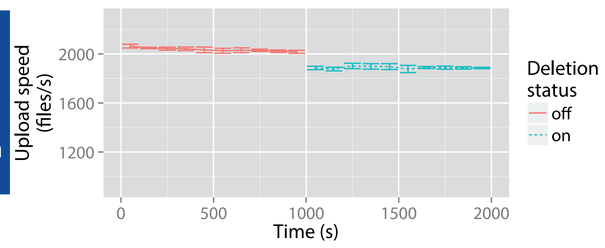
3 S3 INTEGRATION AND TEST DEPLOYMENT IN CVMFS



The CvmFS repository size growth with the first UDS generation (S0) and the second UDS generation (S1)

The CvmFS with S3 storage back-end was tested together with the LHCb experiment by publishing daily the latest version of the LHCb experiment's software stack. The test was run 110 days and in average 162 ± 28 thousand new files occupying 7.45 ± 3.78 GB of data was uploaded daily to the Huawei cloud storage through the S3 API.

Maximum upload performance before and during a deletion process



The S3 back-end ability was evaluated to support the garbage collection feature of CvmFS by deleting files with 400 parallel requests while simultaneously uploading les with the maximum speed. The delete requests decreased the maximum upload speed only slightly, which is expected due to handling of the delete requests.

ACKNOWLEDGMENTS

The authors would like to thank the team of engineers from Huawei, who helped to setup and maintain the UDS system at CERN.

We also would like to thank the CvmFS team and LHCb team, especially Rene Meusel, Jakob Blomer and Ben Couturier, for the help with the CvmFS adaptation and tests.