

Workshop on Cloud Services for File Synchronisation and Sharing

Monday 17 November 2014 - Tuesday 18 November 2014

CERN

Report of Abstracts

Site reports / 10**”Dropbox-like” service for the University of Vienna**

Christian Kracher¹ ; Ralph Staudigl¹ ; Raman Ganguly¹

¹ *University of Vienna*

Corresponding Author(s): raman.ganguly@univie.ac.at

The increasing popularity of dropbox and at the same time increasing awareness for data security did create the demand for an onsite “Dropbox-like” “sync and share” service at the University of Vienna. It has been decided that ownCloud would be a good start, since other academic institutions have been working on an ownCloud based solution as well.

Based on ownCloud enterprise Version 6 the service is currently in test operation with campus wide availability for staff only planned for 12/2014.

Major concerns were the scalability of the storage backend. So instead of using an enterprise storage solution we use Scality’s RING as backend. The RING is an object storage based solution using local storage nodes. Since the ownCloud architecture does so far not allow a RESTbased storage backend we use Scality’s FUSE connector to simulate a virtually limitless filesystem (POSIX). Based on the experiences reported by other academic facilities and our own, our main concerns have been database performance-scalability, storage backend architecture and general software design. Some of which might already have been addressed by ownCloud community version 7. It’s also noteworthy, that ownCloud’s support team responds properly to submitted bug reports. The admittedly limited user feedback has been quite positive so far.

A other issue which have to be solved is the legal issue: What happened with the data, which are shared, after a staff member leaves the University? We want to establish special terms of use for this service, which everybody who want to use the ownCloud service have to accept.

Technology and research / 31**Adaptive Query Processing on RAW Data**

Corresponding Author(s): miguel.branco@epfl.ch

Introduction and Keynotes / 26**Big Data Storage Technologies**

Corresponding Author(s): andreas.joachim.peters@cern.ch

Technology and research / 16**Combining sync&share functionality with filesystem-like access**

Author(s): Maciej Brzezniak¹

Co-author(s): Adam Zawada² ; Gracjan Jankowski² ; Kamil Guryń³ ; Michał Jankowski² ; Stanisław Jankowski⁴ ; Sławomir Zdanowski⁵

¹ *p*² *Poznan Supercomputing and Networking Centre*³ *Bialystok Univesity of Technology*⁴ *Poznań Supercomputing and Networking Centre*⁵ *Poznań Supercompuing and Networking Centre***Corresponding Author(s):** maciekb@man.poznan.pl

In our presentation we will analyse approaches to combine the sync & share functionality with file system-like access to data. While relatively small data volumes (GBs) can be distributed by sync&share application across user devices such as PCs, laptops and mobiles, interacting with really large data volumes (TBs, PBs) may require additional remote data access mechanism such as filesystem-like interface. We will discuss several ways for offering filesystem-like access in addition to sync & share functionality. Today's sync & share solutions may employ various data organisation in the back-end including local and distributed file systems and object stores. Therefore various approaches to providing the client with filesystem-like access are necessary in these systems. We will present possible options to integrate the filesystem-like access with sync&share functionality in the popular sync&share system. We will also show a NDS2 project solution where data backups and archives are kept secure in the distributed storage system and are available through virtual drives and filesystems for Windows and Linux, while users are enabled to securely sync&share selected data using dropbox-like application.

Technology and research / 22

Current practical experience with the distributed cloud data services

George Jaroslav Kremenek¹¹ *u*

We are currently witnessing data explosion and exponential data growth. I will talk about real world experience with very large data sets storage and services. We are storing Peta and in the near future Exa bytes and hundreds or thousands millions of data sets. The one problem is very large number of data objects. File systems were not created to effectively manage thousands of million data items. Inode space is often limited. Storing such large data sets is costly when using rotating storage. Electricity bills for cooling and spinning disks can be prohibitive. Therefore we prefer using magnetic tape technology for the lowest tier of our HSM solution. At OPESOL we employ the LTFS to overcome the old tape technology based limitations and we can provide full POSIX I/O capabilities even for data stored on magnetic tapes

The management of masses of data is a key issue: should both ensure the availability and continuity of access for ever. Scientific collaborations are usually geographically dispersed, which requires the ability to share, distribute and manage efficiently and securely. The media, hardware and software storage systems used can differ greatly from one treatment center to another. At this heterogeneity, are added the continuing evolution of storage media (which induces physical migration) and technological developments in software (which may involve changes in the naming or data access protocols the latter).

Such environments can take advantage of middleware for the management and distribution of data in a heterogeneous environment, including virtualizing storage, that is to say, hiding the complexity and diversity of systems underlying storage while federating data access.

Virtual distributed hierarchical storage system, data grids or data clouds require using and re-using existing underlying storage systems. Creating completely new vertically integrated system is out of the question. iRODS based solutions can take advantage of existing HSM like IBM's HPSS and TSM, SGI DMF, ORACLE (SUN) SAM QFS or emerging cloud storage system like Amazon S3, Google, Microsoft Azure and other Good middleware distributed cloud service for very large data sets should work with all main existing such system and be extensible enough to support main future systems. iRODS (integrated Rule based Data System) is being developed for over 20 years mainly by the DICE

group bi-located at the University of California at San Diego and the University of North Carolina at Chapel Hill. iRODS provides a rich palette of management tools (metadata extraction, data integrity and more). iRODS can interface with virtually unlimited existing and even future storage technologies (mass storage systems, distributed file systems, relational databases, Amazon S3, Hadoop and more). iRODS is company agnostic and the users have all the source code. Migration from one storage resource to another (new) one is just one iRODS command regardless of the data size or number of data objects.

But what makes iRODS particularly attractive is its rules engine that has no equivalent among its competitors. The rules engine allows complex tasks at data management. These policies are remote management of the server side: for example, when data is stored in iRODS, background tasks can be triggered automatically on the server side such as replication across multiple sites, data integrity checks, post-treatment on them (metadata extraction) without specific action on the client side. So, the management policy data is virtualized. This virtualization ensures strict rules set by users, regardless of location data or application that accesses iRODS.

iRODS like systems can deliver full vertical data storage stack including complex tape system management using the existing standard LFTS technology. OPESOL (Open Solutions Inc.) delivers such a system for free. This solution is using LTFS (http://en.wikipedia.org/wiki/Linear_Tape_File_System) which exists on all modern tape drives and tape libraries

I will talk today about sever sites which have chosen a data grid system based on the iRODS (Rule-Oriented Data management) system. iRODS provides a rule-based system management approach which makes data replication much easier and provides extra data protection. Unlike the metadata provided by traditional file systems, the metadata system of iRODS is comprehensive and extensible by user and allows users to customize their own application level metadata. Users can then query the metadata to find and track data.

The iRODS is used in production at L'Institut national de physique nucléaire et de physique des particules (IN2P3).

The Computing Center of IN2P3 (CCIN2P3) offers iRODS service IN2P3 since 2008. This service is open to all who wish to use it. Currently, 34 groups in the fields of particle physics (BaBar, dChooz ...), nuclear physics (Indra, Fasia ...), astroparticle physics and astrophysics (AMS, Antares, Auger, Virgo, OHSA ...), Science Human and Social (Huma-Num) and biology, using the iRODS service CC-IN2P3 for the management and dissemination of data. The CC-IN2P3 also provides hosting the central catalog of iRODS newly created service of France Grille, as well as support administrators to France on the use of Grid technology.

The iRODS service has its own disk servers and is interfaced with our HPSS mass storage (storage on magnetic tape) currently managing over 8 petabytes of data, making it the largest volume service identified internationally.

The service is federated with other services such as iRODS SLAC example. In this perspective, it is also quite possible to federate storage servers available in laboratories with iRODS service CC-IN2P3.

The BNF (Bibliothèque nationale de France)

The Bnf is using iRODS together with open (closed) SAM QFS to store hundreds of million books in its Long-term data preservation. BnF is using iRODS to provide a distributed private data cloud where multiple replicas of data sets are kept at primary BnF site in Paris and secondary site about 40 km From Paris. BnF created a tool to implement its policies for digital preservation SPAR System (Distributed Archiving and Preservation), launched in May 2010, and is continually updated with new collections and feature. BnF employs SPAR and Gallicca for the WEB interface to the distributed private data cloud in iRODS.

The NKP and NDK site and project (Czech National Library, Czech National Digital Library).

I have helped to implement iRODS together with Fedora Commons and other tools at NKP in Prague Czech republic as a base for the EU funded digital library project. The system is now in a full production. It is using IBMS GPFS and TSM as a base layer for its HSM.

The system stores over 300 million data objects. Its data comes nonstop from scanning paper books, electronic data input from Born Digital documents, constant WEB archiving of the ".cz" domain and from all Czech TV and radio broadcasts among others.

Data Management Services for VPH Applications

Author(s): Marian Bubak¹

Co-author(s): Adam Belloum² ; Dmitry Vasyunin² ; Piotr Nowakowski¹ ; Spiros Koulouzis²

¹ AGH University of Science and Technology, Krakow, Poland

² University of Amsterdam

The VPH-Share project [1] develops a cloud platform for the Virtual Physiological Human (VPH) research community. One of the key challenges is to share and access large datasets used by medical applications to transform them into meaningful diagnostic information. VPH researchers need advanced storage capabilities to enable collaboration without introducing additional complexity to the way data are accessed and shared.

In the VPH-Share cloud platform [2], the data storage federation [3] is achieved by an aggregation of data resources in a client-centric manner and exposing it via a standardized protocol that can be also mounted and presented as a local storage so a kind of a file system abstraction is provided. There is a common management layer that uses loosely coupled and independent storage resources and with such a layer a variety of storage resources such as simple file servers, storage clouds and data grid may be aggregated exposing all available storage. As a result, distributed applications have a global shared view of the whole available storage space. This feature makes the development, deployment and debugging of applications considerably easier. The platform is also equipped with mechanisms for validating data reliability and integrity in a heterogeneous distributed cloud environments.

The VPH-share cloud platform has become a full-fledged environment in which VPH data and applications are deployed and shared. The system is protected by an integrated security solution which uses BiomedTown [4] as its OpenID Identity Provider.

Applications which have already been deployed within the VPH-Share platform include a.o. the ViroLab Comparative Drug Ranking System, the @neurIST morphological workflow, and the OncoSimulator tool [5].

Acknowledgment: This work was supported by the EU FP7 project VPH-Share (269978).

References

1. VPH-Share project website: <http://www.vph-share.eu/>
2. Nowakowski P, Bartynski T, Gubala T, Harezlak D, Kasztelnik M, Malawski M, Meizner J, Bubak M: Cloud Platform for Medical Applications, eScience 2012
3. Spiros Koulouzis, Dmitry Vasyunin, Reginald Cushing, Adam Belloum, Marian Bubak: Cloud Data Federation for Scientific Applications. Euro-Par Workshops 2013: 13-22
4. BioMedTown Biological Research Community: <https://www.biomedtown.org/>
5. Atmosphere platform: <http://dice.cyfronet.pl/products/atmosphere>

Technology and research / 19

DataNet: A flexible metadata overlay over file resources

Bartosz Wilk¹ ; Daniel Harezlak² ; Maciej Pawlik¹ ; Marek Kasztelnik¹ ; Marian Bubak²

¹ ACC CYFRONET AGH

² A

Corresponding Author(s): d.harezlak@cyfronet.pl

Managing and sharing data stored in files results in a challenge due to data amounts produced by various scientific experiments [1]. While solutions such as Globus Online [2] focus on file transfer and

synchronization, in this work we propose an additional layer of metadata over file resources which helps to categorize and structure the data, as well as to make it efficient in integration with web-based research gateways. A basic concept of the proposed solution [3] is a data model consisting of entities built from primitive types such as numbers, texts and also from files and relationships among different entities. This allows for building complex data structure definitions and mix metadata and file data into a single model tailored for a given scientific field. A data model becomes actionable after being deployed as a data repository which is done automatically by the proposed framework by using one of the available PaaS (platform-as-a-service) platforms and is exposed to the world as a REST service, which can be accessed from any computing site or a personal computer through the HTTP protocol. Data stored in such a repository can be shared by using various access policies (e.g. user-based or group-based) and can be managed from a wide range of applications. The repository is a self-contained application which can be scaled to improve transfer throughput and can integrate many underlying file storage technologies (currently it supports the GridFTP protocol). The generated REST interface allows data querying and file transfers directly from user web browsers without going through additional servers (this is possible thanks to using the CORS mechanism which is now supported by all major web browsers including mobiles).

Using a PaaS platform as a deployment base for the repository gives an advantage of extending it with different metadata storage backends which can be more suitable for handling metadata schema of certain data models while keeping the source model unchanged. The framework supports it by a plugin system for different storage backends. Such flexible approach allows to adapt the platform to specific requirements without rewriting everything from scratch.

Using a single web endpoint for a repository gives the impression of using a cloud-based service to end users and other services (user credential delegation is also supported) while reusing existing storage facilities maintained in computing centers.

Acknowledgements. This research has been supported by the European Union within the European Regional Development Fund program POIG.02.03.00-00-096/10 as part of the PL-Grid Plus project.

References

- [1] Witt, S.D., Sinclair, R., Sansum, A., Wilson, M.: Managing large data volumes from scientific facilities. *ERCIM News* 2012(89) (2012)
- [2] Foster, I.: Globus Online: Accelerating and Democratizing Science through Cloud-Based Services, *Internet Computing, IEEE*, vol. 15, no. 3, pp. 70,73, May-June 2011
- [3] Haręźlak, D., Kasztelnik, M., Pawlik, M., Wilk, B., and Bubak, M.: A Lightweight Method of Metadata and Data Management with DataNet, *eScience on Distributed Computing Infrastructure*, Eds. Bubak, M., Kitowski, J., Wiatr, K., Springer International Publishing, Lecture Notes in Computer Science, vol. 8500, 2014, pp. 164-177

Technology and research / 20

Dynamic Federations: scalable, high performance Grid/Cloud storage federations

Author(s): Fabrizio Furano¹

Co-author(s): Adrien Devresse¹ ; Oliver Keeble¹

¹ CERN

Corresponding Author(s): fabrizio.furano@cern.ch

The Dynamic Federation project aims to give tools and methods to on-the-fly federate different storage repositories whose content satisfies some basic requirements of homogeneity.

The Dynafeds have been designed to work in WANs, have given so far excellent results also in LAN, and are well adapted to working with the HTTP/WebDAV protocols and derivatives, thus including a broad range of Cloud storage technologies.

In this talk we will introduce the system and its recent larger deployments, and discuss the improvements and configurations that can make it work seamlessly with Cloud storage providers.

Among the deployment possibilities we cite:

- seamlessly using different cloud storage providers at once, thus creating a federation of personal cloud storage providers
- boosting client data access performance by optimizing redirections to data that is globally replicated
- easy, catalogue-free insertion/deletion of transient endpoints
- seamlessly mixing cloud storage with WebDAV-enabled Grid storage
- giving a WAN-distributed WebDAV access backend to services like OwnCloud/CERNBox, to enable collaboration across administrative domains/sites.

Introduction and Keynotes / 29

Evolving CERN data management services

Corresponding Author(s): massimo.lamanna@cern.ch

Software solutions / 5

IBM Software Defined Storage and ownCloud Enterprise Edition - a perfect match for hyperscale Enterprise File Sync and Share

Harald Seipp¹ ; Oliver Oberst¹

¹ IBM

Corresponding Author(s): seipp@de.ibm.com

IBM Software Defined Storage, in particular the technology offering codenamed Elastic Storage (based on GPFS technology) has proven to be an ideal match for Enterprise File Sync and Share (EFSS) solutions that need highly scalable storage. The presentation will provide insight into the integration of Elastic Storage with the ownCloud Enterprise Edition (based on Open Source technology) software that showed impressive scalability and performance metrics during a proof-of-concept phase of an installation that is supposed to serve 300000 users when fully deployed.

Technology and research / 11

OpenStack Swift as Multi-Region Eventual Consistency Storage for ownCloud Primary Storage

David Jericho¹

¹ AARNet

As more users adopt AARNet's CloudStor Plus offering within Australia, interim solutions deployed to overcome failures of various distributed replicated storage technologies haven't kept pace with

the growth in data volume. AARNet's original design goal of user proximal data storage, combined with national and even international data replication for redundancy reasons continues to be a key driver for design choices.

AARNet's national network is over 90ms from end to end, and accommodating this has been a key issue with numerous software solutions, hindering attempts to provide both original design goals in a reliable real-time manner.

With the addition of features to the ownCloud software allowing primary data storage on OpenStack Swift, AARNet has chosen to deploy Swift in a nation spanning multi-region ring to take advantage of Swift's eventual consistency capabilities and the local region quorum functionality for fast writes.

The scaling capability of Swift resolves the twin problems of geographic redundancy, and user proximal access while scaling into the petabyte range.

Significantly, the ring and policy capabilities allow overflow into short or medium term use secondary storage, such as the Australian RDSI project nodes, or Amazon S3. This helps deliver predictable near linear growth of capital and operational expenditures while allowing a higher rate of growth of data volume.

Additionally, the policy capabilities within Swift combined with the ability to grant a user multiple storage targets within ownCloud, allow us to honour data sovereignty rules with respect to the physical location of the data on a per top level folder basis.

Finally, using the combined read and write affinity features of the Swift proxy, AARNet is presently experimenting with deployment of flash cache backed site local application nodes, providing the perception to the user of near instant data ingestion, while the node trickle uploads data to the redundant bulk storage ring using private layer 3 networks.

By switching to true object store systems, AARNet is able to achieve two of its original design goals for cloud storage services, chiefly being user proximal data storage, and continent-spanning geographic redundancy, from hundreds of terabytes into the petabyte scale.

Site reports / 32

Polybox at ETH Zurich

Corresponding Author(s): tilo.steiger@id.ethz.ch

Software solutions / 9

PowerFolder – Peer-to-Peer powered Sync and Share

Christian Sprajc¹

¹ *p*

Corresponding Author(s): sprajc@powerfolder.com

PowerFolder is a peer-to-peer (P2P) sync and share solution which started as spin-off from the University of Cologne and University of Applied Science Niederrhein in 2007. It is available as commercial and open-source solution and in use by hundreds of education and research organization and several thousand businesses. The software enables datacenter providers, NRENs or any education and research organization to operate its own PowerFolder cloud as alternative to public clouds while preserving the same end-user experience: Access to data anywhere on any device (Windows, Linux, Apple, Web, Android and iOS). While approaches to sync and share data from/to a single

central location have several drawbacks the PowerFolder solution offers a unique peer-to-peer algorithm to replicate and transfer data between sites, users and devices with the freedom to choose whether to store or not to store files at the central hub. This is achieved by intelligent; decentralize meta- and binary-data handling between nodes in a self-organizing hybrid peer-to-peer network. The approach increases the overall performance and scalability of the architecture and reduces time to replicate datasets. For security reasons data-access permissions and (federated) AAI remain centrally managed. The talk will introduce technical details of the implementation, outline existing scientific use-cases/installations and further developments of PowerFolder.

Software solutions / 1

Presentation of Pydio

Charles du Jeu¹

¹ *u*

Corresponding Author(s): charles@pydio.com

Pydio is an open-source platform for file sharing and synchronisation. It is widely used by enterprises and organizations worldwide, including major universities.

Pydio comes with a rich web-based interface, native clients for both iOS and Android mobile devices, and lately a new sync client for desktop platforms. Pydio v6 is of great useability and introduces many new features that makes it the ultimate sharing machine.

Its architecture makes it a perfect fit for either easily deploying an on-premise dropbox solution, or building more complicated solutions where the box feature is integrated as OEM. Storage-agnostic, Pydio is a simple layer that can be deployed on top of any storage backend, thus providing scalability and high-availability out-of-the-box.

Charles du Jeu is the lead developer of the solution and will briefly present its feature and how it can fit for research and engineering purpose.

See <https://pyd.io/> for more information.

Introduction and Keynotes / 28

Principles of Synchronization

Corresponding Author(s): bcpierce@cis.upenn.edu

Technology and research / 17

Programmatic access to file syncing services

Patrick Haworth Owen¹

¹ *Imperial College Sci., Tech. & Med. (GB)*

Corresponding Author(s): patrick.haworth.owen@gmail.com

Ganga is a python API for submitting jobs to distributed computing systems. When a user submits processing jobs to a Grid or a Cloud resource, these

jobs will often depend on input as well as producing output. To make the job gain access to the input and how to get easy access to the output afterwards is often a challenge. Within the Ganga framework, we have implemented an abstraction layer for file access. In addition to the traditionally used grid storage solutions, it is now possible to retrieve and store files to Google Drive and ownCloud implementations. The latter, which also provides access to CERNbox, is provided via WebDav access that also opens up many other possibilities. Specific use case will be discussed as well as the wider issues related to authentication.

Site reports / 12

SURFdrive, a sync and share service for Dutch higher education and research

Ron Trompert^{None}

During the first three months of this year we have setup an Owncloud-based sync and share service for Dutch higher education and research. This service uses a SAML2-based federated login. In this presentation we will discuss the requirements, choices we have made, tests we have done, the technical setup and the experiences we have had so far not only with the software and our setup but also setting up this service.

Software solutions / 3

Seafile open source cloud storage, technology and design

Johnathan Xu¹

¹ *Seafile Ltd.*

Seafile (<http://seafile.com>) is an open source cloud storage system. Seafile is already used by over 100 thousand users worldwide.

Seafile provides file syncing, sharing and team collaboration features. It is designed with to make file collaboration for teams more efficient.

The core technology of Seafile is an application-level file system, built on top of object storage systems. The architecture can scale out horizontally by adding more servers. It supports S3, Swift or Ceph/RADOS as storage back end. The file syncing algorithm is designed to be efficient and reliable.

In this talk, we will present:

- Seafile's features, including file syncing, sharing, and team collaboration
- Seafile's architecture and technology, including file system design, cluster, and syncing algorithm

Site reports / 0

Site report: CERNBOX

Hugo Gonzalez Labrador¹

¹ *University of Vigo (ES)*

Corresponding Author(s): hugo.gonzalez.labrador@cern.ch

Cloud sharing and data synchronization over modern backend storages, like EOS ...

Technology and research / 18

Storage solutions for a production-level cloud infrastructure

Author(s): Giacinto Donvito¹

Co-author(s): Marica Antonacci ¹ ; Spinoso Vincenzo ¹

¹ *INFN-Bari*

We have set-up an OpenStack-based cloud infrastructure in the framework of a publicly funded project, PRISMA, aimed at the implementation of a fully integrated PaaS+IaaS platform to provide services in the field of smart-government (e-health, e-government, etc.). The IaaS testbed currently consists of 18 compute nodes providing in total almost 600 cores, 3550 GB of RAM, 400 TB of storage (disks). Connectivity is ensured through 2 NICs, 1Gbit/s and 10Gbit/s. Both the backend (MySQL database and RabbitMq message broker) and the core services (nova, keystone, glance, neutron, etc.) have been configured in high-availability using HA clustering techniques. The full capacity available by 2015 will provide 2000 cores and 8 TB of RAM.

In this work we present the storage solutions that we are currently using as backend for our production cloud services.

Storage is one of the key components of the cloud stack and can be used both to host the running VMs (“ephemeral” storage), and to host persistent data such as the block devices used by the VMs or users’ archived unstructured data, backups, virtual images, etc..

The storage-as-service is implemented in Openstack by the Block Storage project, Cinder, and the Object Storage project, Swift. Selecting the right software to manage the underlying backend storage for these services is very important and decisions can depend on many factors, not only merely technical, but also economic: in most cases they result from a trade-off between performance and costs.

Many operators use separate compute and storage hosts. We decided not to follow this mainstream trend aiming at the best cost-performance scenario: for us it makes sense to run compute and storage on the same machines since we want to be able to dedicate as many of our hosts as possible to running instances. Therefore, each compute node is configured with a significant amount of disk space and a distributed file system (GlusterFS and/or Ceph) ties the disks from each compute node into a single file-system.

In this case, the reliability and stability of the shared file-system is critical and defines the effort to maintain the compute hosts: tests have been performed to assess the stability of the shared file-systems changing the replica factor. For example, we observed that GlusterFS in replica 2 cannot be used in production because highly unstable even at moderate storage sizes.

Our experience can be useful for all those organizations that have specific constraints in the procurement of a compute cluster or need to deploy on pre-existing servers for which they have little or no control over their specifications. Moreover, the solution we propose is flexible enough, since it is always possible to add external storage when additional storage is required.

We currently use GlusterFS distributed file system for:

- storage of the running VMs enabling the live migration,
- storage of the virtual images (as primary Glance image store),
- implementation of one of the Cinder backends for block devices.

In particular, we have been using Cinder with LVM-iSCSI driver since Grizzly release when the GlusterFS driver for Cinder did not support advanced features like snapshots and clones, fundamental

for our use-cases. In order to exploit GlusterFS advantages even using LVM driver, we created the Cinder volume groups on GlusterFS loopback devices.

Upgrading our infrastructure to Havana, we decided to enable Ceph as additional backend of Cinder in order to compare features, reliability and performances of the two solutions.

Our interest for Ceph derives also from the possibility to consolidate the infrastructure overall back-end storage into a unified solution. To this aim, currently we are testing Ceph to run the Virtual Machines, both using RBD and Ceph-FS protocols, and to implement the object storage.

In order to test the scalability and performance of the deployed system using test cases which are derived from the typical pattern of storage utilization. The tools used for testing are standard software widely used for this purpose such as: iozone and/or dd for block storage and specific benchmarking tools like Cosbench, swift-bench and sbench for the object storage. Using different tools for testing the file-system and comparing their results with the observation of the real test case, is also a good possibility for testing the reliability of the benchmarking tools.

Throughput tests have been planned and conducted on the two system configurations in order to understand the performance of both storage solutions and its impacts to applications aiming at achieving the better SLA and end-users experience.

Implementing our cloud platform, we focused also on providing transparent access to data using standardized protocols (both de-iure and de-facto standards). In particular, Amazon-compliant S3 and the CDMI (Cloud Data Management Interface) interfaces have been installed on top of the Swift Object Storage in order to promote interoperability also at PaaS/SaaS levels.

Data is important for businesses of all sizes. Therefore, one of the most common user requirement is the possibility to backup data in order to minimize their loss, stay compliant and preserve data integrity. Implementing this feature is particularly challenging when the users come from the public administrations and the scientific communities that produce huge quantities of heterogeneous data and/or can have strict constraints.

An interesting feature of the Swift Object Storage is the geographic replica that can be used in order to add a disaster-recovery feature to the set of data and services exposed by our infrastructure.

Also Ceph provides a similar feature: the geo-replication through RADOS gateway.

Therefore, we have installed and configured both a Swift global cluster and a Ceph federated cluster, distributed on three different geographic sites. Results of the performance tests conducted on both clusters are presented along with a description of the parameters tuning that has been performed for optimization. The different replication methods implemented in the two middlewares, Swift and Ceph, are compared in terms of network traffic bandwidth, cpu and memory consumption.

Another important aspect we are taking care of is the QoS (Quality of Service) support, i.e. the capability of providing different levels of storage service optimized wrt the user application profile. This can be achieved defining different tiers of storage and setting parameters like how many I/Os the storage can handle, what limit it should have on latency, what availability levels it should offer and so on.

Our final goal is also to set-up a (semi-)automated system that is able of self-optimising. Therefore we are exploring the cache tiering feature of Ceph, that handles the migration of data between the cache tier and the backing storage tier automatically. Results of these testing activities will be shown too in this presentation.

Summary and discussion / 33

Summary

Corresponding Author(s): miguel.branco@epfl.ch

Introduction and Keynotes / 25

The CERNBox project for science

Corresponding Author(s): jakub.moscicki@cern.ch

Technology and research / 14

The File Sync Algorithm of the ownCloud Desktop Clients

Klaas Freitag¹

¹ *ownCloud GmbH*

Corresponding Author(s): freitag@owncloud.com

The ownCloud desktop clients provide file syncing between desktop machines and the ownCloud server, available for the important desktop platforms.

This presentation will give an overview of the sync algorithm used by the clients to provide a fast, reliable and robust syncing experience for the users.

It will describe the phases a sync run will go through and how it is triggered. It also will provide an insight on the algorithms that decided if a file is uploaded, downloaded or even deleted on either on the local machine or in the cloud. Some examples of non obvious situations in file syncing will be described and discussed.

As the ownCloud sync protocol is based on the open standard WebDAV the resulting challenges and the solutions will be illustrated.

Finally a couple of frequently proposed enhancements will be reviewed and assed for the future development of the ownCloud server and syncing clients.

Site reports / 23

The Sync&Share project in North Rhine-Westphalia

Holger Angenent¹

¹ *urn:Google*

Corresponding Author(s): h.angenent6@googlemail.com

Site reports / 7

The cloud storage service bwSync&Share at KIT

Author(s): Alexander Yasnogor¹ ; Nico Schlitter¹

Co-author(s): Andreas Petzold ²

¹ *KIT*

² *KIT - Karlsruhe Institute of Technology (DE)*

Corresponding Author(s): alexander.yasnogor@kit.edu

The Karlsruhe Institute of Technology introduced the bwSync&Share collaboration service in January 2014. The service is an on-premise alternative to existing public cloud storage solutions for

students and scientists in the German state of Baden-Württemberg, which allows the synchronization and sharing of documents between multiple devices and users. The service is based on the commercial software PowerFolder and is deployed on a virtual environment to support high reliability and scalability for potential 450,000 users. The integration of the state-wide federated identity management system (bwIDM) and a centralized helpdesk portal allows the service to be used by all academic institutions in the state of Baden-Württemberg. Since starting, approximately 15 organizations and 8,000 users joined the service. The talk gives an overview of related challenges, technical and organizational requirements, current architecture and future development plans.

Technology and research / 8

The dCache scientific storage cloud

Author(s): Paul Millar¹

Co-author(s): Albert Rossi²; Christian Bernardt¹; Dmitry Litvintsev³; Karsten Schwank; Patrick Fuhrmann⁴; Tigran Mkrtchyan Mkrtchyan⁵

¹ *Deutsches Elektronen-Synchrotron (DE)*

² *FERMILAB*

³ *FNAL*

⁴ *DESY*

⁵ *Deutsches Elektronen-Synchrotron DESY*

Corresponding Author(s): paul.millar@desy.de

For over a decade, the dCache team has provided software for handling big data for a diverse community of scientists. The team has also amassed a wealth of operational experience from using this software in production.

With this experience, the team have refined dCache with the goal of providing a “scientific cloud”: a storage solution that satisfies all requirements of a user community by exposing different facets of dCache with which users interact.

Recent development, as part of this “scientific cloud” vision, has introduced a new facet: a sync-and-share service, often referred to as “dropbox-like storage”. This work has been strongly focused on local requirements, but will be made available in future releases of dCache allowing others to adopt dCache solutions.

In this presentation we will outline the current status of the work: both the successes and limitations, and the direction and time-scale of future work.

Technology and research / 4

WebFTS: File Transfer Web Interface for FTS3

Author(s): Andrea Manzi¹

Co-author(s): Alejandro Alvarez Ayllon¹; Christina Skarpathiotaki²; Michail Salichos¹; Michal Kamil Simon¹

¹ *CERN*

² *National Technical Univ. of Athens (GR)*

Corresponding Author(s): andrea.manzi@cern.ch

WebFTS is a web-delivered file transfer and management solution which allows users to invoke reliable, managed data transfers on distributed infrastructures. The fully open source solution offers a simple graphical interface through which the power of the FTS3 service can be accessed without the installation of any special grid tools. Created following simplicity and efficiency criteria, WebFTS allows the user to access and interact with multiple grid and cloud storage. The “transfer engine” used is FTS3, the service responsible for distributing the majority of LHC data across WLCG infrastructure. This provides WebFTS with reliable, multi-protocol, adaptively optimised data transfers. The talk will focus on the recent development which allows transfers from/to Dropbox and CERNBox (CERN ownCloud deployment)

Introduction and Keynotes / 27

Welcome address

Corresponding Author(s): frederic.hemmer@cern.ch

Introduction and Keynotes / 30

Workshop logistics

Software solutions / 13

ownCloud

Christian Schmitz¹ ; Thomas Müller²

¹ *ownCloud Inc*

² *o*

Corresponding Author(s): thomas.mueller@owncloud.com, cschmitz@owncloud.com

OwnCloud technical overview

Site reports / 2

ownCloud project at CNRS

David Rouse¹

¹ *C*

Corresponding Author(s): david.rousse@dsi.cnrs.fr

CNRS will launch next November an ownCloud based service with the intend to serve CNRS research units. The first step is to deploy this service as a beta solution for 2 months and 2 000 end users, and then to generalize this offer to the whole CNRS users (potentialy 100 000 users).

Our platform is based on ownCloud 7 community edition, with VMWare for virtualization, a Galera/MariaDB cluster database and Scalify for the distributed storage backend.

We will try to present during this workshop our service implementation in detail, and discuss about our choices, our concerns, ... our troubles :)