A Staging Storage Sharing System for Data Handling in a Multisite Scientific Organization

F.Iannone*, I.Bellagamba, G.Bracco, B.Calosso, G.Giovanetti, S.Migliori, M.Mongelli, A.Perozziello, S.Pierattini, A.Quintiliani, F.Ambrosino, D.Di Mattia, A.Funel, G.Guarnieri, G.Ponti, F.Simoni, M.Steffé

ENEA Frascati RC - ICT Division, via E.Fermi,45 00044 Frascati, Rome, Italy
*corresponding author: francesco.iannone@enea.it

Abstract— In large scientific organizations, the laboratory experiments produce a huge amount of data and their processing and storage management are a challenging issue. Cloud architectures are exploited here and there for storage solutions and data sharing as well, in order to realize a collaborative world wide distributed platform. Large experimental facilities manage by themselves Information and Communications Technology (ICT) resources such as compute, networking and storage, while small experimental laboratories are demanding more and more departmental ICT resources for their own scientific instruments aided by data acquisition and control systems, specially in terms of storage and sharing/publishing data solutions. ENEA Staging Storage Sharing (E3S) system has been developed over the ENEA ICT infrastructure using Owncloud as architectural component for file syncing and sharing, trying to answer to these needs. E3S provides an homogeneous platform able to store and share heterogeneous data produced by many different laboratories geographically spread on several sites and working on collaborative projects. A first deployment of E3S works in a project for cultural heritage diagnostics involving several laboratories in different ENEA sites producing schema-free data. The paper presents the first deployment of E3S and a performance analysis of the architectural components. The performance analysis has been carried out with customized benchmark tools on a test bed consisting of a HPC cluster over Infiniband mounting a high performance storage.

Index Terms — Cloud Storage, Linux, AFS, GPFS

INTRODUCTION

The cloud storage services are now widespread in large scientific organisations. They are mainly used for file sharing and scientific datasets access.

The storage service at ENEA is based on the distributed filesystem AFS since 1998. The AFS cell *enea.it* [1] provides users with home-dir as well as storage areas for software and database projects. It can be accessed from geographically distributed ENEA premises. ENEA storage service includes also GPFS, the high performance filesystem for HPC multi-core clusters, InfiniBand based. GPFS is able to provide massive IO required by parallel jobs running on HPC clusters as well as to store big data projects. ENEAGRID is the computational infrastructure integrating all storage and computing resources distributed over WAN between several ENEA research centres .

A cloud solution named E3S (ENEA Staging Storage Sharing) has been developed to extend the storage services in ENEAGRID. E3S is based on OwnCloud framework and allows to sync the staging areas of small diagnostic laboratories data acquisition systems with the ENEA ICT storage area, AFS and GPFS, that are used as back-end. The paper describes the E3S architecture, the package tools that have been developed and finally reports the performance analysis results in order to estimate the workload of the cloud sync service.

STAGING STORAGE SHARING ARCHITECTURE

State-of-the-art diagnostic systems in the cultural heritage domain include computer aided instruments, producing a huge amount of raw data requiring post-processing analysis. The different diagnostic instruments generate a large variety of data, spanning from slow time-series to high frequencies sampling, to 2D and 3D scans and so on. That's why storing raw and post-processed data is a challenging issue. Often the experiments are carried out on the cultural heritage site and the data acquisition systems are remotely controlled. Hence the structure of the raw and post-processed data of all diagnostic instruments has a high complexity, and this is why trying to design a common data model for an uniform data access is not a convenient strategy. A solution was designed in order to hide the data structure complexity, allowing users to store data in local staging areas, synchronized with the distributed filesystems for world wide sharing. A general schema of the E3S architecture is depicted in Fig.1 and it was designed adopting the concepts of data integrity and security, scalability and reliability as fundamental requirements.

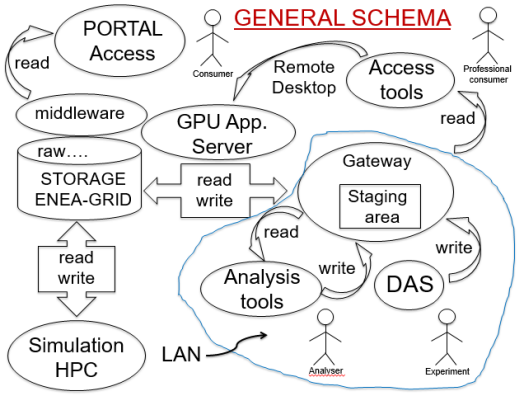

Fig.1: ENEA Staging Storage Sharing system architecture

The data integrity is guaranteed by file synchronization between staging and storage areas. The security is provided by the single-signed on authentication/authorisation system based on Kerberos 5 and Active Directory, that includes the Access

Control List for storage areas under AFS and posix-unix access rights for GPFS filesystem. The scalability is achieved separating the hardware resources of the staging components from the ones of storage and sharing. The storage area components are already scalable, since they are computing and storage resources of ENEAGRID infrastructure, and the staging, as well as the sharing components, are based on cloud solutions that are scalable by definition. The reliability requirement is satisfied as data are stored in the local staging area and managed by hardware resource components that are close to the data acquisition systems and that are independent from the storage services over Wide Area Network out of ENEAGRID.

The staging areas are handled by the component named Gateway Node (GWN). It provides a cloud service able to sync filesystems on the Data Acquisition Systems (DAS) with its own local disks, configured as staging areas. In such a way the data acquisition systems can store data quickly even in the case of a network link failure, and then upload files in asynchronous mode, by means of the HTTP methods PUT/GET on GWN local disks, used as storage back-end.

On one hand the GWN is a server node for the cloud storage service, on the other hand it is an AFS/GPFS client node, in order to sync the local staging areas with the distributed storage areas of the ENEAGRID infrastructure. Synchronization is achieved by means of a batch process running periodically on the node with the authorisations to access the storage areas in read/write mode. Once the data have been copied in the ENEAGRID storage areas, they can be shared with different access tools: *i)* a web portal on the component Middleware Node (MWN), open to any user: it provides cloud sharing services using storage ENEAGRID as a back-end; *ii)* the Application Servers for professional users running specific applications to analyse and visualise data using Graphic Processing Units (GPU); *iii)* HPC clusters running user parallel codes that are able to mount AFS/GPFS filesystem and access to the storage areas.

COMPUTING – STORAGE – PACKAGE TOOLS

Both the GWN and MWN are virtual computers configured in the ENEA Cloud infrastructure, and are based on the VMware framework. The operating system is Linux Centos 7 with OpenAFS client of the cell *enea.it* installed. The GPFS filesystem is imported by means NFS via GEthernet.

The OwnCloud server has a local disk space as back-end storage for staging area of the GWN while the MWN, used for sharing service, has AFS or GPFS filesystem as back-end storage area, accessible only in read mode.

The MWN runs an OwnCloud server with the filesystem AFS or GPFS as backend. Conversely from GWN, the file sharing app of the MWN of OwnCloud server is enabled, and the LDAP app is configured to use the ENEA ICT access service based on Active Directory. The ENEA users in charge of the experimental labs, can share filesystems AFS or GPFS synchronized with the staging area of the GWN defining their own data-policy for any file or folder.

E3S PERFORMANCE ANALYSIS

The aim of the following section is to carry out a performance analysis on the main components of the E3S architecture, in a typical massive data I/O flow scenario where thousands of data acquisition channels produce big data flow in large size experiment facilities. Hence an emulation test-bed has been set up, with many data providers interconnected by means of a low latency network, writing in a high performance storage system using the common OwnCloud tools based on files syncing techniques.

The test loads an OwnCloud server installed on the server node by means of a cluster of sixteen clients, capable of generating a high number of concurrent files uploads.

Three different use cases were tested: 10, 100 and 500 files of 1 MB were uploaded by each client node simultaneously using 1, 4, 8 and 16 nodes.

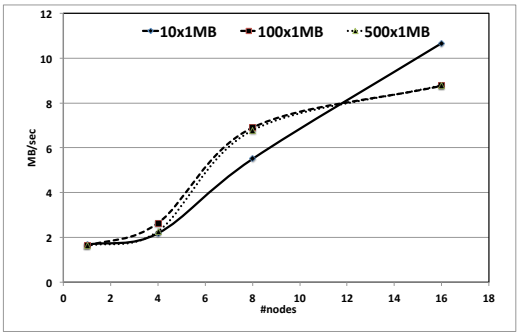

Fig.2: OwnCloud test: files upload throughput vs. #nodes

The OwnCloud server throughput rate is around tens of MB/sec (fig.2). It can be scaled up only in a multi-server OwnCloud architecture.

CONCLUSIONS

The E3S system of staging, storage and sharing has been developed in ENEA adopting files syncing techniques based on open-source cloud technologies. It tries to offer to scientific data producers all the reliability and performances of a distributed filesystem like AFS or GPFS through an easy-to-use interface, that leave unchanged the existing data acquisition procedures. This can be useful in a scenario where dispersed laboratories, producing different kind of data, use a large ICT facility for storage and sharing of their data.

ACKNOWLEDGEMENTS

This work was supported by the ENEA *COBRA* project on the advanced tools for the conservation of cultural heritage, based on the application of Radiation and Enabling Technologies, funded by the Latium Regional Government.

REFERENCE

[1] D.Abate, F.Ambrosino, et al., *ENEA-GRID Infrastructure.* In High performance computing on CRESCO Infrastructure: research activities and results 2010-2011,pg.9 ISBN: 978-88-8286-268-8.