Using IKAROS to form Scalable Storage Platforms

Filippidis Christos [a,b], Cotronis Yiannis [b] & Markou Christos [a]

[a] National Center for Scientific Research Demokritos, [b] National & Kapodistrian University of Athens

I. IKAROS

Is a *data centric, energy aware platform* seeking *synergies* between wider communities, permits ad-hoc nearby storage formations and is able to use a *huge number* of low spec, low power consumption I/O nodes in order to *increase* the available bandwidth and *decrease* the overall power consumptions.

Enable us to *virtually connect*, at the users level, the several different computing facilities (Grids, Clouds, HPCs, Data Centers, Local computing Clusters and personal storage devices), being utilized, *on-demand* based on the needs, by using well know standards and protocols like *HTTP*.

Permit us to create more *user-driven* computing facilities with application users and owners playing a decisive role in governance and focusing on placing computer science and the harvesting of '*big data*' at the center of scientific discovery.

II. Introduction

International collaborative scientific experiments are generating Datasets which are increasing exponentially in both complexity and volume, making their analysis, archival, and sharing one of the grand challenges of the 21st century.

These experiments, in their majority, adopts computing models consisting of Tiers (each Tier is made up of several computer centres and provides a specific set of services) and for the different steps of data processing (simulation, filtering, calibration, reconstruction and analysis) several software packages are utilized.

The computing requirements are extremely demanding and, usually, spans from serial to multi-parallel or GPU-optimized jobs. The collaborative nature of these experiments demands very frequent WAN data transfers and data sharing among individuals and groups. Typically, such a computing model utilizes several different computing infrastructures like: *Grids, Clouds, HPCs, Data Centers and Local computing Clusters*.

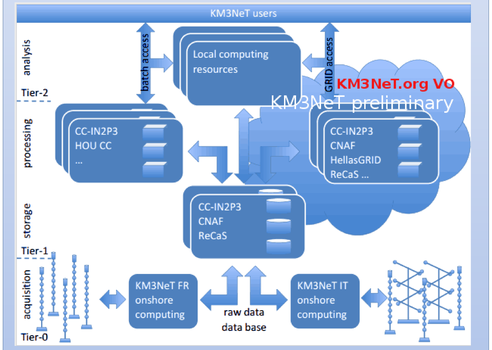
A researcher member of these experiment, in his/her every day working process will have to move Datasets from Data Centers to a Local computing Cluster by using secure copy tools (SCP). Then s/he will conduct a short local analysis and upload the output to a grid User Interface in order to start more coherent analysis, using a Tier 0 facility. Finally s/he will have to copy the output back to the local clusters and probably replicate data to a Data Center.

III. KM3NeT Experiment

The KM3NeT experiment (both an ESFRI project and a CERN recognized experiment) will introduce a distributed network of neutrino telescopes with a total volume of several cubic kilometers at the bottom of the Mediterranean Sea.

Additionally, the networked KM3NeT facility will become part of the EMSO/ESONET (also an ESFI project) seafloor observatory network for long-term real-time measurements in the extreme environment of the deep Mediterranean Sea in the Ligurian Sea near Toulon, the Ionian Sea near Sicily and Pylos.

Computing Model:



IV. Elastic Transfer (eT)

Publicly available module derived from IKAROS
Part of the KM3NeT computing model (Data Management)
Running on Mac, Linux, Windows

http://www.et-js.org

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING
MINISTRY OF EDUCATION & RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co-financed by Greece and the European Union
NSRF 2007-2013 EUROPEAN SOCIAL FUND

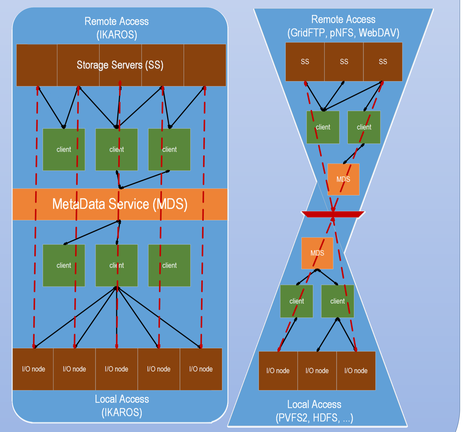
V. IKAROS Approach

IKAROS unifies remote and local access in the overall data flow, by permitting direct access to each I/O node. In this way we can handle the overall data flow at the network layer, which is much faster, while the other combined systems (local-remote access) demand a more active involvement of the operating system layer.

By further analyzing the overall data flow we can easily understand that there is a lack of coordination between the several different software layers. Parallel I/O systems commonly treat access to distributed data as a collective operation, and collective communication operations seek to optimize data transformations and transfers by coordinating related activities, In two-phase I/O and in Remote I/O (3-phase I/O), data is read and then reorganized via interprocess communication prior to transfer.

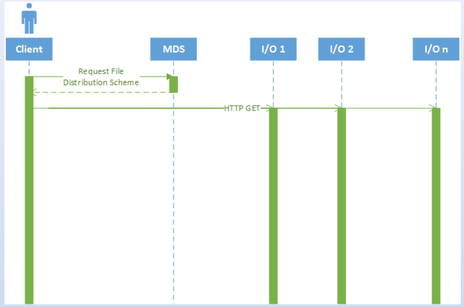
IKAROS approach reduces the required I/O phases, regardless of the Tier, in the computing model, we are acting at and the underlying storage implementation. At the same time we are preserving the separation of each level.

Overall Data Flow:

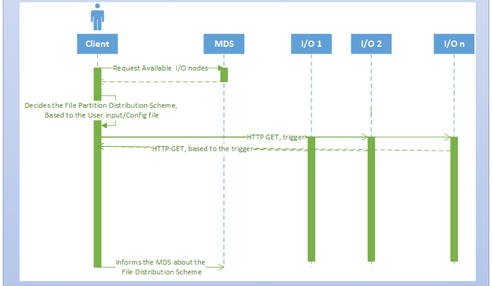


VI. IKAROS Use Case

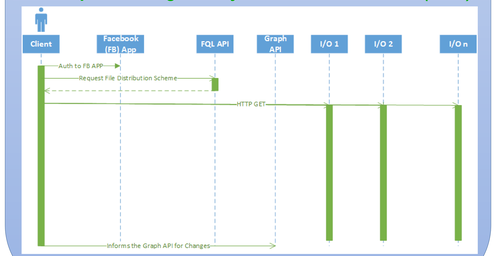
UML Sequence Diagram : Read Operations



UML Sequence Diagram : Write Operations



UML Sequence Diagram: Hybrid MetaData Service (MDS)

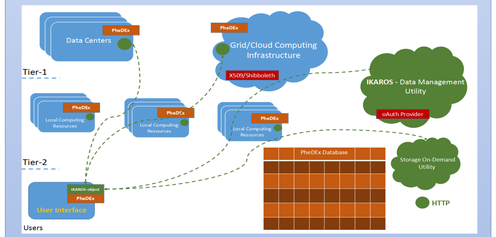


VII. Related Work

Existing Grid Computing tools and frameworks like PhEDEx (the official Data transfer & Management tool for CMS-LHC experiment at CERN) cannot provide the flexibility that the next generation collaborative projects demands. The main criticism against these tools and frameworks, being developed the past decade, is that at the Virtual Organization level they transform the "Grid" to a centralized infrastructure. This approach violates the decentralized concept of a Grid Infrastructure, which is the real advantage of the Grid computing model.

On the other hand, current Data as a Service (DaaS) solutions present a lack of flexibility in terms of allowing users to customize the underlying data models by including new concepts or functionalities (lack of data model *enrichment*). The platform providers usually offers a set of global APIs to enable users to access, download, or upload the data based on Create, Read, Update, Delete (CRUD) commands, or "sell" and transfer data to clients so they can do whatever they want with them. Thereby, collaboration becomes limited and sometimes is not even feasible.

VIII. Presence of IKAROS & PheDEx at all Levels



IX. Future Work

As a future work, we intent to develop IKAROS further in order to add value to current DaaS solutions by sharing not only the data (between the different computing facilities) but also the data models themselves (data model *enrichment*). We argue that in order to succeed that we must first deploy a storage infrastructure that follows the IKAROS approach.