AGH University of Science and Technology

Storage Management Systems for Organizationally Distributed Environments – PL-Grid Case Study

Renata Słota, Łukasz Dutka, Bartosz Kryza, Darin Nikolow, Dariusz Król, Michał Wrzeszcz, Jacek Kitowski

AGH University of Science and Technology,
Faculty of Computer Science, Electronics and Telecommunications,
Department of Computer Science

ACC Cyfronet AGH

PPAM 2013, Warsaw, September 8-11, 2013

Agenda

- Introduction – problem identification
- Storage in Grid Infrastructure
- Data Access Requirements of Domain Grids – Use Cases
  - Identification of problems and users' expectations
- Storage management systems
- Look Into The Future
- Conclusions

Introduction

- Status:
  - The 4th paradigm is believed to be the most important research paradigm in the coming years.
  - The „Big Data” revolution – data too difficult to process using database management tools or traditional data processing applications.

- The major problem is how to access, store, search and share the data.

Storage in Grid Infrastructure

- Variety of storage systems.

- The reasons for high heterogeneity of storage:
  - The users need storage which have different characteristics.
  - The local sites make autonomic decisions about the storage hardware and software, depending on the requirements of their key users.
  - The sites adopt some of the spare storage resources which already exist at the given location.

PL-Grid Infrastructure Current Case

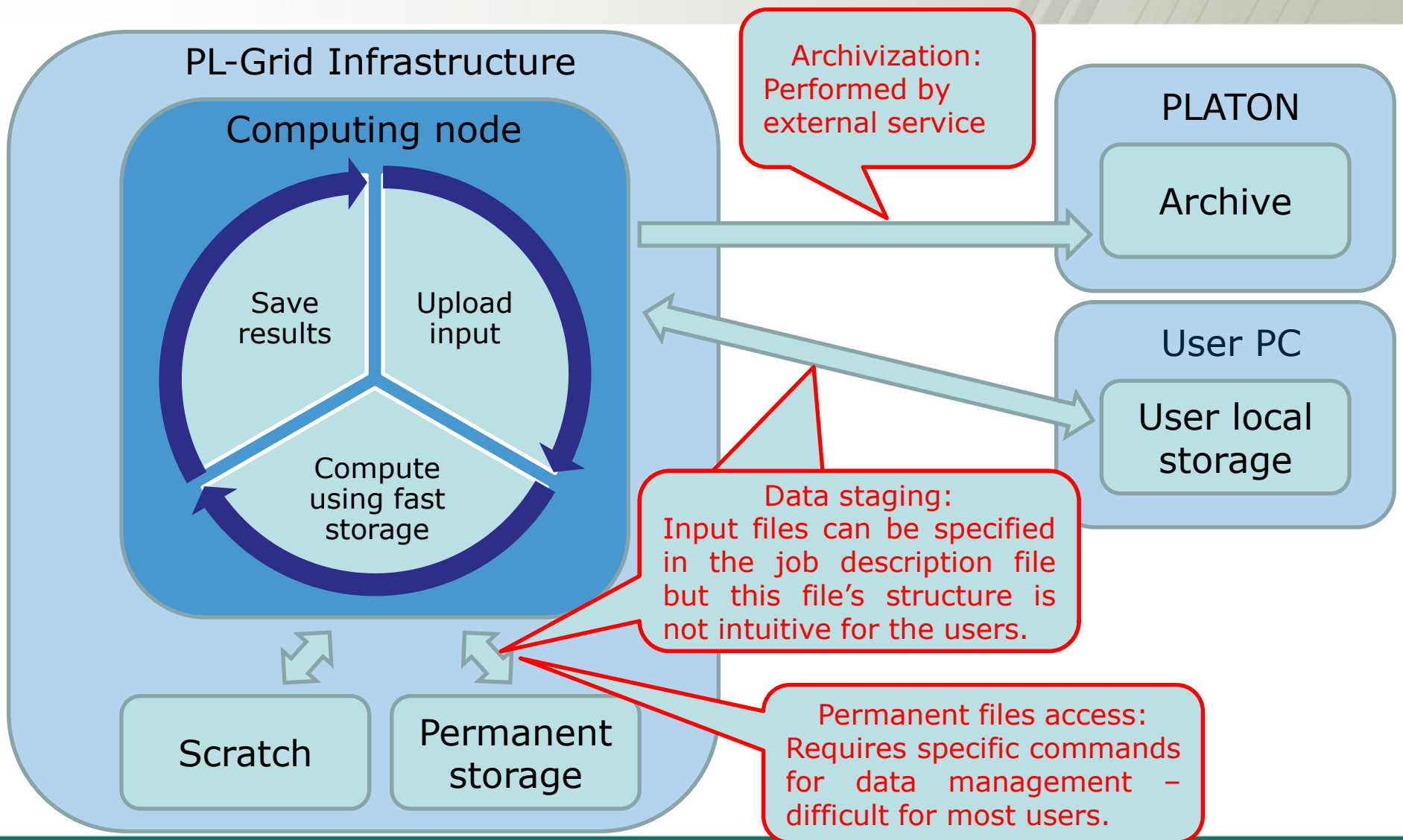
- Sample storage management systems:
  - Lustre
    - The typical use of the Lustre file system is a high performance scratch.
    - Different Lustre instances on the different sites - the data stored on this file system can only be shared within the local cluster of the given site.
  - LFC
    - Provides common file system functionality for distributed storage resources.
    - Supports user initiated file replication.
  - QStorMan
    - Continuous monitoring of the Lustre nodes and dynamically forwarding data access requests to the most appropriate storage resources.
    - Aimed at delivering storage QoS and resource usage optimization for applications which use the Lustre file system.

Data Access Requirements of Domain Grids – Use Cases (1/3)

- Domain Grids are services dedicated to selected areas of science that are using P-Grid Infrastructure.
- Users of different services have different requirements but use cases that are important and difficult for most of them can be found:

Use case	Importance	Difficulty	Comment
Data staging	High	Average	Users can specify necessary job input files in the job description file, provided these files are already registered in the permanent Grid storage (e.g. LFC).
Permanent files access	High	Average	Files can be accessed through DPM [21], LFC [5], Unicore [6]. Block access is possible through RFIO protocol. However DPM still requires specific commands for data management, which makes many users fall back to manual transfers.
Archivization	High	High	Archivization is only available outside of the Pl-Grid infrastructure through PLATON infrastructure [2].

Data Access Requirements of Domain Grids – Use Cases (2/3)



Data Access Requirements of Domain Grids – Use Cases (3/3)

- Other use cases which are believed to have growing importance in future have been identified:

Use case	Importance	Difficulty	Comment
Temporary files access	High	Low	Temporary files are stored locally at sites using NFS or Lustre, however different sites have different data deletion policy and different paths.
Data transfer from/to Grid	High	Easy	Users can use manual SSH/SCP commands.
Data transfer between sites	Average	Average	Users can manually transfer files between sites, either using Grid middleware (e.g. LFC) or manually over SSH based protocols.
Relational database access	Average	Easy	Users can create a custom database on a central MySQL server, this however poses certain performance and scalability issues.
Metadata	Low	Average	Users can only search using the files logical names, however no mechanism for structured metadata descriptions is available.

The problem summary

The user is typically interested in results.

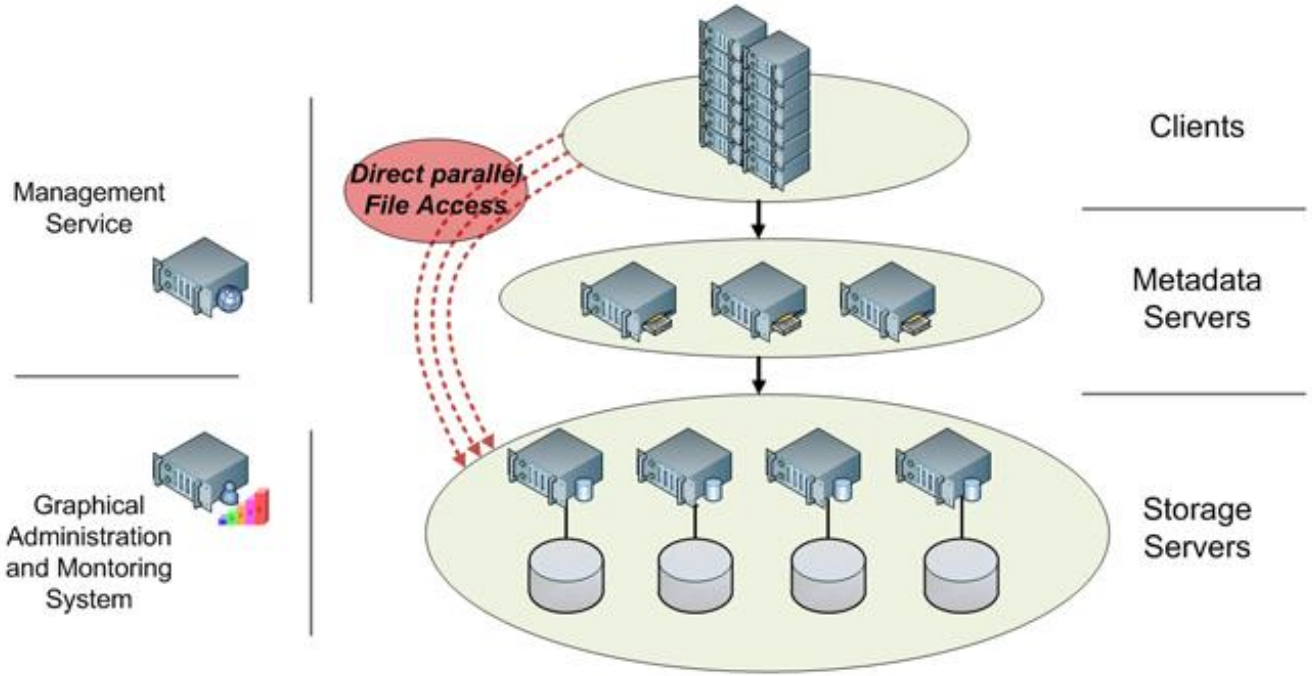
He expects that data access will be simple.

The access to the data in large scale computational environment should be the same as on personal computer.

How to do that?

Storage management systems (1/4)

- 1st possible solution:
  - Use of high-performance parallel file system that supports geographically distributed organizations.
  - Many file systems present on the market e.g. FhGFS, PanFS, DDN's WOS, Scality RING.



FhGFS Architecture

Figure from: http://www.fhgfs.com/wiki/wikka.php?wakka=SystemArchitecture

Storage management systems (2/4)

- Advantages of 1st solution:
  - Common object namespace across distributed locations
  - Minimal metadata access latency.
  - Massively scalable.
  - Secure.
  - Automatic data verification and correction
  - Automatic replication and recovery

- Disadvantages of 1st solution - creation of one simple system is often impossible in infrastructures such as Grid:
  - Weak support for different storage systems (many parallel file systems need dedicated hardware)
    - Existing storage systems in sites should be covered by parallel file system
    - Different expectations of users (large capacity data storage system for archiving and fast system for temporary files) – the system that meets them all would be very expensive
  - Centrally managed storage system makes impossible autonomy of sites in organization

Storage management systems (3/4)

- 2nd possible solution:
  - Use of data management system that supports federations and operates over storage systems at each site.
  - This type of systems already exists e.g. Parrot, iRODS.

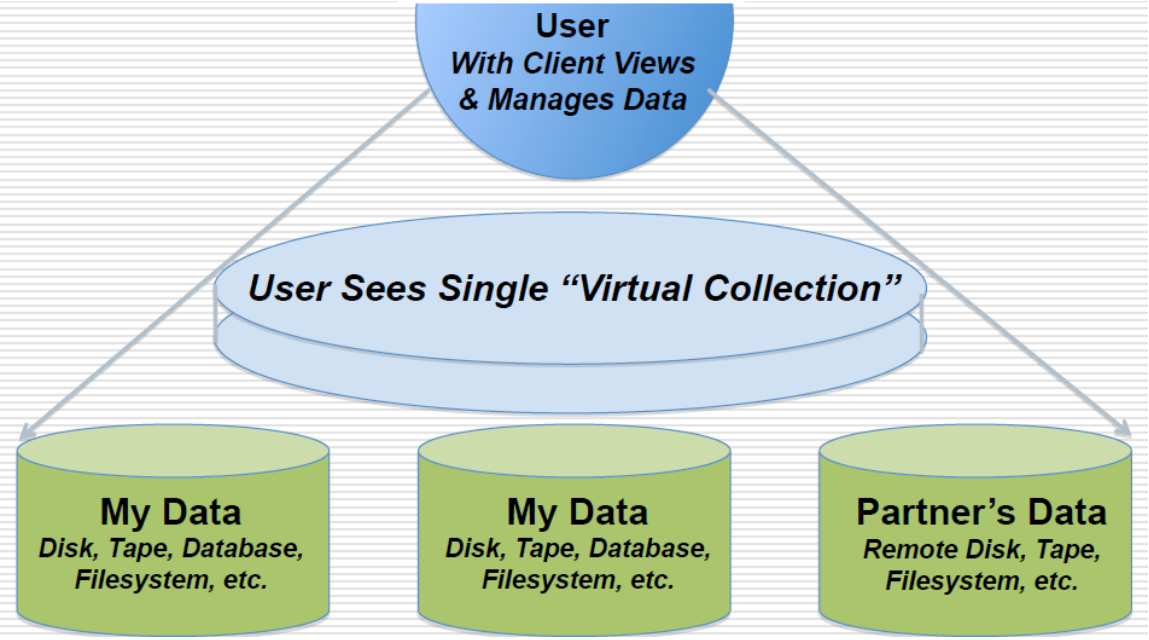


Figure from:
Overview of Data Management and IRODS, Humanities Workshop, Paul Tooby, San Diego, CA, March 19-20, 2009

Storage management systems (4/4)

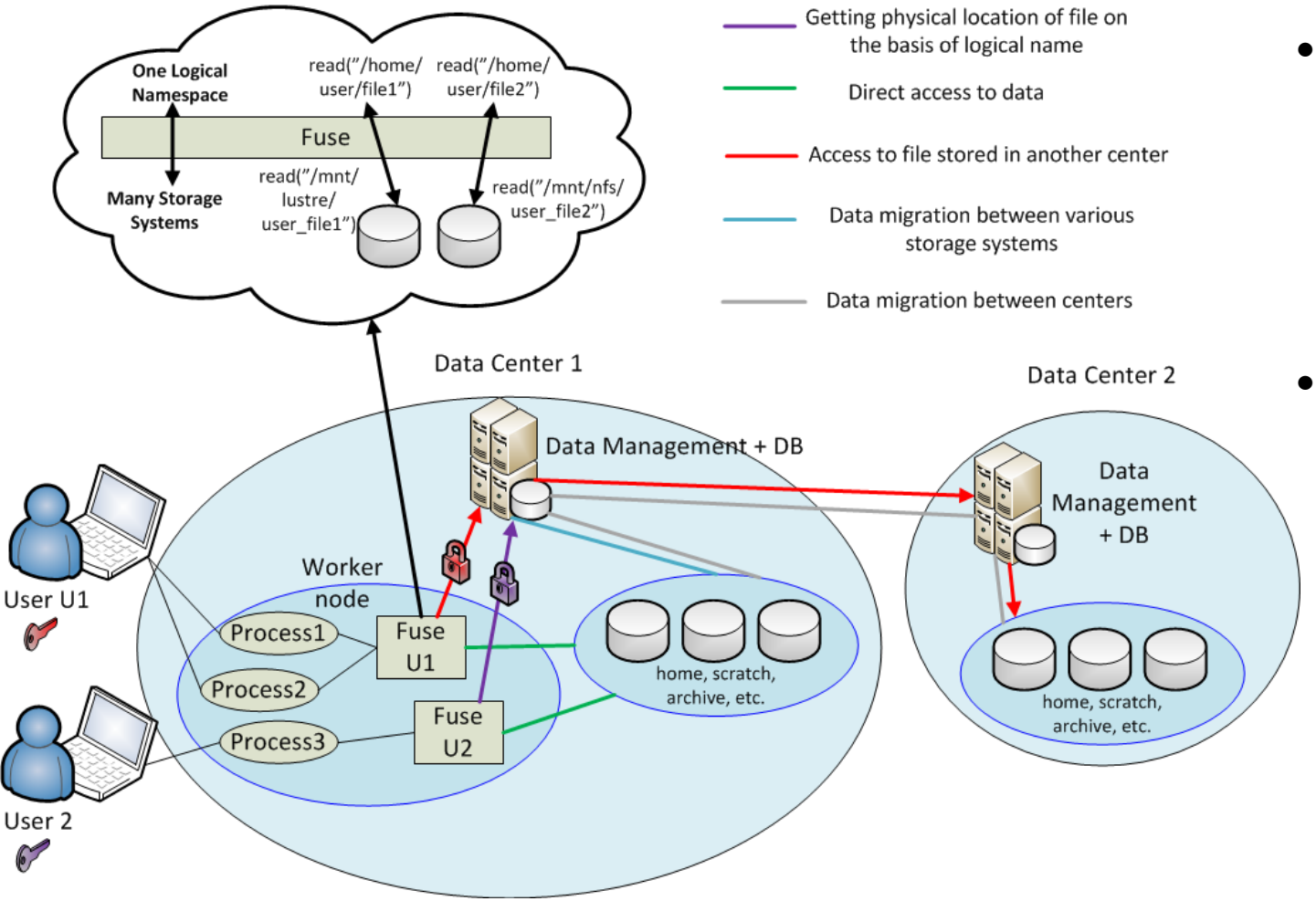
- Advantages of 2nd solution:
  - Many existing storage systems can be hidden.
  - Flexible rules
  - Metadata Catalog
  - Autonomy of sites

- Disadvantages of 2nd solution - no transparency in data center selection in multiple data centres scenarios:
  - User has to know in which site his file is.
  - The optimization of files' location by migration of data between sites is impossible (the user would have problem with finding his data).

Outcome

- Analysis of users requirements and existing solutions showed that, beyond the obvious expectations that most of the systems fulfil, data management system is expected, to offer:
  - access to the data in the same way as personal computers (simplicity),
  - cooperation with different existing storage systems (different systems for different expectations),
  - hide real location of files from users – the system should choose best location for data to meet the requirements of user and administrators.

Look Into The Future



- User point of view:
  - ✓ The user does not see any borders.
  - ✓ The user is able to process his data from anywhere.
- Administrator point of view:
  - ✓ Possibility of data migration provides high utilization of all storage resources.
  - ✓ The system coordinates data access and buffering to minimize network traffic.

Conclusions

- Data processing and managing in large scale environments will be the major problem which the research community will face in the near future.
- The user should be able to process his data anytime/anywhere.
- Analysis of users' requirements showed that everything should be simple:
  - Storage system should cover the data management that is performed by the system and administrators.