

PRACE iRODS Workshop

iRODS experience in EUDAT

Giuseppe Fiameni (g.fiameni@cinca.it),

Giacomo Mariani (g.mariani@cinca.it),

SuperComputing, Application and Innovation – CINECA

Linköping September 2012, Sweden



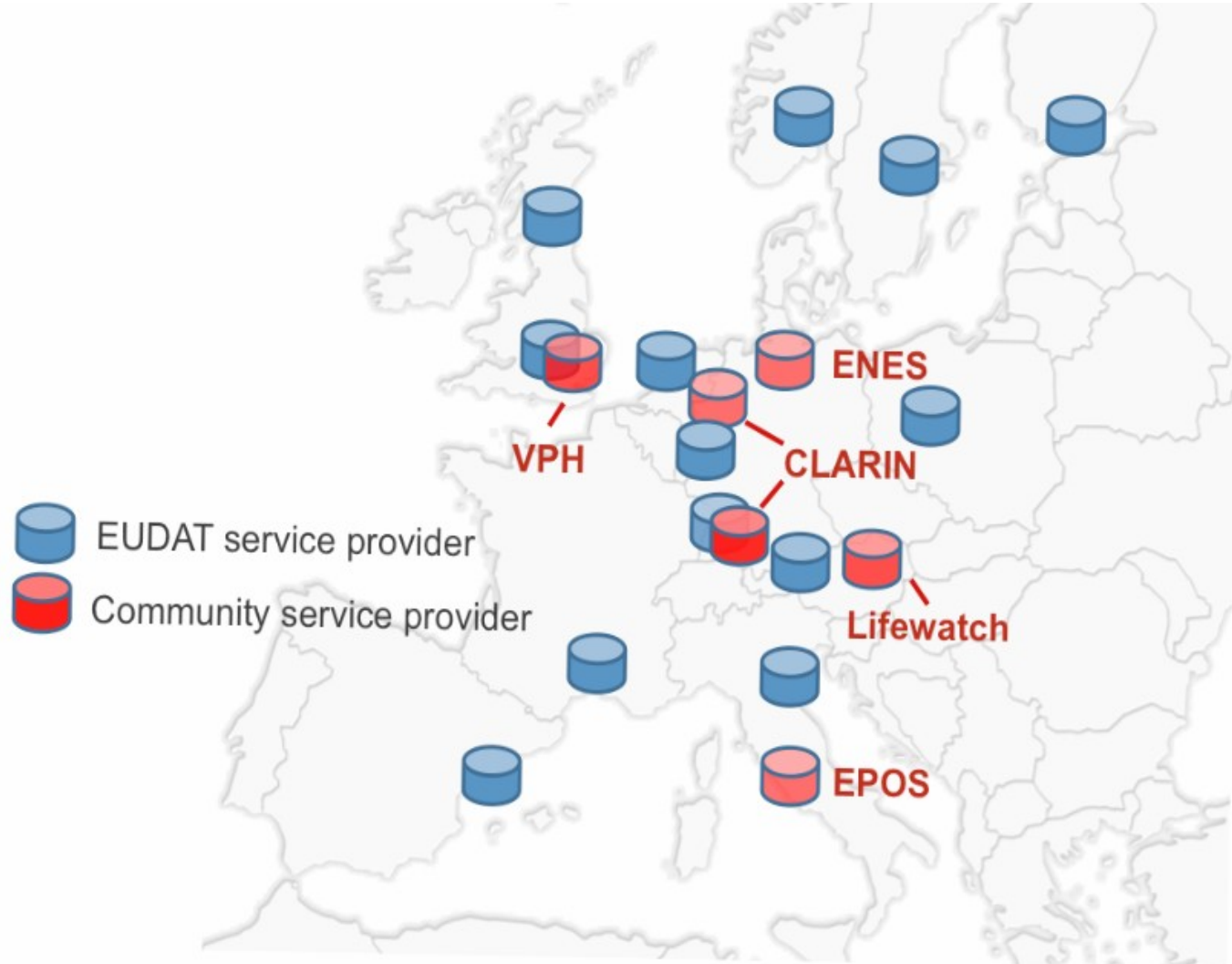
Agenda

- EUDAT
 - EUDAT members
 - EUDAT core services
 - EUDAT approach
- EUDAT and iRODS
 - Safe replication
 - Data Staging
- EUDAT and PRACE

EUDAT: European Data Infrastructure

| | |
|--------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Project Name | EUDAT – European Data |
| Start date | 1st October 2011 |
| Duration | 36 months |
| Budget | 16,3 M€ (including 9,3 M€ from the EC) |
| EC call | Call 9 (INFRA-2011-1.2.2): Data infrastructure for e-Science (11.2010) |
| Participants | 25 partners from 13 countries (national data enters, technology providers, research communities and funding agencies) |
| Objectives | “To deliver cost-efficient and high quality Collaborative Data Infrastructure (CDI) with the capacity and capability for meeting researchers’ needs in a flexible and sustainable way, across geographical and disciplinary boundaries.” |

EUDAT members



Communities

CLARIN
Common Language Resources and Technology Infrastructure

About CLARIN
Services
Publications
Activities

Events
Links
Glossary
Contact
Join CLARIN

Internal Web Site

Semantic data description and descriptive metadata are vital factors for determining if the data can be reused in the future. These metadata are still dependent on rapidly changing ontologies and terminologies.

John Marks
ESF 2008

Activities
Publications

Virtual Language Observatory
Consultancy
Laboratory
Solutions

January 20, 2011
New Virtual Language Observatory launch

LIFEWATCH
e-science and technology infrastructure for biodiversity data and observatories

Home Contact About Participants Get Involved News Cases Events Press Documents

LIFEWATCH COUNTRIES
Austria Belgium Denmark Finland France Greece Hungary Italy Netherlands Norway Poland Portugal Romania Slovak Republic Slovenia Spain Sweden Turkey United Kingdom

LIFEWATCH NEWS
2011-02-16 LIFEWATCH RESEARCH INFRASTRUCTURE STARTS CONSTRUCTION IN 2011 - "The initial country consortium establishing the Lifewatch research infrastructure agreed to finance..."
2011-01-19 LIFEWATCH CLOSING EVENT - "On this page you can download all the slides presented at the closing event of the Lifewatch preparatory project..."
2011-01-17 LIFEWATCH CONSTRUCTION KICKS OFF ON JANUARY 19TH - "On 19 January 2011, at the closing conference of the Lifewatch preparatory project a first group of..."

LIFEWATCH FOCUS
Lifewatch research infrastructure starts construction in 2011. The initial country consortium establishing the Lifewatch research infrastructure agreed to finance the start-up activities for the infrastructure construction. These countries will host the Common Facilities of Lifewatch. On 29th January 2011 representatives from organisations in Hungary, Italy, the Netherlands, Romania and Spain signed a Memorandum of Understanding to cooperate for an early start of the Lifewatch infrastructure for biodiversity and ecosystem research. The Lifewatch Stakeholders Board, representing the ten countries aiming at establishing the Lifewatch ERIC, welcomed the initiative to start early construction.

Newsletter
Subscribe to our newsletter. Send an email to newsletter@lifewatch.org to receive Lifewatch's newsletter.

Quote
"Through our Memorandum of Cooperation GBIF and Lifewatch, based on our respective complementary mandates, now have a formal framework for co-operation and collaboration on infrastructural developments, building on GBIF's 10 years of investment to date."
Dr. Mick King
Director Global Biodiversity Information Facility (GBIF)

ENES
European Network for Earth System Modelling

Welcome

News
IS-ENES
The Rationale
The Aims
The MOU
Related Projects
Partners
Meetings
Sitemap
Contact
Imprint

search...

print page

Welcome

ENES Townhall Meeting at EGU 2010: Here is the [announcement!](#)

For latest news on IS-ENES click [here!](#)

A major challenge for the climate research community is the development of comprehensive Earth system models capable of simulating natural climate variability and human-induced climate changes. Such models need to account for detailed processes occurring in the atmosphere, the ocean and on the continents including physical, chemical and biological processes on a variety of spatial and temporal scales. They have also to capture complex nonlinear interactions between the different components of the Earth system and assess, how these interactions can be perturbed as a result of human activities.

Accurate scientific information is required by government and industry to make appropriate decisions regarding our global environment, with direct consequences on the economy and lifestyles. It is therefore the responsibility of the scientific community to accelerate progress towards a better understanding of the processes governing the Earth system and towards the development of an improved predictive capability. An important task is to develop an advanced software and hardware environment in Europe, under which the most advanced high resolution climate models can be developed, improved, and integrated.

EPOS
EUROPEAN PLATE OBSERVING SYSTEM

Research Infrastructure and E-Science for Data and Observations on Earthquakes, Volcanoes, Surface Dynamics and Tectonics

SEARCH

Project
Home
News
Objectives
VPH-2010
Queries
VPH Links
Contact Us
Activities

Activities
Example Projects
News
Publications
Classification Material
VPH-2010
Event calendar

Mission & Vision Objectives Architecture Partners Preparatory Phase Data Products

VPH NoE
Virtual Physiological Human Network of Excellence

Home WP1 WP2 WP3 WP4 WP5 VPH MP Login

Search...

Welcome to the home page of the Virtual Physiological Human Network of Excellence (VPH NoE) and information portal for the VPH initiative.

Building a new VPH Community

HIGHLIGHTS
Interface Focus special issue with best papers from the VPH-2010 Conference
VPH NoE and the Princes Alliance
Example Project Call 3!
Join the Public Forum of the VPH-FET Support Action
Multi-Institutional Graduate Programme for Virtual Physiological Human Sciences (VPH-MIP)
VPH (Users & Strategy) Paper II
VPH NoE VPH Newsletter (Jan 2011) now available

LATEST VPH EVENTS
01.08.2011 - 02.08.2011 ICCS 2011 (Tbilisi)
08.08.2011 - 09.08.2011 IRODS
08.08.2011 - 09.08.2011 IRODS

VPH NoE
BioMed Town
VPH for the Public
VPH for Researchers
VPH for Clinicians

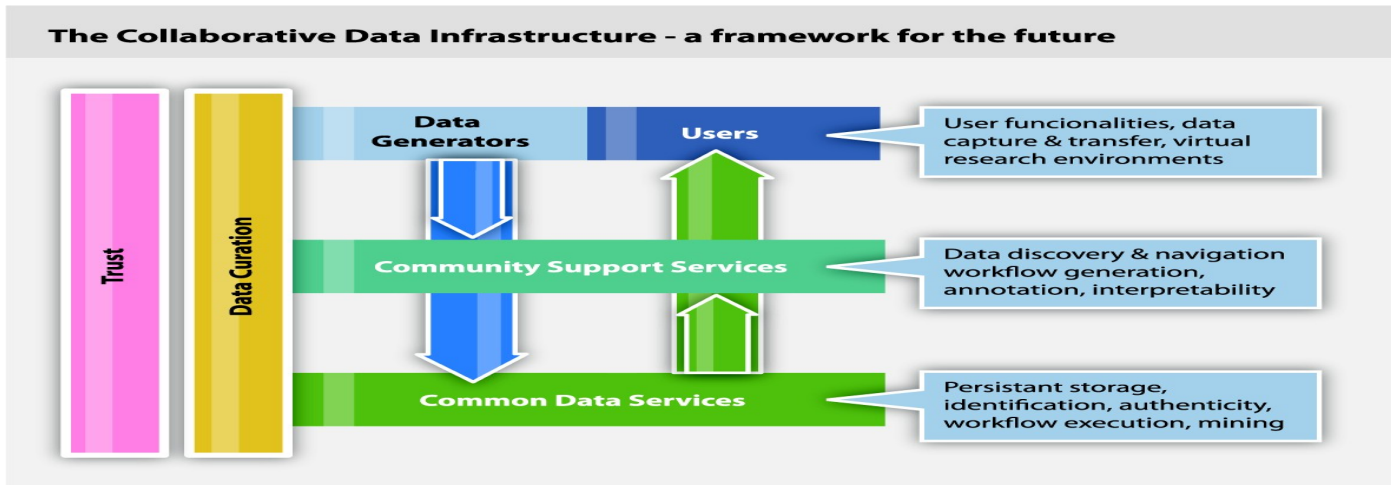
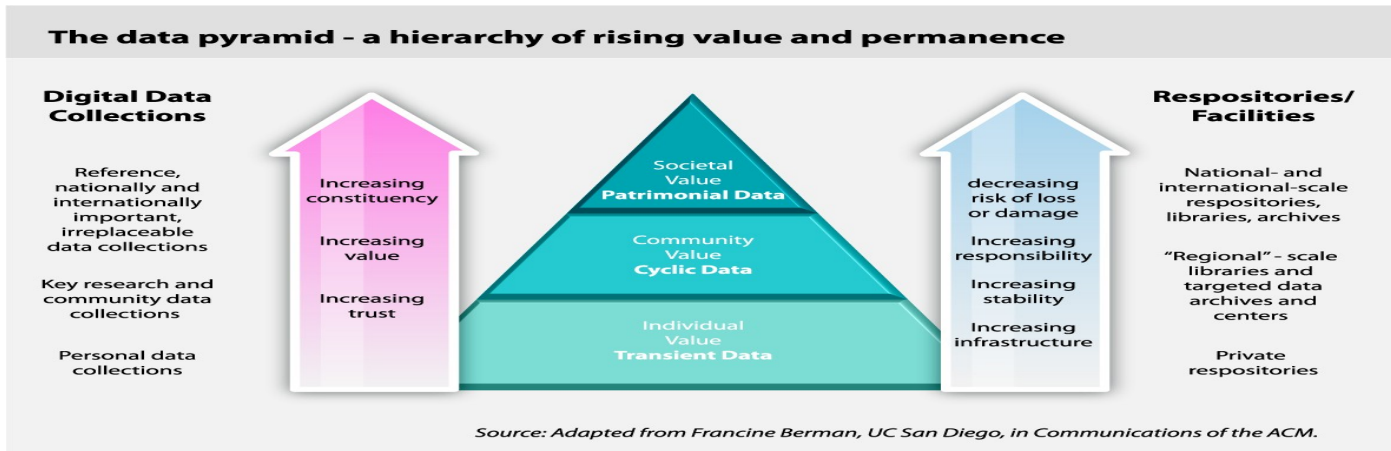
VPH Toolkit
VPH Training
Integrative Research
Dissemination & Networking

The VPH NoE is a project which aims to help support and progress European research in biomedical modelling and simulation of the human body. This will improve our ability to predict, diagnose and treat disease, and have a dramatic impact on the future of healthcare, the pharmaceutical and medical device industries.

VPH 2010 ABH 3010
September 30th - October 1st 2010
Brussels, Belgium



Common Data Infrastructure



EUDAT Core Service Areas

Community-oriented services

- Simple Data Access and upload
- Long term preservation
- Shared workspaces
- Execution and workflow (data mining, etc.)
- Joint metadata and data visibility

Enabling services (making use of existing services where possible)

- Persistent identifier service (EPIC, DataCite)
- Federated AAI service
- Network Services
- Monitoring and accounting

Core services are building blocks of EUDAT's Common Data Infrastructure

They are mainly included on bottom layer of data services





How to achieve this?

EUDAT service design activities

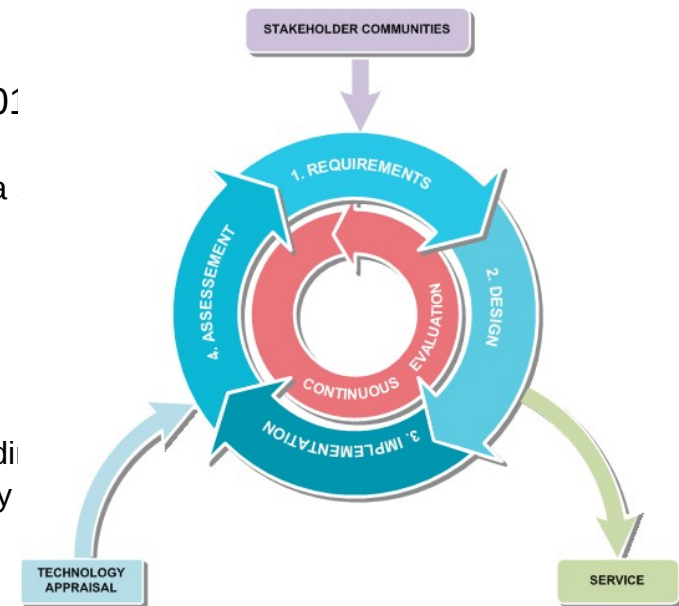
1. Capturing Communities Requirements (WP4)

- 1st round of interviews with the five initial communities (Oct.-Dec. 2010)
 - Understand how data is organised in each community
 - Collect first wishes and specific requirements from a common data layer

• Next phase: refine analysis and expanding it to other communities

2. Building the corresponding services (WP5)

- Technology appraisal (ongoing)
 - What is already available at partners's sites to build the corresponding services
 - What are the gaps and market failures that should be addressed by
- Next phase: Developing candidate services
 - Adapt services to match the requirements
 - Integrate with community and SP services
 - Test and evaluate with communities



3. Deploying the services and operating the federated infrastructure (WP6)

- Designing the federated infrastructure and the interfaces for cross-site operations (ongoing)
- Next phase: integrating and coordinating resource provision, operations and support

Community Service Wishes

- **In Progress (High priority)**
- Safe Replication of data (for Bit-stream Preservation & Access Optimization)
- Common Authentication/Authorization Infrastructure
- Staging of data onto HPC resources

- **In Progress (Medium priority)**
- Aggregated EUDAT Metadata Domain
- Researcher Data Store (Simple Upload, Share and Access)

- **In Progress as Research Issues (WP7)**
- More elaborate policy rules and federation scalability
- Generic workflow execution framework

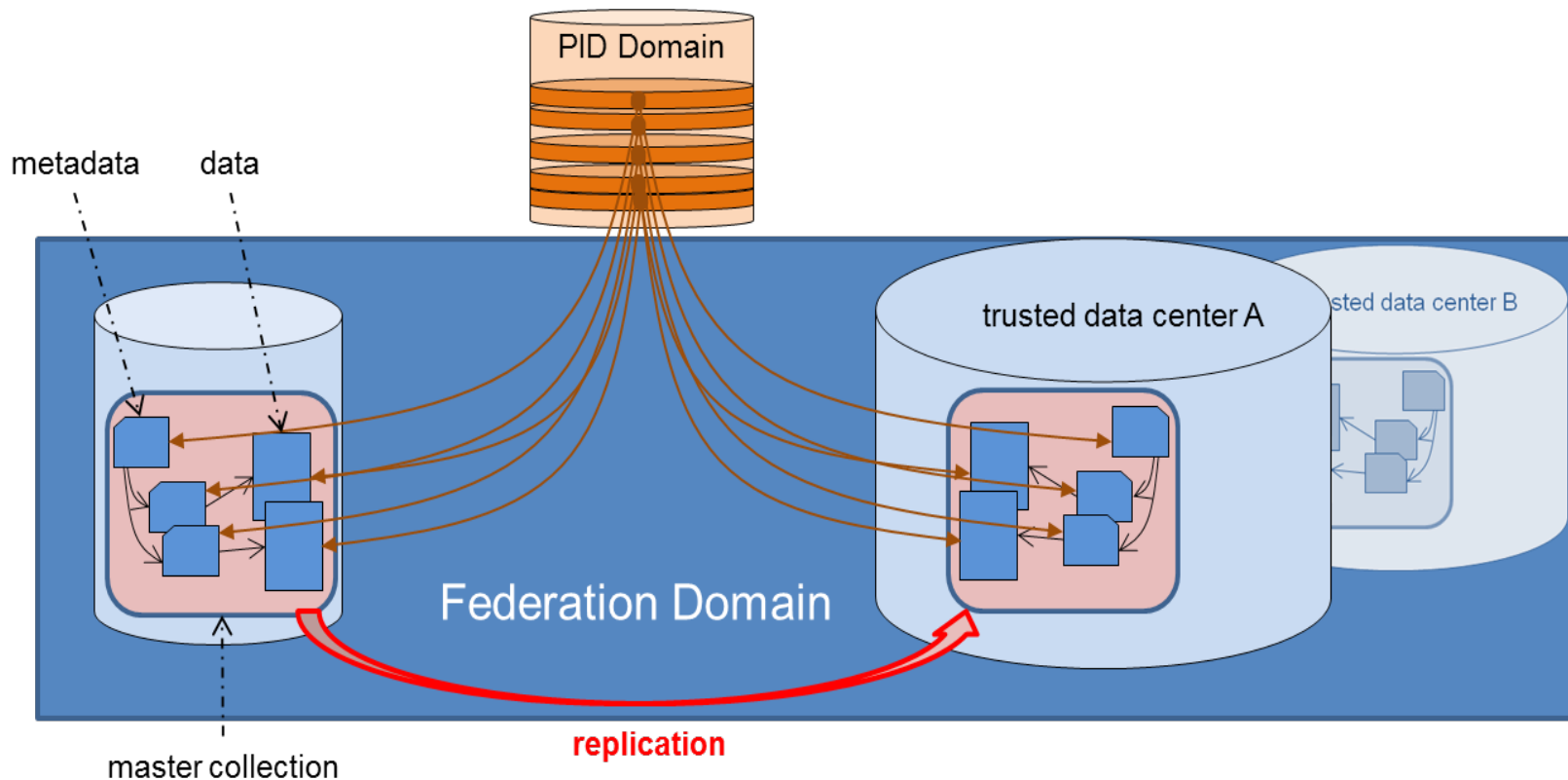
Safe Replication 1/3

- Relevant data needs to be replicated from community centers to a number of data centers in a safe way with several purposes in mind:
 - data bit-stream preservation;
 - more optimal data curation;
 - better accessibility of data;
 - identification of data through Persistent Identifiers (PIDs).
- Common functionality:
 - Create M replicas (identified by a PID record) at different data centers for N years, exclude certain centers, maintaining the given access permissions.

Safe Replication 2/3

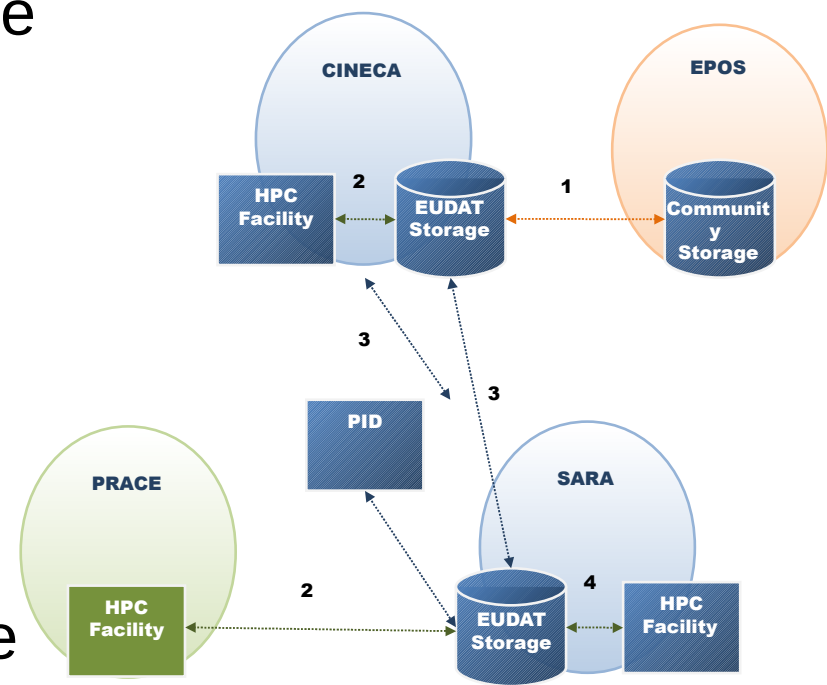
- **Technologies:**
 - Long Term Archives → Community specific technologies, through **iRODS UMSS**
 - Policy-based Replication → **iRODS** with rules and microservices
 - Persistent Identifiers → EPIC/Handle
- **Orthogonal aspects:**
 - AAI
 - Monitoring
 - Center Registry
 - Metadata

Safe Replication 3/3



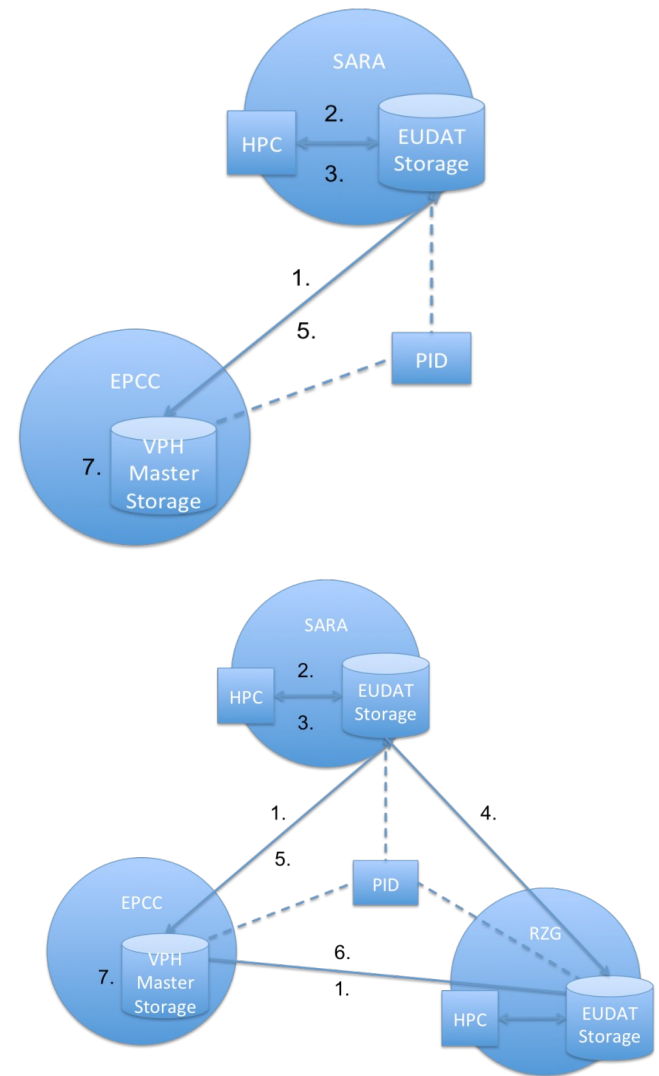
Data Staging 1/3

- **Objective:**
Allow communities to stage data between EUDAT resources and HPC/HTC resources, including the PRACE Infrastructure, for computational purposes.
- **Description:**
This service will allow the communities to dynamically replicate a subset of their data stored in EUDAT to an HPC machine workspace in order to be processed.



Data Staging 2/3

- Intention is to make use of HPC machines for computations on stored data.
- Different configurations possible:
 - computations on a single HPC node where data already is;
 - computations on multiple nodes - use of PRACE fast distributed file system.
- Principles:
 - user issues a compute command;
 - script pushes data into the HPC workspace, results go into workspace;
 - data, identified by the PID, are retrieved from the “best” resource, transparently;

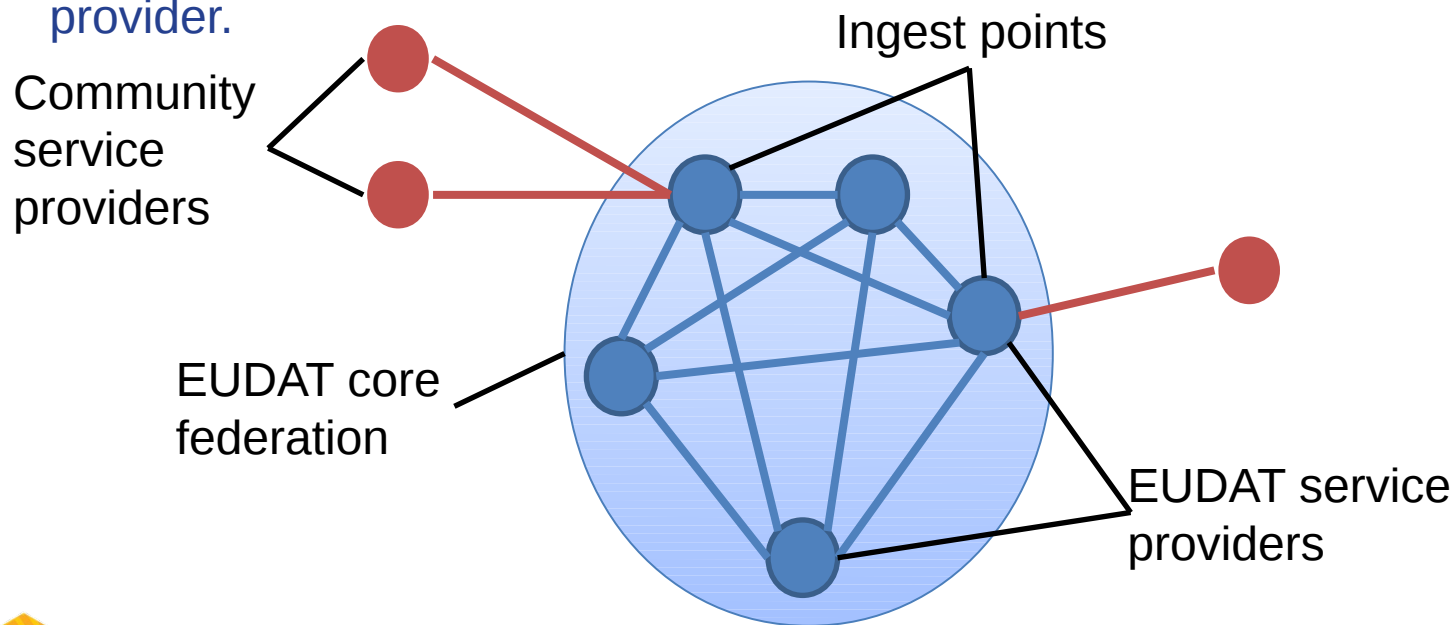


Data Staging 3/3

- **Technologies:**
 - Data transfers -> GridFTP interface to **iRODS** (**Griffin**)
 - Data access and selection -> various interfaces/clients (XSEDE File manager, Globus Online)
- **Orthogonal aspects:**
 - AAI
 - Monitoring

EUDAT and iRODS, the core technology

- iRODS is installed in all EUDAT service providers and community service providers.
- Data are stored in iRODS Vault or available as mounted collection.
- The iRODS instances are federated following the “snow flake” approach: each community service provider is federated with an ingest point, i.e. an EUDAT service provider federated with all the others EUDAT service provider.



iRODS used functionalities

- iRODS, version 3.0 or 3.1, was tested during the first technology appraisal activity.
- The services will be moved in production with iRODS 3.1 and probably be moved to iRODS 3.2 as soon as it will be available.
- The metadata structure definition is on going, so metadata are not used at the moment.
- embedPython has been used with good results, but, due to a recently discovered security issue, its usage has been suspended.
- Griffin, a JAVA GridFTP server able to exploit iRODS as a resource, is in use.
- Davis, a JAVA WebDAV interface has been tested, but it shows some issues (for example some crashes moving big files).
- iDROP web will be probably considered for the web interface.

TF goals and used iRODS functionalities

| | iRODS ACL | iRODS rules and microservices | Embed python | griffin | UMSS |
|------------------------------|-----------|-------------------------------|--------------|---------|------|
| Staging to HPC | Yes | | | Yes | |
| Replicate in different zones | Yes | Yes | Optional | | |
| Share data | Yes | Yes | | | |
| LTA | | | | | Yes |
| PID | | Yes | Optional | | |

CINECA experience

- iRODS 3.0 installed with support for GSI, FUSE and OS authentication
- External MySQL farm
- Data stored in a Vault exploiting GPFS over NFS
- Compound resource exploiting TSM via UMSS
- Ingestion of more than 5 TB to TSM
- Daily sync of about 5TB with a federated zone
- Installed and tested Griffin (solved bugs with GO)
- Installed and tested Davis
- Developed a script for Data Staging exploiting the GO

API

EUDAT experience

- Tens of iRODS installations and federations
- Evaluation of irods configurations
- Vault and imcoll deeply investigated
- Various federations scenarios evaluated
- Various LTA technologies connected
- Performance evaluations moving data and synchronizing data
- Experiences transferring TB of data (big files as well as losf)
- Rules and microservices for replication and PIDding developed
- Griffin evaluation

Technical details at CINECA

- Debian 6.0 on a 8 processors (x86_64)
- iRODS 3.0
- Python 2.6 (2.7 from pythonbrew)
- MySQL 5.1
- Perl 5.10.1
- Griffin 0.8.5
- Davis 0.9.4

EUDAT wishes

- LDAP integration
- Packaging for easier updates
- Optional use of system packages (such as system DB or GSI libraries)
- Migration tool able to import an archive already existing on a shared FS

Other, small wishes

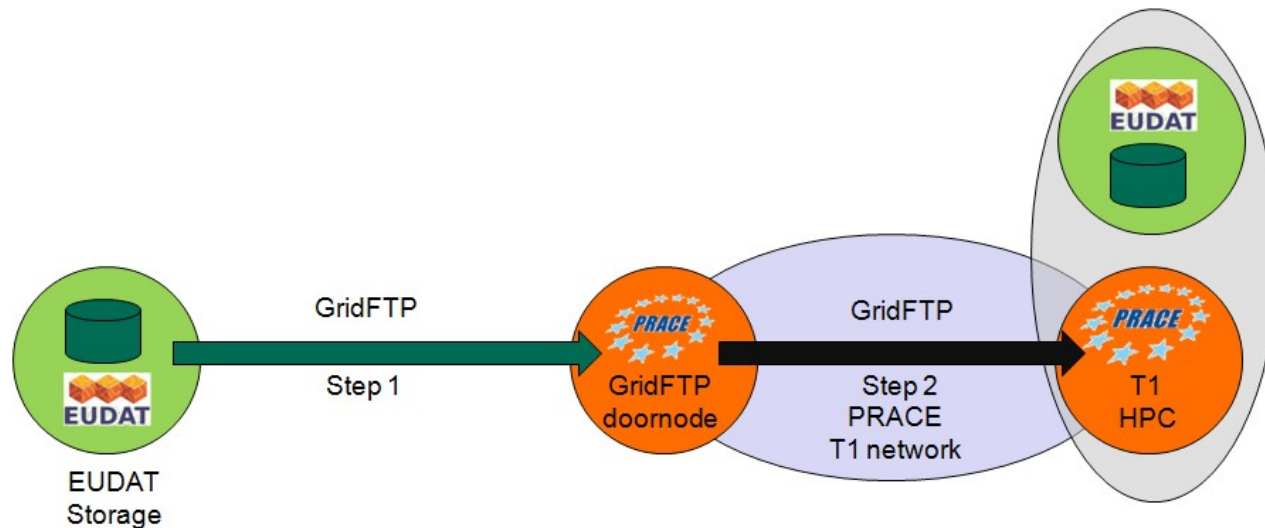
- Fixing the awkward behavior of remote users in rules execution: for example, a user defined in zona A cannot invoke rules in zone B where a irsync from zone B to zone A is involved.
- Fixing third party transfer: a user defined in A and defined as remote in zone B and zone C cannot move data from B to C.

Collaboration with PRACE

(background)

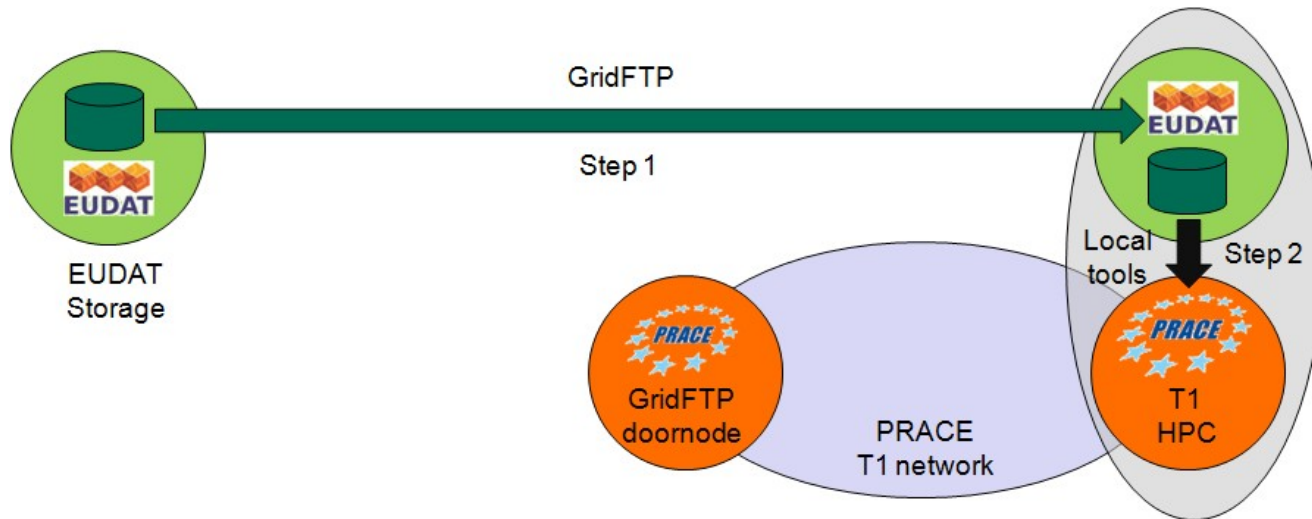
- Some EUDAT Communities (EPOS, VPH) manifested their interest in staging data onto and from PRACE resources.
- Many EUDAT sites are member of PRACE either.
- Staging of data within the PRACE Infrastructure is enabled through the GridFTP protocol under precise conditions: a) staging of data from the public network is permitted only through PRACE door nodes, namely CINECA and LRZ; b) staging of data within the private network is permitted without any restriction through

Two possible scenarios (1)



- Data transit through the PRACE door node before reaching the target HPC facility

Two possible scenarios (2)

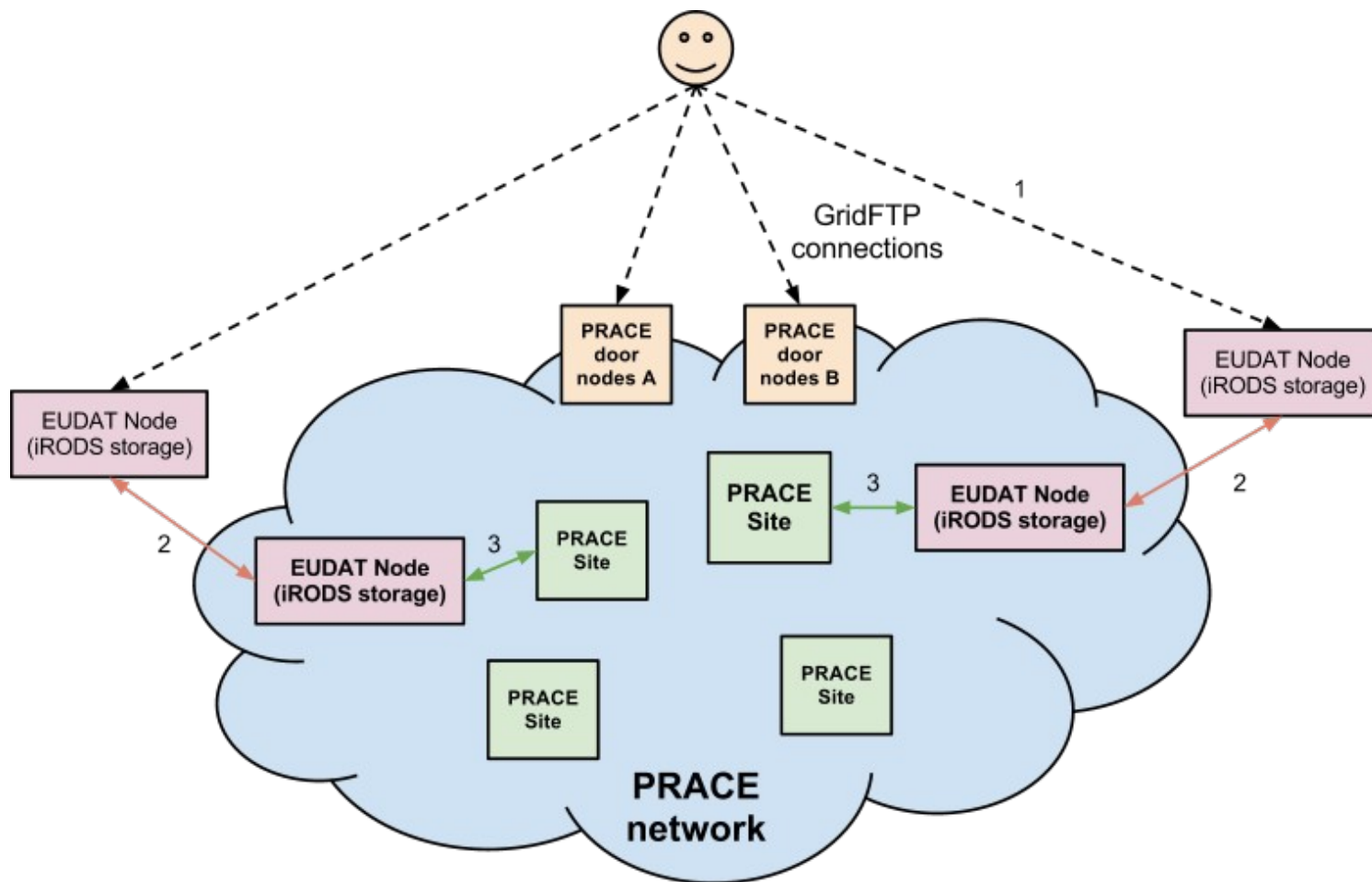


- Data is replicated onto an EUDAT node which is “closer” to the target HPC facility. What happens to those sites being not part of PRACE?

Possible collaboration plan (1/2)

- Few EUDAT nodes enter the PRACE internal network. Data to be staged on PRACE facilities are first replicated to one of these nodes and then staged to the final destination. This scenario will open a new collaboration perspective as, on one hand, EUDAT/PRACE users would be granted with the possibility to stage their data to PRACE facilities, on the other, PRACE users could be provided with the possibility to deposit their results into EUDAT storage and request EUDAT to preserve them. *An ideal scenario would be that*

Possible collaboration plan (2/2)





Thanks for your attention

- Q&A
- :wq