Dynamic replication

How to stage large data sets to HPC systems

Giuseppe Fiameni, Claudio Cacciari
CINECA, Italy
Mark van de Sanden
SARA, The Netherlands

EUDAT User Forum, Barcelona 7/8th March 2012

SEVENTH FRAMEWORK PROGRAMME

EUDAT

Outline of the talk

- Community Service Wishes
- Overview
- Principles
- Scenario
- On-going activity

Community Service Wishes

In Progress as Services (Task Forces set up)

- Safe Data Replication (for Bit-stream Preservation & Access Optimization)
- Dynamic Data Replication into HPC Workspace

In Specification/Discussion as Services

- Aggregated EUDAT Metadata Domain
- Researcher Data Store (Simple Upload, Share and Access)
- Common Authentication/Authorization Infrastructure

In Progress as Research Issues (WP7)

- more elaborate policy rules and federation scalability
- generic workflow execution framework
  (automatic annotation, data mining, etc.)

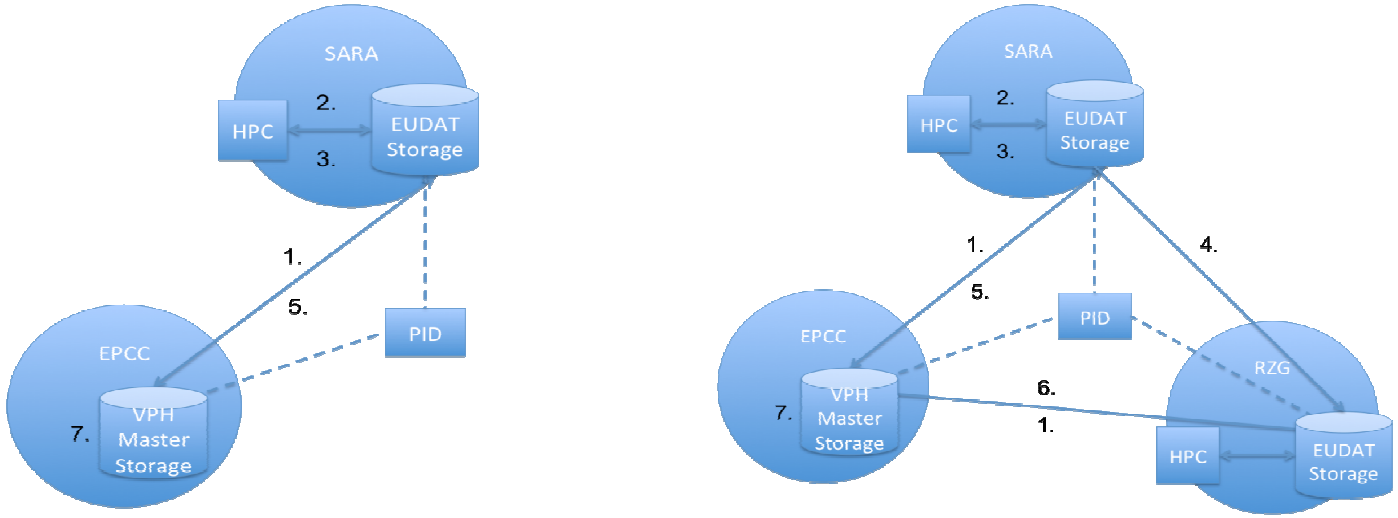
Dynamic data replication

Overview

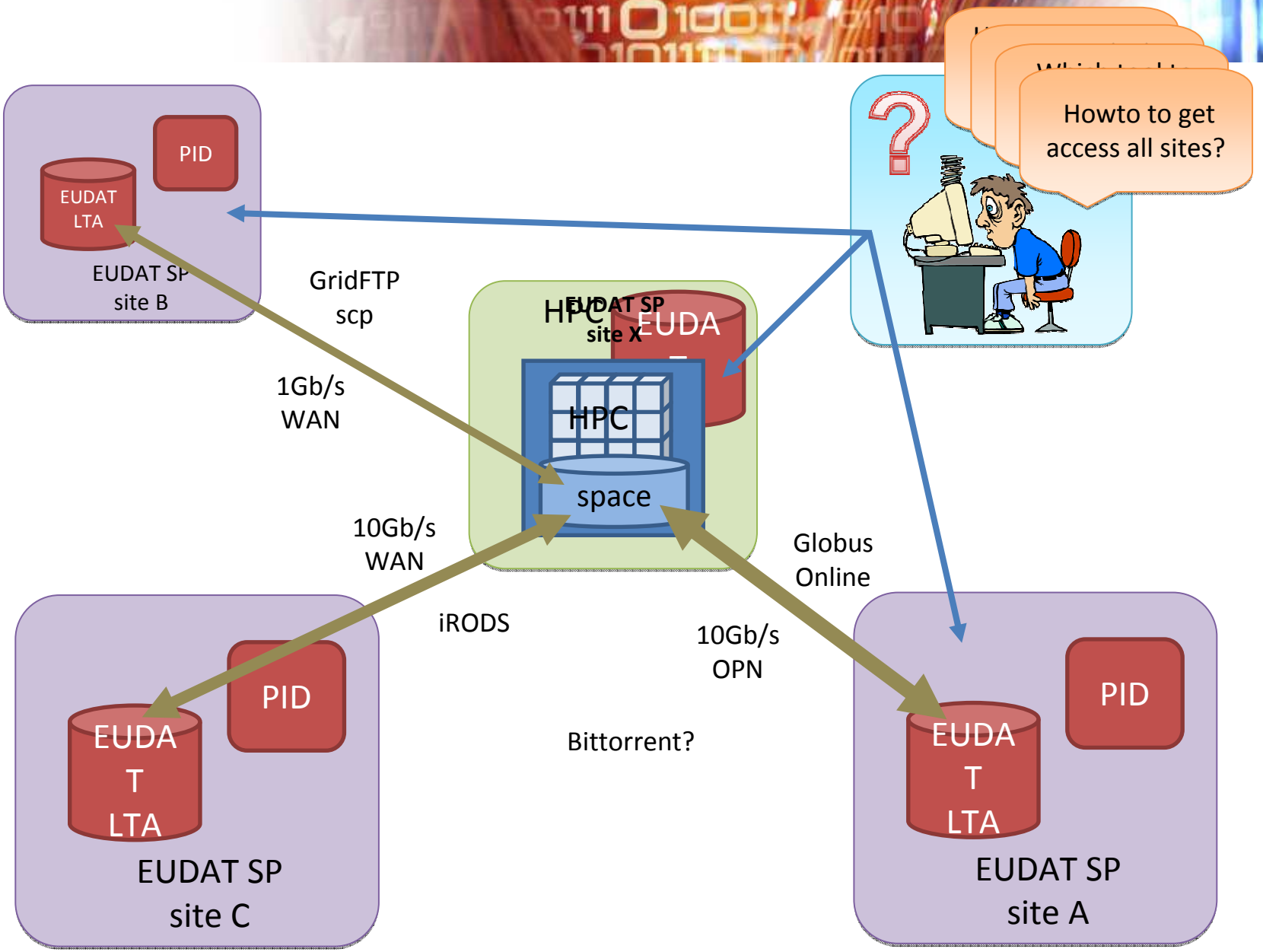
- Easily replicate date sets nearby or onto a HPC workspace for being further processed
- Various communities interested in this service case (VPH, EPOS, ENES)
- Data sets can be very large and composed of many files
- Existing research infrastructures (i.e. PRACE) already provide their own data staging services
- Use case is not new, PRACE users have the same problems between Tier 0 and Tier 1 systems
- Not only HPC system but other data mining cluster could be integrated in workflow framework

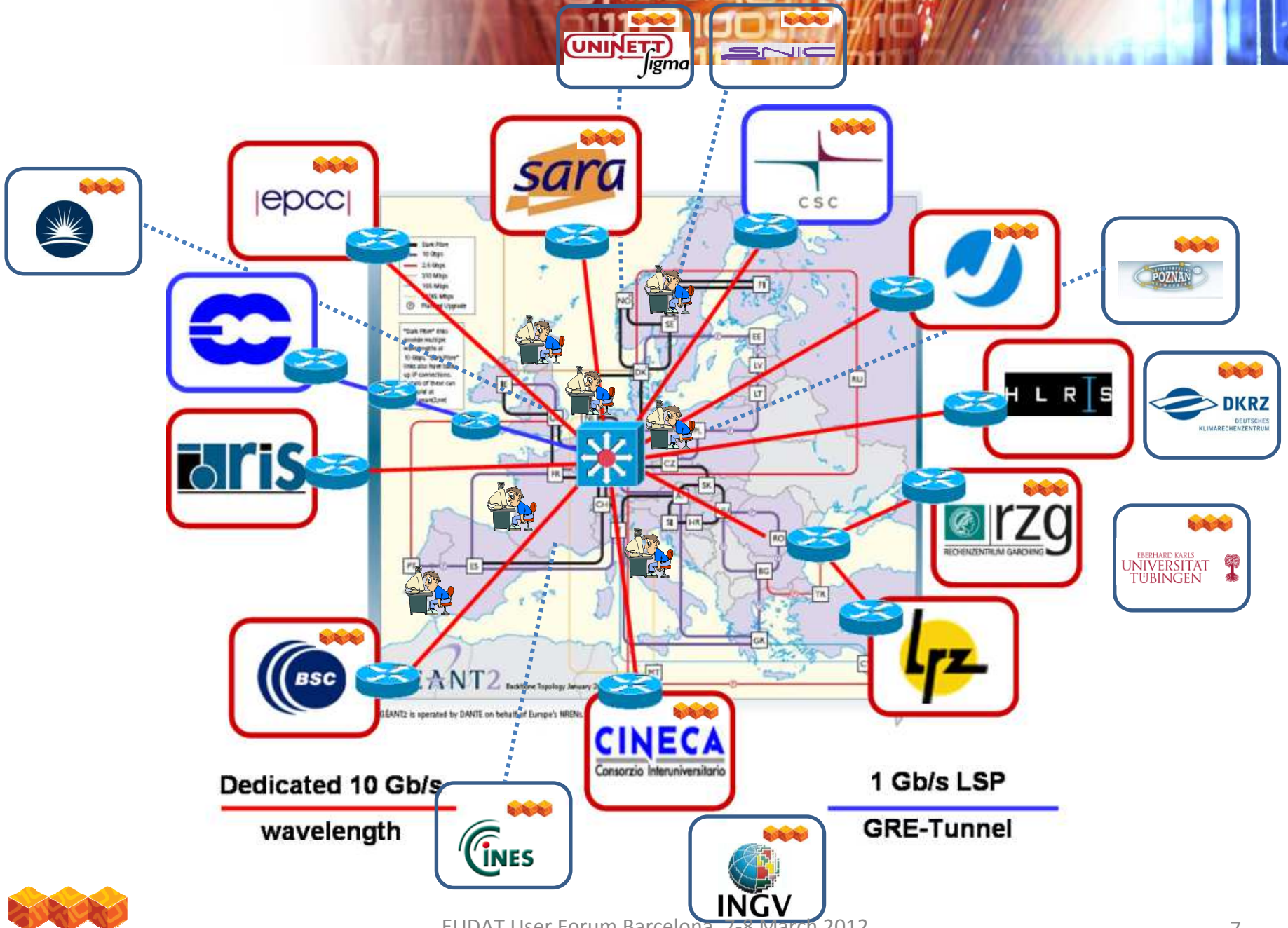
Dynamic data replication

- 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 are
  - 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
  - input data is discarded after job end
  - results are stored back into the storage


PID
EUDAT LTA
EUDAT SP site B
GridFTP scp
1Gb/s WAN
HPC
EUDAT SP site X
EUDA
HPC
space
10Gb/s WAN
iRODS
Globus Online
10Gb/s OPN
Bittorrent?
Howto to get access all sites?
PID
EUDA T LTA
EUDAT SP site C
PID
EUDA T LTA
EUDAT SP site A


UNINETT sigma
SNIC
epcc
sara
CSC
POZNAN
HLRS
DKRZ
DEUTSCHES KLIMARECHENZENTRUM
rzg
RECHENZENTRUM GARCHING
EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN
BSC
CINECA
Consorzio Interuniversitario
INES
INGV
Dedicated 10 Gb/s
wavelength
1 Gb/s LSP
GRE-Tunnel

First Scenario

WAN

Data transport protocols:
GridFTP, HTTP. SCP, FTP, RSYNC, OpenDap?

Which tool to manage the transfers:
iRODS, Globus Online, FTS, RSYNC, Others…….?

WAN
Third party transfers

PID

EUDA
T
LTA

EUDAT SP
site A

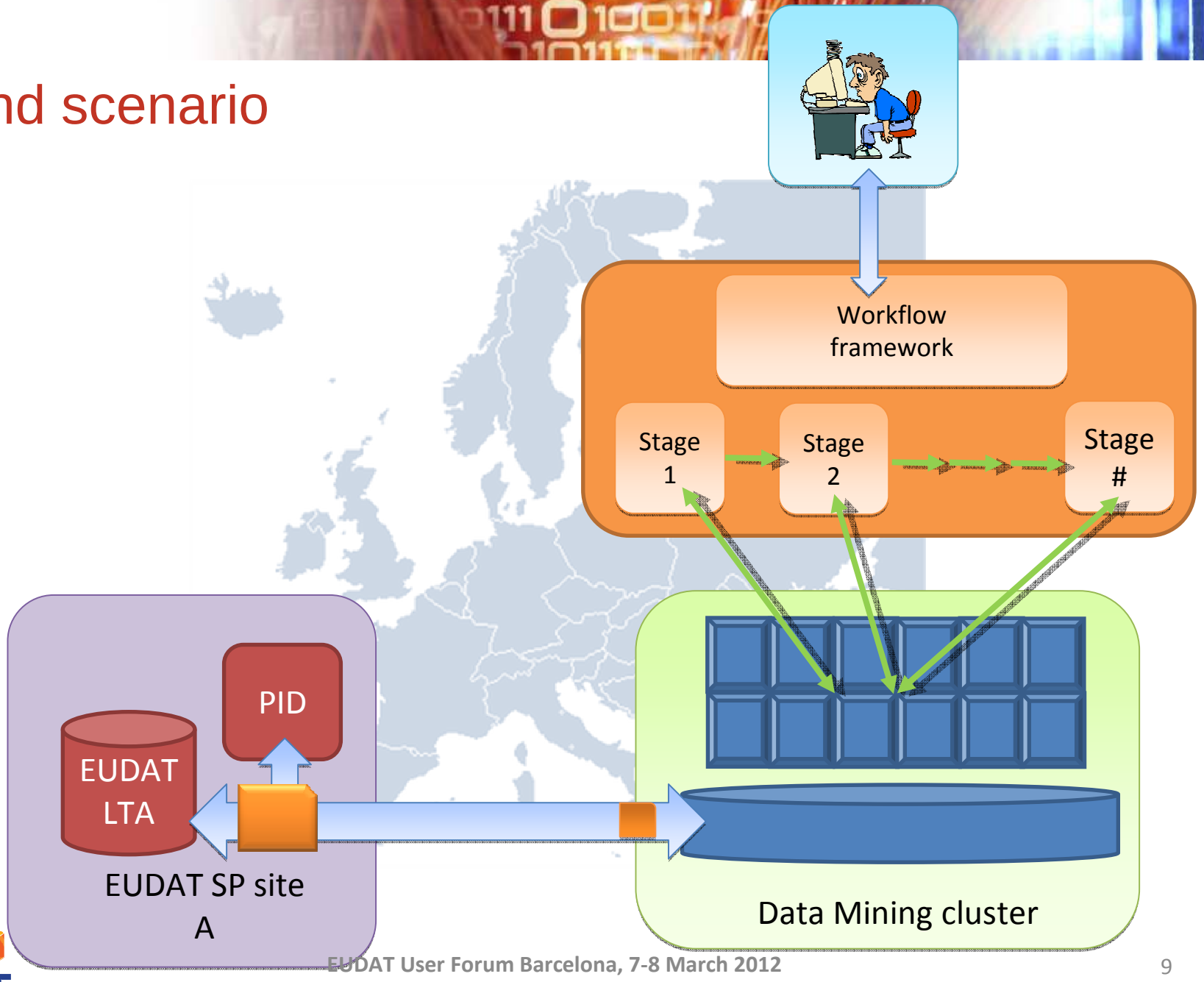
HPC site X

HPC

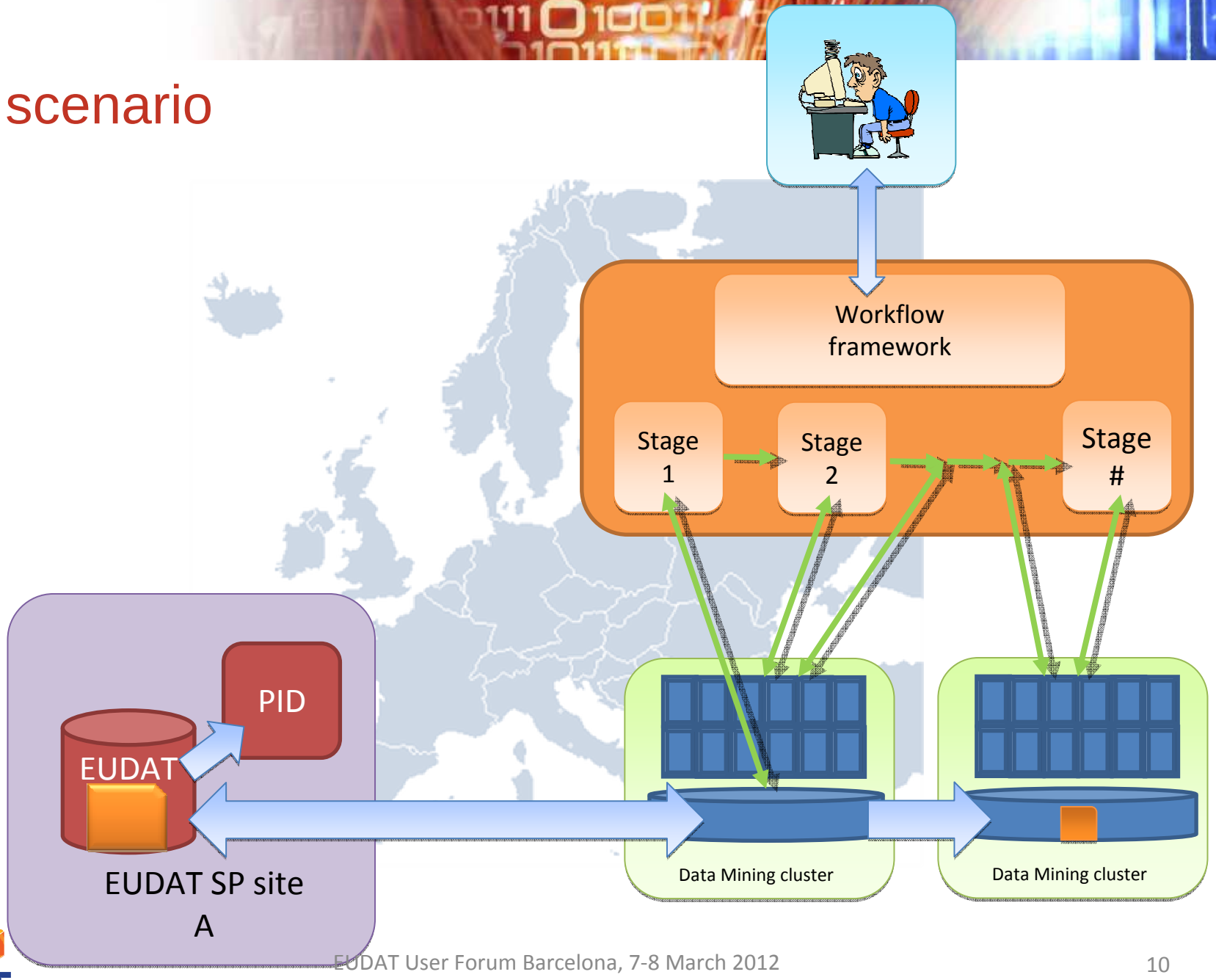
space

Dedicated Network
Third party transfers

Second scenario



Third scenario



PRACE

WAN

Data transport protocols:
GridFTP, HTTP. SCP, FTP, RSYNC, OpenDap?

Which tool to manage the transfers:
iRODS, Globus Online, FTS, RSYNC, Others…….?

PRACE T0

HPC

space

PRACE T1

HPC

Dedicated Network
Third party transfers

Ongoing activity

- A task force is working to implement two different pilots
  - Pilot 1 – The target HPC centre is an EUDAT node
    - CSC and VPH involved
    - Integration of low level data transfer tool (i.e. rsync, cp) within the EUDAT node
  - Pilot 2 – The target HPC centre is part of the PRACE infrastructure
    - INGV and CINECA involved
    - Integration of PRACE tools (i.e. GridFTP) within the EUDAT node
- First outcomes to be ready in summer

Many thanks for you attention!