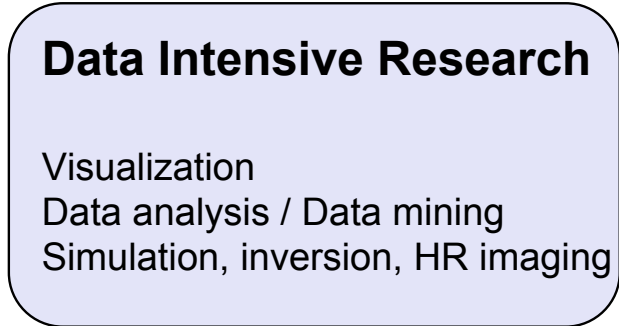# The VERCE Case

Iraklis Klampanos

iraklis.klampanos@ed.ac.uk
University of Edinburgh

12 March 2013

Overview

VERCE

e-Science environment for data intensive research based on an extensive service-oriented architecture

HPC/GRID Infrastructures

egi

PRACE

Data Intensive Research

Visualization
Data analysis / Data mining
Simulation, inversion, HR imaging
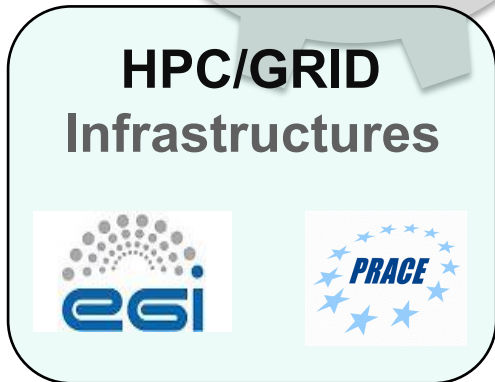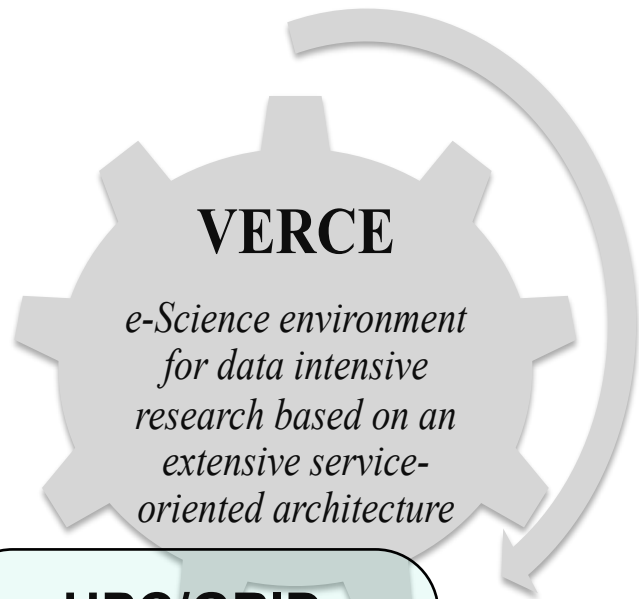
- Empower seismologists with technology to
  - Execute common research-oriented tasks
  - Manage data in a flexible and meaningful way (semantics and metadata)
  - In a unified, user-friendly framework

VERCE
www.verce.eu

VERCE Use-Cases
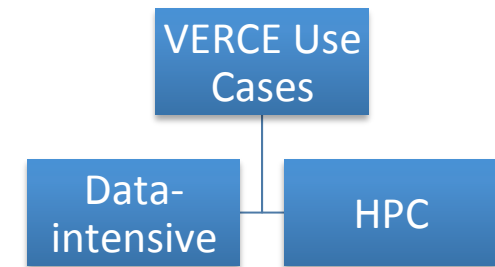
- Broad splitting of tasks based on
  - Current and best practices and software req's
  - Project resources at hand (private, Grid, HPC)
  - Prior technology investment (e.g. Dispel)

- Data movement, staging and archiving

- Meta-data
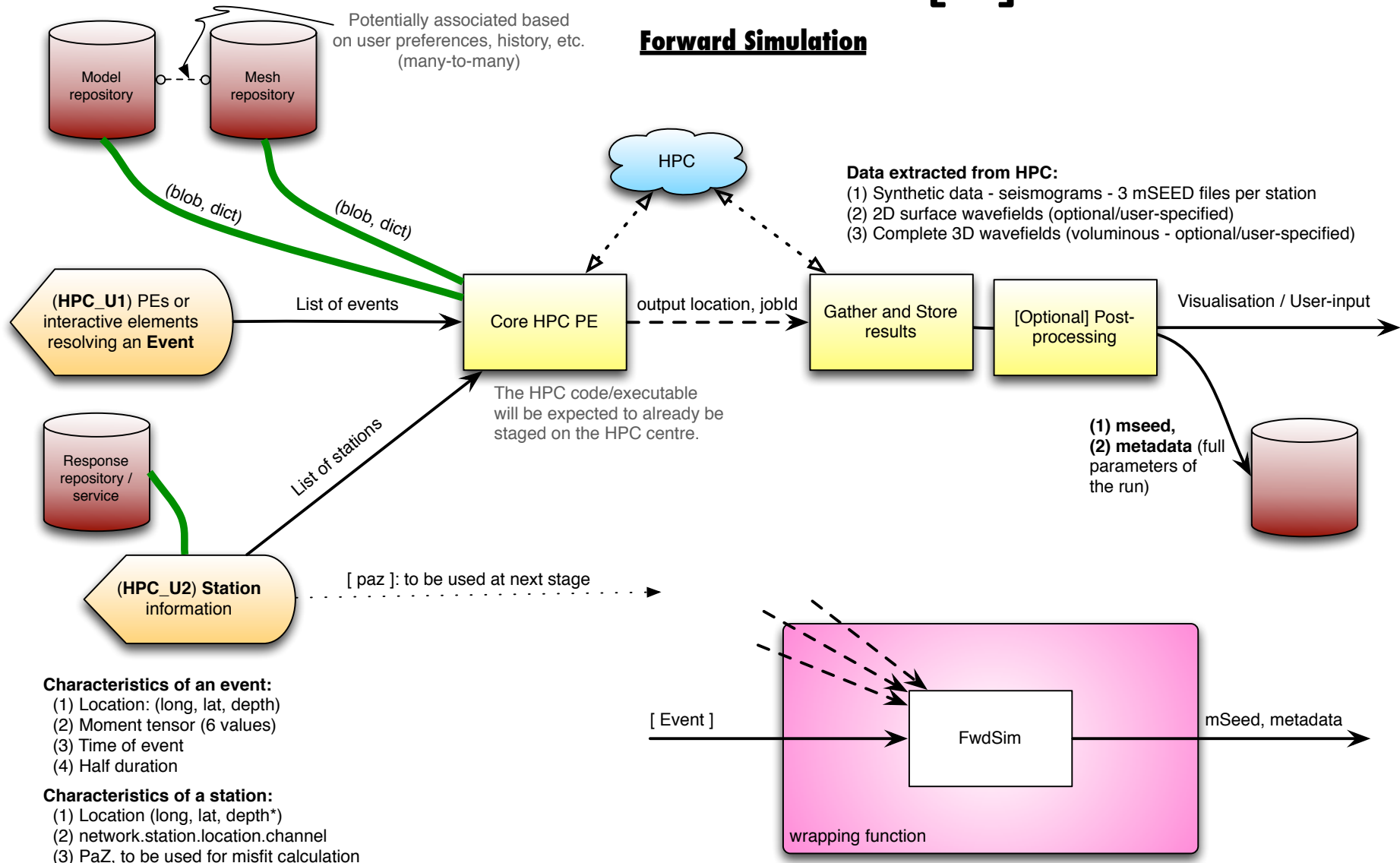 - Annotation, look-up
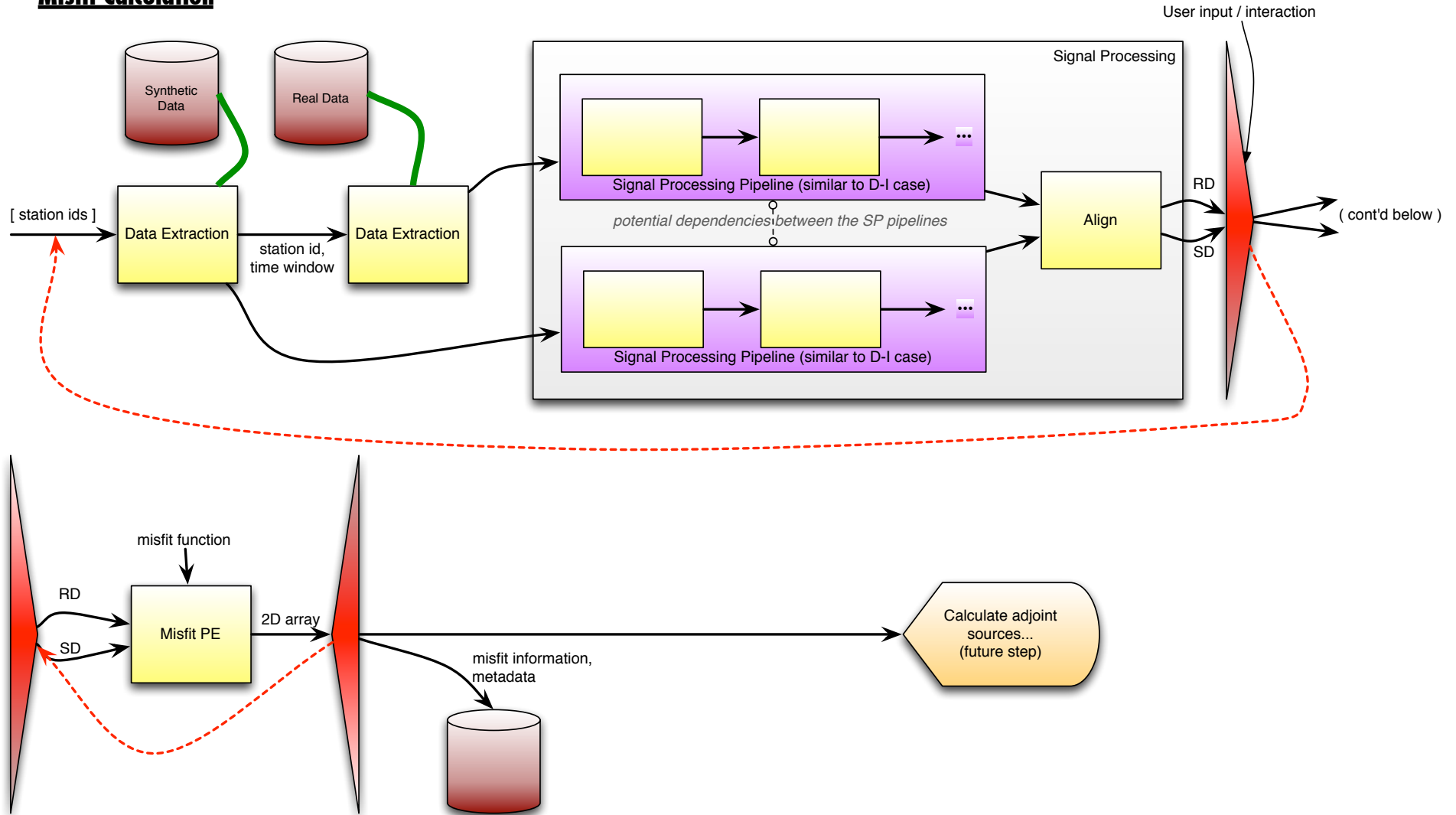 - VERCE-wide common view of meta/data

EGI/PRACE/EUDAT Collaboration

London, 12/03/2013

HPC Use-Case [1]

Forward Simulation

Model repository

Mesh repository

Potentially associated based on user preferences, history, etc. (many-to-many)

(blob, dict)

(blob, dict)

HPC

Data extracted from HPC:
(1) Synthetic data - seismograms - 3 mSEED files per station
(2) 2D surface wavefields (optional/user-specified)
(3) Complete 3D wavefields (voluminous - optional/user-specified)

(HPC_U1) PEs or interactive elements resolving an **Event**

List of events

Core HPC PE

output location, jobId

Gather and Store results

[Optional] Post-processing

Visualisation / User-input

The HPC code/executable will be expected to already be staged on the HPC centre.

Response repository / service

List of stations

(1) mseed,
(2) metadata (full parameters of the run)

(HPC_U2) Station information

[paz]: to be used at next stage

Characteristics of an event:
(1) Location: (long, lat, depth)
(2) Moment tensor (6 values)
(3) Time of event
(4) Half duration

Characteristics of a station:
(1) Location (long, lat, depth*)
(2) network.station.location.channel
(3) PaZ, to be used for misfit calculation

[Event]

FwdSim

mSeed, metadata

wrapping function

VERCE

www.verce.eu

EGI/PRACE/EUDAT Collaboration

London, 12/03/2013

HPC Use-Case [2]

Resources

- PRACE
  - PLX, FERMI, SuperMUC
  - To be executing solvers such as Seissol, SpecFem3D
    - Already in place – integrated with Dispel
    - jSAGA, DCI-Bridge

- EGI
  - SCAI, LRZ, CINECA
  - Not yet integrated – near-future work

VERCE
www.verce.eu

Raw data management strategy

- Data layer among VERCE sites
  - Initially IRODS
  - Including PRACE, EGI and **private** VERCE resources

- Domain-specific metadata extraction layer
  - Each site will be able to query data based on properties relevant to seismologists
  - Used through Dispel or the Science Gateway

- Initial prototype: 3-6 months

Immediately Useful Actions

- VERCE is now starting to integrate resources within the adopted workflow paradigm (Dispel)

- Computing facilities (PRACE, EGI, private) are fixed – we don't need to discover them (yet)
  - Perhaps in the future

- Long-term storage/archiving of final results, medium term for intermediate results.

- More unified policy w.r.t. authentication/ authorisation

Testing and Evaluation

- Long-term storage
  - Could start after ~6 months
  - Integration of VERCE data-management layer with EUDAT storage resources
  - PID and minimal metadata
    - Would be associated with VERCE internal records with potentially more domain-specific metadata

- Authentication/Authorisation
  - Within VERCE's HPC and Data Intensive use cases