Koninklijk Nederlands
Meteorologisch Instituut
Ministerie van Infrastructuur en Milieu

User Centered Provenance Management for Data-Intensive Platforms

Alessandro Spinuso
spinuso@knmi.nl

Respect des Fonds - Historical Facts.

France 1790, Establishment of "Les Archives Nationales" trying to merge private and public archives throughout the country

... 50 years of archivists headache trying to regroup and classify records ...

France 1841**, formulation of the **principle of provenance:

The unity of the archive took precedence over the material objects.

Classification based on administration, organisation, individual, or entity by which they were created.

http://ingmarbergman.se/en/category/tags/principle-provenance
http://en.wikipedia.org/wiki/Respect_des_fonds

The Ambiguous Origins of the Archival Principle of "Provenance"

Shelley Sweeney

"Broadly speaking, the word "provenance," whether used by a rare book librarian, an archaeologist, an art curator, or an archivist, refers to the origins of an information-bearing entity or artifact. But there the consensus ends."

Archeology: *"the place where an object was found or recovered in modern times; the findspot"*

Librarians: *"Information concerning the transmission or ownership, as of a book". (Never used for retrieval)*

Geology: *"The reconstruction of the history of sediments movements over time" (Many types of detrital records to unveil regional tectonic history)*

Wines: *"A documented history of wine cellar conditions". (transactions of old wine with the potential of improving with age)*

How does this apply to the curation of Scientific Data?

How do we use it and which are the Costs?

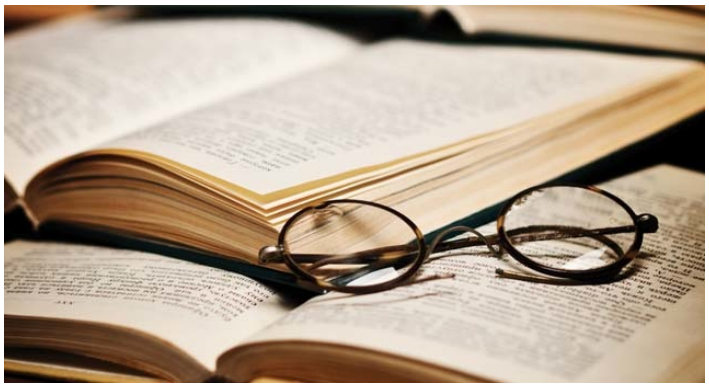
- **Who are the archivists ?**
- Which percentage of the actual provenance recordings is relevant ?
- **Unified models (W3C PROV-O) are useful, how about the content ?**
- How do we produce provenance data (Automated / Manual)
- **Now that CPU time and data bandwidth is a business model, shall we benefit from what we use and pay ?**

Computer Science

*"The **lineage of data or processes**, as per **data provenance**"*

*"Data sets are **reliable** when the **process** used to create them are **reproducible** and analysable for **defects**"*

*"**Scientific workflows** assist scientists and programmers with **tracking their data through** all transformations, analyses, and **interpretations**."*



Researchers are part of the archiving process.

They know what is relevant to understand their results.
Automated system should provide support for a consistent and effective acquisition of provenance metadata - ***Selective and extensible Provenance***. [A. Misra] [I. Foster.]

Reuse of workflow settings and data in *Collaborative Environments* [P. Missier and B. Ludascher].

Data stream processing engines for Data Intensive computation, present expensive requirements for provenance collection, either in terms of size or I/O [W. D. Pauw]

Virtual Earthquake and Seismology Research Community in Europe

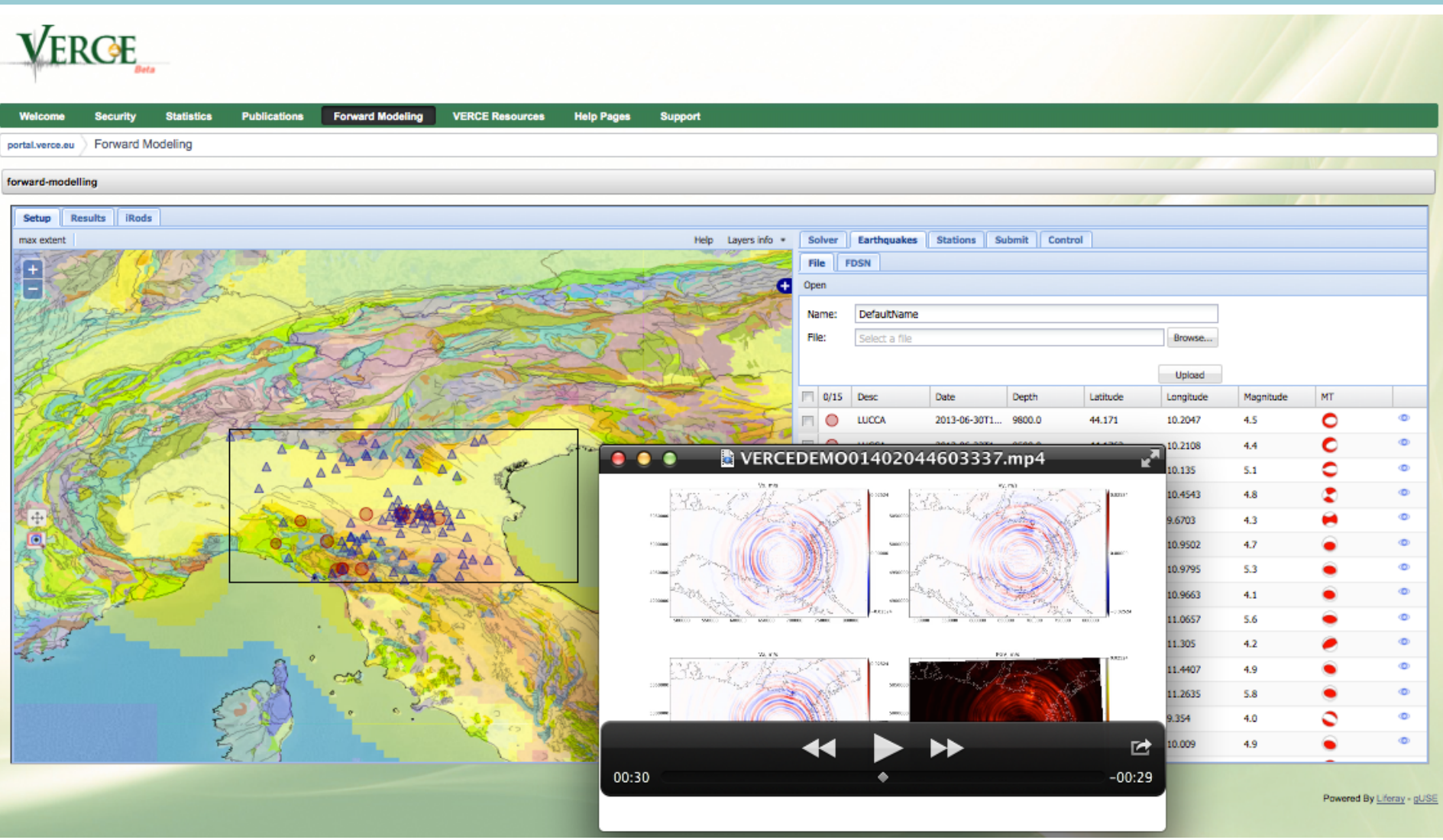
Virtual Environment for of **Earthquakes Simulations** and evaluation of **Earth Models**

http://portal.verce.eu

Combined access to **computing infrastructures** (EGI, PRACE, Local Clusters), for development and execution of large **HPC** computations

Access and use of **European data archives** and services adopting International standards (FDSN, GCMT, OneGeology, EFEHR, QuakeML)

Adoption of Workflow Technologies, Data Management and **Provenance System**

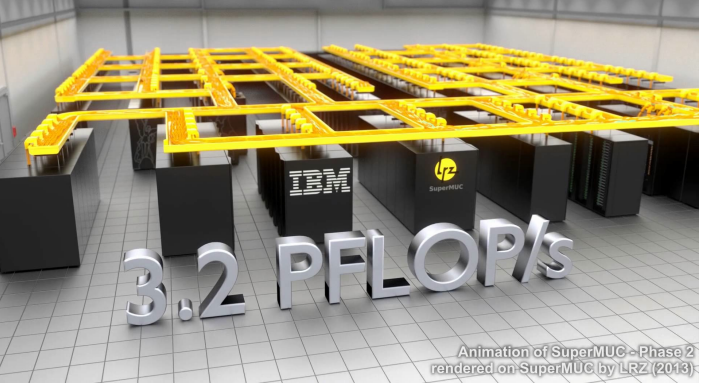

Welcome
Security
Statistics
Publications
Forward Modeling
VERCE Resources
Help Pages
Support
portal.verce.eu
Forward Modeling
forward-modelling
Setup
Results
iRods
max extent
Help
Layers info
Solver
Earthquakes
Stations
Submit
Control
File
FDSN
Open
Name:
DefaultName
File:
Select a file
Browse...
Upload
0/15
Desc
Date
Depth
Latitude
Longitude
Magnitude
MT
LUCCA
2013-06-30T1...
9800.0
44.171
10.2047
4.5
10.2108
4.4
10.135
5.1
10.4543
4.8
9.6703
4.3
10.9502
4.7
10.9795
5.3
10.9663
4.1
11.0657
5.6
11.305
4.2
11.4407
4.9
11.2635
5.8
9.354
4.0
10.009
4.9
VERCEDEMO01402044603337.mp4
00:30
-00:29
Powered By Liferay - gUSE

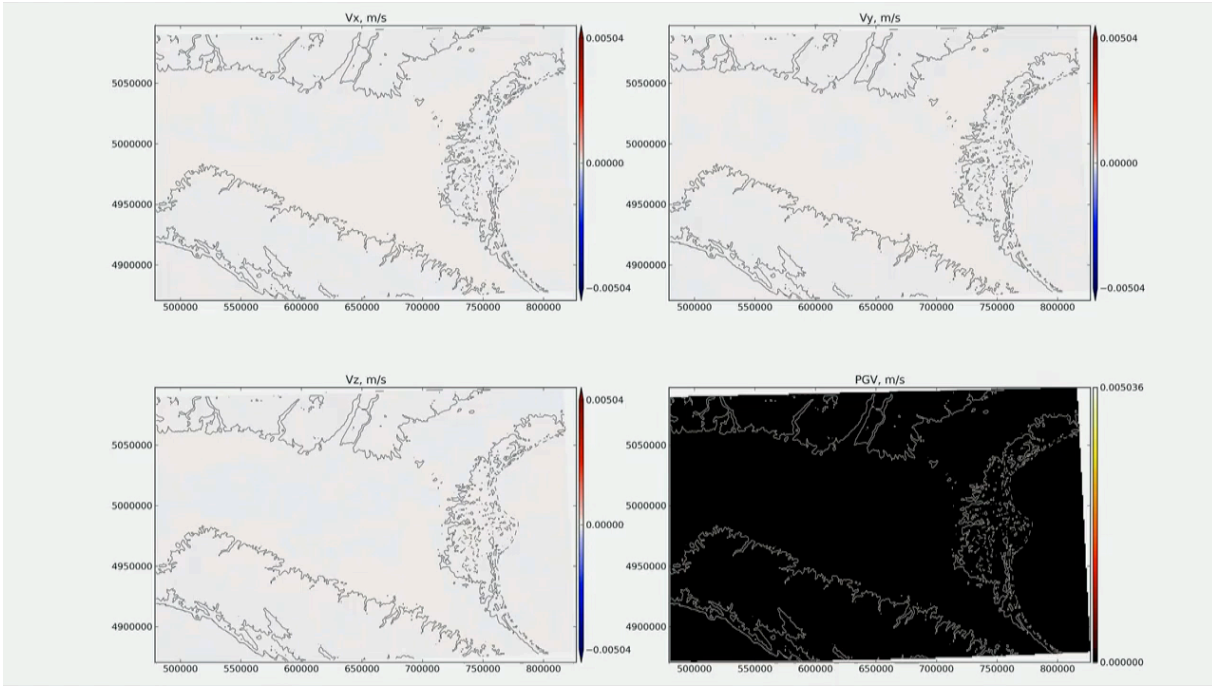
Forward Modelling Use Case

1. Production of **synthetic seismograms** for various **Earth models and earthquakes** on a continental scale Requires the execution of **HPC simulation codes** called solvers **(Simulation)**.

2. The synthetic data may be **compared with real observations (Raw Data Acquisition, MISFIT)**

3. **Model updates and improvement (Inversion)**

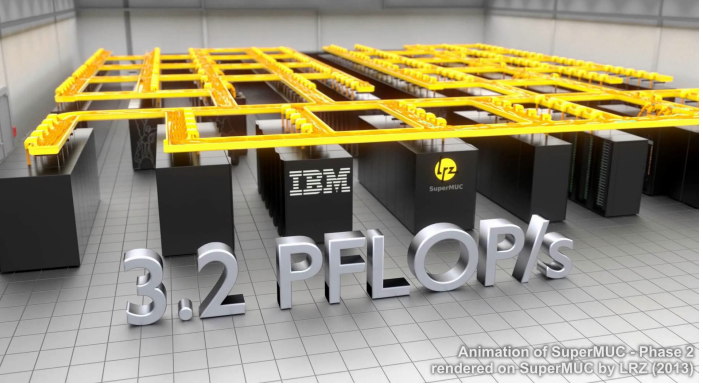


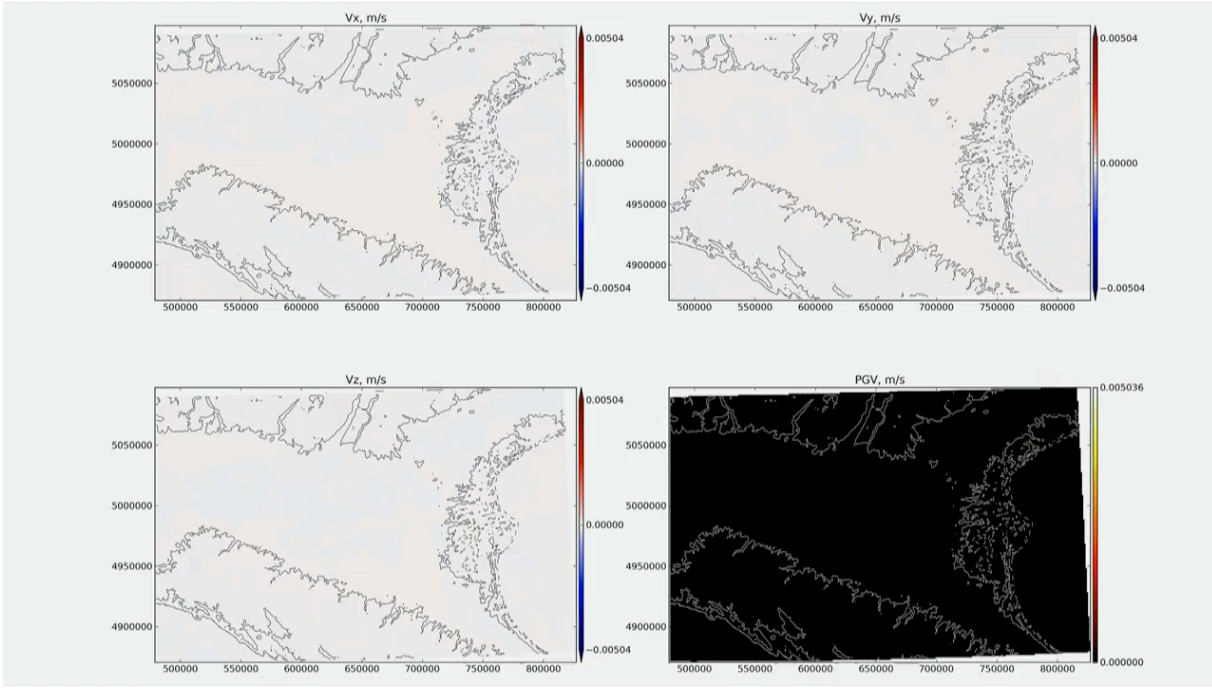


lrz

INGV

KNMI

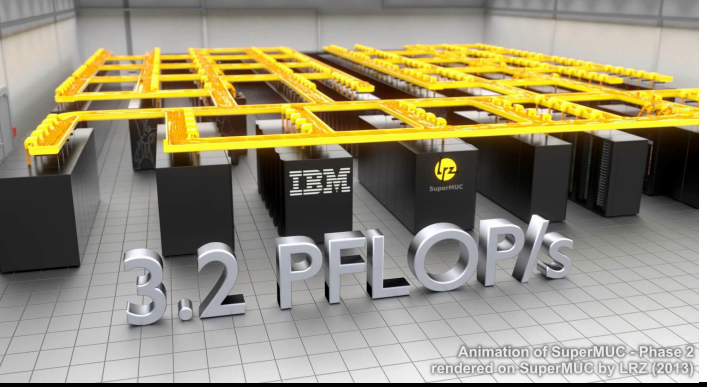


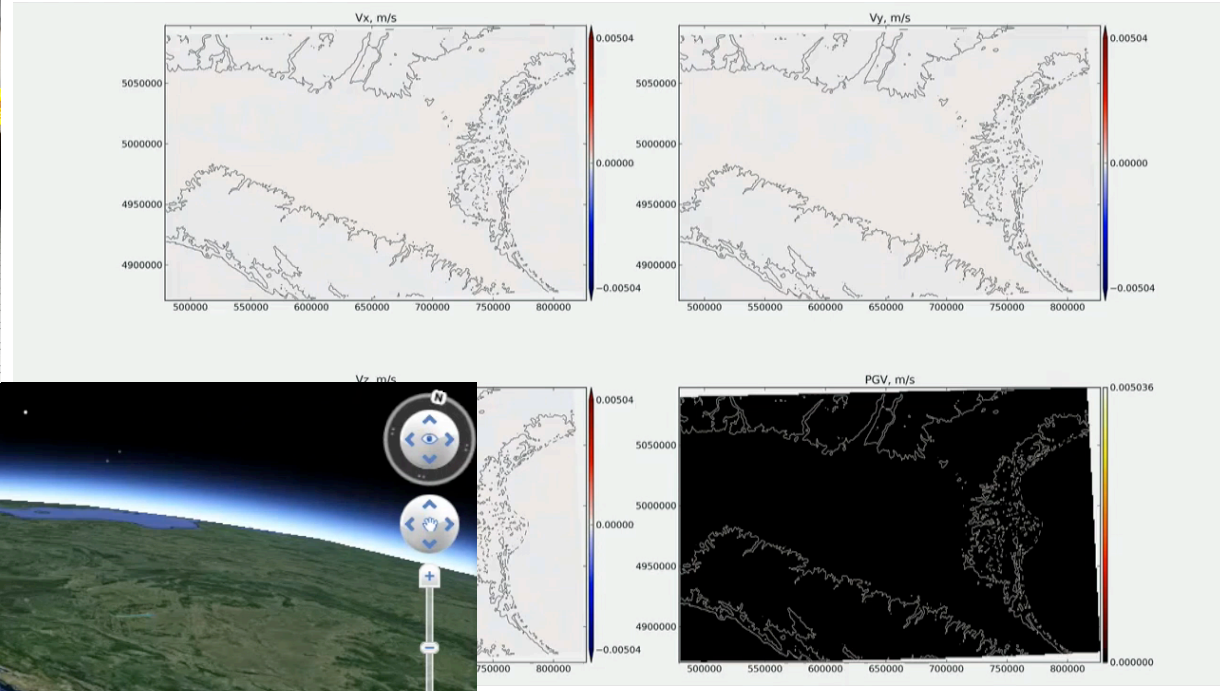


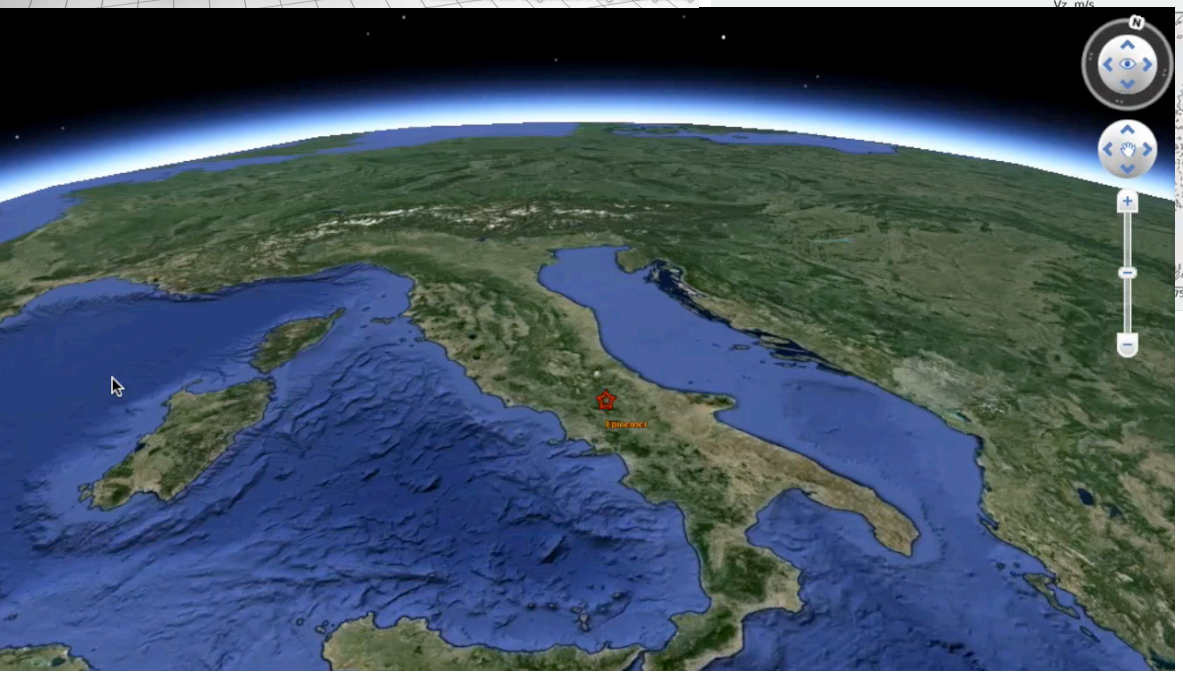
lrz

INGV

KNMI



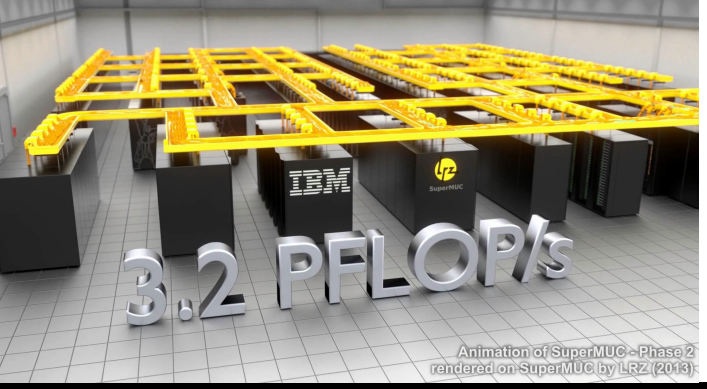


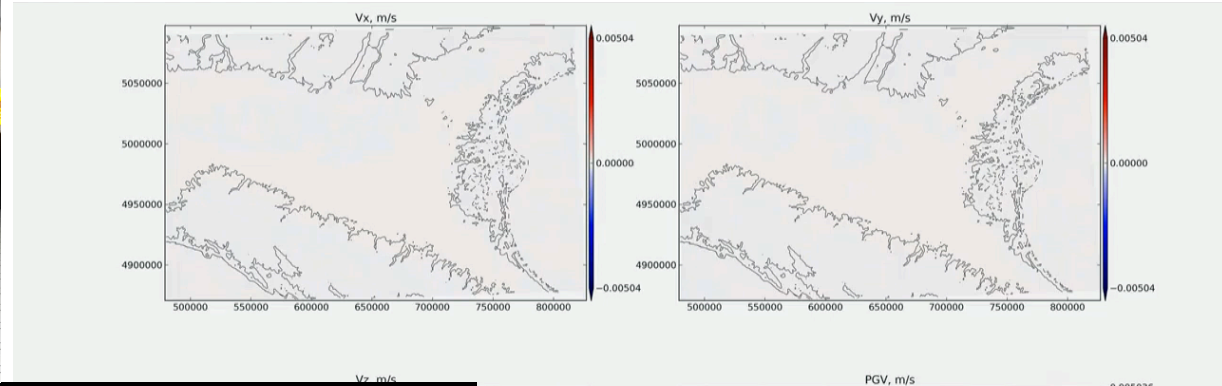


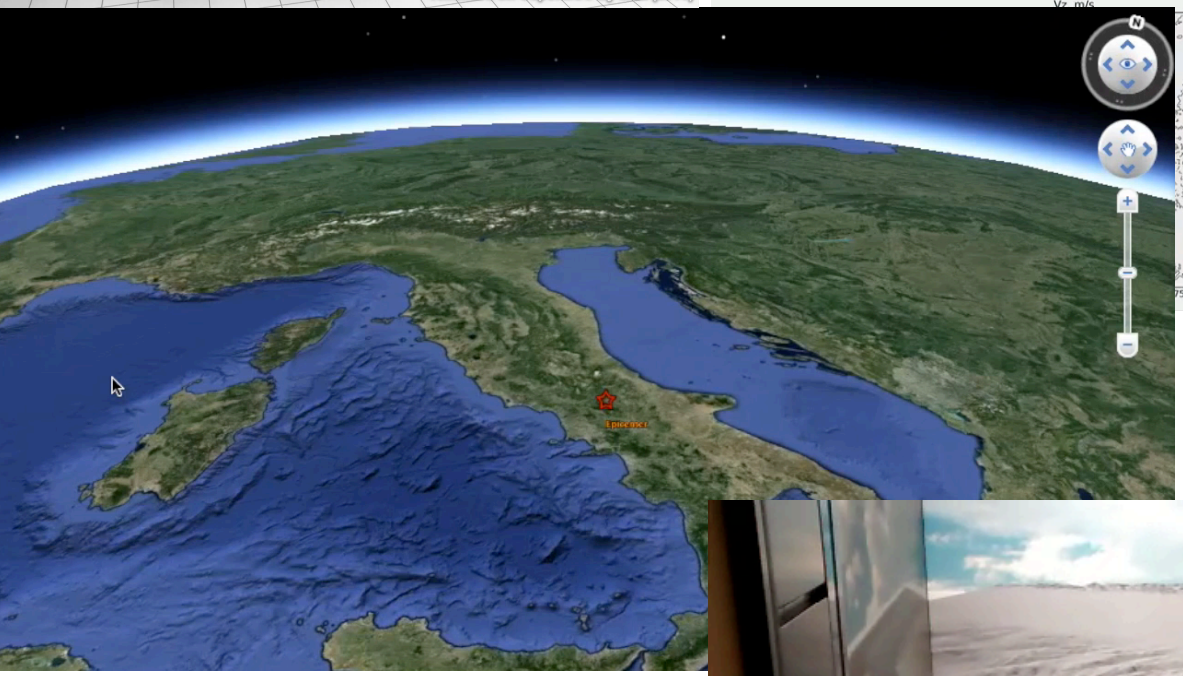
lrz

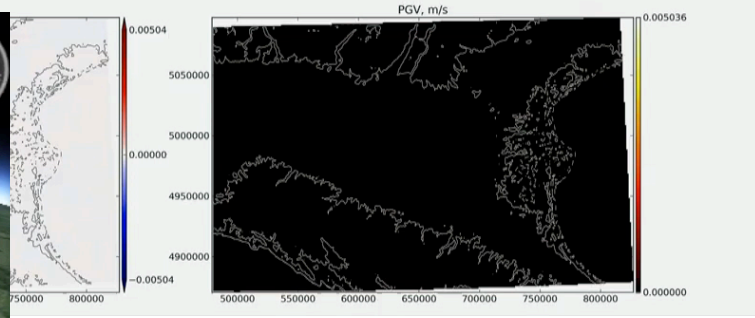
INGV

KNMI









lrz

INGV

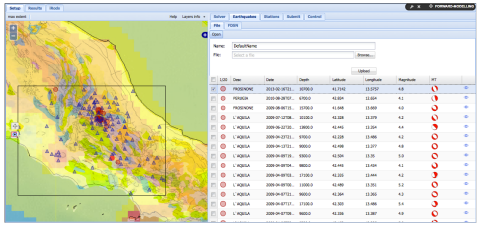
KNMI

VERCE Platform, Components Interaction

Data Management

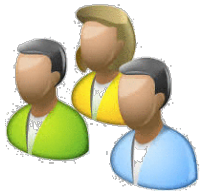
Data Management

Data Archives



Science Gateway
Community Applications

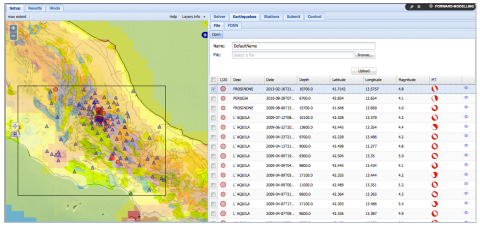
SCI-BUS



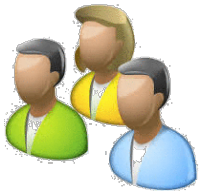
VERCE Platform, Components Interaction

Data Management

Data Management



Science Gateway
Community Applications



SCI-BUS

Data Archives

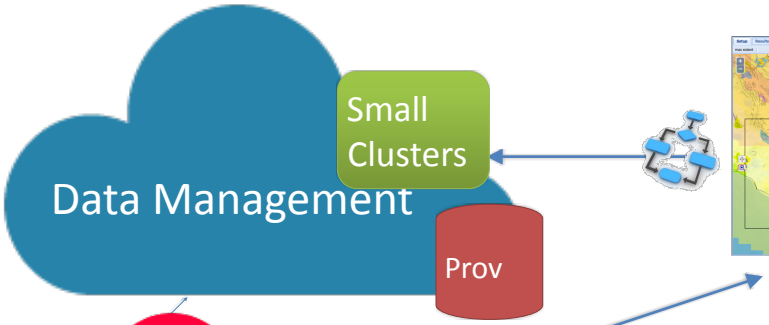
OGC® Open Geospatial Consortium, Inc

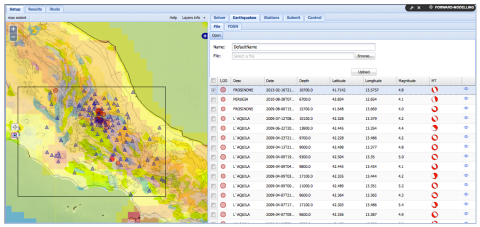
FDSN

VERCE Platform, Components Interaction

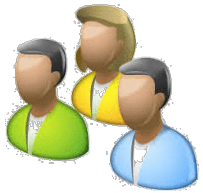
1 - Raw data acquisition

3 - MISFIT

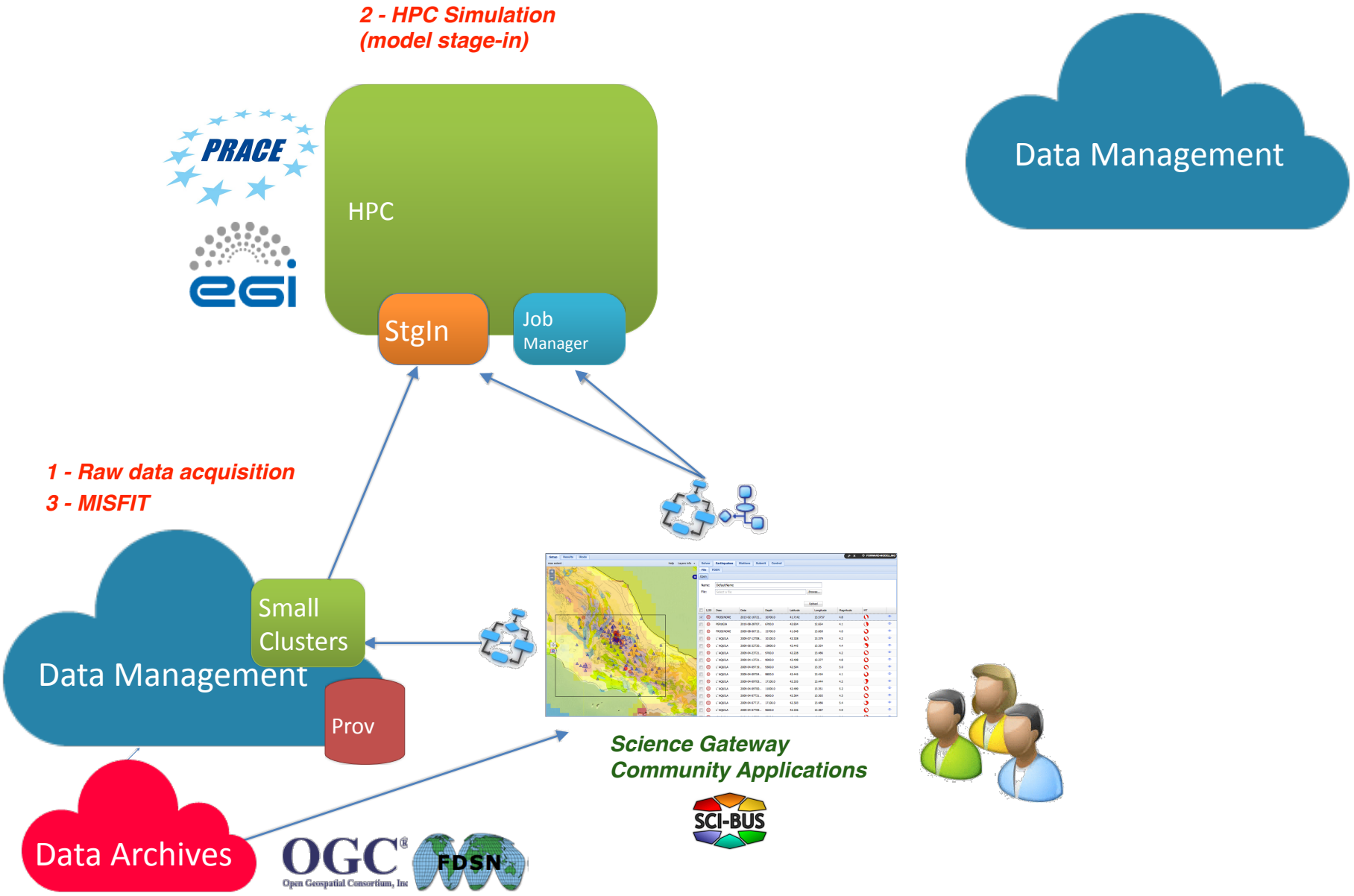




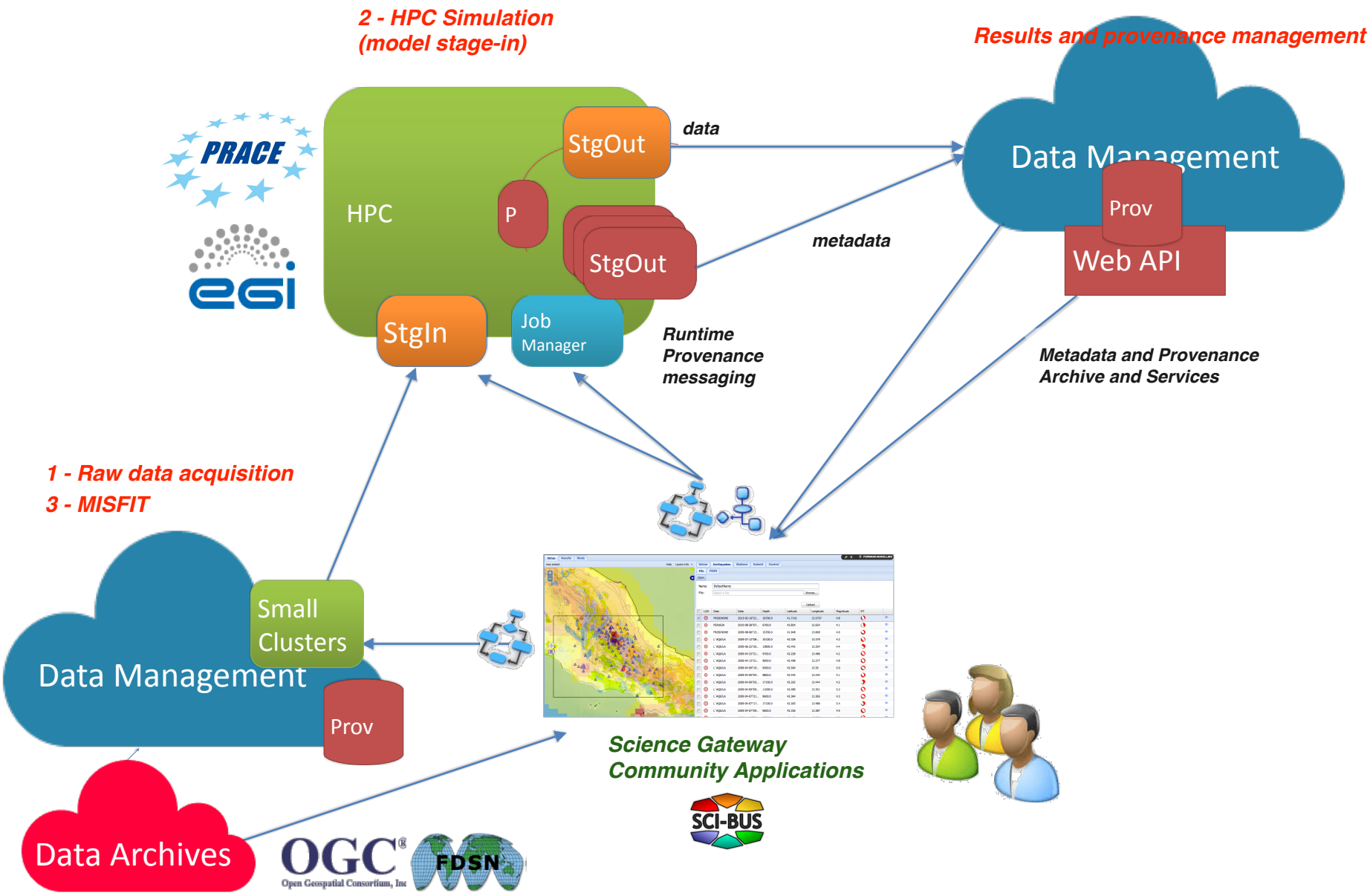
Science Gateway
Community Applications



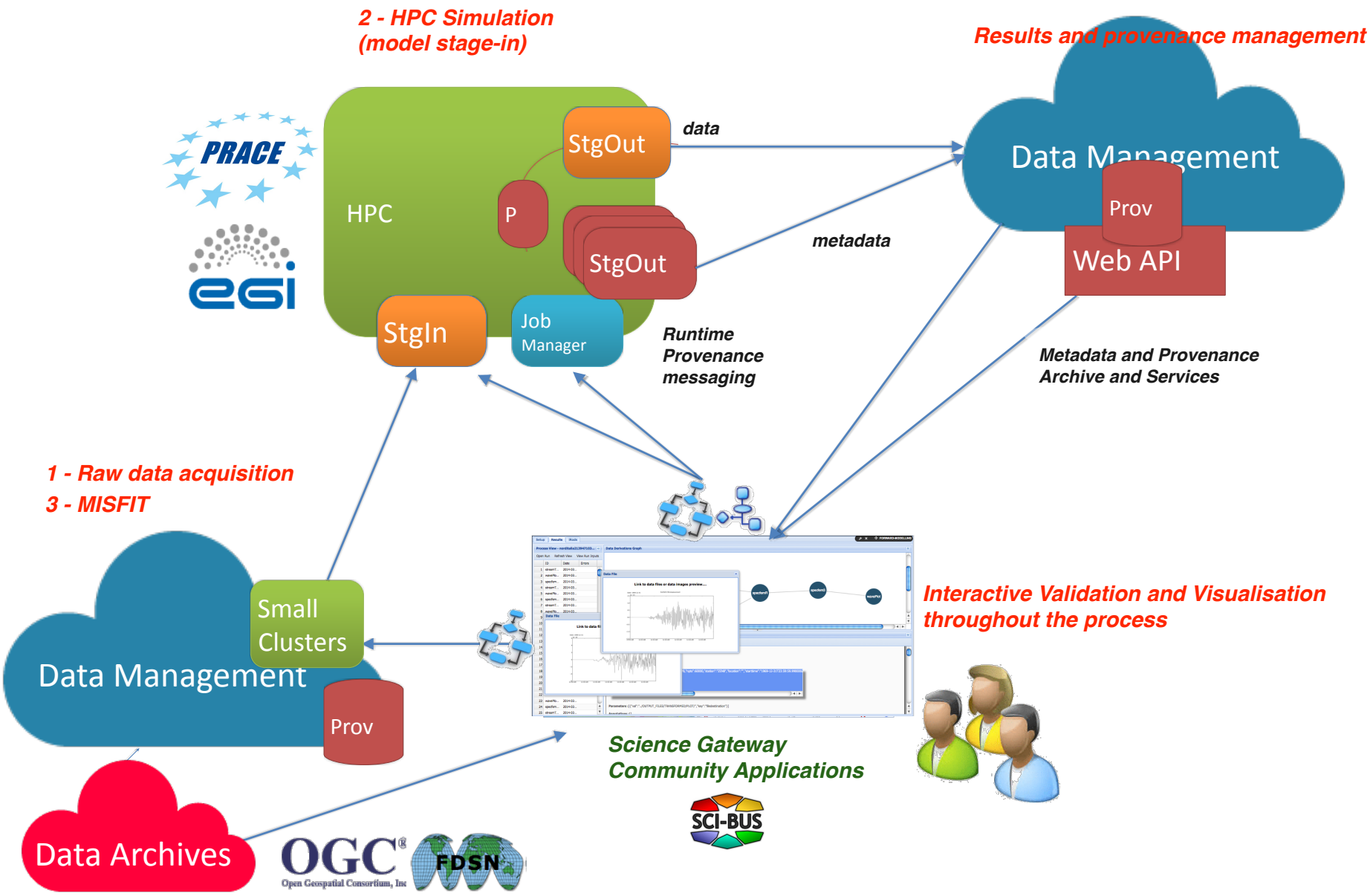
VERCE Platform, Components Interaction



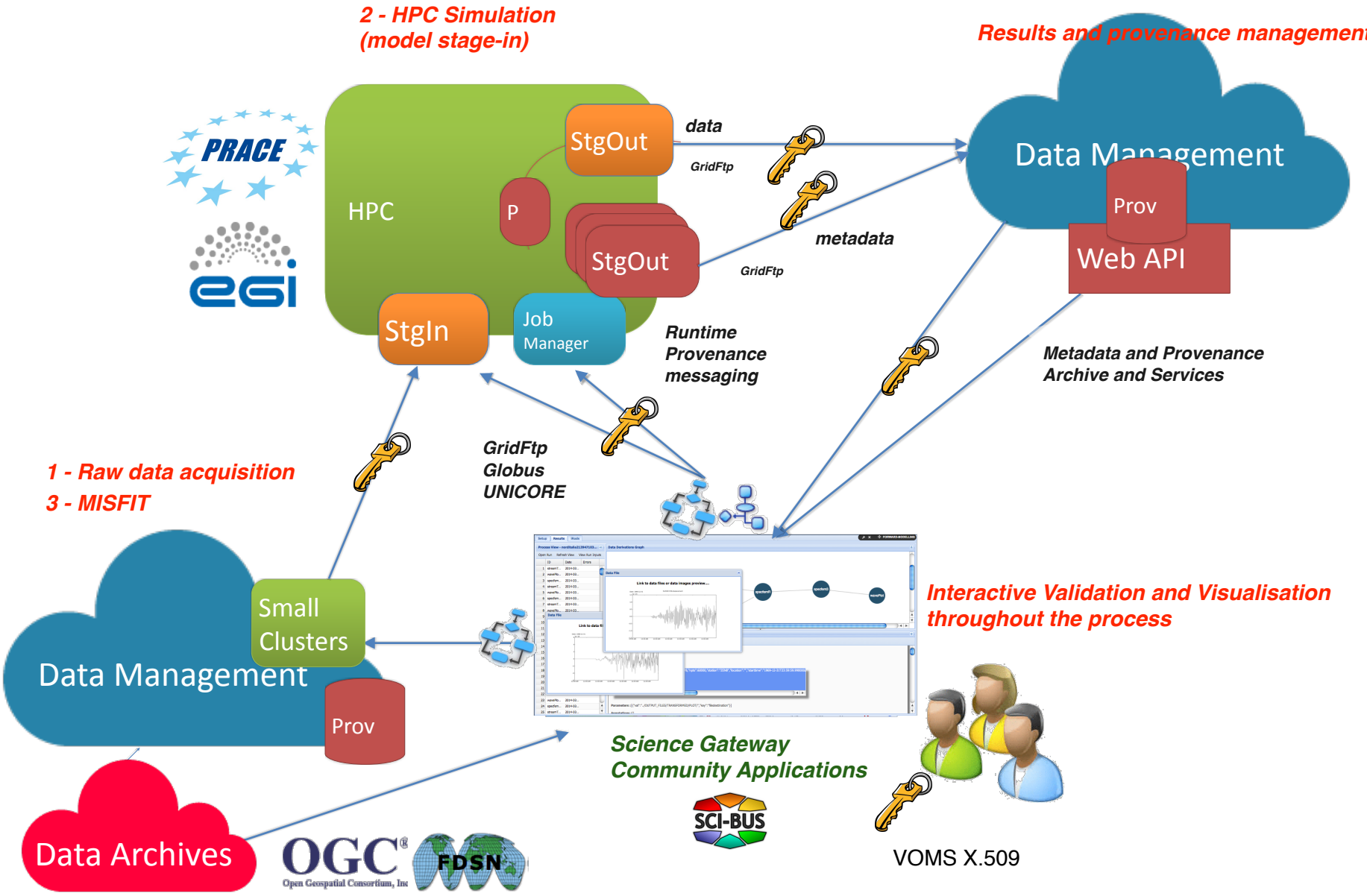
VERCE Platform, Components Interaction



VERCE Platform, Components Interaction



VERCE Platform, Components Interaction



dispel4py

dispel4py.org

Python library used to describe **abstract workflows** for distributed data-intensive applications.

Support for composition: PEs may be defined by having their own internal workflows.

Abstract data-flows described in Dispel4Py can be automatically executed in **numerous parallel environments.**

dispel4py

dispel4py.org

Python library used to describe **abstract workflows** for distributed data-intensive applications.

Support for composition: PEs may be defined by having their own internal workflows.

Abstract data-flows described in Dispel4Py can be automatically executed in **numerous parallel environments.**

MPI

EGI

PRACE

dispel4py

dispel4py.org

Python library used to describe **abstract workflows** for distributed data-intensive applications.

Support for composition: PEs may be defined by having their own internal workflows.

Abstract data-flows described in Dispel4Py can be automatically executed in **numerous parallel environments.**

Storm
Distributed and fault-tolerant realtime computation

MPI

EGI

PRACE

Deployed on local Clouds

dispel4py

dispel4py.org

Python library used to describe **abstract workflows** for distributed data-intensive applications.

Support for composition: PEs may be defined by having their own internal workflows.

Abstract data-flows described in Dispel4Py can be automatically executed in **numerous parallel environments.**

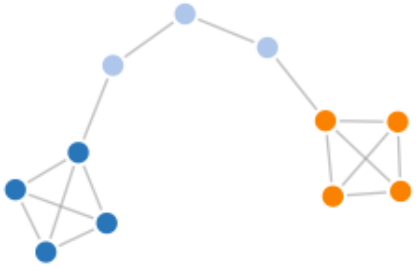
Storm
Distributed and fault-tolerant realtime computation

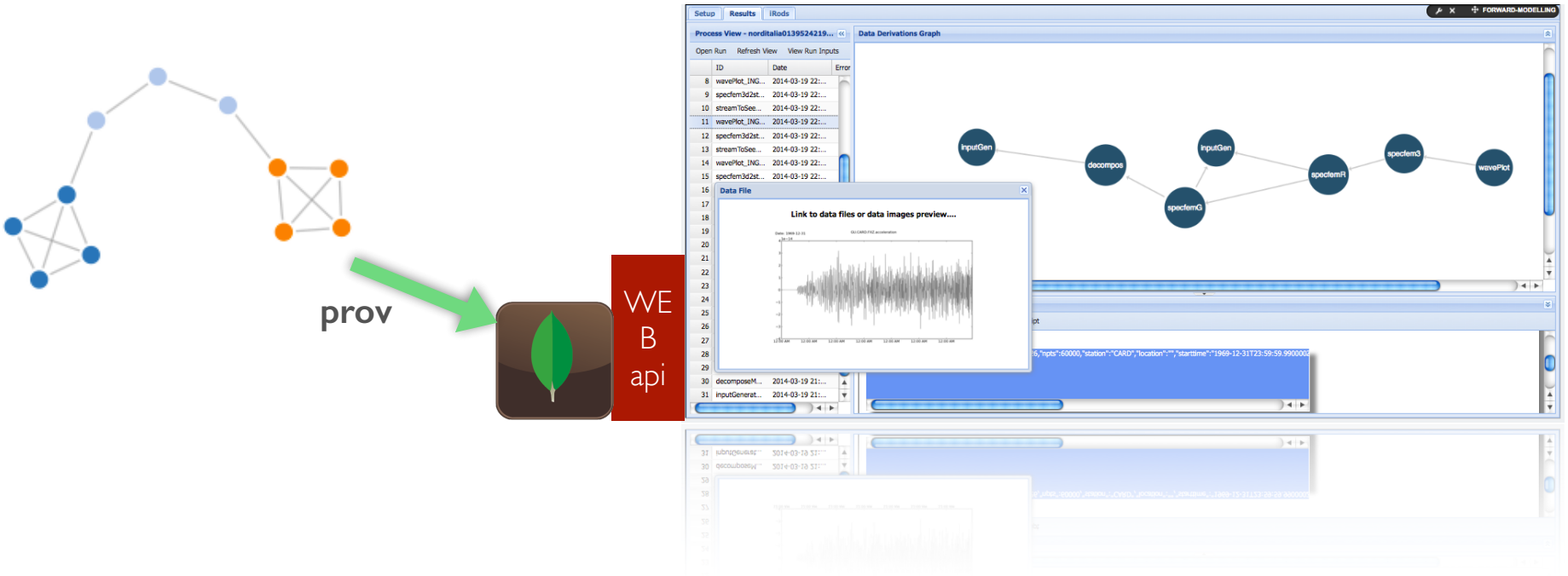
MPI

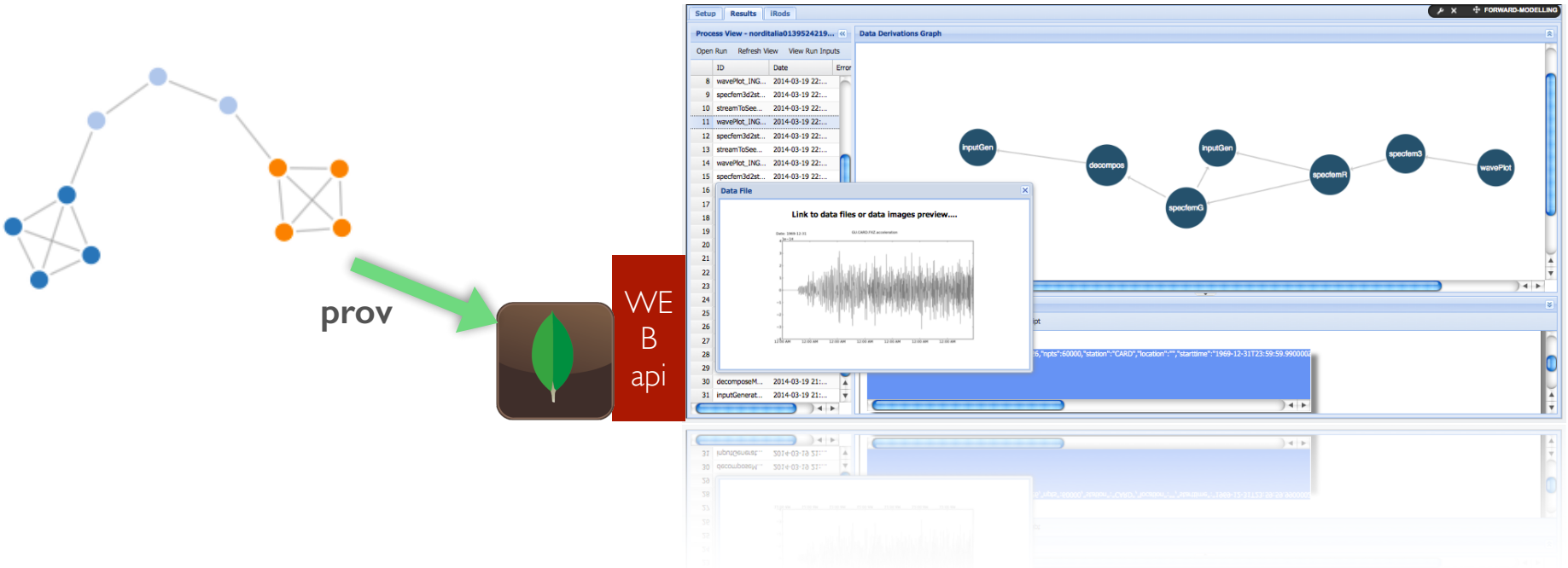
EGI

PRACE

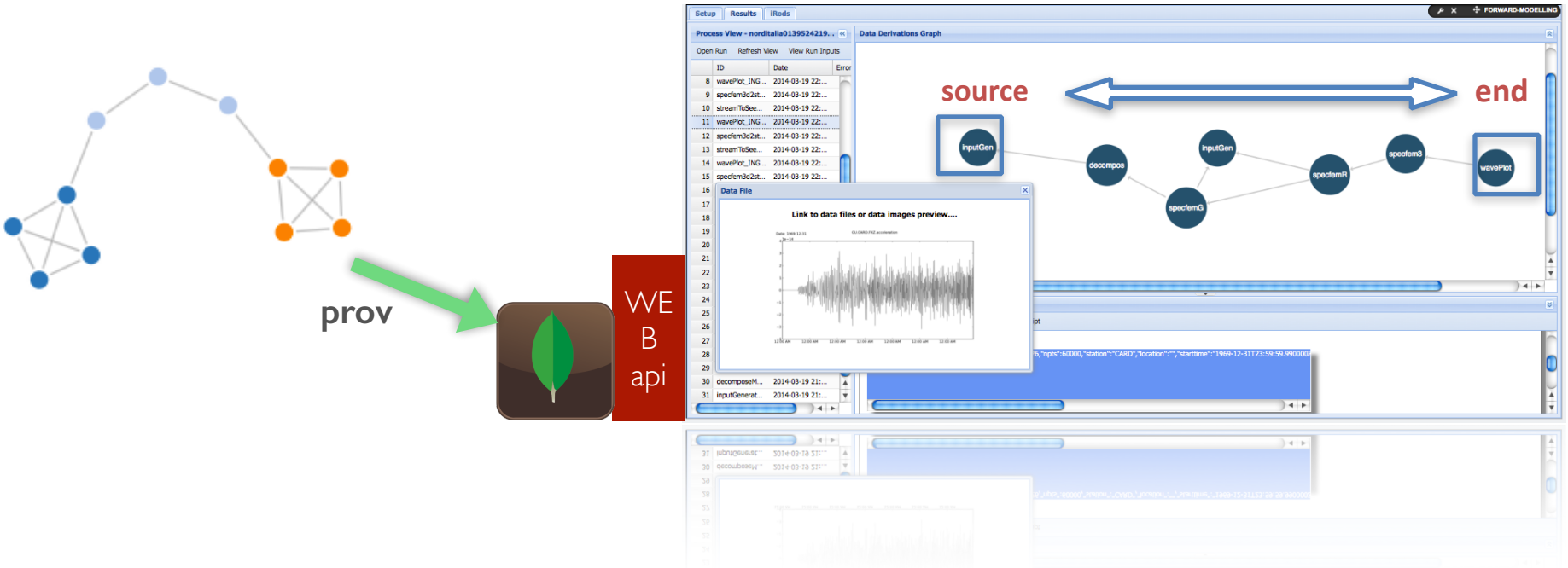
Deployed on local Clouds




prov
WE
B
api
Setup
Results
iRods
FORWARD-MODELLING
Process View - norditalia0139524219...
Open Run
Refresh View
View Run Inputs
ID
Date
Error
Data Derivations Graph
inputGen
decompos
inputGen
specfemR
specfem3
wavePlot
specfemG
Data File
Link to data files or data images preview....



- **Searches** over products metadata within and across runs
- **Data download** and preview
- **Diagnostic** / errors
- **Lineage:** Multi directional navigations across data dependencies
- W3C PROV-DM as reference model.

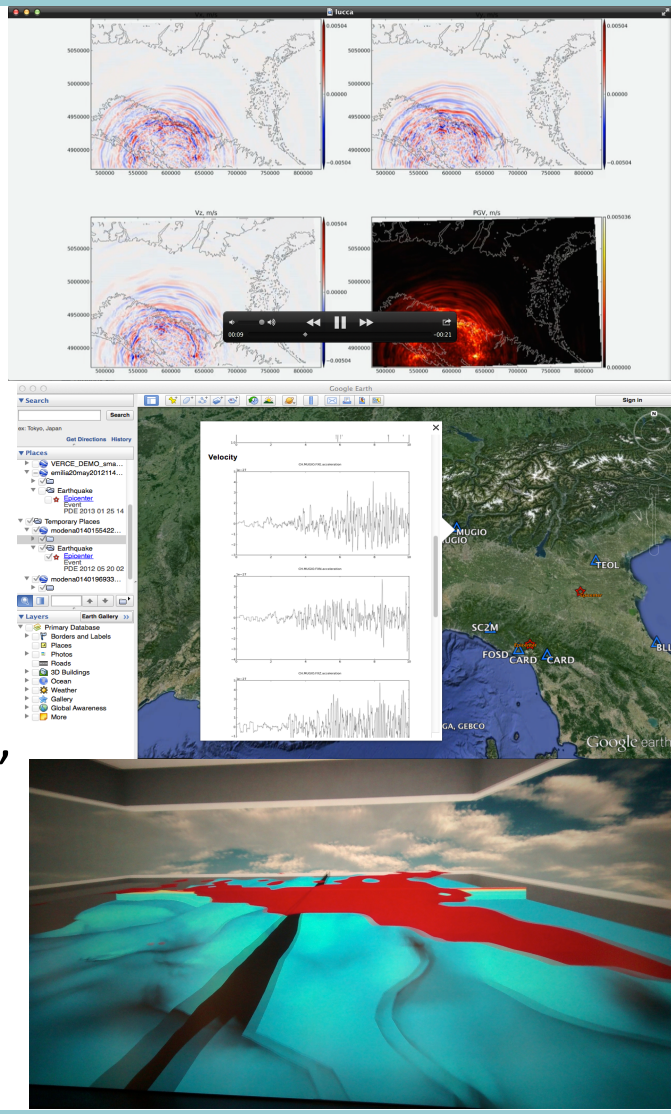


- **Searches** over products metadata within and across runs
- **Data download** and preview
- **Diagnostic** / errors
- **Lineage:** Multi directional navigations across data dependencies
- W3C PROV-DM as reference model.

Earthquake Simulation and Misfit Calculation Data overview:

Synthetic Data: seismograms, plots, 3D Geometry, Videos, KMZ packages, **meshes and models**.
(100 stations = 900 products and metadata)
6-10 GB for a SPECFEM3D simulation on 1000 cores

Raw Data: on demand **access** and **staging** of **observational data** from **EIDA**: Earthquake Metadata, Sensors Metadata, waveform on regional scale.
(At the moment all via the FDSN WEB API)



Products

Earthquake

_event_id: smi:local/ndk/C201302162116A/event
m_tp: -15000000000000000
m_tt: 5420000000000000
m_rr: -31300000000000000
m_rp: 2990000000000000
m_pp: 25900000000000000
m_rt: 8730000000000000
description: SOUTHERN ITALY, C201302162116A
depth_in_km: 21.5
latitude: 41.7
magnitude: 4.95
longitude: 13.51
origin_time: 2013-02-16T21:16:13.300000

Waveform

MY_TERM:VALUE
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
network: IV
longitude: 14.347
latitude: 42.2273
npts: 10000
station: CAFR
starttime: 1969-12-31T23:59:59.990000Z
delta: 0.001
calib: 1
sampling_rate: 1000
endtime: 1970-01-01T00:00:09.989000Z
type: velocity

Processes

name= SPECFEM3D
NPROC=1000
STARTTIME=2015-03-06 16:50:59.654342

Products

Earthquake

_event_id: smi:local/ndk/C201302162116A/event
m_tp: -15000000000000000
m_tt: 5420000000000000
m_rr: -31300000000000000
m_rp: 2990000000000000
m_pp: 25900000000000000
m_rt: 8730000000000000
description: SOUTHERN ITALY, C201302162116A
depth_in_km: 21.5
latitude: 41.7
magnitude: 4.95
longitude: 13.51
origin_time: 2013-02-16T21:16:13.300000

Waveform

MY_TERM:VALUE
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
network: IV
longitude: 14.347
latitude: 42.2273
npts: 10000
station: CAFR
starttime: 1969-12-31T23:59:59.990000Z
delta: 0.001
calib: 1
sampling_rate: 1000
endtime: 1970-01-01T00:00:09.989000Z
type: velocity

Processes

name= SPECFEM3D
NPROC=1000
STARTTIME=2015-03-06 16:50:59.654342

Products

Earthquake

_event_id: smi:local/ndk/C201302162116A/event
m_tp: -15000000000000000
m_tt: 5420000000000000
m_rr: -31300000000000000
m_rp: 2990000000000000
m_pp: 25900000000000000
m_rt: 8730000000000000
description: SOUTHERN ITALY, C201302162116A
depth_in_km: 21.5
latitude: 41.7
magnitude: 4.95
longitude: 13.51
origin_time: 2013-02-16T21:16:13.300000

Waveform

MY_TERM:VALUE
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
network: IV
longitude: 14.347
latitude: 42.2273
npts: 10000
station: CAFR
starttime: 1969-12-31T23:59:59.990000Z
delta: 0.001
calib: 1
sampling_rate: 1000
endtime: 1970-01-01T00:00:09.989000Z
type: velocity

Processes

name= SPECFEM3D
NPROC=1000
STARTTIME=2015-03-06 16:50:59.654342

name= WavePlot
line-color=blue
error=FileNotFound

Products

Earthquake

_event_id: smi:local/ndk/C201302162116A/event
m_tp: -15000000000000000
m_tt: 5420000000000000
m_rr: -31300000000000000
m_rp: 2990000000000000
m_pp: 259000
m_rt: 873000
description: S
depth_in_km
latitude: 41.7
magnitude: 4
longitude: 13
origin_time: 2

Images

format: image/png
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
network: IV
longitude: 14.347
latitude: 42.2273
npts: 10000
station: CAFR
starttime: 1969-12-31T23:59:59.990000Z
delta: 0.001
calib: 1
sampling_rate: 1000
endtime: 1970-01-01T00:00:09.989000Z
type: velocity

Waveform

MY_TERM:VALUE
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
network: IV
longitude: 14.347
de: 42.2273
10000
: CAFR
me: 1969-12-31T23:59:59.990000Z
0.001
1
ng_rate: 1000
e: 1970-01-01T00:00:09.989000Z
elocity

Processes

name= SPECFEM3D
NPROC=1000
STARTTIME=2015-03-06 16:50:59.654342

name= WavePlot
line-color=blue
error=FileNotFound

Products

Earthquake

_event_id: smi:local/ndk/C201302162116A/event
m_tp: -15000000000000000
m_tt: 5420000000000000
m_rr: -31300000000000000
m_rp: 29900000000000000
m_pp: 259000
m_rt: 873000
description: S
depth_in_km
latitude: 41.7
magnitude: 4
longitude: 13
origin_time: 2

Waveform

MX_TERM:VALUE
-bcaec52d20d4
59.990000Z
e: 1970-01-01T00:00:09.989000Z
elocity

Video

format: video/mpg
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
Duration: 30s
Sampling-rate: 25fps

Images

format: image/png
id: b9a3cbcc-060a-11e
network: IV
longitude: 14.347
latitude: 42.2273
npts: 10000
station: CAFR
starttime: 1969-12-31T23:59:59.990000Z
delta: 0.001
calib: 1
sampling_rate: 1000
endtime: 1970-01-01T00:00:09.989000Z
type: velocity

Processes

name= SPECFEM3D
NPROC=1000
STARTTIME=2015-03-06 16:50:59.654342

name= WavePlot
line-color=blue
error=FileNotFound

Products

Earthquake

_event_id: smi:local/ndk/C201302162116A/event
m_tp: -15000000000000000
m_tt: 5420000000000000
m_rr: -31300000000000000
m_rp: 2990000000000000
m_pp: 259000
m_rt: 873000
description: S
depth_in_km
latitude: 41.7
magnitude: 4
longitude: 13
origin_time: 2

Waveform

MX_TERM:VALUE
-bcaec52d20d4

Video

format: video/mpg
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
Duration: 30s
Sampling-rate: 25fp

Images

format: image/png
id: b9a3cbcc-060a-11e
network: IV
longitude: 14.347
latitude: 42.2273
npts: 10000
station: CAFR
starttime: 1969-12-31T23:59:59.990000Z
delta: 0.001
calib: 1
sampling_rate: 1000
endtime: 1970-01-01T00:00:09.989000Z
type: velocity

Mesh

format: binary
id: b9a3cbcc-060a-11e4-96c2-bcaec52d20d4
maxlat: 66.55
minlat:14.5
maxlon: 33
minlong:-12

Processes

name= SPECFEM3D
NPROC=1000
STARTTIME=2015-03-06 16:50:59.654342

name= WavePlot
line-color=blue
error=FileNotFound

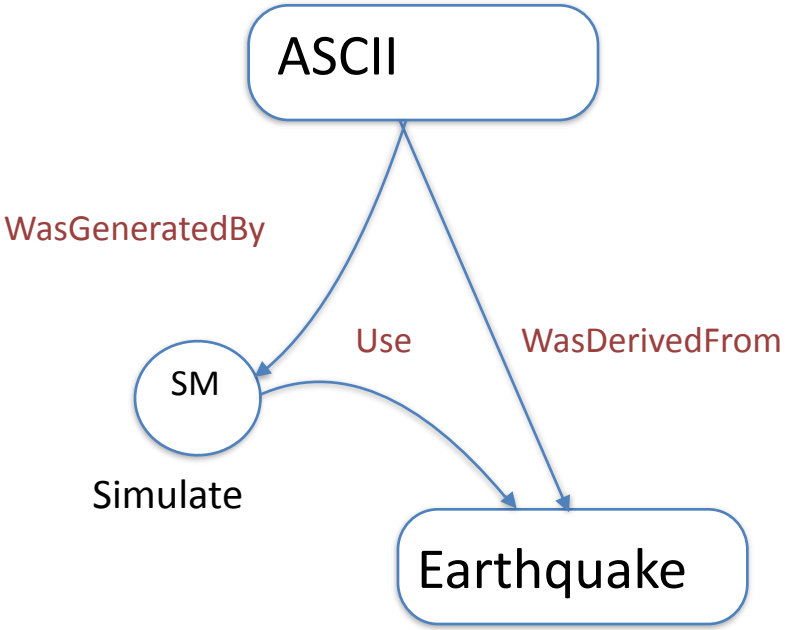
W3C PROV-O offers a generic data model describing relations between data products and processes (provenance)

W3C PROV-O offers a generic data model describing relations between data products and processes (provenance)

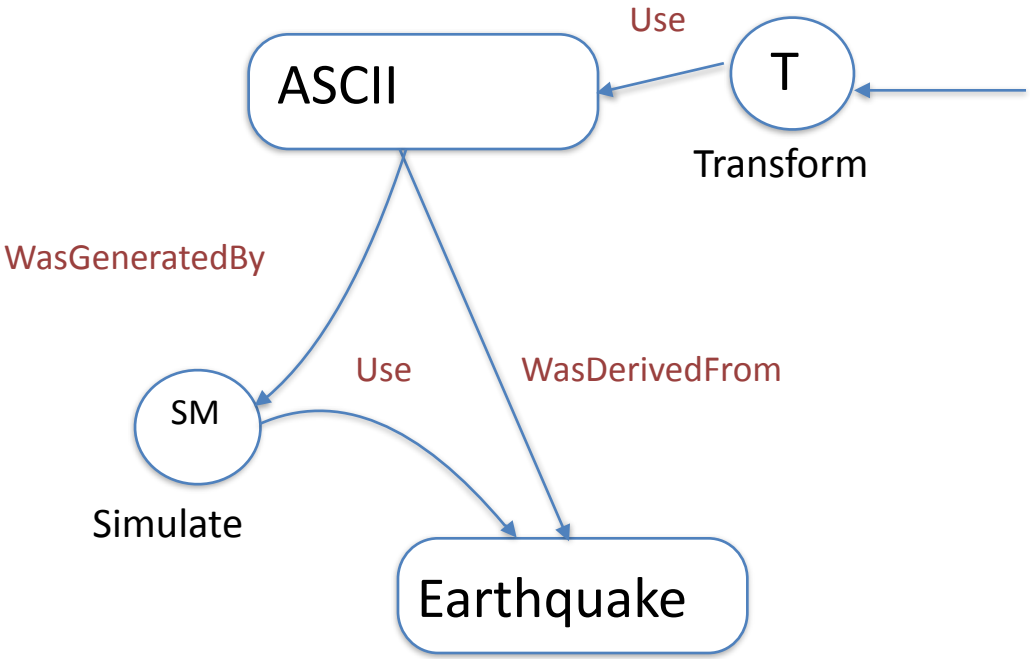
W3C PROV-O offers a generic data model describing relations between data products and processes (provenance)

Earthquake

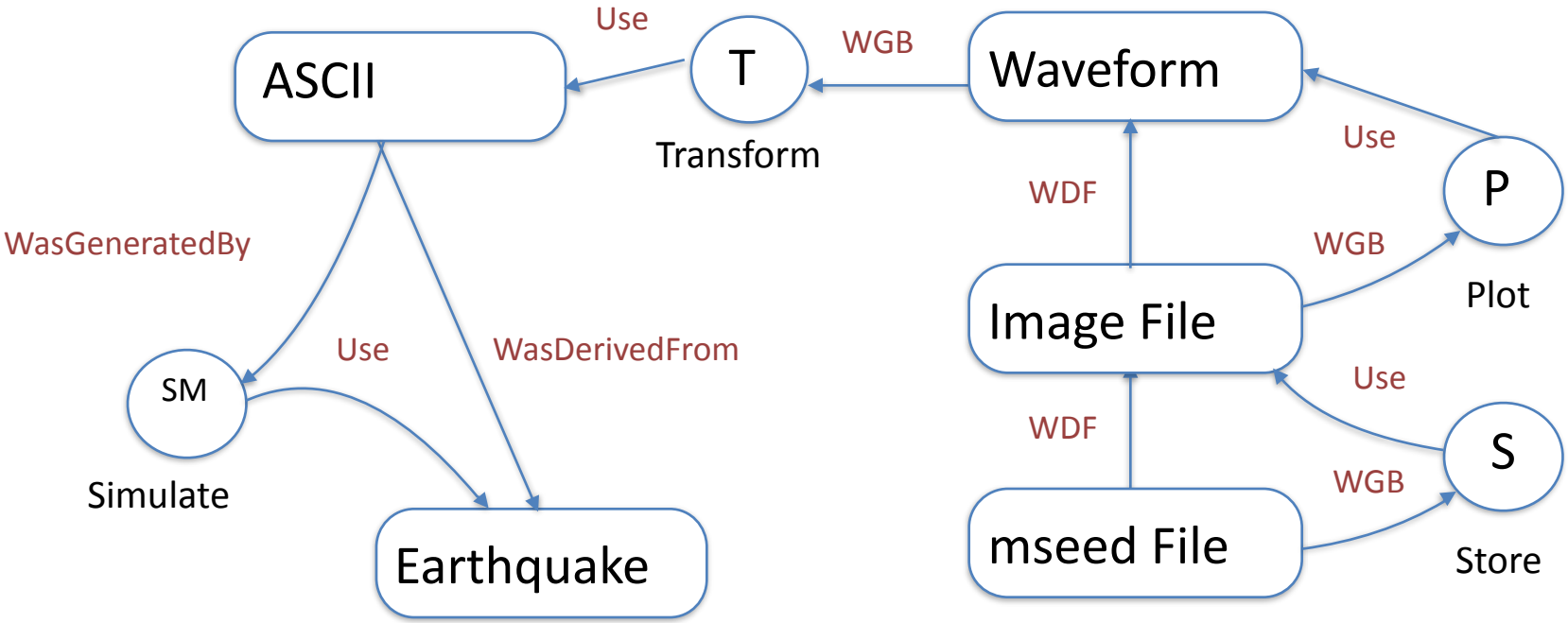
W3C PROV-O offers a generic data model describing relations between data products and processes (provenance)

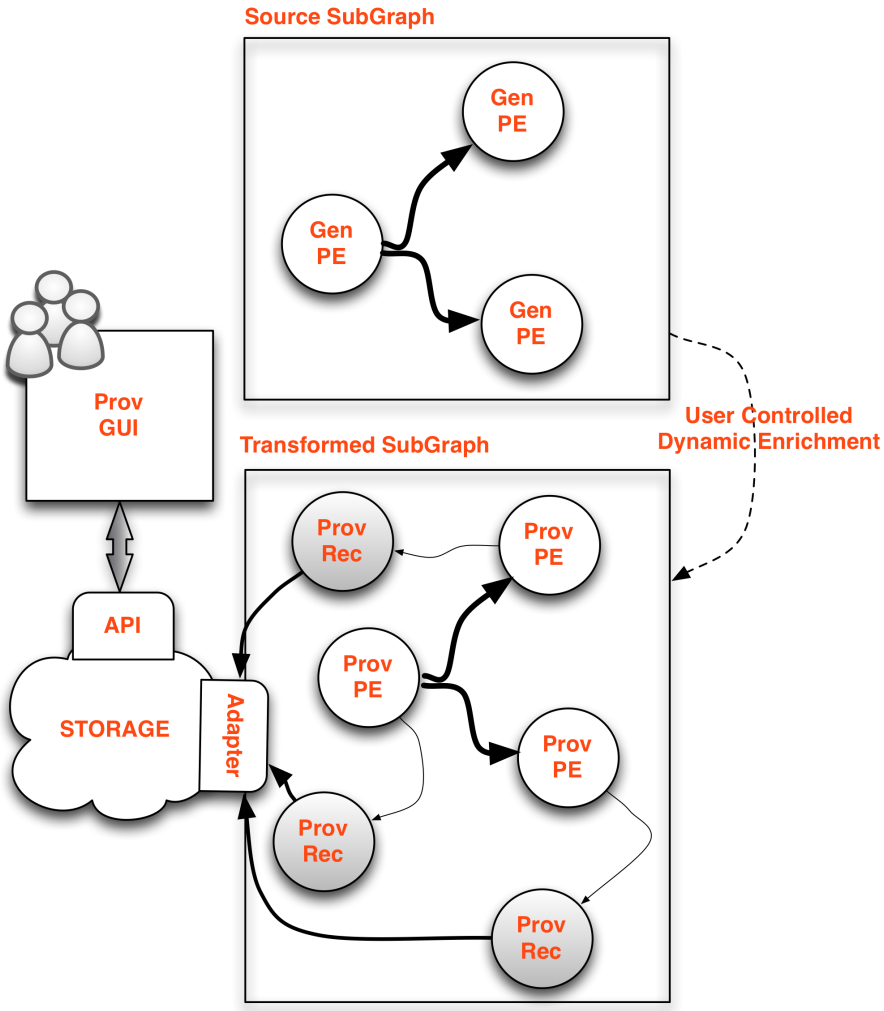


W3C PROV-O offers a generic data model describing relations between data products and processes (provenance)



W3C PROV-O offers a generic data model describing relations between data products and processes (provenance)





Selective and extensible Provenance

The ***GenericPEs*** of a workflow subgraph are extended at runtime assuming the ***ProvenancePE*** type *(Dynamic Polymorphism)*

It gets connected to instances of **ProvenanceRecorderPE** for rapid analysis and transfer of the provenance traces to a dedicated catalogue.

The provenance data is then explored via a **GUI**, which connects to a web service interface.

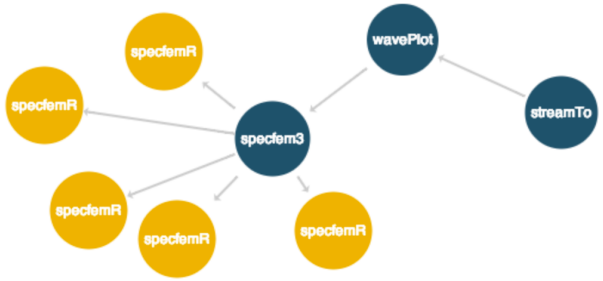
Selective and extensible Provenance

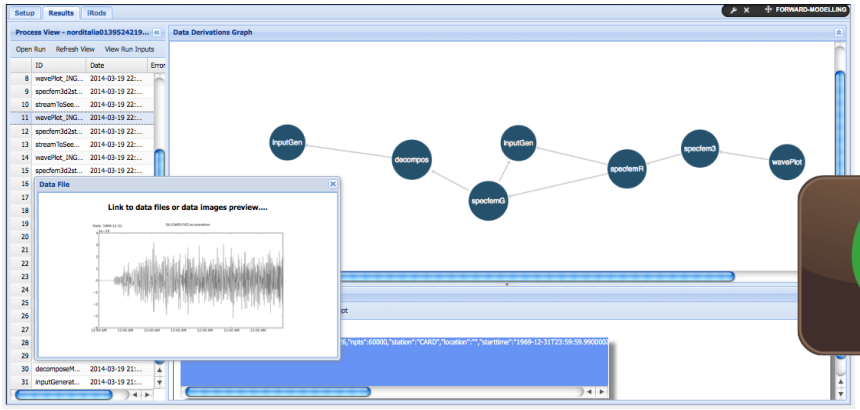
The ***GenericPEs*** of a workflow subgraph are extended at runtime assuming the ***ProvenancePE*** type (Dynamic Polymorphism)

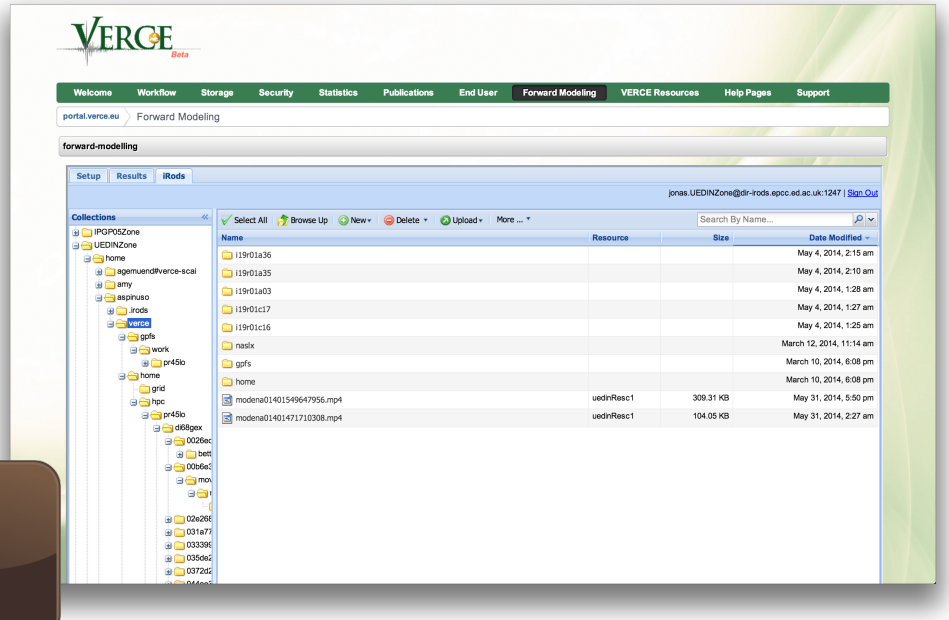
***ProvenancePE extended functions*:**

- ***Write(port, data, metadata, format, errors, control)***
- ***ExtractItemMetadata(data)***

port: output port which is connect to the adjacent PEs.
data: stream element to send to those adjacent PEs.
metadata: dictionary of metadata describing the data.
format: typically contains the mime-type of the data.
errors: erroneous situations and its description
control: are the control instructions like ***"con:skip"*** and ***"con:immediateAccess"***







Browse provenance and workflows output

Access Data Products shipped from clusters to iRODS (e.g. **“con:immediateAccess”**)

(Integrated Rule-Oriented Data System)

Why ProvenanceRecorderPEs?

For a comprehensive provenance framework across DI-HPC models

1

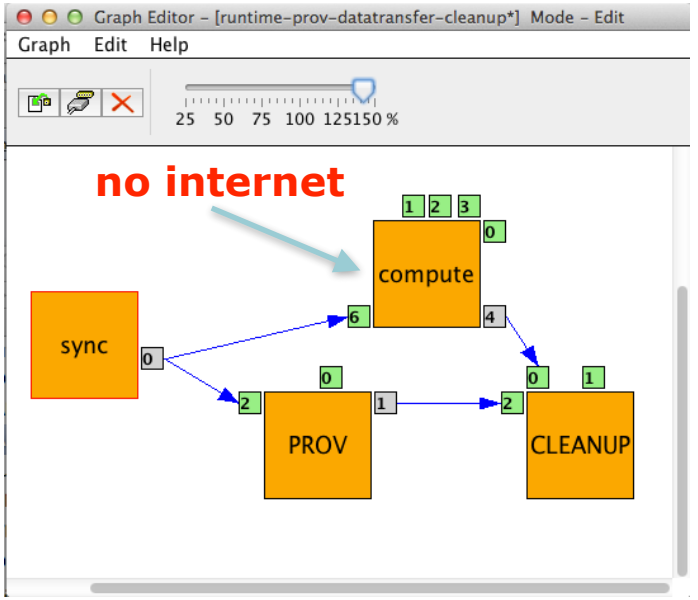
Platform Control Workflow:

sync: metadata preparation and staging

compute: actual computation

prov: reads metadata, updates prov repository, intermediate data stage-out based on prov (**con:ImmediateAccess**)

cleanup: full data stageout and cleanups



Why ProvenanceRecorderPEs?

For a comprehensive provenance framework across DI-HPC models

1

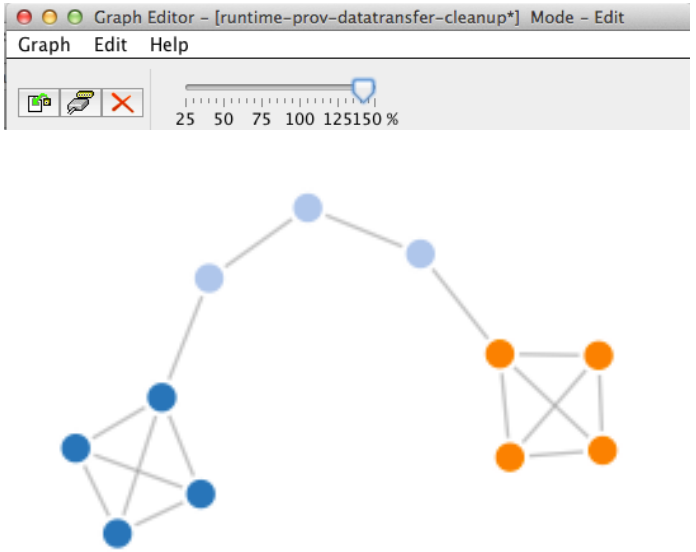
Platform Control Workflow:

sync: metadata preparation and staging

compute: actual computation

prov: reads metadata, updates prov repository, intermediate data stage-out based on prov (**con:ImmediateAccess**)

cleanup: full data stageout and cleanups



compute receives and runs scientific workflows

2

Science Case Workflow (According to community preferences)

Extraction of **user defined metadata** and **fine-grained Provenance**

Use of *con:skip* for Selective provenance extraction

For a comprehensive provenance framework across DI-HPC models

E.g. Large Streaming Cross-Correlations: Data recorded by sensors are preprocessed in pipelines and cross correlated: *n(n-1) / 2 correlations*
Full Provenance coverage would impact for 7-8% on our tests on ~350 sensors

con:skip **enables provenance extraction and recording only for a subset of stations and for a specific set of metadata.**

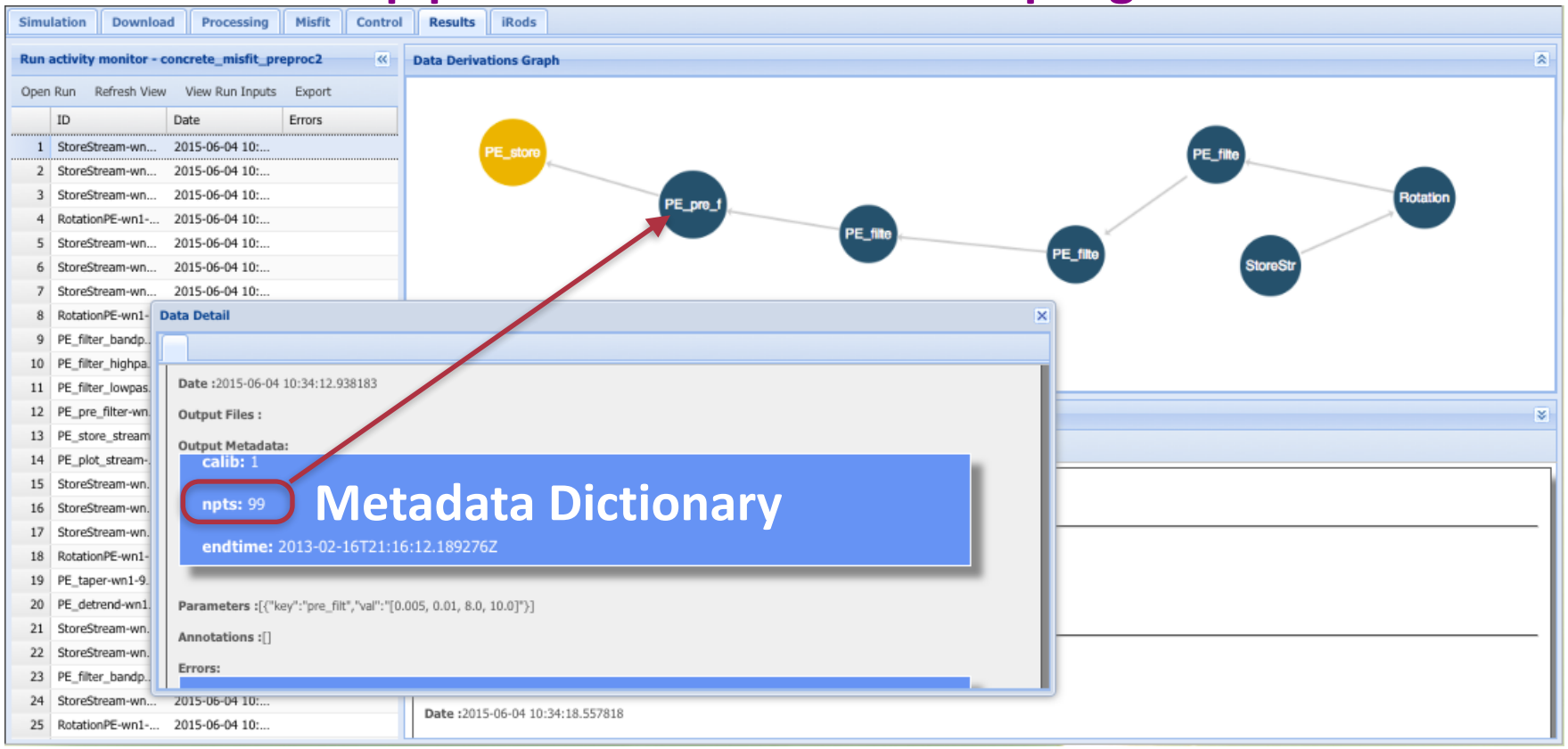
```
if stream.name not in list4prov:
    self.write('output',stream,metadata=dic,control="con:skip")
```

Note: When using custom implementations of a ProvenancePE's function ExtractItemMetadata(data), the metadata in *dic* will be added to the default.

Use of Provenance for Diagnostics and Validation

For a comprehensive provenance framework across DI-HPC models

Which PE in the pipeline is affection the sampling rate?



Use of Provenance for Interactive Workflow Preparation

For a comprehensive provenance framework across DI-HPC models

Simulations Runs

Simulation runs

runId	Workflow name	Description	Date
foggia01410428...	PRACE-LRZ-DIS...	foggia small	2014-09-11T09:...
foggia01410426...	PRACE-LRZ-DIS...	foggia con video	2014-09-11T09:...
modena014104...	PRACE-LRZ-DIS...	modena	2014-09-11T08:...
abruzzo100001...	PRACE-LRZ-DIS...	d4py	2014-09-11T08:...
abruzzo014104...	PRACE-LRZ-DIS...	d4py	2014-09-11T08:...
abruzzo014103...	PRACE-LRZ-DIS...	d4py	2014-09-10T12:...
modena014103...	PRACE-LRZ-DIS...	modena	2014-09-10T12:...

Observed Data Pre-staging Runs

raw-data download runs

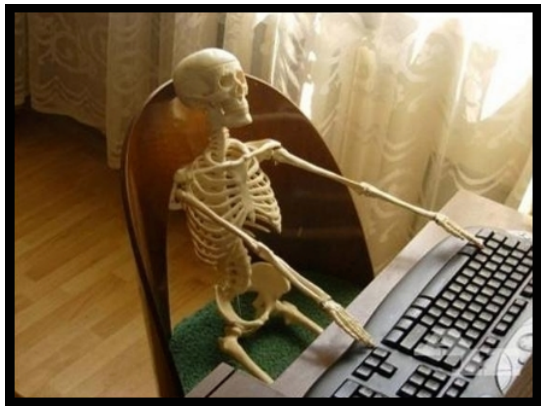
runId	Workflow name	Description	Date
test_download5	download	test	2015-05-28 13:...
test_download2	download	test	2015-05-22 12:...
data_download_8	download	test	2015-05-21 13:...
data_download_5	download	test	2015-05-21 13:...
data_download...	download	test	2015-05-06 14:...
data_download...	download	test	2015-05-06 14:...
data_download...	download	test	2015-04-02 08:...

commons stations

☐	IV	TRIV
☑	**IV**	**RNI2**
☐	IV	T0204
☐	IV	T0201
☐	IV	
☐	IV	SIRI
☐	IV	SLCN
☐	IV	
☐	IV	SOR
☑	**IV**	**VVLD**
☐	IV	T1011
☐	IV	VITU
☐	IV	VULT
☑	**IV**	**SGG**

Sensors' data with common properties for MISFIT analysis (time-range, event)

Runtime and Selective Provenance in Data Intensive Platforms



- Provides immediate feedback on the produced data with sensible and tuneable metadata.
- Useless waits for long and unfruitful runs can be reduce
- Fosters Dynamic Steering, Failure Recovery and Errors
- Detection, saving computing cycles, storage ($$) and energy!…

References

P. Missier, S. Dey, K. Belhajjame, V. Cuevas-Vicenttin, and B. Ludascher. D-prov: extending the prov provenance model with work ow structure. In Proceedings of the 5th USENIX Workshop on the Theory and Practice of Provenance, TaPP '13, pages 9:1-9:7, Berkeley, CA, USA, 2013. USENIX Association.

M. Anand and S. Bowers. Exploring scientific workflow provenance using hybrid queries over nested data and lineage graphs. Scientific and Statistical Database Management, pages 237-254, 2009.

P. Missier and B. Ludascher. Linking multiple workflow provenance traces for interoperable collaborative science. Workf;ows in Support of Large-Scale Science (WORKS), 2010 5th Workshop on, 2010.

S. Madougou and S. Shahand. Characterizing workflow-based activity on a production e-infrastructure using provenance data. Future Generation Computer Systems, (Dci), 2013.

D. Crawl and I. Altintas. A provenance-based fault tolerance mechanism for scientific workflows. Provenance and Annotation of Data and Processes, pages 152-159, 2008.

P. Buneman, J. Cheney, and E.V. Kostylev. Hierarchical models of provenance. In Proceedings of the 4th USENIX conference on Theory and Practice of Provenance, TaPP'12, pages 10–10, Berkeley, CA, USA, 2012. USENIX Association.

E. Griffis, P. Martin, and J. Cheney. Semantics and provenance for pro- cessing element composition in dispel workflows. WORKS '13 Proceedings of the 8th Workshop on Workflows in Support of Large-Scale Science, 2013.

A. Spinuso, J. Cheney, and M. Atkinson. Provenance for seismological processing pipelines in a distributed streaming workflow. Proceedings of the Joint EDBT/ICDT 2013 Workshops, 2013.