Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

Xavier Briand

Bruno Bzeznik

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Use Case of Data-Intensive processing with iRODS

Collaboration between:

IT part of Whisper:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

➠ **Sofware development, computation (Xavier Briand)**

Platform Ciment:

CIMENT

CIGRI

➠ **IT infrastructure of the University of Grenoble**
➠ **High Performance Computing HPC**
➠ **Data-Grid: grid manager CiGRi and iRODS storage (Bruno Bzeznik)**

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Whisper: a European seismological project

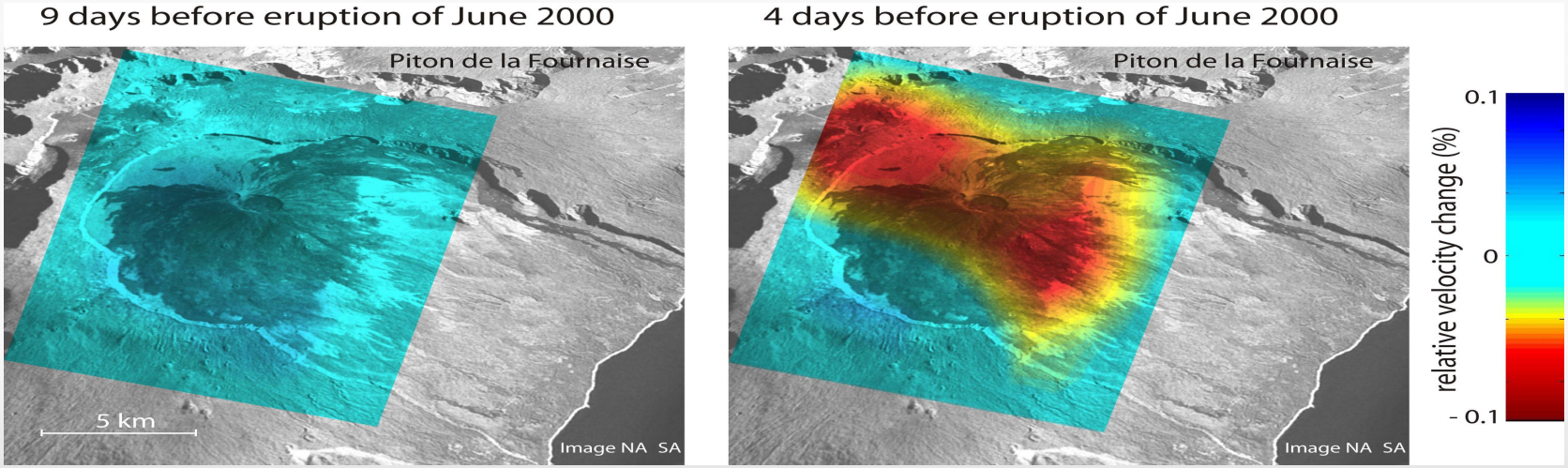
European Research Council

Main **Goal**:

- Detect sligth changes of properties in the solid Earth
- Kind of ”Datamining” of seismograms

A **Motivation**:

- Variation of wave speed before eruption (La Reunion 2000)



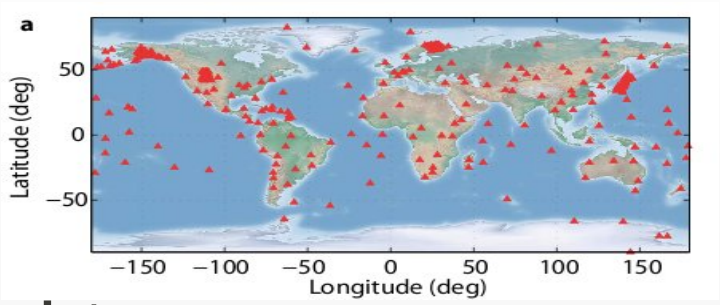
Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Main constraint for Whisper: Massive data processing

Data provenance:

- Noise Continuously recorded by seismic stations worldwilde.
  Europe: Alpes
  China: LMS
  USA: USArray (mobile Network)
  Japan Networks (NIED)
  ...
- The computations produce even more data ...



How much data are we talking about?

Japanese Network: 20 TB per year
Typical processing: 8 TB in 5 days
Another: 'read' 3 TB and 'write' 1 TB in 6 hours
Many processings have to be tested ...
200 TB managed at the same time

Main operations: Raw data processing and computation of the correlations

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Other constraints for Whisper:

Time:

Organize a scientific workflow
Duration PHD and Posdoc
Available IT Infrastrucure
Ease of access for researchers

➠ Implementing a generic model, technical support

Specific tools:

Organize distributed data
Access to computational ressources

➠ Need a Data Grid environment:

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Plan

1- Presentation of the IT infrastructure Ciment:
Grid computing with CiGri and iRDOS

2- Presentation of the Whisper Code:
Software for data intensive processing

3- Results and feedback:
Whisper Use Case
iRODS experience in the context of Ciment

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

IT infrastructure of the Ciment platform

Platform: HPC centre of the University of Grenoble
Partial pooling of computing resources
- ➠ 10 computing clusters, 6600 cores + GPUS

Local data grid environement: Resources in a local grid of supercomputers

Distributed storage storage 700 TB
- ➠ Centralized controller with total observation: monitor computation

Grid manager CiGri (OAR project)
- ➠ Short parametric jobs
- ➠ Best effort mode
- ➠ Improves resource usage

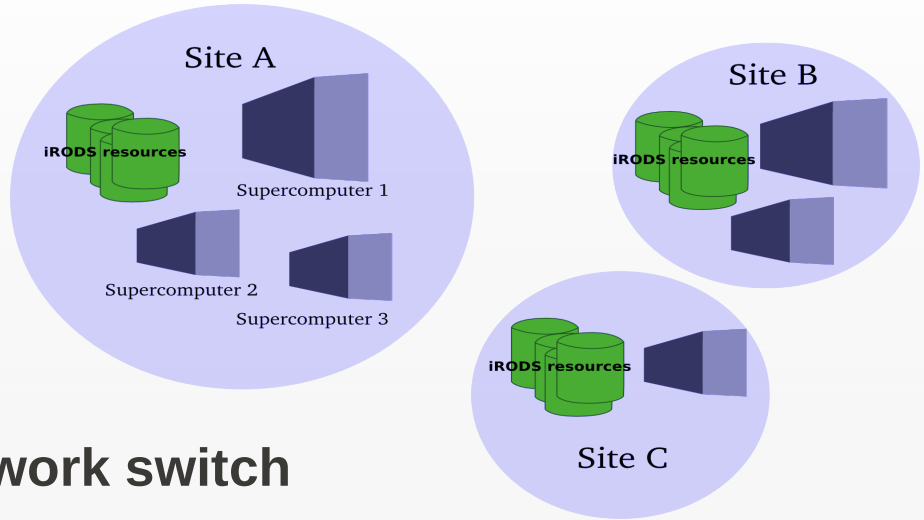
Whisper case: most perfectly (aka embarrassingly) parrallel but data intensive processing

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

IRODS infrastructure

3 sites close to supercomputers:



An unique zone with the iCat server
Dozen of Irods nodes
Heterogenous WAN connexions
Each site has its own 10Gbe local network switch

Site specification:

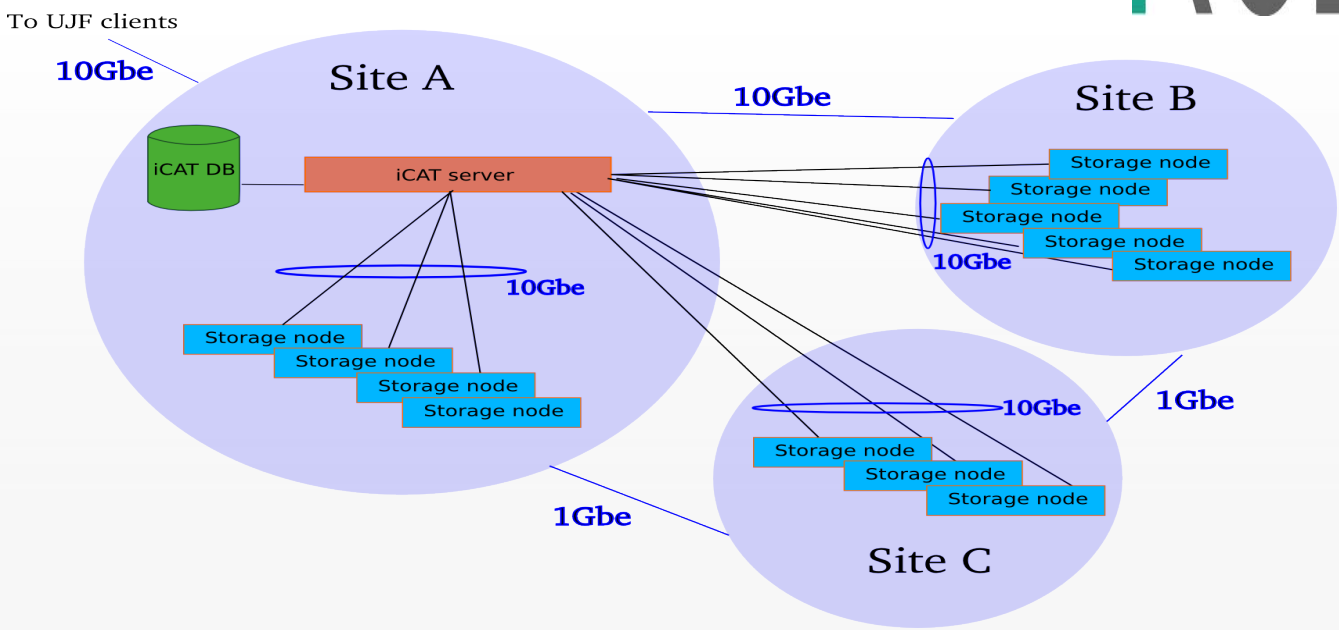
Each site has its own iRODS resource group
Store results of computation randomly at the **same site**
➠ Use of LAN and not WAN

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

IRODS infrastructure

More precisely:



Site A and B have 10Gbe WAN connexion
Site C has only 1Gbe WAN

Automatic staging for site C

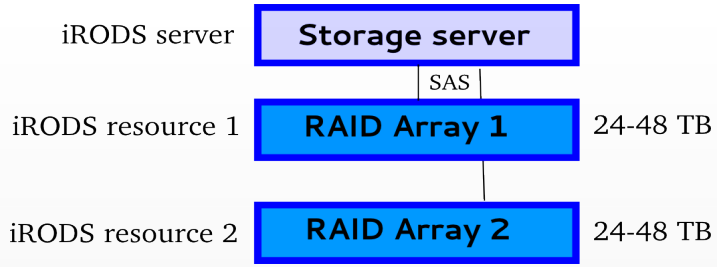
- **Automatic replication of resoures from sites A and B to site C when data get from site C**

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

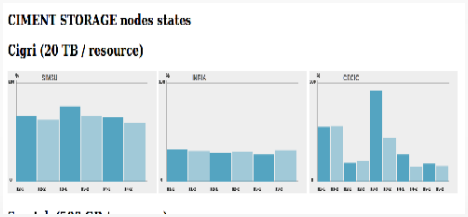
IRODS infrastructure

Focus on storage nodes:



Two RAID arrays 24 TB – 48 TB
Irods Nodes with Debian GNU/Linux
(Easy synchronisation of the system admin with Kanif)

Ciment provides a web interface to check resource status



Now 700 TB, constant evolution
Irods offers great scalability

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Cigri middleware

Cigri achieves access to all 6600 cores of 10 clusters within Ciment.
Perfectly parallel jobs on idle processors

Principles
Each cluster uses the ressource manager OAR
Cigri communicates with the cluster through OAR's RESTful API
Cigri submits jobs without exhausting local queues

Best effort mode
Jobs on idle processors of every computing nodes.
➠ **Improves computational ressources**
Jobs have lowest priority but automatic resubmission
(CiGri works also in normal mode)

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Cigri middleware

Run a set of jobs: a campaign

User defines the campaign parameters in a JSON file
For each cluster (or all clusters):
- Accepted clusters, ressources needed
- Maximum duration
- Maximum number of jobs
- Location of the code
- Prologue, epilogue

And iRODS?

Code and input are retrieved from iRODS with i-commands (or API)
No direct connexion between iRODS and CiGri
But completely complementary through job scripts.

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Cigri middleware

Technologies, functionalities

RestFul API, Ruby, around a PostgreSQL database.
Apache with SSL authentication

Statistics: execution rate, resubmission rate
Email notification of failed jobs
Retrieve the jobs' standard output and standard error from CiGri RestFul API.

Authentication, security

A centralized LDAP infrastructure for Ciment
Password method with synchronization with LDAP authentication.

Transparent for users: automatic initialization in home directory
(.irods/.irodsEnv and .irods/.irodsA files)

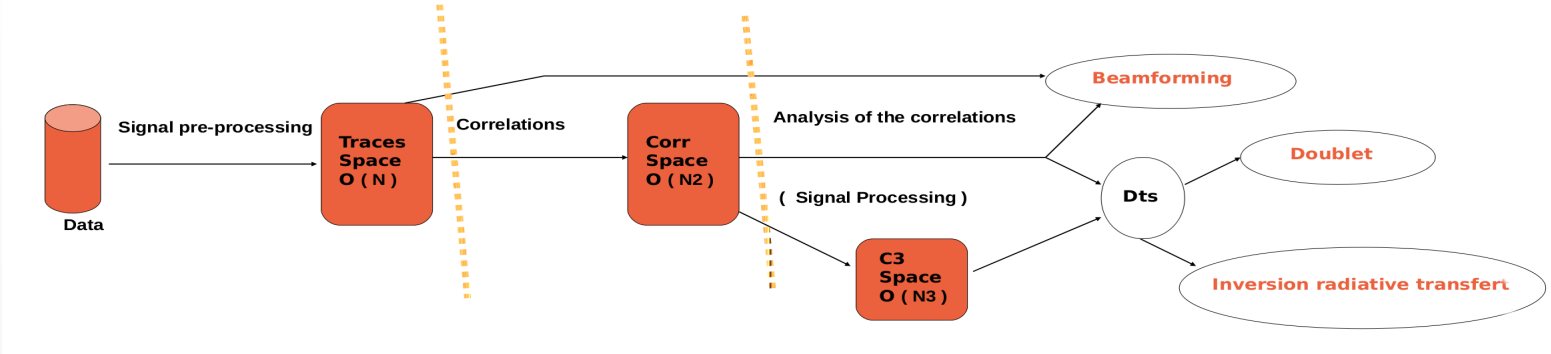
Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Whisper Code

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

General design



Grouped into 3 parts:
- **Seismograms (arrays) processing**
- **Computation of correlations (couple of seismograms)**
- **Analysis of the correlations (differents methods)**

Python language, libraries Scipy and Obspy (seismology)

Development driven by performance, evolution and required support.

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Whisper Code

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

Raw data (seismograms) processings

A flexible way of specifying a pipeline of processing

User specifies:
- A directory
- A set of seismic stations
- A set of dates
- A sequence of treatments

Users can enter their own code into the processing

Code:
- Scan a directory
- Extract all pieces of seismograms
- Organize a specific architecture of files of daily seismograms

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

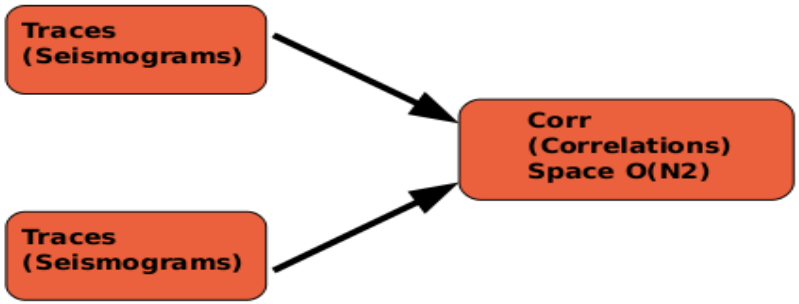
2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Whisper Code

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

Computation of the correlations

Operation with 2 seismograms
that provides a new virtual seismogram
➠ Get 'virtual' new observational data



Architecture of files that corresponds to all the couple of seismograms.
➠ Quadratic space complexity (linear for seismogram processing)
➠ Need to store seismograms processing

Quadratic space complexity can be critical ➠ Optimization

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Whisper Code

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

Focus on optimization of the computation of correlations

Numerical optimization:

- Algorithm uses the fast fourier transform
- Pre-calculating 'Good' combinations of small prime numbers
- 40% improvement for good cases

Main optimization: try to catch the cache

- Test the GC behavior in order to follow the cache heuristics
- Don't use the Garbage Collector directly!
- Write code towards 'optimal' use of the cache.. (and find good unfolding)

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

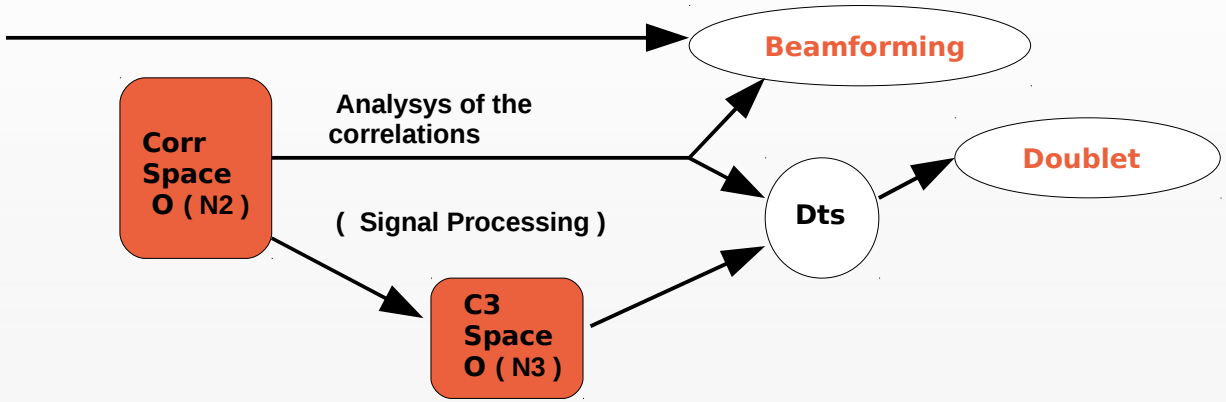
Whisper Code

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

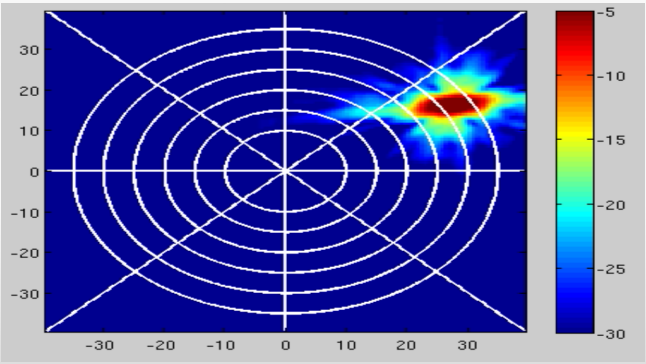
Last part of Code: Analysis of correlations

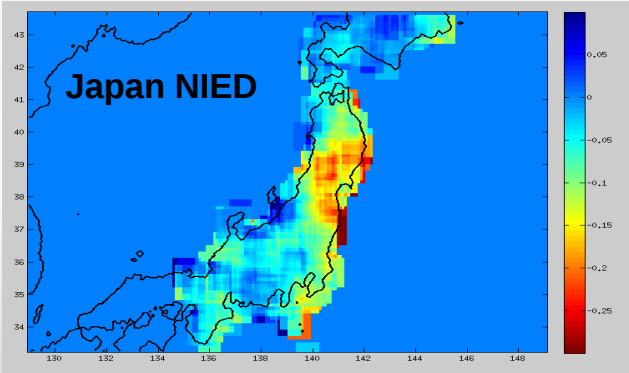
Further methods

- Beamforming
- Doublet
- Inversion



➠ New filter the signal, computation of the variation velocity of waves
(Note C3 is cubic in Space: distribution for quadratic space complexity)





Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Results:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

CIMENT

Organization of a computation

Either local computation with a dedicated bay or grid computation

- Retrieve a dataset
- Anticipate computation time
- Evaluate required storage capacity
- Assess users' computer skills

The specification must evolve

- New requirements from researchers
- Not generic enough …
- Very complex to evaluate

Adaptation to the grid environment: development time?

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Results:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

CIMENT

Whisper computation with Ciment data grid

First step: convert data into seismic standard formats
➠ **Can be a very data intensive processing**

Minimize concurrency on Irods
➠ **IT code for randomly spreading files on resources**

Find a 'good' distribution model

Fine tuning distribution of computation and transfer

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Results:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

CIMENT

Seismogram processing on the grid:
Ouput of the Japanese Network over one year

Seismograms processing: simplest case for data grid process
Convert 9 TB into 20 TB of seismic standard format
Modeling the distribution: subsets of dates and subsets of stations

Adaptation for iRODS
Add python modules to retrieve data on iRODS:

Provide encapsulation of the iget and iput: number of try, waiting time, 'else' command

```
...
get parameter
...
build the iget commands
perform iget (encapsulation)
...
Codes whisper
for seismogram processing
....
build the iput commands
perform iput (encapsulation)
...
```

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Results:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

CIMENT

Adaptation for Cigri: Define a file of parameters: each line corresponds to a job

Campaign with 1280 jobs
160 sublists of dates and 8 sublists of stations

```
cat param.txt

traces160_0_8_0  160 0 8 0
traces160_0_8_1  160 0 8 1
...
traces160_1_8_0  160 1 8 0
traces160_1_8_1  160 1 8 1
...
```

'traces160_0_8_1 160 0 8 1'
➠ Sublist of dates of index 0 and sublist of stations of index 1

Define the paramter of the campaign (jdl file JSON)
- Campaign named 'test_processing'
- Run the 'start.bash' with 'param.txt'
- Use clusters 'c1' and 'c2'
- Best effort mode
- Campaign associated to 'whisper'
- Maximum number vof jobs
- Retrieve 'start.bash' script from iRODS

...

```
cat processingSeismogram.jdl

{
  "name": "test_processing",
  "resources": "core=1",
  "exec_file": "$HOME/start.bash",
  "param_file": "param.txt",
  "type": "best-effort",

  "clusters": {
    "c1": {
        "prologue": [
        secure_iget -f /IrodsColl/start.bash,
        secure_iget -f /IrodsColl/param.txt,
        ... other lines of commands ],
        "project": "whisper",
        "walltime": "00:20:00"
        },
    "c2": {
        "prologue": [
        ... lines of commands ],
        "project": "whisper",
        "max_jobs": "450",
        "walltime": "00:30:00"
        },
    ...
    }
...
}
```

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Results:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

CIMENT

Correlation computation on the grid:

Ouput of the Japanese Network over one year

Add similar modules to retrieve data on Irods
Two types of processes: with one sublist and with two disjoint sublists
➠ **Transfer is proportional to the distribution for the stations**

Maximize the distribution of dates
Minimize the distribution of stations

Fine tuning the size of files transferred between iRODS and computing nodes:

- **Correlations into dictionary**
- **Size file between 100 MB and 500 MB**
- **Appropriate for iRODS infrastructure**

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Results:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

CIMENT

One year of the Japan Network on the grid
Processing all seismograms now takes half a day (down from 4 months)

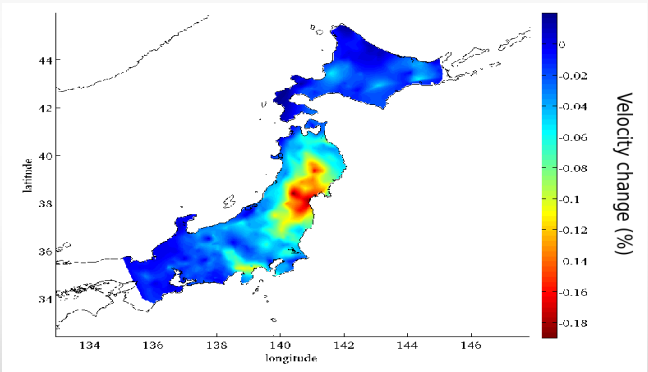
Computation of 350 Million correlations for the Japanese Network (down from 2 years ...)

- **Between 9h and 20h (It depends on stacking, overlap)**
- **'iget' command correponds to 11 TB**
- **'iput' command correspons to 3.3 TB**
- **Best-effort can increase also transfer...**

➠ **Big change in opportunities to test processing**

Many scientific results

About variation of velocity change for the tohoku earthquake in Japan: in Science, June 2014



Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Results:

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

CIMENT

iRODS experience in the context of Ciment

Opportunity to test and improve Ciment Infrastructure for Big Data

Add new functionalities:

- **Limit the number of jobs avoid overload iRODS Infrastructure**
- **Limit the number of simultaneous connexions**
- **Develop wrapper for i-commands: secure_iget …**

For iRODS: unique namespace over distributed sites

➠ **Only the meta Catalog can block the system (asynchronous)**

Simply add a server as an iRODS resource and define access policy for each project

Challenge with small files (< 32 MB)
Completly overload 1Gbe link between 2 sites for conversion processing during automatic staging.

Data Intensive processing with iRODS and the middleware CiGri for the Whisper project

2014 06
Irods User Meeting
Harvard University
Xavier Briand
Bruno Bzeznik

Thanks so much!

Whisper Seismic ambient noise.
Towards continuous monitoring of the continuously changing Earth.

http://whisper.obs.ujf-grenoble.fr

https://ciment.ujf-grenoble.fr

http://cigri.imag.fr

http://irods.org

http://oar.imag.fr

kanif

http://taktuk.gforge.inria.fr/kanif

http://www.liglab.fr