edinburgh
data-intensive
research

EFFORT

Earthquake and Failure Forecasting in Real Time from controlled laboratory test to volcanoes and earthquakes

Rosa Filgueira Vicente.
University of Edinburgh.
Data Intensive Research.

EFFORT

Edinburgh Geosciences: Ian Main, Andrew Bell.

Edinburgh Informatics: Malcolm Atkinson, Rosa Filgueira.

UCL Rock Physics: Steve Boon, Chris Kilburn, Phil Meredith.

Index

Introduction

- Brittle failure of rock samples in the laboratory is analogous to brittle failure associated with volcanic eruptions, earthquakes …

- Signals observed in the laboratory, could be used to forecast the timing of hazard events.

EFFORT goals

- Determine the **predictability** of brittle failure of rock samples in the **laboratory experiments.**
- Determine how this **predictability** scales to the greater complexity, physical size, and slower strain-rates of **natural-world** phenomena.
- The project will **develop methodologies based on archive data and then apply them in "near real-time"** to a variety of synthetic, experimental and natural data.

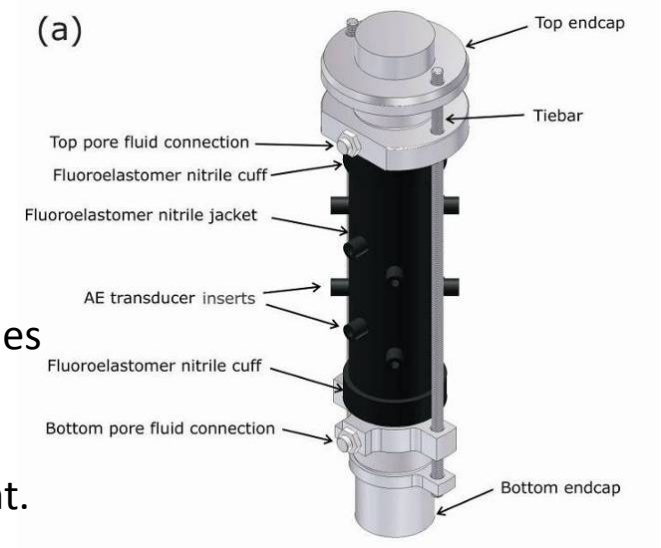
Types of Experiments

- Controlled Laboratory Experiments or Creep 2 experiments:
  - Data from the NERC funded **UCL/Edinburgh Creep2 project** which will undertake a series of low strain-rate brittle creep experiments in a deep-sea laboratory off the coast of Sicily.
  - The deep-sea equipment is currently under development with a likely first deployment in early summer 2013.
  - Traditional laboratory brittle creep experiments will also be performed in the Rock physics labs at UCL as part of Creep2.

Types of Experiments

- Natural Experiments:
  - Work with data collated from real-world brittle failure phenomena.
  - Volcanic Data : Agreement with the **Icelandic Meteorological Office** to use their volcanic earthquake data. Much of this data is already published daily on observatory websites, and is readily harvestable.

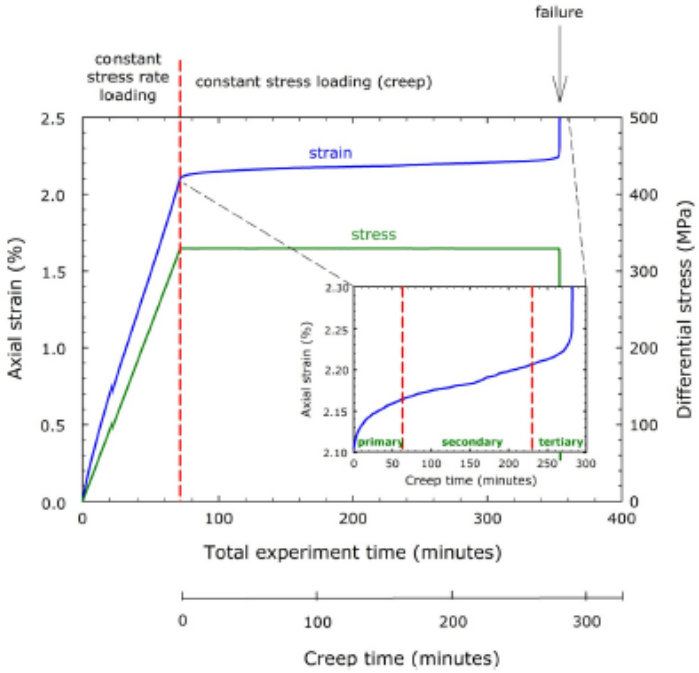
Laboratory creep experiments



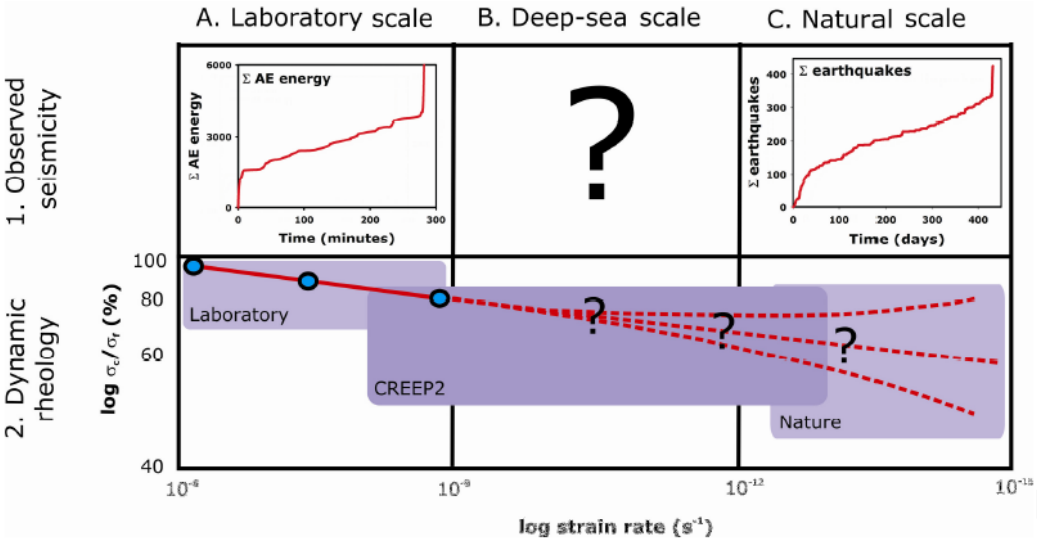
1. Address the time dependent sub-critical deformation that precedes dynamic failure.

2. Measure short term peak stress in constant strain rate experiment.

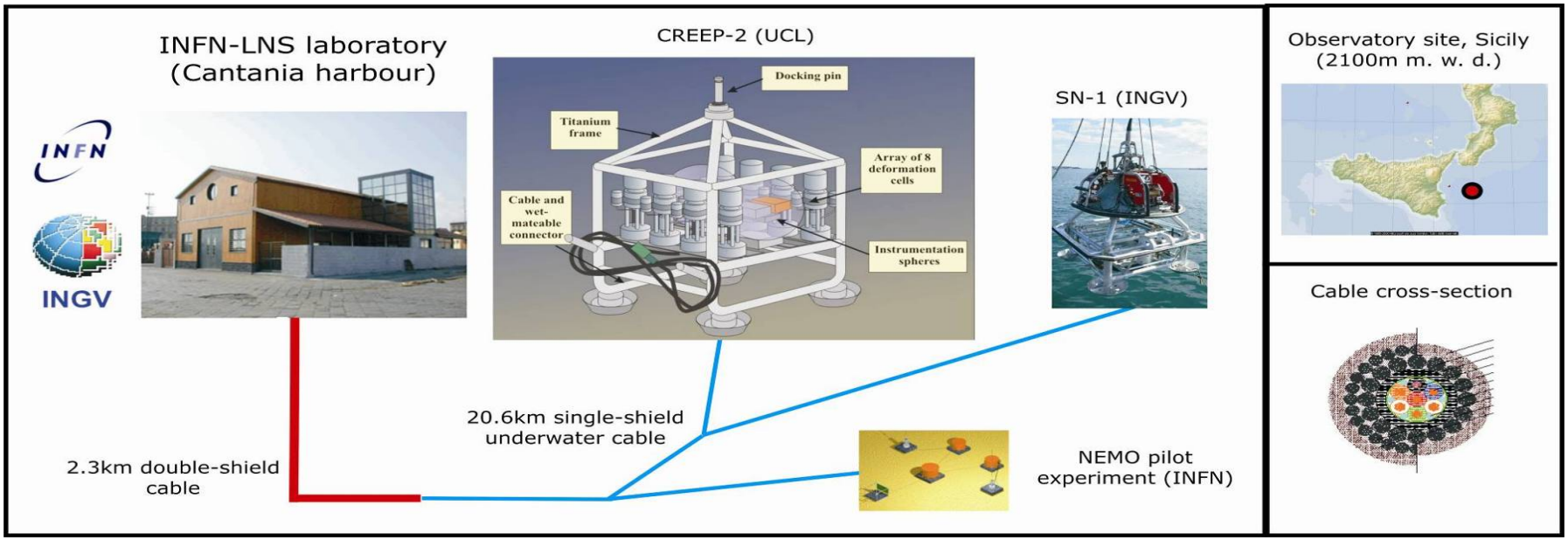
3. Simultaneous measure proxies for crack damage: Axial strain, porosity change, Output AE energy



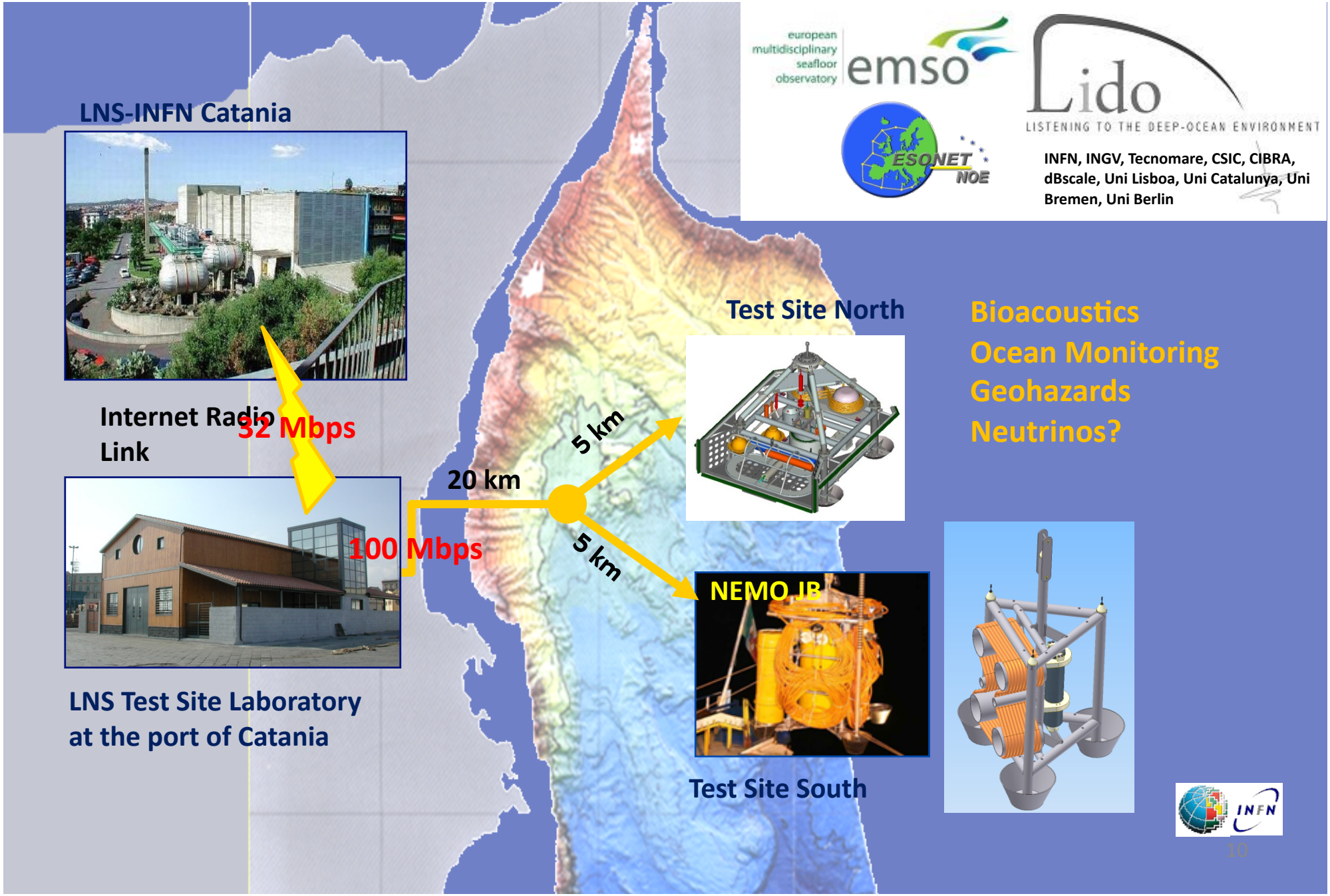
CREEP II - bridging the gap



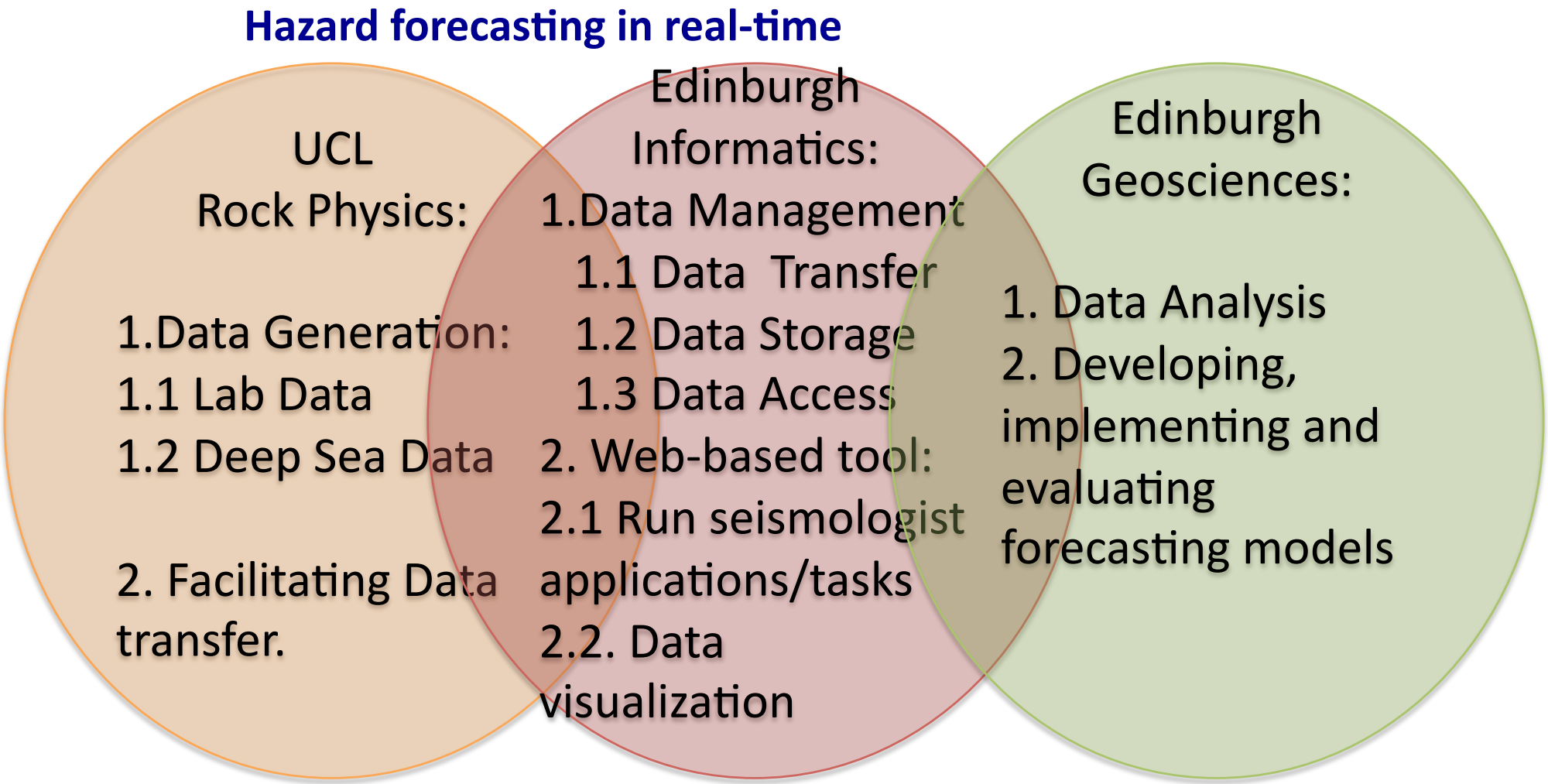
In deep sea: Constant temperature, pressure, long term experiment



Creep 2 Project.

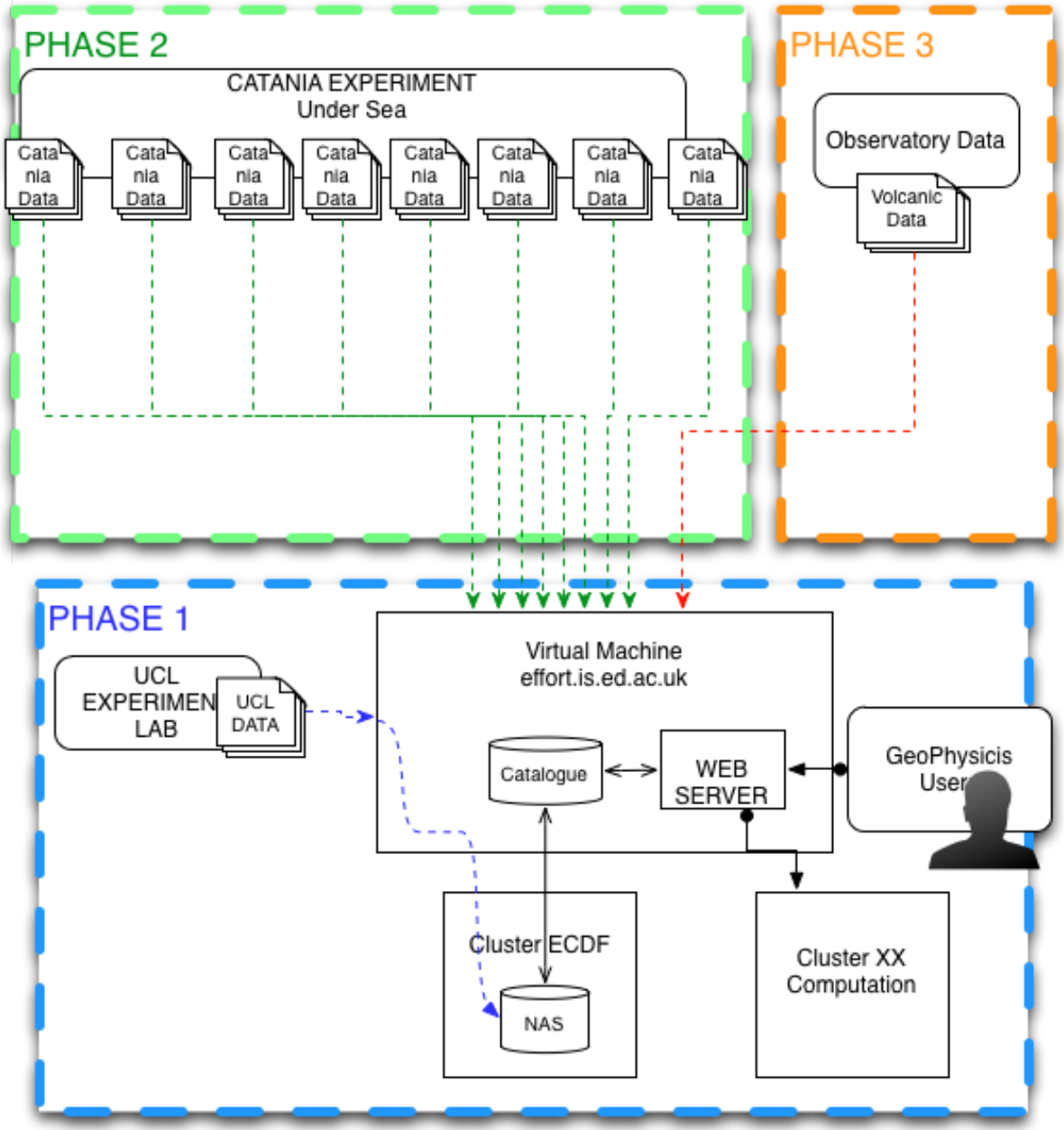


EFFORT roles and tasks



Edinburgh Informatics tasks for EFFORT

- Provide an efficient implementation of data management (inter and intra site).
- Provide a portal interface for EFFORT by using RAPID to allow data access, analysis and visualization.
  - Automatic job submission portlet generation system.
    - This system allows an expert to specify information about a job submission portlet in an XML file. The Rapid system takes this XML file and translates it into a fully porlet.


PHASE 2
CATANIA EXPERIMENT
Under Sea
Catania Data
Catania Data
Catania Data
Catania Data
Catania Data
Catania Data
Catania Data
Catania Data
PHASE 3
Observatory Data
Volcanic Data
PHASE 1
UCL EXPERIMEN LAB
UCL DATA
Virtual Machine
effort.is.ed.ac.uk
Catalogue
WEB SERVER
GeoPhysicis User
Cluster ECDF
NAS
Cluster XX Computation

Edinburgh vision of EFFORT

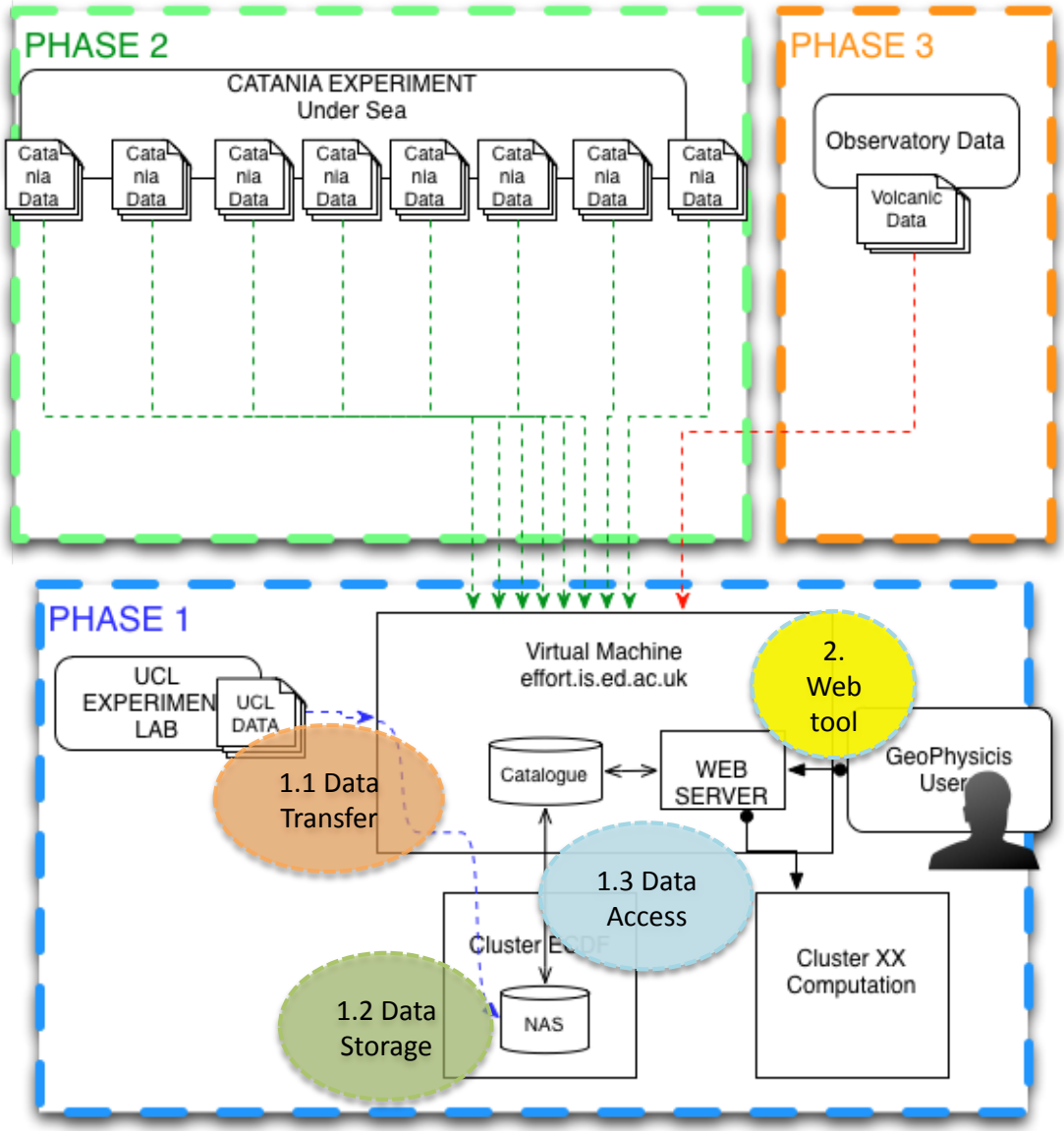
Edinburgh Informatics

1.Data Management
- 1.1 Data Transfer
- 1.2 Data Storage
- 1.3 Data Access

2. Web-based tool:

2.1 Run seismologist applications/tasks

2.2. Data visualization



Data Features

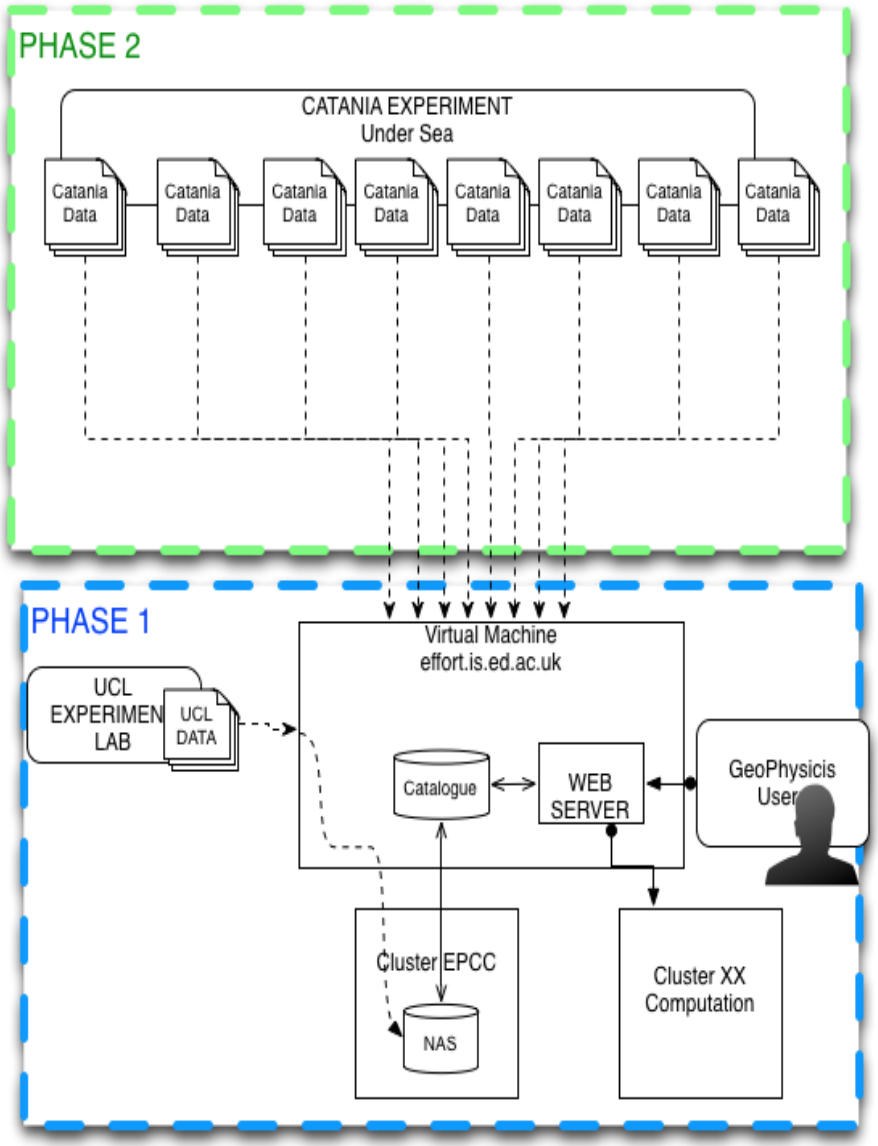
- Synthetic data:
  - Simulate the Controlled laboratory data.
  - Ideal for testing methodologies and exploring the limits of forecasts in “perfect” conditions.
  - The data are generating in the UCL laboratory machine, and transferred to Edinburgh in order to test as well the data management methodology.
- Controlled laboratory experiments:
  - Data from the UCL-Lab.
  - Data from Deep Sea in Catania.
- Volcanic data
  - Data from the Volcano Observatory web-site.

Controlled laboratory data Features

- Two data types:
  - Time Driven Data (TDD): Reflect changes in the bulk properties of the sample.
    - Time driven data has *constant adjustable interval.* One file per minute.
    - **1 TDD file ≈ 113 KB**
  - Acoustic Emission Data (AE): AE waveform data will be processed to pick individual AE events and compile a catalogue.
    - AE is generated only when the sample starts to get damaged, at very small intervals (micro seconds. Example: 10000 AE events in a second)
    - **1 AE file ≈ 204 MB**
- Automatic transfer files considering:
  - From different sources: UCL and Catania
  - Without manual interaction

Data size

- Phase2 :
  - TDD ≈ 160MB *6 experiment ≈ **960 MB/day**
    - **1440 files * 6 experiments**
  - AE ≈ 284GB * 6 experiment ≈ **1.6 TB/day (in the worst case)**
- Phase 1:
  - TDD ≈ **160 MB /day**
    - **1440 files of 113 KB each**
  - AE ≈ **284 GB/day**



Data Transfer challenges
Controlled laboratory experiments

- Chose and set up a mechanism on the server machine to receive data from UCL and Catania.
- The mechanism must to have the next characteristics:
  - Automatic, without human interaction.
  - Compatible with different operating systems:
    - Host machines: Windows
    - Server machine: Linux (Debian)
  - Support sending data every minute during a long period of time.
  - Catch up with transfers if there has been an intermission in connection, a reboot or a data loss.

Data Transfer solutions

- First prototype:
  - Run two periodically scripts (one in host and other in the server machine) by using winscp tool and SFTP protocol.
  - The host machine initiates the data transfer.
- Second prototype:
  - Run a periodically script in the server machine to transfer the data by using Globus online
  - The server machine initiates the data transfer.

Globus Online

- Globus Online is a convenient interface for transferring files between two endpoints.
- Globus Online offers a feature called Globus Connect, which enables you to move files easily to and from your laptop or desktop computer and other endpoints.
- After you set up Globus Connect, you can use Globus Online to transfer files to and from your computer.

Why use Globus Online

- **Easy:** simplified signup, login and use – and minimized manual intervention
- **Fast:** Globus Online can move large filesets in hours
- **Secure:** No need to worry about security configs or certificates, and one-time-passwords just work
- **Reliable:** Users can fire and forget their transfers
- **Research-focused:** Globus Online is the only file transfer service built with the scientific researcher in mind
- We use **command-line interface** to generate the periodically script in the server machine

Data Transfer Challenges
Volcanic Data

- Run a periodically script in the server machine, to connect to the Observatory web-site, and download data by using FTP.
  - **File** ≈ 20 KB.
- This scripts must to be executed once per day.

Data Storage challenges
Controlled laboratory experiments

- Store each experiment's files on Edinburgh service.
  - Files coming to the server, near real-time.
- Set up a database to store the catalogue of experiments
  - Create a catalogue in the database with the different experiments (Lab experiment + 6 Deep –Sea experiments).

Data Storage Solution

- Files stored in the NAS repository provided by ECDF
  - Different directories depending on:
    - Source, Number of Experiment, Data type.
- When a new experiment starts:
  - New enter in the catalogue (database) specifying the characteristic of the experiment, and the path where the files will be stored.
  - Database chosen: MYSQL

Storage challenges
Volcanic Data

- Record the volcanic data in the database.

Data Access challenges

- Use a Data Base to access more easy to the data:
  - Controlled experiments: Use the catalogue to find the files.
  - Volcanic data: Use the data from the database.
- Adapt the RAPID portlet to allow seismologists access the data and run different tasks/applications:
  - Select the data of one experiment using the database.
    - In the case of volcanic data, work with MYSQL and Spatial Extensions.
  - Develop a Task:
    - Apply a forecasting model to the data selected.
  - Execute and monitor the task as it is executing.
  - Copy the results from the server to the host.
  - Visualise the result of the computation in the portal

Technologies Under Consideration

Data Transfer

GridFTP, Globus Online, RSYNC, FDT, WINSCP

Data Storage

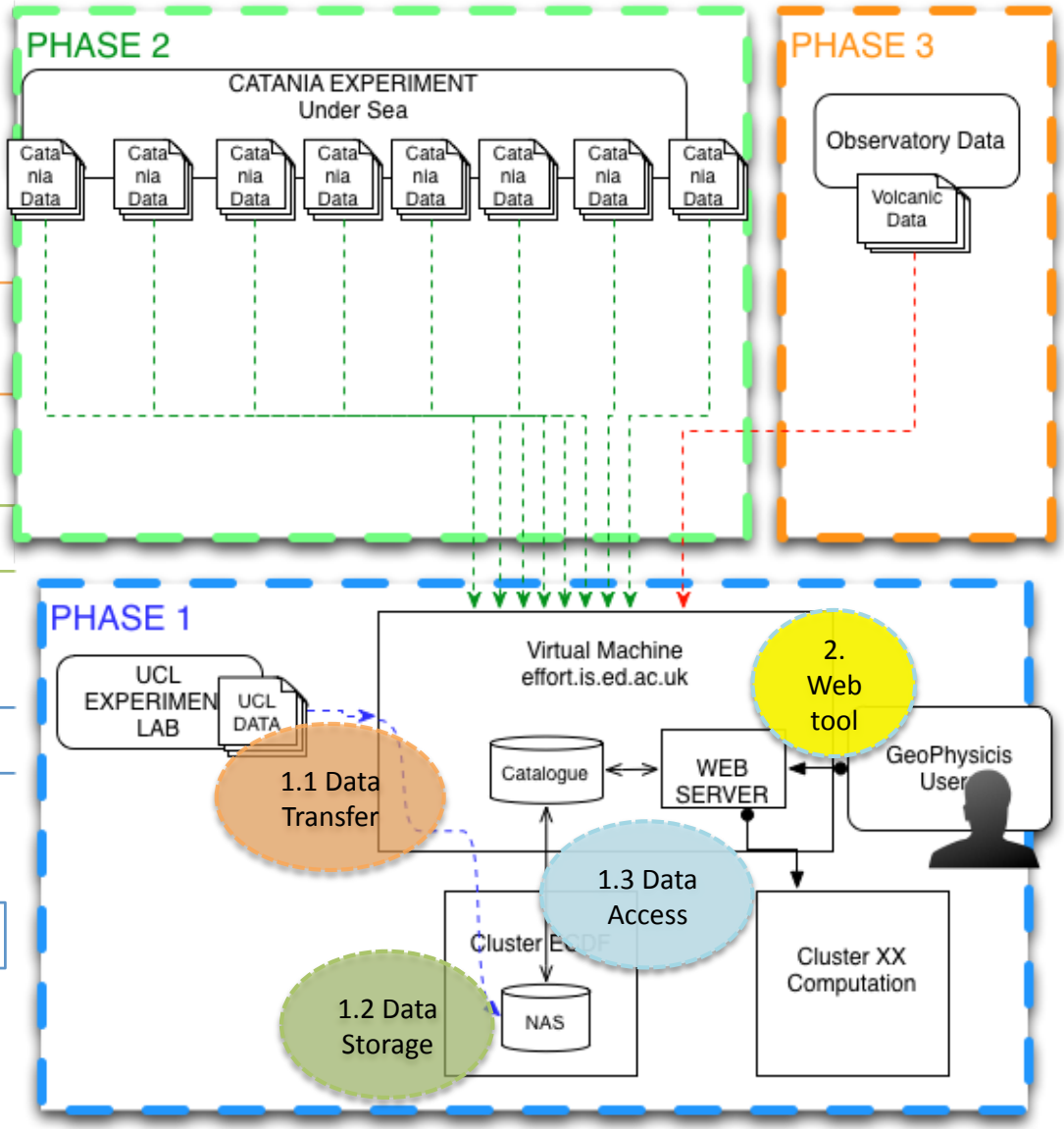
Irods, SeisHub, MonetDB, Mysql

Data Access

Mysql

Web tool

Liferay + Rapid.



First Prototype

Data Transfer

Periodical Script by using WINSCP tool :
→FTP connexion
→Send data
→Check data

Data Storage

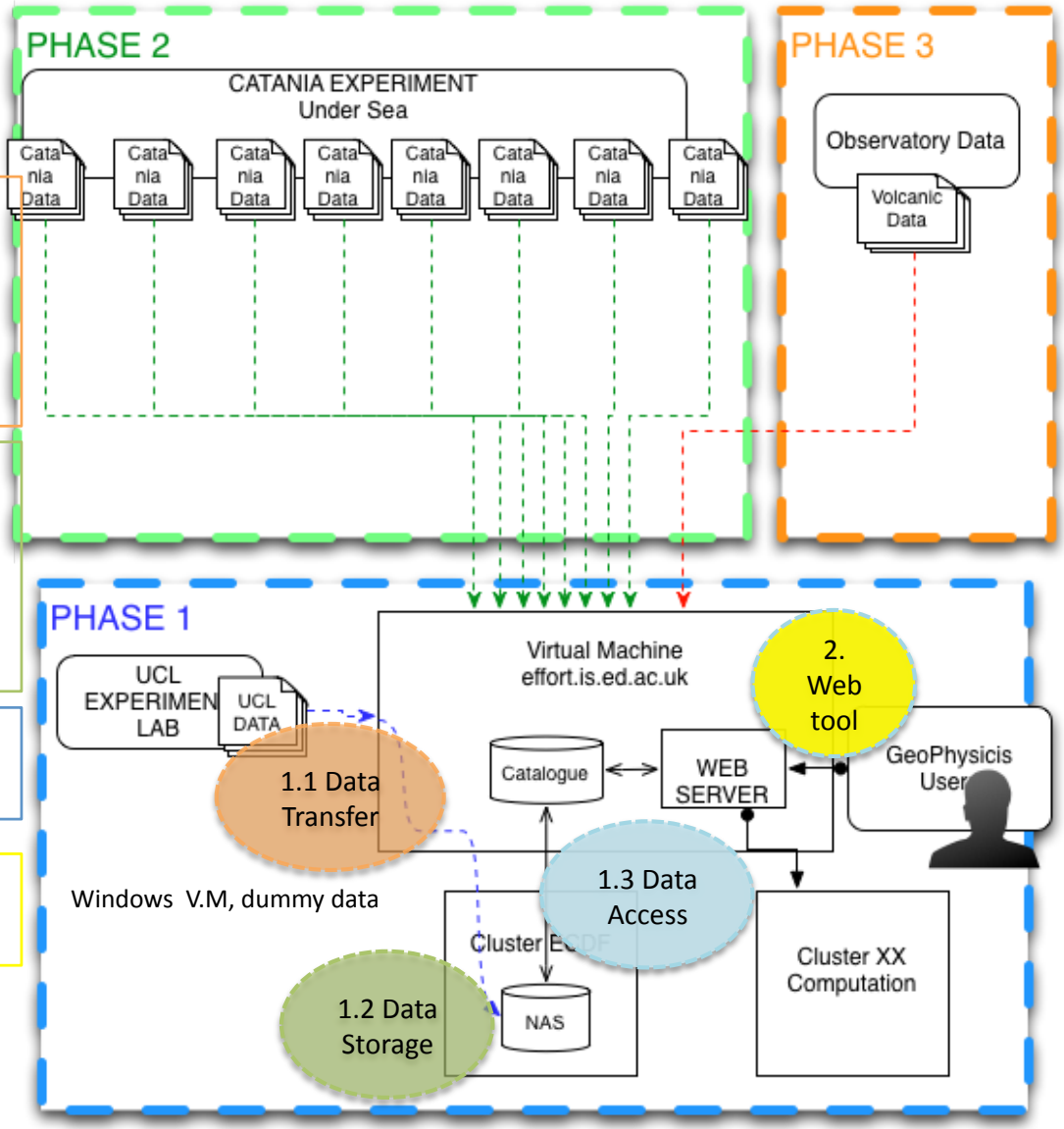
Store the files in the NAS in the correspond directory.
Transform files format.
Store the data in a Database (MYSQL)

Data Access

By using the catalog find the data associate with a experiment.

Web tool

Prototype of Web tool by using a Rapid porlet. Access to the data.



Second Prototype

Data Transfer

*Creep-2: Periodical script by using Globus Online.
*Volcanic data: Periodical script by using FTP.

Data Storage

*Creep-2: Store the Creep-2 files in the NAS in the correspond directory. Store the catalogue of the creep2 experiment in the Database.
*Volcanic data: Store the volcanic data in the database.

Data Access

By using the database find the data associate with a experiment.

Web tool

Several portlet to apply forecast models and visualize the results in the web site.