





## PRAS-DT: Portable, Reliable, and Automatic Streaming Data Transfer

Christine Harvey Dr. Rosa Filgueira

# **OSDC PIRE Fellowship**

- Four week internship
- University of Edinburgh, School of Informatics
- Dr. Rosa Filgueira The Effort Project



#### EFFORT goals

- To determine the predictability of brittle failure of rock samples in the laboratory experiments.
- To determine how this predictability scales to the greater complexity, physical size, and slower strainrates of natural-world phenomena.
- To develop a Forecasting Model Testing Centre for archiving and monitoring Rock Physic data.

#### EFFORT roles and tasks

UCL Rock Physics:

Data Generation: •Lab Data •Deep Sea Data Edinburgh Informatics:

Forecasting Model Testing Centre \*Data Transfer \*Web Data Access \*Metadata & Data Storage Edinburgh Geosciences:

Data Analysis \*Forecasting models

### EFFORT experimental data

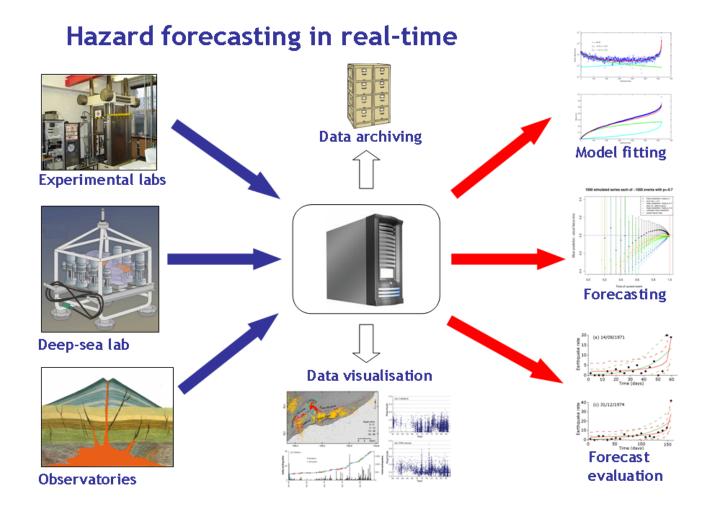
- Experiments at UCL- Rock Physics Laboratory
  - Data focus for the internship
- Deep- sea experiments, produced by the Creep2 project
- Synthetic data
- Volcanic data

#### Experiments at UCL- Rock Physics Laboratory

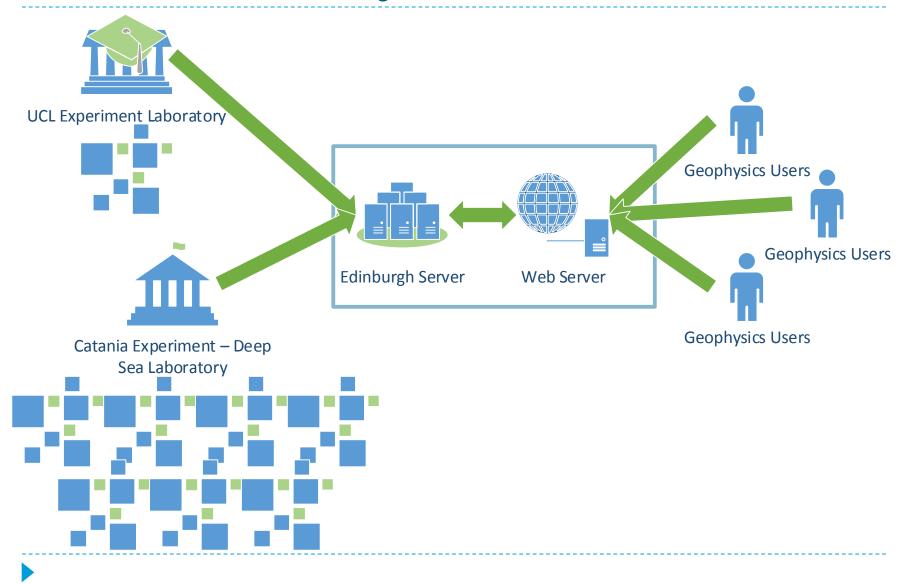
- Data Transfer features of UCL laboratory experiments
- Time Driven Data (TDD) up to 500 KB/Day
  - Every day a new TDD file is generated. Every minute, this file is modified with new information. This file must be synchronized to Edinburgh resources every minute.
- Acoustic Emission (AD) Data up to 75KB/Day
  - As soon as the rock (sample) starts to break, an AD file is generated per day. This file is updated during small intervals (microseconds). This file must be synchronized to Edinburgh resources every minute.
- The duration of an experiment is undefined. The last one was 45 days. Could be longer.



# Data Transfer Outline



## The EFFORT Project Data Outline



## Data Transfer challenges Controlled laboratory experiments

- Chose and set up a mechanism on the server machine to receive data from UCL and Catania.
- Necessary Characteristics:
  - Automatic, without human interaction.
  - Compatible with different operating systems:
    - Host machines: Windows
    - Server machine: Linux ( Debian)
  - Support sending data every minute over along period of time.
  - Ability to 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 scripts periodically (one in host and other in the server machine) using a 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 with Globus Online
  - The server machine initiates the data transfer
  - PRAS-DT: Portable reliable adaptive streaming data transfer

## PRAS-DT

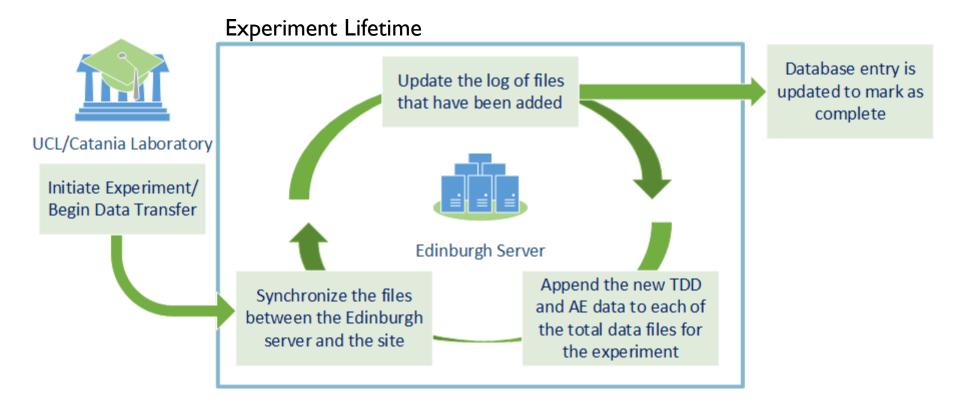
#### Globus Online

- Fast setup time
- Automated Script
  - Run for length of experiment



Endpoint rosa#Ch	ristinePC V			Er	ndpoint	t rosa#effort		▼
Path /~/desk	op/UCL-Christine/	Go			Path /~/			
select all   none	<b>t</b> up one folder	🖒 refresh list	<b>\$</b> -	select all   none	t_ up	one folder	🖒 refresh list	<b>\$</b> -
dir_results_10			Folder	💼 CEH_globusOnli	ne-script			Fo
dir_results_101			Folder	CProg				Fo
dir_results_11			Folder	🚞 ChristineTest				Fo
dir_results_12			Folder	🚞 Receive_Data				Fo
dir_results_14			Folder	TDD_SIMULATO	R			Fo
dir_results_6			Folder	🚞 globusOnline-sc	ript			Fo
dir_results_7			Folder	CEHPoster.pdf				1.0
dir_results_8			Folder	_				
dir_results_9			Folder					
HelloWorld.txt			9b					
TDD_Generator.py	r		4.42kB					
more options	Label This Tr	anster Transfer Ex	periment 10 Res	ults				
		This will be	displayed in your tra	no for a clivity				

### PRAS-DT



# Project Outcomes

Poster presentation at Supercomputing 2012

- Education Program
- MITRE interview presentation
- Continued work by Dr. Filgueira

#### Contact

- Christine Harvey
  - cehavrey@mitre.org
- Dr. Rosa Filgueira
  - rosa.filgueira.vicente@gmail.com