

ESGF Dashboard Working Team

P. Nassisi, S. Fiore, G. Aloisio
Euro Mediterranean Center on Climate
Change (CMCC)

Marriott Hotel
Dec 8-11 2015, Monterey, CA

Outline

- ❖ Working Team mandate, leads and members
- ❖ Architecture in the large of the system
- ❖ Work achieved over the past year – Intro
 - ❖ Coarse grain system - Design view
 - ❖ Downloads statistics
 - ❖ Downloads by Identity Provider
 - ❖ Downloads by User activity
 - ❖ Fine grain system - Design view
 - ❖ Federated and cross-project statistics
 - ❖ Project-specific statistics
- ❖ Prioritized development and roadmap for the next year



Working Team mandate, leads and members

- ❖ *Working Team Acronym*

esgf-dwt (Dashboard Working Team)

- ❖ *Mandate*

Design and implementation of a distributed and scalable system for (i) monitoring the Earth System Grid Federation and (ii) providing data usage statistics in a comprehensive way and through a simple and intuitive web interface.

- ❖ *Leader*

Sandro Fiore (until 2014), Paola Nassisi (since 2015)

- ❖ *Members*

Paola Nassisi, Alessandra Nuzzo, Maria Mirto, Sandro Fiore

- ❖ *Modules*

esgf-dashboard and esgf-desktop



Architecture in the large of the system

The main modules of the monitoring system are, for the **back-end**:

- ❖ the Information Provider

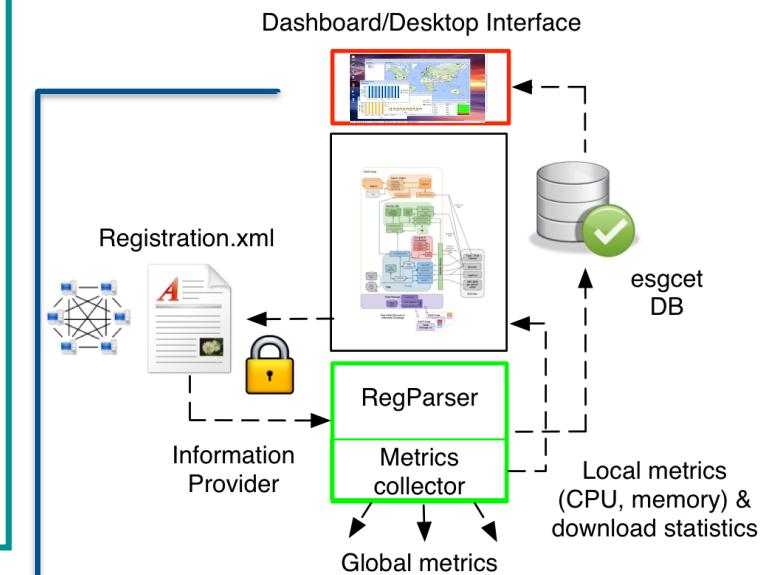
It's responsible for retrieving all the metrics and storing them in the *esgcet* catalog and binary files (for long term statistics).

- ❖ the dashboard catalog

A system database which stores all the information about hosts, peer-groups, services, users, availability, deployment, etc.

- ❖ the sensors

Global and local metric sensors to retrieve and manage information about *node type*, *registered users*, *downloaded data*, *system metrics* both a single site and federation level

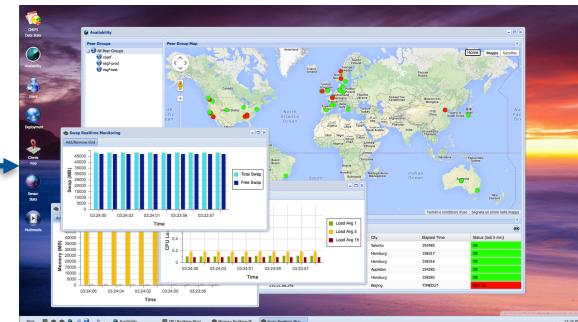


For the **front-end**:

- ❖ a web-based environment GUI: the ESGF Desktop

modular web application relying on a strong adoption of Web 2.0 concepts and providing several views at different granularity levels

- ❖ a set of configuration files



Work achieved over the past year - Intro

❖ Activity

- ❖ Complete implementation of the first data usage statistics system (**coarse grain**), relying on the existing access logging system.
- ❖ Status: Under testing. Final bug fixing
- ❖ Delivery date: December 2015

❖ Metrics

❖ Download stats

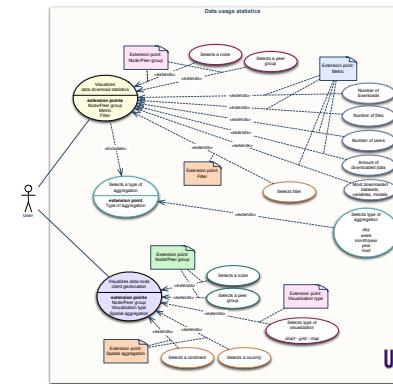
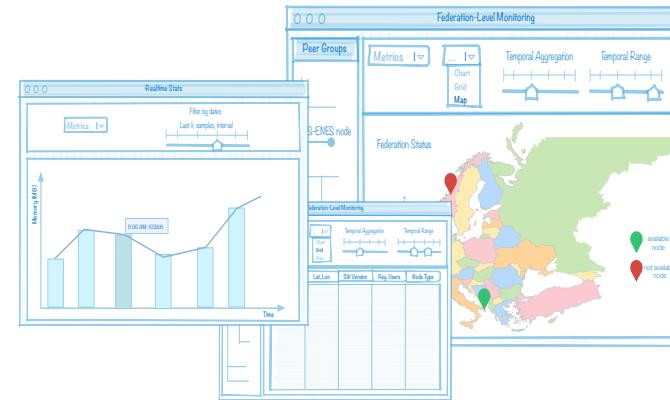
- ❖ Data downloaded (GB/TB)
- ❖ Number of downloads
- ❖ Number of distinct files
- ❖ Number of distinct users
- ❖ Downloads by user
- ❖ Downloads by Identity Provider

❖ Clients statistics

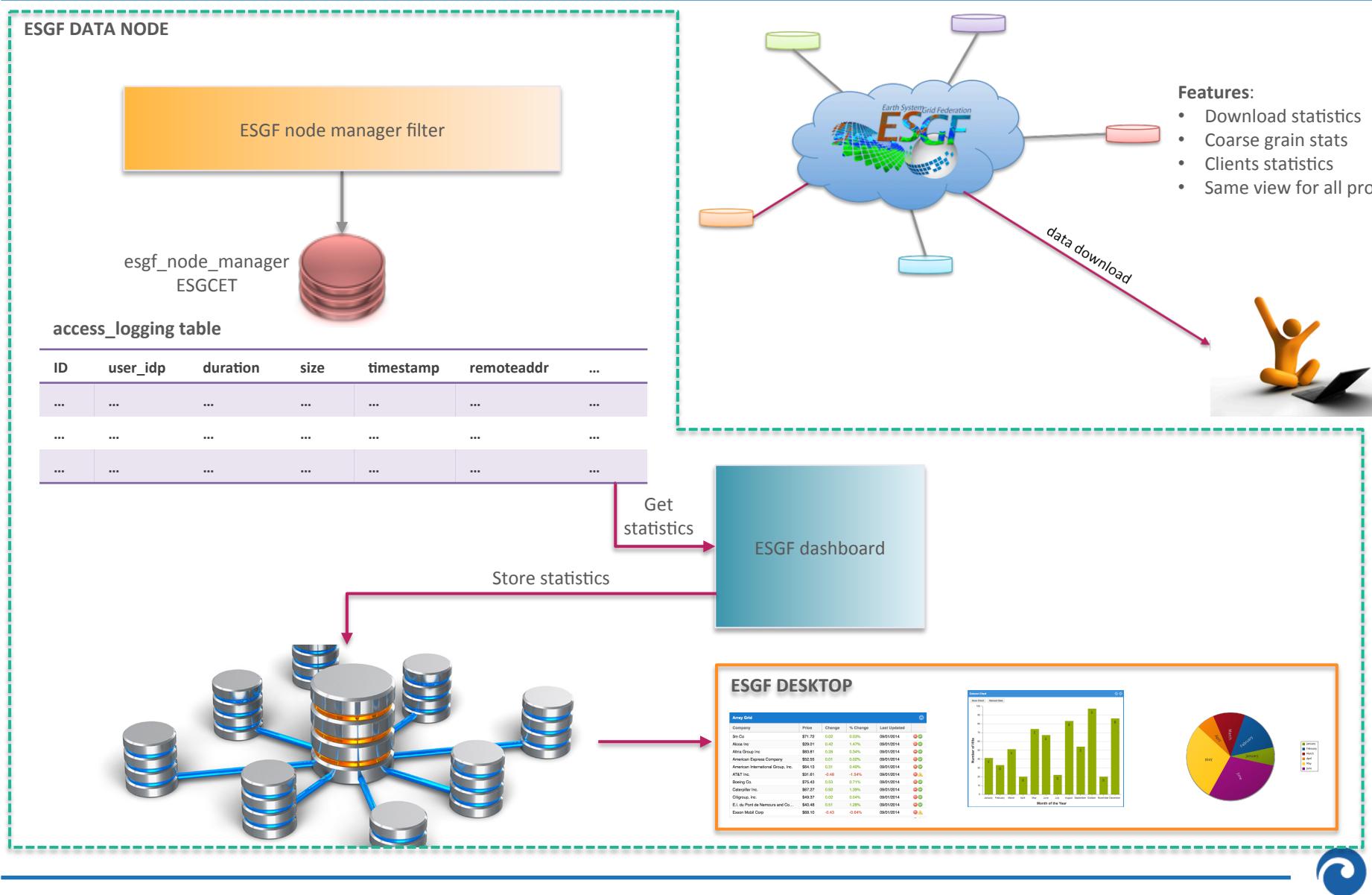
- ❖ Geographic map
- ❖ Country/continent distribution

❖ Design and implementation of the new system (**fine grain**)

- ❖ Prototype Jan 2016
- ❖ First release Feb 2016

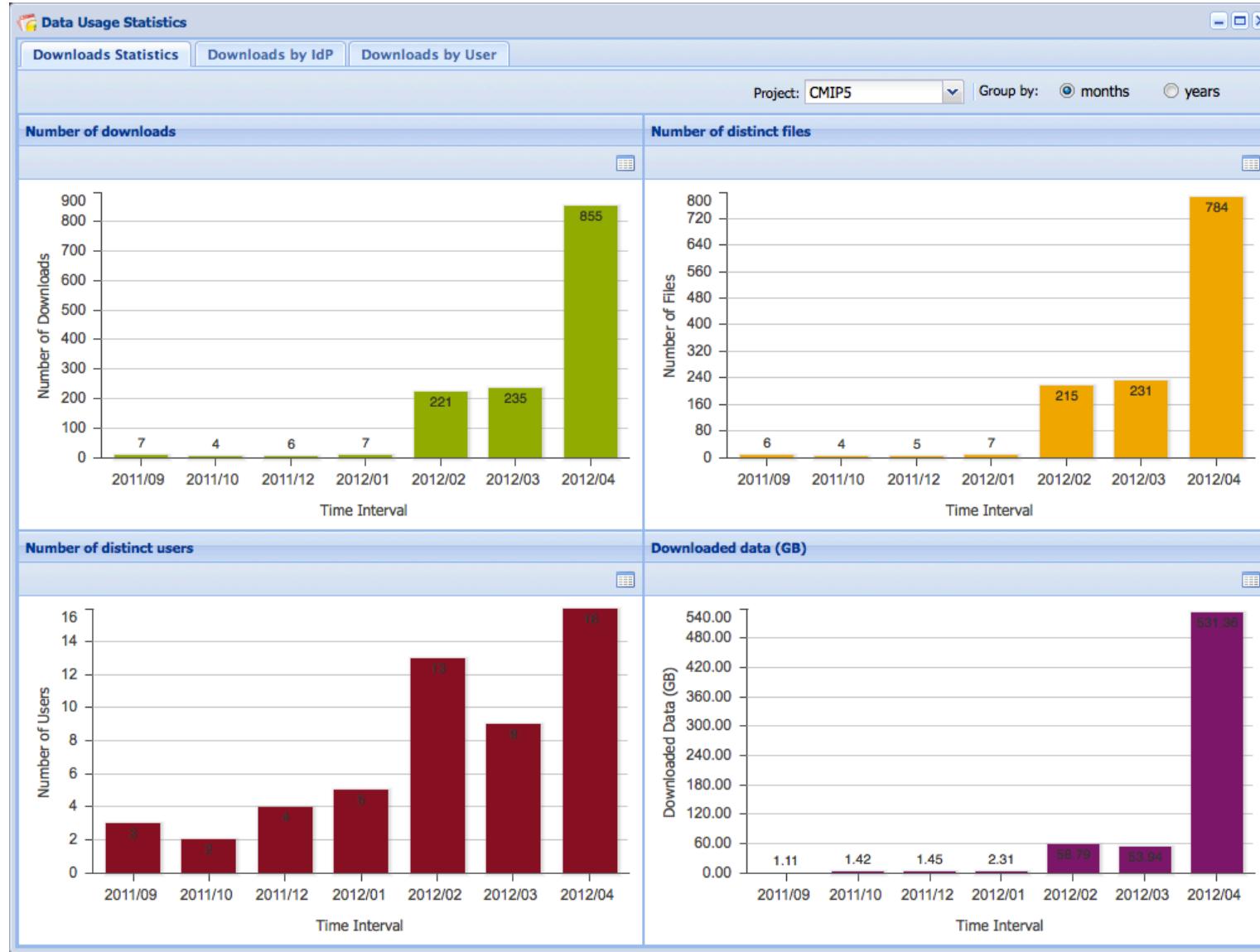


Work achieved over the past year – Coarse grain system Architectural view



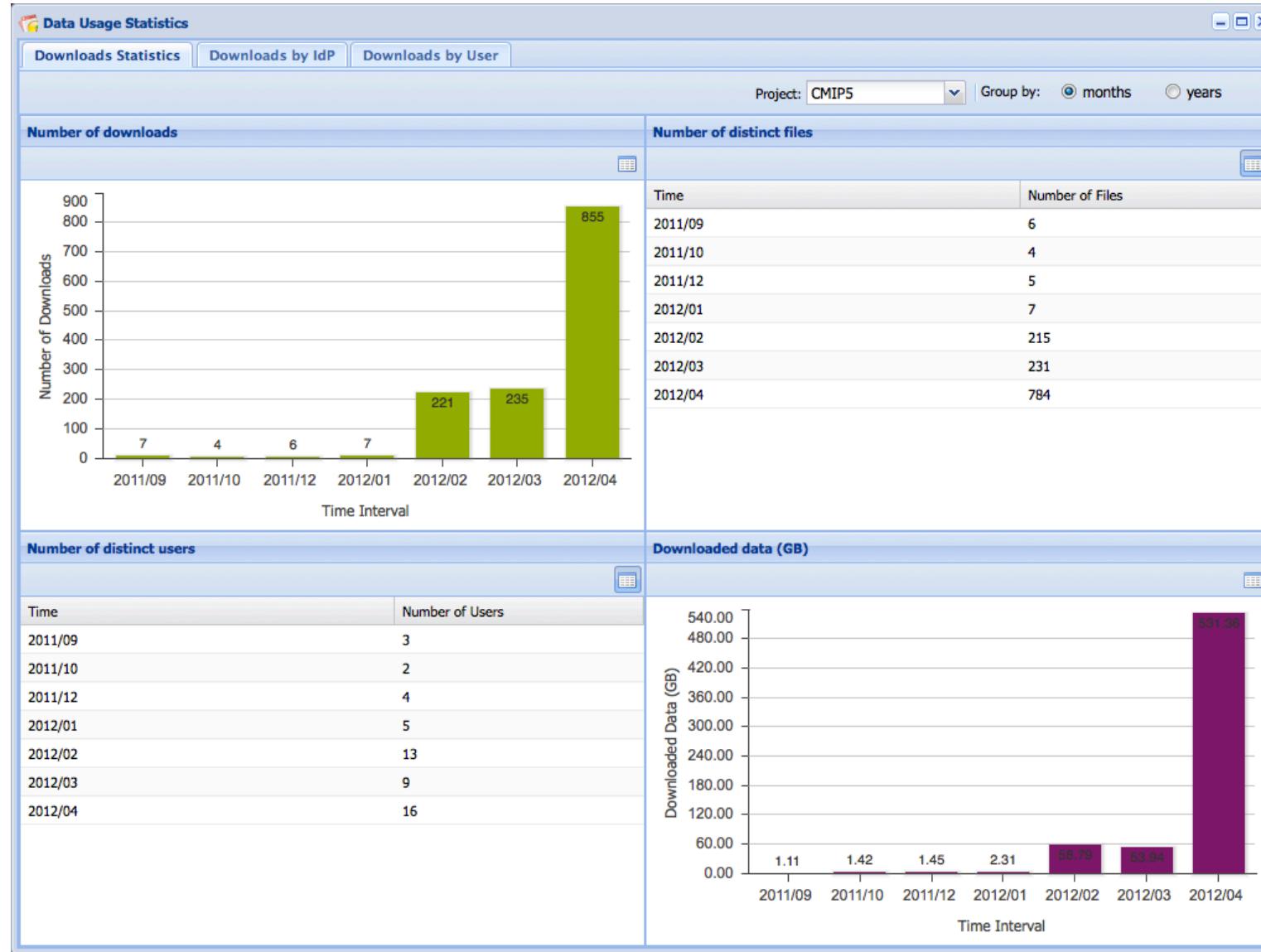
Work achieved over the past year – Coarse grain system

Download statistics (I)



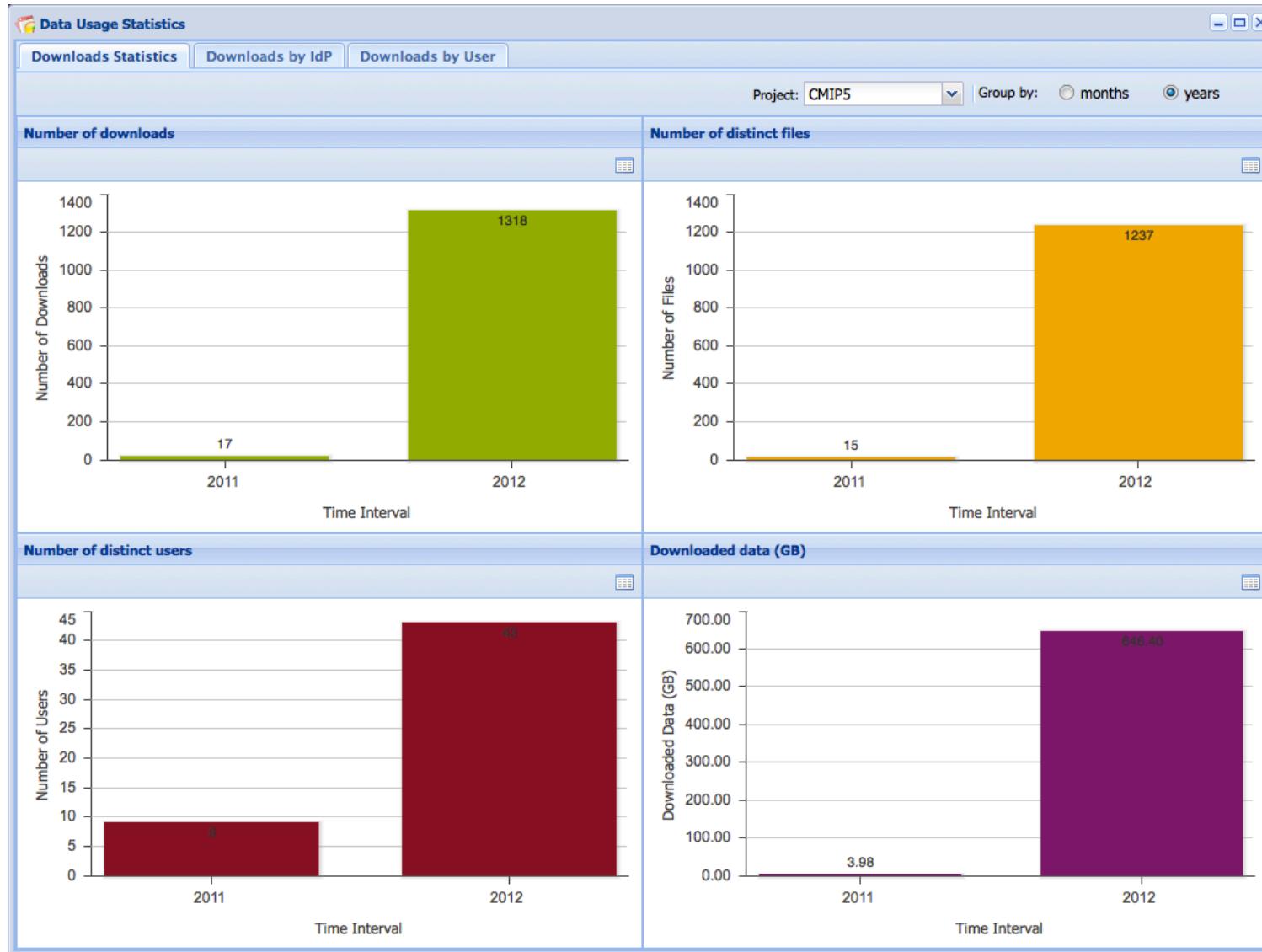
Work achieved over the past year – Coarse grain system

Download statistics (II)



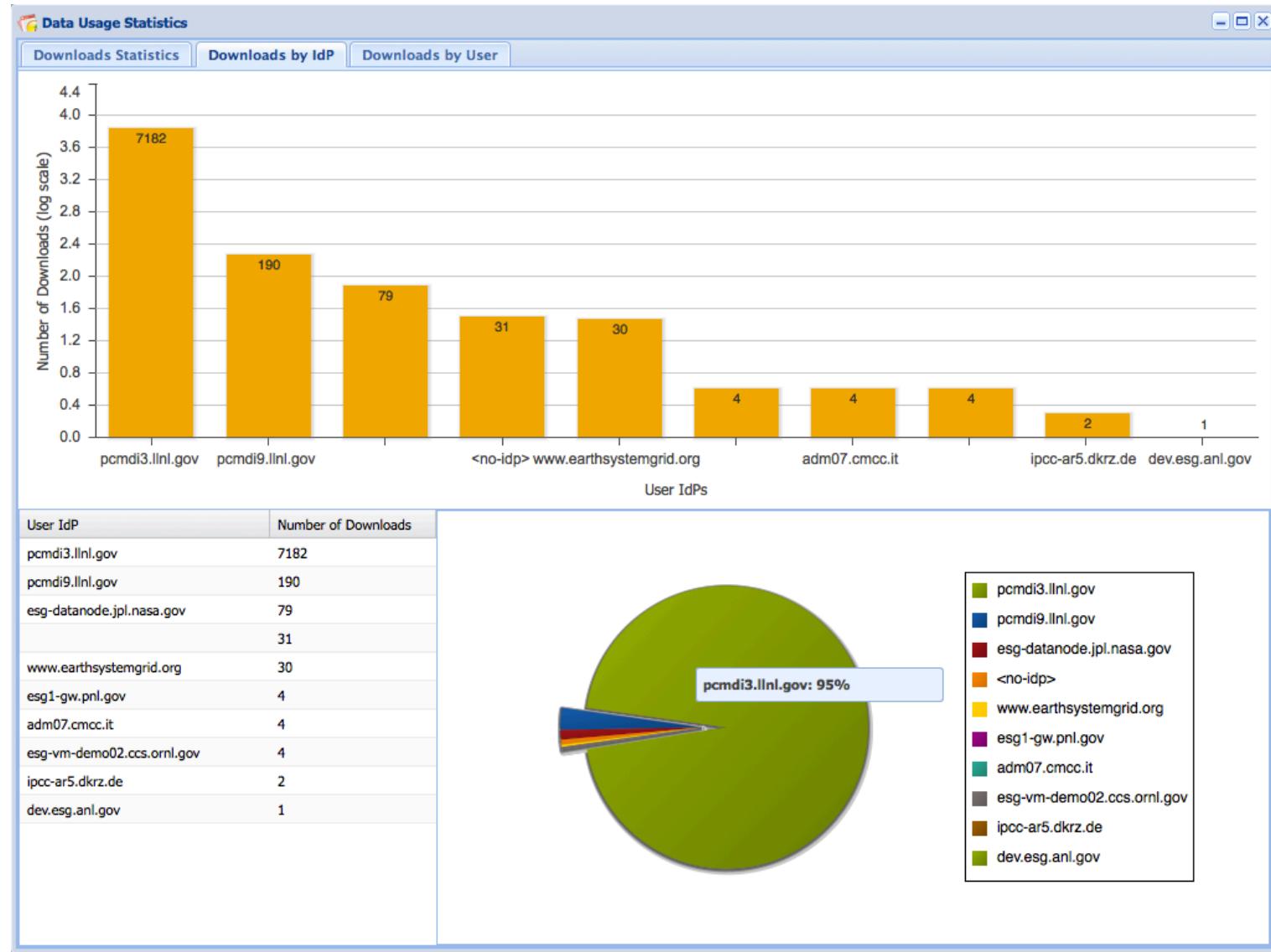
Work achieved over the past year – Coarse grain system

Download statistics (III)



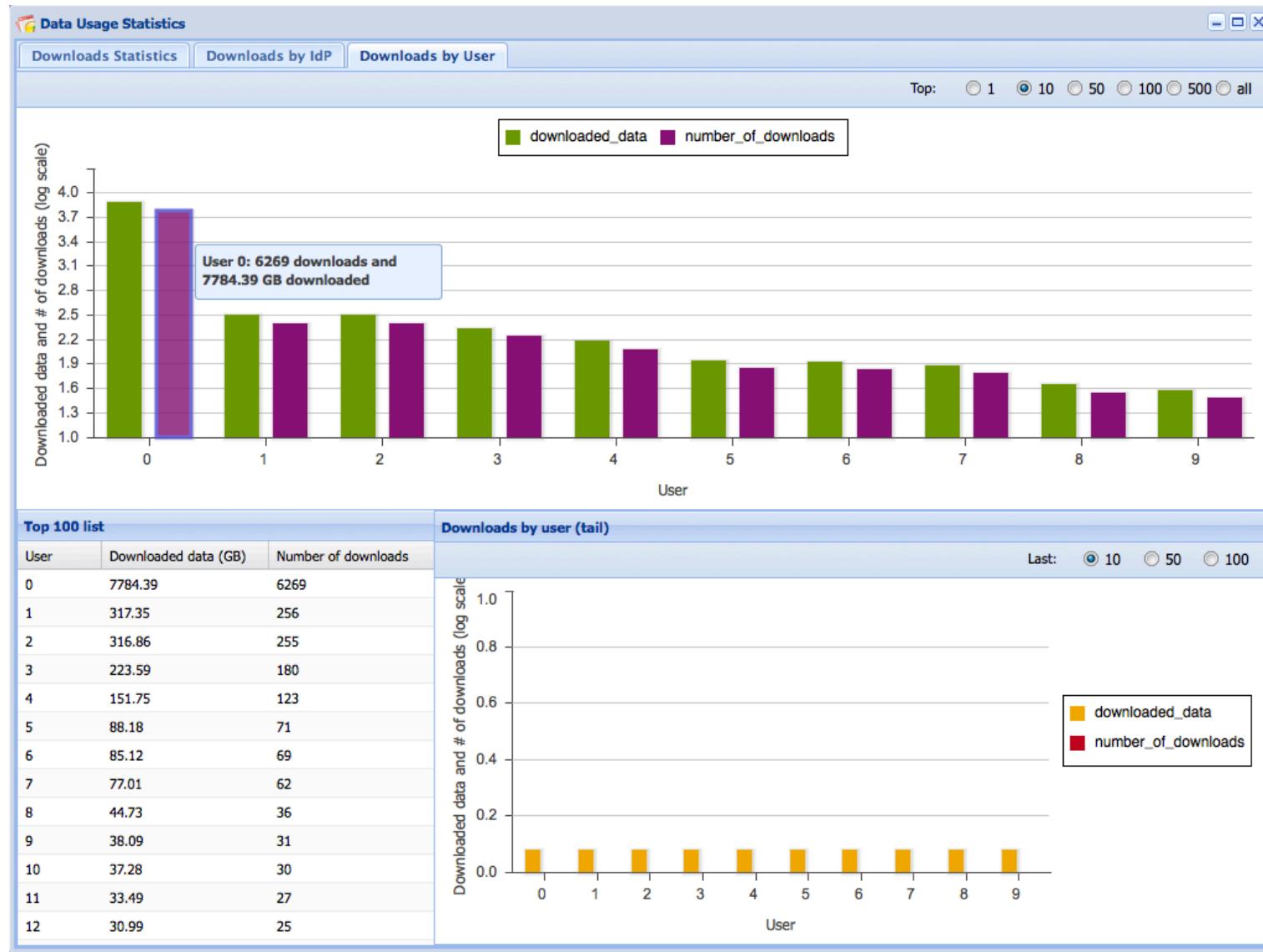
Work achieved over the past year – Coarse grain system

Downloads by Identity Provider



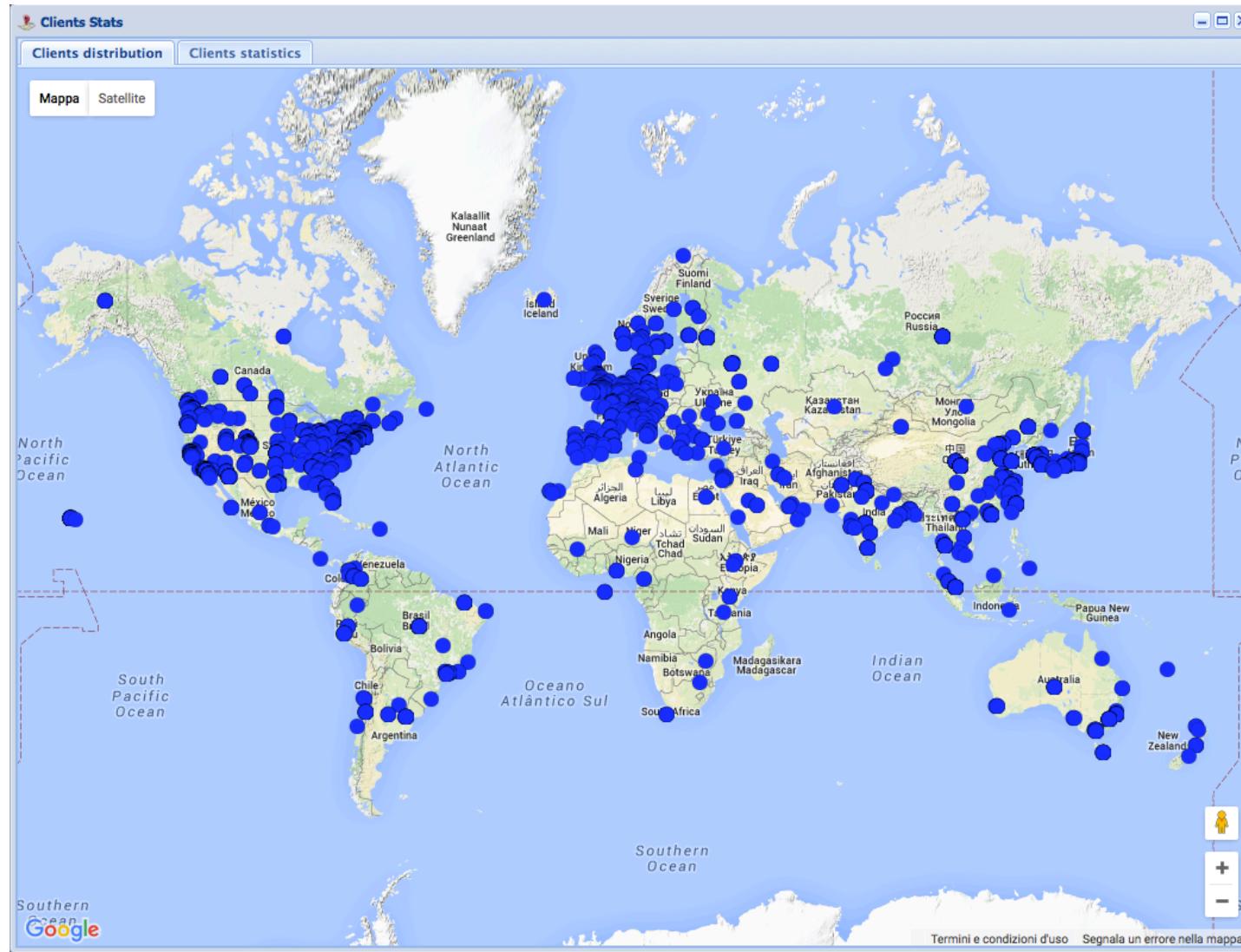
Work achieved over the past year - Coarse grain system

Downloads by User (activity)



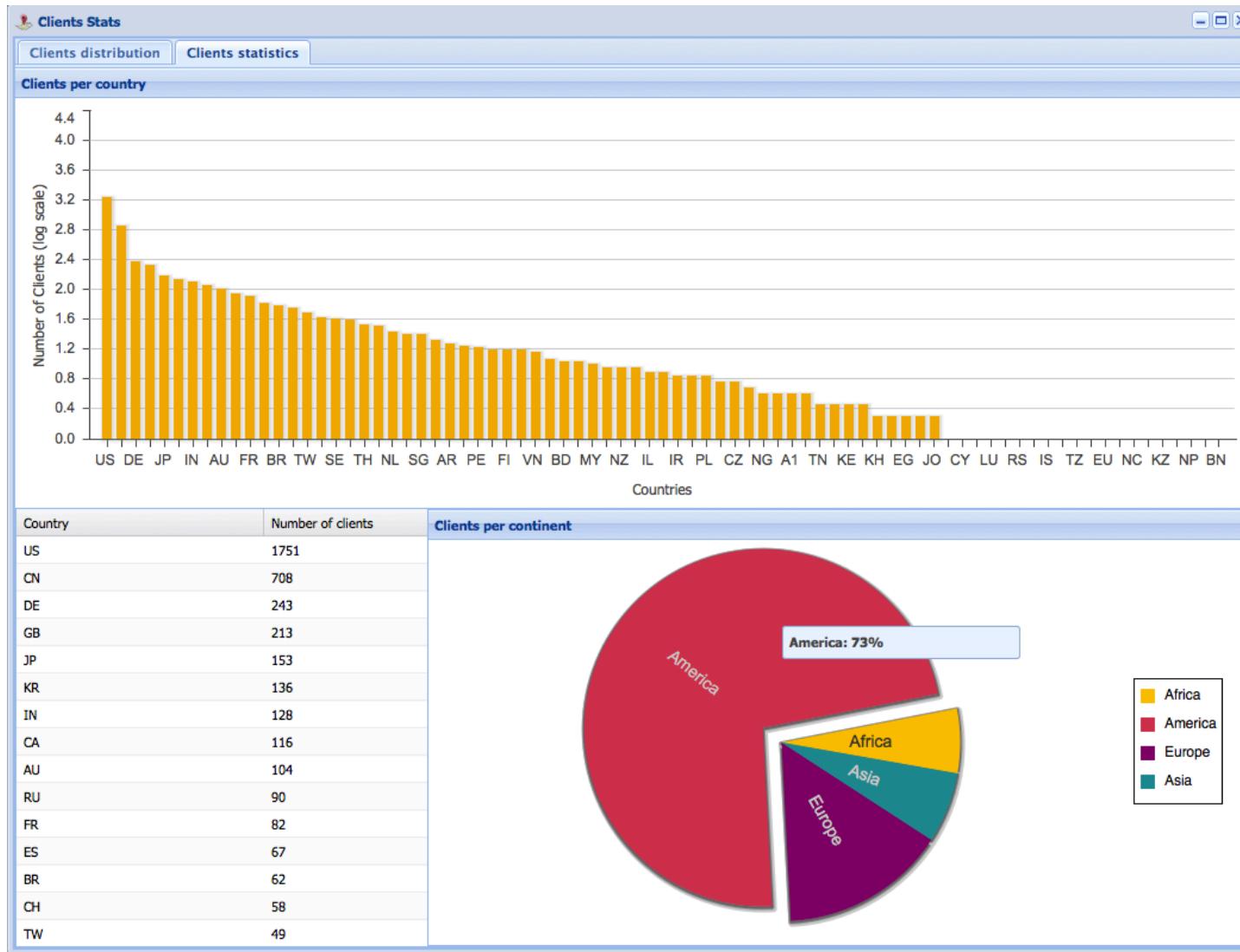
Work achieved over the past year - Coarse grain system

Client distribution

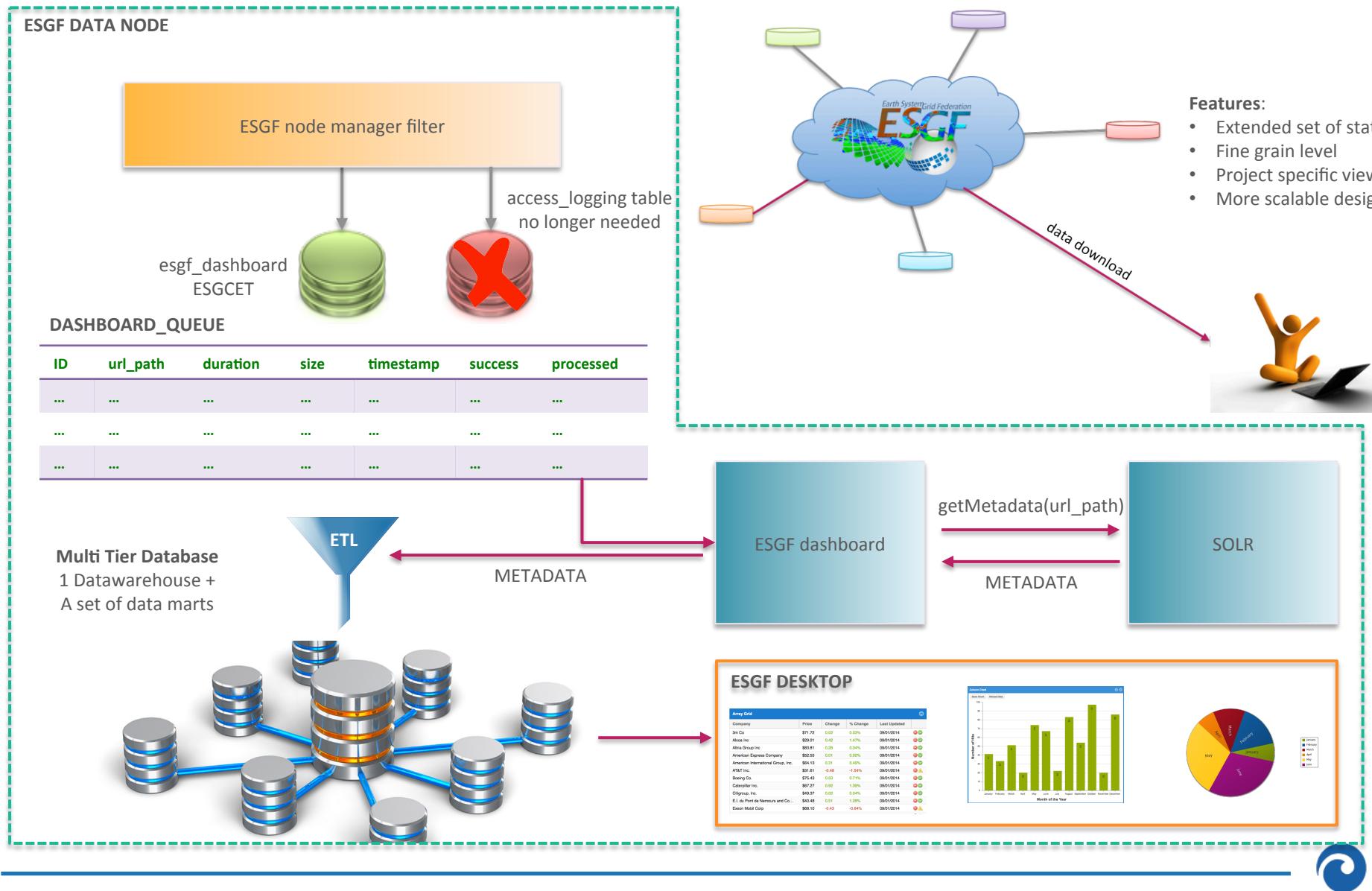


Work achieved over the past year - Coarse grain system

Client statistics



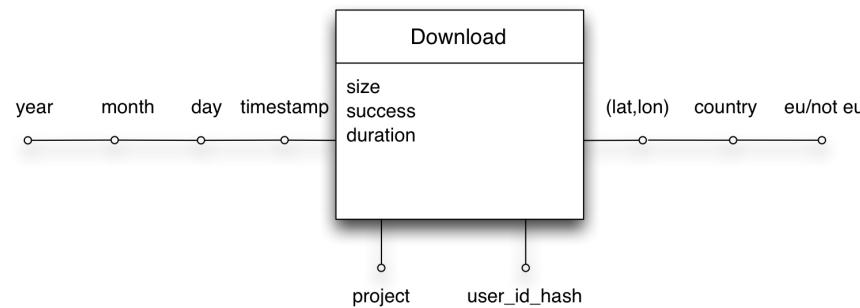
Work achieved over the past year – Fine grain system Architectural view



Work achieved over the past year – Fine grain system Statistics database design at glance

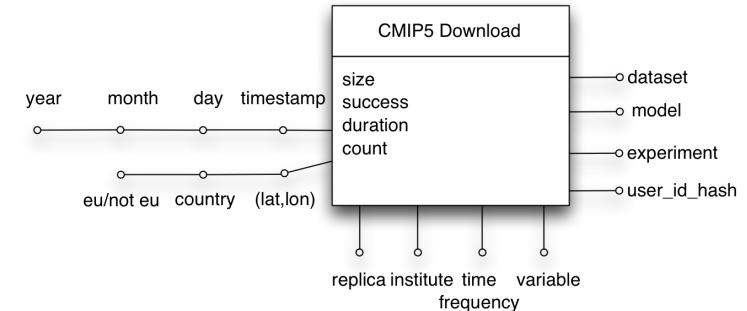
Dimensional Fact Model

Level 1: data warehouse cross project

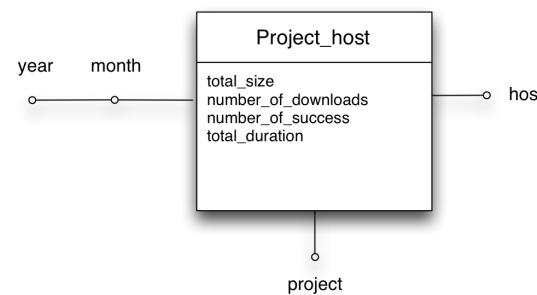


Dimensional Fact Model

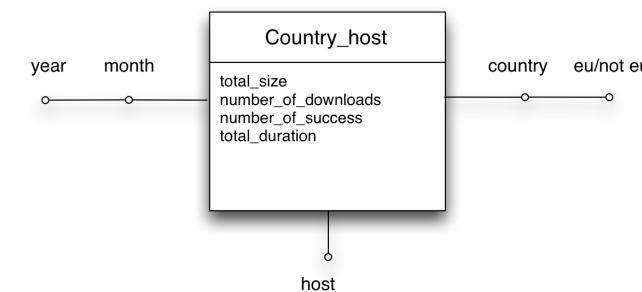
Level 2: data warehouse project specific



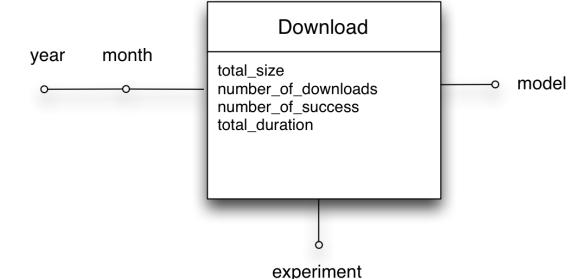
Data mart 1



Data mart 2

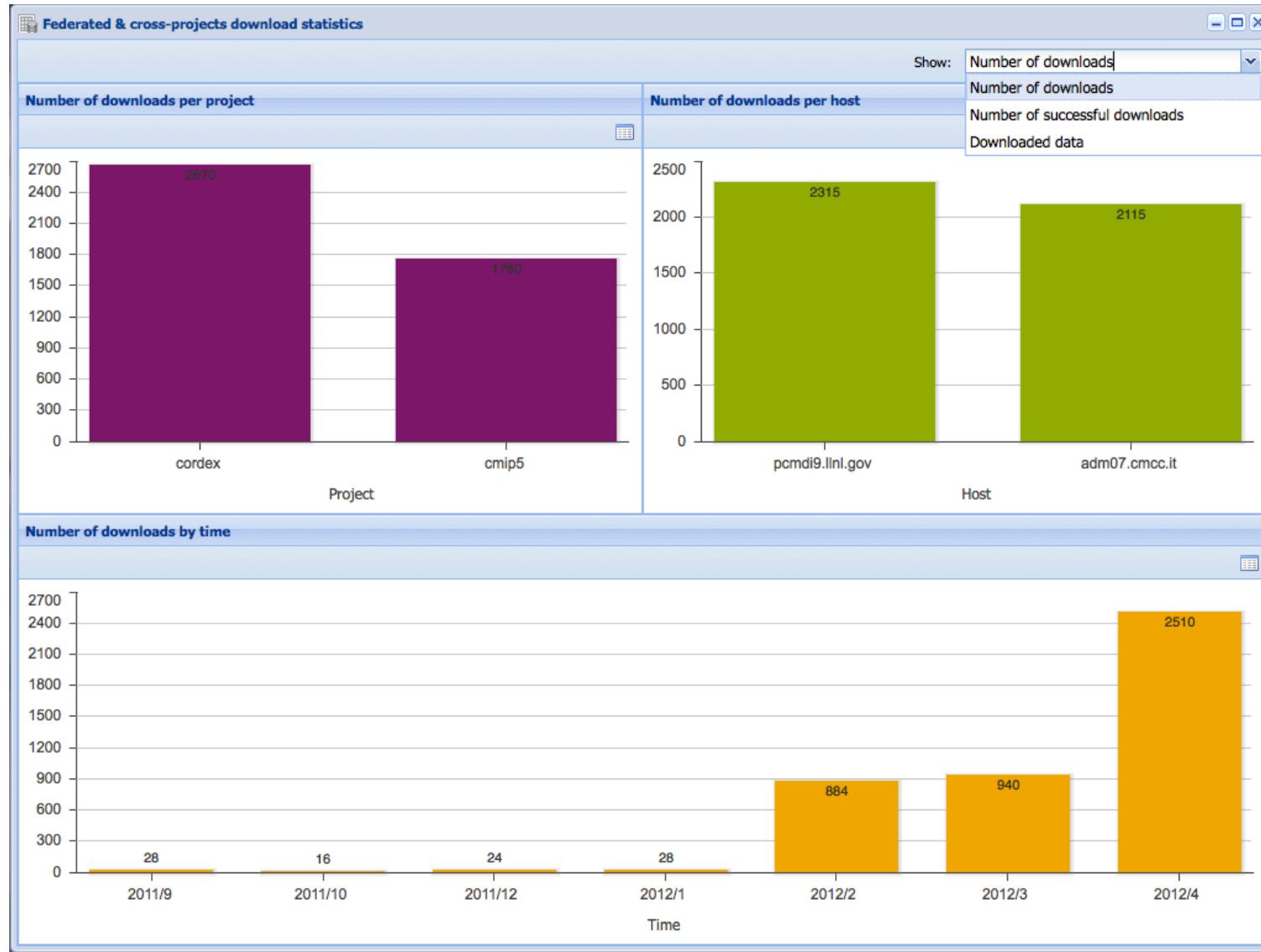


Data mart 1



Work achieved over the past year – Fine grain system

Federated and cross-project statistics



Work achieved over the past year – Fine grain approach

Project-specific statistics (CMIP5)



Prioritized development and roadmap for the next year

- ❖ Status of the system and future roadmap for the fine grain system
 - ❖ Database back-end design completed
 - ❖ Preliminary implementation available
 - ❖ Single site & Federated
 - ❖ Single and cross-project
 - ❖ Search lib in progress to extract project-specific information
 - ❖ e.g. obs4MIPS: cf_standard_name, processing level, realm, time frequency, variable, datetime start, datetime stop, etc.
 - ❖ New project specific views to be defined in the context of the ESGF-DWT
 - ❖ Delivery date: February 2016 for CMIP5, CORDEX, obs4MIPs
 - ❖ Ext. deps: link to the esgf-node-manager for gathering federation-level information
 - ❖ Stronger link with: node-manager, network and search working teams.
 - ❖ REST APIs
 - ❖ Single node level – Jan 2016
 - ❖ Federation level (March 2016)
 - ❖ Extended set of views with geo-location and federation-level statistics by May 2016
 - ❖ New front-end presentation layer (Aug 2016 first release, Dec 2016 final release)



Thank you

