Contribution ID: **22** Type: **not specified**

# The gCube Data Transfer Facilities

*Tuesday, 27 March 2012 17:00 (30 minutes)*

## Description of the Work

CERN as responsible of the Data Transfer task will integrate and enhance a set of facilities for reliable data transfer mechanism between the nodes of the Data e-Infrastructure. The iMarine data providers manage a large set of multi-type datasets distributed across different repositories. This task will provide a secure, reliable, and efficient solution for other deployed services to move different data types between remote infrastructure nodes under different transfer protocols (e.g. srm-copy, gridFTP, HTTPS, BitTorrent, OPeNDAP, WCS, WMS, WFS, etc) and the combination/optimization of state-of-the-art technologies (i.e. high-bandwidth networks, peer-to-peer). This task will also work on a data transfer mechanism to pass data by reference between infrastructure services by relying on a list of records that are part of a specific record set. The data transfers facilities will support multiple transfer requests, minimize network load, not cause storage overload, manage transfer shares at service and user level, and allow data transfer parameterization.
CERN is already responsible of the FTS service ( File Transfer Service) which is used in both WLCG and EGI infrastructures as main actor for file transfer between grid storage element. FTS as planned in EMI project is going to be partially re-implemented ( FTS 3) mostly to remove some limitations of the current production version (i.e.Oracle DB dependency, static transfer channel configuration). For this reason the CERN iMarine team and the EMI team are going to work in junction to extend the planned FTS 3 to be integrated in the gCube framework, by relying on the data access mechanisms extensions under development in iMarine.

## Conclusions

As CERN being involved both in the context of EMI and iMarine project in file and data transfer activities, the integration between the technologies to be implemented has to be considered straightforward. The 2 teams are already in the process of analyzing the requirements and designing the first software architectural draft, a starting point for the adaptation of the existing solution within gCube. The first iMarine project milestone for Data Transfer facilities specification and the first software release are planned for April 2012, and they will be further consolidated till the iMarine project end on April 2014.
Based on this, the conference talk will outline the work done in iMarine for the first data transfer software specification and report the experience of collaborating with EMI on the adaptation of existing services.

## Impact

iMarine will integrate, maintain, extend, and operate the D4Science e-Infrastructure.
This e-Infrastructure will not work in isolation but it will benefit from and rely on resources and services operated by existing infrastructures and service providers of various types. Many of these existing infrastructures are available under the D4Science Federation. This federation of infrastructures is providing access to a large set of multi-type heterogeneous resources coming from different research and scientific domains (humanities,

high energy physics biodiversity, environmental science, and others). Relevant examples of such infrastructures are: GENESI-DR, AquaMaps, EGI, etc.
The implementation of a data transfer service in gCube to be exploited in the iMarine infrastructure, will guarantee members of the iMarine communities to be able to transfer data in a secure and reliable way further crossing the boundaries between e-infrastructures. In addition the collaboration with the EMI and EGI projects (already started during the D4Science-II project) will be further consolidated. By relying on the experience gathered thanks to the MoU between EMI and D4Science-II, which has permitted the exploitation of some of the EMI services (CREAM, VOMS, WMS, DPM) in the gCube software, the FTS service integration would guarantee:
• The availability of the gCube data transfer extension to new communities
• A first integration of gCube technologies in the context of EMI and the possibility to deploy the new FTS 3 with gCube extensions on the EGI infrastructure.

## URL

http://www.gcube-system.org/

## Overview (For the conference guide)

The iMarine EU project, started on the 1st November 2011 as a follow up of the D4Science II project, aims to establish and operate a data Infrastructure supporting the principles of the Ecosystem Approach to Fisheries Management and Conservation of Marine Living Resources.
The Data-Infrastructure enabling technology, gCube, developed throughout three different EU projects (DILIGENT, D4Science I and II) is a software framework which enables the declarative and interactive creation of transient Virtual Research Environments that aggregate and deploy on-demand content resources and application services by exploiting computational and storage resources of grid and cloud infrastructures.
We report on the results of an evaluation of data management solutions from the EMI software portfolio and the experience of collaborating with EMI on the adaptation of existing services to implement data transfer mechanism within the gCube framework, a key element in sustainability for the middleware.

**Primary author:** Mr MANZI, Andrea (CERN)

**Co-authors:** Dr KEEBLE, Oliver (CERN); Dr PAGANO, Pasquale (ISTI-CNR)

**Presenter:** Mr MANZI, Andrea (CERN)

**Session Classification:** Data Technologies

**Track Classification:** Middleware services