Portal – User Interface

The C3Grid Portal acts as a graphical user interface and allows users the access to C3Grid and its resources. It makes the C3Grid metadata catalogue searchable and sends workflow descriptions to the Workflow Scheduling Service (WSS).

Data Retrieval

The Data Information System (DIS) is the catalogue where the C3Grid metadata (ISO 19115) information are stored. It is built upon panFMP, a generic and flexible framework for building data portals independent of metadata formats and protocols. panFMP is based on Apache Lucene and the OAI Protocol for Metadata Harvesting. With the portal a user is able to search and browse for data within the DIS easily. The portal also provides the staging of files and gives the possibility to download the staged results.
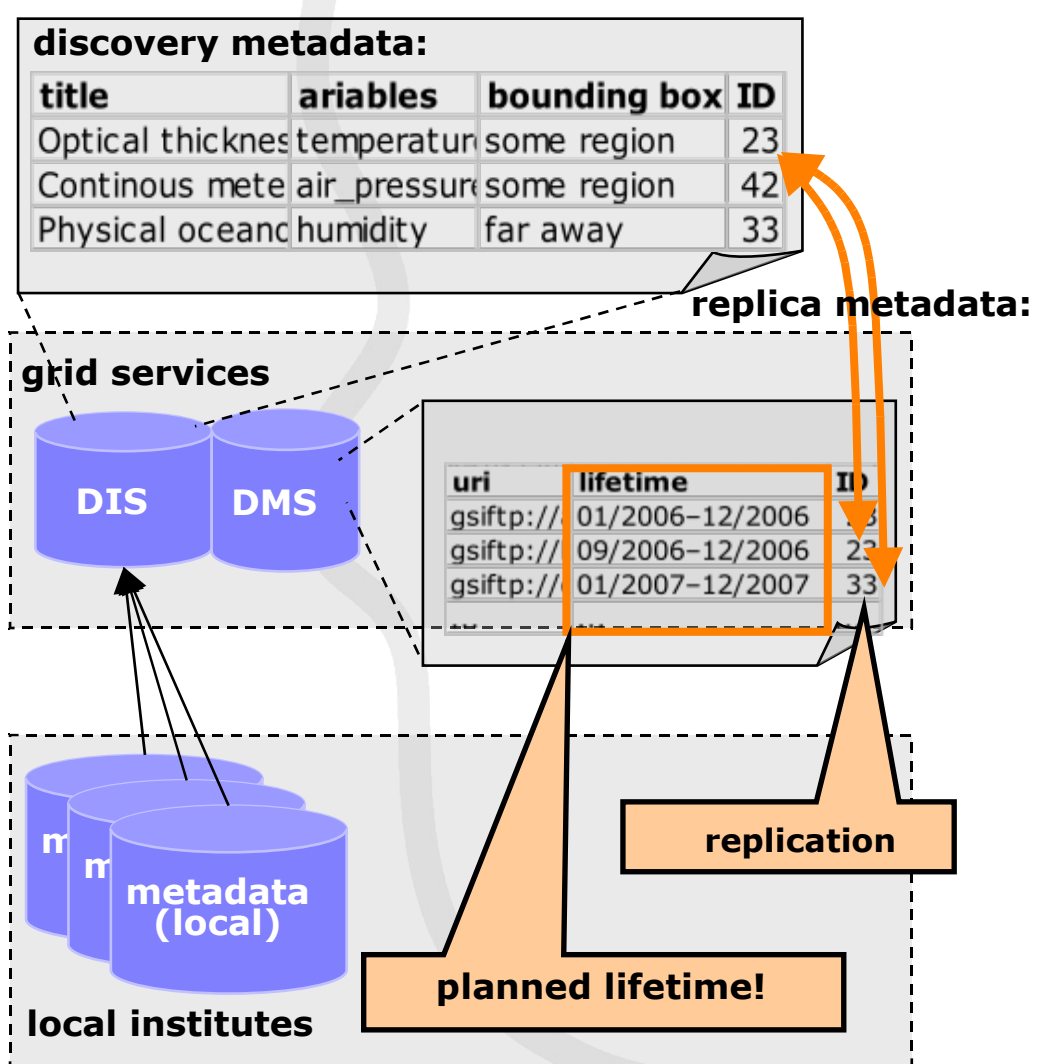
Workflow Submission

The other main application of the portal provides the submission of workflows. There are some specific workflows which contain predefined datasets (this means that they are compatible with the corresponding workflow) and proper parameters. Based on the the selections made by the user the portlet builds a job description object and sends it to the Scheduler's WebService interface. After submission the WSS sends status messages to a notification service (based on WS-Notification) which reacts for example with sending an email to the user to inform him about the current state of his job.

C3Grid Architecture Overview
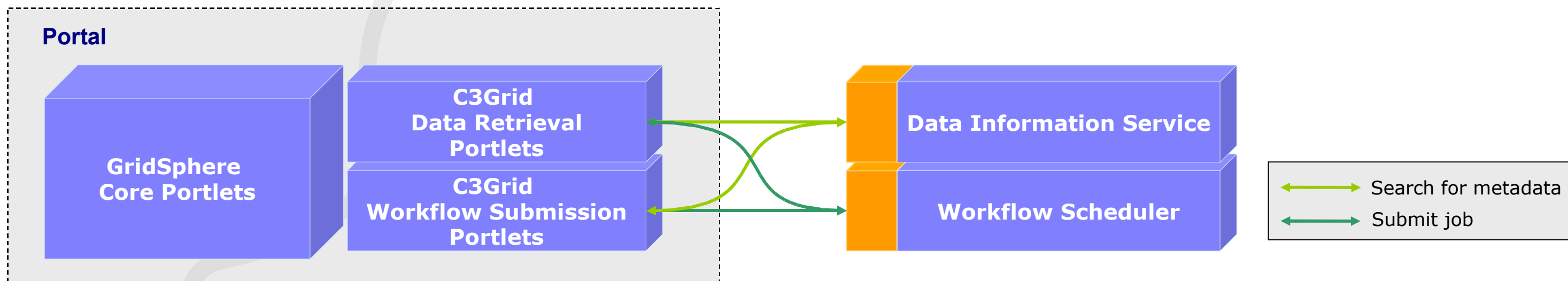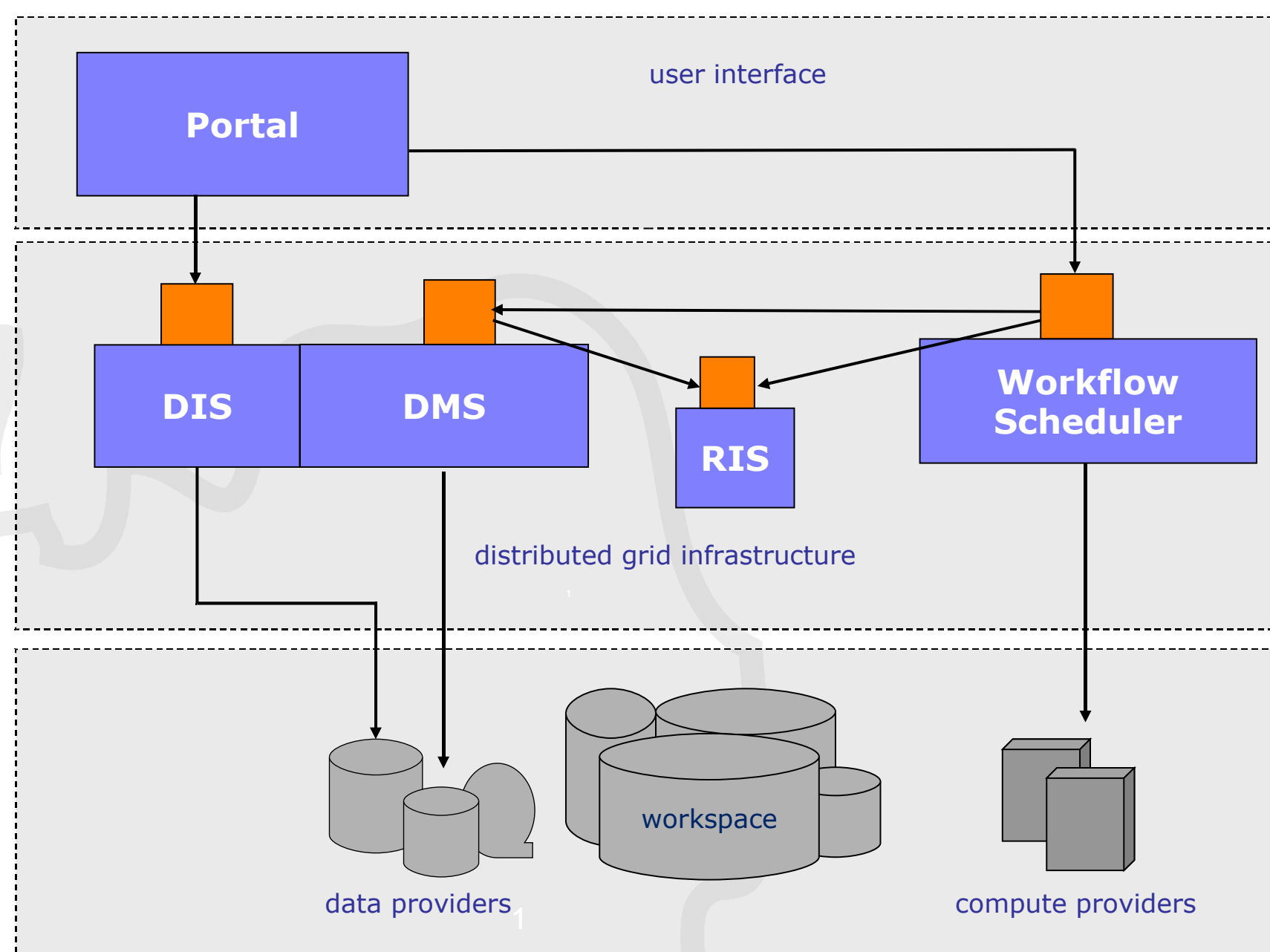
Data Management System

The Data Management System (DMS) stages data from primary sources, manages replicas and inter-workspace data transfers.

discovery metadata:

title	ariables	bounding box	ID
Optical thickne	temperatur	some region	23
Continous mete	air_pressur	some region	42
Physical ocean	humidity	far away	33

replica metadata:

uri	lifetime	ID
gsiftp://	01/2006–12/2006	
gsiftp://	09/2006–12/2006	23
gsiftp://	01/2007–12/2007	33

Import/Export

The DMS provides support for importing and exporting data into and out of the grid

Dynamic Resource Discovery

DMS dynamically integrates new providers using proven Globus Toolkit resource discovery services like MDS.

Implementation:
based on GridFTP, GSI, MDS, Apache Axis and Lucene
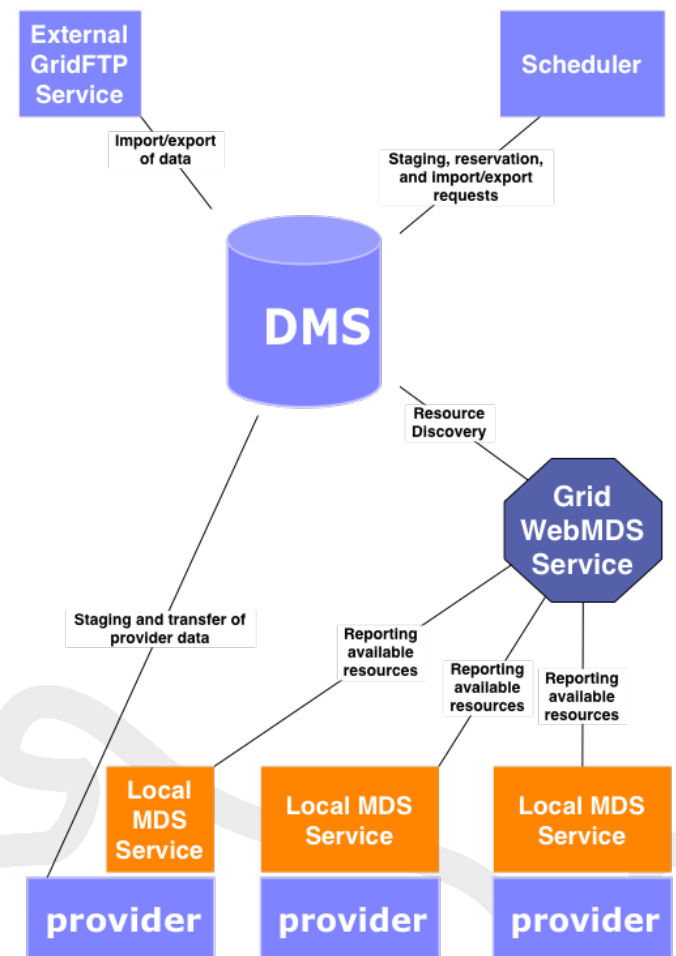
Challenges: Plan the future!
- at which sites will file A exist tomorrow at 11:00 am?
- copy file A from X to Y today between 6 pm and 10 pm
- How long will a transfer take?
- When can file X be restored from tape to disk earliest?

Co-scheduling

The Workflow Scheduler relies on the transfer plans of the DMS for its own scheduling.

Provenance tracking

Answers: How were these data produced? Which data were derived from this?

Workflow Scheduling Service

The Workflow Scheduling Service accepts workflow descriptions from the Portal and coordinates resource allocation and data provision in collaboration with the DMS.

Workflow description

Climate workflows consist of data staging, transfer, and execution tasks which are described in JSDL. The dependencies between those tasks, are given by the proprietary C3Grid Workflow Specification Language (WSL).

Workflow scheduling

Based on the modular workflow concept, the scheduler decides where to transport data and when to execute analysis tasks with respect to the defined task inter-dependencies:
- negotiation-based coordination of data provisioning with the DMS
- the selection of adequate resources for job execution is based on information stored in the Resource Information Service
- the C3Grid Production Environment (CPE) provides a modular, individual, and dynamical environment for user applications
- job execution is done via Globus WS-GRAM

C3Grid Production Environment

Supported by the DEISA modules technologies to dynamically load required software and tools. Capabilities of a resource are published to the Resource Information Service.

Status Messaging to Portal

Implementation of the OASIS WS-Notification standard for workflow and task status reports.