Icecore – a portal for workflow execution

Sami Maisala, University of Helsinki

ADASS XXI, Paris, 9 November 2011

BOF: Scientific Workflows in Astronomy

Workflow approach

- ## Sampo Project for ESO
  - As part of Finland's joining fee a contribution "in kind" of 6 people for three years (18 FTE) was made available. Sampo started in January 2005.

  - Aim: to assess the requirements for ESO data reduction and analysis software infrastructure over the next decade and beyond and produce a series of pilot softwares.

- ## Project result: Pilot of ESO Reflex
  - pilot project was developed on top of Taverna
  - released version is developed on top of Kepler

Problems in traditional desktop oriented data analysis

- Complexity of installing analysis software:
    - Data analysis pipelines for various instruments
    - Collection of data analysis tools
    - Collection of visualization tools.

- Huge amount of data

- Computing resources

- Platform dependency

Sami Maisala, ADASS XXI, Paris

IceCore Portal

- Executes scientific workflows in a workflow engine.

- Controls administrative tasks of workflow execution and resources.

- Manages Databases for workflows and results.

- Distributed System.

Portal elements

- Workflow models

- Workflow instances

- Executions

- Results

IceCore Portal

Sami Maisala, ADASS XXI, Paris

Architectural design

- Three –tier architecture extended with knowledge bus
- REST communication between tiers
- Spring Framework
- Hibernate for mapping of model to a relational persistence database
- JSP Portlets with GWT components

ICECORE ARCHITECTURE

Client1   Client2

Internet

User Tier

IceCore UI (ICUI)
- GWT UI components
- Communication with IceCore

Network

Business Tier

IceCore Controller (ICC)
- Administration
- Registry

Network

IceCore Engine (ICE)

Network

Data Tier

IceCore DataManager (ICDM)
- Service for access IceCore data
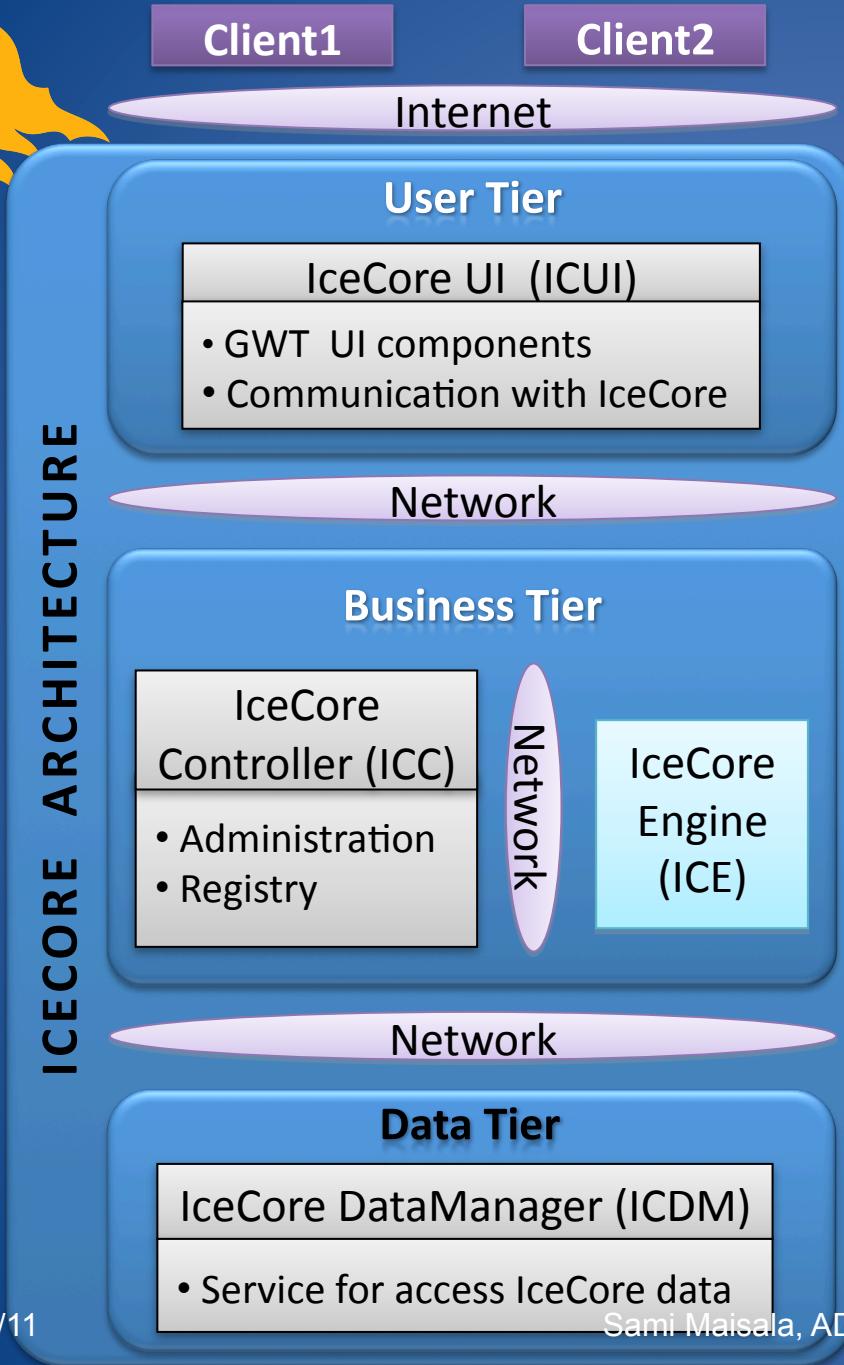
The IceCore system used for workflow execution, follows the three-tier architectural scheme:

User Tier:
The user connects to the system through the portal UI (ICUI), where the execution of various workflows can be monitored or controlled, and data presented to the user.

Business Tier:
ICUI communicates with other components via the IceCore Controller (ICC) middleware component which holds a component registry.
IceCore Engine (ICE) is a web service extension to the execution engine (e.g. Kepler or Taverna).

Data Tier:
On the bottom-level tier IceCore DataManager (ICDM) handles the data management for both ICUI and ICE.

IceCore Knowledge Bus

- Knowledge Bus is a mixture of Web/Message-oriented middleware

- Provides message interoperability between applications which are working in both server-side and client-side.

- Knowledge Bus receives messages from different types of applications and forwards messages to applications which are registered to receive this particular type of data.

Sami Maisala, ADASS XXI, Paris

Distributed system

Client 1

Client 2

Kepler engine 1

IceCore UI Server

IceCore Knowledge bus

Kepler engine 2

IceCore Data Manager

IceCore Controller Server

Sami Maisala, ADASS XXI, Paris
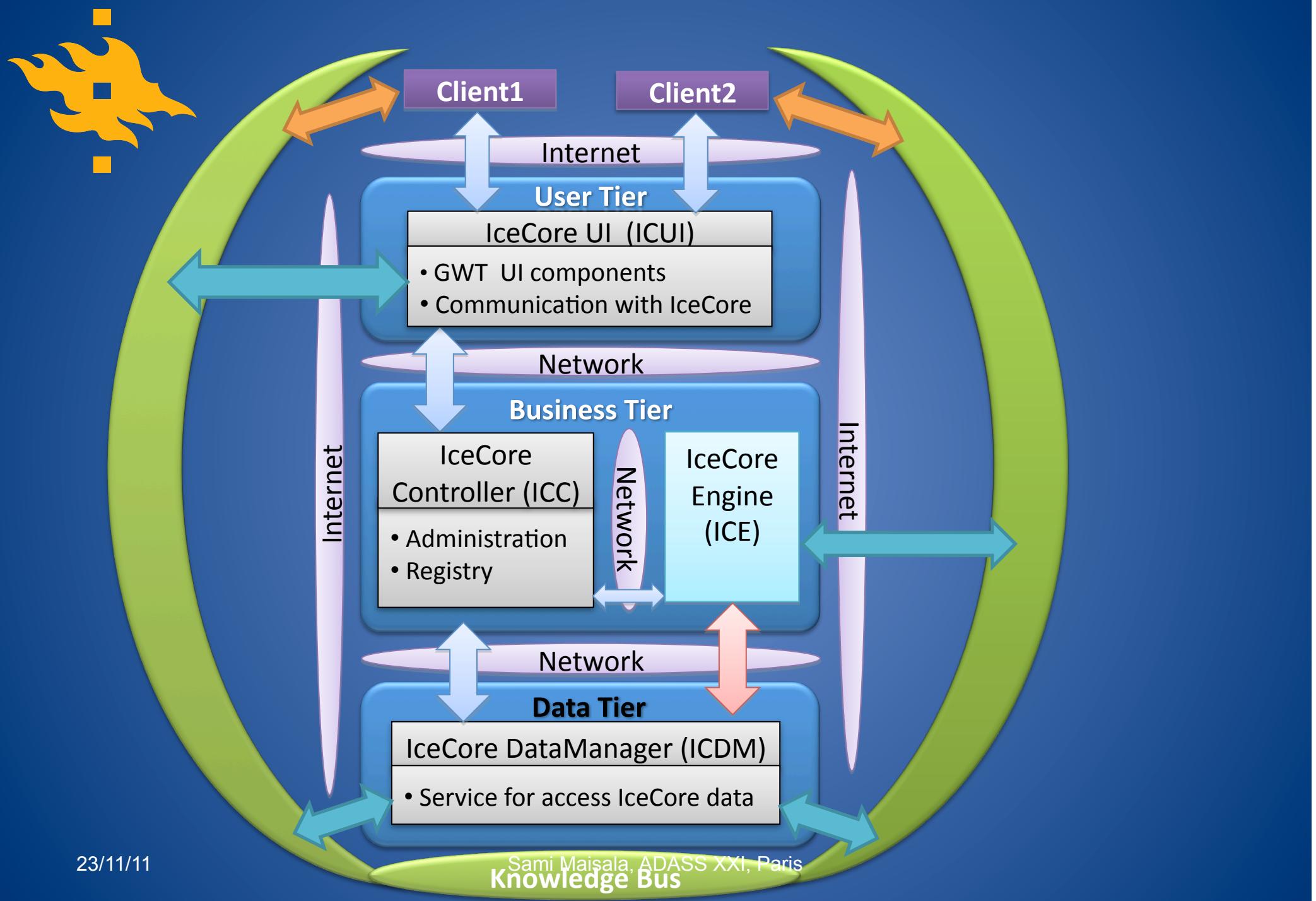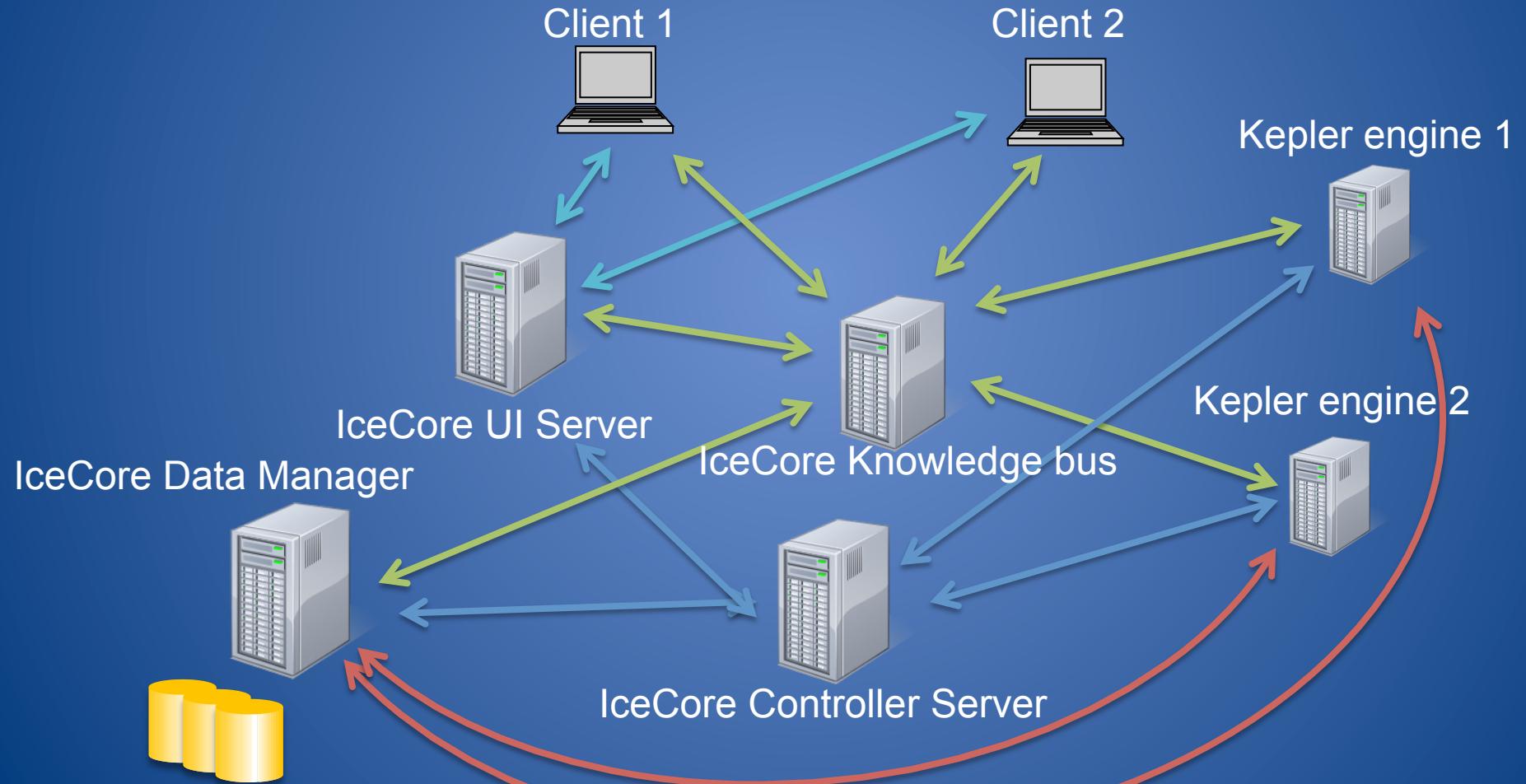
Star Cluster Workflow

This workflow runs an algorithm that aims at locating stellar clusters in the UKIDSS near infrared astronomical database. When the mining has been done, the images of the cluster candidates are fetched from the image database, with SIMBAD database information.

IceCore Portal

Sami Maisala, ADASS XXI, Paris

IceModels portlet

Users can view available models and their description and create new workflows if the descriptive workflow template user can give to the model.

IceCore Portal

IceWorkflows portlet

List of actual workflow instances which are created from existing models. Wizard style portlet will guide user to select workflow to be added in execution list of runnable workflows.

Selected workflow from workflow list and selection of the engine which is used to execute the workflow.

IceCore Portal

IceExecutions Portlet

Displays all available workflows with status information. Each workflow can be executed to chosen workflow engine.

Engine	Workflow	Status	ExeId	Create time	Start time	Stop time
mifsasEngine	ICE dataoutput multi 0	COMPLETE	76	13.3.2011 9:16:27	13.3.2011 9:16:33	13.3.2011 9:16:39
mifsasEngine	testi mifsas	COMPLETE	95	15.3.2011 9:39:02	15.3.2011 9:39:57	15.3.2011 9:40:03
mifsasEngine	ADASS-demo-workflow	COMPLETE	157	8.11.2011 10:51:01	8.11.2011 15:24:23	8.11.2011 15:25:04

1-3 of 3

update status   Execute   ReRun

update status   Execute   ReRun

IceResults

Results can be examined by selecting an application from the list. The list contain... acts as mediat... ...lts portlet.
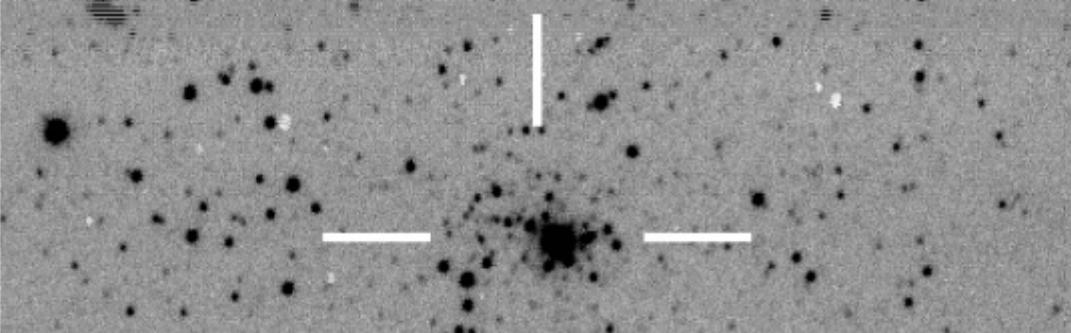
The results can be text, hyperlinks or images.

IceCore Portal

Sami Maisala, ADASS XXI, Paris