RENKEI: REsources liNKage for E-scIence

Kento Aida
National Institute of Informatics
Resource Federation for e-Science

- **e-Science**
  A new scientific research method to process multidisciplinary data using advanced IT in order to achieve new scientific discoveries.

- **Key technology**
  federation of various resources on network, e.g. computers, data storages, databases, applications, ...
Is (current) grid enough for e-Science?

- There still exist barriers for more federation.
  - supercomputer center grid vs. laboratory/department
  - grids with different middleware

Kento Aida, National Institute of Informatics
more resource federation

- Resource federation in a single grid is not enough.
- More resource federation is needed for e-Science.

Goal

- research and development of middleware technology to utilize computing powers, files, DBs, ... on different operational environments, e.g. local machines and grids with different middleware.

Members

- AIST, Fujitsu, KEK, NII, Osaka Univ. Tamagawa Univ., Tokyo Tech., Tsukuba Univ.
RENKEI Project Overview

(1) computing resource federation, and application hosting

(2) file sharing, and file catalogue federation

(3) DB federation, and federation with ID management systems

(4) API for multiple grid middleware

LLS: Laboratory Level System
NIS: National Infrastructure System

Kento Aida, National Institute of Informatics
Difference of Compt. Env.

- We want to run an application on both local machines (in a laboratory or a department) and machines on a supercomputer center grid, but ...
  - The user needs to run jobs on the machines in different way. (no single sign-on)

NII

Kento Aida, National Institute of Informatics
We want to run an application on machines in multiple grids, but ....

- grid environments with difference grid middleware
- The user can run jobs on his/her grid environment but not on other grid.
Computing Resource Federation

- workflow system
  - job submissions to LLS/NIS
- interfaces for grid interoperation
  - job submission and resource information sharing through OGF standards, e.g. HPCBP

Kento Aida, National Institute of Informatics
Application Hosting

- sharing applications in a research community
- deployment and registration of applications by application developers
- running registered application

source: Hitohide Usami, Tamagawa University
We want to access files on a supercomputer center grid from local machines, but ...

- The user needs to manually transfer/copy files between local machines and machines on the grid.
Difference of File Access Method (cont’d)

- We want to access files on multiple grids, but ...
  - multiple file catalog implementations
  - The user needs to learn multiple operations for the file catalog implementations.
File Sharing

- distributed file system
  - high performance file access from LLS/NIS with automatic replica file allocation
  - fault tolerance with automatic file replication

source: Hideo Matsuda, Osaka University
File Catalogue Federation

- federation of file catalogues running on different grid middleware (international interoperation)
  - development of the file catalogue system with RNS (OGF standard)
  - standardization of interfaces in OGF

source: Hideo Matsuda, Osaka University
We want to query data on multiple DBs in multiple research communities, but ...

- multiple DB implementations
- The user needs to learn multiple operations for the DB implementations.
Database Federation

- unified and easy access to multiple DB with different implementation, e.g. RDB, XML, Web, RDF, search engine,…

- high resolution height data from satellites
- geological data
- rainfall data, e.g. AMEDAS
- hazard map
- landslip simulation by using multiple DBs and HPC systems

source: Yoshio Tanaka/Isao Kojima, AIST
Federation with ID Management Systems

- Generation of grid auth. information federating with various ID management systems
  - Username/password, OpenID, Shibboleth, Kerberos, PKI (X.509 certificates)

Source: Yoshio Tanaka/Isao Kojima, AIST
Difference of App. Development Env.

- We want to develop an application running on multiple grid middleware, but ...
  - multiple API implementation
  - The user needs to modify application programs for each middleware.
API for Multiple Grid Middleware

API to develop/run applications on multiple grid middleware with OGF standards, e.g. SAGA and RNS

End users
“Non-GRID”

GRID

RNS Server

UI-CLI

NAREGI-CLI

lLite

NAREGI

Browsers

PC

UG-API (Python Library)

SAGA C++ Engine

Adaptors

Adaptors

Adaptors

WGS

Local

source: Takashi Sasaki, KEK
Evaluation and Community Support

- **testbed**
  - building testbeds with LLS/NIS
- **collaboration with user communities**
  - collecting needs from application users for e-science infrastructure
  - evaluation of middleware technologies developed in the “RENKEI” project and feedback to development teams

source: Manabu Higashida, Osaka University
Summary

More resource federation is needed for enabling e-Science.

Removing barriers for more resource federation.
- local environment (laboratory/department) vs. supercomputer center grid
- grids with different middleware

The middleware development project, “RENKEI”, will remove the barriers.
http://www.e-sciren.org/