Grid Standardization from the NorduGrid/ARC perspective

Balázs Kónya, Lund University, Sweden
NorduGrid Technical Coordinator

ETSI Grid Workshop on Standardization,
May 24, 2006, Sophia Antipolis, France
The NorduGrid Collaboration

- 2001–2002: a research project of the NORDUNet2 program aimed to enable Grid in the Nordic countries
- Since end–2002 NorduGrid is a research collaboration between Nordic academic institutes
  - Open to anybody, non-binding
- Focuses on middleware
  - Develops own Grid middleware: Advanced Resource Connector (ARC)
  - Provides middleware to research groups and national Grid projects
- ARC is now installed on ~50 sites (~5000 CPUs) in 13 countries all over the World

NorduGrid is a collaboration established by five Nordic academic institutes with the purpose of pursuing Grid technology research. The collaboration is at the core of a worldwide Grid Research and Development community that develops, maintains and supports a free Grid middleware, known as the Advanced Resource Connector (ARC).

The aim is to deliver a fault-tolerant, scalable, portable, light-weight and yet fully featured solution for a global computational and data Grid system. This set of tools and services developed by NorduGrid comprises the ARC middleware. ARC is an Open Source software, available for a wide range of Linux systems.

The goals:
- Develop and support ARC middleware
- Coordinate contributions to the core ARC code
- Define strategic development directions following the latest tendencies in Grid technologies
- Promote ARC middleware solutions in such areas as Grid development, deployment and usage
- Contribute to development and spread of worldwide Grid standards
**The NorduGrid Collaboration**

- From ...                  ... To
  - EDG
  - Tesbed
  - HEP
  - 4 Nordic
  - 20 cpu’s
  - 2001
  > >50 sites
  > Bio,Chem,...
  > >13 countries
  > >5000 cpu’s
  > >2003

...from a research project to a research collaboration
...from a Grid testbed to a **major middleware provider**
ARC: components overview

NorduGrid ARC
Middleware Components

Goal: no single point of failure

www.nordugrid.org
**ARC: functionality overview**

- Provides reliable implementation of fundamental Grid services:
  - The usual grid security: single sign on, Grid ACLs (GACL), VOs (VOMS)
  - Job submission: direct or via matchmaking and brokering
  - Information services: resource aggregation, representation, discovery and monitoring
  - Implements core data management functionality
    - Automated seamless input/output data movement
    - Data Indexing (RLS, Fireman), client-side data movement
  - Job monitoring & management
  - Logging service

- Builds upon standard open source solutions and protocols
  - Globus Toolkit® pre–WS API and libraries (no services!)
  - OpenLDAP, OpenSSL, SASL, SOAP, GridFTP, GSI
ARC: most important facts

- General purpose Open Source European Grid middleware
  - Being developed & maintained by the NorduGrid Collaboration
  - Deployment support, extensive documentation
- Lightweight architecture for a dynamic heterogeneous system
- User & performance driven development
  - Production quality software since May 2002
  - First middleware ever to contribute to HEP data challenge
- Middleware of choice by many national academic projects due to its technical merits
  - SWISS Grid(s), Finnish M-Grid, etc...
  - Majority of ARC users are NOT from the HEP community
- Involvement in Interoperability initiatives
  - LCG <-> ARC gateway
- Strong commitment to provide implementations of standards:
  - JSDL, GGF Usage Record support with the coming release

www.nordugrid.org
For simplicity, let’s “forget” most of the major middlewares (UNICORE, Avaki, SRB, GT4, Condor, etc) and take a look at only on the Glite(LCG) vs ARC interoperability issue:
- both are production level middlewares
- both are pre–standard implementations (non WS–based systems)

<table>
<thead>
<tr>
<th>Service/component</th>
<th>Glite (LCG)</th>
<th>ARC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic software stack</td>
<td>GT2 solutions from VDT</td>
<td>pre–WS GT libraries, own patches</td>
</tr>
<tr>
<td>Data transfer</td>
<td>GridFTP, SRM v? (DPM)</td>
<td>GridFTP, SRM v1.1 client</td>
</tr>
<tr>
<td>Data management</td>
<td>EDG RLS, Fireman &amp; Co, LFC</td>
<td>RC, RLS, Fireman</td>
</tr>
<tr>
<td>Information</td>
<td>LDAP, GLUE1.1, BDII–GIIS, R–GMA, interested (2) “Glue2”</td>
<td>LDAP, ARC schema, ARC–GIIS, interested in “Glue2”</td>
</tr>
<tr>
<td>Job description</td>
<td>JDL (based on classAds)</td>
<td>RSL, soon JSDL support</td>
</tr>
<tr>
<td>Job submission</td>
<td>mixture of Condor &amp; GRAM</td>
<td>ARC protocol via GridFTP</td>
</tr>
<tr>
<td>Security</td>
<td>GSI, VOMS, MyProxy, CAS (?)</td>
<td>GSI, VOMS, GACL</td>
</tr>
</tbody>
</table>
Standardization: the NorduGrid perspective

- Interoperability efforts: good starting point in understanding different existing systems and documenting their interfaces
- Fragmentation of the standardization landscape and process
  - by far too many (non-implemented) proposals of numerous Standard Development Organizations
  - integration concerns: can all these proposals be implemented and used together?
- It is desirable that major middleware providers become more dedicated in influencing and implementing emerging standards
  - major players: Globus, Condor (OSG), Unicore, Glite (EGEE), OMII, Chinese middleware(s), ObjectWeb (Proactive), ARC (NorduGrid)
  - there is a threat that Grid community will keep producing non-implemented, conflicting standards
- NorduGrid would like to avoid emerging “de facto standards” defined by implementations
- Interfaces, Interfaces, Interfaces
Standardization: the NorduGrid perspective

- Areas where standardization & implementations are urgently needed:
  - Job description language (JSDL 1.x?)
    - JSDL 1.0 is deliberately narrowly focused
  - Representation of Grid-related objects (Glue2?, CIM?)
    - Computing & storage resources
    - Grid jobs (simple things such as “job states” are not defined!!!)
  - Standard interface to computing resources (execution service)
    - Job submission, Job management, Job monitoring (OGSA–BES?)
  - Standard interface to Storages (SRM?)
  - (pre-deployed) Grid Application environments/frameworks
  - Description of Virtual Organizations
  - Standards related to Grid economy (e.g. Usage info)
    - GGF–UR group declared that “... will not be a Grid usage record”
Conclusion

- NorduGrid Collaboration develops, maintains & supports an open source Grid middleware

- ARC is a reliable, robust, easy-to-use Grid middleware, enabling distributed production facilities already for almost 3 years, non-stop. ARC is a choice of middleware for numerous national Grid projects.

- The NorduGrid Collaboration is involved in and devoted towards Grid interoperability and standardization efforts
  - Currently, only site and user certification is standardized, and to some extent – data transfer