NorduGrid's ARC: A Grid Solution for Decentralized Resources

Oxana Smirnova
(Lund University/CERN)
for the NorduGrid collaboration

GDB, June 22, 2005, CERN
2001–2002: a part of the NORDUNet2 program, aimed to enable Grid middleware and applications in the Nordic countries
- **Middleware**: EDG
- **Applications**: HEP (ATLAS), theoretical physics
- **Participants**: academic groups from 4 Nordic countries
  - **Denmark**: Research Center COM, DIKU, NBI
  - **Finland**: HIP
  - **Norway**: U. of Bergen, U. of Oslo
  - **Sweden**: KTH, Stockholm U., Lund U., Uppsala U. (ATLAS groups)

Since end–2002 is a research collaboration between Nordic academic institutes
- Open to anybody, non-binding

Hardware: mostly rental resources and those belonging to users

Since end–2003 focuses **only** on middleware
- Develops own Grid middleware: the *Advanced Resource Connector (ARC)*
  - 6 core developers, many contributing student projects
  - 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 had strong links with EDG
- WP6: active work with the ITeam; Nordic CA
- WP8: active work with ATLAS DC1
- WP2: contribution to GDMP
- Attempts to contribute to RC, Infosystem

Had to diverge from EDG in 2002
- January 2002: became increasingly aware that EDG
  - Is not suitable for non-dedicated resources with a non-CERN OS
  - Won’t deliver a production-level middleware in time
- February 2002: developed own lightweight Grid architecture
- March 2002: prototypes of the core services in place
- April 2002: first live demos ran
- May 2002: entered a continuous production mode

Since 2004, used by more and more national Grid projects, not necessarily related to NorduGrid or HEP/CERN
ARC Grid

- A Grid based on ARC middleware
  - Driven (so far) mostly by the needs of the LHC experiments
  - One of the world’s largest production-level Grids
- Close cooperation with other Grid projects:
  - SWEGRID, DCGC …
  - NDGF
  - LCG
  - EGEE
- Assistance in Grid deployment outside the Nordic area
- Recently introduced: the ARC Community VO to join those who share their resources
Goals

1. The system must be:
   a) Light-weight
   b) Portable
   c) Non-intrusive:
      • Resource owners retain full control; Grid is effectively a yet another user (with many faces though)
      • No requirements w.r.t. OS, resource configuration, etc.
      • Clusters need not be dedicated
      • Runs independently of other existing Grid installation
   d) Client part must be easily installable by a novice user
      • Trivial tasks must be trivial to perform

2. Strategy: start with something simple that works for users and add functionality gradually
Architecture

- Each resource has a front-end
  - Authenticates users, interprets tasks, interacts with LRMS, publishes information, moves data
- Each user can have an independent lightweight brokering client (or many)
  - Resource discovery, matchmaking, job submission and manipulation, monitoring
- Grid topology is achieved by an hierarchical, multi-rooted set of indexing services
- Monitoring relies entirely on the information system
- Ad-hoc data management, for the beginning
Components

NorduGrid ARC
Middleware Components

Goal: no single point of failure
Components

- **Computing resources**: Linux clusters/pools or workstations
  - Addition of non-Linux resources is possible via Linux front-ends

- **Front-end**:
  - Runs custom pluggable GridFTP server for all the communications
    - Accepts job requests and formulates jobs for LRMS/fork
    - **Performs most data movement** (stage in and out), cache management, interacts with data indexing systems
    - Manages user work area
  - Performs all kinds of job management upon client request
  - Publishes system and job information
Components

- **Client**: a lightweight *User Interface* with the built-in Resource Broker
  - A set of command line utilities
  - Minimal and simple
  - Under the hood: resource discovery, matchmaking, optimization, job submission
  - Complete support for single job management
  - Basic functionality for multiple job management
  - Support for single file manipulations

- Portals and GUI clients are being developed
**Information System**: based on Globus-patched OpenLDAP: it uses GRIS and GIIS back-ends

- Keeps strict registration hierarchy
- Multi-rooted
- Effectively provides a pseudo-mesh architecture, similar to file sharing networks
- Information is only kept on the resource; never older than 30 seconds
- Own schema and providers
Components

- **Storage**: any kind of storage system with a disk front-end
  - Own GridFTP server implementation with pluggable back-ends
    - Ordinary file system access
    - Grid Access Control Lists (GACL) based access
  - *Smart* Storage Element – WS based data service with direct support for Indexing Services (Globus’ RC, RLS)
  - tape storage systems are being acquired
Functionality

- Single-login access to a heterogeneous, distributed set of computing and storage resources
  - Based on pre-WS Globus libraries and API
- Complete up-to-date information on the available resources
- Serial batch job submission to best resources available
  - Matchmaking, brokering
- Job monitoring and management
- Seamless input/output data movement
- Basic data management
  - Indexing, movement
Features

- Light-weight
- Portable
- Dynamic, heterogeneous
- Non-intrusive:
  - Resource owners retain full control
  - No requirements w.r.t. OS, resource configuration, etc.
  - Clusters need not be dedicated
  - Runs independently of other existing Grid installation
- Client part is easily installable by a novice user
- Optimized for large data processing, bulk job submission
- Simple monitoring of the entire system and each job
- Specialization: oriented towards serial batch jobs
  - Parallel jobs are perfectly possible, but only within a cluster; no optimization
  - Interactive tasks are not supported
  - No optimization for “Hello World” short jobs
Distribution, availability

- At ftp.nordugrid.org:
  - Stable releases, including:
    - Binary RPMs and tar-balls are available for most Linux platforms
    - Source RPMs
    - Standalone client tar-ball for installation by a non-privileged user
      - Only 13 MB when unpacked
      - Contains all the EU Grid PMA approved CA keys
      - Includes all the basic Globus client tools
  - Weekly development builds
  - Nightly builds

- CVS at cvs.nordugrid.org
- License: GPL
- More info, complete documentation, contacts at www.nordugrid.org
On interoperability

- **Some historical perspective**
  - **1988:** Condor. No data management, no security infrastructure
  - **1993:** Legion/Avaki. Also calculation-oriented
    - First object-oriented model; commercialized
  - **1997:** UNICORE. No data management still
  - **1997:** SRB. Only data management
    - Source unavailable
  - **1998:** Globus Toolkit 1. GSI, GridFTP, but still no data management, very mixed success, MDS – a bottleneck
    - Was about to die, but received some appreciation in Europe and became a de-facto standard:
  - **2001:** EDG. Makes use of GT2, Condor. Some data management (RC, GDMP)
  - **2002:** VDT. A “delivery channel” offering GT2, Globus etc – no data management
  - **2002:** NorduGrid/ARC. Makes use of GT2 libraries and data management (RC). Implements Web Services for some components
  - **2002:** OGSA. Object-oriented approach, modularity, **STANDARDS**
    - UNICORE quickly moves towards OGSA standards
  - **2003:** LCG–1. Largely inherits EDG
  - **2003:** Grid3. Based on VDT. No data management
  - **2005:** GT4. Implements Web Services for some components
  - **2005:** gLite. EDG–line, but implements Web Services for some components, comprehensive basis for data management
  - ... What more?...

- **Note:** most of the mentioned solutions are incompatible with each other
Some quotes:

- Rob Baxter (EPCC, NeSC), October 2002:
  - “do we have a real chance of a single, uniform architecture for distributed systems integration?” (speaking of OGSA)

- Mark Linesch (GGF Chairman), June 2005:
  - “OGSA is in the early stages of development and standardization“
  - “GGF distinguishes between the OGSA architectural process, OGSA profiles and specifications, and OGSA software. All of these are important to maintain coherence around OGSA and Grid standards”
  - “At the time of writing, we have OGSA Use Case documents, OGSA Architectural documents and drafts of an OGSA Basic Profile document and OGSA Roadmap document. We do not yet have any OGSA–compliant software implementations or OGSA compliance tests”
On interoperability–3

- For simplicity, let’s “forget” about UNICORE, Avaki, SRB, GT4 etc
  - But for how long?
    - UNICORE enjoys EU support, adhere to OGSA – maybe this is the standard to follow?
    - GT4 has improved security infrastructure, Reliable File Transfer etc – perhaps this is the standard?
  - LCG–2 and gLite are much closer to ARC than to e.g. UNICORE

- Short list of “our” services and solutions:

<table>
<thead>
<tr>
<th>Service/component</th>
<th>LCG–2, gLite</th>
<th>ARC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basis</td>
<td>GT2 from VDT</td>
<td>GT2 own patch, GT3 pre-WS</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, MDS+BDII, R–GMA</td>
<td>LDAP, ARC schema, MDS–GIIS</td>
</tr>
<tr>
<td>Job description</td>
<td>JDL (based on classAds)</td>
<td>RSL</td>
</tr>
<tr>
<td>Job submission</td>
<td>Condor–G to GRAM</td>
<td>GridFTP</td>
</tr>
<tr>
<td>VO management</td>
<td>VOMS, gLite VOMS, CAS (?)</td>
<td>VOMS</td>
</tr>
</tbody>
</table>
On interoperability–4

Notes:
- Cooperation between ARC and Condor led in October 2004 to Condor–G version that can submit jobs to ARC GridFTP (translation from ARC infosystem schema to GLUE was developed by Rod Walker). Was meant to be used by LCG – but nobody configured an RB this way yet
  - ARC does not use Condor components, can not submit jobs vice versa
- GLUE2 schema is expected to be developed soon, with participation of NorduGrid, OSG and others. All chances to get a common resource representation.
- “Rome” Common Resource Management initiative (includes Globus, UNICORE, LCG, EGEE, NorduGrid, NAREGI) converged on usage of GGF JSDL for job description
  - JSDL v1.0 is still rudimentary, but is the least common denominator

Bottom line:
- LCG and NorduGrid are quite close already and have all chances to get closer
  - It might make sense to move together to a worldwide standard in cooperation with other Grid projects, instead of inventing new ways or converging to each other
  - Respecting GGF recommendations (OGSA, SRM, JSDL etc) might be a good starting point
  - Apparent substantial differences between gLite 1.1 and LCG–2, plus non-standard SRM are confusing – in the interoperability perspective.
Conclusion

- NorduGrid’s ARC is a reliable and robust Grid middleware, supporting distributed production facilities already for almost 3 years, non-stop
- The middleware is in development, everybody is welcomed to use and contribute
- ARC is meant to offer Grid solution for decentralized, opportunistic resource usage:
  - Using ARC does not give an automatic access to any resource: please negotiate with the resource owners (create Virtual Organizations)
  - Deploying ARC does not open doors to all the users: only resource owners decide whom to authorize
- ARC developers are deeply interested and willing to take part in global Grid standardization and interoperability efforts
  - Currently, only site and user certification is standardized, and to some extent – data transfer