NorduGrid

3 years of building Grid-like infrastructure in Nordic countries

Presented by
Aleksandr Konstantinov on behalf of NorduGrid collaboration
Vilnius University/Lithuania and University of Oslo/Norway

GlobusWORLD 2005
February 7–11, Boston, Massachusetts, USA
NorduGrid

- **NorduGrid** is a research collaboration established by universities in Denmark, Estonia, Finland, Norway and Sweden
  - Focuses on providing production-capable Grid-like middleware for academic researchers
  - Currently supports one of the largest Grid production systems
    - 10 countries, 40+ sites, ~4000 CPUs, ~30 TB storage
ARC (Advanced Resource Connector) is the Grid middleware developed by the NorduGrid

- Collection of tools and services
- Based on Globus Toolkit™ 2 libraries and services
  - Can be built using GT3 pre-WS code too
  - Using available utilities whenever possible
  - Some services which could/can not provide functionality outlined in NorduGrid architecture were replaced
  - Some other services not provided in Globus Toolkit™ 2 were developed

Initial development principles:
- Simple
- Stable
- Non-invasive
ARC components

- **Job management**
  - Lightweight **User Interface** with built-in Personal Resource Broker
    - Basic and complete support for single job management
    - Basic functionality for data management
  - Resource frontend – **Grid Manager** – accessible through GridFTP interface

- **Information**
  - Information System based on Globus MDS 2 with a modified schema
  - WS based logging service – Logger

- **Data management**
  - Supported Indexing Services include legacy Globus' **Replica Catalog (RC)** and **Replica Location Service (RLS)**
    - All operations involving Indexing Services are done automatically
    - Looking for better solutions
  - GridFTP server implementation with pluggable backends
  - "**Smart**" Storage **Element (SSE)** – Web Service based data service with direct control of Indexing Services
  - Every piece of code has support for full set of protocols

- **System monitoring**
  - Web interface to current state of the system – Grid Monitor
  - System's history and statistics – NGLogger
• Solution similar to Globus' gatekeeper reimplemented in order to have needed functionality
• GridFTP interface for job control
  – Each job is presented as virtual subdirectory
  – FTP commands are mapped to job management operations
• Handles pre- and post-staging of files with integrated support for data indexing services (RC, RLS).
• Shared cache of pre-staged files with automatic registration in Indexing Services
• Application-specific Runtime Environments
• Limitations due to architecture – data staging only at beginning and end of job
**ARC - Information System**

- Uses Globus’ MDS 2.2
  - Soft-state registration allows creation of dynamic structure
  - Multi-rooted tree
  - GIIS caching is not used by the clients due to buggy implementation
  - Several patches and bug fixes are applied
- Mostly cluster-oriented
  - A new schema and information providers were developed, to serve clusters
- All queries are anonymous
  - Authenticated queries are very inflexible
- Not very scalable
- Looking for new solution now

---

**NorduGrid Hierarchy**

- Top Level: GIIS, GIIS, GIIS, GIIS
- Country Level: Users, Jobs, Jobs, Users, Jobs, Users, Jobs, Users
- Host Level: User-1, User-2, User-3, User-2, User-2, User-1, User-1, User-1
• Own implementation of GridFTP server
  – Protocol is implemented using globus_ftp_control library from Globus Toolkit™
  – User-dependent virtual tree of directories with flexible access rules evaluated against
    • Grid identity of the user
    • Virtual Organization Membership Service credentials
    • Any external module
  – Local file/object access is through pluggable modules
    • fileplugin – ordinary file access with static access rules
      – Based on Grid–UNIX identity mapping
      – Based on Grid identity only
    • gaclplugin – access to each object is controlled through GACL object maintained by object's owner
    • jobplugin – control/access user's job
ARC – "Smart" Storage Element

- Uses HTTPS/HTTPG + SOAP
  - Firewall friendly
- Integrated flexible access control
  - Per stored object
  - Evaluated against Grid identity of a client
- Direct interface to data indexing services
  - Indices are kept more consistent

- Data transfer tasks
  - Integrated support for data replication
- SRM v2 interface being developed
  - Still waiting for any SRMv2 enabled client
ARC – User Interface

• Collection of command line utilities:
  - `ngsub` – find suitable resource and start job
  - `ngstat` – check status of job
  - `ngcat` – monitor job by looking at its stdout/stderr
  - `ngget` – get results of finished job
  - `ngkill` – stop job
  - `ngclean` – delete job from computing resource
  - `ngsync` – find user's jobs
  - `ngrenew` – update remote credentials
  - `ngls` – list files on storage element or in job's directory
  - `ngcopy` – data transfer
  - `ngrequest` – third-party transfers or data tasks
  - `ngremove` – delete remote files
ARC – User Interface (cont.)

- Contains a Personal Resource Broker – job submission sequence:
  - Collects information about Computing Resources (clusters) through MDS network
  - Collects information about requested input data (size and availability at Computing Resources if available)
  - Collects Information about resources at Computing Resources
  - Selects suitable Computing Resource and submits job request
    - The user must be authorized to use the cluster and the queue
    - The cluster’s and queue’s characteristics must match the requirements specified in the RSL string (max CPU time, required free disk space, installed software etc.)
    - If the job requires a file that is registered in a data indexing service, the brokering gives priority to clusters where a copy of the file is already present
    - From all queues that fulfills the criteria one is chosen randomly, with a weight proportional to the number of free resources or shortest queue
  - Uploads locally available input data
ARC – Grid Monitor

• Web based interface captures current state of the system
  – Implemented in PHP
  – Very rich interface
  – From statistics to detailed view
    • Summary per cluster
    • Jobs per cluster
    • Jobs per user
    • etc.

2005-02-09 www.nordugrid.org
ARC – Logger and NGLogger

• Logger
  – Central MySQL database with WS interface
  – Information is pushed from Grid Managers

• NGLogger
  – Web interface to Logger's database
  – Complementary to Grid Monitor
  – Shows statistics computed from history
    • Can be tuned for particular project
    • Provides statistics per cluster
    • Provides statistics per user
ARC – Ongoing efforts

• arcLib – effort to create well defined set of C++ classes to interface ARC-enabled resources
  – Currently being developed
  – Already minimal functionality available
  – Will provide all UI functionality in a flexible and portable way
  – Pluggable brokering algorithms

• jarcLib – set of Java classes
  – Because ARC is mostly written in C/C++ Java developers had problems with adding ARC support into their products
  – Implements only basic functionality
  – Uses Java CoG – causes problems due to incompatibility with mainstream Globus Toolkit™

• Distributed data and replicas indexing service.
ARC – Middleware status

• Current stable release: 0.4.4 (2004–10–18)
  – GPL license
  – Binaries available for 12+ Linux distributions
  – Builds on top of Globus Toolkit 2 and 3 – NorduGrid-patched Globus Toolkit is preferred for stability
  – EDG VOMS integrated (voms–1.1.39–5ng)
  – Being maintained

• Current development series: 0.5.x
  – Contains the “Smart” Storage Element and other newly introduced features
  – Quite stable already – some sites are using it

• Anybody is free to use; best–effort support is guaranteed
  – Support: nordugrid–support@nordugrid.org
  – Bug reports: http://bugzilla.nordugrid.org

• Any constructive contribution is welcome
  – Join nordugrid–discuss@nordugrid.org (very busy list!)