NorduGrid: Collaboration status update

Oxana Smirnova (Lund, NDGF/KnowARC/NGIn/NorduGrid)
OSG Consortium All Hands Meeting
March 6, 2007
Outline

- ARC middleware status
- Overview of NorduGrid–spawned projects
  - EU KnowARC, NDGF
- Interoperation and standardization
- **NorduGrid**: a research collaboration
  - Based on MoU
  - Supports ARC m/w development via different projects
- **ARC**: Advanced Resource Connector, the middleware
  - GPL software, coordinated by NorduGrid
- **KnowARC**: an EU FP6 R&D project (STREP)
  - Core contributor to ARC development
- **NDGF**: The Nordic DataGrid Facility, a Nordic infrastructure and R&D project
  - Builds the Nordic Tier-1
  - Contributes to ARC development
- **NGIn**: Innovative Services and Tools for NorduGrid, an educational and R&D project by Nordunet3
- **NGN**: Nordic Grid Neighborhood, a networking project
  - ARC community in Nordic countries, Baltic states and Russia
December 2006: ARC 0.6 is out

- A major milestone: first stable release since 2004
- New external software, new packaging and distribution
  - Globus 4 based (pre-WS components, e.g. GSI, gridftp libraries, LDAP backends)
  - Latest VOMS, GSOAP
  - Natively works on 64bit architectures
  - Streamlined configuration
  - Software repositories (e.g. apt, yum)
- Major re-write of the client code base
  - Modularity
  - Multithreaded implementation
  - Extensibility
- Client libraries, modules available for C++ (native), Python, Perl etc
  - via SWIG, and Java
  - Easy to use API for clients that need to interface to ARC, e.g. GUIs, portals, CLIs
- Many new attributes in the information system
  - Better monitoring and decision-making
  - Improved job description
More new ARC 0.6 features

- Even more extensions on the server side:
  - Authorization, security: VOMS, MyProxy support, GACL for jobs
  - JSDL support
  - SRM support (also in the data movement client)
  - LSF support; PBS, SGE and Condor interfaces significantly improved
- “Smart Storage Element” (SSE) is a part of the release
  - Files instantiated at an SSE are registered in e.g. RLS automatically
  - Has a basic SRM interface
- Non-root ownership of services
- Performance improvements of the core services
- Logging infrastructure: new Usage Record format, better performance
- Improved monitoring
- Localization of clients (user interface, monitor – 8 languages)
- Improved usability: notifications, logs, command line options
What is still missing

- ARC 0.6 is still based on 5-years old technologies
  - Too many non-standard dependencies, esp. pre-WS Globus & Co
  - Hit the limits of extensibility
- Desirable functionalities:
  - Execution services:
    - Job migration
    - Semi-interactive applications, parallel jobs
    - Fine-grained policies, priorities (VO-based)
  - Data management:
    - Reliable data indexing
    - Support for data collections
    - Fine-grained, consistent access control (VO-based)
    - SRM interface to distributed storage facilities
  - Monitoring, logging and accounting
    - User-friendly interface to logging/accounting systems
    - Lightweight logging/accounting
    - Security layer over information and monitoring services (VO-based access)
  - Advanced clients
    - GUI: a Java-based prototype (the “Arconaut”) exists
    - Clients on popular systems like MS Windows, Mac OS
    - More versatile brokering
    - “Off-line” job management functionality (“babysitting”)
- Standards compliance, interoperability
  - Still a lot of moving targets
The KnowARC project

- EU FP6 Specific Targeted Research Project
  - "Grid-enabled Know-how Sharing Technology Based on ARC Services and Open Standards"
  - Duration: June 1 2006 – June 1 2009
  - Partners: NorduGrid members and research teams in medicine, bioinformatics, physics, engineering, automotive industry apps, IT (10 partners from 7 countries)

- Objectives:
  - to create a novel, powerful Next Generation Grid middleware based on ARC, widely respected for its simplicity, non-invasiveness and cost-efficiency;
  - to promote Grid standardization and interoperability;
  - to contribute to Grid technologies take-up, bridging the gaps between business and academia in Grid development

- Will develop the middleware that will be the next step after ARC 0.6, addressing current limitations and shortcomings
KnowARC today

- Design stage resulted in a number of documents, see [http://www.knowarc.eu/documents/](http://www.knowarc.eu/documents/)
- October 2006: Standards Conformance Roadmap
  - Overview of existing and emerging standards
  - Highly relevant standards are: OGSA, OGSA WSRF Basic Profile, JSDL, GridFTP v2, OGSA–BES, OGSA Security Basic Profile – secure channel, GLUE, X.509 PKI, WS–I Basic Profile, WS–I Basic Security Profile, SRM
- January 2007: Design Document is released
  - Lays out high-level design principles
  - Elaborates on key capabilities
    - Hosting Environment
    - Information Capability
    - Execution Management Capability
    - Data Capability
    - Resource Management Capability
    - Security Capability
    - Self Management Capability
- Coding has started
  - Hosting Environment Daemon, Dynamic Runtime Environments
NDGF == “Nordic Data Grid Facility”
- Idea conceived by NorduGrid members in 2002
- Goal: create a Nordic Grid infrastructure, primarily for LHC Grid computing (Tier1)
- 2003–2006: pilot project funded by the 4 Nordic countries (Denmark, Finland, Norway, Sweden)
- NorduGrid/ARC middleware chosen as the basis

June 1st 2006: NDGF is launched
- Nordic production Grid, leveraging national grid resources
- Common framework for Nordic production Grid
- Co-ordinates & hosts major Grid projects (e.g. the Nordic LHC Tier-1)
- Develops Grid middleware (ARC contributor)
- Single Point of Entry for collaboration, middleware development/deployment, e-Science projects
- Represents the Nordic Grid community internationally

NDGF 2006–2010
- Funded (2 MEUR/year) by National Research Councils of the Nordic countries (NOS–N)

NDGF coordinates activities – does not own resources or middleware
NDGF’s distributed Tier1 under construction
NDGF resources

- **People:**
  - 5 managers
  - 4 National Coordinators
  - 6 Middleware Developers
    - VOBoxes integration
    - Distributed dCache development
    - Interoperation
  - Plus many helpful local sysadmins and power users

- **Hardware:**
  - Large part of what is known as NorduGrid
    - Shared heterogeneous resources running ARC
  - Some own resources (servers)
  - Dedicated WLCG hardware is on its way
    - Still shared, heterogeneous, and running ARC
Both are of key importance for NorduGrid and all the related projects
- Interoperation will work when based on widely accepted well documented open standards
- Neither should compromise desired functionality and performance

NorduGrid, KnowARC and NDGF are all involved in various related activities
- GIN activities
- JSDL
- Usage Record, Resource Usage Service
- Information–related standardization
  - OGF Glue2 group
  - Hosted Grid Service Information Discovery Meeting in December
  - Contacts with U. of Iowa group
- GridFTPv2 contribution
  - Triggered by needs to set up a distributed dCache => dCache code contribution
- Some gateway– and portal–based approaches in the pipeline

Pre–WS Globus used by major Grid projects is a mature but ageing technology
- OGF is becoming a true standardization body
- Let’s move together from de facto to de jure standards!