NorduGrid’s ARC middleware: status and future

Balázs Kónya
Lund University, NorduGrid Collaboration
ARC meets SwiNG, 25th June 2008, Bern
Outline

- Brief history of the middleware
- Current production ARC: "ARC Classic (ARC0)"
  - Overview of main components
  - ARC Classic limitations
  - Status and plans
- Next generation ARC: "ARC1"
  - What is new? What is/will be inside the box?
    - A-REX service
  - Backward Compatibility
  - Interoperability
  - Standard conformance
  - Status & Plans
Advanced Resource Connector
in a nutshell

- General purpose Open Source European Grid middleware
  - One of the major production grid middlewares
  - Developed & maintained by the NorduGrid Collaboration
  - Deployment support, extensive documentation, available on most of the popular Linux distributions

- Lightweight architecture for a dynamic heterogeneous system following Scandinavian design principles
  - Start with something simple that works for users and a functionality gradually
  - Non-intrusive on the server side
  - Flexible & powerful on the client side

- User- & performance-driven development
  - Production quality software since May 2002
  - First middleware ever to contribute to HEP data challenge

- Strong commitment to standards & interoperability
  - JSDL, GLUE, Active OGF player

- Middleware of choice by many national grid infrastructures due to its technical merits
  - SweGrid, SWISS Grid(s), Finnish M–Grid, NDGF, etc...
  - Majority of ARC users are NOT from the HEP community

Illustrations: “Scandinavian Design beyond the Myth”
www.scandesign.org
Brief history of the middleware

- **Y2K**: Grid Hype, European Data Grid (EDG), re-discovery of Globus Toolkit (version 1.1.4)
- Back in 2001...HEP Institutes from Scandinavia wanted to share their computing resources and jointly contribute to CERN/LHC computing
  - The born of the “NorduGrid”, a research project of the NORDUNet2 program aimed to enable Grid in the Nordic countries
- **2002 February**: decision to develop an alternative middleware by making use of Globus libraries. NorduGrid design, architecture, philosophy.
- **2002 May**: 3rd NorduGrid Workshop, Helsinki demonstration of the first version of the middleware.
  - Since then the NorduGrid middleware has been used in production, first middleware ever to contribute to a production HEP data challenge.
- **2004 April**: announcement of release 0.4 of NorduGrid middleware (also known as the Advanced Resource Connector), the first official release of this software.
- **2006 June**: Development of the next generation ARC via the KnowARC project started.
- **2007 May**: After a long hardenning process ARC version 0.6, the second stable release of the middleware was released.
- **2008 June 17**: Version 0.6.3 of ARC Classic released
- **2008 August**: a prototype version of the next generation ARC is expected

All over the years: ARC has become one of the major grid middlewares used in production all over the world.
ARC Classic: 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
  - Job monitoring & management
  - Information services: resource aggregation, representation, discovery and monitoring
  - Implements core data management functionality
    - Automated seamless input/output data movement
    - Interfacing to Data Indexing, client-side data movement
    - Storage Elements
  - 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
Computing resources: Grid-enabled via ARC layer on head node (front-end):
- Custom GridFTP server for all the communications
- Grid Manager handles job management upon client request, interfaces to LRMS
- Performs most data movement (stage in and out), cache management, manages user work area
- Publishes resource and job information via LDAP
ARC Classic 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
  - Built upon ARCLIB

- Portals and GUI clients are being developed
ARC Classic components

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
- **Storage:** any kind of storage system with a disk front-end
  - **Conventional Storage:**
    - 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)
  - no tape storage systems in use so far
ARC Classic: limitations

- Custom, non-standard based service interfaces
  - ARC does not (will not) work together with other Grids
- Non-service oriented architecture, monolithic software, nordugrid-dependent data models
  - Middleware-dependent client-side library
- Strong Globus dependency
  - limited portability
  - Firewall issues
  - Support issues
- Security limitations:
  - Clumsy and poorly documented configuration
  - Security concerns due to grid mapping, root services
  - Non-perfect isolation of Grid jobs running on the same resource
  - No security layer in Information System
  - No or very limited support for grid economy
- Usability concerns
  - No support for bulk job management, one-job-focus
  - Managing Software Environments (Applications) is a pain
  - Poor support for parallel applications
  - Why can not I use the grid from Windows?
  - What do I gain running ngsub instead of qsub?
  - Too complex for end-users: certificates, XRSL, GACL syntax, new set of commands
  - No easy-to-use GUI (only a prototype)
- Limited data management
  - Clumsy input/output data management
  - No coherent and consistent data management
- Scalability concerns
  - Data staging process, information monitoring, application deployment
  - Support for clusters with ten thousands of CPU cores

Read more: Section 4.2 "Strengths and limitations" in "Design document"
ARC Classic: status and plans

- ARC is a matured software which has proven its strengths in numerous areas
- Production release 0.6.3 is out
  - Stability improvements, bug fixes
  - LFC (file catalogue) support
- ARC faces a scalability challenge posed by "ten–thousand–core" clusters
  - Thursday meeting
  - Cache redesign
  - Uploading/downloading, increased load on frontend
  - Local information system
- "projects" on top of ARC Classic
  - Job manager
  - Automated software installation: Janitor
  - Lunarc portal
  - GUI: ARCONAUT
  - "Get rid of Globus": use native software (openssl, openldap)
- Release plans based on ARC Classic
  - 0.6.x stability releases will be periodically released
  - Preliminary planning for an 0.8 release incorporating new major features (e.g. Janitor)
- Migration plan to ARC1
  - gradually replace components with ARC1 modules initially co-deploying both versions.
What are the main differences?

- **ARC Classic**
  - Robust solution
  - Non-standard interfaces
  - Not modular
  - Limited portability
  - Difficult to ride

- **ARC1**
  - No-crazy design (will not be an extreme middleware)
  - Early prototype
  - Standard interfaces
  - Modular structure
  - Windows, MAC OS
  - *Improved performance*
    - *building on ARC Classic*
  - *Enhanced user experience*
ARC1 Design: service decomposition

- Job Manager (job pool)
- Job Manager (caretaker)
- Policy Decision Point
- Selection Service
- Selection Service
- Policy Decision Point

Detailed services include:
- A-Rex
- Authorization
- Information Registration Service
- Janitor
- VMMS
- Information Registration Service
- Hosting Environment
- External Information Link service
- Monitoring Accounting Recording Service
- Monitoring Service
- Information Indexing Service
- RTE Catalog
- Storage Manager
- Storage Manager
- Storage Manager
- Storage Element
- Storage Element
- Storage Element
- Data Catalogue
- Data Catalogue
- Data Catalogue

This diagram illustrates the architecture and services involved in the ARC1 design, emphasizing the integration of various components and their interconnections.
What is/will be inside the ARC1 box?

- **Already inside the box**
  - **HED framework**
    - Flexible service hosting and development framework
    - Takes care all the networking-layer (e.g. SOAP)
    - Available on windows and mac–os as well
    - Offers python and java language bindings in addition to the native c++
  - Test service (echo)
  - A–REX (computing element)
  - Central ISIS (information indexing service)
  - Storage System services
  - Security framework
    - Including delegation
  - **ARCLIB and powerful command–line tools**
    - Including plugins for ARC Classic, glite

- **Will be added to the box**
  - Accounting service (MARS)
  - broker service supporting pull operation as well (scheduller)
  - P2P–based information system backbone (p2p ISIS)
  - Consistent storage system solution
  - *some sort of support for dynamic application management and job sandboxing using virtualization techniques*
  - Sysadmin–friendly, functionality–centric configuration
  - Technical documentation
ARC1 flagship service: A–REX

- ARC Resource–coupled EXecution service
- Job execution service implemented within ARC1 HED
- Clever Grid layer on top of a computing resource
- Based on the Grid Manager modul of Classic ARC
  - Numerous enhancement in code base
- Standard interface:
  - BES–compliant service (only the BES–Factory supported)
  - JSDL
  - NorduGrid extensions for both BES and JSDL
  - WSRF for information queries
- Demonstrated standard–based interoperability with large number of prototypes including glite, Unicore
  - Supercomputing 2007 HPCP interoperability fiesta
  - ARC/Glite/Unicore Tutorial at OGF3, Barcelona
On backward compatibility

- On the server side, almost everything will have new interface/format
  - Information model
  - Information query
  - Job management
  - Security language
  - Configuration

*OLD clients (ARC Classic) will not work with new servers (ARC1), nevertheless we’ll provide co-deployable services*

- On the client side, ARCLIB will offer a smooth transition, all the client utilities will support old servers

*OLD servers (ARC Classic) will be accessible from new clients (ARC1)*
Standards supported by ARC1 (1/2)

- "Standards document"
  - A comprehensive survey which presents a short description, relevance, support status of 66 specifications
  - www.knowarc.eu/documents/Knowarc_D3.3-1_08.pdf

- GFD.108 OGSA® Basic Execution Service Version 1.0
  - Mostly implemented

- GFD.99 OGSA® Security Profile 1.0 – Secure Channel
  - HTTP over TLS

- GFD.56 Job Submission Description Language (JSDL) Specification v1.0
  - Subset of elements is supported
  - Work is in progress for achieving compliant behavior
Standards supported by ARC1 (2/2)

- **GFD.114**  HPC Basic Profile, Version 1.0
  - Except BasicFilter extension
- **GFD.111**  JSDL HPC Profile Application Extension, Version 1.0
  - Supported elements: *Executable, Argument, Input, Output, Error*
- **OASIS Web Services Resource Framework (WSRF)**
  - WS-BaseFaults (WSRF-BF)
  - WS-ResourceProperties (WSRF-RP)
  - GetResourcePropertyDocument
  - GetResourceProperty
  - GetMultipleResourceProperties
  - QueryResourceProperties
- **W3C Web Services Addressing (WS-Addressing)**
- **OASIS Web Services Security (WSS)**
  - Username Token Profile 1.1
On interoperability

- Our main goal is to achieve interoperability via standards embraced by all the major grid middleware players
  - BES, JSDL, GLUE, SRM, GridFTP, X509, SAML
  - Agreed extensions are critical (profiles)!
- For a short term (?) transitional period gateway-like solutions are necessary
  - ARCLIB addresses the ARC -> other middleware direction
  - ng* commands will offer transparent access capability to 3rd party services
- Target platforms: Glite, UNICORE
Addressing ARC0 limitations in ARC1 (1/2)

- Custom, non-standard based service interfaces
  - ARC1 will exhibit standard-based interfaces
- NO SOA, monolithic software, nordugrid-dependent data models
  - HED and ARCLIB is very much modular
- Strong Globus dependency
  - The necessary minimal globus dependency is factored out into optional modules
- Security limitations:
  - Clumsy and poorly documented configuration
  - Security concerns due to grid mapping, root services
  - Non-perfect isolation of Grid jobs running on the same resource
  - No security layer in Information System
    - Will be part of the new information system
  - No or very limited support for grid economy
    - Partially will be covered by the MARS service
Usability concerns
- No support for bulk job management, one-job-focus
- Managing Software Environments (Applications) is a pain:
  - Some work started with dynamic RTE
- Poor support for parallel applications
  - Improvements are on the way due to glue2
- Why can not I use the grid from Windows?
  - Windows client is on the way, so is MAC–OS
- What do I gain running ngsub instead of qsub?
- Too complex for end-users: certificates, XRSL, GACL syntax, new set of commands
- No easy-to-use GUI (only a prototype)

Limited data management
- Clumsy input/output data management:
  - Some promising work with FUSE,
- No coherent and consistent data management
  - new storage system with global namespace

Scalability concerns
- Data staging process, information monitoring, application deployment
- Support for clusters with ten thousands CPU cores
- To be addressed together with ARC Classic
ARC1 overall status

- Not ready yet for the general public
  - So far only developers involved can deploy and run it
  - Requires lots of tuning
  - Documentation, integration, testing need lots of additional effort needed
  - We are looking forward the "August release"

- HED is quite robust, unfortunately not that many services are available yet
  - Service development framework is not ready yet for ordinary developers either

- Key service: A-REX is OK,
  - full job submission chain works
  - Successful DEMOs on SCxx, OGF events

- ARC1 clients will appear soon (capable talking to ARC Classic too)

- Information system, Security framework and configuration are the biggest black holes

- Detailed component status: "feature file" in the svn
  - http://svn.nordugrid.org/trac/nordugrid/browser/arc1/trunk/FEATURES
ARC1 (immediate) plans

- Deliver the August release
- Bring A-REX on the level so that it could replace the ARC Classic Computing Element
- Implement BES/JSDL/HPCP/GLUE standard conformance in A-REX according to our "standard roadmap"
- Complete the new client based on new ARCLIB
- Fill the pressing gap: information system in ARC1
- Make the security framework useable,
- Add security layer to storage system, support additional transfer methods (not just byteio)
- Perform scalability measurements, work together with ARC Classic to address the thousand-core clusters problem
- TESTING, TESTING, TESTING
- DOCUMENTATION, DOCUMENTATION, DOCUMENTATION
Further information

- Wealth of information on www.nordugrid.org and www.knowarc.eu
- ARC Classic "white paper":
- ARC1 design:
- Code:
  - svn.nordugrid.org -> arc0 and arc1 directories of the nordugrid repository
  - download.nordugrid.org -> official source and binary packages (mostly for arc classic), external software
- Join the community!
  - Check out, sing up for the nordugrid-discuss mailing list
  - Attend some of the Technical Meetings or Conferences
Conclusion: the big questions

Core facts:
- ARC Classic is there and doing a very good job
  - The scalability issues are relevant for ARC1 as well
- ARC1 is pursuing an intensive and very promising development program

Questions:
- Can ARC1 replace ARC Classic?
  - Delivers at least the same level of reliability, scalability
  - Comes with an extended feature set
- Will ARC1 ready by time?
  - When will that time arrive?