# ARC for Admins, Introduction

*Marko Niinimaki, HEPIA, Geneva*

*GridKA school 5$^{th}$ - 9$^{th}$ Sept 2011*

*Karlsruhe, Germany*

# Outline

ARC (Advanced Resource Connector)

- Introduction

- How does it work?

- How do you install/manage it?

*These slides are based on Balazs Konya's presentation ARC's view on the European*
*(Grid) Middleware Initiative*

8/23/2011

www.nordugrid.org

# ARC Introduction: general

- Well supported, lightweight open source middleware
- Best suited for high-throughput distributed computing
- Independent, portable code base
  - GSI-based

- Clear separation of cluster and grid layer
  - No grid layer on the nodes (unless required by users)
  - Input/output grid data handling by the front-end
    - Increases CPU utilization
    - Automatically allows for data caching
  - ARC frontend: all grid related operations
  - ARC infosys: based on LDAP
- Resource discovery and brokering encapsulated in the client
  - No single point of failure, ARC clients act as "agents"
  - Based on a client API, ARCLIB

# ARC Introduction: features

Best suited for **shared** **community** resources:

- **Portable** (Linux/Unix, OSX, MSWin..)
- Has interfaces to **most major batch systems**
  - ○ Any new batch system can easily be plugged in
- **Minimal intrusiveness**, minimal footprint
- **Quite simple** installation and maintenance (compared to other solutions)
  - ○ Suits a 1-CPU "site" and a 3000+ cores cluster
- Versatile, **portable CLI**: ~14 MB in size, needs no root privileges
  - ○ Does "everything" from SRM storage listing to brokering

# ARC introduction: history

Y2K: Grid Hype, European Data Grid (EDG), re-discovery of Globus Toolkit (version 1.1.4)

2001 HEP Institutes from Scandinavia wanted to share their computing resources and jointly contribute to CERN/LHC computing -> "NorduGrid", a research project of the NORDUNet2 program aimed at enabling Grid in the Nordic countries

2002 February: decision to develop an alternative middleware by making use of Globus libraries. 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 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 (-2009) project started.

2007 May: After a long hardening process ARC version 0.6, the second stable release of the middleware was released.

2009 Oct: ARC 0.8 released: classic ARC that already includes features of the "next gen" ARC.

2011 May: ARC 11.05, a first release of the "next gen" web service oriented ARC.

2011 May: First release by the EMI project that contains "a complete and consolidated set of middleware components from ARC, dCache, gLite and UNICORE."

8/23/2011
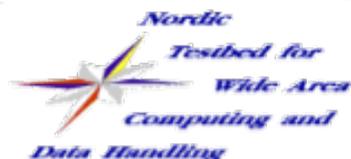
# Introduction: Applications

*Disclaimer: information shown here is incomplete and was collected by Oxana Smirnova in half an hour by asking people around and googling*

- Biophysics
- Biochemistry
- Computational chemistry
- Quantum chemistry
  - GAMESS
- Molecular dynamics
  - GAUSSIAN, DALTON, MOLDEN
- Bioinformatics
  - Taverna
  - BLAST, HMMER
  - eQTL
- Solid state physics
- Computational physics

- Mathematical crystallography
- Informatics, mathematical logic clause solving
- Automatic malware comparison
- Medical imaging
- Simulation of avalanche dynamics
- HEP
  - ATLAS, IceCube, CMS, ALICE, LHCb tested
- Other materials sciences

8/23/2011

# How does it work?

- ARC 1 is a modular software whose core is a WS service container called HED (hosting env).

- Other components: Information system, file transfer (storage).

- HED provides standards-conforming services, including the job execution service, A-REX

- HED provides bindings/interfaces for programmers to

STORAGE ETC

HOSTING ENV

A-REX
AUTHORIZATION
INFOSYS

INF. INDEX
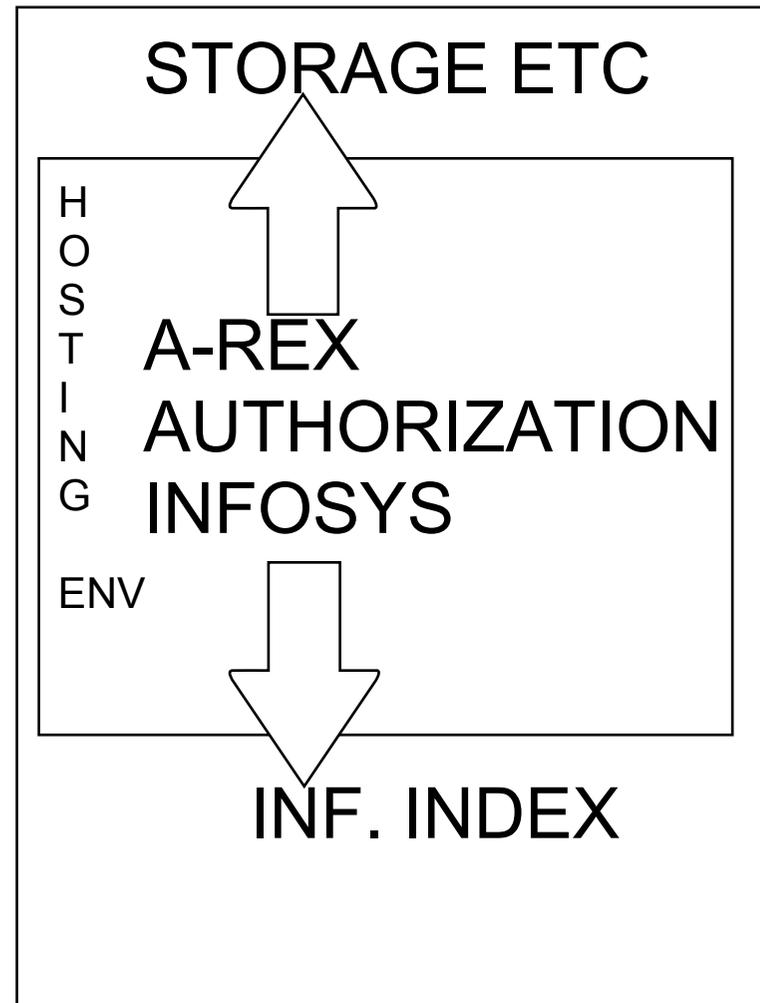
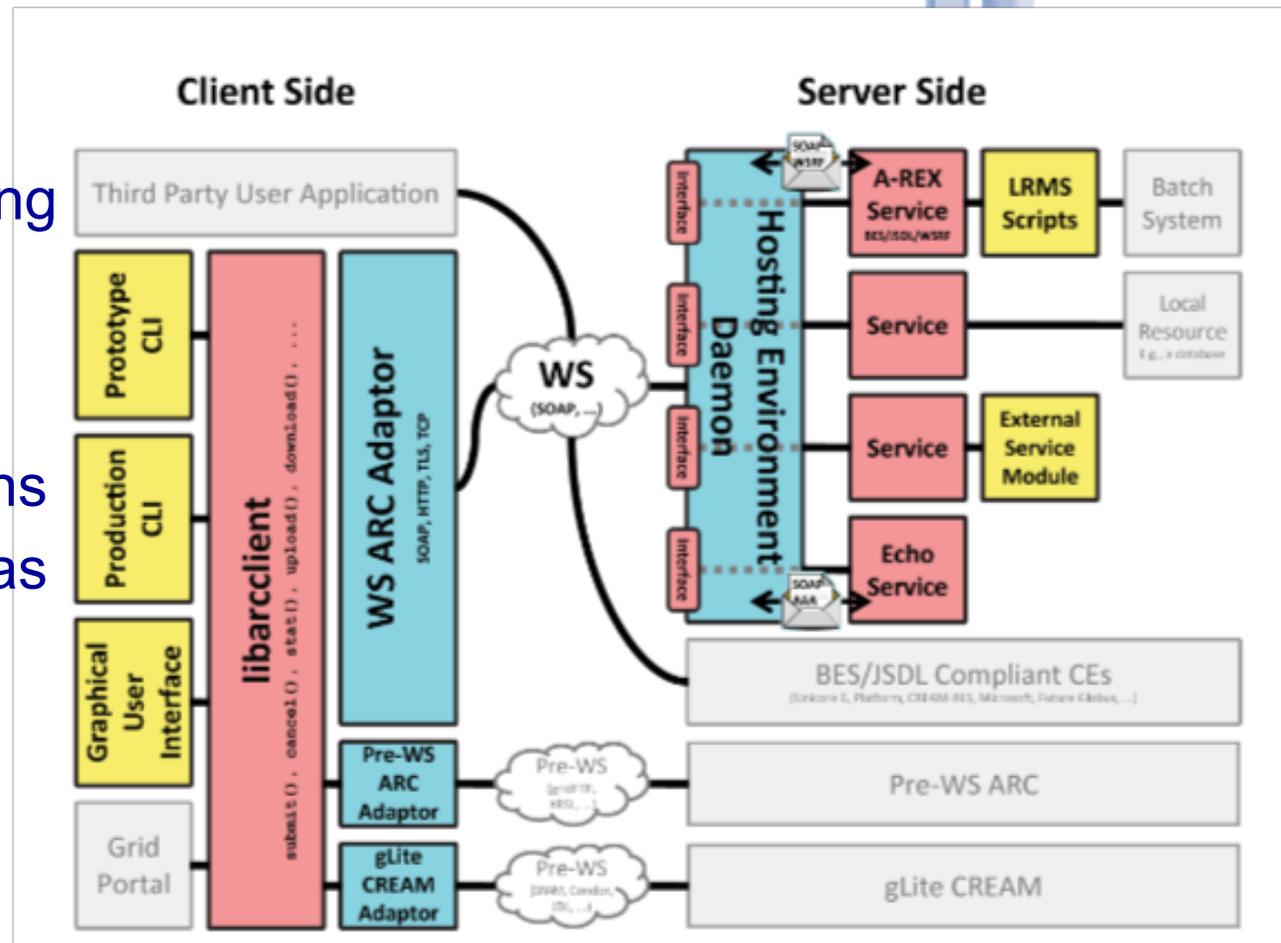# *How does it work? .. in detail ..*
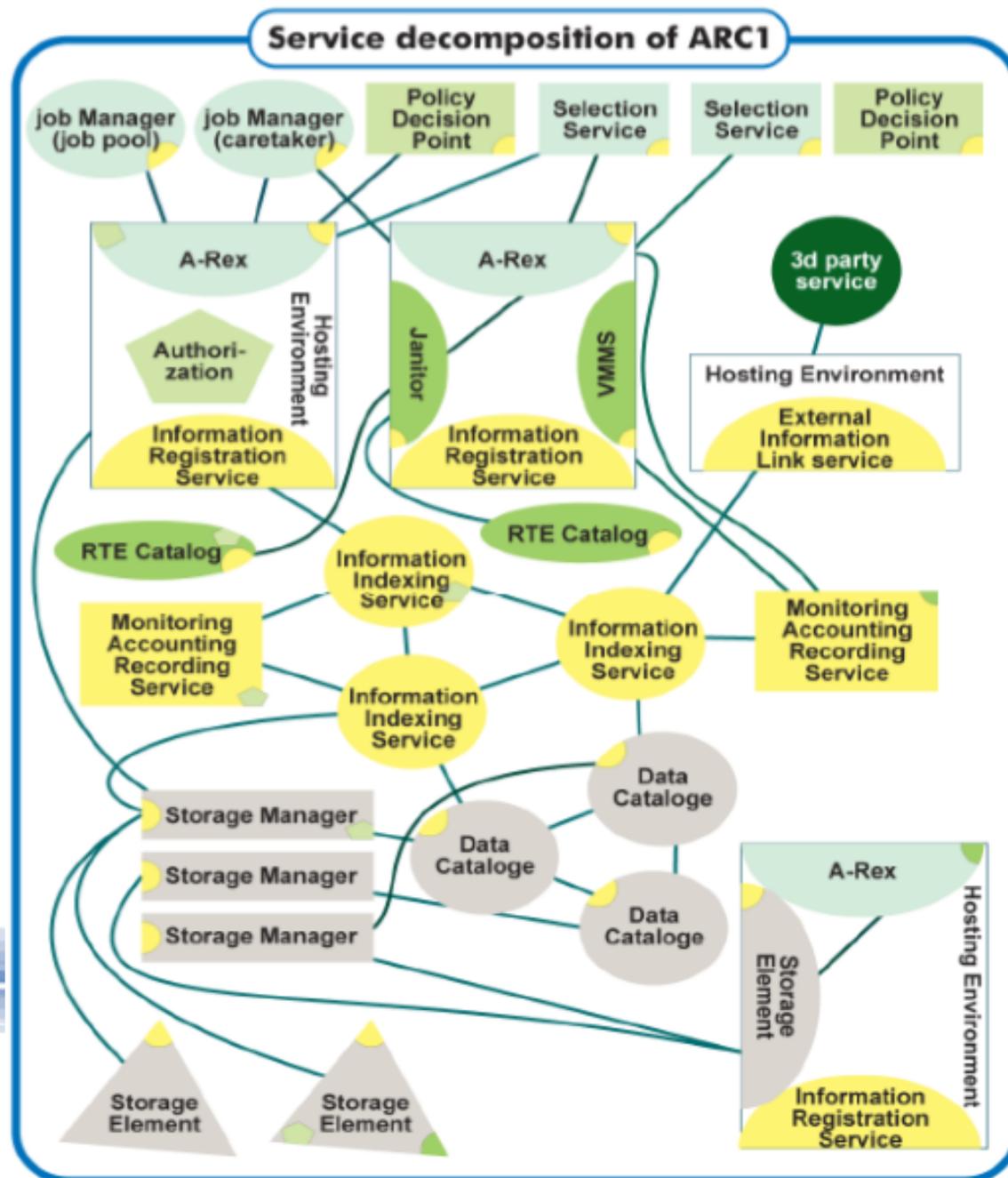


Service decomposition of ARC1

# How does it work: grid jobs

● **A-REX**

○ The main HED service implementing a Computing Element (CE)

○ JSDL/BES/GLUE2 with ARC extensions

○ Became available as part of the 0.8 production ARC release

○ Based on the Grid-Manager of classic ARC.



8/23/2011

# How does it work: running the server

/etc/init.d/gridftpd start
/etc/init.d/grid-infosys start
/etc/init.d/a-rex start

| | | | | | | |
|---|---|---|---|---|---|---|
| ◈ gridftpd | Sleeping | 0 | 0 | 6862 | 968.0 KiB | inet_csk_wait_for_connect |
| ⊟ ◈ grid-info-soft-register | Sleeping | 0 | 0 | 30318 | 108.0 KiB | do_wait |
| ◈ ldapadd | Sleeping | 0 | 0 | 1527 | 348.0 KiB | inet_wait_for_connect |
| ⊟ ◈ grid-info-soft-register | Sleeping | 0 | 0 | 30311 | 100.0 KiB | do_wait |
| ⊟ ◈ arched | Sleeping | 0 | 0 | 1062 | 500.0 KiB | hrtimer_nanosleep |
| ⊟ ◈ scan-fork-job | Sleeping | 0 | 0 | 1837 | 80.0 KiB | do_wait |
| ⊘ sleep | Sleeping | 0 | 0 | 1845 | 64.0 KiB | hrtimer_nanosleep |

8/23/2011

# How does it work: client

**Linux:** redhat fedora ubuntu debian

Platform selection:

| Vendor | Version | Architecture |
|--------|---------|--------------|
| debian | 8.10 | i386 |
| fedora | 9.04 | amd64 |
| redhat | 9.10 | |
| ubuntu | 10.04 | |
| | 10.10 | |
| | 11.04 | |

Download

http://download.nordugrid.org/base.html

tar xvzf nordugrid-arc-standalone..
cd nordugrid-arc-standalone
source setup.sh

+ Quick edit of 1 file (.arc/client.conf) +
certificates in place = a functional grid.

marko@man-inspire:~$ arc/bin/arcproxy
Your identity: /DC=ch/DC=switch/DC=slcs/O=Haute Ecole Specialisee de Suisse occidentale
(HES-SO)/CN=Markopekka Juhani Niinimaeki E0ED1216
Enter pass phrase for /home/marko/.globus/userkey.pem:
.....++++++
......++++++
Proxy generation succeeded
Your proxy is valid until: 2011-08-09 22:45:58
marko@man-inspire:~$ arc/bin/arctest -J 1
Test submitted with jobid: gsiftp://arctest.unige.ch:2811/jobs/105511312879872177620756

# How to install/manage ARC

Installation is relatively easy thanks to software repositories, see http://www.nordugrid.org/documents/arc-server-install.html

Example:

Edit software sources

apt-get install nordugrid-arc-compute-element

Put host certificates in place.

Create and edit /etc/arc.conf

Enable and start daemons gridftpd, grid-infosys and a-rex.

8/23/2011

# Further information on ARC

- Wealth of information on
  - www.nordugrid.org and www.knowarc.eu
- The original ARC "white paper":
  - *"Advanced Resource Connector middleware for lightweight computational Grids"*. M.Ellert et al., Future Generation Computer Systems 23 (2007) 219-240.
- An update containing information about new components:
  - *"Recent ARC developments: through modularity tointeroperability"* , O. Smirnova et al., J. Phys. 219 (2010).
- Code:
  - svn.nordugrid.org
  - download.nordugrid.org -> official source and binary packages, external software
- The community:
  - Check out, sign up for the **nordugrid-discuss** mailing list
  - Technical Meetings or conferences typically 2-3 times a year

# Notes for hands-on at GridKA

Purpose: guided installation and configuration of an ARC server.

https://ocikbapps.uzh.ch/gc3wiki/sgs2011.html

8/23/2011

# Backslides [do not use]

Purpose: guided installation and configuration of ARC server.

How: use VirtualBox + a VM image available at ..

Requirements: a 64bit computer +OS, 2GB mem, VirtualBox

Installing and starting the image:
Start VirtualBox, in Preferences enable a "host-only" network connector
Import GridSchoolAllInOne.ova into VirtualBox and start it.
Log in as root, password gridtempl.
Optional: run ifconfig in the GridSchoolAllInOne window, check the IP
address, login using ssh from your computer.

8/23/2011

Next steps (details given during the hands-on):
Certificates
Testing torque
Installing the ARC server
Configuring ARC (typically edit just 1 file, /etc/arc.conf)
Test job submissions

yum install nordugrid-arc-gridftpd nordugrid-arc-arex nordugrid-arc-client

create a host and user certificate using https://arc-emi.grid.upjs.sk/instantCA

get a minimal fork-based arc.conf here: http://www.cern.ch/man/arc.conf-min

mv arc.conf-min /etc/arc.conf

start your grid

/etc/init.d/gridftpd start

/etc/init.d/grid-infosys start

/etc/init.d/a-rex start

add your cert subject in /etc/grid-security/grid-mapfile

Checking: files in /var/log/arc

ldapsearch -h lscf.nbi.dk -p 2135 -x -b "mds-vo-name=local,o=grid"

8/23/2011

Log in as user theuser, create dir .globus
Use the same CA file that you got from instant CA : install the
user files (cert, key) in .globus
arcproxy
arcinfo -I  aio.grid.zoo
arctest -J 2 -c aio.grid.zoo
-> job submitted

Thoubleshooting:
chmod 400 /etc/grid-security/hostkey.pem