Using the EMI testbed
ARC middleware

Marek Kočan

University of P. J. Šafárik, Košice

Outline

- ARC services available
- ARC Information system (how to query/register)
- ARC infrastructure usage

ARC services available

- Instant CA – demo user/host certificates generation (https://arc-emi.grid.upjs.sk/instantCA/)

- Classic ARC and WS-ARC Grid-Monitors (http://arc-emi.grid.upjs.sk/gridmonitor.html)

- Classic ARC GIIS and WS-ARC ISIS Cloud (http://arc-emi.grid.upjs.sk/howtoconnect.html)

- Classic ARC CE

- WS-ARC CE

- Unixacl Storage Element

- ARC echo service

ARC Information system

- ARC GIIS
  - Scalable, production quality dynamic distributed information system allowing registration of Classic ARC Services
  - OpenLDAP-based system which is derived from the Globus Monitoring and Discovery Services framework
  - It uses BDII together with a updated MDS LDAP Schema
  - More info at http://www.nordugrid.org/documents/arc_infosys.pdf
- *WS-ARC* Information System Indexing Services (ISIS) Cloud
  - Forms a set of information containers
  - Every generic ARC Service pushes information about itself into a nearby ISIS container
  - The information stored in every ISIS is propagated among all ISIS services.
  - More info at http://www.nordugrid.org/documents/infosys_technical.pdf

ARC Information system (cont.)

- ## Querying information from Classic ARC GIIS
  - ldapsearch -h arc-emi.grid.upjs.sk -p 2135 -x –b 'mds-vo-name=ARC-EMI,o=grid' -s base giisregistrationstatus

- ## Querying information from ISIS Cloud

  - isistest -i http://lem.grid.niif.hu:60002/isis -m Query "/RegEntry"

ARC Information system (cont.)

- Registering resource to Classic ARC GIIS
  - needs to add following block to *arc.conf*

    #cluster registration block

    [infosys/cluster/registration/clustertoARC-EMI]

    targethostname="arc-emi.grid.upjs.sk"

    targetport="2135"

    targetsuffix="mds-vo-name=ARC-EMI,o=grid"

    regperiod="120"

ARC Information system (cont.)

- Registering WS-ARC service to WS-ARC ISIS Cloud

```
<infosys:InfoRegister>
  <infosys:Period>PT75S/infosys:Period>
  <infosys:Endpoint>"your service endpoint"</infosys:Endpoint>
  <infosys:Expiration>PT1M/infosys:Expiration>
  <infosys:Registrar>
   <URL>http://arc-emi.grid.upjs.sk:60000/isis/URL>
   <Retry>4</Retry>
  </infosys:Registrar>
</infosys:InfoRegister>
```

or

```
[register]
index_server1=http://arc-emi.grid.upjs.sk:60000/isis
["service type"]
endpoint="your service endpoint"
```

ARC infrastructure usage

- Generating proxy credentials

  arcproxy -S testers.eu-emi.eu

  man arcproxy for more info

ARC infrastructure usage

- *arcsub*
  - Submits jobs to grid enabled computing resources
  - Supports submission to Classic ARC CE, WS-ARC CE, gLite Cream, Unicore (in experimental version currently)
  - Submits jobs described in XRSL, JSDL and JDL formats
  - *man arcsub* for more info

arcsub jsdl_hostname.xml
  - Submits job to randomly chosen cluster (from set of registered clusters)

arcsub -c ARC0:ldap://rc0.grid.upjs.sk:2135/nordugrid-cluster-name=rc0.grid.upjs.sk,Mds-Vo-name=local,o=grid jsdl_hostname.xml
  - Submits job to concrete Classic ARC cluster

arcsub -c ARC1:https://lem.grid.niif.hu:60000/arex jdl_hostname.jdl
  - Submits job to concrete WS-ARC cluster

ARC infrastructure usage (cont.)

- *arcstat*
  - obtains the status of jobs that have been submitted to grid enabled resources
  - *man arcstat* for more info

arcstat <jobID>
  - Retrieves status of concrete job

arcstat -a
  - Retrieves status of all jobs submitted (and not already cleaned)

ARC infrastructure usage (cont.)

- *arccat*
  - Can be used to view the stdout or stderr of a running job(s)
  - *man arccat* for more info

arccat <jobID>
  - Prints stdout of concrete running job

arccat -a
  - Prints stdout of all running jobs

ARC infrastructure usage (cont.)

- *arcget*
  - Retrieves the results from a job
  - *man arcget* for more info

arcget –k <jobID>
  - Retrieves the results of finished job and keeps files on cluster (-k option)

arcget –a
  - Retrieves results of all finished jobs and cleans all files on cluster

ARC infrastructure usage (cont.)

- *arckill*
  - Kills running job(s)
  - *man arckill* for more info

arckill <jobID>
  - Kills concrete running job

arckill –a
  - Kills all running jobs

ARC infrastructure usage (cont.)

- *arcclean*
  - Removes a job(s) from the remote cluster
  - *man arcclean* for more info

arcclean <jobID>
  - Removes job from the remote cluster

arcclean –a
  - Removes all jobs from the remote cluster

ARC infrastructure usage (cont.)

- arcls
  - Lists files and directories in grid storage elements and file index catalogues
  - Following protocols may be used
    file (file:// prefix may be omitted), http, https, httpg, ftp, gsiftp, lfc, rls, srm
  - *man arcls* for more info

arcls –l <jobID>

  - Lists files/directories generated by running job (-l for long format)

arcls gsiftp://rc0.grid.upjs.sk:2811/unixacl/

  - Lists content of Unixacl SE

ARC infrastructure usage (cont.)

- arccp
  - copies files to, from and between grid storage elements
  - Following protocols may be used
    file (file:// prefix may be omitted), http, https, httpg, ftp, gsiftp, lfc, rls, srm
  - *man arccp* for more info

arccp fileToGSI gsiftp://rc0.grid.upjs.sk:2811/unixacl/uploadedFile
  - Uploads file to an Unixacl SE

arccp gsiftp://rc0.grid.upjs.sk:2811/unixacl/10mb-file downloadedFile
  - Downloads file from Unixacl SE

ARC infrastructure usage (cont.)

- *arcrm*
  - deletes files on grid storage elements
  - Following protocols may be used
    file (file:// prefix may be omitted), http, https, httpg, ftp, gsiftp, lfc, rls, srm

  - *man arrm* for more info

arcrm gsiftp://rc0.grid.upjs.sk:2811/unixacl/uploadedFile
  - removes file from Unixacl SE

ARC infrastructure usage (cont.)

- Additional commands
  - *arcinfo* - obtains status and information of clusters on the grid
  - *arcecho* - client for the ARC echo service
  - *arcmigrate* - migrates queued jobs from one cluster to another cluster. Currently migration only works between clusters running A-REX (WS-ARC CE)
  - *arcresume* – resumes a job that has been submitted to grid enabled resources.
  - *arcsync* - synchronizes local jobs list with the information at a given cluster or index server
  - *arcresub* - resubmits jobs to grid enabled computing resources

Thank you!