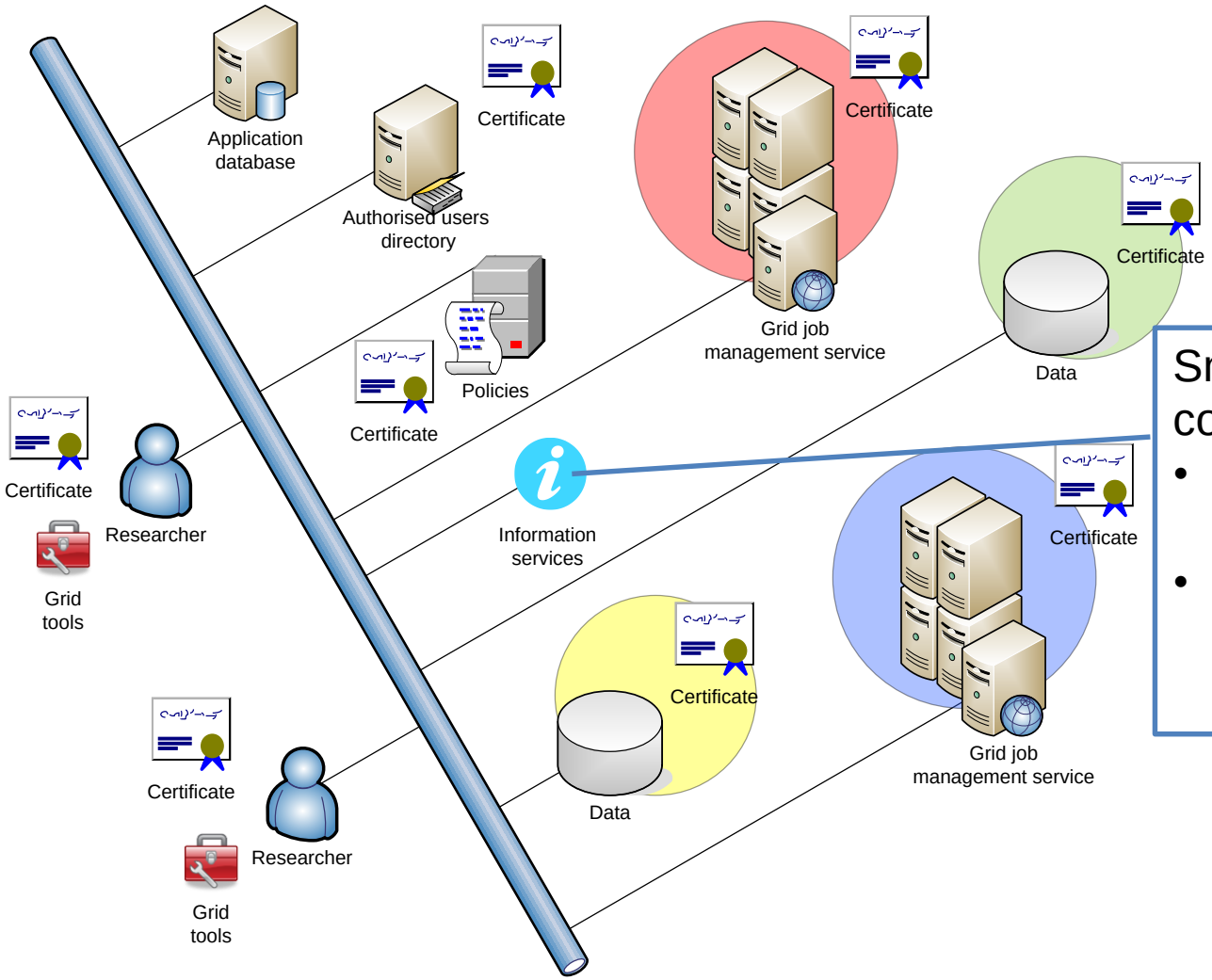
Information and monitoring

Information is essential



Small but most important component:

- List of Grid services and their properties
- Necessary for service discovery and status monitoring

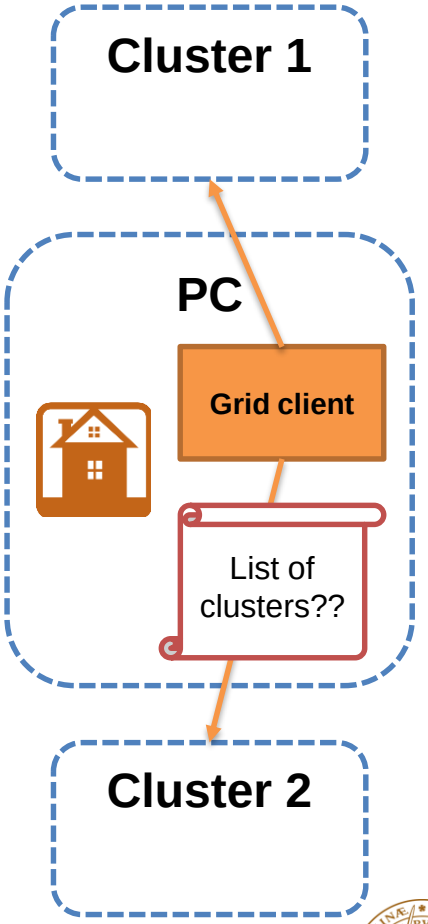
LUND UNIVERSITY

Grid bookkeeping

- We now can submit jobs to different clusters, e.g.:

```
> arcsub -c cluster1.lab.org myjob.xrsl
> arcsub -c bigcluster.inst.org myjob.xrsl
```

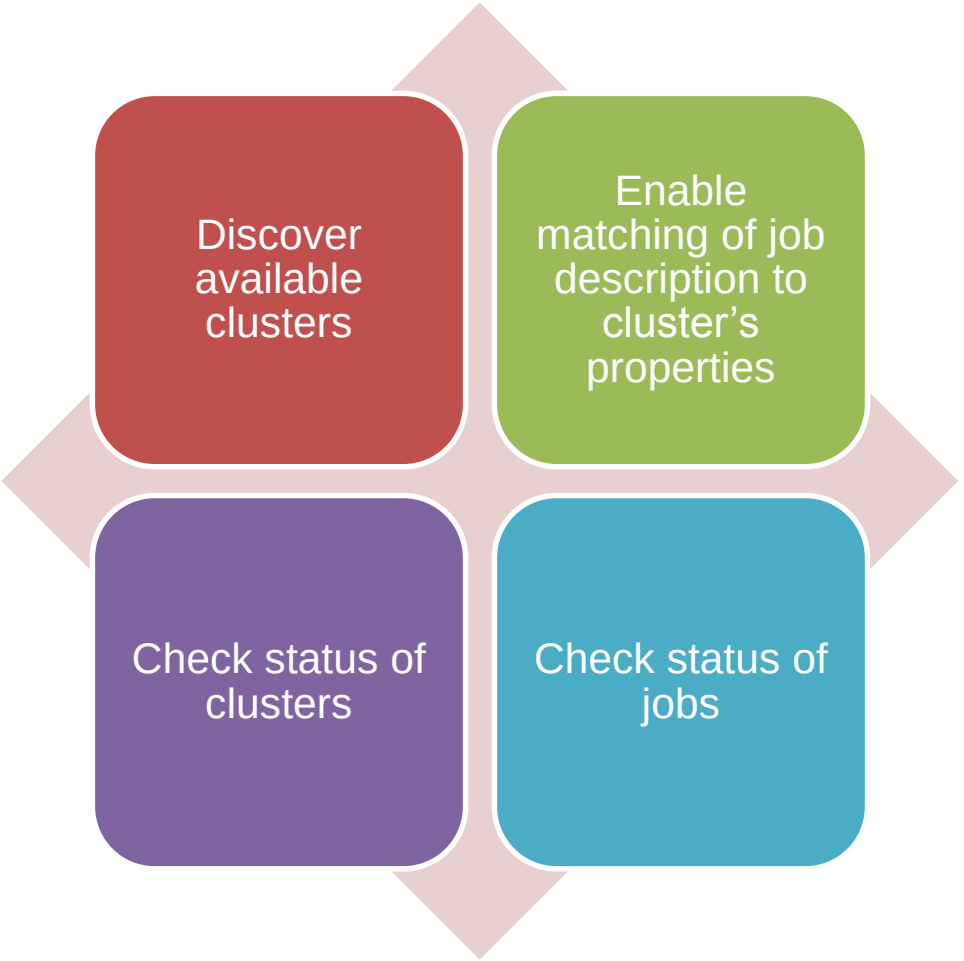
- Doing this, we benefit from 2 Grid advantages:
  - Common login using proxy credentials
  - Common job description
- But how do we know that these clusters exist?
  - Keeping a list on a piece of paper sounds silly
  - Should Grid client keep such a list?
  - Who should keep an eye on changes?
- What if these clusters do not fit my job description?



Information system concept

- Each Grid service must provide information about itself on request
  - True for all services, but currently implemented only for computing services, and partially – for storage
- A request can be anonymous or authorised
  - Anonymous requests are useful for overall monitoring
  - Authorised requests allow to provide user-specific information
    - » Details of user's or VO's jobs
  - In practice, only anonymous requests are fully implemented
    - » You can check status of any service or anybody's job on the Grid
    - » You don't even need a proxy to do this
- In what follows, only clusters' information is considered

Purposes of information system



LUND UNIVERSITY

Types of information published by clusters

- Information published by clusters can be used for matchmaking and monitoring
- Each of these can be either semi-static or dynamic, for example:

	Matchmaking	Monitoring
Semi-static	• Type of interface • Maximum memory • Performance benchmark	• Cluster name • Cluster location • Total CPU count
Dynamic	• Free CPU count • Free storage • Queue length	• Occupied CPU count • Number of jobs • Queue status (up/down)

Many dilemmas

Minimum vs maximum

- Job descriptions state minimal requirements
- Clusters advertise maximal capacity

Centralised vs distributed

- Distributed dynamic information is wrong for everybody (users are unaware of each other)
- Centralised databases have synchronisation and scalability issues

Generic vs specific

- Generic information is not enough for specific needs
- Specific information changes from user to user and from cluster to cluster

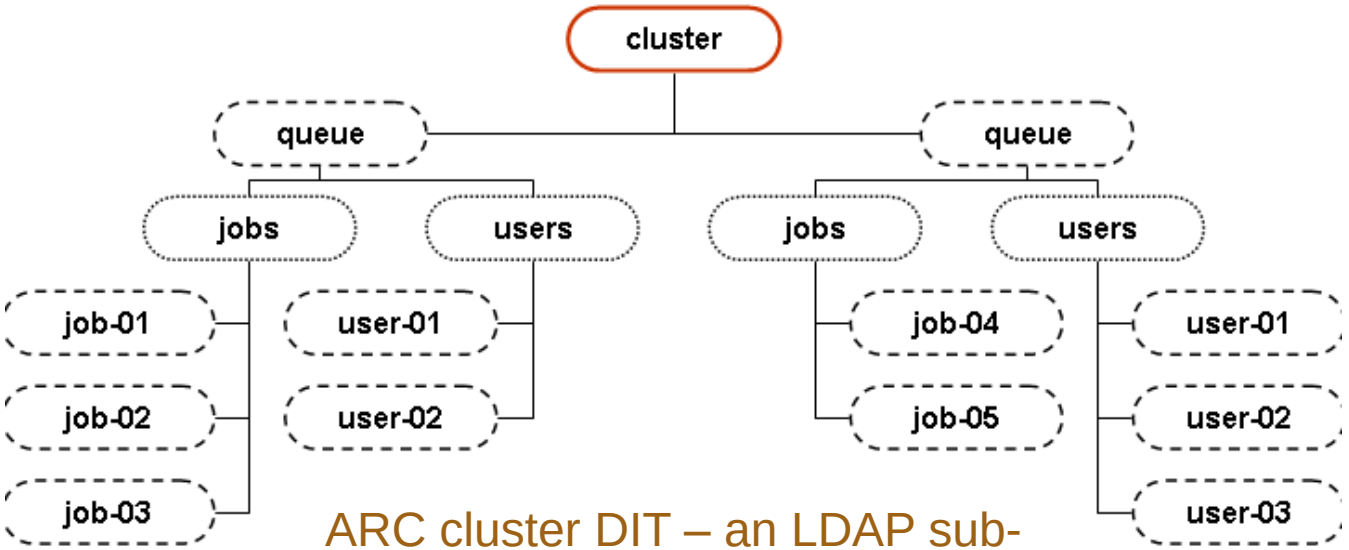
Common approach: LDAP

- LDAP: Lightweight Directory Access Protocol
  - A hierarchical database (directory) and a protocol to interact with it (`ldap://`…)
  - Perfect for storing information like phone directories
    - » Widely used to store personal information, even in LU
  - Clusters have hierarchical structure, hence LDAP is a perfect fit
    - » Each cluster has queues, each queue has jobs and users, etc.
- LDAP is a subset of the X.500 standard
  - The directory consists of **objects** (entries), each of which has a number of **attributes**
    - » Attributes are defined in a **schema**
    - » Several schemas exist on the Grid
  - Objects are organised in **trees and sub-trees**
  - Each object has a **Distinguished Name** (DN)
    - » LDAP DNs have same structure as X.509 certificate DNs, which is no coincidence
    - » DN is not an attribute – it is just a name

LUND UNIVERSITY

ARC schema

- Schema defines a set of rules defining the structure of a Directory Information Tree (DIT)
- ARC comes with an own schema and DIT
  - Not a standard, but quite logical



ARC cluster DIT – an LDAP sub-tree representing a computing resource

LUND UNIVERSITY

Some ARC cluster attributes

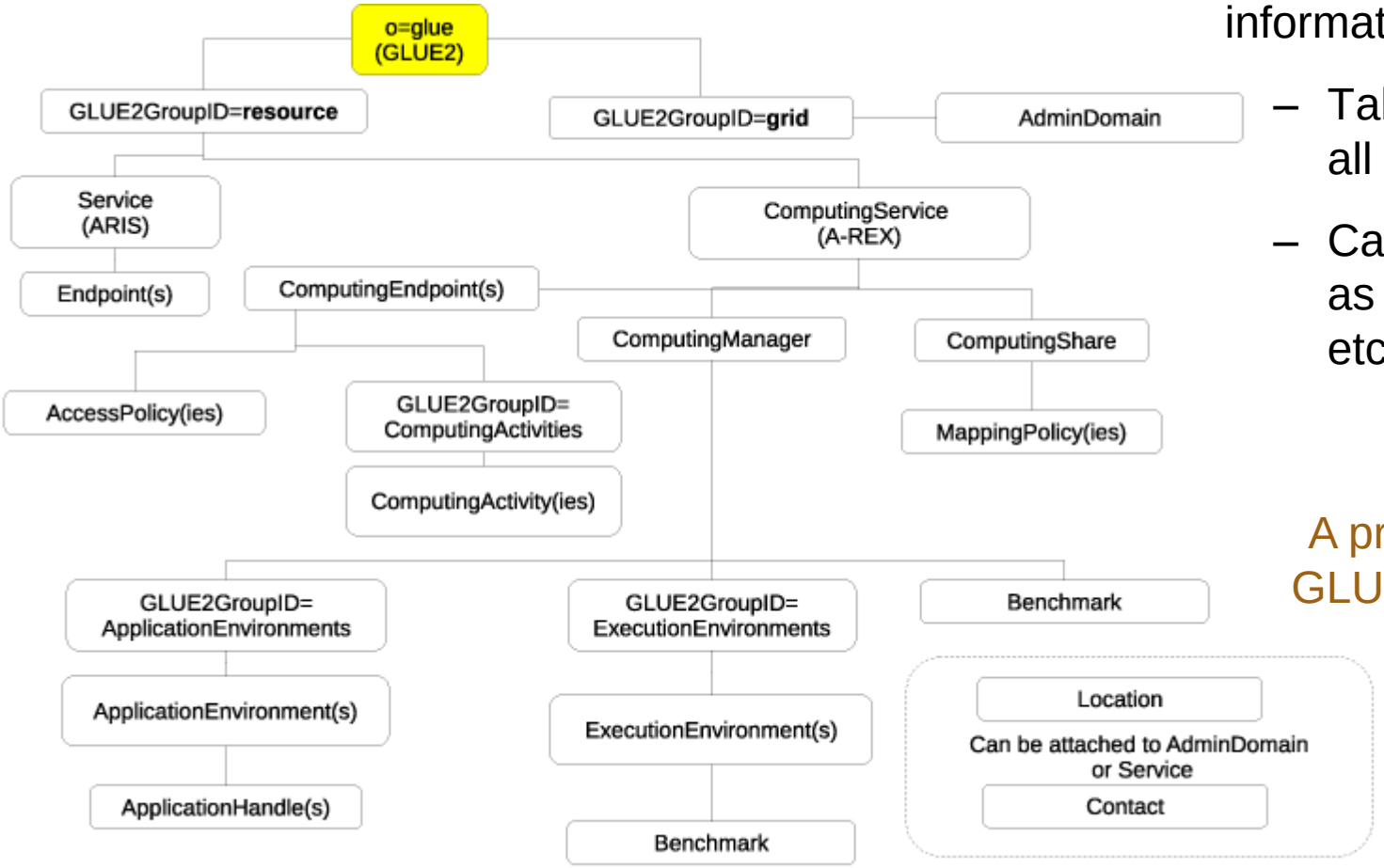
Distinguished Name

Certificate Authority

```
dn: nordugrid-cluster-name=arc-iridium.lunarc.lu.se,Mds-Vo-name=local,o=grid
nordugrid-cluster-totalcpus: 192
nordugrid-cluster-name: arc-iridium.lunarc.lu.se
nordugrid-cluster-aliasname: Iridum Cluster ARC
nordugrid-cluster-lrms-type: SLURM
nordugrid-cluster-lrms-version: 2.6.6-2
nordugrid-cluster-middleware: nordugrid-arc-4.0.0
nordugrid-cluster-cpudistribution: 16cpu:12
nordugrid-cluster-issuerca: /C=NL/O=TERENA/CN=TERENA eScience SSL CA
nordugrid-cluster-sessiondir-total: 1330688
nordugrid-cluster-architecture: x86_64
nordugrid-cluster-prelrmsqueued: 0
nordugrid-cluster-cache-free: 4194167
nordugrid-cluster-contactstring: gsiftp://arc-iridium.lunarc.lu.se:2811/jobs
nordugrid-cluster-totaljobs: 0
nordugrid-cluster-usedcpus: 0
nordugrid-cluster-sessiondir-free: 1303773
objectClass: Mds
objectClass: nordugrid-cluster
Mds-validfrom: 20140407234414Z
Mds-validto: 20140407235414Z
```

Ordered by hand: LDAP doesn't care much about ordering

Standard in works: GLUE2

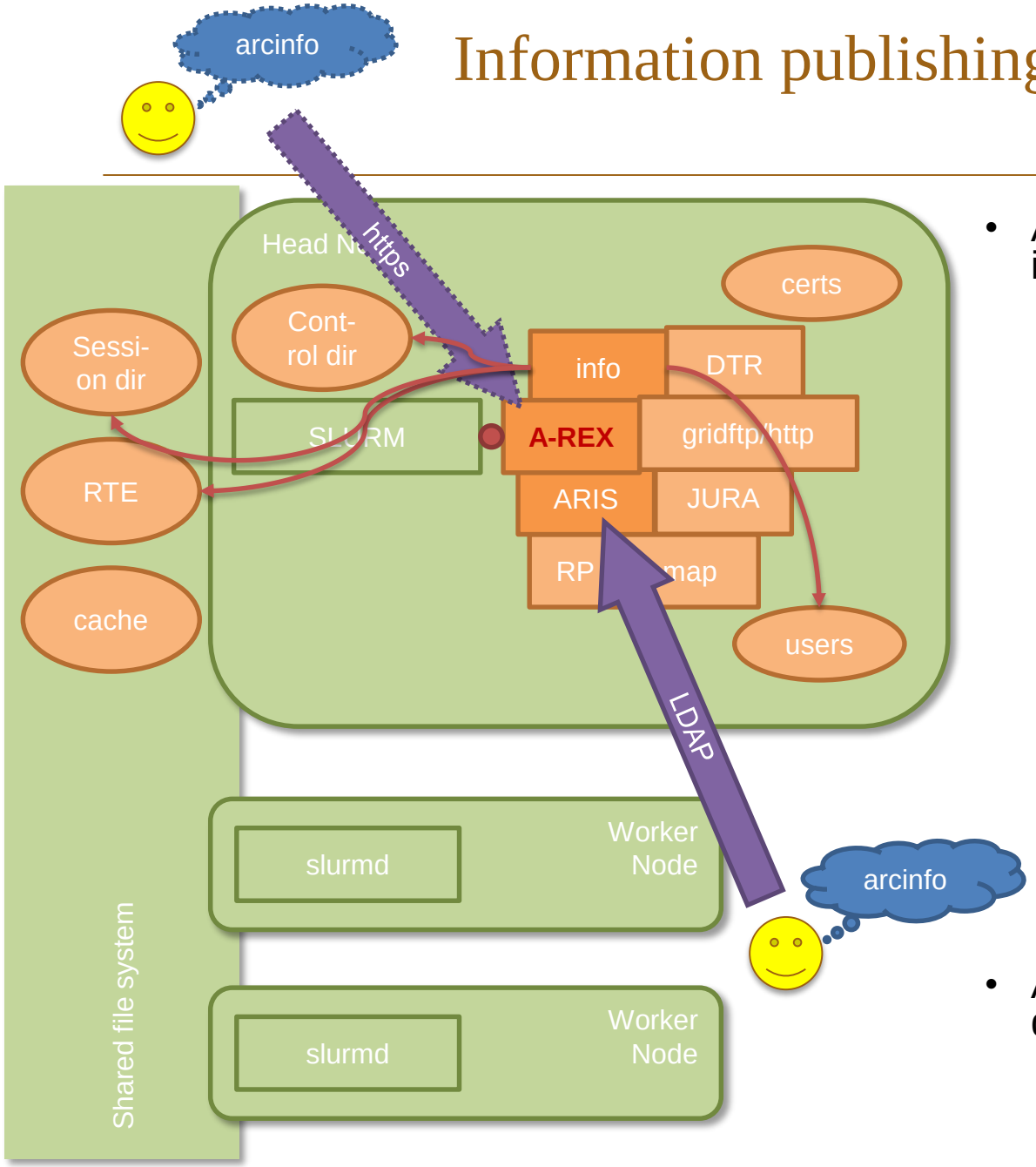


- GLUE2 is a common Grid information standard
  - Takes *__long__* time for all the players to agree
  - Can be implemented as LDAP, or as XML etc

A proposed ARC GLUE2 LDAP tree

LUND UNIVERSITY

Information publishing for an ARC cluster

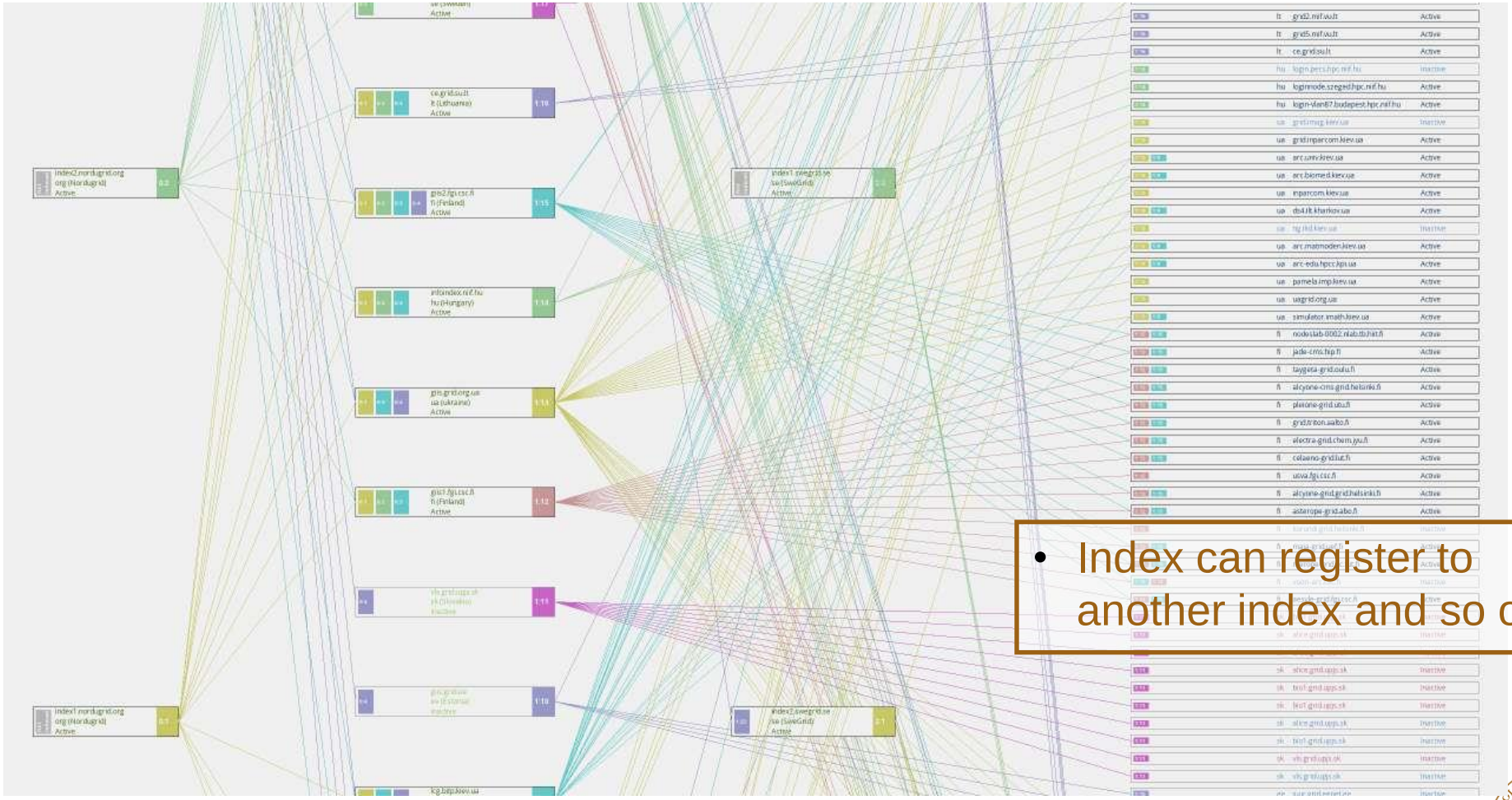


- A-REX periodically launches information providers which:
  - Collect all details defined by relevant information schemas, such as
    - » Hardware details
    - » LRMS details
    - » Available application software (RTE)
    - » Authorised users (DNs)
    - » etc
  - Create formatted output ready to be served on request
  - Populate LDAP databases of ARIS
- ARIS serves information when queried via LDAP

Information indexing

- So far, so good: each cluster provides a list of its properties
- One cluster is not really a Grid, even two clusters is not too impressive
- What provides the list of clusters when there are hundreds of them?
- Answer: each cluster regularly announces itself to information indexing servers
  - Certain (minimal) amount of information is being sent
    - » Most importantly, cluster's contact details
- Information indexing servers usually make use of LDAP, too
  - Indexing servers can even store entire set of information from each cluster, but it does not scale well

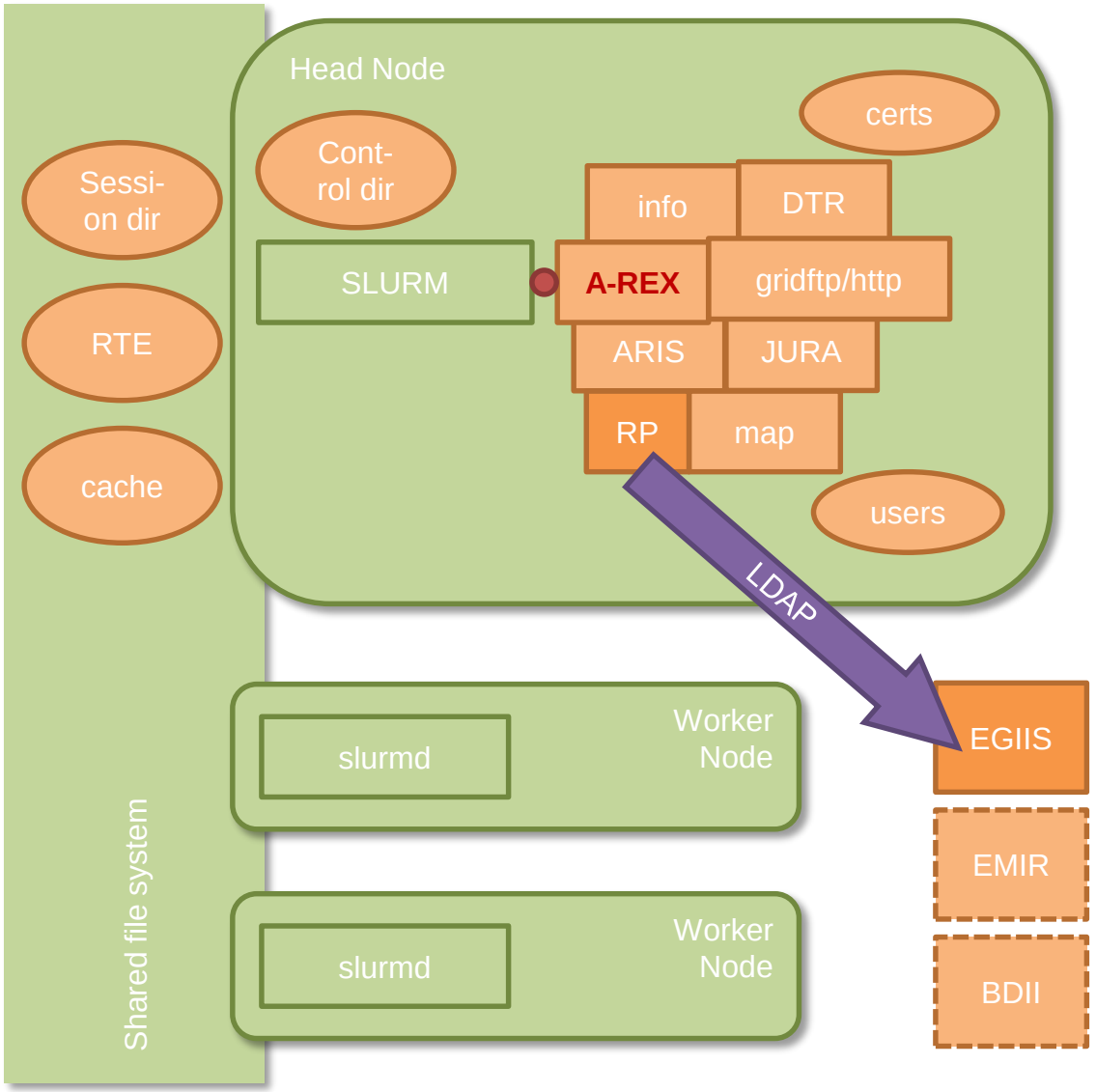
Information indexing hierarchy



- Index can register to another index and so on

LUND UNIVERSITY

Registration Process on an ARC cluster



- Information Registration Process periodically sends pre-configured information to one or more pre-configured information registries
  - Service type (cluster in this case)
  - Service contact details (contact string, port etc)
- Currently all such data are communicated via LDAP
  - Other technologies are possible, as primary data are stored internally as regular files

Extended information

- Many users and VOs want to know additional information, e.g.
  - E-mail of a security officer
  - Planned downtimes
  - Performance history for specific kind of tasks
  - Details of particular software
- Such information is often user-specific, and often can not be automatically generated
- Several additional approaches are used:
  - VOs create their own information systems
    - » Harvest information from standard Grid channels
    - » Add own details
  - Jobs are instructed to collect information themselves, and change workflow depending on the result
    - » Pilot/agent jobs: “empty” jobs that collect information and pull in matching tasks from VO databases

Monitoring

- Current status of the Grid can be inspected via the native information system
- Historical details need an additional database
  - Various VO-specific “dashboards” exist to monitor changes of status
- Some information can only be collected by special monitoring jobs – probes
  - A probe can test whether a specific environment is set up in a way that suits a VO
  - Results of such jobs can be used by standard monitoring systems, e.g. Nagios

Exercises

- Discover **the local information tree** published by the LDAP server running on the CE using the **`ldapsearch` CLI**
  - Read the **`ldapsearch`** man pages, then connect to the root entries of the LDAP trees on the Iridium computing service:

```
ldapsearch -h arc-iridium.lunarc.lu.se -p 2135 -x -b "mds-vo-name=local,o=grid" -s base
```

  Or try the GLUE2 tree:

```
ldapsearch -h arc-iridium.lunarc.lu.se -p 2135 -x -b "o=glue" -s base
```

  - Get the child elements containing all the attributes

```
ldapsearch -h arc-iridium.lunarc.lu.se -p 2135 -x -b "o=glue" -s one
```

  - Or the "full tree skeleton" only with the DNs:

```
ldapsearch -h arc-iridium.lunarc.lu.se -p 2135 -x -b "mds-vo-name=local,o=grid" dn
```

  - Check the dynamic behaviour of the information system. Run the following **`ldapsearch`** for fetching information about Grid jobs BEFORE and AFTER submitting a couple of "hello grid" jobs:

```
ldapsearch -h arc-iridium.lunarc.lu.se -p 2135 -x -b "mds-vo-name=local,o=grid"
objectclass=nordugrid-job
```

  Can you find out e.g. the submission time of certain job? (hint: check the **`nordugrid-job-submissiontime`** attribute value)

Exercises

- Browse the **local information LDAP Trees** with a **Graphical LDAP browser**
  - On Windows, install either **`ldapadmin`** (preferred) or **`jxplorer`** (requires Java):

    http://www.ldapadmin.org/download/ldapadmin.html

    http://jxplorer.org/downloads/users.html
  - On Linux, install the **`jxplorer`** (can be installed without root privileges)

    http://sourceforge.net/projects/jxplorer/files/jxplorer/version%203.3.1/jxplorer-3.3.1-linux-installer.run/download
  - Define a new connection with following settings:

    `server: arc-iridium.lunarc.lu.se port: 2135, base: mds-vo-name=local, o=grid` or **`base: o=glue`**
  - Walk through the tree, check the different LDAP entries and their attributes

Exercises

- Discover the **hierarchical registration system** implemented by the NorduGrid **Index Services** (top and country EGIISes):
- Check the content of a Top NorduGrid EGIIS:

```
ldapsearch -h index1.nordugrid.org -p 2135 -x -b "mds-vo-name=NorduGrid, o=grid" -s base giisregistrationstatus
```

- Check the content e.g of the Country EGIIS of Norway:

```
ldapsearch -h <host name obtained from the top GIIS> -p 2135 -x -b "mds-vo-name=Norway, o=grid" -s base giisregistrationstatus
```

  - Unfortunately, there are no graphical browsers for the EGIIS hierarchy

Exercises

- Monitoring & dashboards
  - Fire up the Grid monitor from www.nordugrid.org
  - Browse the ATLAS dashboard: http://dashboard.cern.ch/atlas/
  - SAM: http://operations-portal.in2p3.fr/ (“other tools” ->…)
  - Yet another site monitoring database, the GOCDB: http://goc.egi.eu/