



## EGEE / ARC Interoperability Status

Michael Grønager, PhD  
(UNI-C / NBI / EGEE)  
for the NorduGrid collaboration  
*September 28, 2005, Culham*

- Interoperability background
- NorduGrid/ARC in 2 minutes
- Interoperability Status:
  - Multi middleware clusters
  - Common Interfaces
  - Mediated/adapted interoperability: Gateways
- Roadmap



## Interoperability background

- Many sites have deployed ARC (~50sites ~5000CPUs)
- Site wish to provide compute power via ARC
- Experiment support are keyed towards LCG
  
- Interoperability meeting in Rome in February 2005
- Interoperability meeting at CERN in August 2005

# Interoperability background

- Proposed work plan from the CERN meeting:
  - Short term: Multiple Middlewares at large sites
  - Medium term: Gateways between grids
  - Long term: Common Interfaces
  
- Short term is being addressed already
- Work plan made for Medium term tasks
- Long term: CRM Initiative, GLUE2, GGF, GT4



# NorduGrid/ARC in 2 minutes

- History
- The present ARC grid
- Architecture and components
- Distribution



# NorduGrid history

- 2001–2002: a part of the NORDUNet2 program, aimed to enable Grid middleware and applications in the Nordic countries
  - Middleware: EDG
  - Applications: HEP (ATLAS), theoretical physics
  - Participants: academic groups from 4 Nordic countries
    - **Denmark:** Research Center COM, DIKU, NBI
    - **Finland:** HIP
    - **Norway:** U. of Bergen, U. of Oslo
    - **Sweden:** KTH, Stockholm U., Lund U., Uppsala U. (ATLAS groups)
- Since end–2002 is a research collaboration between Nordic academic institutes
  - Open to anybody, non-binding
- Hardware: mostly rental resources and those belonging to users
- Since end–2003 focuses *only* on middleware
  - Develops own Grid middleware: the *Advanced Resource Connector (ARC)*
  - 6 core developers, many contributing student projects
  - Provides middleware to research groups and national Grid projects
- ARC is now installed on ~50 sites (~5000 CPUs) in 13 countries all over the World

# ARC history

- NorduGrid had strong links with EDG
  - WP6: active work with the ITeam; Nordic CA
  - WP8: active work with ATLAS DC1
  - WP2: contribution to GDMP
  - Attempts to contribute to RC, Infosystem
- Had to diverge from EDG in 2002
  - January 2002: became increasingly aware that EDG
    - Is not suitable for non-dedicated resources with a non-CERN OS
    - Won't deliver a production-level middleware in time
  - February 2002: developed own lightweight Grid architecture
  - March 2002: prototypes of the core services in place
  - April 2002: first live demos ran
  - May 2002: entered a continuous production mode
- Since 2004, used by more and more national Grid projects, not necessarily related to NorduGrid or HEP/CERN

# ARC Grid



- A Grid based on ARC middleware
  - Driven (so far) mostly by the needs of the LHC experiments
  - One of the world's largest production-level Grids
- Close cooperation with other Grid projects:
  - EU DataGrid (2001-2003)
  - SWEGRID, DCGC ...
  - NDGF
  - LCG
  - EGEE
- Assistance in Grid deployment outside the Nordic area
- Recently introduced: the **ARC Community VO** to join those who share their resources

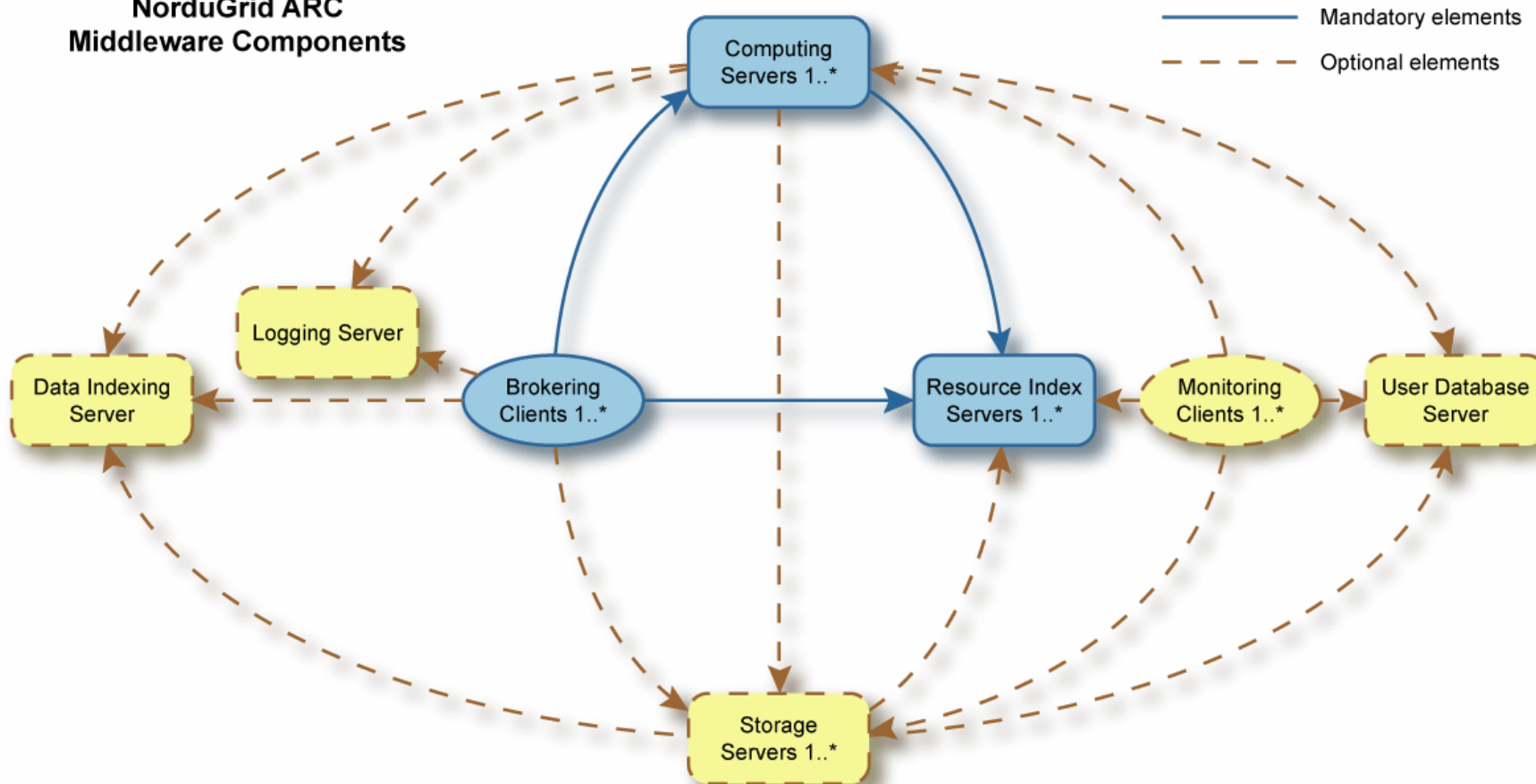


# Architecture

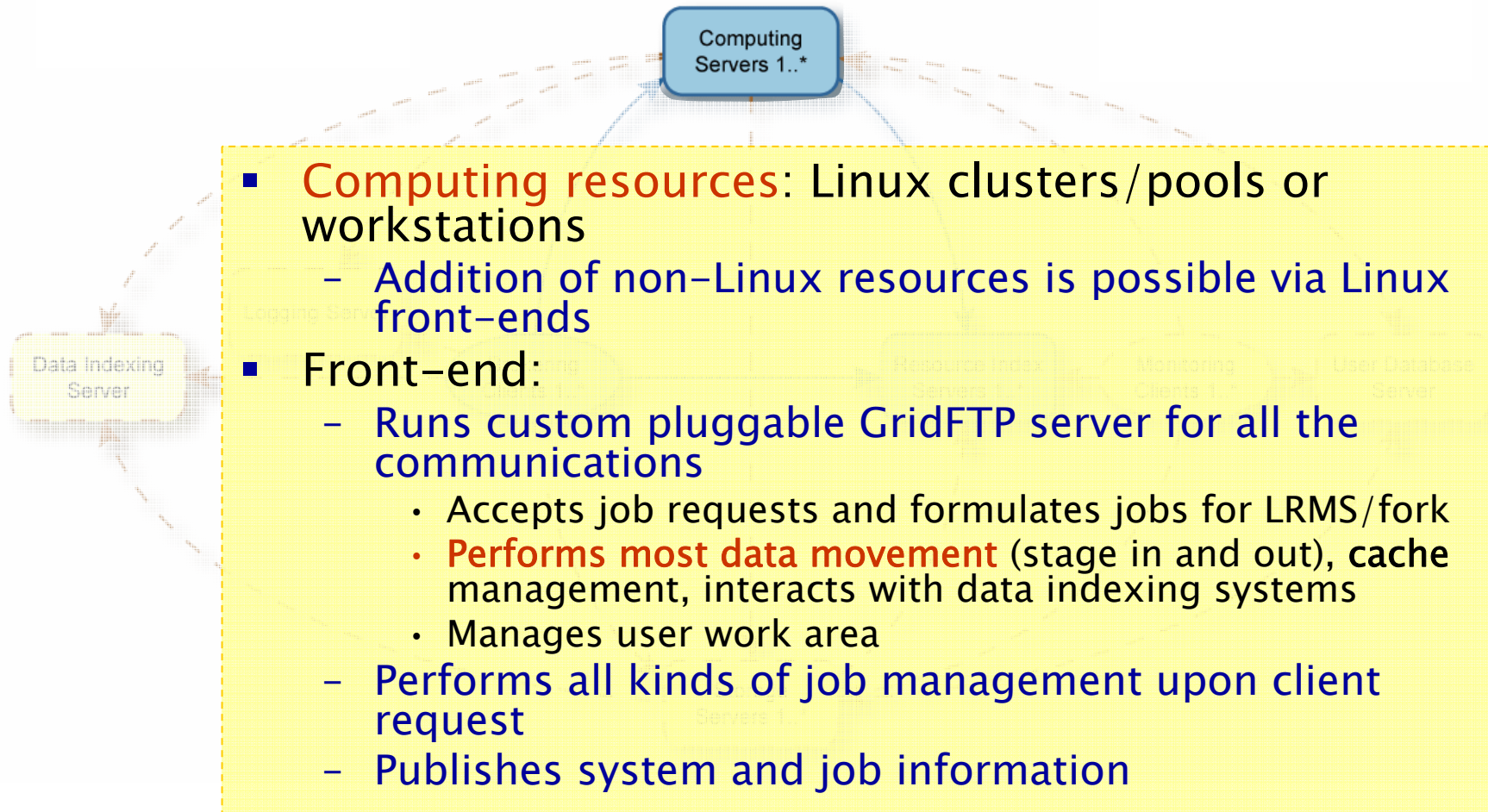
- Each resource has a front-end
  - Authenticates users, interprets tasks, interacts with LRMS, publishes info, moves data, supports non-Linux WNs
- Each user can have an independent lightweight brokering client (or many)
  - Resource discovery, matchmaking, job submission and manipulation, monitoring
- Grid topology is achieved by an hierarchical, multi-rooted set of indexing services
- Monitoring relies entirely on the information system
- Ad-hoc data management, for the beginning

# Components

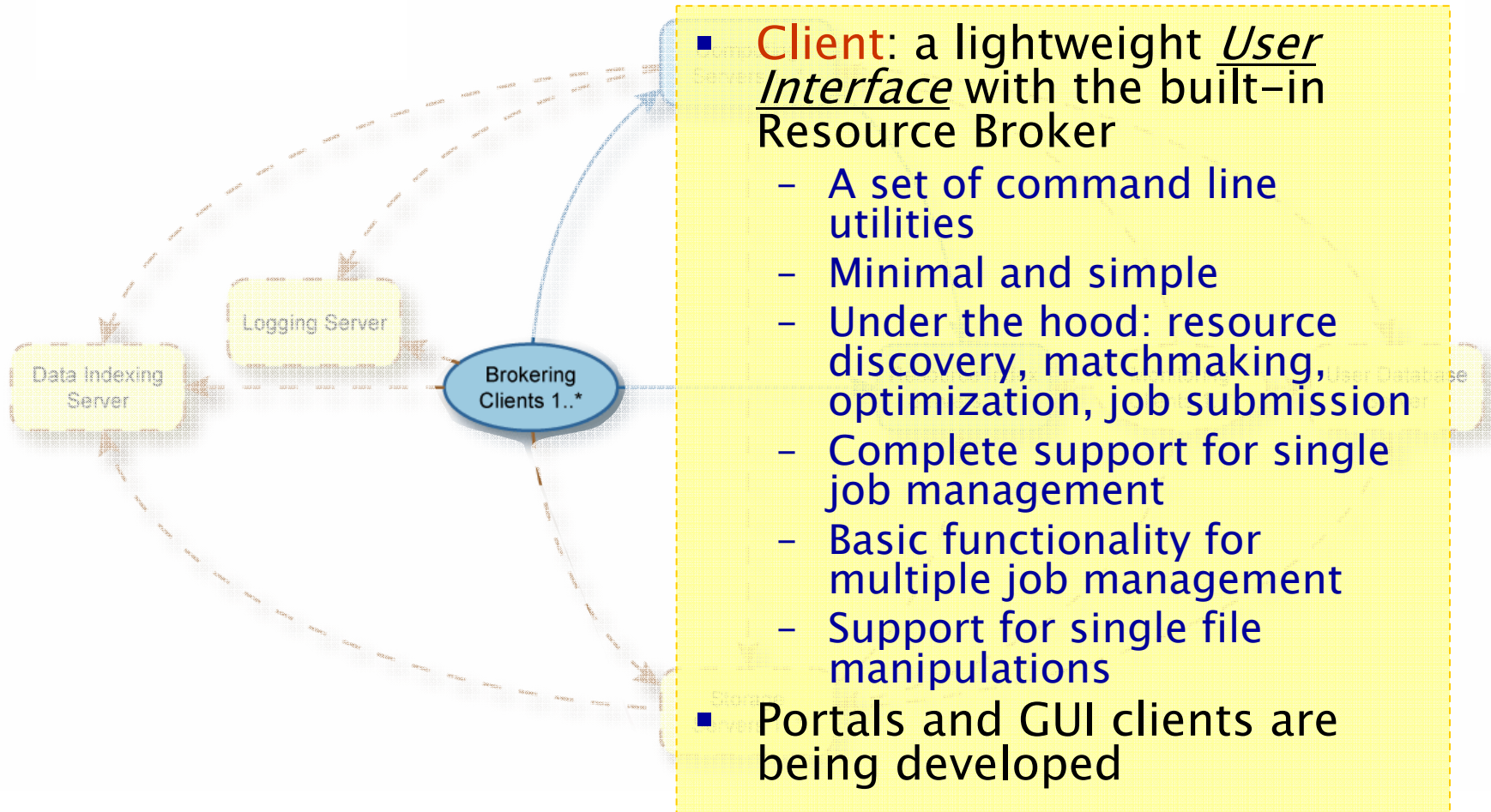
## NorduGrid ARC Middleware Components



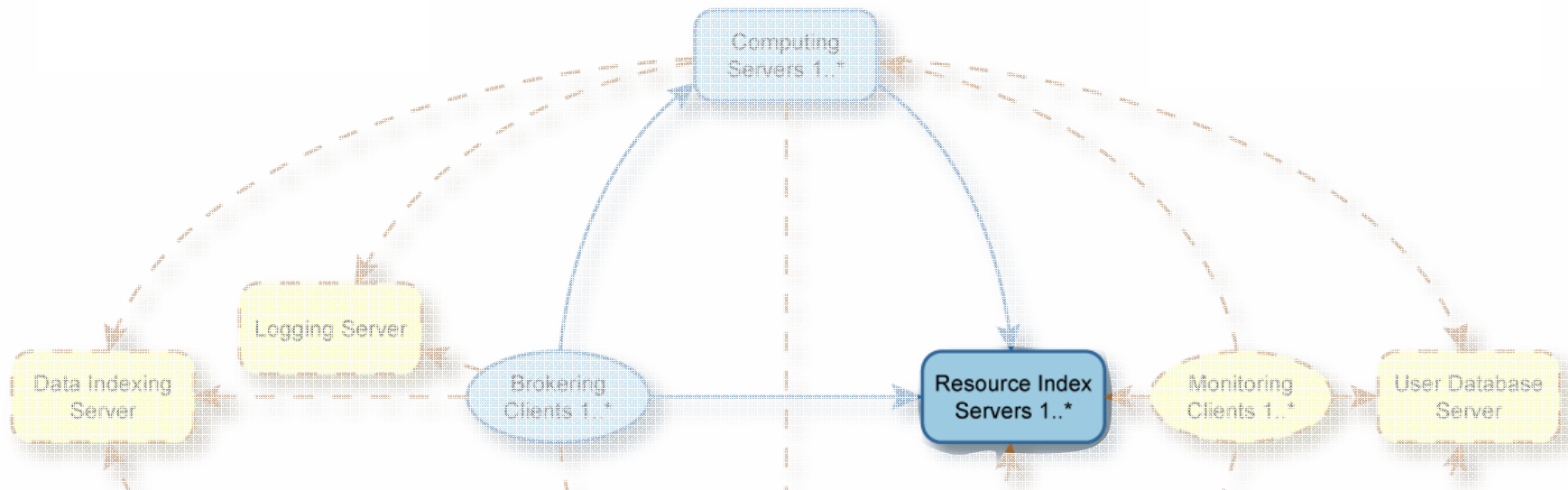
# Components



# Components



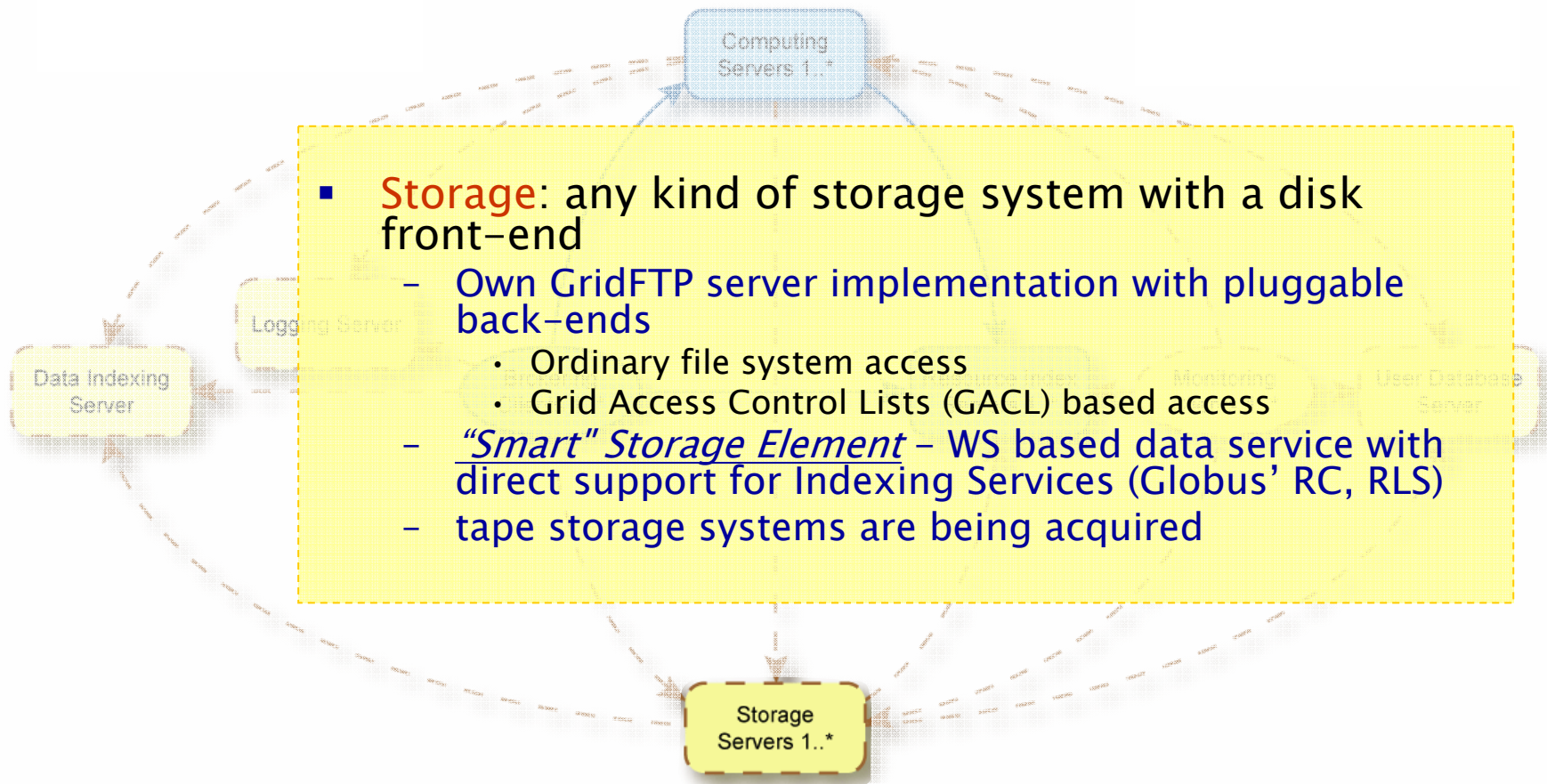
# Components



- **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



# Components





## Distribution, availability

- At [ftp.nordugrid.org](http://ftp.nordugrid.org):
  - Stable releases, including:
    - Binary RPMs and tar-balls are available for most Linux platforms
    - Source RPMs
    - Standalone client tar-ball for installation by a non-privileged user
      - Only 13 MB when unpacked
      - Contains all the EU Grid PMA approved CA keys
      - Includes all the basic Globus client tools
  - Weekly development builds
  - Nightly builds
- CVS at [cvs.nordugrid.org](http://cvs.nordugrid.org)
- License: GPL
- More info, complete documentation, contacts at [www.nordugrid.org](http://www.nordugrid.org)

# Multi middleware clusters

- Simplest form of interoperability is co-existence
  
- Clusters with ARC/LCG co-existing
  - Germany: FZK (1200CPUs)
  - Switzerland: PHOENIX (35CPUs)
  - Sweden: Ingrid, Bluesmoke (200CPUs)
  - Denmark: Morpheus(almost there), Steno(scheduled) (400CPUs)
  
- Not nice for site administrators though –need to install/maintain two middlewares...



# Common Interfaces

- Diverges mainly on:
  - Job submission and description
    - However, work on JSDL is in progress
  - Information system

Service/component	LCG-2, gLite	ARC
Basis	GT2 from VDT	GT2 own patch, GT3 pre-WS
Data transfer	<b>GridFTP</b> , SRM v? (DPM)	<b>GridFTP</b> , SRM v1.1 client
Data management	EDG RLS, <b>Fireman</b> & Co, LFC	RC, RLS, <b>Fireman</b>
Information	<b>LDAP</b> , GLUE1.1, <b>MDS</b> +BDII, R-GMA	<b>LDAP</b> , ARC schema, <b>MDS</b> -GIIS
Job description	JDL (based on classAds)	RSL
Job submission	Condor-G to GRAM	GridFTP
VO management	<b>VOMS</b> , gLite VOMS, CAS (?)	<b>VOMS</b>

# Common Interfaces

## ■ Diverges mainly on:

### Note:

-“Rome” Common Resource Management initiative (includes Globus, UNICORE, LCG, EGEE, NorduGrid, NAREGI) converged on usage of GGF JSDL for job description  
 •JSDL v1.0 is still rudimentary, but is the least common denominator

-ARC now supports JSDL at Grid Manager level.

-xRSL<->JSDL in client is in alpha stage.

Service		
Base		pre-WS
Data transfer	<b>GridFTP</b> , SKM v? (DPM)	<b>GridFTP</b> , SKM v1.1 client
Data management	EDG RLS, <b>Fireman</b> & Co, LFC	RC, RLS, <b>Fireman</b>
Information	<b>LDAP</b> , GLUE1.1, <b>MDS</b> +BDII, R-GMA	<b>LDAP</b> , ARC schema, <b>MDS</b> -GIIS
Job description	JDL (based on classAds)	RSL
Job submission	Condor-G to GRAM	GridFTP
VO management	<b>VOMS</b> , gLite VOMS, CAS (?)	<b>VOMS</b>

# Common Interfaces

## Diverges mainly on:

Note:

- “Rome” Common Resource Management initiative (includes Globus, UNICORE, LCG, EGEE, NorduGrid, NAREGI) converged on usage of GGF JSDL for job description
  - JSDL v1.0 is still rudimentary, but is the least common denominator

Note:

- GLUE2 schema is expected to be developed soon, with participation of NorduGrid, OSG and others. All chances to get a common resource representation.

Service		
Base		e-WS
Data		t
Data management	EDG RLS, <b>Fireman</b> & CO, LFC	RC, RLS, <b>Fireman</b>
Information	<b>LDAP</b> , GLUE1.1, <b>MDS</b> +BDII, R-GMA	<b>LDAP</b> , ARC schema, <b>MDS</b> -GIIS
Job description	JDL (based on classAds)	RSL
Job submission	Condor-G to GRAM	GridFTP
VO management	<b>VOMS</b> , gLite VOMS, CAS (?)	<b>VOMS</b>

# Common Interfaces

## Diverges mainly on:

Note:

-“Rome” Common Resource Management initiative (includes

Note:

-Mark Linesch (GGF Chairman), June 2005:

- “OGSA is in the early stages of development and standardization”

- “GGF distinguishes between the OGSA architectural process, OGSA profiles and specifications, and OGSA software. All of these are important to maintain coherence around OGSA and Grid standards”

- “At the time of writing, we have OGSA Use Case documents, OGSA Architectural documents and drafts of an OGSA Basic Profile document and OGSA Roadmap document. *We do not yet have any OGSA-compliant software implementations or OGSA compliance tests*”

Service		
Basic		WS
Data transfer		
Data management		
Information		GIIS
Job description		
Job submission	Condor-G to GRAM	GridFTP
VO management	<b>VOMS</b> , gLite VOMS, CAS (?)	<b>VOMS</b>

- **Condor / Condor-G**
  - LCG supports submission via Condor-G natively
  - LCG supports Condor as a queuing system
  - ARC supports Condor as a queuing system
  - Cooperation between ARC and Condor led in October 2004 to Condor-G version that can submit jobs to ARC GridFTP (translation from ARC infosystem schema to GLUE was developed by Rod Walker). Was meant to be used by LCG – but nobody configured an RB this way yet
  
- Perhaps the most important common interface?

- Work plan proposed at the CERN meeting:
  - WP1: Document LCG CE (David Smith)
  - WP2: LCG to ARC job submission (Laurence Fields)
  - WP3: ARC to LCG job submission (Balazs Konya)
  - WP4: Service discovery (Laurence Fields)
  - WP5: GLUE2 (long term...)
  
- Work plan about to initiate – next status meeting: CERN October 31<sup>st</sup>

- Possible submission scheme from LCG to ARC
  - Setup an LCG CE using Condor LRMS
  - Setup Condor-G queue to submit to ARC
  
- Possible submission scheme from ARC to LCG
  - Setup an ARC CE using Condor LRMS
  - Setup Condor-G queue to submit to LCG

# Conclusions

- Initiate gateway work plan
  - (rather sooner than later)
  
- Tight schedule – LHC is closing in...
  
- ARC Interoperability part of EGEE–II SA3
  - Enables easy use of several extra CPUs in LCG
  - Ensures gateway support



End of Slideshow

Extra Slides...



# Major Architectural Differences

- Problem: Globus GRAM submission does not work for many jobs:
  - Each new queued job spawns a process on the gate keeper, which regularly executes a Perl script
  - Does not perform for more than 400 jobs
- Solutions:
- LCG:
  - Condor-G Grid Monitor: kills listening process and replaces it with a per user process forked on the gate keeper
- ARC:
  - Own Grid Manager: Interacts with local queuing system, submission via GridFTP.

# Major Architectural Differences

- Problem: Globus MDS badly implemented:
  - Problems with caching
  - Schema not complete
- Solutions:
- LCG:
  - Berkeley DB MDS implementation: BDII
  - Introduces GLUE schema
- ARC:
  - Not using cache information system but MDS in file sharing like network with super nodes.
  - Additions to MDS schema

# Major Architectural Differences

- Problem: Globus has no broker:
- Solutions:
- LCG:
  - Central broker service
- ARC:
  - Brokering done on the fly by client