



KnowARC

# NorduGrid and ARC

*Katarina Pajchel*

*katarina.pajchel@fys.uio.no*

*University of Oslo,*

*Outreach officer KnowARC*



- NorduGrid and the origins of ARC
- ARC architecture and development
- User communities



- ❖ ARC – Advanced Resource Connector
  - ❖ Grew out of High Energy Physics (HEP) but now majority of users are not from HEP community
  - ❖ Production quality software since 2002
  - ❖ Continuous use since 2002
  - ❖ Usage spans 17 countries
- ❖ NorduGrid
  - ❖ A Grid Research and Development collaboration
  - ❖ Aiming at development, maintenance and support of the ARC Grid middleware
- ❖ KnowARC – EU founded project 2006 – 2009
  - ❖ Extent and re-design the ARC-middleware





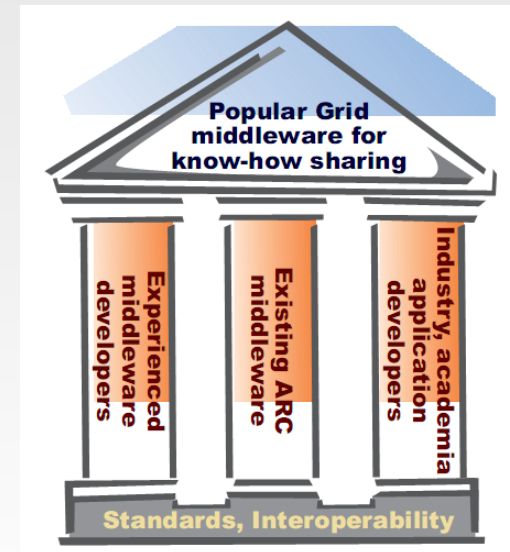
- ❖ The mission of KnowARC is
  - ❖ Create a novel, powerful next generation Grid middleware
  - ❖ To promote Grid standardisation and interoperability
  - ❖ To contribute to Grid technology take-up
- ❖ Currently 9 partners across 8 countries
  - ❖ Both academic and commercial partners, with a wealth of experience in scientific and Grid computing



emergence  
tech



- ❖ General purpose Open Source European Grid middleware
  - ❖ Lightweight, robust, reliable, scalable
  - ❖ Non-intrusive on the server side
  - ❖ Flexible & powerful on the client side
  - ❖ Easily portable
  - ❖ Strong commitment to standards and interoperability
- 
- ❖ ... many nice words, but there are working solutions behind

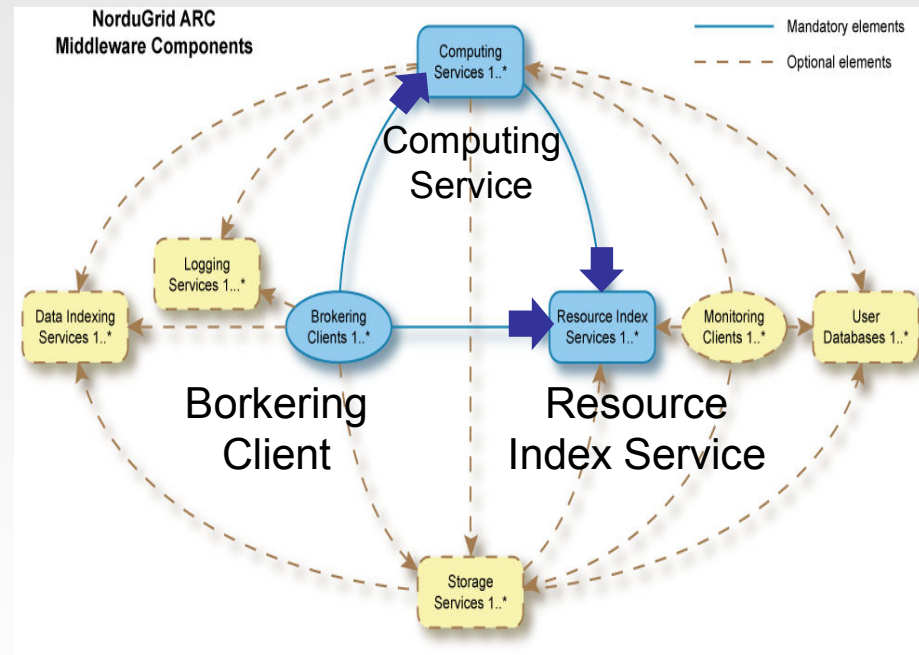


# ARC design principles



KnowARC

- ❖ User- and performance-driven development
  - ❖ start with something simple that works for users and add functionality gradually
- ❖ Simplicity is the key-word
- ❖ **Today: ARC Classic**
  - ❖ Reliable implementation of fundamental Grid services
  - ❖ Avoid centralized services
  - ❖ Only three mandatory components
    - Computing Service
    - Resource Index Service
    - Brokering Client





## ❖ From a users point of view

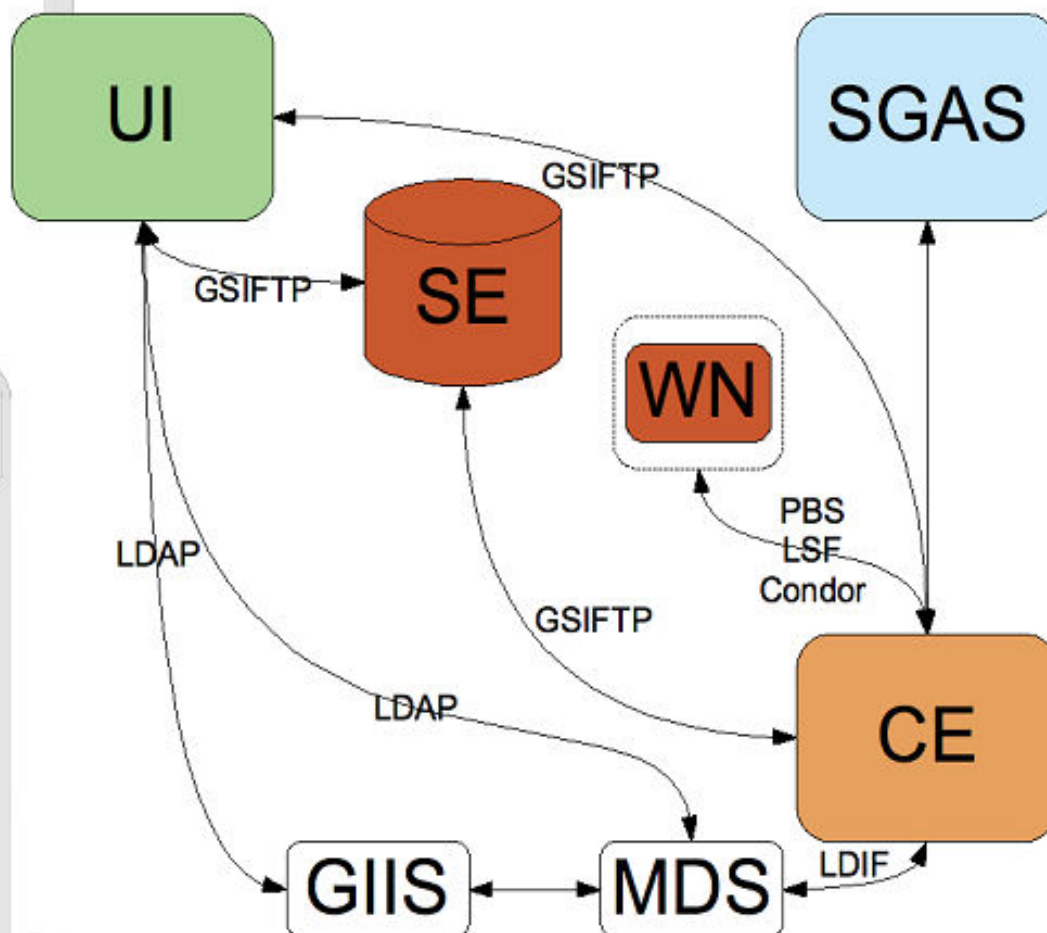
- ❖ Lightweight simple to install client (.tar), also for non – privileged users
- ❖ Simple to use
- ❖ Powerful brokering capability and job management

## ❖ From sysadm point of view

- ❖ Computing Service – GridFTP-Grid Manager pair of services
- ❖ Data movement and interface to local resource management through backends
- ❖ Clear separation between the local batch system and the grid
  - No middleware on the nodes
- ❖ Non-intrusive, secure, easy to deploy and minimal maintenance effort

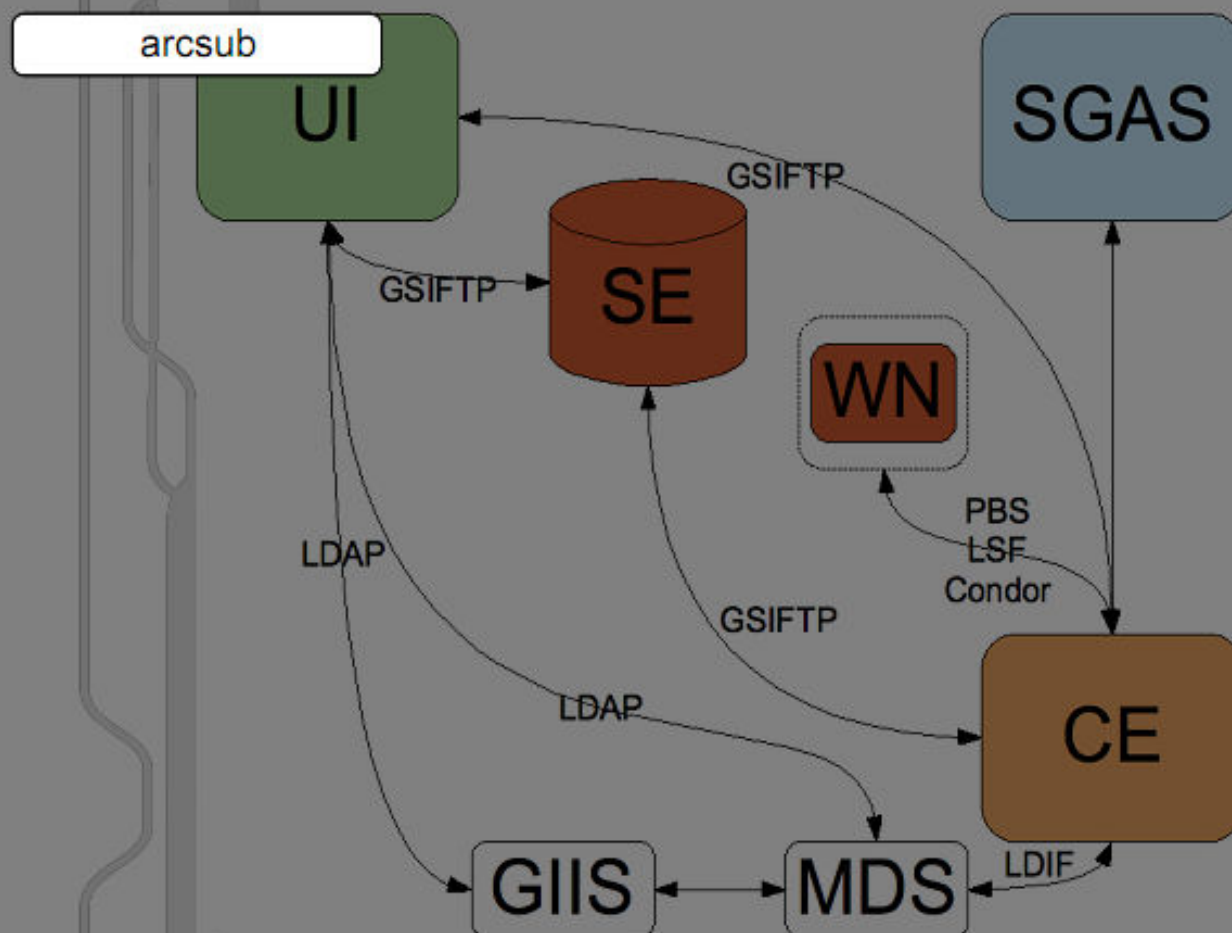


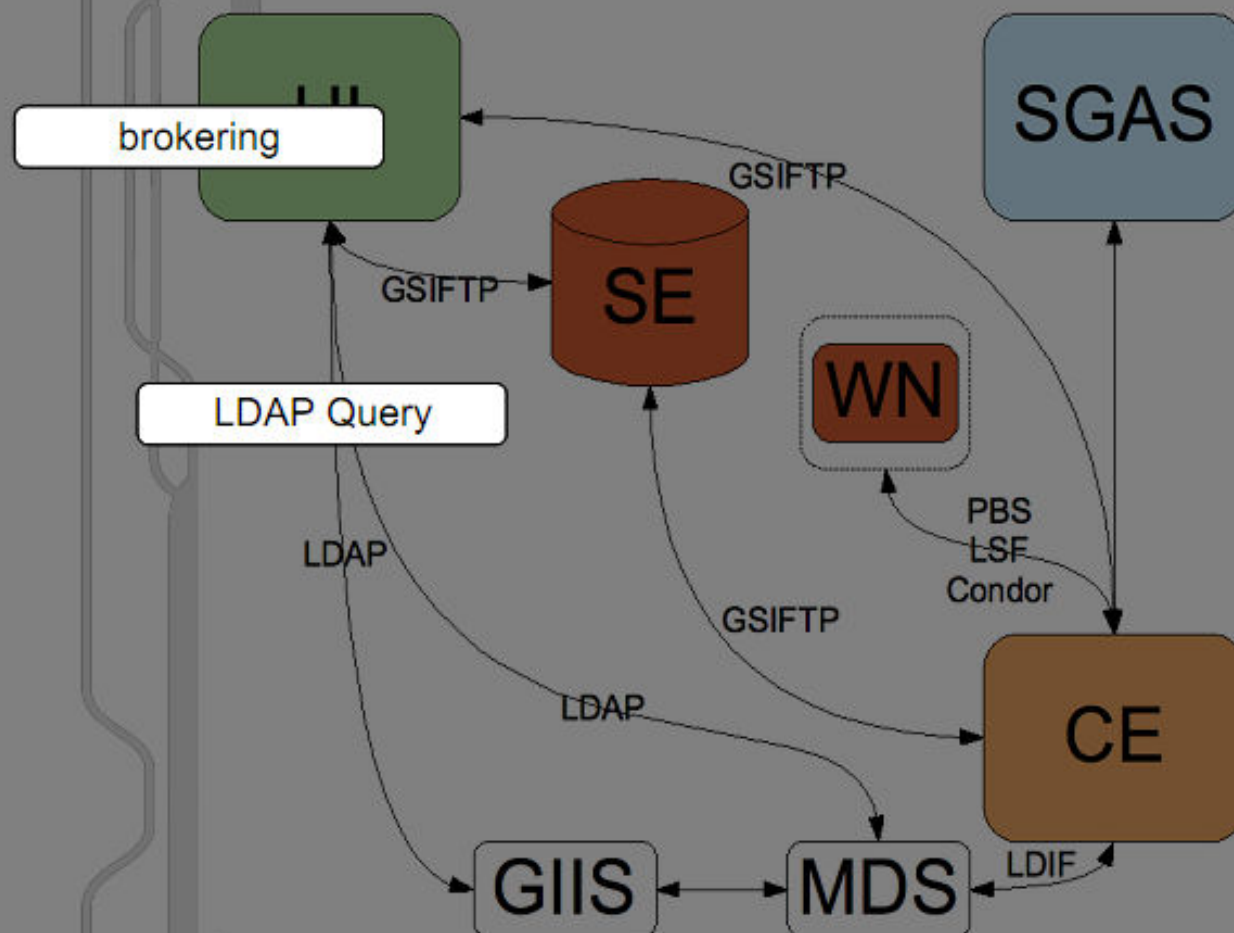
<http://download.nordugrid.org/>



Slides from  
Michael  
Gronager,  
NDGF

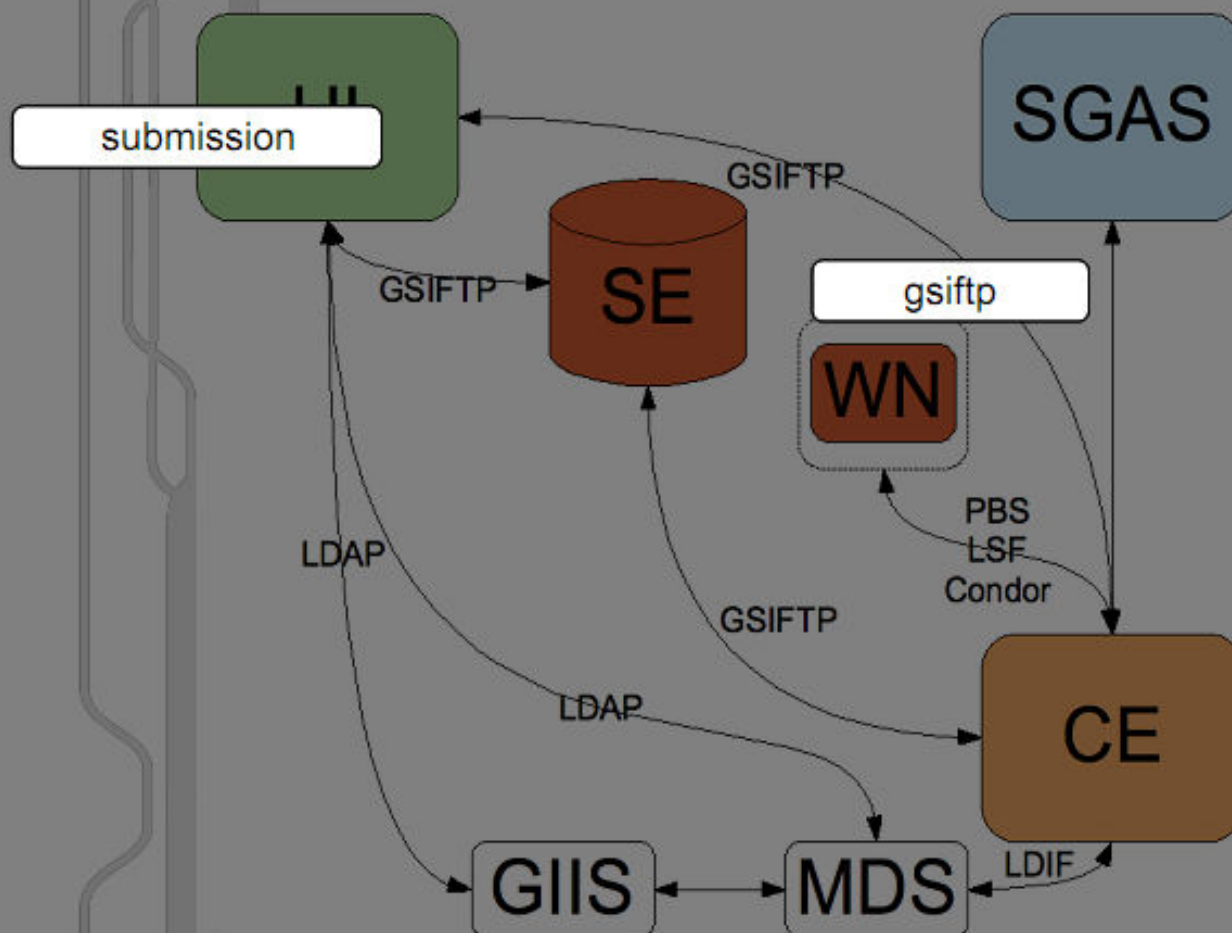






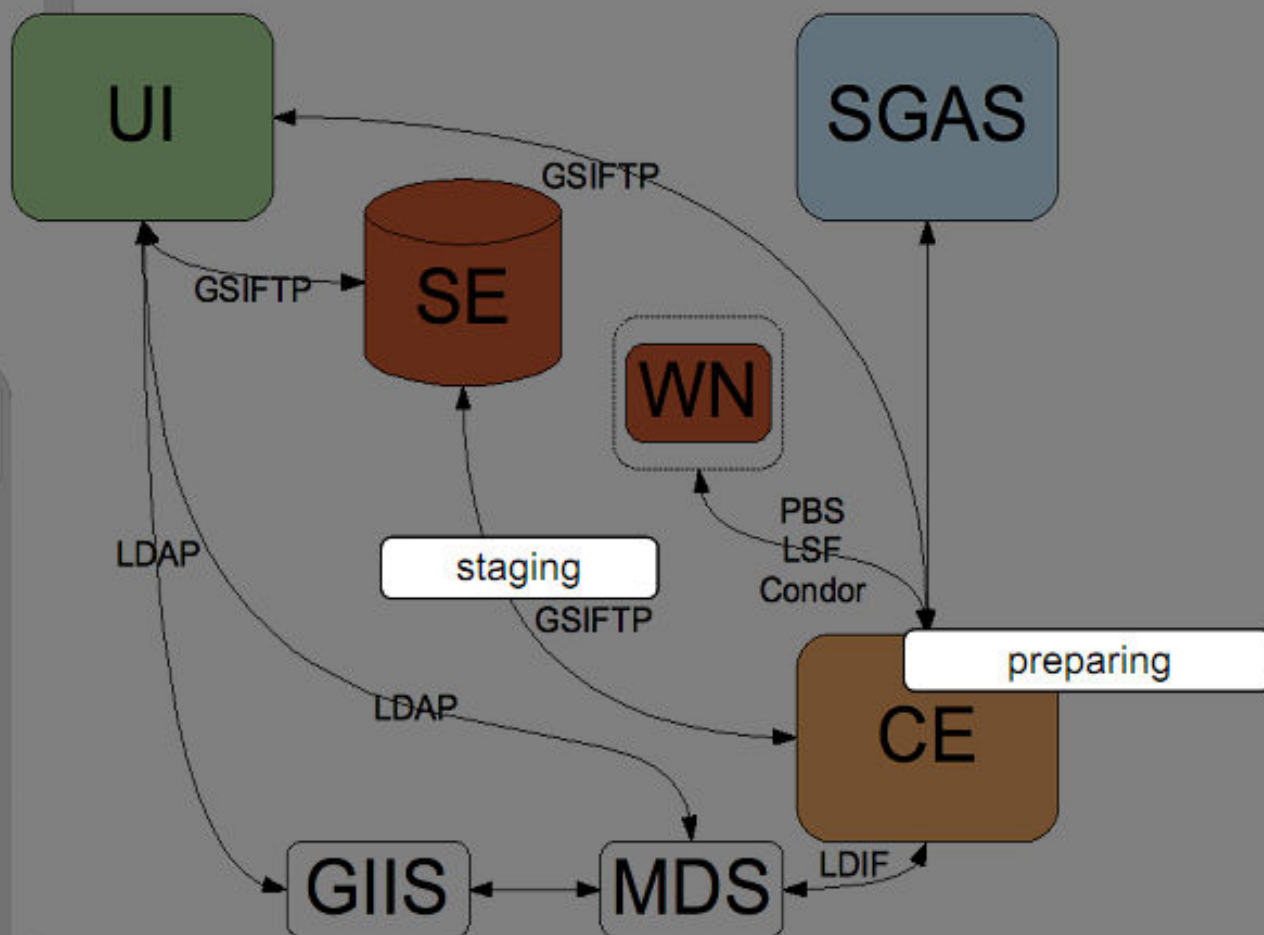


# JobCycle: ARC



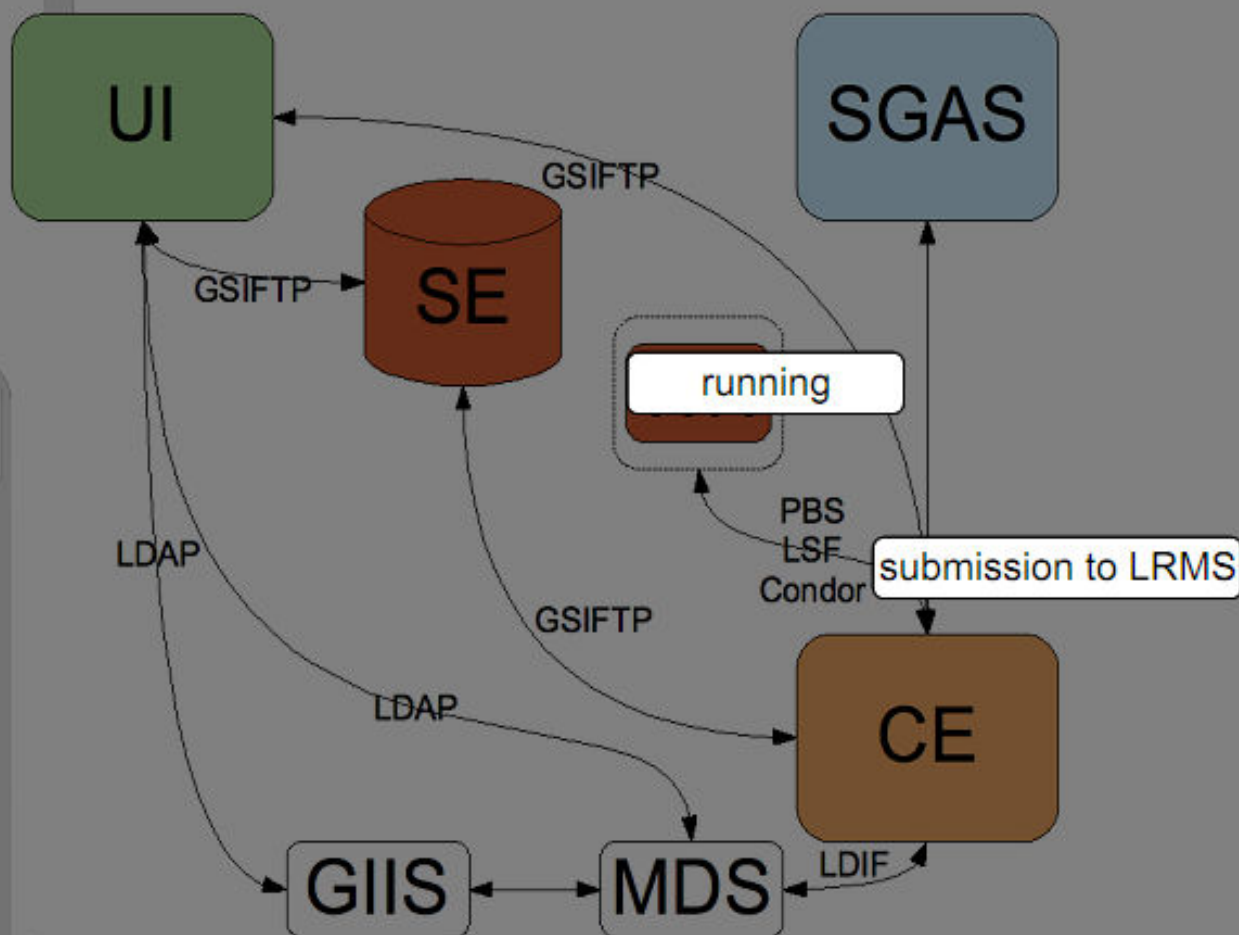


# JobCycle: ARC



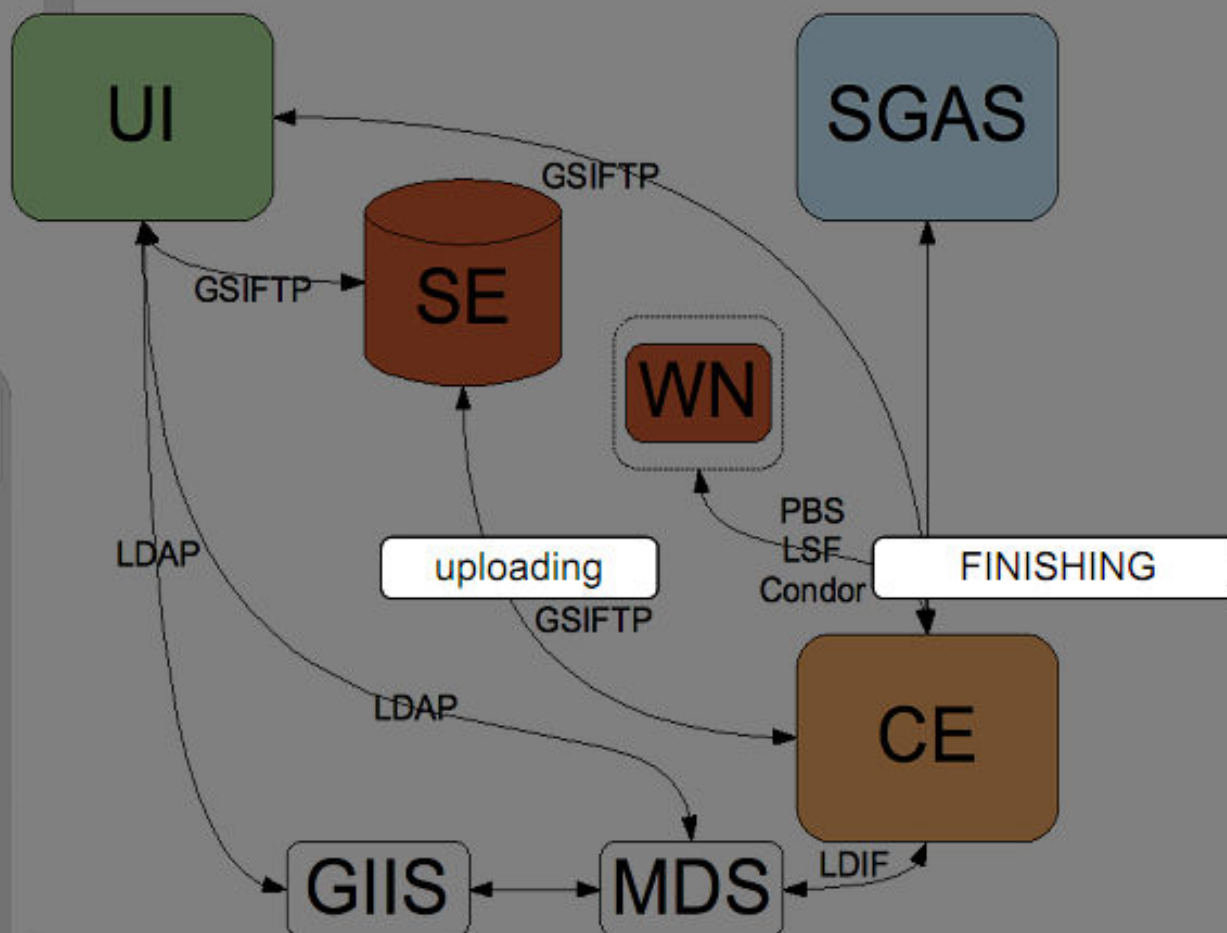


# JobCycle: ARC



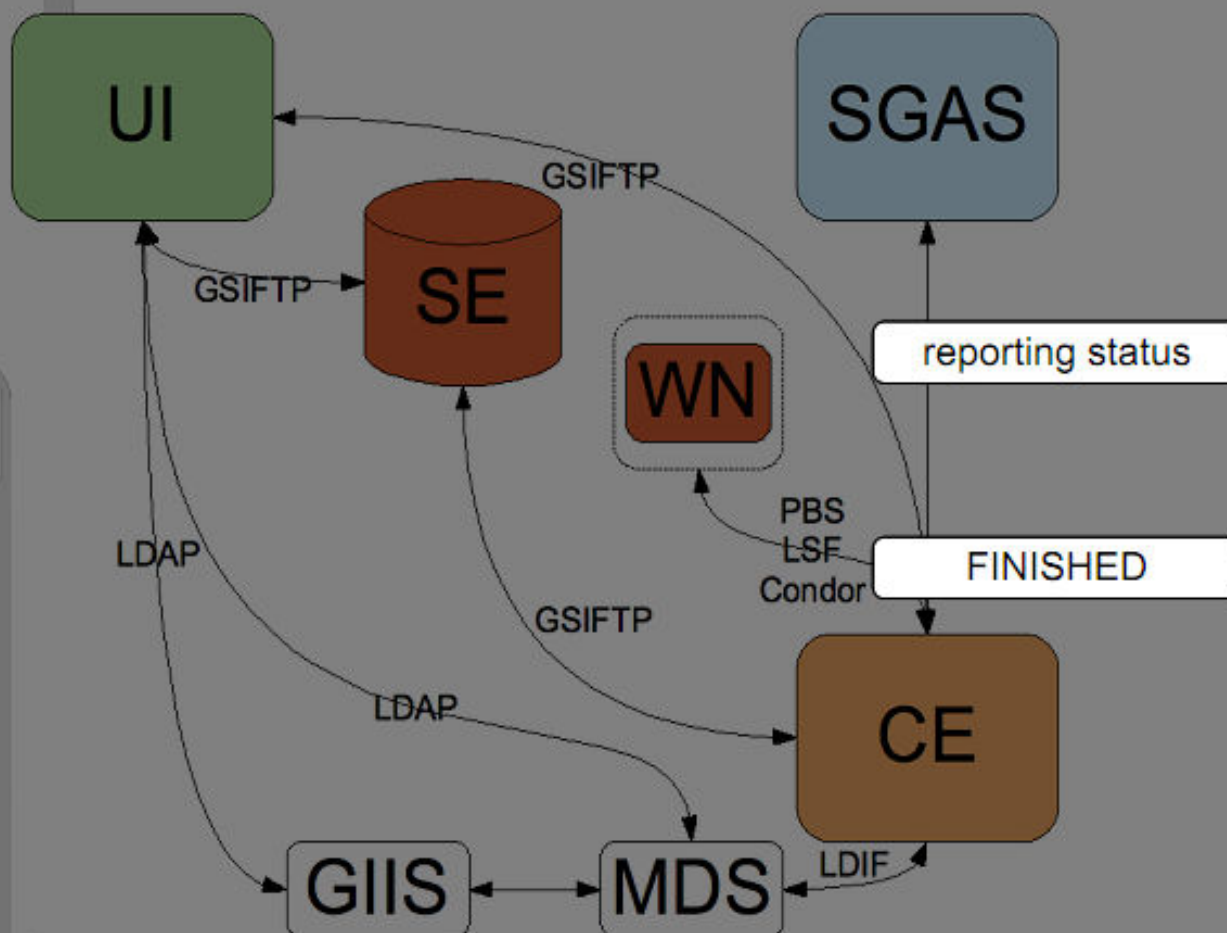


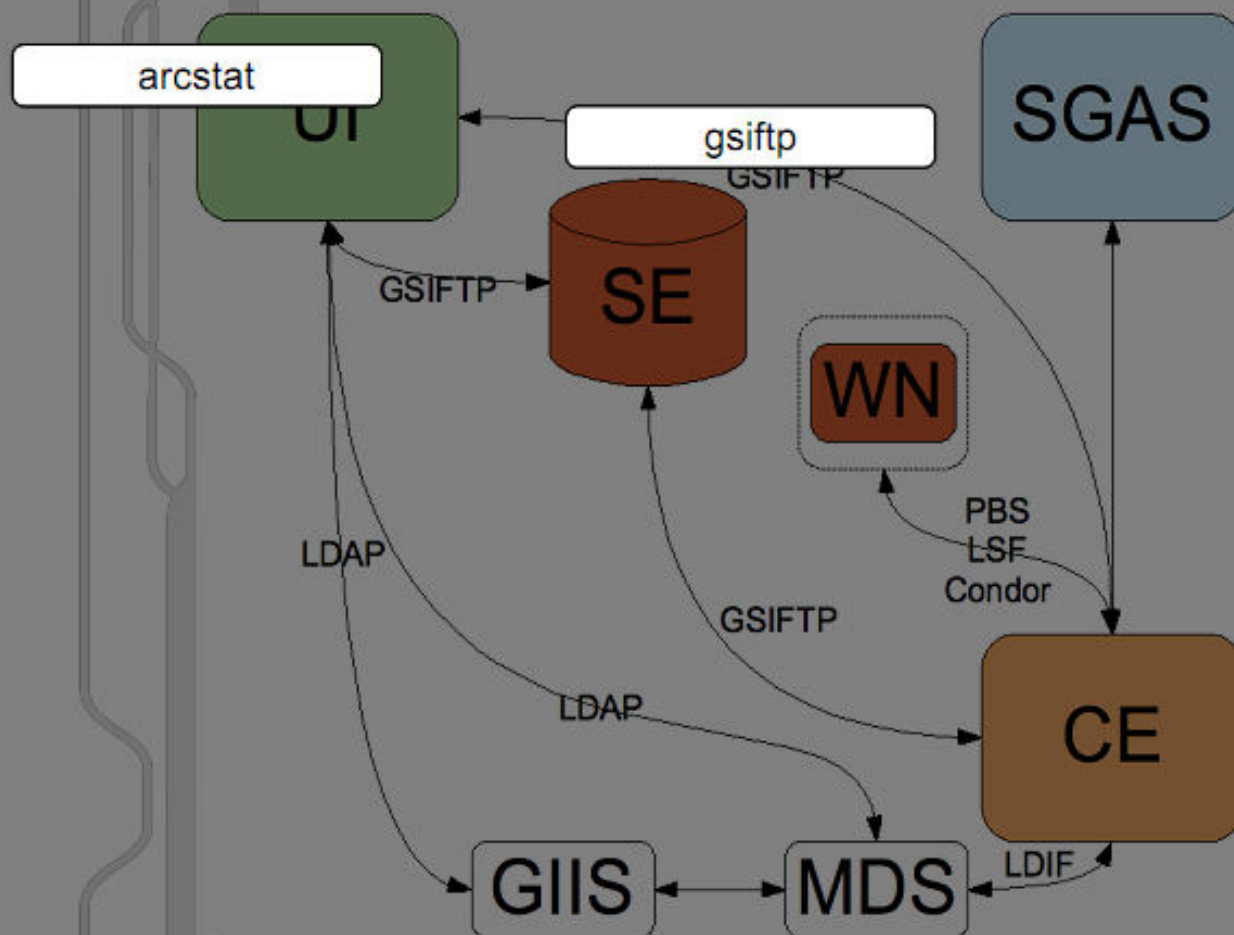
# JobCycle: ARC





# JobCycle: ARC







# ARC – Upcoming release



KnowARC

- ✦ KnowARC is developing the next generation ARC
- ✦ Design principles
  - ❖ Build on the success of current ARC middleware
  - ❖ No single point of failures, no bottlenecks – **stable by design**
  - ❖ Non-intrusive : tools and utilities respect local owner policies, in particular, security related
  - ❖ Easy to maintain and utilize
  - ❖ Standards compliant

From



towards



... and no  
exaggerations

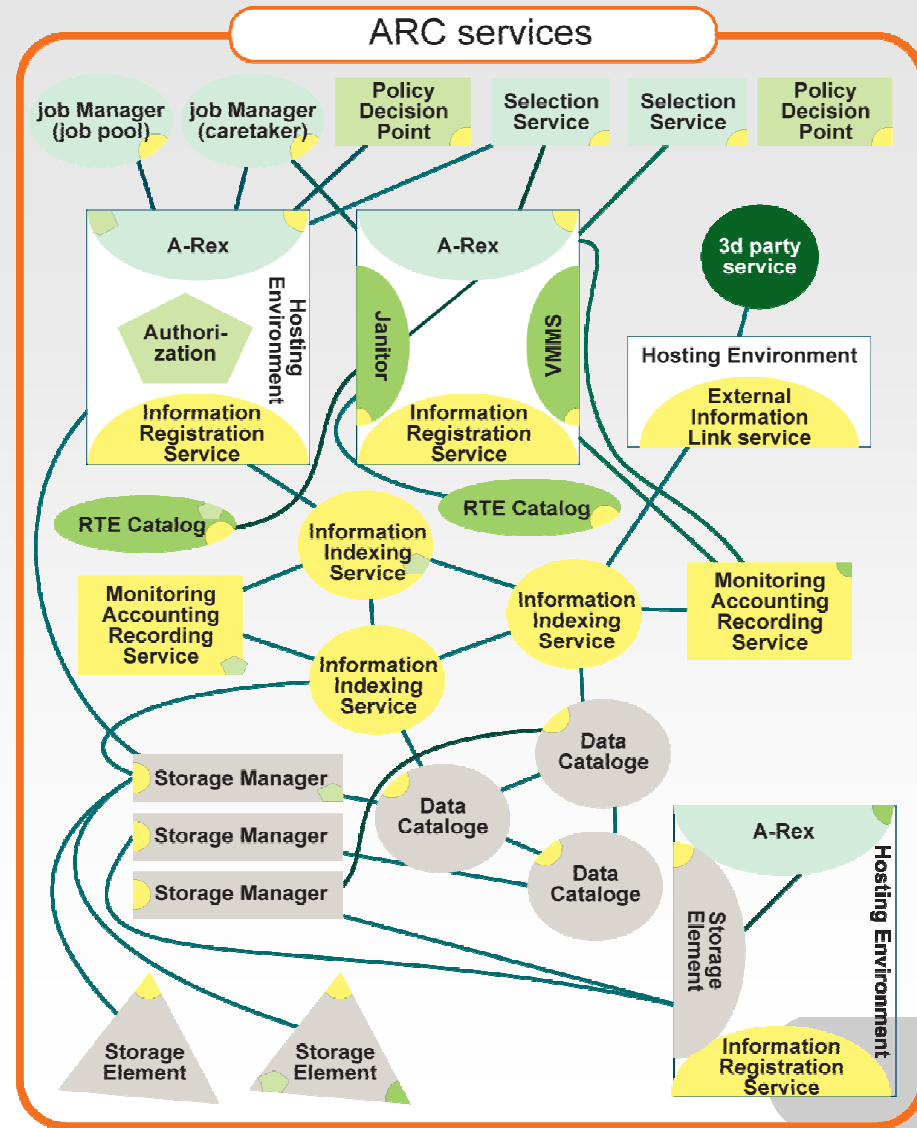


# What is new I



KnowARC

- Service Oriented Architecture
- Modular structure
- Self-sufficient core components



# What is new II



KnowARC

- ❖ Interoperability built on standard
- ❖ User and developer friendly
- ❖ Business friendly Open Source license Apache 2.0
- ❖ Portable
  - ❖ runs on almost all Linux variants, Solaris
  - ❖ porting to Windows and Mac OS in progress
- ❖ Aiming at integration into Fedora, Debian and Ubuntu



7/4/2009



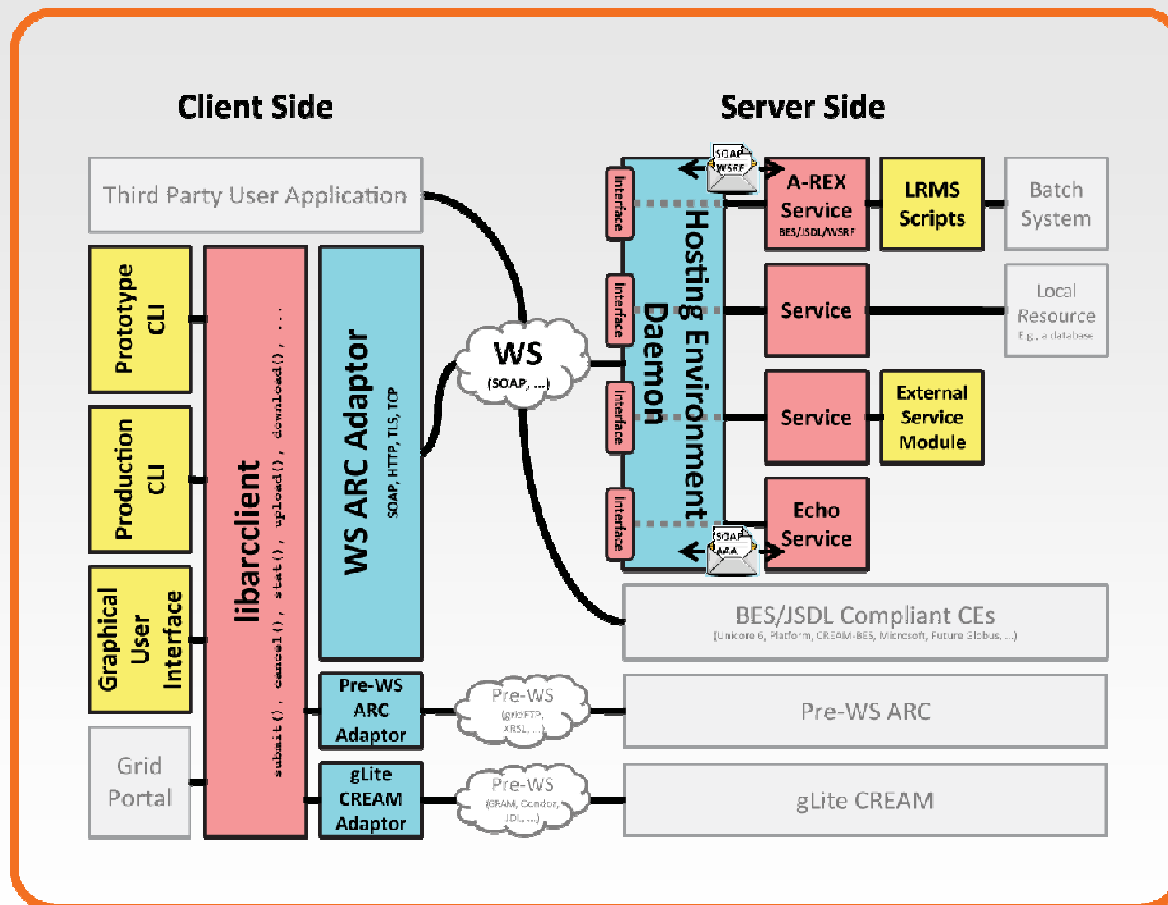
[www.knowarc.eu](http://www.knowarc.eu)

# ARC WS-based components



KnowARC

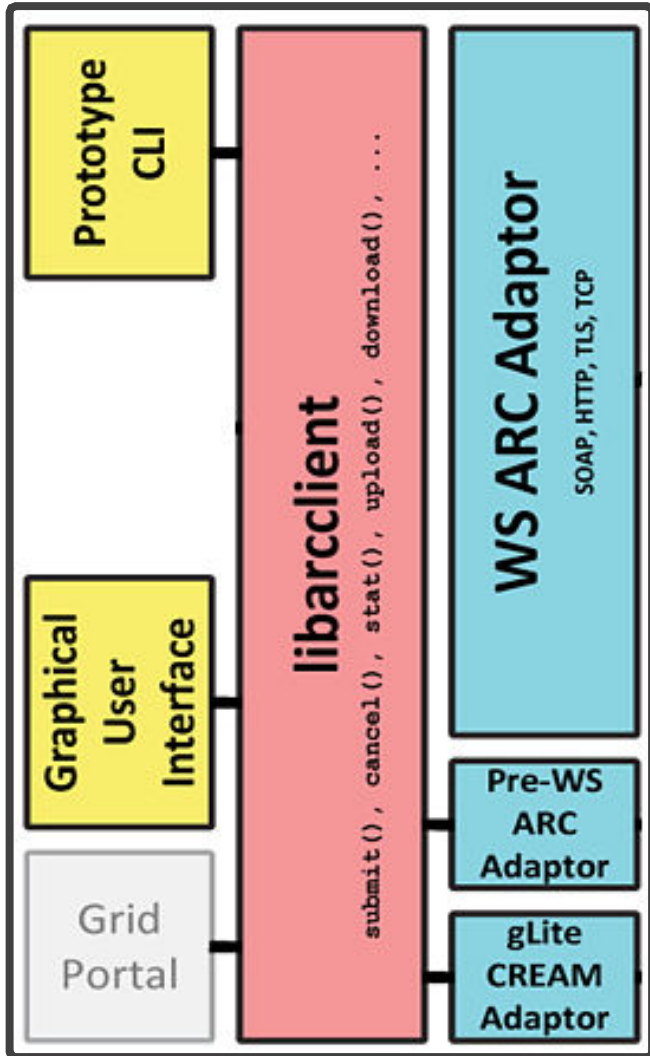
## Internal structure of ARC components



# Key Feature - New ARC client



KnowARC



- ✦ Relies on dedicated library
  - ❖ Implemented in C++
  - ❖ Python and Java bindings
  - ❖ Allows easy development of application-specific clients
- ✦ Implements a user Grid toolbox
  - ❖ Handling of user & host credentials
  - ❖ Resource discovery & matchmaking & brokering & job submission
  - ❖ Input/output data handling

- ✦ The new library and arc\* commands can handle glite-CREAM and UNICORE

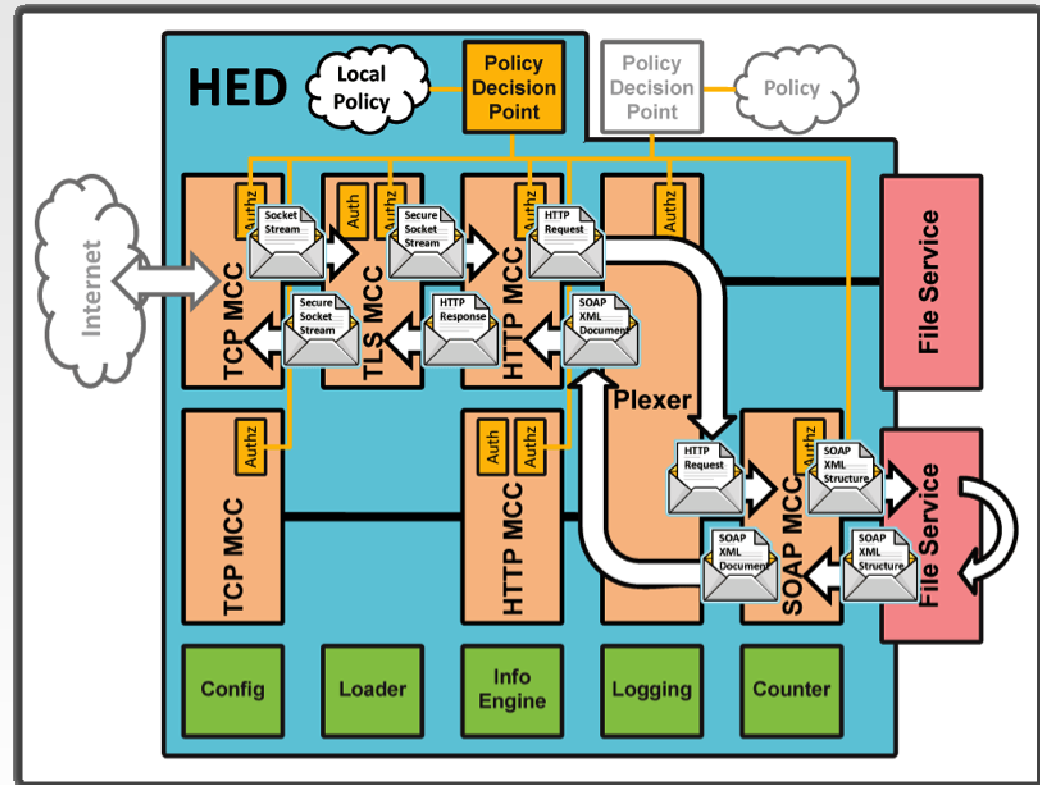


- ✦ Windows and Mac OS client
- ✦ GUI – user interface, just delivered !

# Key Feature - HED



- ✦ HED – The Hosting Environment Daemon
- ✦ Container for all the server-side functional components
- ✦ Main functions:
  - ❖ Rout messages between the services and the outside world
  - ❖ Provide inter service communication
- ✦ Provides a basic security infrastructure (SSL/GSI)
- ✦ Consists of pluggable modules
- ✦ Light-weight
  - ❖ no Apache, no Axis





- ✦ ARC Resource-coupled Execution Service
  - ❖ Provides Execution Management capability
  - ❖ The Grid Manager from ARC Classic as core
  - ❖ Extended with WS interface implementing Basic Execution Service (BES)
  - ❖ Accepts Job Submission Description Language (JSDL)
  - ❖ Information and resource discovery – GLUE 2 schema

- ✦ Support for wide range of Local Resource Management Systems:

- ❖ Torque, PBS/OpenPBS, SGE, LoadLeveler, LSF, Condor and SLURM

- ✦ Released in ARC 0.8, available at:

[http://wiki.nordugrid.org/index.php/ARC\\_v0.8](http://wiki.nordugrid.org/index.php/ARC_v0.8)

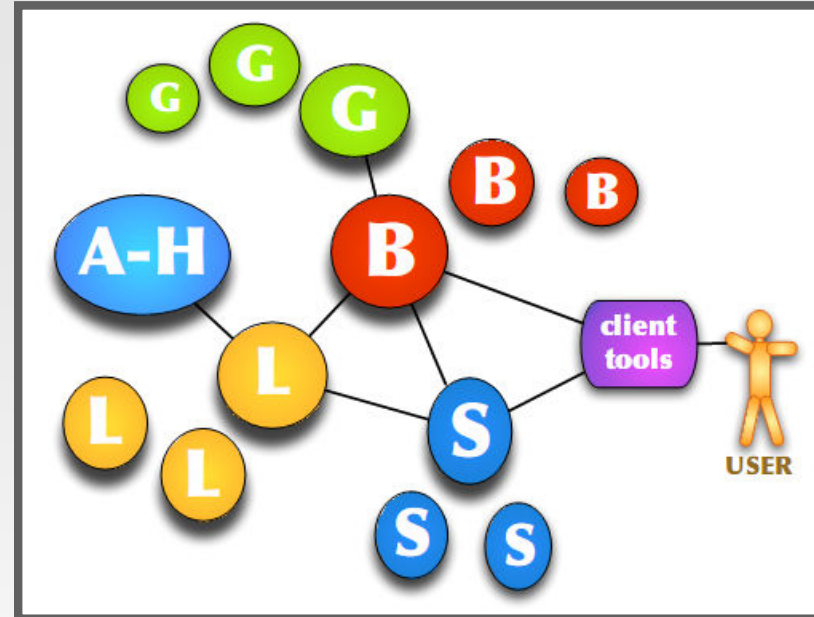


# Key Service – New Storage



KnowARC

- ❖ 'Distributed by Design' storage system
  - ❖ Global namespace
  - ❖ Supports collections and subcollections to any depth
- ❖ **A-Hash** – a replicated database to store metadata
- ❖ **Librarian** – handles:
  - ❖ Metadata and hierarchy of collections and files
  - ❖ The location of replicas
  - ❖ Health data of the shepherd services
- ❖ **Shepherd** – manages storage services, and provides a simple interface for storing files on storage nodes



- ❖ **Gateways** – used to communicate with external storage managers
  - ❖ interoperability
- ❖ **Bartender** - high-level interface for the users and for other services





- ❖ Our main goal is to achieve **interoperability via standards**
- ❖ KnowARC is heavily involved in Open Grid Forum [www.ogf.org](http://www.ogf.org)
- ❖ Next generation ARC implemented using
  - ❖ SRM, GridFTP, X509 (already in use) and BES, JSDL, GLUE 2, SAML, VOMS (VO management)
  - ❖ WS technology standards, common interfaces
- ❖ Target interoperability platforms: gLite, UNICORE
- ❖ Reference document:  
[http://www.knowarc.eu/documents/Knowarc\\_D3.3-1\\_08.pdf](http://www.knowarc.eu/documents/Knowarc_D3.3-1_08.pdf)



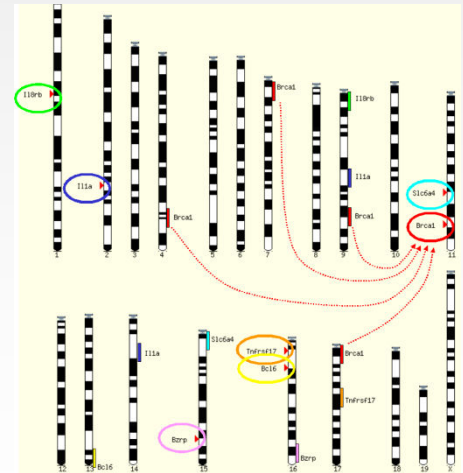
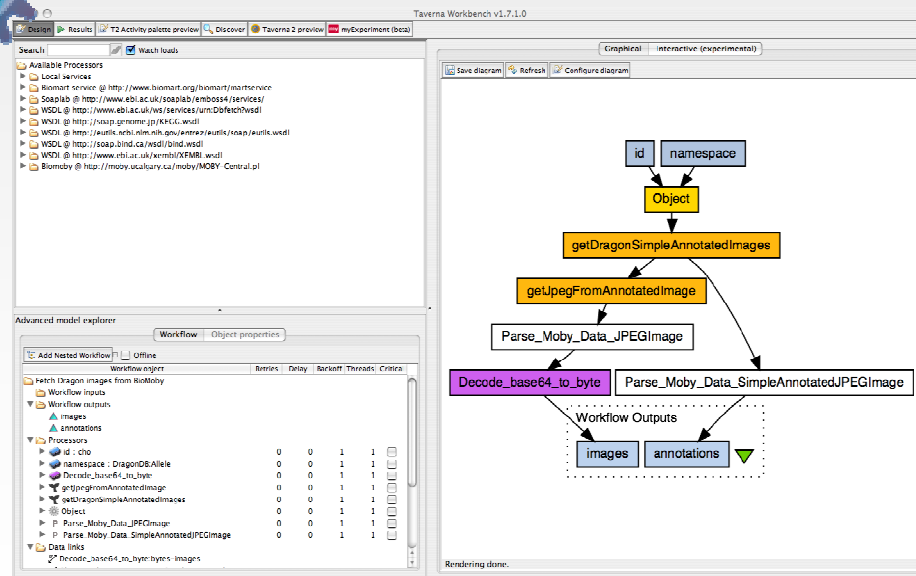
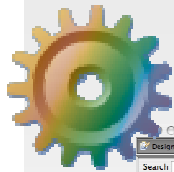


- ❖ ARC is the middleware of choice for a number of Grid infrastructures
  - ❖ [M-Grid](#) materials science project in Finland (Finnish NGI)
  - ❖ [SwiNG](#) Swiss Multiscience Computing Grid
  - ❖ [SweGrid](#) Swedish national computational resource
  - ❖ [BalticGrid-II](#) regional Grid infrastructure (plans to use ARC)
- ❖ Bioinformatics partner in KnowARC
  - ❖ Lubeck University – developing plug-in for Taverna, a workflow management software
- ❖ Healthcare – medical imaging finding tool
  - ❖ University of Geneva
- ❖ High Energy Physics – Distributed Tier1 centre for wLCG
  - ❖ Facilitated by the Nordic DataGrid Facility NDGF
- ❖ ... and other ...



## ❖ Bioinformatics

- ❖ Lubeck University study of polygenic autoimmune disease
- ❖ ARC plug-in for Taverna, a workflow management tool
- ❖ ARC-enabled calculation shows a **3-5 fold better usage of resources** due to uniform accessibility & retrieval of results



# Applications – MedGIFT



KnowARC

- ✦ [MedGIFT](#) demo
- ✦ >70 000 images/day in Geneva Hospital radiology
- ✦ Image retrieval and comparison aid diagnosis



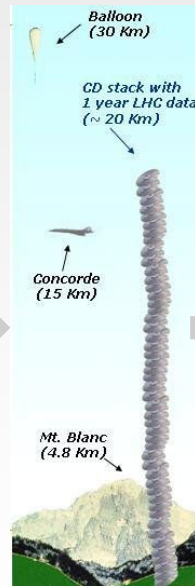
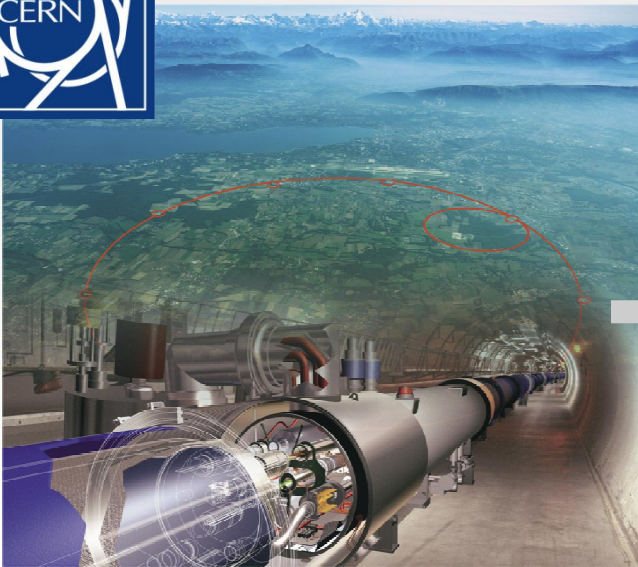
The screenshot displays two browser windows. The left window is the MedGIFT search interface, showing a query of three images and a grid of search results. The right window shows a detailed case view for 'Association de chondrocalcinose et d'une polyarthrite rhumatoïde', including a diagnosis, clinical presentation, and description. An arrow points from the search results to the case view.

# Applications – LHC Tier1



KnowARC

- ✦ NDGF facilitates the Nordic distributed Tier1 for CERN LHC experiments ATLAS, ALICE, CMS
- ✦ ATLAS Monte Carlo production in 2008 > 93% efficiency
  - ❖ Best performing ATLAS Tier center through in 2008





## ❖ International Impact

- ❖ KnowARC actively participates in European level discussion on e-Infrastructures



## ❖ Part of the European Grid Roadmap

- ❖ ARC middleware selected by the European Grid Initiative (EGI) for their Universal Middleware Distribution (UMD)
- ❖ UMD is a plan for an interoperable set of European middleware component backed by the European Commission
- ❖ More information available at <http://knowledge.eu-egi.eu/knowledge/index.php/UMD>





- ❖ The established and well performing ARC middleware is gradually renewed by improved and extended components
  - ❖ New modular WS based architecture
  - ❖ Strong emphasis on standards and interoperability
  - ❖ ... and the same old good quality
- ❖ **The future is now – final deliverables this year – 2009**
  - ❖ Wealth of information on [www.knowarc.eu](http://www.knowarc.eu), [www.nordugrid.org](http://www.nordugrid.org) and [www.ndgf.org](http://www.ndgf.org)
  - ❖ [ARC Community Wiki](#) releases and installation info
  - ❖ ARC code, also the latest news [svn.nordugrid.org](http://svn.nordugrid.org)
  - ❖ Technology Preview ARC design document [http://www.knowarc.eu/documents/Knowarc\\_D1.1-1\\_07.pdf](http://www.knowarc.eu/documents/Knowarc_D1.1-1_07.pdf)