

# GRID RESOURCE BROKER

**ABHISHEK SINGH**

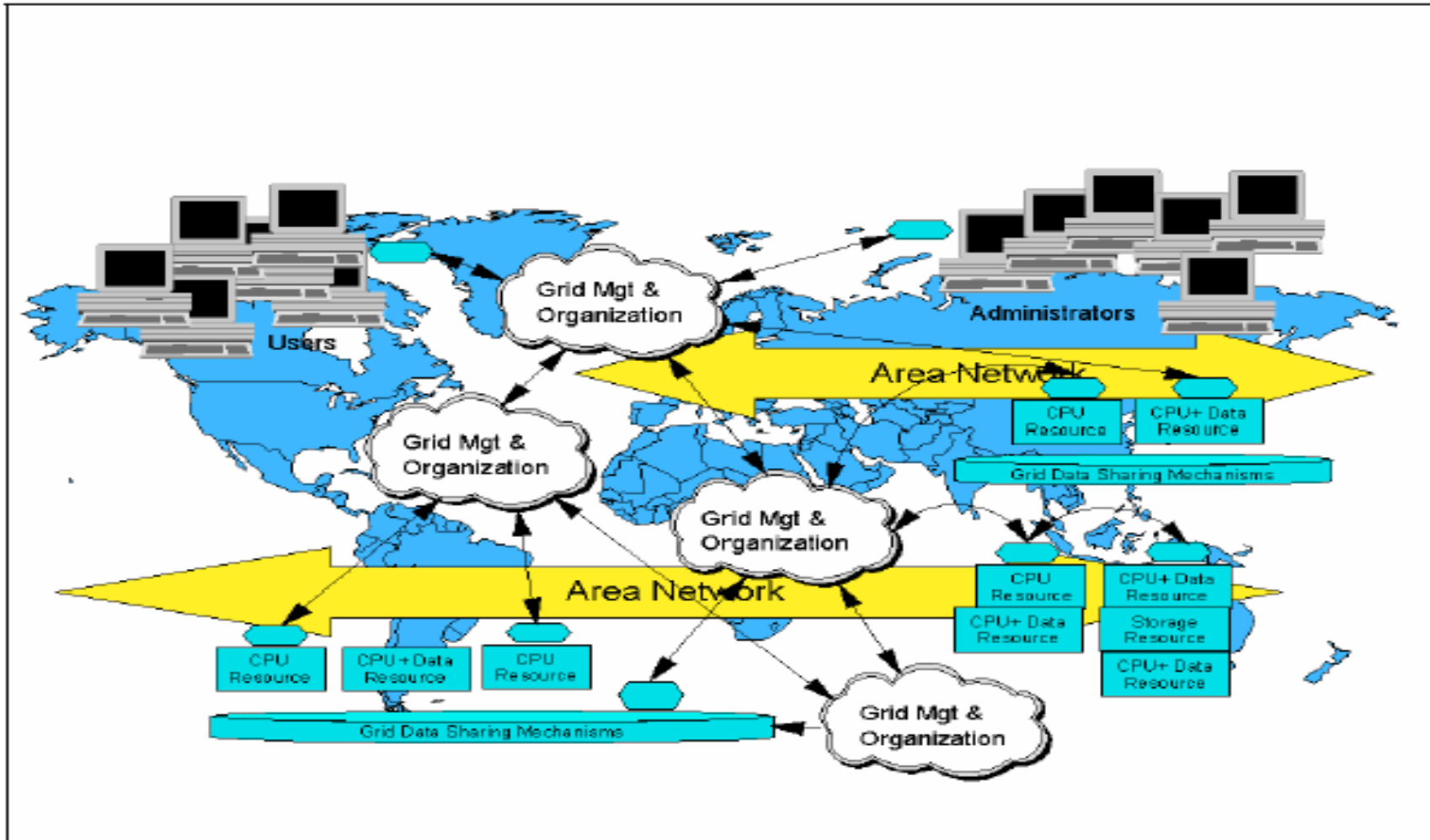
CSE Deptt. University at Buffalo

Fall - 2006

# Agenda

- Grid Resource Broker- Architecture and Algorithm
- Adaptable Resource Broker ( Drawback and Proposed solution)
- SDSC Storage Resource Broker (SRB)
- Future Enhancements

# A Sample Grid



# Grid Resource Broker

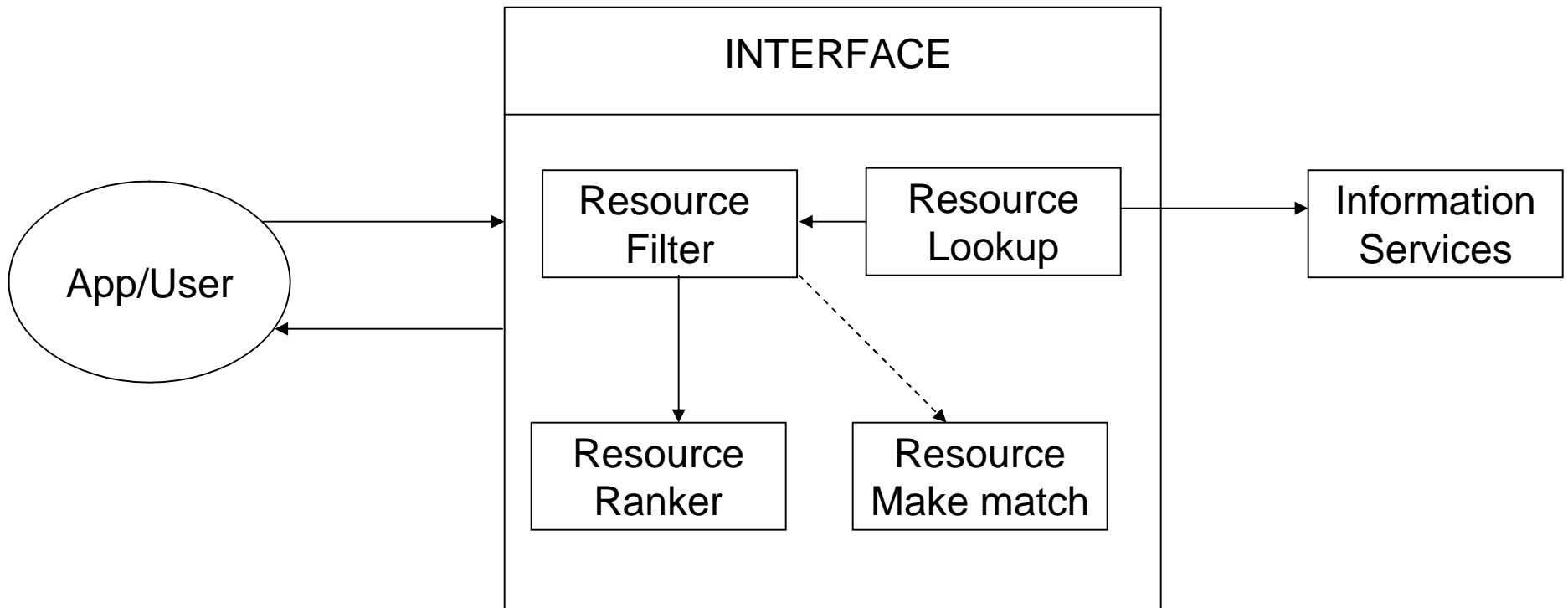
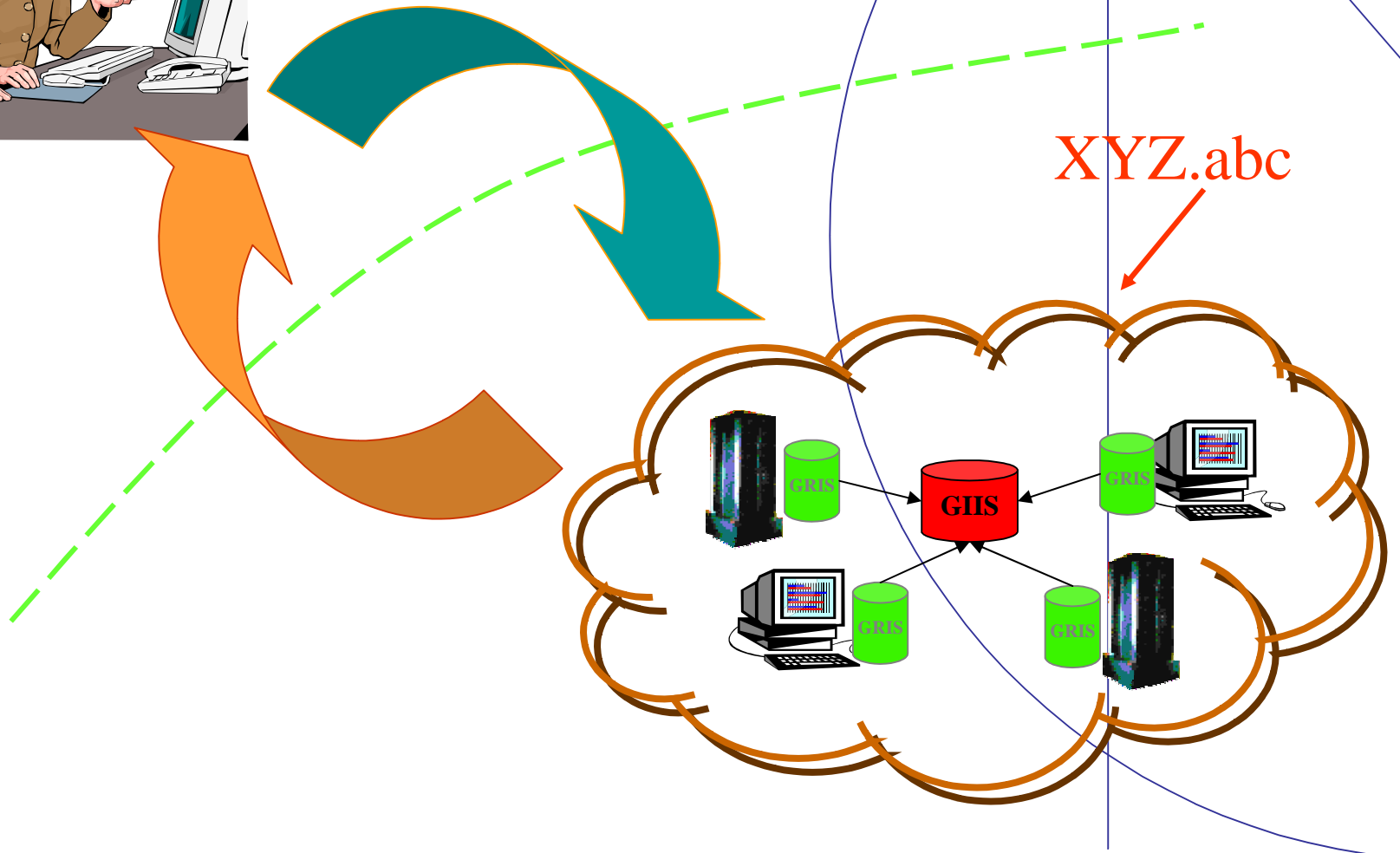


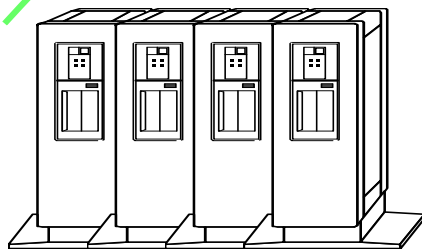
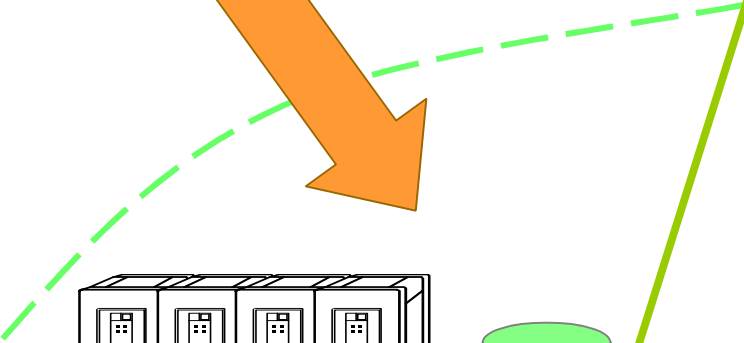
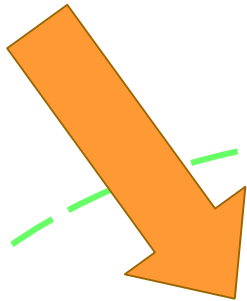
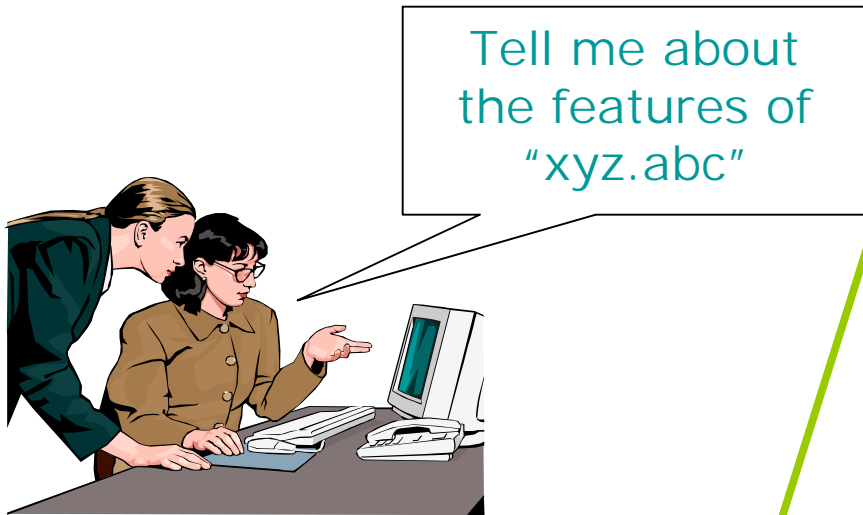
Fig 1. Resource Broker Architecture showing main components of the Design.

# Information Service : GIIS

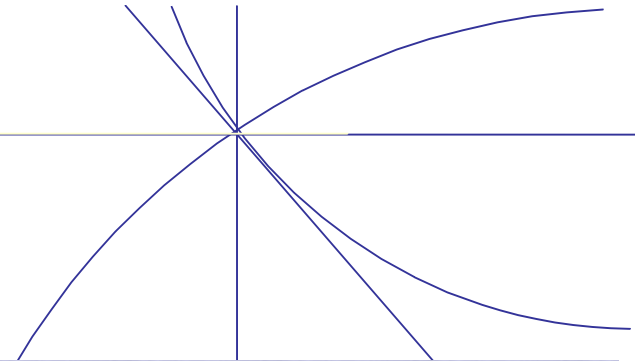
Tell me about the computing resources belonging to the HPC Lab that are uniprocessor Linux workstation, with low CPU load and available memory < 250 Mbyte



# Information Service: GRIS



**XYZ.abc**



## Information about a machine

dn: sw=Globus, hn=wyglaf.unile.it, dc=unile, dc=it, o=Grid

dn: hn=wyglaf.unile.it, dc=unile, dc=it, o=Grid

rn	Host wyglaf.unile.it
canonicalsysteemname	i686 unknown linux 2.2.14-15mdk
manufacturer	unknown
model	Pentium II (Deschutes)
type	workstation
osname	linux
osversion	i686 unknown linux 2.2.14-15mdk
osrelease	2.2.14-15mdk
ostype	linux
physicalmemorysize	251 MBytes
cputype	GenuineIntel Pentium II (Deschutes) 6 (step 1) (fpu vme de pse tsc msr pae mce cx8 sep mtrr pge mca cmov pat pse36 mmx fxsr)
cpucount	1
cpuspeed	350 MHz
cpuload1	1.03
cpuload5	0.28
cpuload15	0.10

dn: service=jobmanager, hn=wyglaf.unile.it, dc=unile, dc=it, o=Grid

cputype	i686
osversion	2.2.14-15mdk
osname	linux
manufacturer	unknown

# Grid Resource Broker

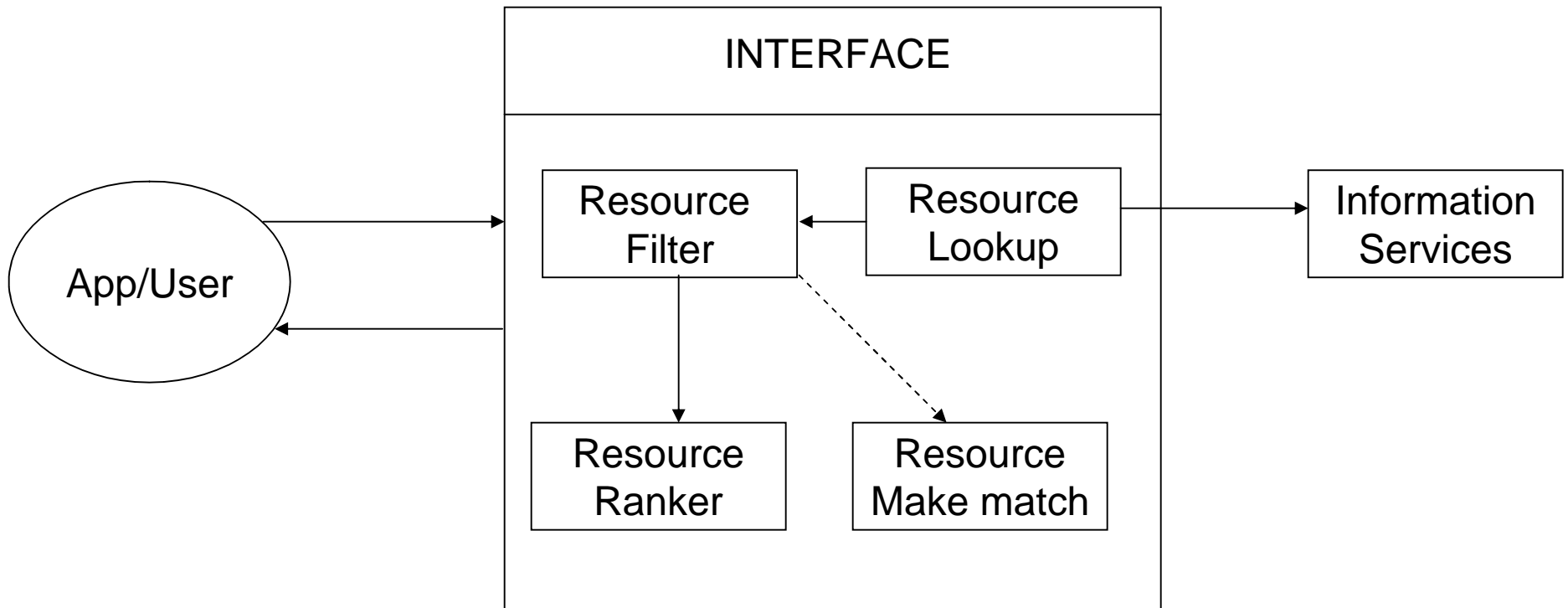


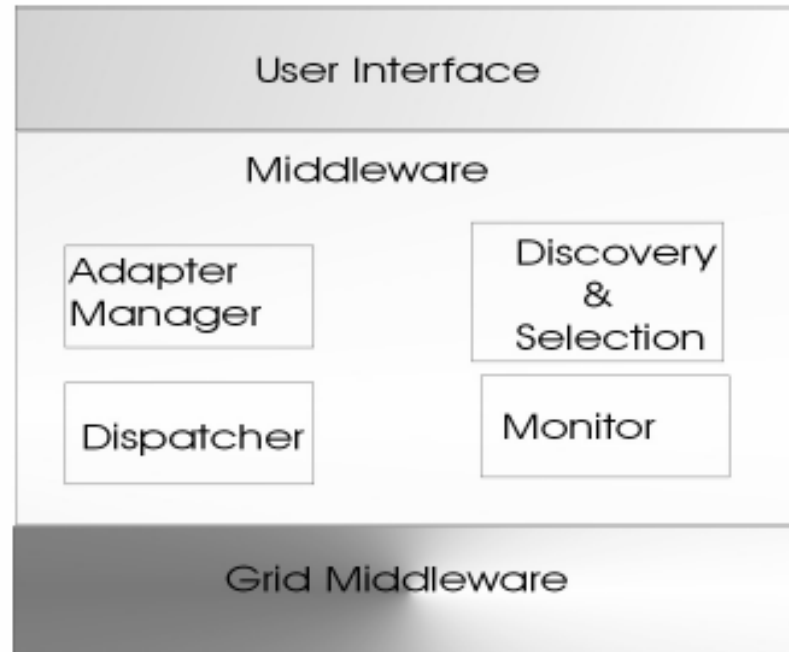
Fig 1. Resource Broker Architecture showing main components of the Design.

# Resource Brokering Algorithm

- Input :- one or more job request
  - Action :- Select and submit job to most appropriate resources.
  - Output: none
1. Contact GUIS server(s) to obtain a list of available clusters.
  2. Contact each resource's GRIS for static and dynamic resource information( hardware and software characteristics, current queue and load, etc.)
  3. For each job :-
    - (a) Select the cluster to which the job will be submitted.
      - i. Filter out clusters that do not fulfill the requirements on memory, disk space architecture etc, and clusters that the user are not authorize to use.
      - ii. Estimate the Total Time to Delivery (TTD) for each remaining resource.
      - iii. Select the cluster with the shortest predicted TTD.
    - (b) Submit the job to the selected resource.
    - (c) Release any reservation made to non selected clusters.



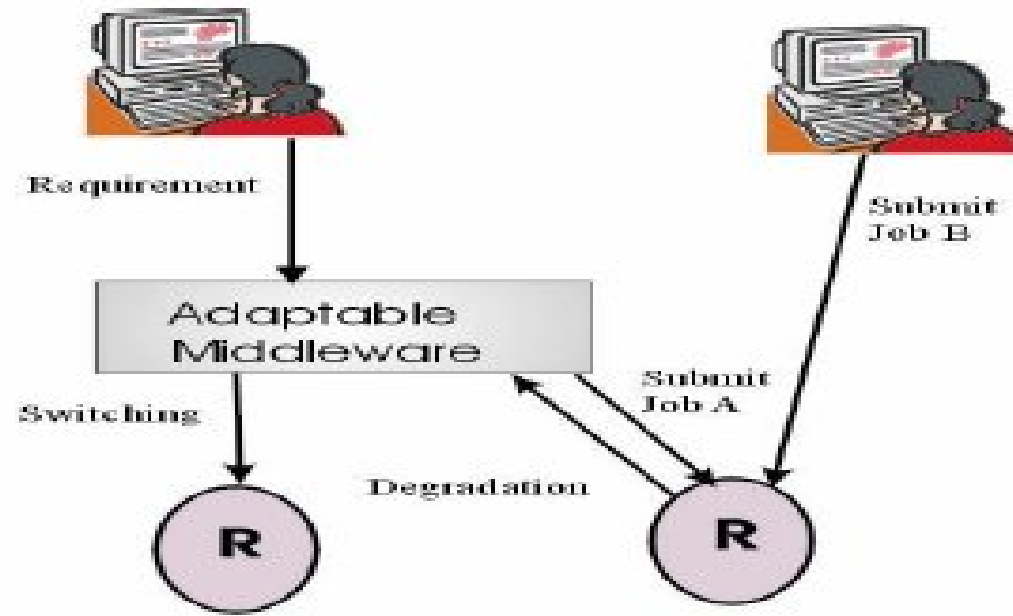
# Adaptive Grid Resource Broker



$$T_{remaining} = \frac{T_{100\%} - \sum_{i=1}^{n-1} T_i F_i}{F_{estimate}}$$

$$F_{estimate} = \bar{F}_n = \frac{1}{n} \sum_{i=1}^n F_i$$

# Adaptable Broker Scenario

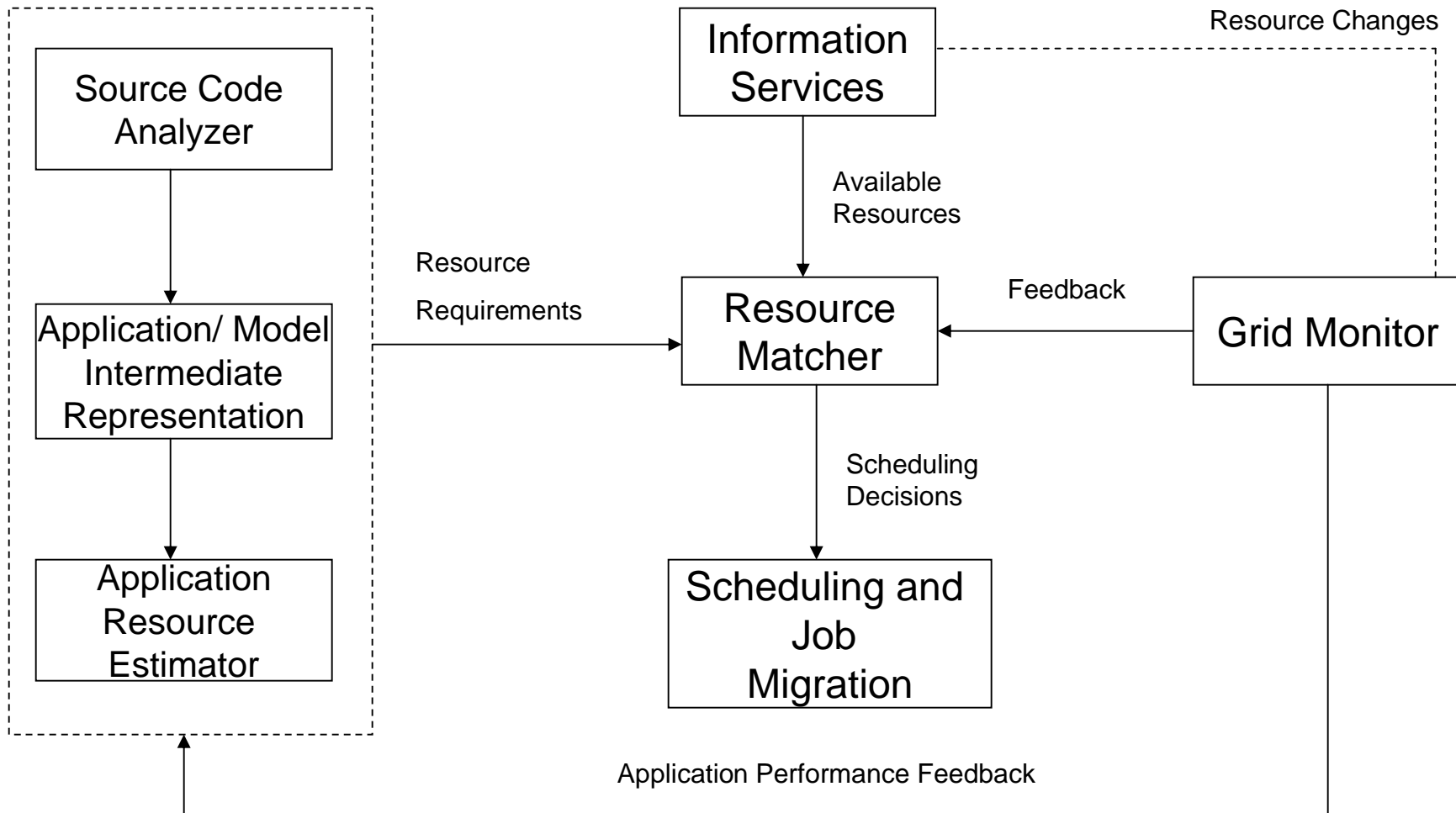


**Adaptable Broker Scenario**

# USE OF PROFILER

- Profiler can be used to by the monitor for predicting the parameters like CPU usage memory requirement.
- Sample Profiler (if the incoming process is a Program) can be built by the use of Halstead Complexity Algorithm along with MacCabe cyclomatic complexity algorithm.
- Halstead algorithm can be predict on the size of the program, the difficulty and effort required. This difficulty and effort can be mapped to the processor speed required and the size to the memory required.
- Cyclomatic complexity:- Given a program one can draw its control flow graph  $G$ , where each node is a block of code and each branch is decision of branching in that block of code.

# Application Performance Feedback



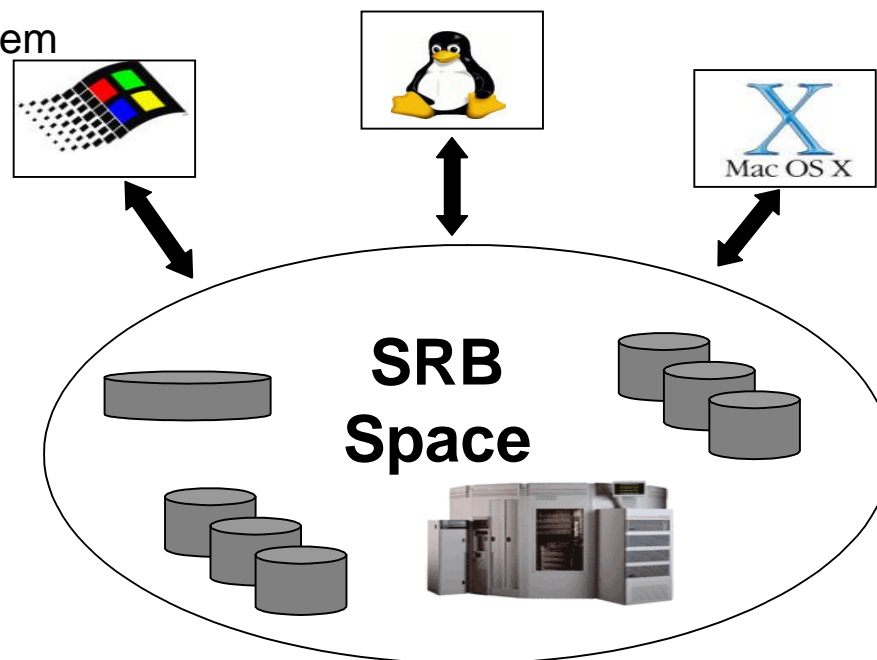
# SDSC Storage Resource Broker (SRB)

- **What is SRB?**

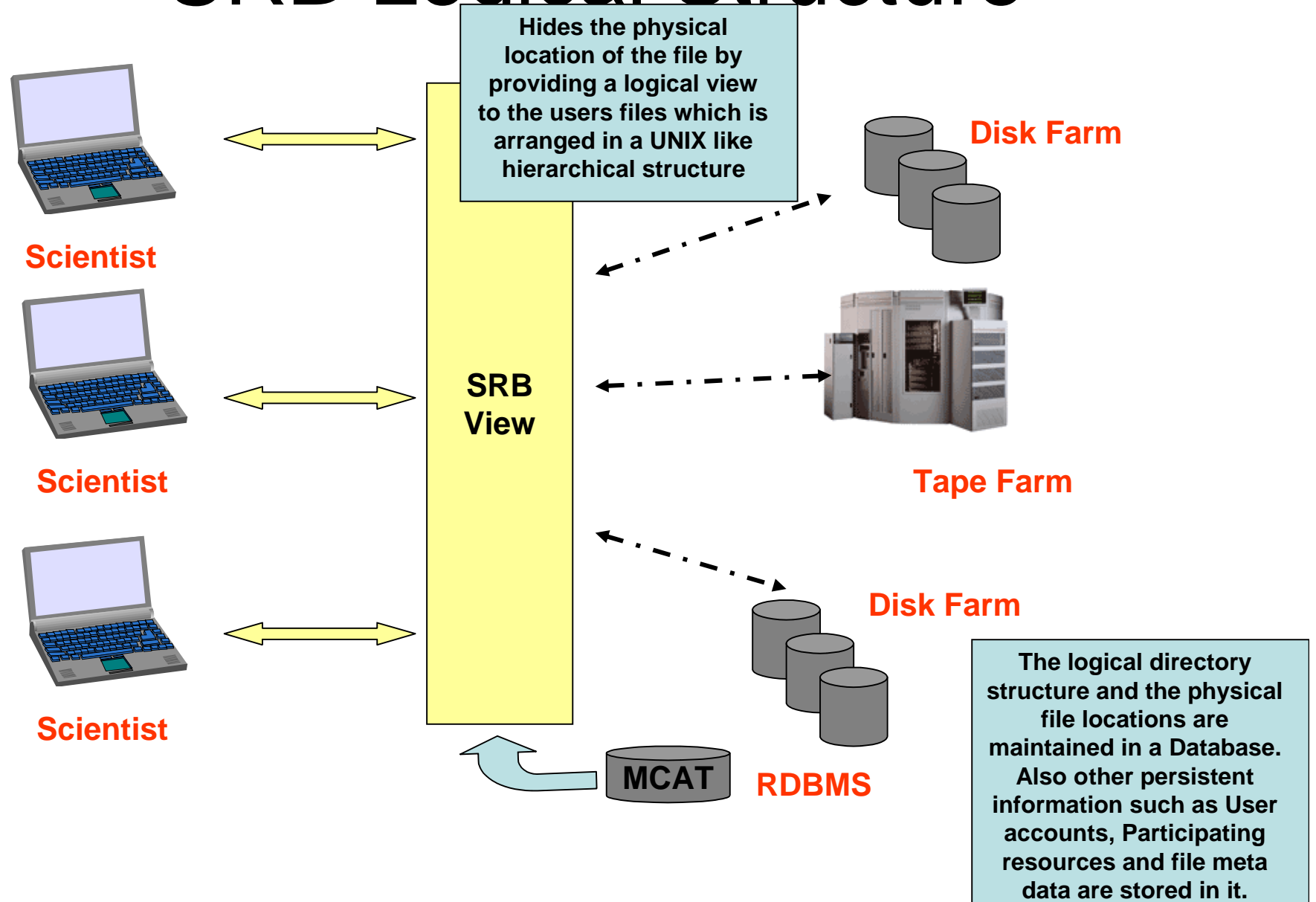
- It is a software that sits in between users and resources and provides a storage service by managing users, file locations, storage resources and metadata information
- The SDSC Storage Resource Broker (SRB) is client-server middleware that provides a uniform interface for connecting to heterogeneous data resources over a network and accessing unique or replicated data objects.
- SRB, in conjunction with the Metadata Catalog (MCAT), provides a way to access data sets and resources based on their logical names or attributes rather than their names and physical locations.

- **SRB Features:-**

- It is a distributed but unified file system
- It is a database access interface
- It is a digital library
- It is a semantic web
- It is a data grid system
- It is an advanced archival system



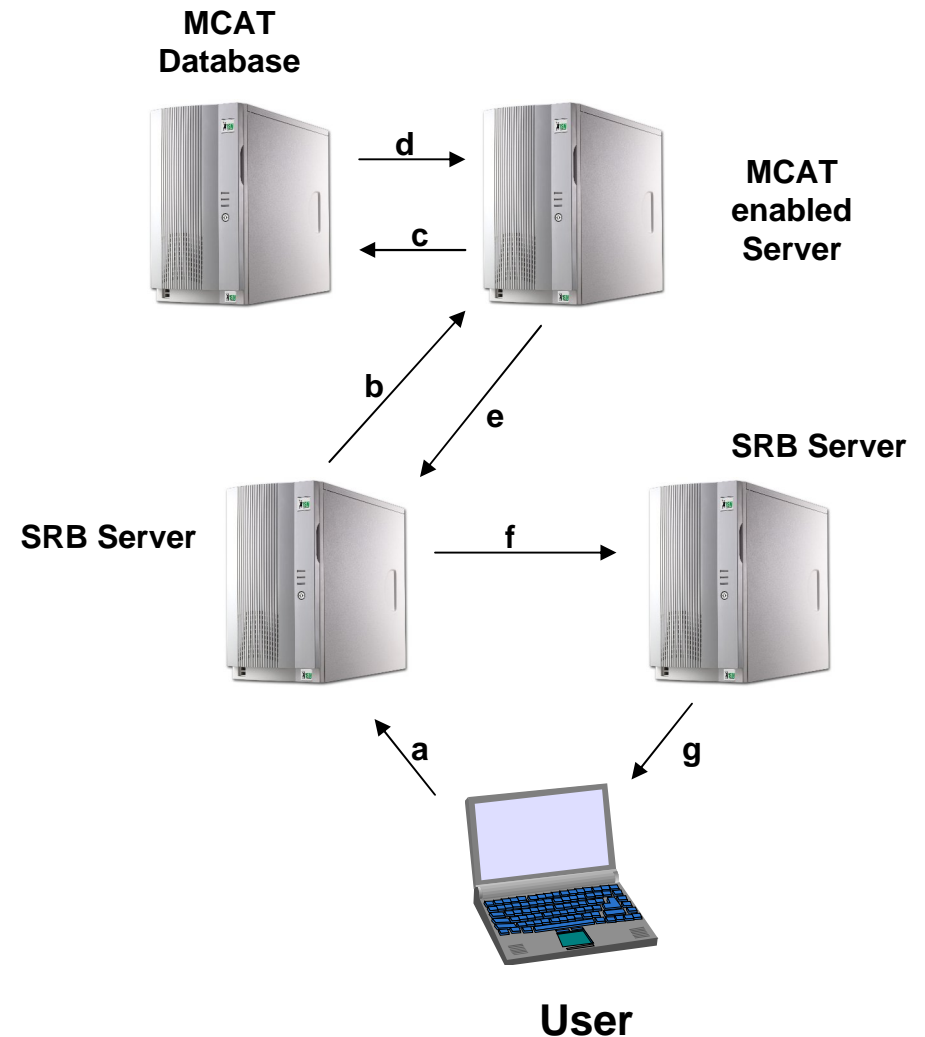
# SRB Logical Structure



# SRB Communication

## User File Request

- a) SRB Client sends request for file to SRB server.
- b) SRB Server contacts MCAT Enabled Server (MES).
- c) MES translates query into SQL and sends to database hosting MCAT
- d) Database query returned to MES
- e) Location of file etc returned to SRB Server A.
- f) SRB Server A contacts SRB Server B hosting data file.
- g) Data file transferred to user.



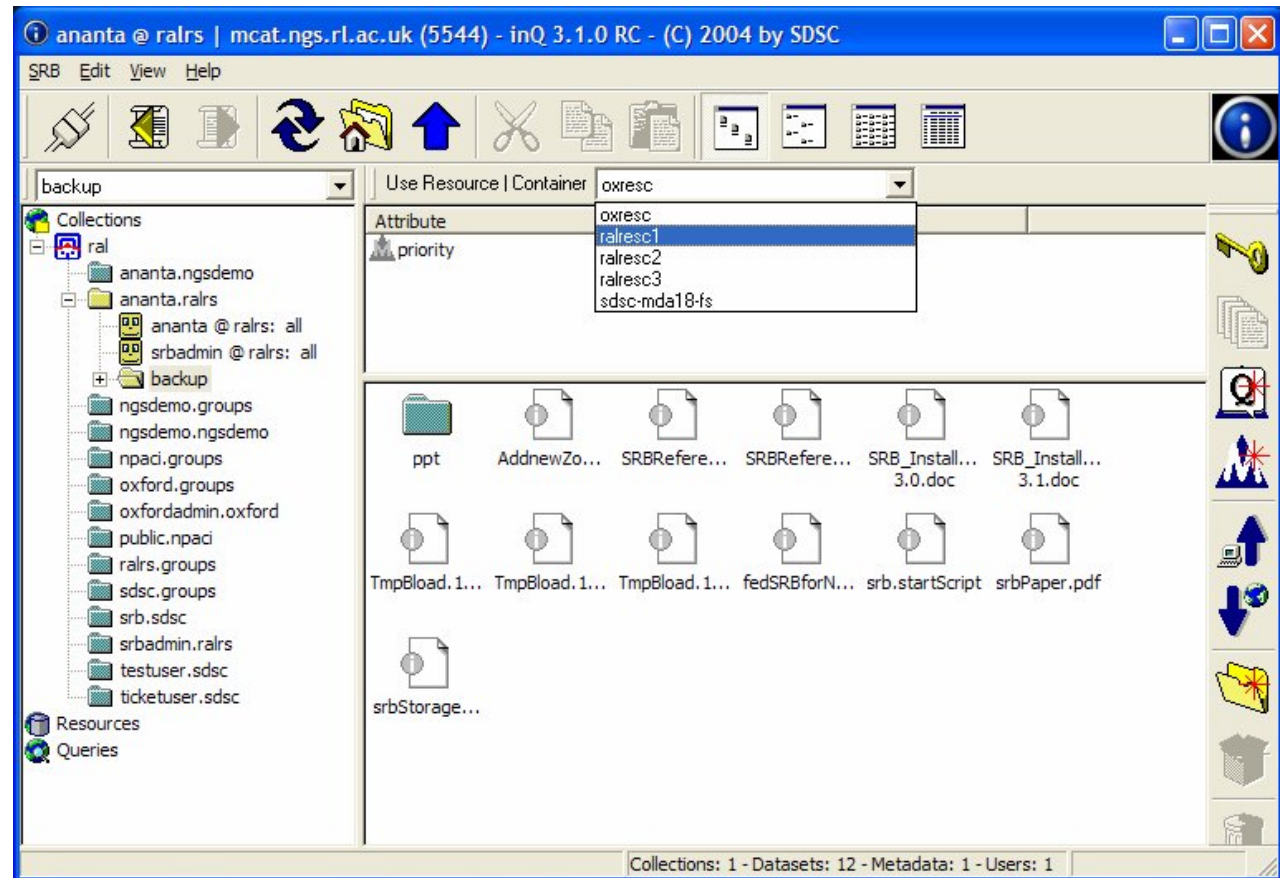
# SRB User Interfaces

- InQ – Windows Explorer Style interface to SRB

•Primary highlights includes

•Support for drag and drop between Windows mounted filesystems

•Provisions for access control enforcements, file replication, metadata entry and metadata query.





# SRB User Interfaces

- My SRB Interface – Web browser interface to SRB

• Web based interface to SRB space.

• Works through port 80 and hence works around firewall issues.

The screenshot shows a web browser window titled 'MYSRB - A TransSystem Data Explorer - Mozilla Firebird'. The address bar shows the URL 'http://forth.dl.ac.uk/cgi-bin/mysrb2.cgi/ananta'. The page content includes a 'View All Metadata' section with the following details:

- Collection: ananta.eminerals
- Parent Collection: /home
- Owner: ananta@eminerals

Below the details is a toolbar with icons for Move, Ingest, Create, Register, Register, Register, Register, Register, Register, Make, Make, Browse, Other, Online, and Exit. A secondary row of links includes Up, File, File, File, Directory, URL, SQL, ORBData, Command, Collection, Container, Query, Info, Help, and MySRB.

The main content area displays a table of metadata for the path '/home/ananta.eminerals'. The table has columns for Function, Data Name, Creation Time, Owner, Replica Version Number, Size, and D: Ty.

Function	Data Name	Creation Time	Owner	Replica Version Number	Number	Size	D: Ty
Get File	<a href="#">Acm description.doc</a>	2004-03-16-16.20.26	ananta@eminerals	0	0	37376	gen
Get File	<a href="#">Acm description.doc</a>	2004-04-28-20.32.34	ananta@eminerals	1	0	37376	gen
Get File	<a href="#">Acm description.doc</a>	2004-04-28-20.48.08	ananta@eminerals	2	0	37376	gen
Get File	<a href="#">AnneDissertation1.pdf</a>	2004-05-06-09.59.01	ananta@eminerals	0	0	418963	gen
Get File	<a href="#">SRBStorageServer1.3.pdf</a>	2003-10-13-15.21.52	ananta@eminerals	0	0	167796	gen
Get File	<a href="#">ccircTemplate</a>	2004-04-28-21.25.00	ananta@eminerals	0	0	20992	gen
Get File	<a href="#">plotTest</a>	2004-04-28-20.59.06	ananta@eminerals	0	0	68677	gen
Get File	<a href="#">readme</a>	2004-04-29-12.21.25	ananta@eminerals	0	0	730	gen
Get File	<a href="#">test</a>	2004-03-31-16.30.00	ananta@eminerals	0	0	17148	gen

Below the first table is a second table with columns for Function, SubCollection, Creation Time, and Owner.

Function	SubCollection	Creation Time	Owner
Open Collection	<a href="#">newColl</a>	2004-04-28-21.07.45	ananta@eminerals
Open Collection	<a href="#">pictures</a>	2004-04-27-16.11.10	ananta@eminerals

# Resource Brokering Algorithm

- Input :- one or more job request
  - Action :- Select and submit job to most appropriate resources.
  - Output: none
1. Contact GUIS server(s) to obtain a list of available clusters.
  2. Contact each resource's GRIS for static and dynamic resource information( hardware and software characteristics, current queue and load, etc.)
  3. For each job :-
    - (a) Select the cluster to which the job will be submitted.
      - i. Filter out clusters that do not fulfill the requirements on memory, disk space architecture etc, and clusters that the user are not authorize to use.
      - ii. Estimate the Total Time to Delivery (TTD) for each remaining resource.
      - iii. Select the cluster with the shortest predicted TTD.
    - (b) Submit the job to the selected resource.
    - (c) Release any reservation made to non selected clusters.

# Enhancements

## Steps to optimize the algorithm

- A Resource broker to find the memory and complexity of a incoming process which can be a program.
- Estimating the Total Time to Delivery
- Job queue adaptation. Keep searching for better resources.
- Advance resource reservations.
  
- Presently there are particular resource brokers like:
  - SRB – for storage
  - Application Level Scheduler (AppLeS) – used for scheduling and deploying where tasks have no or little inter-task communication.
  - Condor, Nimrod-G etc....
  
- Take a case where a process needs to execute some database query and need to send the result to some other user. Then the result from the user needs rigorous mathematical calculation which requires high CPU speed and lots of memory.
- ???????

Questions?

# References

- Overview of the SDSC Storage Resource Broker By Wayne Schroeder May,2004
- Storage Resource Broker By Arcot Rajasekar, Director, Data Grid Technologies Group SDSC/UCSD.
- A Grid Resource Broker Supporting Advance Reservations and Benchmark-Based Resource Selection :- Erik Elmroth and Johan Tordsson
- A High-Performance Computational Resource Broker for Grid Computing Environments :- Chao-Tung Yang Po-Chi Shih Kuan-Ching Li