

Running BaBar jobs on the grid using gsub and AliBaBa

Mike AS Jones

- *BaBar job life-cycle*
- *gsub* – *to submit to the grid*
- *alibaba* – *to monitor the submissions and help the user*
- *morgiana* – *to look pretty*
- *bfgrits* – *to test the grid nodes*
- *afs suitability*
- *open issues and future directions*

Date 7 April 2004
Event HEP IoP
Venue Birmingham



THE UNIVERSITY
of MANCHESTER

submitting BaBar jobs to local farms

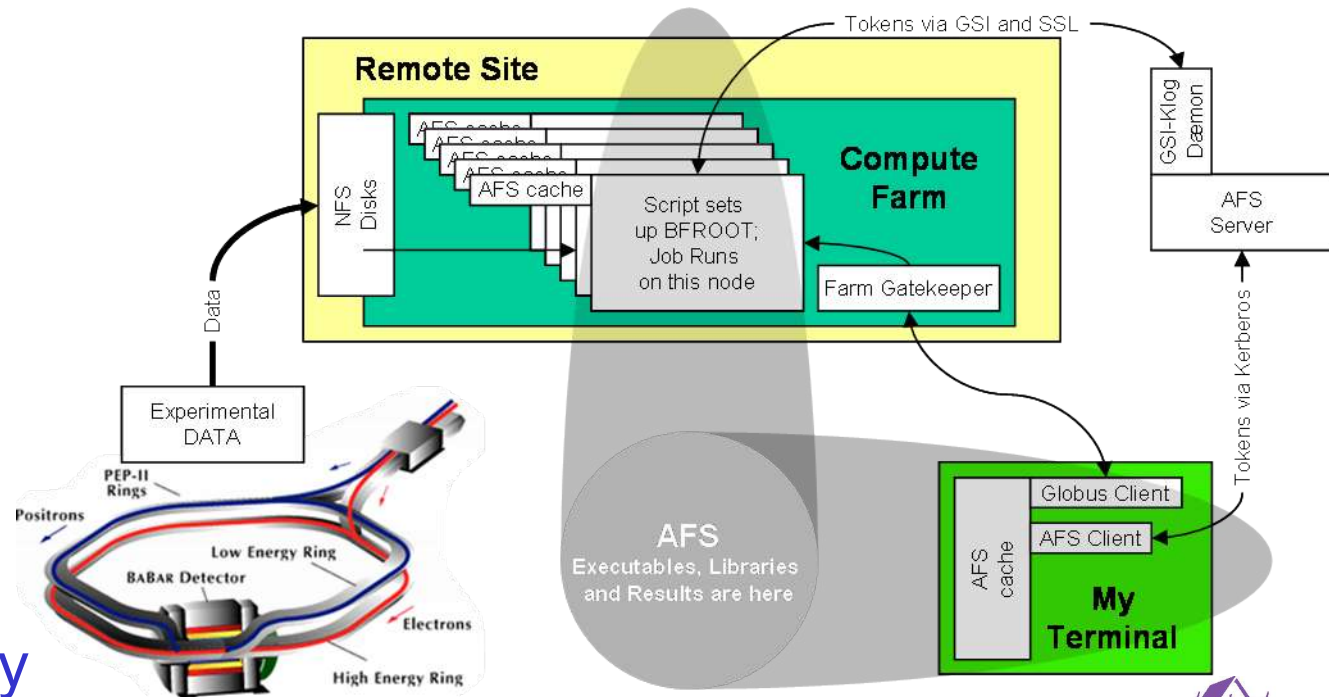
- start in directory which is mounted on the farm
- check out code
 - CVS repository somewhere
- write more code
- set up environment, compile and link code
- find data and create index
 - skimData --blah --otherblah
- set up environment and qsub executable
 - Job runs locally, finds local data and saves files locally
- results returned to files on local file system
- grid?
 - globus/dg-globus, SRB, Dump, Software – hard to use – hard to install
 - gsub, SkimData portal – follows scheme familiar to user

Submitting jobs to BaBar Farms with gsub

- compute farms are distributed throughout GB
- large datasets which are located only at specific farms
- executables with client
- results wanted by client
- maybe write a complex resource broker and use complicated middleware to transfer data

~or~

- distributed file system moves user data and executables transparently
- data reduces RB task



Submitting jobs to BaBar Farms

- what gsub does
 - 1) checks lots of things
 - 2) gets the current list of gatekeepers etc
 - 3) creates a script (to wrap the executable on farm PC)
 - 1) sets up a normal environment
 - 2) notifies alibaba
 - 3) gets (pag separated) AFS credentials using gsi klog
 - 4) creates BFROOT – BaBar environment
 - 5) changes to directory submitted from
 - 6) starts a shepherd process
 - 1) this will look after job's grid stuff and talk to alibaba
 - 7) runs user's executable (script or binary)
 - 8) unlogs
 - 4) uses globus to stage and submit the script to a queue on a local/remote machine
 - 5) uses curl over ssl to tell a website the status of the job (alibaba)

gsub usage

gsub [Options] command args...

AFS related:

[{-a|-afs} <user@cell>]
[+a|+afs} <extra user@cell>]+
[{-c|-cell} <cell>]
[{-p|-principal} <principal>]

If not specified by one method above, gsub will try to guess principal and realm.

Globus related:

[{-g|-gate} <gatekeeper>]
[{-j|-jobman} <jobmanager>]
[{-x|-proxy} <non-standard proxy location>]

local machines related:

[{-bf|-bfroot} <local BFROOT>]
[{-d|-display} <DISPLAY>]

remote machine related:

[{-S|-site} <BABAR-SITE>]
[{-s|-source} <RemoteSourceFile1> [{-s|-source} <File2>] ...]
[{-rb|-rbfroot} <Path to Remote BFROOT on Remote Machine>]
[-nb]
[-t|-tmp]
[{-CA|-capath} <path to CA's>]
[{-queue|-q} <queuename>]

user interaction related:

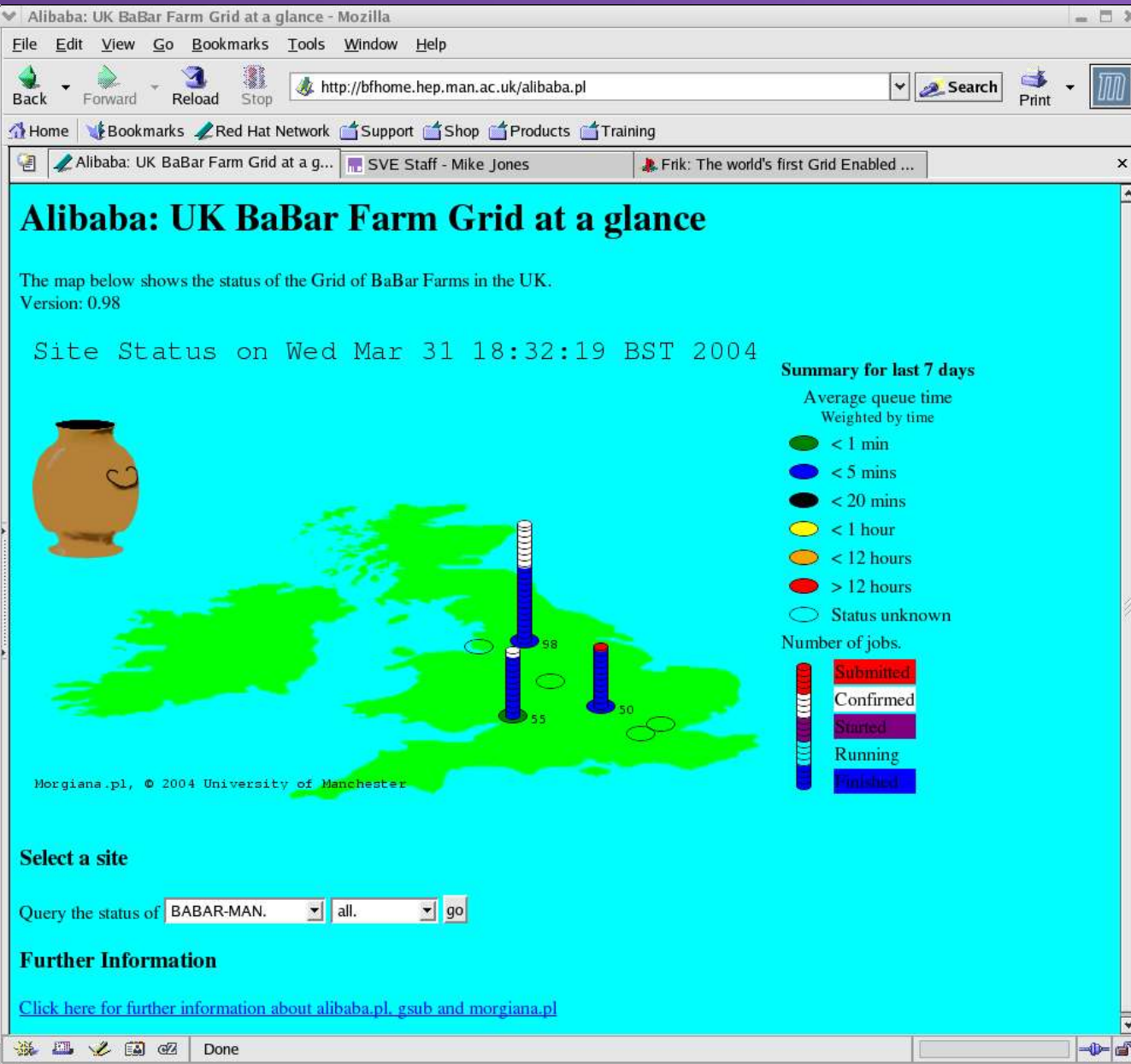
[-i|-int [-e|-err <errorfile>] [-o|-out <outfile>]]
[-l] [-v|-verbose] [-vv|-vverbose] [-D|-dump] [-T|-dry] [-C|-cat]
[-h|-?|-help] [-u|-usage] [-V|-version]

etc.

- is a CGI perl script
- is hosted by a Gridsite 1.0+
- takes several variables in get method
 - Default returns a web page with status map
 - Links to specific sites' statuses
 - Methods for running jobs to upload their statuses securely
 - Methods for using the server to retrieve globus status and output
- records job statuses
- draws pretty pictures

AliBaBa front page

- site queue status
 - jobs submitted
 - jobs running
 - jobs finished
- image not cached
- links to more details



<http://bfhome.hep.man.ac.uk/alibaba.pl?action=query>

Fine Detail

- `action=query`
 - status for each site can be viewed in http and https
 - unauthenticated
 - authenticated
 - extra information
 - job status can be sorted into successful jobs, failed jobs and stale jobs
 - action (status and retrieve)

Alibaba Jobs@BABAR-MAN - Mozilla


File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop <http://bfhome.hep.man.ac.uk/alibaba.pl?action=query&site=BABAR-MAN> Search Print

Home Bookmarks Red Hat Network Support Shop Products Training


Alibaba Jobs@BABAR-MAN SVE Staff - Mike Jones Friik: The world's first Grid Enabled ...

Alibaba Jobs@BABAR-MAN



Contact URL	Stages					Exit Status	Action
	Submitted	Confirmed	Started	Running	Finished		
https://bfb.hep.man.ac.uk:50103/14701/1080753881/	Wed Mar 31 17:24:29 2004	+0:00:15	+0:00:53	+0:00:55	+0:01:27	0	
https://bfb.hep.man.ac.uk:50100/14637/1080753863/	Wed Mar 31 17:24:14 2004	+0:00:14	+0:00:31	+0:00:35	+0:01:05	0	
https://bfb.hep.man.ac.uk:50095/14453/1080753840/	Wed Mar 31 17:23:51 2004	+0:00:22	+0:00:52	+0:00:55	+0:01:25	0	
https://bfb.hep.man.ac.uk:50074/14325/1080753826/	Wed Mar 31 17:23:37 2004	+0:00:12	+0:01:05	+0:01:09	+0:01:39	0	

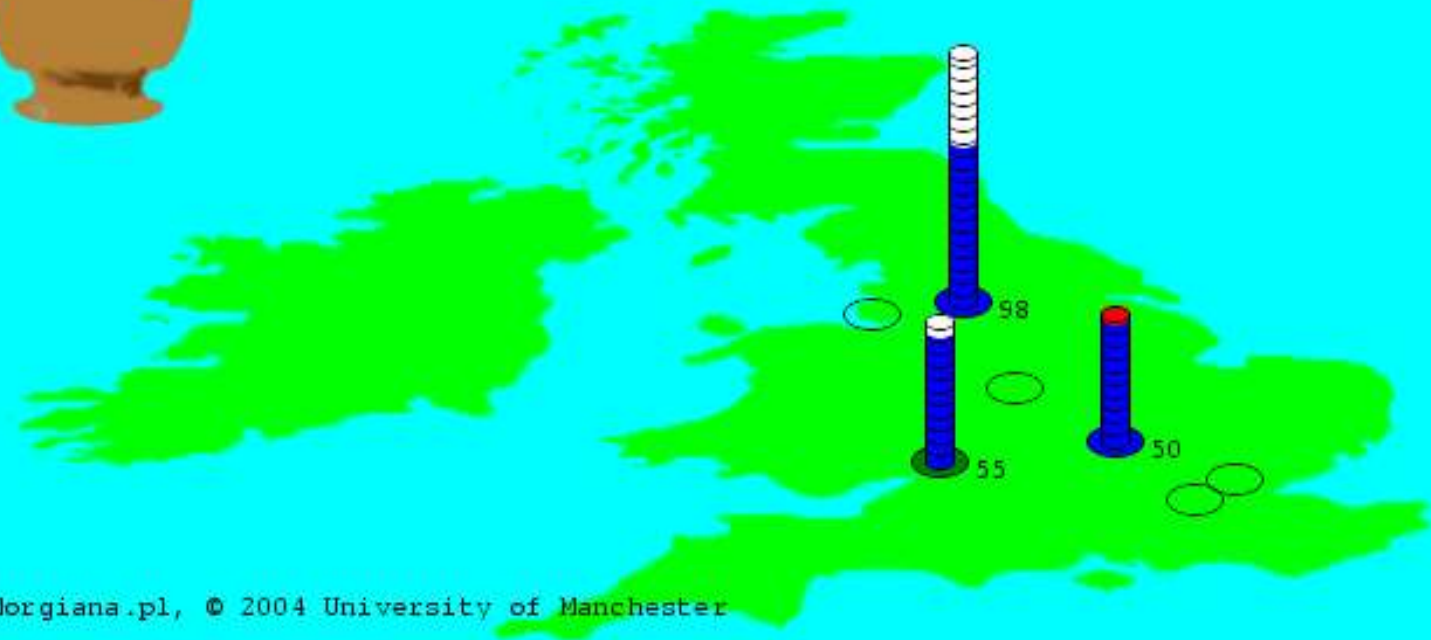
Alibaba Jobs@BABAR-MAN



Contact URL	Stages					Exit Status	Action
	Submitted	Confirmed	Started	Running	Finished		
https://bfb.hep.man.ac.uk:50103/14701/1080753881/	Wed Mar 31 17:24:29 2004	+0:00:15	+0:00:53	+0:00:55	+0:01:27	0	
https://bfb.hep.man.ac.uk:50103/14701/1080753881/	Submitted by masj@130.88.1.130 Submitted to x-gram://bfb.hep.man.ac.uk:2119/jobmanager-pbs Started as masj@194.36.3.78 The proxy certificate 673 minutes						
https://bfb.hep.man.ac.uk:50100/14637/1080753863/	Wed Mar 31 17:24:14 2004	+0:00:14	+0:00:31	+0:00:35	+0:01:05	0	
https://bfb.hep.man.ac.uk:50100/14637/1080753863/	Submitted by masj@130.88.1.130 Submitted to x-gram://bfb.hep.man.ac.uk:2119/jobmanager-pbs Started as masj@194.36.3.18 The proxy certificate 673 minutes						
https://bfb.hep.man.ac.uk:50095/14453/1080753840/	Wed Mar 31 17:23:51 2004	+0:00:22	+0:00:52	+0:00:55	+0:01:25	0	
https://bfb.hep.man.ac.uk:50095/14453/1080753840/	Submitted by masj@130.88.1.130 Submitted to x-gram://bfb.hep.man.ac.uk:2119/jobmanager-pbs Started as masj@194.36.3.16 The proxy certificate 673 minutes						
https://bfb.hep.man.ac.uk:50074/14325/1080753826/	Wed Mar 31 17:23:37 2004	+0:00:12	+0:01:05	+0:01:09	+0:01:39	0	
https://bfb.hep.man.ac.uk:50074/14325/1080753826/	Submitted by masj@130.88.1.130 Submitted to x-gram://bfb.hep.man.ac.uk:2119/jobmanager-pbs Started as masj@194.36.3.78 The proxy certificate 673 minutes						
https://bfb.hep.man.ac.uk:50073/13908/1080753763/	Wed Mar 31 17:22:31 2004	+0:00:31	+0:01:11	+0:01:13	+0:01:43	0	
https://bfb.hep.man.ac.uk:50073/13908/1080753763/	Submitted by masj@130.88.1.130 Submitted to x-gram://bfb.hep.man.ac.uk:2119/jobmanager-pbs Started as masj@194.36.3.78 The proxy certificate 673 minutes						
https://bfb.hep.man.ac.uk:50100/14117/1080753794/	Wed Mar 31 17:23:04 2004	+0:00:32					
https://bfb.hep.man.ac.uk:50100/14117/1080753794/	Submitted by masj@130.88.1.130 Submitted to x-gram://bfb.hep.man.ac.uk:2119/jobmanager-pbs The proxy certificate 673 minutes						

- **action=submitted | confirmed | started | running | update | finished**
 - must be authenticated https (a GSI proxy will do)
 - designed for gsub to use not for user!
 - allows uploading of job's progress
 - stored in individual job xmls file on web server
 - status data only accessible to owner of the GSI credential

Site Status on Wed Mar 31 18:32:19 BST 2004



Morgiana.pl, © 2004 University of Manchester

Status Map

- image updated on server every time state changes
- site blob colour
 - time jobs spend in queue
 - weighted by age of result
- extremely easy to add a new sites
 - add directory on server
 - create xml file with xy position of site!

Interoperability Tests

- based on the UK eScience GITS
 - which are based on teragrid's original tests
- bash (or ksh) cf perl – for job control reasons
- GITS centric
- contains extra test for gsub
- writes results in text to stdout, in html and xml to files
 - xml files are compatible with UK eScience GITS database
- Is wrapped in a script: bftests
 - uses gatekeepers.xml rather than GITS
 - writes xml and html to Bftests.(xml|html) on bfhome if run by authorised user

http://bfhome.hep.man.ac.uk/BFtests.html

Home Bookmarks Red Hat Network Support Shop Products Training

Grid Resource Interoperability Results

Grid Resource Interoperability

Run on 2004-01-02 at 16:43:12 by /C=UK/O=eScience/OU=Manchester/L=MC/CN=michael.jones/CN=proxy from pc53.hep.man.ac.uk

Test Hosts obtained from http://bfhome.hep.man.ac.uk/gatekeepers.xml

PING	GPING	RSL	GJR	GJRS	GJSUB	GJST	GJCL	GJRET	GASS	GFTP	GSSH	GSUB
Timeouts (in seconds) set to:												
5	10	60	60	60	60	60	120	90	90	40	30	600
gpnce04.gridpp.rl.ac.uk:2119/jobmanager-pbs:/C=UK/O=eScience/OU=CLRC/L=RAL/CN=gpnce04.gridpp.rl.ac.uk/Email=s.traylen@rl.ac.uk												
PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	FAIL	FAIL	FAIL	PASS
bottom.phy.bris.ac.uk:2119/jobmanager-pbs:												
PASS	PASS	PASS	PASS	PASS	PASS	PASS	FAIL	PASS	FAIL	PASS	FAIL	PASS
bfh.hep.man.ac.uk:2119/jobmanager-pbs:												
PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	FAIL	PASS

Keys

- PASS** Passed!
- FAIL** Failed
- N.T.** Not Tested
- unAuthz** Remote user not Authorised
- Unknown** Remote Authentication failed
- Bad Gate** Gatekeeper's certificate from http://bfhome.hep.man.ac.uk/gatekeepers.xml is different to certificate supplied by gatekeeper
- No 3rd party** GSI Authentication to ftp server ok but failed port or passive command.
- T.O.** Time Out

PING # ping <host>

GPING # globusrun -a -r <gatekeeper>

RSL # globusrun -o -r <gatekeeper> '&(executable="/bin/echo")(arguments="Hello World")'

GJR # globus-job-run <gatekeeper> /bin/echo Hello World

GJRS # globus-job-run <gatekeeper> -s \$TEMPFILE

GJSUB # globus-job-submit <gatekeeper> /bin/sleep 600

GJST # globus-job-status <contact>

GJCL # globus-job-clean -f <contact>

GJRET # globus-job-submit <gatekeeper> /bin/echo HelloWorld
globus-job-status <contact>
globus-job-get-output <contact>

GASS # globusrun -s -r <gatekeeper> "&(executable=\\$(GLOBUS_LOCATION)/bin/globus-url-copy)(arguments=\\$(GLOBUSRUN_GASS_URL)\\$(TEMPFILE) \\
\\file:\\$(TEMPFILE).g1\")(environment=(LD_LIBRARY_PATH \\$(GLOBUS_LOCATION)/lib))"
globusrun -s -r <gatekeeper> "&(executable=\\$(GLOBUS_LOCATION)/bin/globus-url-copy)(arguments=\\"file:\\$(TEMPFILE).g1\" \\
\\\$(GLOBUSRUN_GASS_URL)\\$(TEMPFILE).g2)(environment=(LD_LIBRARY_PATH \\$(GLOBUS_LOCATION)/lib))"
diff -q \\$(TEMPFILE) \\$(TEMPFILE).g2

GFTP # globus-url-copy file:\\$(TEMPFILE) gsiftp://<host>\\$(TEMPFILE).1
globus-url-copy gsiftp://<host>\\$(TEMPFILE).1 file:\\$(TEMPFILE).2
diff -q \\$(TEMPFILE) \\$(TEMPFILE).2
globus-url-copy gsiftp://<host>\\$(TEMPFILE).1 gsiftp://<host>\\$(TEMPFILE).2

GSSH # (gsi)ssh -o "BatchMode=yes" -p \$PORT <host> /bin/echo HelloWorld

GSUB # gsub -site <BABAR-SITE> "/bin/echo Hello World > afstempfile" ; echo afstempfile | grep "Hello World;"

BFgits web page

afs read write and append tests

- is AFS slow?
 - not really
 - BaBar jobs seem to run (if they get through the queue)
- what does AFS do?
 - transfers files
 - list file, read file, write file, create file, delete file, lock file, dir admin
 - time consuming components
 - actual transfer
 - obtaining locks
 - cache
- script to test AFS speed
 - tests - use gsub, script measures times:

• read	~ 250-500KB/s small files	~ 2-10MB/s large files
• write	~ 50-100 KB/s small files	~ 1-3MB/s large files
• append	~ 1-3 KB/s small files	~ 1-3MB/s large files

Open Issues and Future Directions

- **gsiklog/gssklog**
 - move to gssklog
 - expand gssklogd take-up
- **more automated data discovery**
 - skimData grid service (OGSI-LITE) or web service
 - LDAP or new BaBar computing thing
- **resource discovery**
 - in-house, LDAP, GIIS/MDS, RGMA, BDII
- **grid credential movement**
 - user push: globusrun -refreshproxy / Job pull: MyProxy
- **SRB and data movement**
 - AFS stuff fine for small (<1GB) transactions
 - what if I want to run at any grid enabled farm
 - Data must be present or moved
 - GridFTP, Bit Torrent, MBNG, ...