

Practical GRIDportal

Working with mpiBLAST

Martin Matusiak

The NTNU High Performance Computing Project
Norwegian University of Science and Technology

10 November 2005

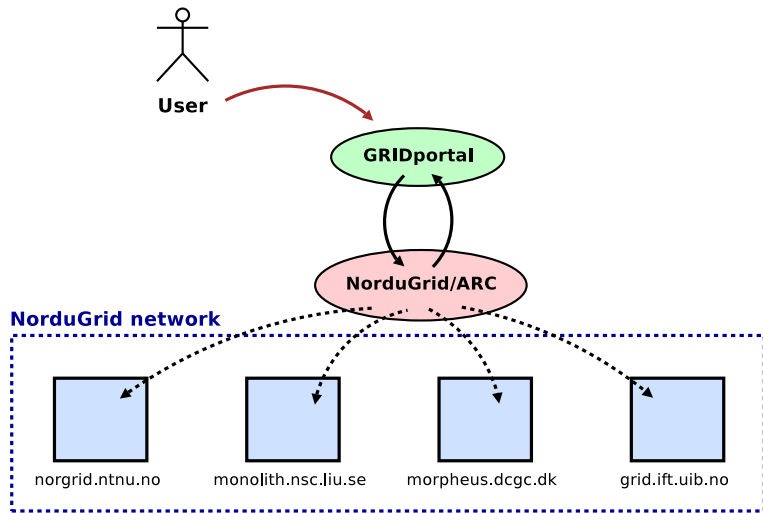
1 Introducing GRIDportal

2 Working with mpiBLAST

1 Introducing GRIDportal

2 Working with mpiBLAST

GRID: a network of computer resources



A quick look at NorduGrid/ARC

```
alex@frasier:~  
marmat@norgrid blast $ grid-proxy-init  
Your identity: /O=Grid/O=NorduGrid/OU=ntnu.no/CN=Martin Matusiak  
Enter GRID pass phrase for this identity:  
Creating proxy ..... Done  
Your proxy is valid until: Sat Aug 13 07:08:39 2005  
marmat@norgrid blast $ ngsync -f job.xrsl -c norgrid.ntnu.no -d0  
run.sh      0 s:      0.2 kB      0.0 kB/s      0.0 kB/s      . . .  
martin_testset.txt  0 s:      0.9 kB      0.0 kB/s      0.0 kB/s      . . .  
  
Job submitted with jobid gsiftp://norgrid.ntnu.no:2811/jobs/2100011238665261699770589  
marmat@norgrid blast $ ngsync -fa -c norgrid.ntnu.no  
marmat@norgrid blast $ ngstat -a  
Job gsiftp://norgrid.ntnu.no:2811/jobs/2100011238665261699770589  
  Jobname: blastpJob4  
  Status: FINISHED  
marmat@norgrid blast $ ngget -a  
ngget: downloading files to /home/marmat/blast/2100011238665261699770589  
ngget: download successful - deleting job from gatekeeper.  
marmat@norgrid blast $ ls 2100011238665261699770589/  
blast_out.QUERY  gmlog/  martin_testset.txt  run.sh  stderr.txt  stdout.txt  
marmat@norgrid blast $
```

GRIDportal: a web-based grid application portal

The screenshot shows a Mozilla Firefox browser window titled "ManageJobPage - Mozilla Firefox". The address bar displays the URL "https://norgrid.ntnu.no/gridportal/ManageJobPage". The page content includes the "GRID portal" logo and the text "the NTNU Application Portal" with the subtitle "High Performance Computing Project". A navigation menu contains links for "Information", "Authorization", "Session", "Preferences", "Create", "Manage", and "About...". The main content area is titled "Manage jobs" and lists three jobs: "blastpJob", "tblastxJob", and "MatlabJob", each with a radio button. Below the list are buttons for "Edit", "View files", "Submit", "Delete", and "Reset". The status bar at the bottom shows "Done" and "norgrid.ntnu.no" with an "Adblock" extension icon.

1 Introducing GRIDportal

2 Working with mpiBLAST

What is mpiBLAST?

A flavor of BLAST designed to run on a cluster (a number of computers [nodes] connected together). **BLAST and mpiBLAST are near identical in use.**

The way mpiBLAST works is that it

- 1 divides the database into equal segments,
- 2 distributes each segment onto a node,
- 3 performs BLAST search on each node in parallel,
- 4 and merges the results from each node into a common result set.

Creating a BLAST job

PLMpiBlastJobPageblastp - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://norgrid.ntnu.no/gridportal/PLMpiBlastJobPage

PLMpiBlastJobPageblastp

Create a blastp job

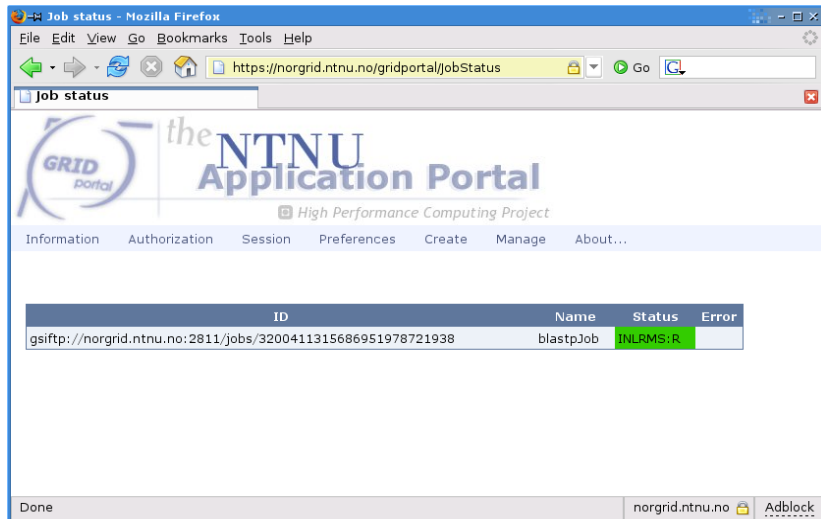
* CPU time (min)	<input type="text" value="30"/>
* Job name	<input type="text" value="blastpJob"/>
Email notification	<input type="text"/>
* Enter query sequence(s)	<input type="text"/>
Sequences input file	<input type="text" value="/home/alex/kombinert_tre.txt"/> <input type="button" value="Browse..."/>
Location on query sequence [10,400]	<input type="text" value="0,100"/>
Database	<input type="text" value="refseq_protein"/>
Filter query sequence	<input type="text" value="Yes"/>
Expectation value (E)	<input type="text" value="10.0"/>
Word Size	<input type="text" value="3"/>
Matrix	<input type="text" value="BLOSUM62"/>
HTML format on output	<input type="text" value="No"/>
Number of alignments reported	<input type="text" value="250"/>
Number of scores/descriptions reported	<input type="text" value="500"/>

Done norgrid.ntnu.no

Submitting a BLAST job

The screenshot shows a Mozilla Firefox browser window titled "ManageJobPage - Mozilla Firefox". The address bar contains the URL "https://norgrid.ntnu.no/gridportal/ManageJobPage". The page content includes the logo for "the NTNU Application Portal" and "High Performance Computing Project". A navigation menu contains links for "Information", "Authorization", "Session", "Preferences", "Create", "Manage", and "About...". The main content area is titled "Manage jobs" and lists three jobs: "blastpJob" (selected), "tblastxJob", and "MatlabJob". Below the list are buttons for "Edit", "View files", "Submit", "Delete", and "Reset". The status bar at the bottom shows "Done" and "norgrid.ntnu.no" with an "Adblock" extension icon.

Checking up on a job



The screenshot shows a Mozilla Firefox browser window titled "Job status - Mozilla Firefox". The address bar contains the URL "https://norgrid.ntnu.no/gridportal/jobStatus". The page header features the "GRID portal" logo and the text "the NTNU Application Portal" and "High Performance Computing Project". A navigation menu includes "Information", "Authorization", "Session", "Preferences", "Create", "Manage", and "About...". Below the menu is a table with the following data:

ID	Name	Status	Error
gsiftp://norgrid.ntnu.no:2811/jobs/3200411315686951978721938	blastpJob	INLRMS:R	

The status "INLRMS:R" is highlighted in green. The browser status bar at the bottom shows "Done" and "norgrid.ntnu.no" with an "Adblock" button.


Getting job results

ManageGridJobPage - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://norgrid.ntnu.no/gridportal/ManageGridJobPag Go

ManageGridJobPage

 the **NTNU**
Application Portal
High Performance Computing Project

Information Authorization Session Preferences Create Manage About...

Manage GRID jobs

JobID	JobName	Status
<input checked="" type="checkbox"/> gsiftp://norgrid.ntnu.no:2811/jobs/3200411315686951978721938	blastpJob	FINISHED

Done norgrid.ntnu.no Adblock

Viewing results of a job

File view - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://norgrid.ntnu.no/gridportal/ViewFilesPage

File view

the NTNU Application Portal
High Performance Computing Project

Information Authorization Session Preferences Create Manage About...

Downloaded job files

Type	File	Size	Last accessed	Last modified
	gmllog	4096	Wed Nov 9 21:41:27 2005	Wed Nov 9 21:41:35 2005
	martin_testset.txt	1846	Wed Nov 9 21:41:20 2005	Wed Nov 9 21:41:21 2005
	.modulesbeginenv	1601	Wed Nov 9 21:41:22 2005	Wed Nov 9 21:41:22 2005
	stdout.txt	0	Wed Nov 9 21:41:23 2005	Wed Nov 9 21:41:23 2005
	stderr.txt	0	Wed Nov 9 21:41:24 2005	Wed Nov 9 21:41:24 2005
	blast_out.QUERY	88420	Wed Nov 9 21:41:26 2005	Wed Nov 9 21:41:26 2005

View Download Download all (tar.gz) Back Reset

Done norgrid.ntnu.no Adblock

Viewing job files

File view - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://norgrid.ntnu.no/gridportal/ViewFilesPage

File view

blast_out.QUERY

BLASTP 2.2.10 [Oct-19-2004]

Reference: Aaron E. Darling, Lucas Carey, and Wu-chun Feng,
"The design, implementation, and evaluation of mpiBLAST."
In Proceedings of ClusterWorld 2003, June 24-26 2003, San Jose, CA

Query= gi|18450252:40-338 replicase [Garlic virus A] Methyltransferase
(868 letters)

Database: refseq_protein
1,804,530 sequences; 665,153,773 total letters

Sequences producing significant alignments:

	Score (bits)	E Value
ref NP_569126.1 replicase [Garlic virus A]	210	6e
ref NP_620648.1 replicase [Shallot virus X]	183	8e

Done norgrid.ntnu.no Adblock

- GRIDportal project website
<<http://gridportal.dynalias.org/>>
- GRIDportal deployment site
<<http://norgrid.ntnu.no/gridportal/>>

Thank you for your attention!