Application faq

Application specific Information FAQs

Recent changes

As of Monday May 18, 2009, how you start interactive sessions on the cluster will change. This is a very minor change needed to account for usage and to streamline application/lsf integration support. Please note this has no effect for batch submissions.

For example, currently you access an interactive Matlab session with the Graphical interface by:
-bash-3.2$ module load matlab
-bash-3.2$ matlab

The new approach is:
-bash-3.2$ module load matlab
-bash-3.2$ bsub -Ip -q int_public6 matlab

Note, you must still load the module corresponding to the application you need to use.
This will launch matlab in queue int_public6, as an interactive session. This is a new queue for interactive performance with a 4 hour limit. Choose this queue for interactive jobs. Contributed node users will have a similar queue. For additional information about new queues click **here**.

Current list of software

You may check the current name of loadable software by:

-bash-3.2$ module available

When in doubt, always be version explicit in the named module you load.

The following table lists the new submission format for interactive sessions. In some cases the program name may have a version specific name. This in not an exhaustive list.

application	actual application name	bsub command to run	Comments
matlab	**matlab**	bsub -Ip -q int_public6 matlab	graphics interface
maple	**xmaple**	bsub -Ip -q int_public6 xmaple	graphics interface
maple	**maple**	bsub -Ip -q int_public6 maple	text interface
mathematica	**mathematica**	bsub -Ip -q int_public6 mathematica	graphics interface
mathematica	**math**	bsub -Ip -q int_public6 math	text interface
comsol	**comsol**	bsub -Ip -q int_public6 comsol	graphics interface
paraview	**paraview**	bsub -Ip -q int_public6 paraview	graphics interface
xstata	**xstata-se**	bsub -Ip -q int_public6 xstata-se	graphics interface to SE version
stata	**stata-se**	bsub -Ip -q int_public6 stata-se	text interface to SE version
xstata	**xstata**	bsub -Ip -q int_public6 xstata	graphics interface to IC version
stata	**stata**	bsub -Ip -q int_public6 stata	text interface to IC version
ansys	**launcher121**	bsub -Ip -q int_public6 launcher121	graphics interface
ansys workbench	**runwb2**	bsub -Ip -q int_public6 runwb2 -mesa	WorkBench interface(Note:you may or may not need the -mesa option depending on your OpenGL support and version of Workbench. Version 13 doesn't need it.
fluent	**fluent**	bsub -Ip -q int_public6 fluent	graphics interface to 2D and 3D programs

fluent icepak	**icepak**	bsub -Ip -q int_public6 icepak	graphics interface to Icepak
ggobi	**ggobi**	bsub -Ip -q int_public6 ggobi	graphics interface
R	**R**	bsub -Ip -q int_public6 R	graphics delivery enabled
abaqus	**abq692.exe**	bsub -Ip -q int_public6 abq692.exe	command line interface
abaqus cae	**abq692.exe cae**	bsub -Ip -q int_public6 abq692.exe cae -mesa	CAE interface
abaqus 6.9-EF2	**abq69ef2.exe**	bsub -Ip -q int_public6 abq69ef2.exe	command line interface
weka	**weka.jar**	bsub -Ip -q int_public6 java -jar /opt/weka/3-5-8/weka.jar	graphics interface
tecplot	**tec360**	bsub -Ip -q int_public6 tec360 -mesa	graphics interface
fe-safe	**FESafe**	bsub -Ip -q int_public6 FESafe	graphics interface
SAS	**sas_u8**	bsub -Ip -q int_public6 sas_u8	graphics interface
RATS	**rats**	bsub -Ip -q int_public6 rats	command line interface

Cluster scaling considerations

The use of commercial packages like Matlab, Ansys, Mathematica, etc., often have usage scenarios for which one solution does not fit all situations. For example:

Compute Needs	**Resoucres**
Development	Laptop/Desktop installation of software
Batch runs, no GUI	cluster queues(Normal, Long, dregs)
SMP parallelism	Exclusive compute node access
MPI parallelism	application dependent, parallel queue
Parametric Simulation	various long running queues

If you are uncertain of a suitable approach, please contact cluster-support@tufts.edu for assistance.

Ansys, Fluent and Polyflow related.

Now that Ansys purchased Fluent, how does one access Fluent functionality?
Most Fluent functionality is incorporated into Ansys WorkBench.

How do I run Ansys WorkBench interface?

-bash-3.2$ module load ansys
-bash-3.2$ bsub -Ip -q int_public6 runwb2

How do I run Ansys Polyflow?
This package is accessed via the Ansys WorkBench product.
-bash-3.2$ module load ansys
-bash-3.2$ bsub -Ip -q int_public6 runwb2

How do I access the PolyFlow tutorial files referenced in the install directory /opt/shared/ansys_inc/v121/polyflow/polyflow12.1.1/ ?

New versions of the missing files are located at:
/cluster/shared/dmarshal/ISO/PolyFlow/

How can I read Ansys/Fluent/WorkBench and other docs while on the cluster?
You may run the Ansys Help application after you set up the module by:
-bash-3.2$ bsub -Ip -q int_public6 anshelp130

How do I run ansys in batch mode without using the Launcher?
-bash-3.2$ bsub -Ip -q normal_public6 ansys130

You will be prompted for an Ansys style input file(with a .dat extension). You may also add command line options for input, output, jobname and more after the **ansys130** command. By doing so you may drop the -Ip bsub option for interactive screen content, suitable for long running unattended jobs.

The following Fluent related info may no longer work due to Ansys incorporation of Fluent into WorkBench.

How do I run a Fluent parallel job?
Define the number of cpus(up to 8) within the bsub and fluent command line. The -n and -t arguments is the location to change the number of cpus.

-bash-3.2$ bsub -I -q queuename -n8 -a fluent fluent 2d -g -t8 -pib -mpi=intel -lsf -i yourfluent.script
Note: replace 2d with the appropriate fluent product, such as 3d and so on...

- What is a useful Fluent resource?*
  CDF-Online

How can I view the Fluent docs?
> module load fluent
> firefox file:///opt/Fluent.Inc/fluent6.3.26/help/index.htm

Where can I find example files?

/opt/shared/ansys_inc/v130/ansys/data/models/

How do I get fluent command line options?
-bash-3.2$ bsub -Ip -q int_public6 fluent -help

How do I run the 2D or 3D fluent without asking the questions interactively?
-bash-3.2$ bsub -Ip -q int_public6 fluent 3d

How do I run Icepak?
Icepak is available from within Ansys WorkBench interface and as a command line standalone application. Start WorkBench in the normal fashion, as above. Icepak will be a choice on the Components Pane, click on it and then click setup. To run just the Icepak interface try:

-bash-3.2$ module load ansys
-bash-3.2$ bsub -Ip -q int_public6 icepak

Abaqus related.

How do I run Abaqus?
There are several ways to run jobs. Most will find the following useful for checking syntax,etc..:
-bash-3.2$ module load abaqus
-bash-3.2$ bsub -Ip -q int_public6 abq682.exe job=mydesign datacheck interactive input=./design.inp

To actually run the code, remove the option **datacheck** and **interactive**.

This will run your abaqus input file called design.inp as a serial job and if successful will create about 14 files in the directory you are working in. One such file is the .odb Abaqus database file. All the files are created with the job name "mydesign" and differ in the filename extensions and type of file.

Note, normally there are several versions of Abaqus available. Therefore one needs to note which one reflects what was set up with Modules. For example, the command **abq682.exe** is version specific.

How do I run a parallel Abaqus job?
-bash-3.2$ bsub -Ip -q int_public6 abq682.exe job=design_4cpus cpus=4 mp_mode=mpi input=./design.inp

Note different Abaqus products may invoke different solvers. Check the documentation for options.

How can I find the abaqus command line options?
Suppose you are using ver. 6.9-2.

-bash-3.2$ module load abaqus/6.9-2
-bash-3.2$ bsub -Ip -q int_public6 abq692.exe -help

How do I run Abaqus version 6.9.2 CAE gui?
-bash-3.2$ bsub -Ip -q int_public6 abq692.exe -cae

Why is OpenGl not working with CAE?
Add the -mesa option to the start of abaqus as mentioned above.

How do I read Abaqus documentation?
If you have X11 forwarding in place you will get a webpage with documentation.

-bash-3.2$ module load abaqus
-bash-3.2$ bsub -Ip -q int_public6 abq682.exe doc

Where do I find the Abaqus 3rd party product FE-Safe documentation?
Look in: /opt/shared/fesafe/version.5.4/documentation/userdocs.pdf

Comsol related.

How do I run Comsol as a serial job?
-bash-3.2$ module load comsol
-bash-3.2$ bsub -Ip -q int_public6 comsol

How do I run Comsol as a parallel job?
Start Comsol as above, but add the command line options as follows:
-bash-3.2$ bsub -Ip -q int_public6 comsol -np 8 -mpmode owner

Then in the gui setup for Solver parameters, change the default Solver to the Parallel Direct Solver(Pardiso) and rerun your model.

To do the same task but on an node with exclusive access try:
-bash-3.2$ bsub -Ip -q int_public6 -x -R "spanptile=8" comsol -np 8 -mpmode owner

The comsol help button doesn't work, how do I find the documentation?
The compute nodes do not have browsers installed. Firefox is available on the headnode of the cluster instead. The install tree for Comsol contains html docs and some pdfs as well.

$ ls /opt/comsol35a/doc/
aco browser chem multiphysics plain sme

For example to view the html chem docs:
$ firefox file:///opt/comsol35a/doc/chem/index.htm

Sometimes you may find that your comsol efforts no longer work as they may have in the past; resulting in java errors of several types. These errors will be noted in your log file near the end. What can one do to get around this?

Close any open Comsol sessions, and remove your .comsol directory located in your home directory.
> cd
> rm -rf .comsol
Everything in that dot directory and below will be erased. Restart comsol and a new set of structures will be created.

How do I run comsol in batch mode?

> module load comsol

To run Comsol's tube1b.mph example:

>bsub -Ip -q normal_public6 comsol batch -inputfile tube1b.mph -outputfile tube1b-output.mph

The solution will update a new version of the .mph file as output. You can read the output file into Comsol later for further post processing. There will be 3 output files and a directory created with name **tube1b_files** containing graphics and other files produced from the batch job.

Note, you may then log off while your job runs. YOu will be notified by email when the job is done. Queue **int_public** has a run time limit of 4 hours. Choose a different queue if you expect longer runs. For example:

> bqueues

> bqueues -l normal_all6
has a runtime limit of 3 days. And so forth.....

DEFORM 3d

Your login shell must be the csh shell.

To test to see what your shell is, login and type:

> echo $SHELL

if it returns **/bin/csh** then you are set. Anything else will need to be altered. This is how:

> chsh -s /bin/csh

put in your password. All subsequent logins will be with **csh** and not **bash** shell.

Note: depending on load, shell changes can take from less than a minute to up to several minutes. This is due to the latency of updating account information.

Then proceed to:

> module load DEFORM

> bsub -Ip -q exint_deform /opt/shared/DEFORM/start-deform3

GAMESS

-bash-3.2$ module load gamess
-bash-3.2$ bsub -Ip -q parallel_public6 -n 16 rungms exam02

This will submit an interactive batch run of exam02.inp on 16 cores.

Or to capture output to a file:
-bash-3.2$ bsub -Ip -q parallel_public6 -n 16 rungms exam02 > exam02.output

GAMESS sample files, exam*.inp, can be found at:
/cluster/shared/dmarshal/ISO/gamess_examples

Imagemagick related.

Where do I find information on Imagemagick graphic tools?

Imagemagick home, usage info and scripting links. Also local man pages on individual commands is available. Note Imagemagick is not a command. It is a suite of programs.

For example to use the convert program:
-bash-3.2$ bsub -Ip -q int_public6 convert ...your convert arguments here...

How does one display a picture using Imagemagick's display command without running the interface
-bash-3.2$ bsub -Ip -q express_public6 display my_test.gif

Matlab related:

- How do I get matlab command line options?*
  -bash-3.2$ bsub -Ip -q int_public6 matlab -help

How do I run a matlab batch job?

-bash-3.2$ module load matlab

-bash-3.2$ bsub -q queuename "matlab -nojvm -nodisplay < your-matlab-code.m"

This will submit your code to the queue you choose. Matlab has several ways to direct output and results by use of matlab commands such as: diary, print and several other commands.

How do I speed up my Matlab calculations?
By default Matlab on the cluster has multi-thread parallelism enabled. Many Matlab functions will automatically use the threaded libraries. The number of threads obtained is equal to the number of cores on the node your job is placed on. Currently it is either 8 or 12. You no longer need to request exclusive node access via bsub.

How do I access the Parallel capability of Matlab?
Matlab can make use of MPI parallelism if you modify your code accordingly(see Parallel computing documentation within Matlab) and setup your session to instruct Matlab to use **LSF**. Currently Distributed Computing Engine Toolbox is licensed for 32 cores. It is strongly recommended to initially use 4, 8 or 16 cores while you explore how your code scales. Depending on what you are doing, execution times may not benefit from additional cores. To setup this feature, import a working configuration found at:

/cluster/shared/dmarshal/ISO/matlab-R2012-lsfconfig16.settings

Use Matlab's Cluster Profile Manager Tool to import the settings. Choose the menu options **Parallel** -> **Import Cluster Profile** and browse to the config file.

To test this out, you may run the Validation tests by choosing **Start Validation**. This will take several minutes and should result in five passed tests.

I want to run hundreds of matlab jobs using my matlab scripts. What do I need to know?
Your jobs will end up on a queue with a limit of 100 concurrent jobs/user. Our matlab license is sized at 65 concurrent seats. These floating license seats serve 3 campuses for desktop use as well. Due to the non predictive nature of random demand, we ask that you do not submit more than 10 jobs at any given time. This is a crude rule of thumb. For one or two jobs there are usually no problems. However during peak usage times this can be a hit or miss situation relative to the available number of matlab licenses.

The solution is to use Matlab's Compiler to compile your code into a standalone executable file. When done in this manner, you submit the compiled file instead of the raw matlab code, for processing and the 65 seat license limitation is then removed. So in principle you may submit 1000 jobs with up to 100 of them available for concurrent execution. The Matlab Compiler is a Matlab Toolbox with its own documentation and is available on the cluster. Please see the documentation for additional examples.

Is Matlab MCR available to run compiled jobs?

Yes. You will need to load the MCR version corresponding to the Matlab version you used to compile your standalone executable. The Matlab example called magicsquare in the documentation can be used as an example. Once compiled the named executable, magicsquare, can be run without loading Matlab nor checking out licenses. Such an example is available below. Note, be sure to quote any needed command line arguments to your named executable.

> module load MCR/R2012a
> bsub -Ip -q int_public6 "/cluster/shared/dmarshal/ISO/magicsquare 3"

How do I access the public domain Matlab toolbox EDGE?

EDGE has been installed into cluster directory /opt/shared/EDGE/1.02. To use it you'll need to launch MATLAB then change into the /opt/shared/EDGE/1.0/Matlab directory by issuing the following command at the Matlab prompt:

cd /opt/shared/EDGE/1.0/Matlab

Or better yet, add this path to your Matlab path within Matlab. Choose **File** -> **Set Path** and **Add with subfolders**.

To launch the **Importer** tool from within Matlab type on the command line:

semiauto

See the **EDGE website** for additional documentation.

How do I access the public domain Matlab toolbox WaveLab?
Load the corresponding module:
> module load Wavelab
> module load matlab

In Matlab choose **File** -> **Set Path** and then **Add with Subfolders**. Add the following: **/opt/shared/Wavelab/Wavelab850**

Then issue the command **startup** at the Matlab command line.

At this point you can issue commands and demos. Check the **website** for documentation.

Mathematica related.

Where can I find more Mathematica documentation?

The vendor's website and the online help in the application. In addition look at the broadcast section on the vendor site to view various movies on how to do many tasks in Mathematica. Also check out some paper based **options**.

My Mathematica notebook file doesn't work like it did before, what is up?
The following explanation is a good starting point to explore this **issue**.

How does one submit a mathematica batch job?
Suppose you have a text file called my_mathematica.m created with a text editor with the following tasks:

AppendTo[$Echo, "stdout"]
Export["test3.gif",Plot[x^2,{x,0,2}]]
Export["test4.gif", Plot[x+x^2,{x,0,2}]]

To submit as a batch job through the interactive queue but not the Mathematica gui interface:
> bsub -Ip -q express_public6 math < my_mathematica.m > my_math.out

Adjust your queue choice for longer running jobs than that support by the **express_public** queue.

Alternatively, try a variation like this:
> bsub -q normal_public6 -o my_math.out -i my_mathematica.m math

How does one submit a batch job using as input a Mathematica notebook and not a text file of Mathematica commands?

It is convenient to use the notebook **.nb** file format as input. However this is done normally in an interactive GUI session. To simulate this X based environment use the Xvfb functionality. For example, to submit a job:

-bash-3.2$ bsub -q normal_public6 xvfb-run mathematica your_notebook.nb

Add to your notebook the correct mathematica save and quit commands at the end.

Take a look at Mathematica documentation on commands:
NotebookSave and **NotebookClose**

Also notice that there are a few other scenarios and commands to do final housekeeping to existing notebooks that might serve you as well.

I get an error about Mathematica missing fonts,etc... What does this mean?

The cluster does not provide a Font Server supporting Mathematica font sets. Whereas a local install of Mathematica on your desktop includes this font set. Remote users need to provide access to the fonts locally.

For linux desktops, the easy way is to obtain the Mathematica helper program called, Mathematica Player. You may download this from the vendor website. Before you start Mathematica on the cluster for display on your computer, start the helper program locally. This will provide the font support.

For MacOSx desktops you need to either install Mathematica or the Mathematica Player and then point to the installed fonts. For example if you have Mathematica installed, to use it's fonts, open a terminal window(shell) and type:
> xset fp+ /Applications/Mathematica.app/SystemFiles/Fonts/Type1/
> xset fp rehash
Assuming that is the path to the fonts.
Then connect to the cluster and start Mathematica normally.

NCAR related.

How do I start NCAR?
NCAR has lots of separate programs. These are all available to you after you setup the environment.

-bash-3.2$ module load NCAR

These executables are located at: /opt/shared/NCAR/5.2.1/bin
To submit a job, for example:

-bash-3.2$ bsub -Ip -q express_public6 ng4ex gsun01n.ncl

Please consult the NCAR website for user documentation and examples.

Python related.

How can I verify if a particular Python package is installed?
Add-on tools such as **numpy** and **scipy** are installed. Others would be under the install tree located at:
/opt/shared/python/
in the version specific **site-packages** directory.

PetSc related.

How do I install PetSc?
PetSc has a complex build environment that allows one to include various applications. We encourage interested users to build the suite of interest in their home directory. Installation instructions for a particular set of applications can be found **here**.

Perl related.

How can I verify if a particular Perl package is installed?
Suppose you wanted to see if package GD is installed:
-bash-3.2$ rpm -q perl-GD gd

PDF related.

How do I view pdf files on the cluster?
Use the X11 based program **evince**. It is already in your PATH.

RATS related.

Why RATS?
THe linux version of RATS is provided for legacy codes access.

How does one submit a RATS command file for processing?
> bsub -Ip -q int_public6 rats your_input_file.rpf > your_stuff.output

Is there a GUI interface?
Yes, but it is broken. Whenever it is fixed this will be updated.

Where can examples and data files be found on the web?
Check the RATS **browser**

Where are example files?
/opt/shared/RATS/8.0.2/examples/

R related.

How do I run a long running R job as a batch job and I require lots of ram?

You may send your R script to the normal_public6 queue requesting access to a 32gig ram node, as an example:

-bash-3.2$ bsub -q normal_public6 -R Mem32 R CMD BATCH your-inputfile.r outputfilename

How can I fine tune or limit memory usage in an R job?
R has a set of command line options for limiting memory. These are documented in the large R reference manual available on the R website. Note, these are meant for Windows machines and not linux!
The switches are:
--min-vsize=vl --max-vsize=vu --min-nsize=nl
--max-nsize=nu --max-ppsize=N mem.limits(nsize = NA, vsize = NA)

On the cluster, ulimit is set to unlimited and these Windows options will not toggle ulimit settings. Instead request minimum memory needed for your job by using the Mem8, Mem16, etc... bsub resource options.

For example, a large R job:
-bash-3.2$ bsub -q normal_public6 -o stdout.txt -e stderr.txt R CMD BATCH --no-save your_batch_job.R youroutput.txt

For example, a larger R job:
-bash-3.2$ bsub -R Mem8 -q normal_public6 -o stdout.txt -e stderr.txt R CMD BATCH --no-save your_batch_job.R youroutput.txt

Where is the R documentation?
Extensive user documentation and tutorials are available on the R **web site**. There are many texts as well, here is nice **example**

And if you are an **Spss or SAS** user check this out.

And another source for comparative statistical package example usage can be found at **UCLA**

Where can I find command line R documentation?
> module load R
> bsub -Ip -q int_public6 R CMD INSTALL --help

Or use the man pages available at the linux shell prompt.

-bash-3.2$ man R

Also of note is the following online CodeSchool **tutorial**

Additional info on R can be found in the admin **guide**

How can I install an R package that is only available as a tar file?
Download the tar file and from within R try the following. Make changes to the utln and path to reflect the location. For example:
> install.packages("/cluster/shared/utln/your_R_package.tar", repos = NULL)

You may also install a package into another location:

-bash-4.1$ bsub -Ip -q int_public6 R CMD INSTALL --library=/cluster/shared/utln /cluster/shared/utln/someRpackage.tar.gz

Change the **utln** to your login name and adjust the filesystem path.

You will need to add the location to R search path for libraries.
Within R type:

> .libPaths()
1 "/opt/shared/R/2.15.0-rhel6/lib64/R/library"

This is the default library system wide.

Now to add the new additional location:

> .libPaths(/cluster/shared/utln/)

> .libPaths()
1 "/cluster/shared/utln"
2 "/opt/shared/R/2.15.0-rhel6/lib64/R/library"

How can I install an R package from the CRAN repository?
From within R issue the commands:

> options(CRAN="http://cran.us.r-project.org/")
> install.packages("name_of_package")

Note that R package names are case sensitive.

Alternatively, try:

> install.packages(name_of_package,~utln/R-libs/,CRAN=getOption("CRAN"),destdir=~utln/R/)

edit the utln to reflect yor Tufts username and directory locations. Locations may be home directory or optional storage locations.

Check the R documentation for additional options.

Stata related.

How to choose between SE and IC editions?
Most users will likely only need IC. Both versions have the same statistical functionality but differ in the **size** of their data structures.

How does one send a Stata batch job to the cluster compute nodes?
Make a stata **do** file with stata commands. To send:
> bsub -q normal_public6 stata -b your_filename.do

What are the command line arguments to Stata?
To find out:
> bsub -Ip -q int_public6 stata -help

SAS related.

To run interactively:
> module load SAS
> bsub -Ip -q int_public6 sas_u8

For quick and short batch use:

>bsub -Ip -q int_public6 sas_u8 yourprogram.sas

You will get **yourprogram.log** and **yourprogram.lst** output files.

Tecplot related.

On the cluster Tecplot's help button doesn't work, where do I find TecPlot documentation?
Html docs can be found on the vendor's **website**
PDF versions are an option as well.

Are there any Tecplot tutorials?
See this **tutorial link**.

Bioinformatic programs and specifics.

How do I access Blast programs on the cluster?
-bash-3.2$ module load blast
-bash-3.2$ bsub -Ip -q int_public6 blastx ...
And likewise for other commands. Also consider what queue is best at meeting your compute needs.

Where are the blast commands located?
/opt/shared/ncbi-blast/bin

How do I start the serial version of MrBayes in interactive mode?
-bash-3.2$ module load mrbayes
-bash-3.2$ bsub -Ip -q int_public6 mb

How do I start the parallel version of MrBayes in interactive mode?
-bash-3.2$ module load mrbayes
-bash-3.2$ module load openmpi
-bash-3.2$ bsub -Ip -q int_public6 mb -openmpi

A similar solution is available for mpi using mvapich2.

How do I start Haploview?
-bash-3.2$ module load java
-bash-3.2$ module load haploview
-bash-3.2$ bsub -Ip -q int_public6 java -jar /opt/CTSA/Haploview4.1/Haploview.jar

How do I start plink?
plink requires the presence of two file with .ped and .med extensions. These correspond to the expected input format and corresponding data information. As an example copy the test.ped and test.med files from the installation directory to your home directory.
-bash-3.2$ cp /opt/CTSA/plink-1.06-x86_64/test* .

-bash-3.2$ module load plink
-bash-3.2$ bsub -Ip -q int_public6 plink --file test --noweb

How do I access online help for plink?
Command line help can be seen by:
-bash-3.2$ plink --help

Also check the **plink** site for additional up to date information.

HOw do I start velvet?
-bash-3.2$ module load velvet

Velvet has two versions: velvetg and velveth
To run one of these:
-bash-3.2$ bsub -Ip -q int_public6 velvetg (velvet arguments)
Without arguments velvet returns some command line options.

Where is velvet installed?
/opt/shared/velvet-1.0.02/

How do I start Merlin?
-bash-3.2$ module load merlin
-bash-3.2$ bsub -Ip -q int_public6 merlin

How do I start Pedcheck?
-bash-3.2$ module load pedcheck
For interactive command line use:
-bash-3.2$ bsub -Ip -q int_public6 pedcheck

Submitting as a batch job:
-bash-3.2$ bsub -Ip -q int_public6 pedcheck -p myfile.ped -d myfile.dat
where -p <pedigree file> and -d <locus file> are your input files.

What is required to run Solar?
This program requires a user specific ID key. Send a request to cluster-support@tufts.edu requesting one. This can take up to 24 hours for the vendor to respond. This is a one time registration event. Once you have received your key:

Key for username abcde01 is: 3gakiuxe
To register, log in with the specified username, run SOLAR, and use the
`register` command, for example:

solar> register 3gakiuxe

This will create a file named .solar_reg in the home directory.

How do I start Solar?
-bash-3.2$ module load solar
For interactive command line use:
-bash-3.2$ bsub -Ip -q int_public6 solar

Note, to follow the Readme file suggestions, run the .tcl scripts this way:
solar> makemibd
solar> doanalysis

How do I run Bowtie?
The install directory is /opt/shared/bowtie/bowtie-0.12.7/ where you will find sample indexes and reads in subdirectories of the same name. As an example, to run bowtie against ecoli indexes with a sample reads file:
-bash-3.2$ module load bowtie
-bash-3.2$ cp /opt/shared/bowtie/bowtie-0.12.7/indexes/e_coli.* .
-bash-3.2$ cp /opt/shared/bowtie/bowtie-0.12.7/reads/e_coli_1000.fq .
-bash-3.2$ bsub -Ip -q int_public6 bowtie e_coli e_coli_1000.fq > my_ecoli.outputfile

This will run an interactive job and dump the output to a text file, called my_ecoli.outputfile

How do I run Cufflinks?
The installed version has 3 binaries: cuffcompare cuffdiff cufflinks

Load the Cufflinks module to access them.
-bash-3.2$ module load cufflinks

To test it out, download the test data file from the cufflinks website or grab a copy from:
-bash-3.2$ cp /cluster/shared/dmarshal/cufflinks-test_data.txt .
To run an interactive job:
-bash-3.2$ bsub -Ip -q int_public6 cufflinks cufflinks-test_data.txt

How to run Tophat?

All tophat binaries are located at:
/opt/shared/tophat/1.0.14/bin/

To access them and the bowtie dependices:
-bash-3.2$ module load tophat
-bash-3.2$ module load bowtie

To test it out, download the test data set from the tophat site or grab the tar file from:
-bash-3.2$ cp /cluster/shared/dmarshal/ISO/tophat_test_data.tar .
-bash-3.2$ tar xvf tophat_test_data.tar
-bash-3.2$ cd test_data/

-bash-3.2$ bsub -Ip -q int_public6 tophat -r 20 test_ref reads_1.fq reads_2.fq

How to run IMa2?
-bash-3.2$ module load ima2/8.27.12

Refer to module output for explicit versions:
-bash-3.2$ module available

How to list command line arguments for IMa2?
-bash-3.2$ bsub -Ip -q int_public6 IMa2

How to run a typical IMa2 example?
-bash-3.2$ bsub -Ip -q int_public6 IMa2 -iyour_input_data -oyour_output_file_name -q2 -m1 -t3 -b10000 -l100 -s123
Note those commandline options are only for illustration, yours may be very different.

Documentation may be found on the IMa website.

How to get started with bioPerl?
Check the corresponding **BioPerl wiki**

How to start Structure and the interfaces?
Check the public Structure website for the manual. A parameter file is needed in both cases of usage below.

For the graphical interface:
-bash-4.1$ module load structure-java/2.3.4
-bash-4.1$ bsub -Ip -q int_public6 java -jar /opt/shared/structure/2.3.4/class/Structure.jar

For the command line text version:

-bash-4.1$ module load structure
-bash-4.1$ bsub -Ip -q int_public6 structure

How to start Desmond?

-bash-4.1$ module load Schrodinger

To see command line arguments for desmond, type:

-bash-4.1$ bsub -Ip -q int_public6 desmond

Several arguments are need to specify a proper desmond run.
Check the manual on the Desmond website for details.

Where do I find the PolyPhred programs and how do I access them?
The program binaries are in /opt/shared/phred/. You must load the phred module for access.

-bash-3.2$ module load phred

How does one work with diyabc on the cluster?
The cluster version is simply two linux executables, **diyabc_sim** and **diyabc_cat**. The first is for simulation generation and the second is for combining output reference table files. Unlike the Windows version, these are command line tools without a gui user interface. The manual pdf available on the diyabc website details usage. The following is how to run the example from the manual.
-bash-3.2$ module load diyabc
Copy the two input files to your local directory:
-bash-3.2$ cp /opt/shared/diyabc/ga_* .
-bash-3.2$ bsub -q short_public6 -o diyabc.output diyabc_sim ga_001 1 1000 4

Output files can be combined for transfer back to you your PC version for post processing and display.

How is QIIME started?
This functionality is built upon python modules and codes. The environment to support it has many dependencies that are satisfied by loading the QIIME module.

-bash-3.2$ module load QIIME

To submit a job from the QIIME tutorial:
-bash-3.2$ bsub -q normal_public6 check_id_map.py -m Fasting_Map.txt -o mapping_output

This example is straight forward. However enhancements to bsub options and to QIIME python functions will vary as will data sources locations.

Standalone HPC math libraries

The following libraries are installed. To check to see if there is an environment module associated with these, try:
-bash-3.2$ module avail

BLACS
BLAS
Lapack
Scalapack
SuperLU
gsl
UMFPack

Corresponding installation locations are found at:

SuperLU: /opt/shared/SuperLU_4.0/lib
Scalapack: /opt/shared/scalapack-1.8.0/
Lapack: /opt/shared/lapack-3.2.1/
BLAS: /opt/shared/BLAS/
BLACS: /opt/shared/BLACS/LIB/
UMFPACK: /opt/shared/UMFPACK/5.5.1
Cula: /opt/shared/cula/

Installed library files for **gsl**, **atlas** and many others are located in **/usr/lib64/**.

Linkage to specific library file(s) can be found in their respective install location at /opt/shared/. Note, you need to specify which libary file, since in general there may be more than one. For example:

> ls /opt/shared/BLACS/LIB/*.a
/opt/shared/BLACS/LIB/blacsCinit_MPI-LINUX-0.a /opt/shared/BLACS/LIB/blacsF77init_MPI-LINUX-0.a
/opt/shared/BLACS/LIB/blacs_MPI-LINUX-0.a

For example:
> gcc mycode.c -o mycode -L /opt/shared/SuperLU_4.0/lib -libsuperlu.a