Running Bioinformatics Applications at CCV

Mark Howison
Interim Executive Director, CCV

CCV/BioMed Bioinformatics Workshop, November 14, 2014
Software modules on Oscar

$ module avail :bio
$ module avail blast
$ module help blast
$ module load blast/2.2.28+

Need to request a software installation?

support@ccv.brown.edu
Interactive vs. batch jobs

$ interact ← MANUAL
    Forwards your session to a compute node
    If your connection is interrupted, the job terminates
    default: 1 core, 4GB memory, 30 minutes

$ interact -n 8 -m 20g
    8 cores, 20GB memory

$ sbatch myscript.sh ← AUTOMATED
    Runs the script on compute nodes in the background
    Writes the output to a file
Essential commands on Oscar

$ myq
   List all of your pending/running jobs

$ allq
   Same, but for the entire queue (all users)

$ nodes
   Summary of available nodes, how many are in use

$ myquota
   How much disk space you have left
Example batch jobs

$ sbatch ~/batch_scripts/blast.sh
   1-core, serial job

$ sbatch ~/batch_scripts/bowtie2.sh
   8-core, threaded job

$ sbatch ~/batch_scripts/raxml.sh
   16-core, distributed/MPI job

$ sbatch ~/batch_scripts/blast-array.sh
   Job array: launches several 1-core, serial jobs
bioinformatics Brew
The cross-platform package manager for command-line bioinformatics tools.

Currently supports:
- Ubuntu 13.10
- CentOS 6.3
- Mac

Installs to /opt/bib
BiB doesn’t get in the way of your other software or anything outside of its prefix. It installs each package in a self-contained directory tree, then links the program into /opt/bib/active/bin. Install your Python packages with pip and your non-bioinformatics software with the package manager of your choice.

Install your favorite tools
BiB provides all the common bioinformatics software you would expect, plus specialized tools that are often left out of other package managers and repositories.

Specify the exact version
And install multiple versions, with one active at a time.

$ bib install blast bowtieraxmlswipe
$ bib list
Installed packages:

<table>
<thead>
<tr>
<th>Name</th>
<th>Version</th>
<th>InstallStatus</th>
<th>ActiveStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>blast</td>
<td>2.2.29+</td>
<td>Installed</td>
<td>Active</td>
</tr>
<tr>
<td>bowtie</td>
<td>1.0.0</td>
<td>Installed</td>
<td>Active</td>
</tr>
<tr>
<td>raxml</td>
<td>7.7.6</td>
<td>Installed</td>
<td>Active</td>
</tr>
<tr>
<td>swipe</td>
<td>2.0.9</td>
<td>Installed</td>
<td>Active</td>
</tr>
</tbody>
</table>

$ bib install blast/2.2.29+ blast/2.2.29+
$ bib activate blast/2.2.29+
$ ls -L "which blastn"

$ bib install blast/2.2.28+ blast/2.2.28+
$ bib activate blast/2.2.28+
$ ls -L "which blastn"
“Stronghold”

Secure computing environment for sensitive data
Self-certified by Brown’s CISO as HIPAA compatible
Brown has recently signed BAA with Lifespan
100+ TB of storage capacity
Typical VM allocation: 2 cores, 8GB memory
Interested in being an early adopter?
   mhowison@brown.edu