1. Versions of Python on the ACCRE Cluster
2. Checking Installed Packages
3. Installing New Packages
4. Example Scripts
5. Managing Packages with Anaconda
6. Jupyter (iPython) Notebooks
7. Contributing New Examples

Python is an interpreted programming language that has become increasingly popular in high-performance computing environments because it's available with an assortment of numerical and scientific computing libraries (NumPy, SciPy, Pandas, etc.), relatively easy to learn, open source, and free.

1 Versions of Python on the ACCRE Cluster

When a user initially logs into the cluster, the system version (Python that comes pre-installed on the operating system) is 2.7.6 and is located in `/usr/local/bin`:

```
[bob@vmps14 ~]$ which python
/usr/local/bin/python
[bob@vmps14 ~]$ python --version
Python 2.7.6
```

This version of Python may be sufficient for simple tasks, but it does not include many of the essential numerical libraries cluster users need. ACCRE administrators have hand-compiled multiple other versions of Python that are linked against highly optimized linear algebras like ATLAS and Intel's MKL, and therefore will in general yield better performance (faster execution time). To see a list of installed versions of Python on the cluster, use Lmod's `spider` command:

```
[bob@vmps14 ~]$ ml spider -r '^Python'
----------------------------------------------------------------------------
  Python:
----------------------------------------------------------------------------
    Description:
    Python is a programming language that lets you work more quickly and int

     Versions:
        Python/2.7.12

----------------------------------------------------------------------------
 For detailed information about a specific "Python" module (including how to
  For example:
```

```
     $ module spider Python/2.7.12
----------------------------------------------------------------------------
```

Per Lmod's instructions, we can get more information about the installed Python version:

```
----------------------------------------------------------------------------
  Python: Python/2.7.12
----------------------------------------------------------------------------
    Description:
    Python is a programming language that lets you work more quickly and int


   You will need to load all module(s) on any one of the lines below before t

      GCC/5.4.0-2.26
      Intel/2016.3.210-GCC-5.4.0-2.26

    Help:

      Description
      ===========
    Python is a programming language that lets you work more quickly and int
       more effectively.


      More information
      ================
       - Homepage: http://python.org/


      Included extensions
      ===================
    arff-2.1.0, bitstring-3.1.5, blist-1.3.6, cryptography-1.4, Cython-0.24,
    dateutil-2.5.3, decorator-4.0.10, ecdsa-0.13, enum34-1.1.6, funcsigs-1.0
    lockfile-0.12.2, mock-2.0.0, netaddr-0.7.18, netifaces-0.10.4, nose-1.3.
    paramiko-2.0.1, paycheck-1.0.2, pbr-1.10.0, pip-8.1.2, pycrypto-2.6.1,
    pyparsing-2.1.5, pytz-2016.4, setuptools-23.1.0, six-1.10.0, virtualenv-
```

We have compiled Python 2.7.12 with both the GCC compiler and with Intel compilers, linking against Intel's MKL library, which should yield better performance on our Intel processors. Lmod tells us that we need to load either GCC or Intel as a dependency of Python, so let's do just that:

```
[bob@vmps14 ~]$ ml Intel/2016.3.210-GCC-5.4.0-2.26
[bob@vmps14 ~]$ ml Python/2.7.12
[bob@vmps14 ~]$ python --version
```

```
Python 2.7.12
[bob@vmps65 ~]$ which python
/opt/easybuild/software/Compiler/intel/2016.3.210-GCC-5.4.0-2.26/Python/2.7.12
```

In addition to Python, we have compiled the performant `numpy` library against GCC and MKL:

```
[bob@vmps12 ~]$ ml spider numpy/1.12.1-Python-2.7.12

----------------------------------------------------------------------------
  numpy: numpy/1.12.1-Python-2.7.12
----------------------------------------------------------------------------
    Description:
    NumPy is the fundamental package for scientific computing with Python. I
    other things: a powerful N-dimensional array object, sophisticated (broa
    tools for integrating C/C++ and Fortran code, useful linear algebra, Fou
    random number capabilities. Besides its obvious scientific uses, NumPy c
    efficient multi-dimensional container of generic data. Arbitrary data-ty
    This allows NumPy to seamlessly and speedily integrate with a wide varie


   You will need to load all module(s) on any one of the lines below before t

      GCC/5.4.0-2.26  OpenMPI/1.10.3
      Intel/2016.3.210-GCC-5.4.0-2.26  IntelMPI/5.1.3.181

    Help:

      Description
      ===========
    NumPy is the fundamental package for scientific computing with Python. I
     a powerful N-dimensional array object, sophisticated (broadcasting) fun
     code, useful linear algebra, Fourier transform, and random number capab
     NumPy can also be used as an efficient multi-dimensional container of g
     defined. This allows NumPy to seamlessly and speedily integrate with a


      More information
      ================
       - Homepage: http://www.numpy.org
```

as well as the `mpi4py` library:

```
[bob@vmps12 ~]$ ml spider mpi4py/2.0.0-Python-2.7.12

----------------------------------------------------------------------------
  mpi4py: mpi4py/2.0.0-Python-2.7.12
----------------------------------------------------------------------------
```

```
    Description:
    MPI for Python (mpi4py) provides bindings of the Message Passing Interfa
    the Python programming language, allowing any Python program to exploit


   You will need to load all module(s) on any one of the lines below before t

      GCC/5.4.0-2.26  OpenMPI/1.10.3
      Intel/2016.3.210-GCC-5.4.0-2.26  IntelMPI/5.1.3.181

    Help:

      Description
      ===========
    MPI for Python (mpi4py) provides bindings of the Message Passing Interfa
     the Python programming language, allowing any Python program to exploit


      More information
      ================
       - Homepage: https://bitbucket.org/mpi4py/mpi4py
```

Examples using these libraries can be found in ACCRE's GitHub repository

When users need to use non-standard Python packages, one highly recommended option is to use Anaconda. Anaconda simplifies managing dependencies and will be discussed in Section 4. Since Anaconda distributes packages as precompiled binaries, its expected that using Python via Anaconda might not be as performant as using the Python/2.7.12, but the advantages are chiefly reproducibility and encapsulation of code. Two versions of Anaconda are installed on the cluster:

```
[bob@vmps12 ~]$ ml spider Anaconda

----------------------------------------------------------------------------
  Anaconda2: Anaconda2/4.3.1
----------------------------------------------------------------------------
    Description:
    Anaconda is a freemium open source distribution of the Python programmin
    large-scale data processing, predictive analytics and scientific computi
      simplify package management and deployment.


    This module can be loaded directly: module load Anaconda2/4.3.1

    Help:
```

```
      Description
      ===========
     Anaconda is a freemium open source distribution of the Python
    programming language for large-scale data processing, predictive analyti
    scientific computing that aims to simplify package management and deploy


      More information
      ================
       - Homepage: https://www.continuum.io/anaconda-overview
```

```
------------------------------------------------------------------------------
  Anaconda3: Anaconda3/4.3.1
------------------------------------------------------------------------------
    Description:
    Anaconda is a freemium open source distribution of the Python programmin
    large-scale data processing, predictive analytics and scientific computi
      simplify package management and deployment.


   This module can be loaded directly: module load Anaconda3/4.3.1

    Help:

      Description
      ===========
     Anaconda is a freemium open source distribution of the Python
    programming language for large-scale data processing, predictive analyti
    scientific computing that aims to simplify package management and deploy


      More information
      ================
       - Homepage: https://www.continuum.io/anaconda-overview
```

Anaconda provides its own Python distribution, as well as the `conda` command for creating and managing environments:

```
[bob@vmps12 ~]$ ml Anaconda3
[bob@vmps12 ~]$ which python
/opt/easybuild/software/Core/Anaconda3/4.3.1/bin/python
[bob@vmps12 ~]$ which conda
/opt/easybuild/software/Core/Anaconda3/4.3.1/bin/conda
```

2 Checking Installed Packages

Each different version of Python has its own packages installed into it, so as you are switching between different versions it is prudent to check the packages that are available. One way to do this is using `pip`:

```
[jill@vmps14 ~]$ ml GCC Python
[jill@vmps14 ~]$ pip freeze | sort
bitstring==3.1.5
blist==1.3.6
cryptography==1.4
Cython==0.24
decorator==4.0.10
ecdsa==0.13
enum34==1.1.6
funcsigs==1.0.2
idna==2.5
ipaddress==1.0.18
liac-arff==2.1.0
lockfile==0.12.2
mock==2.0.0
netaddr==0.7.18
netifaces==0.10.4
nose==1.3.7
paramiko==2.0.1
paycheck==1.0.2
pbr==1.10.0
pyasn1==0.2.3
pycrypto==2.6.1
pyparsing==2.1.5
python-dateutil==2.5.3
pytz==2016.4
six==1.10.0
Theano==0.9.0
```

Using `pip` is convenient because it gives you package version information and also lists any packages you have installed locally (in your home directory). Occasionally you may also want to see a list of Python modules, which you can see by using the `python\_pkginfo.py` command:

```
[jill@vmps14 ~]$ python_pkginfo.py --type both

For output options type 'python_pkginfo.py -h'

Installed Packages and Modules for Python version 2.7.8
[GCC 4.6.1]:
```

```
To see package versions try: pip freeze | sort
after issuing the appropriate "setpkgs -a python" command

abc (Module)                 _multiprocessing (Module)
_abcoll (Module)             multiprocessing (Package)
aifc (Module)                mutex (Module)
alignlib_lite (Module)       _mysql (Module)
antigravity (Module)         _mysql_exceptions (Module)
anydbm (Module)              MySQLdb (Package)
appdirs (Module)             natsort (Package)
argparse (Module)            netrc (Module)
array (Module)               networkx (Package)
ast (Module)                 new (Module)
asynchat (Module)            nis (Module)
asyncore (Module)            nntplib (Module)
atexit (Module)              nose (Package)
audiodev (Module)            ntpath (Module)
audioop (Module)             nturl2path (Module)
backports (Package)          numbers (Module)
base64 (Module)              numpy (Package)
BaseHTTPServer (Module)      oncodriveclust (Package)
Bastion (Module)             oncotator (Package)
BCBio (Package)              opcode (Module)
bdb (Module)                 openpyxl (Package)
BeautifulSoup (Module)       operator (Module)
BeautifulSoupTests (Module) optparse (Module)
binascii (Module)            os (Module)
binhex (Module)              os2emxpath (Module)
Bio (Package)                ossaudiodev (Module)
BioSQL (Package)             _osx_support (Module)
bisect (Module)              pandas (Package)
_bisect (Module)             parse (Module)
bokeh (Package)              parser (Module)
brewer2mpl (Package)         past (Package)
bs4 (Package)                patsy (Package)
_bsddb (Module)              pdb (Module)
bsddb (Package)              pep8 (Module)
builtins (Package)           pickle (Module)
bx (Package)                 pickletools (Module)
bx_extras (Package)          pip (Package)
bz2 (Module)                 pipes (Module)
calendar (Module)            pkg_resources (Package)
Canvas (Module)              pkgutil (Module)
CDROM (Module)               platform (Module)
certifi (Package)            plistlib (Module)
CGAT (Package)               popen2 (Module)
```

```
CGATReport (Package)         poplib (Module)
CGATReportPlugins (Package) posixfile (Module)
CGATScripts (Package)        posixpath (Module)
cgi (Module)                 pprint (Module)
CGIHTTPServer (Module)       profile (Module)
cgitb (Module)               pstats (Module)
chunk (Module)               psyco_full (Module)
click (Package)              psycopg2 (Package)
cmath (Module)               pty (Module)
cmd (Module)                 py (Package)
code (Module)                py_compile (Module)
codecs (Module)              pybedtools (Package)
_codecs_cn (Module)          pyclbr (Module)
_codecs_hk (Module)          pycuda (Package)
_codecs_iso2022 (Module)     pydoc (Module)
_codecs_jp (Module)          pydoc_data (Package)
_codecs_kr (Module)          pyexpat (Module)
_codecs_tw (Module)          pygments (Package)
codeop (Module)              _pyio (Module)
collections (Module)         pylab (Module)
_collections (Module)        pyparsing (Module)
colorama (Package)           pysam (Package)
colorsys (Module)            pysqlite2 (Package)
commands (Module)            pystache (Package)
compileall (Module)          pytest (Module)
compiler (Package)           _pytest (Package)
concurrent (Package)         python_pkginfo (Module)
ConfigParser (Module)        pytools (Package)
configparser (Package)       pytz (Package)
contextlib (Module)          pyximport (Package)
Cookie (Module)              Queue (Module)
cookielib (Module)           queue (Package)
copy (Module)                quopri (Module)
copy_reg (Module)            random (Module)
copyreg (Package)            _random (Module)
corebio (Package)            rdflib (Package)
cPickle (Module)             re (Module)
cProfile (Module)            readline (Module)
crypt (Module)               repr (Module)
cStringIO (Module)           reprlib (Package)
csv (Module)                 requests (Package)
_csv (Module)                resource (Module)
_ctypes (Module)             rexec (Module)
ctypes (Package)             rfc822 (Module)
_ctypes_test (Module)        rlcompleter (Module)
_curses (Module)             robotparser (Module)
```

```
curses (Package)
_curses_panel (Module)
cython (Module)
Cython (Package)
datetime (Module)
dateutil (Package)
dbhash (Module)
dbm (Module)
decimal (Module)
decorator (Module)
deeptools (Package)
Dialog (Module)
difflib (Module)
dircache (Module)
dis (Module)
distutils (Package)
DLFCN (Module)
doctest (Module)
docutils (Package)
DocXMLRPCServer (Module)
drmaa (Package)
dumbdbm (Module)
dummy_thread (Module)
_dummy_thread (Package)
dummy_threading (Module)
easy_install (Module)
_elementtree (Module)
email (Package)
encodings (Package)
enum (Package)
ete2 (Package)
fcntl (Module)
filecmp (Module)
FileDialog (Module)
fileinput (Module)
FixTk (Module)
flask (Package)
fnget (Module)
fnmatch (Module)
formatter (Module)
fpformat (Module)
fractions (Module)
freetype (Package)
ftplib (Module)
functools (Module)
_functools (Module)
rpy2 (Package)
ruffus (Package)
runpy (Module)
sched (Module)
scimath (Package)
scipy (Package)
ScrolledText (Module)
seaborn (Package)
select (Module)
sets (Module)
setuptools (Package)
sgmllib (Module)
sha (Module)
shapely (Package)
shelve (Module)
shlex (Module)
shove (Package)
shutil (Module)
SimpleDialog (Module)
SimpleHTTPServer (Module)
SimpleXMLRPCServer (Module)
site (Module)
six (Module)
skbio (Package)
sklearn (Package)
smtpd (Module)
smtplib (Module)
sndhdr (Module)
socket (Module)
_socket (Module)
SocketServer (Module)
socketserver (Package)
SPARQLWrapper (Package)
sphinx (Package)
spwd (Module)
sqlalchemy (Package)
_sqlite3 (Module)
sqlite3 (Package)
sre (Module)
sre_compile (Module)
sre_constants (Module)
sre_parse (Module)
ssl (Module)
_ssl (Module)
stat (Module)
statsmodels (Package)
```

```
future (Package)
__future__ (Module)
future_builtins (Module)
futures (Package)
gdbm (Module)
genericpath (Module)
geojson (Package)
geopy (Package)
getopt (Module)
getpass (Module)
gettext (Module)
gevent (Package)
geventwebsocket (Package)
ggplot (Package)
glob (Module)
gnupg (Module)
greenlet (Module)
grp (Module)
gzip (Module)
hashlib (Module)
_hashlib (Module)
heapq (Module)
_heapq (Module)
hgapi (Package)
hmac (Module)
_hotshot (Module)
hotshot (Package)
html (Package)
html5lib (Package)
htmlentitydefs (Module)
htmllib (Module)
HTMLParser (Module)
http (Package)
httplib (Module)
idlelib (Package)
ihooks (Module)
imaplib (Module)
imghdr (Module)
importlib (Package)
imputil (Module)
IN (Module)
inspect (Module)
io (Module)
_io (Module)
IPython (Package)
isodate (Package)
statvfs (Module)
string (Module)
StringIO (Module)
stringold (Module)
stringprep (Module)
strop (Module)
_strptime (Module)
struct (Module)
_struct (Module)
stuf (Package)
subprocess (Module)
sunau (Module)
sunaudio (Module)
symbol (Module)
symtable (Module)
sysconfig (Module)
_sysconfigdata (Module)
syslog (Module)
tabnanny (Module)
tarfile (Module)
telnetlib (Module)
tempfile (Module)
termios (Module)
test (Package)
_testcapi (Module)
tests (Package)
textwrap (Module)
this (Module)
_thread (Package)
threading (Module)
_threading_local (Module)
threadpool (Module)
time (Module)
timeit (Module)
Tix (Module)
tkColorChooser (Module)
tkCommonDialog (Module)
Tkconstants (Module)
Tkdnd (Module)
tkFileDialog (Module)
tkFont (Module)
Tkinter (Module)
_tkinter (Module)
tkinter (Package)
tkMessageBox (Module)
tkSimpleDialog (Module)
```

```
itertools (Module)          toaiff (Module)
itsdangerous (Module)       token (Module)
jdcal (Module)              tokenize (Module)
jinja2 (Package)            tornado (Package)
_json (Module)              trace (Module)
json (Package)              traceback (Module)
keyword (Module)            traits (Package)
lib2to3 (Package)           ttk (Module)
libfuturize (Package)       tty (Module)
libpasteurize (Package)     turtle (Module)
linecache (Module)          types (Module)
linuxaudiodev (Module)      TYPES (Module)
locale (Module)             unicodedata (Module)
_locale (Module)            unittest (Package)
logging (Package)           urllib (Module)
lsblock (Module)            urllib2 (Module)
_lsprof (Module)            urlparse (Module)
_LWPCookieJar (Module)      user (Module)
lxml (Package)              UserDict (Module)
macpath (Module)            UserList (Module)
macurl2path (Module)        UserString (Module)
mailbox (Module)            uu (Module)
mailcap (Module)            uuid (Module)
mako (Package)              vcf (Package)
markdown (Package)          warnings (Module)
_markerlib (Package)        wave (Module)
markupbase (Module)         weakref (Module)
_markupbase (Package)       _weakrefset (Module)
markupsafe (Package)        web (Package)
math (Module)               webbrowser (Module)
matplotlib (Package)        weblogolib (Package)
matplotlib_venn (Package)   websocket (Package)
md5 (Module)                werkzeug (Package)
memcache (Module)           whichdb (Module)
mhlib (Module)              winreg (Package)
mimetools (Module)          wsgiref (Package)
mimetypes (Module)          xdrlib (Module)
MimeWriter (Module)         xlwt (Package)
mimify (Module)             xml (Package)
mmap (Module)               xmllib (Module)
mock (Module)               xmlrpc (Package)
modulefinder (Module)       xmlrpclib (Module)
more_itertools (Package)    _yaml (Module)
_MozillaCookieJar (Module)  yaml (Package)
mpi4py (Package)            zipfile (Module)
mpl_toolkits (Package)      zlib (Module)
```

```
mpld3 (Package)              zmq (Package)
_multibytecodec (Module)
multifile (Module)
```

python_pkginfo.py will also list any packages/modules you have installed locally in your home directory.

3 Installing New Packages

If you find that a particular package you need is missing from one of the Python versions, you have a few options. If you believe the package will be widely used by other cluster users, you can open a helpdesk ticket and request that we install the package for cluster-wise access. The other option is to install the package yourself into your home directory. There are multiple ways to install Python packages (`pip`, `easy_install`, from source). Just make sure you have the appropriate version of Python (via `module load`) in your environment when perform the installation. To install a package (e.g. a packaged called `word-count`) into your home directory with `pip`:

```
[jill@vmps14 ~]$ pip install word-count --user
Collecting word-count
  Downloading word_count-0.1.0-py2.py3-none-any.whl
Installing collected packages: word-count
Successfully installed word-count-0.1.0
[jill@vmps14 ~]$ ls ~/.local/lib/python2.7/site-packages/ | grep word_count
word_count-0.1.0.dist-info
word_count.py
word_count.pyc
[jill@vmps14 ~]$ pip freeze | grep word-count
word-count==0.1.0
```

Alternatively, you can build and install a package from source code into your home directory. See our FAQ page for instructions on doing this. Note that it is generally better practice to use Anaconda virtual environments and install new packages into the virtual environment (see instructions below). This helps isolate packages to specific projects and prevents complex dependency problems when, for example, you need a different version of a package for two different projects.

4 Example Scripts

Running a Python script within a SLURM job is generally straightforward. Unless you are attempting to run one of Python's multi-processing packages, you will want to request a single task, load the appropriate version of Python from

your SLURM script, and then redirect your Python file to the Python interpreter. The following example runs Python 2.7.12 on a simple Python script demonstrating the utility of writing vectorized Python code:

```
[bob@vmps14 run1]$ ls
python.slurm  vectorization.py
```

```
[bob@vmps12 vectorization]$ cat python.slurm
#!/bin/bash

#SBATCH --nodes=1
#SBATCH --constraint=haswell
#SBATCH --ntasks=1
#SBATCH --time=00:10:00
#SBATCH --mem=500M
#SBATCH --output=python_job_slurm.out

module load Intel IntelMPI Python numpy

python vectorization.py
```

```
[bob@vmps14 run1]$ cat vectorization.py
#!/usr/bin/env python
#
# Python 2.7 script demonstrating vectorized execution
#
import numpy as np
import time

# 10 million entries
t = np.linspace(-10,10,10000000)
x1 = np.zeros(len(t))
x2 = np.zeros(len(t))

time1 = time.clock()
# naive, non-vectorized implementation
for i,ti in enumerate(t):
    x1[i] = np.sin(ti)
time2 = time.clock()
print '%s: %0.2f seconds elapsed' % ("naive implementation", time2-time1)

# vectorized implementation
time1 = time.clock()
x2 = np.sin(t)
time2 = time.clock()
print '%s: %0.2f seconds elapsed' % ("vectorized implementation", time2-time1)
```

```
if ( np.array_equal(x1,x2) ):
    print "arrays equal!"
```

```
[bob@vmps14 vectorization]$ sbatch python.slurm
Submitted batch job 1832675
```

After waiting a few minutes:

```
[bob@vmps14 vectorization]$ ls
python_job_slurm.out  python.slurm  README.md  vectorization.py
```

```
[bob@vmps12 vectorization]$ cat python_job_slurm.out
naive implementation: 11.45 seconds elapsed
vectorized implementation: 0.15 seconds elapsed
arrays equal!
```

5 Managing Packages with Anaconda

Anaconda and its package manager "conda" aim to simplify package and dependency management. To load Anaconda, simply type

```
[bob@vmps14 ~] ml Anaconda2
```

or

```
[bob@vmps14 ~] ml Anaconda3
```

Both versions of Anaconda support installing any available Python version.

Anaconda encapsulates packages into virtual environments; in this way, packages between different projects do not conflict with one another. To create a virtual environment running Python version 3.4:

```
$ conda create --name myenvironment python=3.4
Fetching package metadata .......
Solving package specifications: ..........

Package plan for installation
in environment /home/bob/.conda/envs/myenvironment:

The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    openssl-1.0.2h             |                1         3.2 MB
    xz-5.2.2                   |                0         644 KB
    python-3.4.5               |                0        15.2 MB
    ------------------------------------------------------------
                                           Total:        19.0 MB
```

```
The following NEW packages will be INSTALLED:

    openssl:    1.0.2h-1
    pip:        8.1.2-py34_0
    python:     3.4.5-0
    readline:   6.2-2
    setuptools: 26.1.1-py34_0
    sqlite:     3.13.0-0
    tk:         8.5.18-0
    wheel:      0.29.0-py34_0
    xz:         5.2.2-0
    zlib:       1.2.8-3

Proceed ([y]/n)? y

Fetching packages ...
openssl-1.0.2h 100% |######################| Time: 0:00:03  1.06 MB/s
xz-5.2.2-0.tar 100% |######################| Time: 0:00:01 561.46 kB/s
python-3.4.5-0 100% |######################| Time: 0:00:03  4.00 MB/s
Extracting packages ...
[      COMPLETE      ]|##########################| 100%
Linking packages ...
[      COMPLETE      ]|##########################| 100%
#
# To activate this environment, use:
# $ source activate myenvironment
#
# To deactivate this environment, use:
# $ source deactivate
#
```

As mentioned in the conda message, to activate the environment, type:

```
[bob@vmps14 ~] source activate myenvironment
```

Then, to view the installed packages:

```
(myenvironment) [bob@vmps14 ~]$ conda list
# packages in environment at /home/bob/.conda/envs/myenvironment:
#
openssl                   1.0.2h                        1
pip                       8.1.2                     py34_0
python                    3.4.5                         0
readline                  6.2                           2
setuptools                26.1.1                    py34_0
sqlite                    3.13.0                        0
tk                        8.5.18                        0
wheel                     0.29.0                    py34_0
```

```
xz                        5.2.2                     0
zlib                      1.2.8                     3
```

Notice that "(myenvironment)" is prepended to the bash command prompt and that all the required packages for the setting up the environment are shown. New packages can be found using the *conda search* command, and they can be added to the current environment using *conda install*, for example:

```
(myenvironment) [bob@vmps14 ~]$ conda install numpy
Fetching package metadata .......
Solving package specifications: ..........

Package plan for installation
in environment /home/bob/.conda/envs/myenvironment:

The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    numpy-1.11.1               |           py34_0         6.1 MB

The following NEW packages will be INSTALLED:

    mkl:   11.3.3-0
    numpy: 1.11.1-py34_0

Proceed ([y]/n)? y

Fetching packages ...
numpy-1.11.1-p 100% |########################| Time: 0:00:00  10.65 MB/s
Extracting packages ...
[      COMPLETE      ]|################################| 100%
Linking packages ...
[      COMPLETE      ]|################################| 100%
```

Executing the code above installed numpy version 1.11.1 for Python 3.4 as well the package it depends on, mkl version 11.3.3. To use *myenvironment* in a batch script, it is necessary to first set the Anaconda package and then activate the environment in the *.slurm* file:

```
ml Anaconda3
source activate myenvironment
```

Other useful conda commands can be found in the Conda Cheatsheet.

6 Jupyter (iPython) Notebooks

Jupyter notebooks (formerly iPython notebooks) enable a user to interactively code in Python from a web browser with support for inline plotting, equation editing, among many other things. In general, a cluster environment is used for batch processing rather than interactive processing, however we do allow users to launch notebooks from compute nodes and then do editing on a local machine via a web browser. In order to do this, first submit a batch job with a SLURM script similar to the following:

```
#!/bin/bash

#SBATCH --ntasks=1
#SBATCH --mem=2G
#SBATCH --time=0-04:00:00      # 4 hours
#SBATCH --output=notebook.out

module load Anaconda3
jupyter notebook --no-browser --ip='*' --port=7777
```

Here, a user is requesting four hours of walltime and is using Anaconda3. The important line is the last one, where a few important options must be passed to the jupyter interpreter. Make sure you include these same options in your SLURM script. If you want to launch multiple notebooks from the same compute node, you would need to use a different port (try incrementing by one). Use *squeue* to monitor when your job has started. Once it has started, write down the name of the node where it is running and open `notebook.out` to see the password for this notebook; it should look something like

```
Copy/paste this URL into your browser when you connect for the first time,
          to login with a token:
           http://localhost:9999/?token=0gt123481a68b53a92490c64a1712
```

Open a Linux terminal/shell on your local machine (lab desktop, laptop, etc.), and enter the following command:

```
ssh -L 9999:vmp506:7777 vunetid@login.accre.vanderbilt.edu
```

Be sure to replace `vmp506` with the name of the node where your job is running, and `vunetid` with your VUNetID. This command binds port 9999 on `localhost` (i.e. your local machine) to port 7777 on `vmp506`. Finally, point a web browser on your local machine to `localhost:9999` (enter this where you would normally put a URL address), and you should get a Jupyter notebook login screen. Enter the token from above, and you'll launch a session in your web browser, where processing within the notebook will take place on the compute node. This ssh session must remain active in order for you to continue accessing the notebook from your local machine. You might consider adding these lines to your `~/.ssh/config` file to ensure your ssh session does not time out:

```
ServerAliveInterval 60
TCPKeepAlive yes
KeepAlive yes
```

Note that your account will be charged based on the length of your job, and not the amount of CPU time used by the job.

For more tips on using Jupyter notebooks, check out ACCRE's Python repository on GitHub.

7 Contributing New Examples

In order to foster collaboration and develop local Python expertise at Vanderbilt, we encourage users to submit examples of their own to ACCRE's Python Github repository. Instructions for doing this can be found on this page. You can also find the examples from Sections 4 and 5 in the ACCRE repository.