



Using Python on Blue Waters

Colin MacLean (NCSA/University of Illinois)
Email: cmaclean@illinois.edu



Python on Blue Waters

Python support on Blue Waters is provided by the BWPY module

```
$ module load bwpy
```

Python on Blue Waters

Python Interpreters

CPython 2.7 (aliases: python2, python)

CPython 3.3 (deprecated)

CPython 3.4 (alias: python3)

Cpython 3.5

Pypy

Pypy3

BWPY submodules

module load

bwpy-mpi

MPI support enabled

bwpy-libsci_mp

BWPY built with OpenMP Cray BLAS libraries (libsci_mp)

bwpy-libsci_acc

BWPY built with auto CUDA BLAS libraries (libsci_acc)

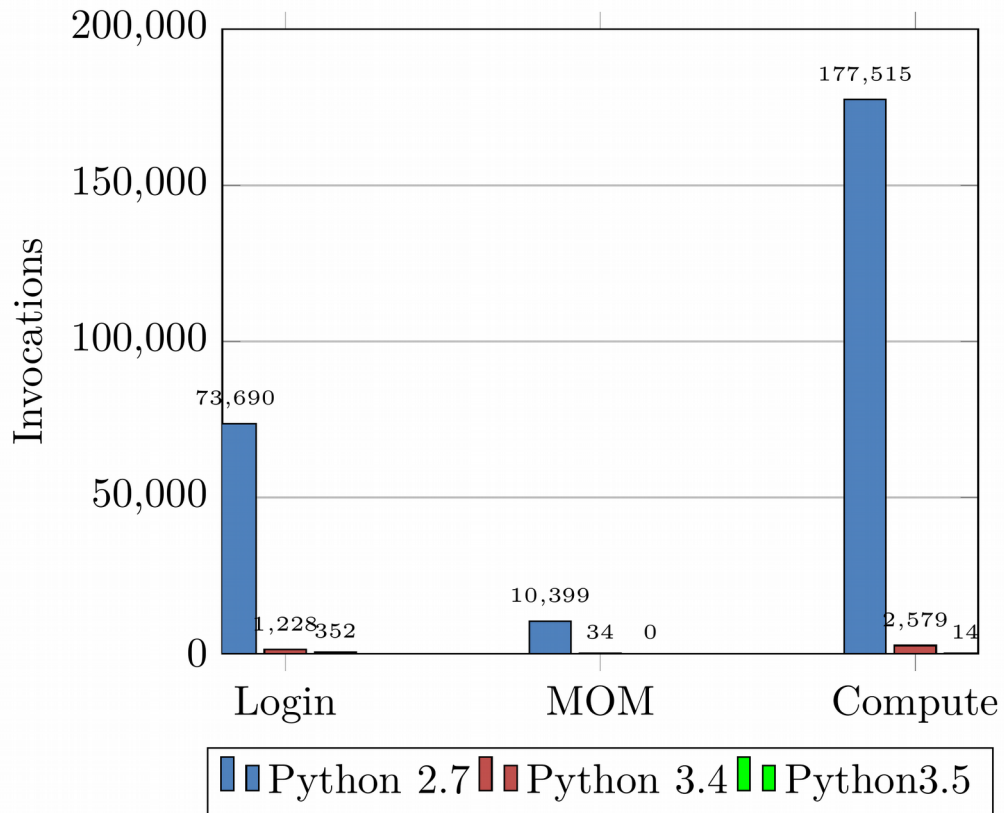
tensorflow-0.12.1

Newer Tensorflow (0.9.0 to 0.12.1)

tensorflow-1.0.0

Newer Tensorflow (0.9.0 to 1.0.0)

Python Versions

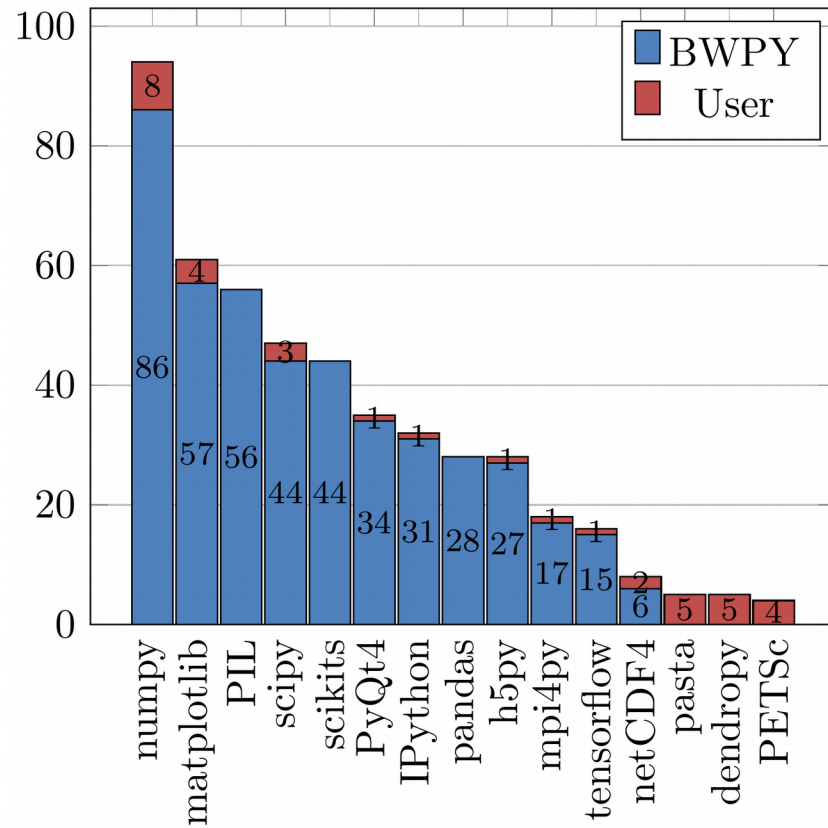


- BWPY supports CPython 2.7, 3.3, 3.4, 3.5, Pypy and Pypy3
- Python 2.7 dominates,
- Python 3.3, Pypy, and Pypy3 not used
- ~270,000 invocations

Running Python

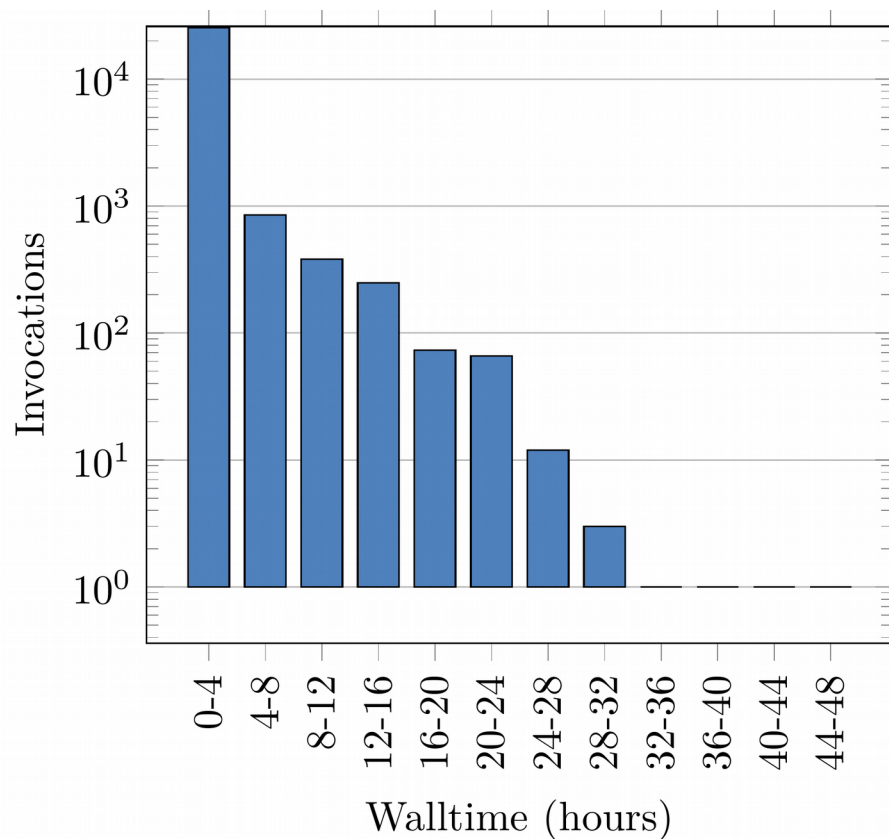
- Use `#!/usr/bin/env python2/3` shebang at start of script for portability
- Start moving to Python 3!
 - Python 2 will no longer be updated past 2020
 - 2 years, 10 months, 23 days from now is EoL
 - <https://pythonclock.org/>

Module Usage: By User



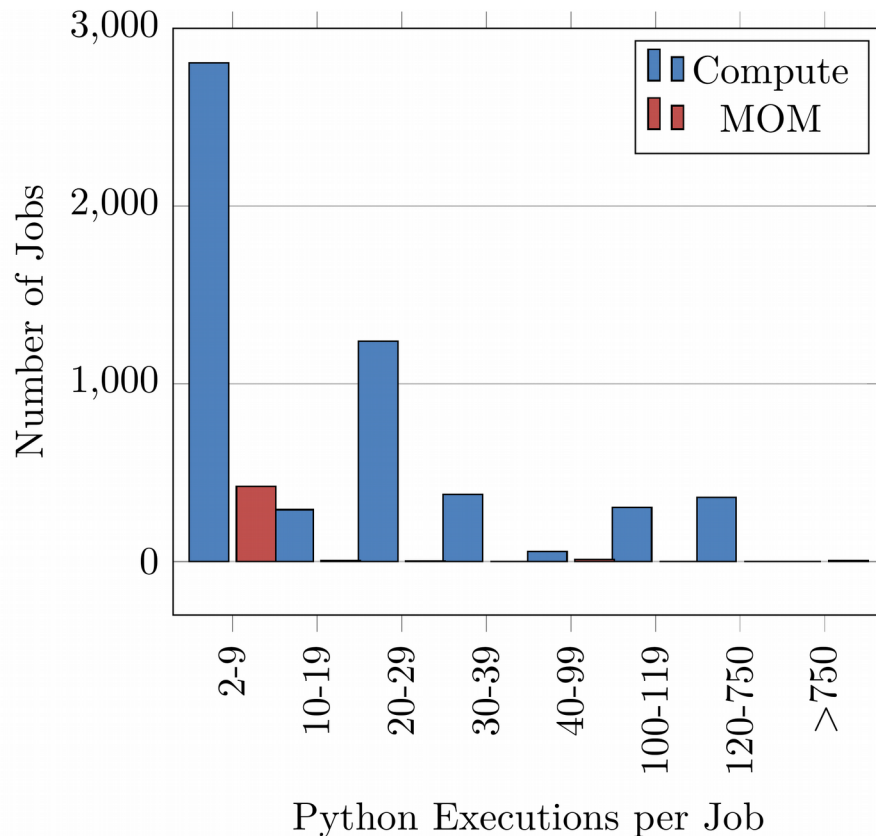
- 131 total users
- Unique module names imported: 663
- Those that do not use scipy, try PyPy
 - Incomplete numpy, but may be sufficient

Compute Node Job Walltimes



- 10893 jobs < 15 minutes
- 7004 jobs 15-30 minutes
- 6288 jobs 1/2-1 hour
- Only 162 jobs >16 hours
- Bundle jobs where possible

Executions per Job



- Some jobs execute Python many times
- Execute Python as few times as possible
- Write a script in Python to iterate through tasks

Compute Job Sizes

# of nodes	Invocations
1	170488
2-32	8555
33-64	53
65-128	222
129-256	600
257-512	94
513-1024	71
1025-2048	25

- The vast majority of Python jobs were single node
- How does slow Python startup time affect job width?



Thank You!

