HTCondor in MacPorts

MacPorts + HTCondor High Throughput Computing

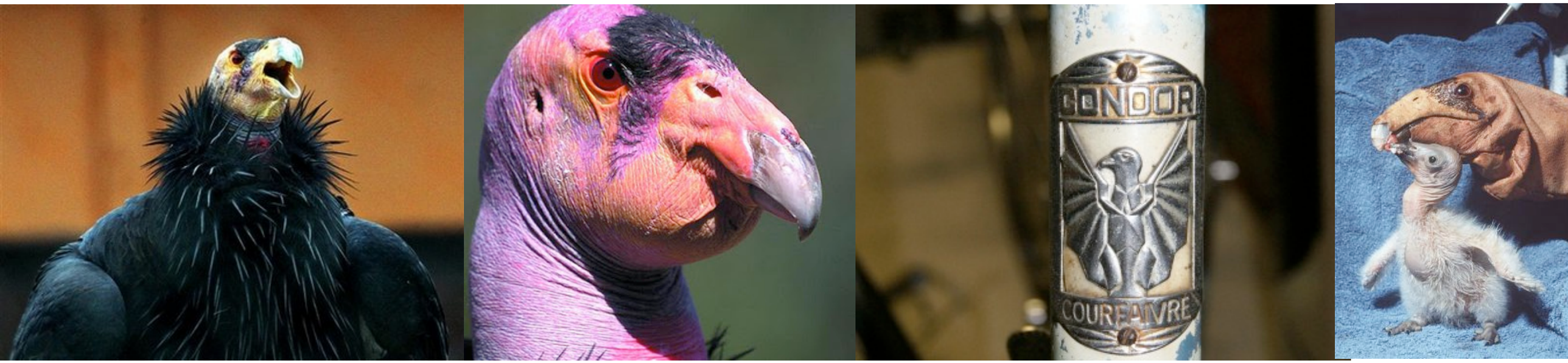


image credits from left to right: AP Photo/Rick Bowmer via Yahoo! News, http://www.weirdtwist.com/2012/12/8-butt-ugly-animals.html, http://www.swissbicycles.com/condor/allegro-tube-example-post/, http://en.wikipedia.org/wiki/File:Gymnogyps_californianus1.jpg

by Leo Singer

LIGO-G1300396-v8

quick reference:

MacPorts

`http://macports.org`

largest collection of open-source software ported to Mac OS (16982 ports)

now features an HTCondor package with built-in personal Condor pool

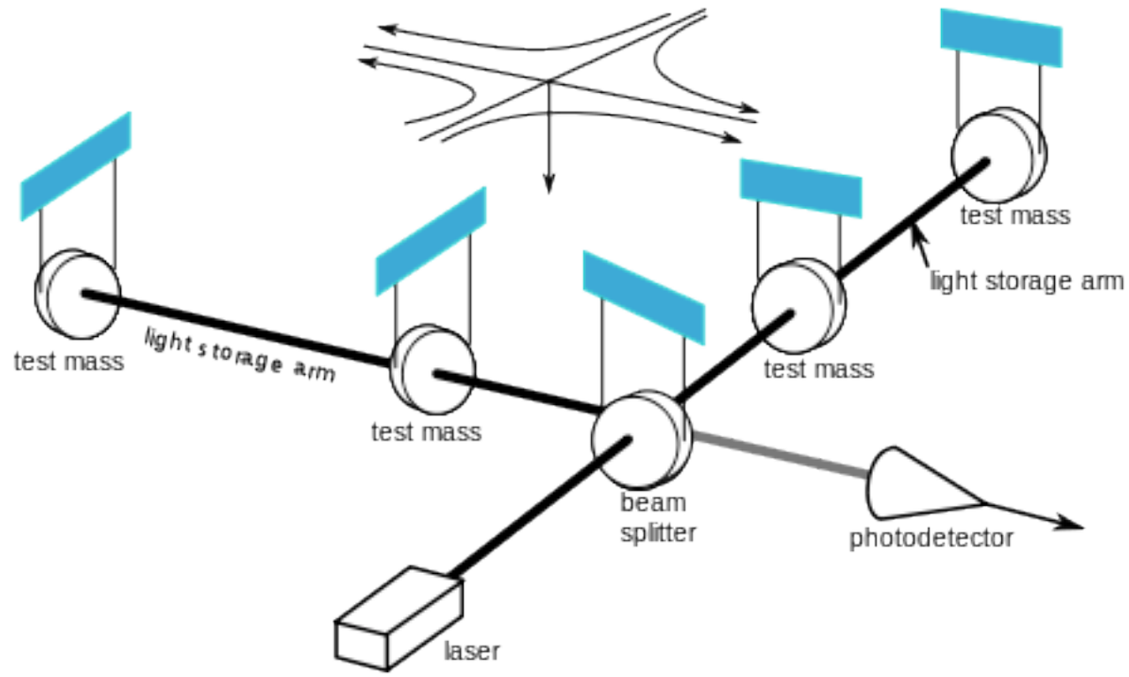
```
$ port install htcondor
$ port load htcondor
$ port unload htcondor
```

Outline

- My day job
  LIGO: Laser Interferometric Gravitational-wave Observatory
- Why HTCondor in MacPorts?
  intended user base: scientists who use HTCondor clusters
- Example project
  with personal HTCondor pool

My day job

Image credit: http://writescience.wordpress.com/2012/11/04/knowing-something-about-everything/



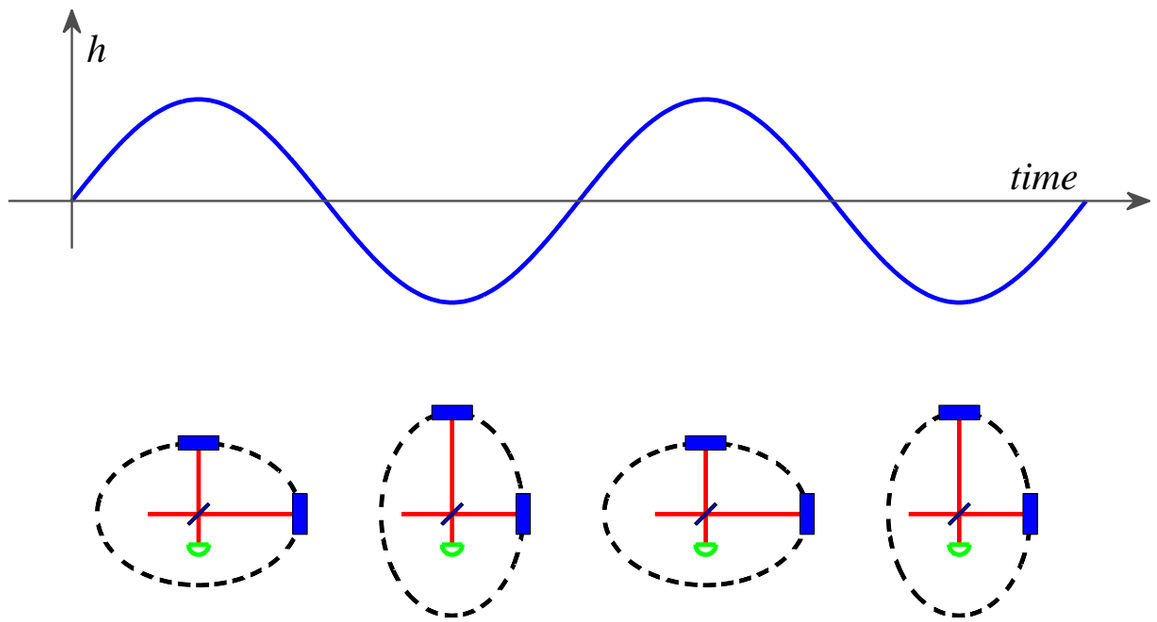


Image credits: http://en.wikipedia.org/wiki/File:Ligo.svg,
Rep. Prog. Phys. 72 (2009) 076901

What does LIGO use HTCondor for?

Matched filter banks, online & offline searches

-

Machine learning & detector characterization

-

Markov-chain Monte Carlo parameter estimation

-

Time-frequency analysis

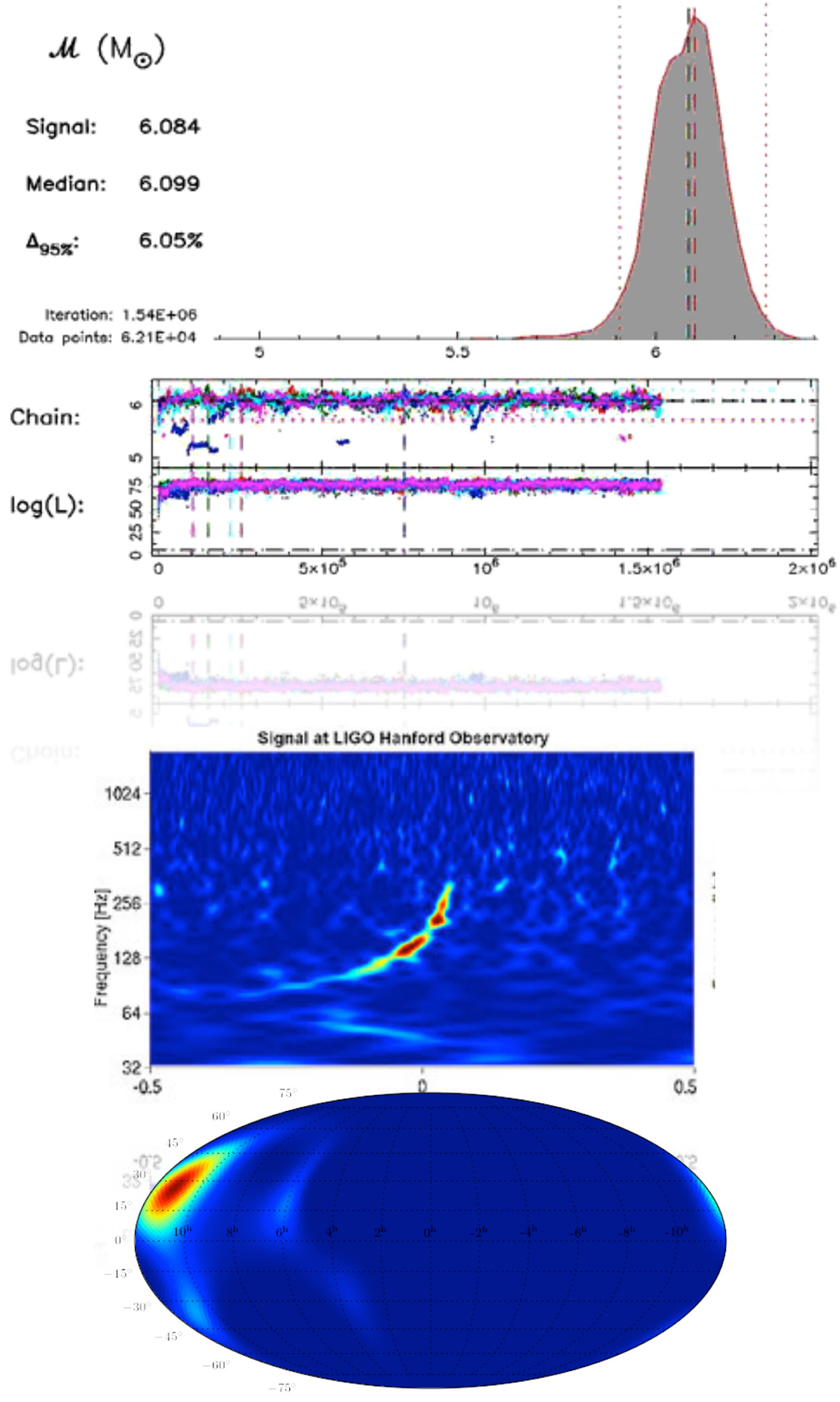


Image credits: Vivien Raymond, http://www.ligo.caltech.edu/~vraymond/, LIGO-Virgo blind injection, http://www.ligo.org/news/blind-injection.php, unpublished graphic, Leo Singer

Why HTCondor in MacPorts?

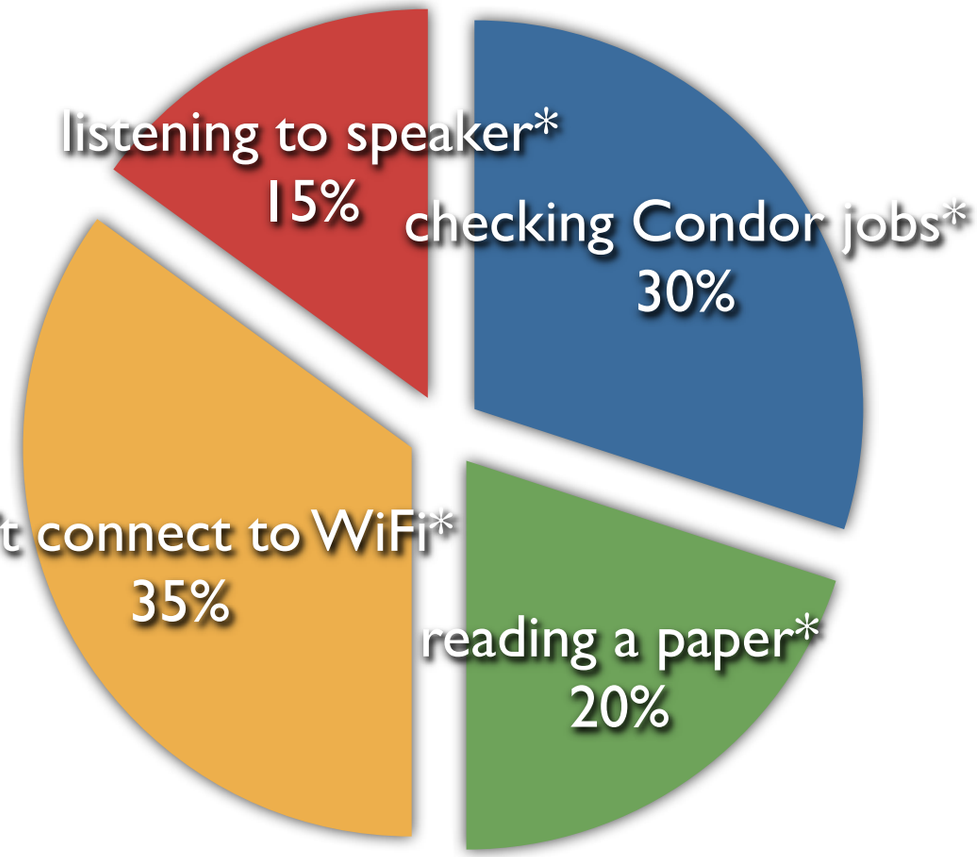
A typical LIGO meeting

What do you notice about this picture?

Image credit: http://aquarius.elte.hu/lv0909/photos.shtml

A typical LIGO meeting

What do you notice about this picture?



- Most people have Macs
- Most people are trying to check on their Condor jobs*

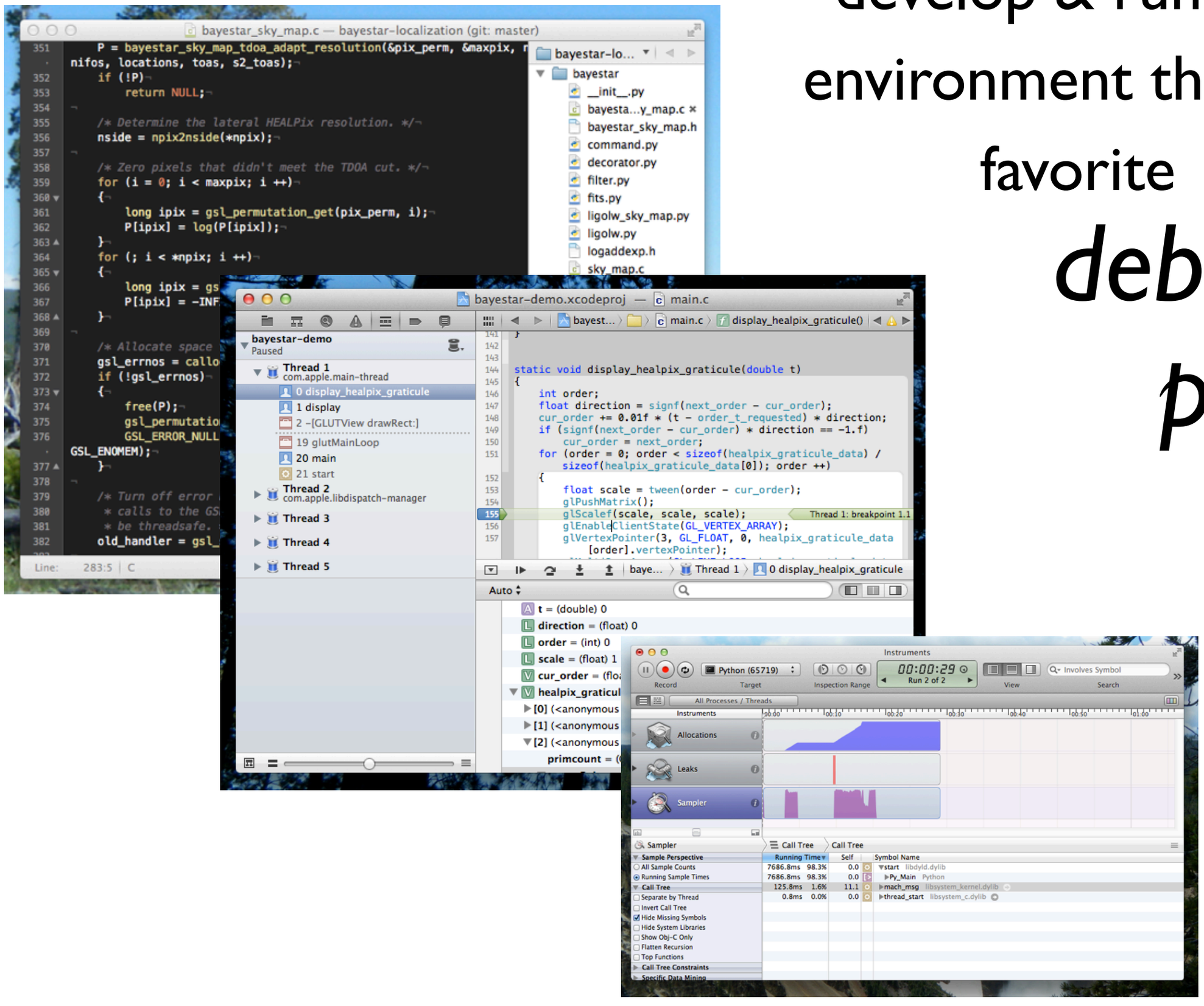
** note: this data is fabricated*

Why would I want to turn my laptop into an HTCondor pool?

ATLAS cluster, AEI–Hannover Massimo Fiorito

Whoa, there. You know, I'm not sure how *I* feel about this.

develop & run analyses in environment that has your favorite *editors, debuggers, profiling tools...*



Why HTCondor in MacPorts?

- "Personal Condor" configuration
  - Test your Condor workflows in the comfort of your own laptop
  - Even in an airplane, without network access
  - No competition for CPU time while testing
  - Test on fast local filesystem (no laggy NFS)

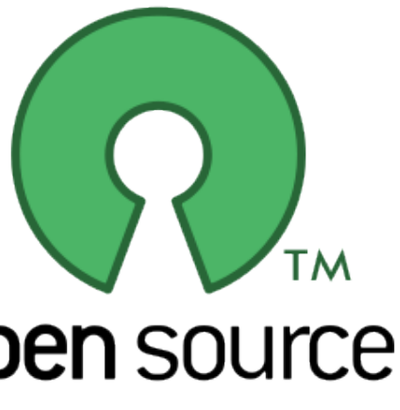


Image credits: Launchpad, https://launchpad.net
Fedora project, https://fedoraproject.org
MacPorts project, http://www.macports.org
Debian project, http://www.debian.org

Getting MacPorts

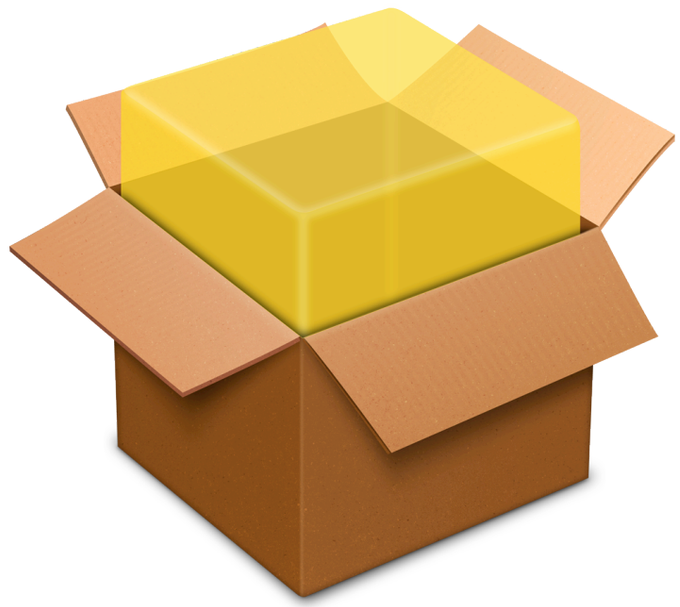
MacPorts is great.

It builds tens of thousands of open-source packages...

But it's definitely for 'power users.'

Step 1: Install Xcode (from App Store, on Lion+).

Step 2: Install MacPorts dmg from http://macports.org.



viz.: **you need a compiler to use MacPorts**, so MacPorts is not *necessarily* a substitute for standalone HTCondor binaries.

The port

```
$ port info htcondor
htcondor @7.8.8 (science, parallel, net)
Variants:             debug, [+]personal, universal

Description:          HTCondor is a specialized workload management system for
                      compute-intensive jobs. Like other full-featured batch
                      systems, HTCondor provides a job queueing mechanism,
                      scheduling policy, priority scheme, resource monitoring,
                      and resource management. Users submit their serial or
                      parallel jobs to HTCondor, HTCondor places them into a
                      queue, chooses when and where to run the jobs based upon a
                      policy, carefully monitors their progress, and ultimately
                      informs the user upon completion.
Homepage:             http://research.cs.wisc.edu/htcondor

Build Dependencies:   cmake, latex2html
Library Dependencies: boost, expat, kerberos5, openssl, pcre
Platforms:            darwin
License:              apache
Maintainers:          aronnax@macports.org
```

To install:

```
$ sudo port install htcondor
```

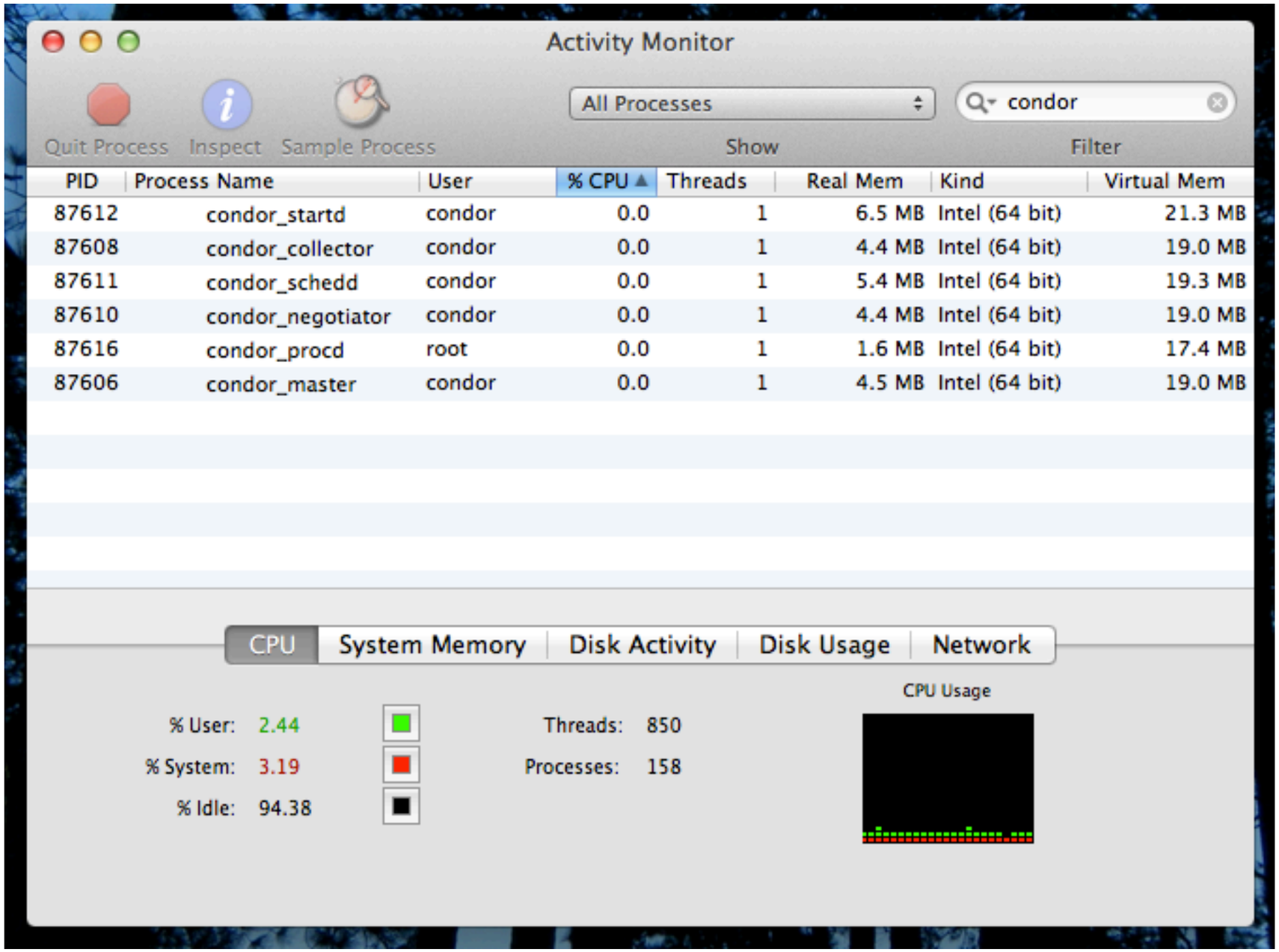
To start your Condor pool:

```
$ sudo port load htcondor
```

To stop your Condor pool:

```
$ sudo port unload htcondor
```

What it looks like



What it looks like

```
$ condor_status

Name                OpSys      Arch   State     Activity LoadAv Mem   ActvtyTime

slot1@gwave-125.li OSX        X86_64 Unclaimed Idle     1.000  1024  0+00:05:04
slot2@gwave-125.li OSX        X86_64 Unclaimed Idle     0.290  1024  0+00:05:05
slot3@gwave-125.li OSX        X86_64 Unclaimed Idle     0.000  1024  0+00:05:06
slot4@gwave-125.li OSX        X86_64 Unclaimed Idle     0.000  1024  0+00:05:07
slot5@gwave-125.li OSX        X86_64 Unclaimed Idle     0.000  1024  0+00:05:08
slot6@gwave-125.li OSX        X86_64 Unclaimed Idle     0.000  1024  0+00:05:09
slot7@gwave-125.li OSX        X86_64 Unclaimed Idle     0.000  1024  0+00:05:10
slot8@gwave-125.li OSX        X86_64 Unclaimed Idle     0.000  1024  0+00:04:38
                     Total Owner Claimed Unclaimed Matched Preempting Backfill

           X86_64/OSX     8     0       0         8       0          0        0

                Total     8     0       0         8       0          0        0
```

Caveats

- No ‘standard’ universe; Condor only supports it on Linux
- Built without Globus support; Globus is not in MacPorts (yet...)
- MacPorts buildbot has pre-built HTCondor for OS X Snow Leopard (10.6) and onward
  (officially, MacPorts supports 3 most recent Mac OS releases; legacy Leopard and Tiger support on best-effort basis)
- Tested only on Mountain Lion (10.8)

Relatively challenging port

- Could not use official source tarballs
  can't download them anonymously; had to use GitHub tags instead
- Difficult livecheck (automatic upstream version discovery)
  ~10k tags in GitHub; default GitHub livecheck fails due to pagination
- Had to fix broken runpath for system libraries
  build expects symlinks/copies of dependencies in `$prefix/lib/condor` (to avoid stomping on host OS' copies?), patched helper script `macosx_rewrite_libs` to be a noop
- Patched to look for config files in install prefix
  look for `condor_config` in `$prefix/etc`, not `/etc`
- Manpages not part of default build target
  reverse-engineered how to get CMake to generate & install them

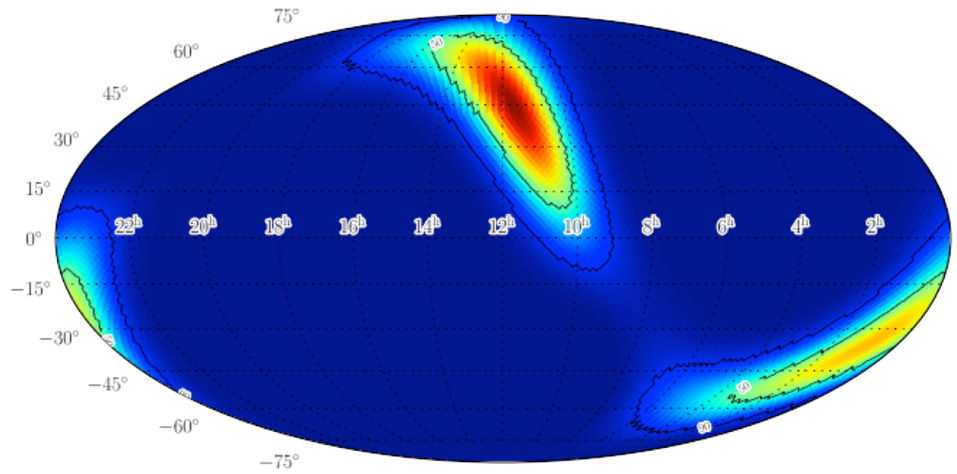
Relatively challenging port (continued)

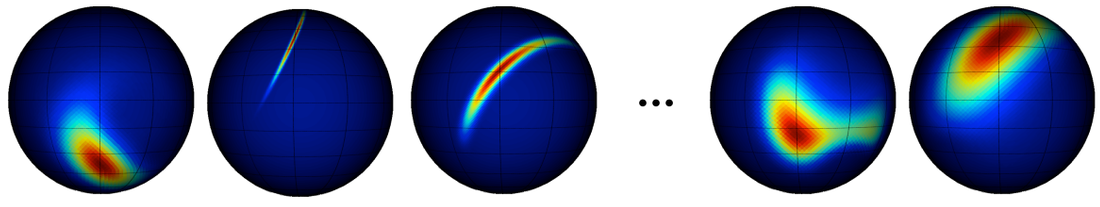
- Disabled `condor_configure`, `condor_install`
  hard to test, also might encourage user w/ sudo privileges to do silly things
- Personal HTCondor pool and IP address changes
  the hardest part: getting HTCondor pool to survive changing wireless networks or work with no internet connection at all—due to FQDN checks
- Idea: put working, zero-configuration personal HTCondor pool in upstream as an example config?
  ...so that it becomes a part of the Debian, Fedora packages as well

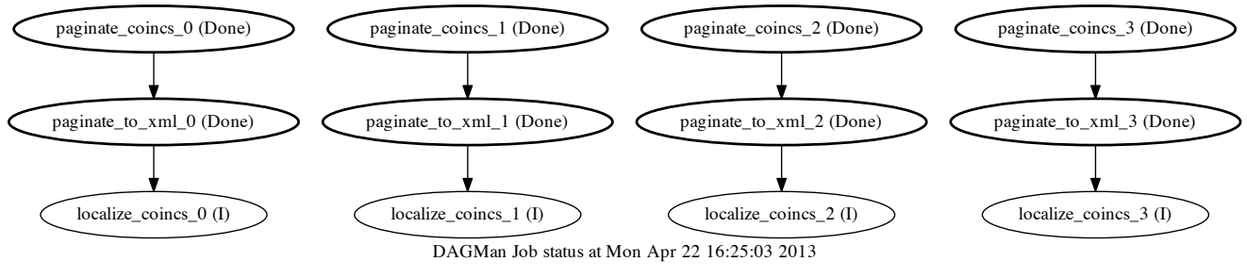
Example project

from my day job, with personal HTCondor pool

- Rapid sky localization for LIGO triggers
- Inject simulated signals into artificial detector noise
- Produce prob. sky maps for all detected injections, in batches of ~100 events
- Post-process, determine sky localization accuracy & study self-consistency of prob. contours



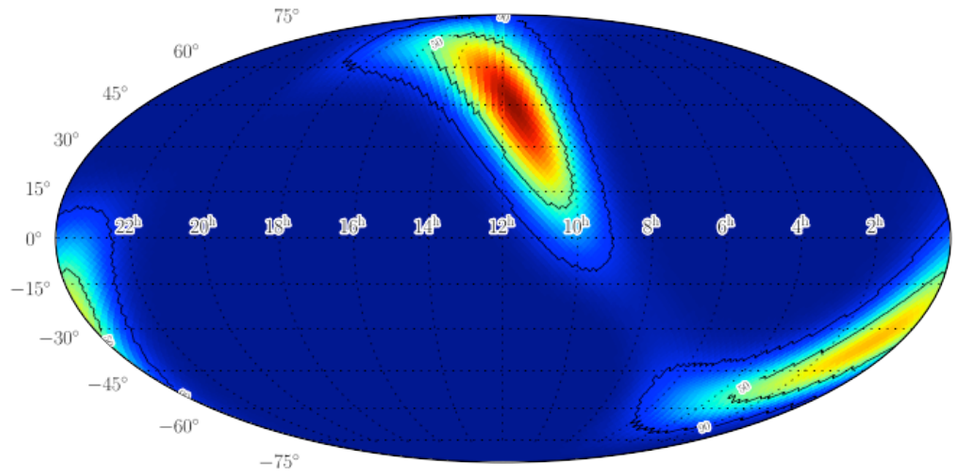


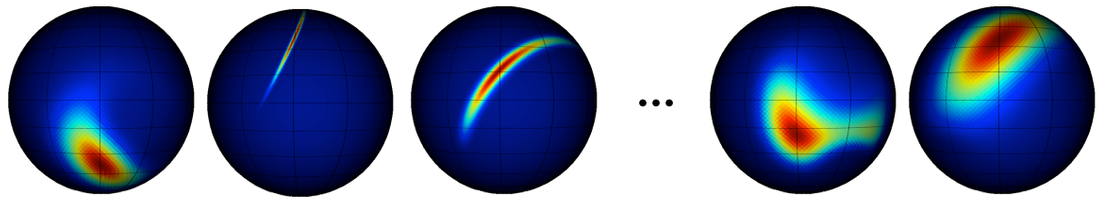


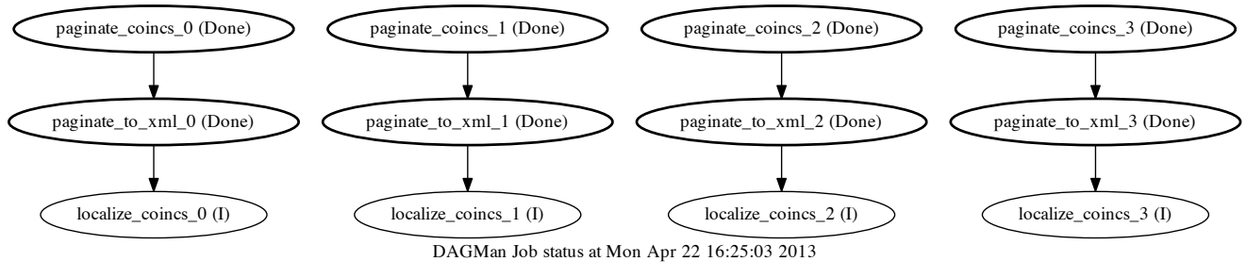
Example project

with personal HTCondor pool

- Cluster head nodes were down for maintenance!
- Wrote DAG on my laptop, analyzed a smaller dataset
- What a breeze! no waiting for CPU time, no slow NFS filesystem, debugged code *in situ*...
- Later ran full workload, O(50) times larger, on cluster







My role

- I'm not an HTCondor expert.
  I'm just a physics grad student. I'm also a MacPorts volunteer. I am happy to maintain the port, but welcome others to contribute.

- Maintenance plan: track stable HTCondor releases
  as soon as I notice them with `port livecheck`, or I am prompted to on the MacPorts mailing list or in a MacPorts update request ticket

- Check out the port and/or get involved!
  Give me feedback, or even volunteer as a MacPorts co-maintainer if you *are* an HTCondor expert.

Thank you!

Acknowledgements

- My Ph.D. advisor, Alan Weinstein
- The NSF Graduate Research Fellowship
- The LIGO Project
- The MacPorts Project
- The HTCondor Team