

Supporting Experiment Workflows in GENI

GENI Experiment Workflow and
Services Working Group

Jeannie Albrecht
Williams College

Revisiting Experiment Workflow

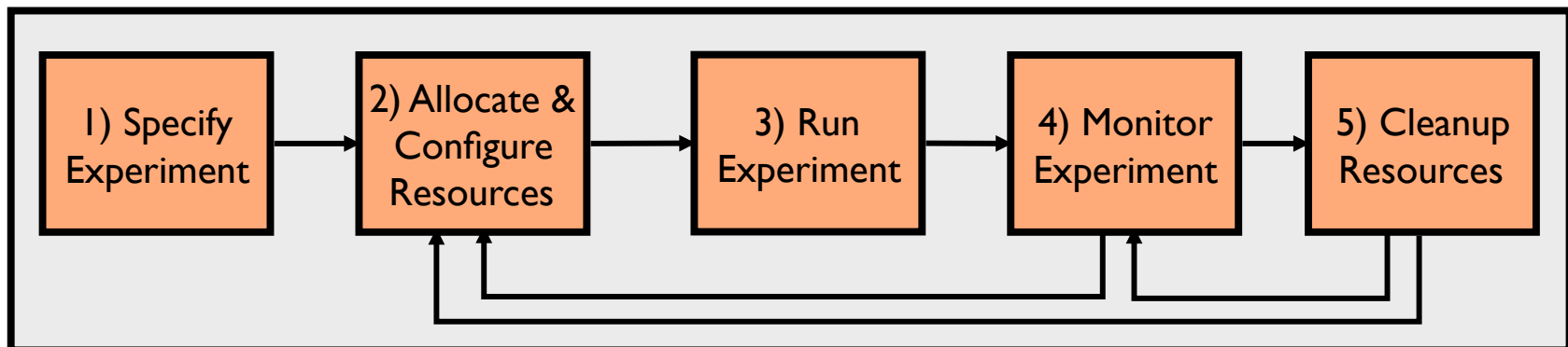
- One of the key purposes of this working group is to focus on “Experiment Workflow”
- How can we make it “easy” for a researcher or user to run an experiment on GENI?
- Solutions must address tasks associated with slice management, resource discovery and configuration, and experiment control (some potential overlap with other WGs)

PlanetLab Experience

- PlanetLab taught us many things about experiment control and maintenance
 - For novice researchers, it's often not “easy”
- Several challenges must be overcome in wide-area distributed environments:
 - Acquiring and configuring “optimal” resources
 - Managing a long-running experiment or service
 - Detecting and recovering from failures
 - Monitoring experiments in real-time
- Many researchers resort to brittle, custom-built, application-specific scripts to accomplish these tasks

Things to Consider

- Experiment control requirements:
 - Extensible experiment specifications
 - Resource discovery and slice configuration
 - Support for potentially multi-phased execution
 - Support for experiment composition
 - Experiment maintenance and monitoring



General Design Goals

- **Extensibility**
 - Define a general set of APIs for controlling experiments
 - Novice vs. experienced researchers
 - Long-running services vs. Short-lived experiments
 - Support evolution
- **Robustness**
 - Resources will fail
 - Researchers will run buggy/broken code
- **Scalability**
 - Experiments may run on thousands of resources (or more)
- **Usability**
 - We need to make GENI “easy to use”

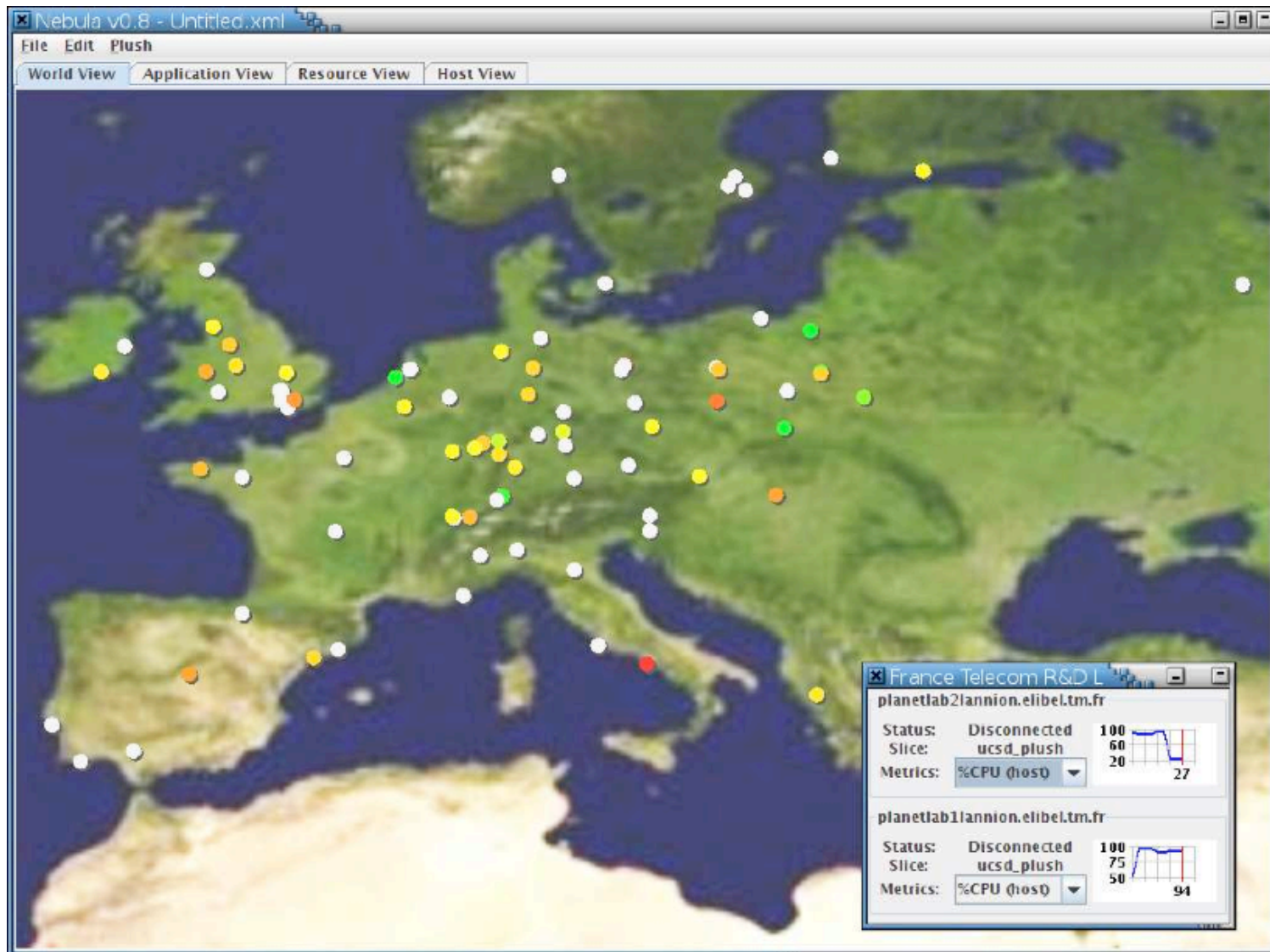
My Goals

- I teach at Williams College
 - Top ranked liberal arts college
 - CS students are bright, ambitious, and like to tinker
- Distributed systems courses are often not taught to undergrads (why?)
- Goal 1: Use GENI as a learning laboratory
 - Help my students appreciate, understand, and experiment with real distributed systems
 - Bring tech-richness of big universities to a small college
- Goal 2: Lower the entry barrier for research in distributed systems
 - Involve undergrads!
 - Skilled undergrads → Smarter grad students and software developers

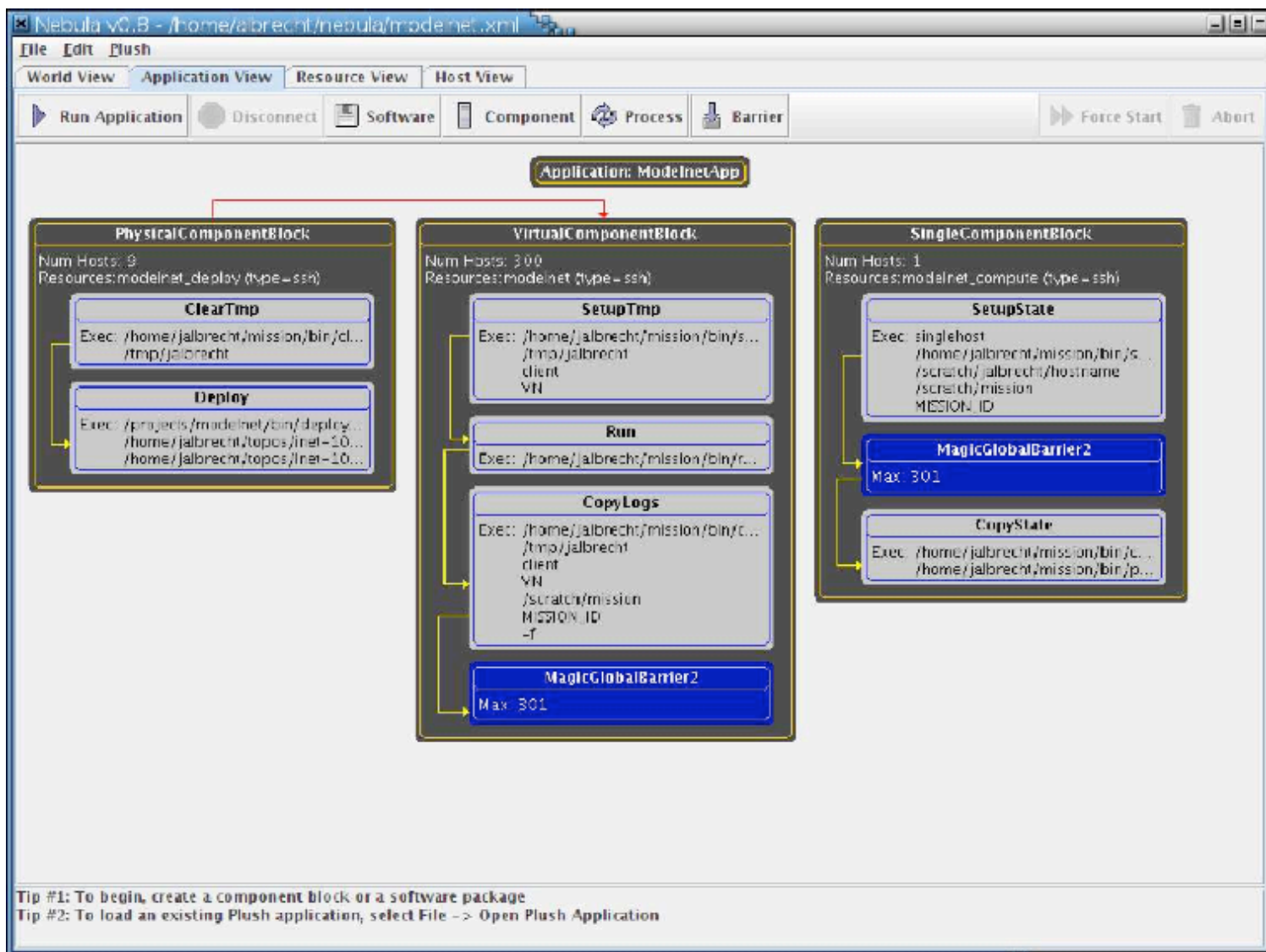
One Potential Solution: Gush

- Gush (GENI User Shell)
 - Scalable experiment control framework for deploying and maintaining GENI experiments
 - Extension of Plush, which was initially designed for PlanetLab application control
 - Uses XML for describing experiments and resources
 - Exports APIs for interacting with resource managers and measurement services
 - Supports three UIs: graphical, command-line, programmatic (currently XML-RPC)

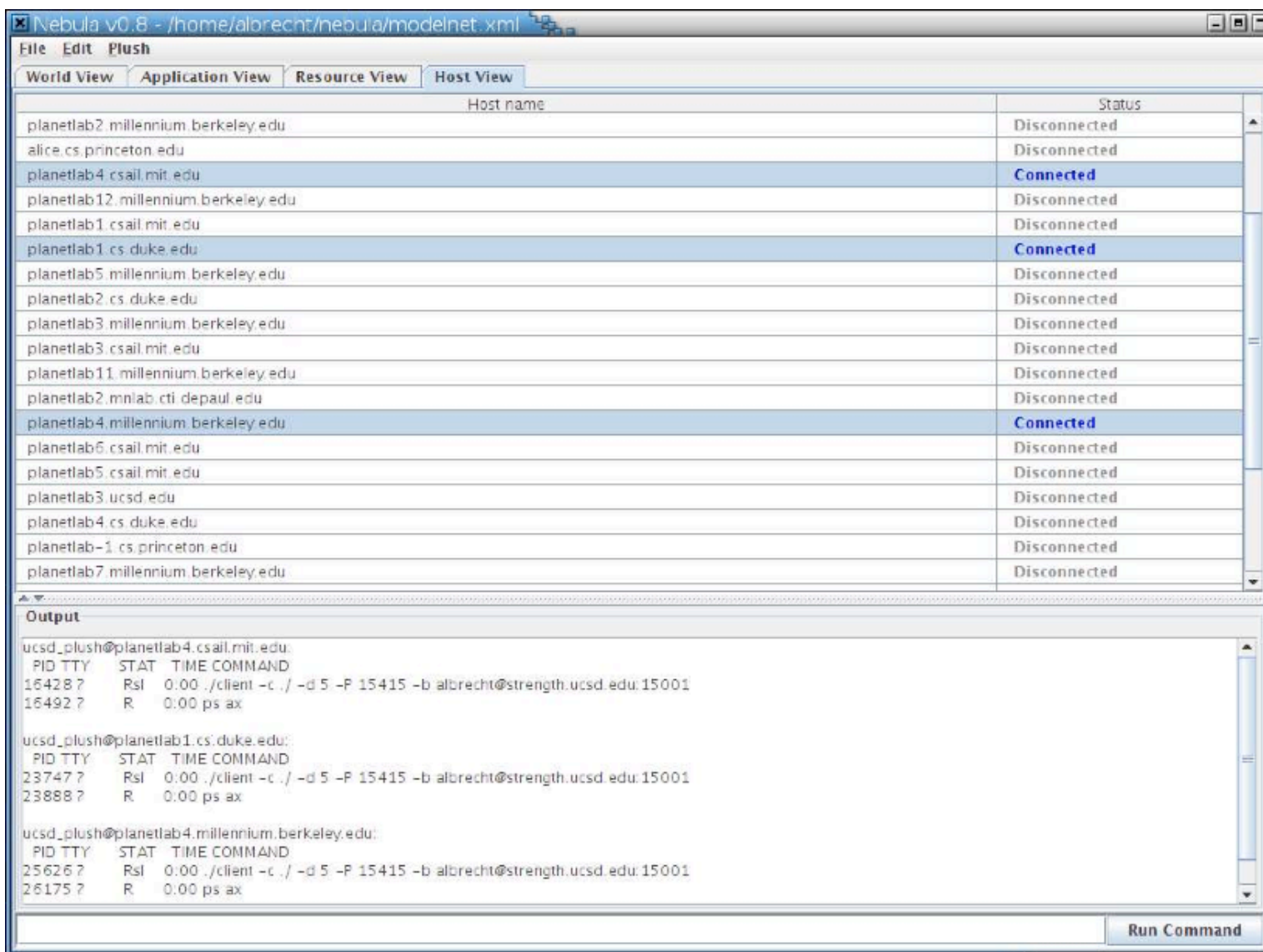
Visualizing Experiments & Resources



Describing Experiments



Issuing Commands to Resources



The screenshot shows the Nebula v0.8 interface with the 'Host View' tab selected. The interface displays a list of hosts and their status. Below the list, there is an 'Output' window showing the results of commands executed on three different hosts.

Host name	Status
planetlab2.millennium.berkeley.edu	Disconnected
alice.cs.princeton.edu	Disconnected
planetlab4.csail.mit.edu	Connected
planetlab12.millennium.berkeley.edu	Disconnected
planetlab1.csail.mit.edu	Disconnected
planetlab1.cs.duke.edu	Connected
planetlab5.millennium.berkeley.edu	Disconnected
planetlab2.cs.duke.edu	Disconnected
planetlab3.millennium.berkeley.edu	Disconnected
planetlab3.csail.mit.edu	Disconnected
planetlab11.millennium.berkeley.edu	Disconnected
planetlab2.mnlab.cti.depaul.edu	Disconnected
planetlab4.millennium.berkeley.edu	Connected
planetlab6.csail.mit.edu	Disconnected
planetlab5.csail.mit.edu	Disconnected
planetlab3.ucsd.edu	Disconnected
planetlab4.cs.duke.edu	Disconnected
planetlab-1.cs.princeton.edu	Disconnected
planetlab7.millennium.berkeley.edu	Disconnected

Output

```
ucsd_plush@planetlab4.csail.mit.edu:
PID TTY STAT TIME COMMAND
16428 ? Rsl 0:00 ./client -c ./ -d 5 -P 15415 -b albrecht@strength.ucsd.edu:15001
16492 ? R 0:00 ps ax

ucsd_plush@planetlab1.cs.duke.edu:
PID TTY STAT TIME COMMAND
23747 ? Rsl 0:00 ./client -c ./ -d 5 -P 15415 -b albrecht@strength.ucsd.edu:15001
23888 ? R 0:00 ps ax

ucsd_plush@planetlab4.millennium.berkeley.edu:
PID TTY STAT TIME COMMAND
25626 ? Rsl 0:00 ./client -c ./ -d 5 -P 15415 -b albrecht@strength.ucsd.edu:15001
26175 ? R 0:00 ps ax
```

Run Command

Connecting to Resources

```
Nebula v0.8 - /home/albrecht/plush/src/tests/planetlab.xml
File Edit Plush
World View Application View Resource View Host View SSH:planetlab1.cs.duke.edu x
Connecting to planetlab1.cs.duke.edu...
[ucsd_plush@planetlab1 ~]$ ls -lat
total 12684
6702473 drwx----- 3 ucsd_plush slices 4096 may 25 21:30 .
6701511 drwxr-xr-x 3 root root 4096 may 11 06:25 ..
6702476 -rw-r--r-- 1 ucsd_plush slices 24 may 11 06:25 .bash_logout
6702474 -rw-r--r-- 1 ucsd_plush slices 191 may 11 06:25 .bash_profile
6702475 -rw-r--r-- 1 ucsd_plush slices 124 may 11 06:25 .bashrc
6702480 -rwxr-xr-x 1 ucsd_plush slices 4764 may 11 06:48 bootstrap.pl
6702504 -rwxr-xr-x 1 ucsd_plush slices 6376124 may 23 03:57 client
6702499 lrwxrwxrwx 1 ucsd_plush slices 41 may 25 21:30 client.txt -> ./logfile-planetlab1-15415-1180128637.txt
6702484 drwxr--r-- 3 ucsd_plush slices 4096 may 8 21:09 helper-scripts
6702483 -rwxr-xr-x 1 ucsd_plush slices 670 Dec 16 00:14 install_yum.sh
6702500 -rw-r--r-- 1 ucsd_plush slices 18541 may 14 21:54 logfile-planetlab1-15415-1179179445.txt
6702523 -rw-r--r-- 1 ucsd_plush slices 41382 may 25 21:31 logfile-planetlab1-15415-1180128637.txt
6702481 -rw-r--r-- 1 ucsd_plush slices 6471680 may 10 23:50 plush-i386-Linux.tar
6702501 lrwxrwxrwx 1 ucsd_plush slices 35 may 25 21:30 plush-logfile.txt -> ./plush-logfile15415-1180128637.txt
6702502 -rw-r--r-- 1 ucsd_plush slices 281 may 14 21:50 plush-logfile15415-1179179445.txt
6702524 -rw-r--r-- 1 ucsd_plush slices 468 may 25 21:31 plush-logfile15415-1180128637.txt
6702482 -rwxr--r-- 1 ucsd_plush slices 241 may 17 22:34 plush.prefs
[ucsd_plush@planetlab1 ~]$ ps auxww
USER      PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
814        5147  0.3  0.1  2304  1288 pts/2    Ss   21:32   0:00 /bin/bash -l
814        5223  0.0  0.0  2404   816 pts/2    R+   21:32   0:00 ps auxww
[ucsd_plush@planetlab1 ~]$ traceroute www.google.com
traceroute: Warning: www.google.com has multiple addresses; using 64.233.169.103
traceroute to www.1.google.com (64.233.169.103), 30 hops max, 38 byte packets
 1  152.3.138.61 (152.3.138.61)  0.722 ms  0.275 ms  0.231 ms
 2  152.3.219.69 (152.3.219.69)  0.390 ms  1.236 ms  0.282 ms
 3  r1gh7600-gw-to-duke7600-gw.ncnren.net (128.109.70.17)  1.386 ms  1.146 ms  1.110 ms
 4  r1gh1-gw-to-r1gh7600-gw.ncnren.net (128.109.70.37)  1.116 ms  1.111 ms  1.101 ms
 5  rtp1-gw-to-rpop-oc48.ncnren.net (128.109.52.6)  3.772 ms  1.526 ms  1.466 ms
 6  * * *
 7  six.onl.atl.google.com (198.32.132.41)  16.357 ms  18.578 ms  16.509 ms
 8  64.233.174.84 (64.233.174.84)  22.702 ms  24.489 ms  64.233.174.86 (64.233.174.86)  15.666 ms
 9  72.14.238.96 (72.14.238.96)  22.983 ms  18.640 ms  15.421 ms
10  64.233.175.217 (64.233.175.217)  24.838 ms  24.367 ms  64.233.175.171 (64.233.175.171)  16.622 ms
11  72.14.232.25 (72.14.232.25)  19.693 ms  19.596 ms  18.790 ms
12  yo-in-f103.google.com (64.233.169.103)  16.621 ms  16.551 ms  17.289 ms
[ucsd_plush@planetlab1 ~]$ hostname
planetlab1.cs.duke.edu
[ucsd_plush@planetlab1 ~]$ whoami
ucsd_plush
[ucsd_plush@planetlab1 ~]$
```

Future Plans

- Get feedback from my students regarding usability
- Involve undergrads in development of Gush
 - Currently working with a freshman
 - Two students will work on Gush this summer (hopefully)
- Future goals:
 - Develop API for interacting with Clearinghouses to find resources on behalf of researchers
 - Need to interact with other WG services (perhaps describe resources using RSpecs)
 - Plug in external debugging, measurements, monitoring tools

Questions?