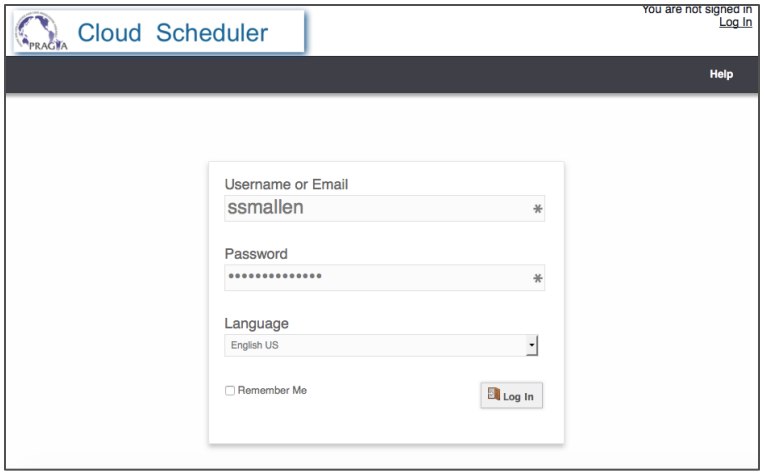
Building a persistent PRAGMA Cloud Testbed

Shava Smallen,
Nadya Williams, Serena Pan,
Weicheng Huang, Philip Papadopoulos

January 29, 2016
PRAGMA30

PRAGMA Cloud Testbed

- Biosciences and other working groups have requested a persistent Cloud Testbed
- Prototype of Lightweight Cloud Scheduler introduced in PRAGMA28 and enhanced in PRAGMA29.
- Migration from prototype to (very) early users



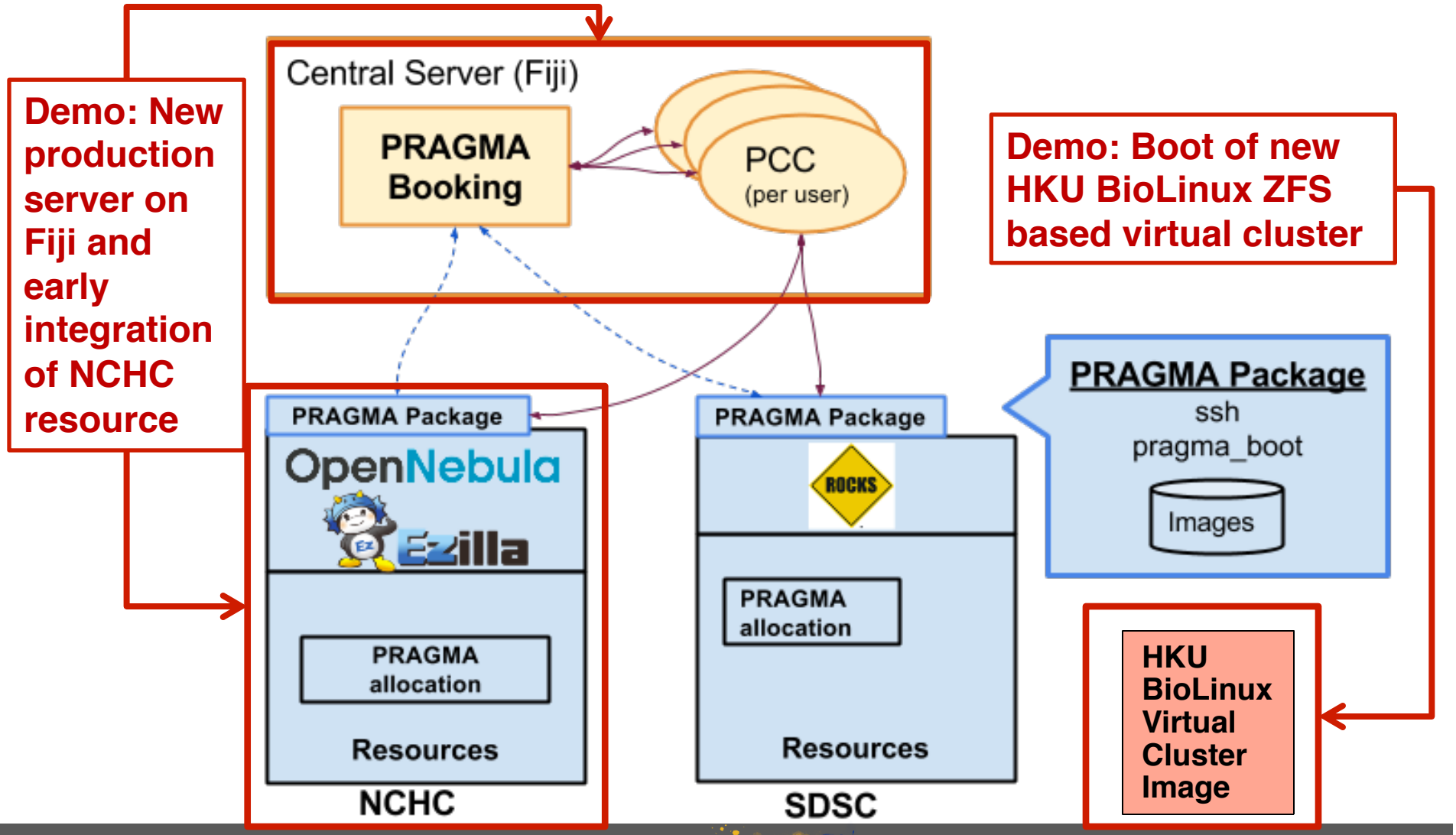
Login screen for PRAGMA lightweight scheduler

PRAGMA Cloud Scheduler

- **Goal: Low participation overhead and easy to use**
  - Sites only have to install a small package (SSH and pragma_boot) to participate
  - Users have convenient web interface to start up and manage their virtual clusters

- **Leverages the following tools:**
  - **Booked**: Open source room reservation software from Twinkle Toes
  - **pragma_boot**: Boots virtual clusters for users across PRAGMA institutions using local VM provisioner. Currently supports Rocks and OpenNebula.
  - **Personal Cloud Controller**: Manages startup, status monitoring, and shutdown of a virtual cluster. Built on top of pragma boot and HTCondor.

Architecture Overview



Building Cloud Testbed Server (Fiji)

- **Step 1**: Catalogued Fiji configuration, backed it up, and developed detailed upgrade plan
- **Step 2**: Consolidated some machines (used some nodes to upgrade memory and disk
  - 10 nodes, each node with 16 CPUs and 32 GB memory
- **Step 3**: Upgraded Fiji to Rocks 6.2 and re-configured virtual clusters
- **Step 4**: Installed cloud-scheduler, pragma_boot, and img-storage rolls

  http://fiji.rocksclusters.org/cloud-scheduler

Integrating NCHC Ezilla Resource

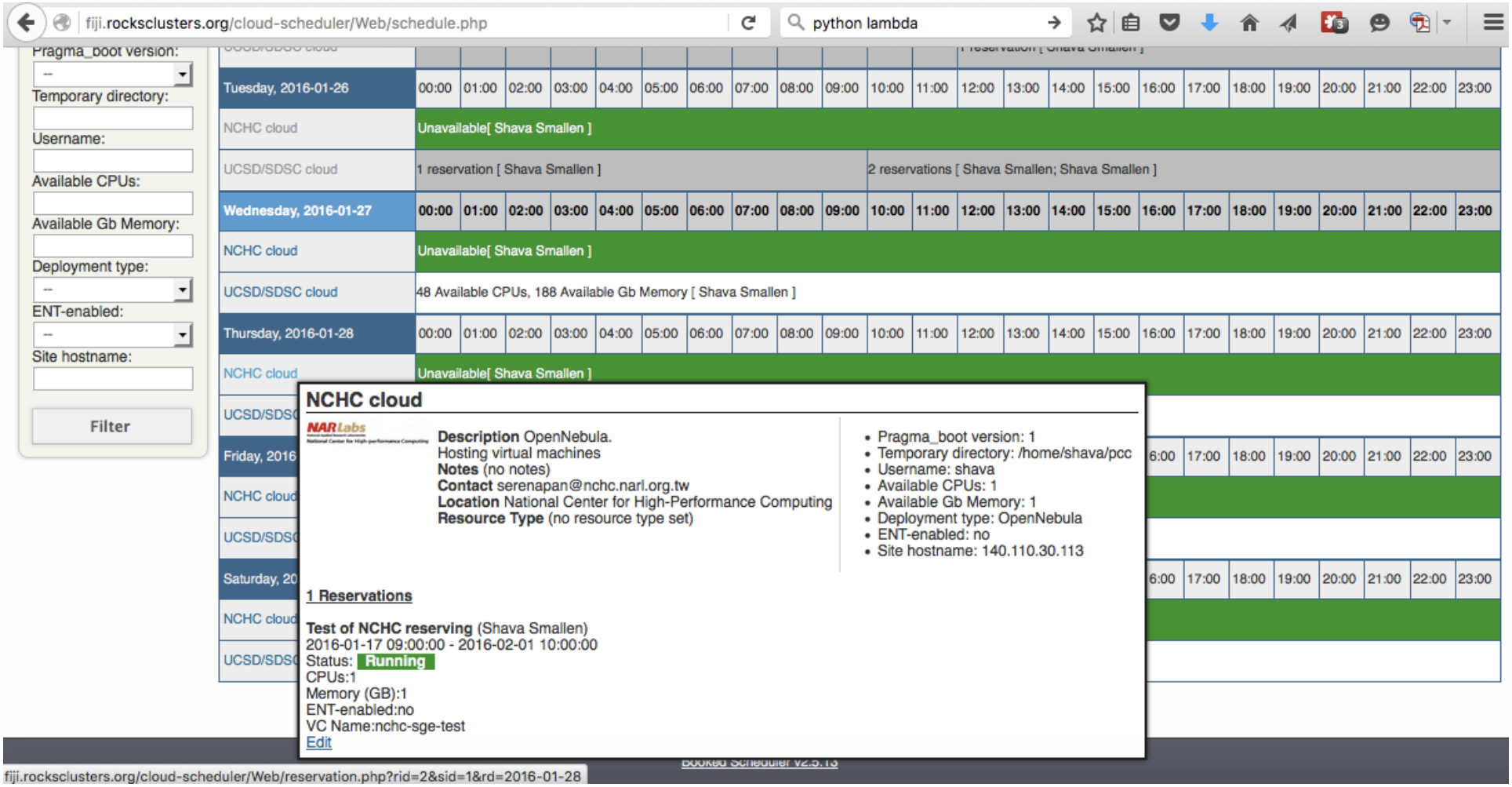
(work with Serena Pan and Weicheng Huang)

NARLabs 承諾·熱情·創新

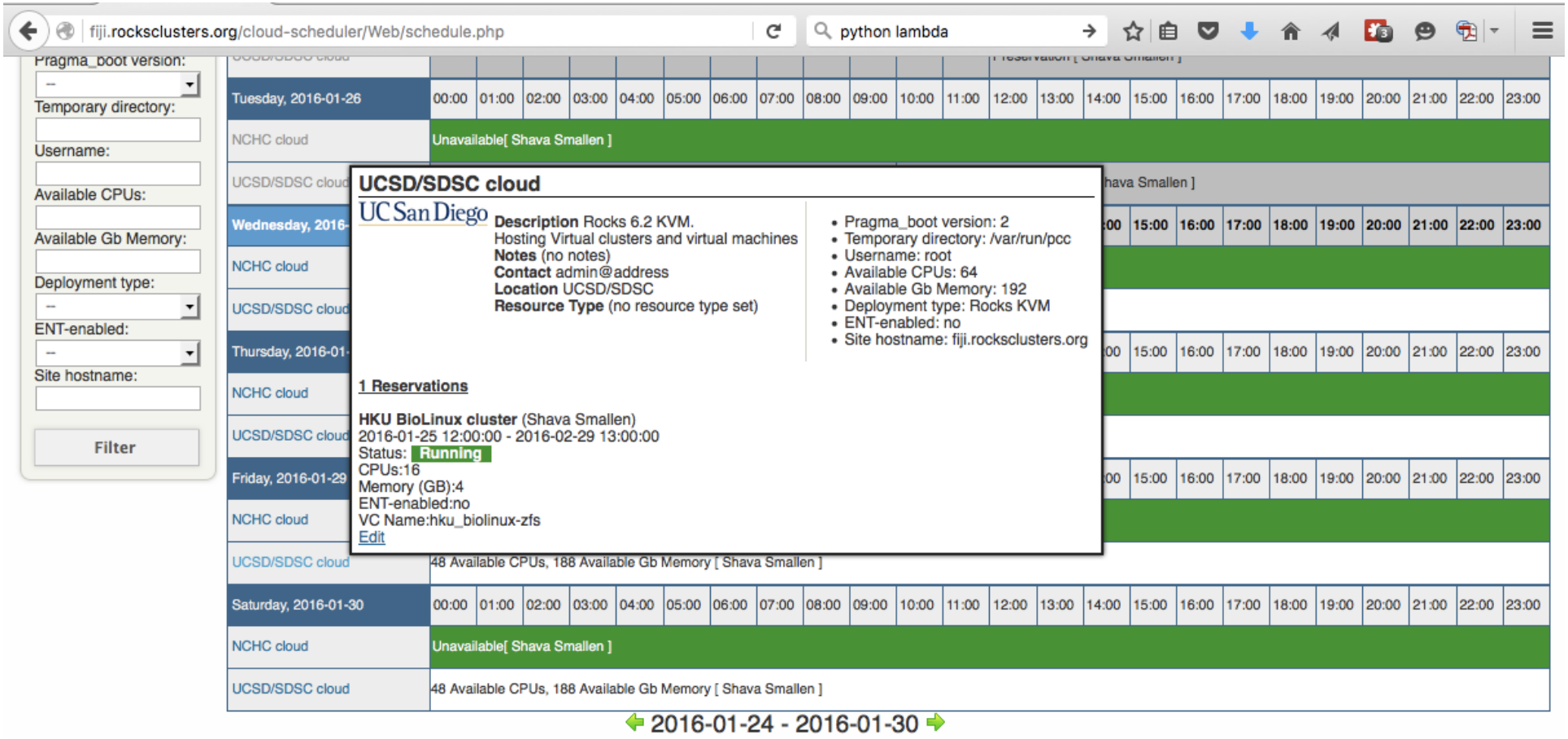
- **Step 1**: Installed pragma_boot from Github
- **Step 2**: Configured opennebula driver
- **Step 3**: Copied over virtual cluster image repository
- **Step 4**: Made bug fix to add wait for image to complete copying before generating template to launch. Still having issues with networking config.
- **Todos:**
  - Rewrite driver to leverage multi-cloud library like boto3 and integrated with new version of pragma_boot
  - Add automated shutdown and cleaning

X

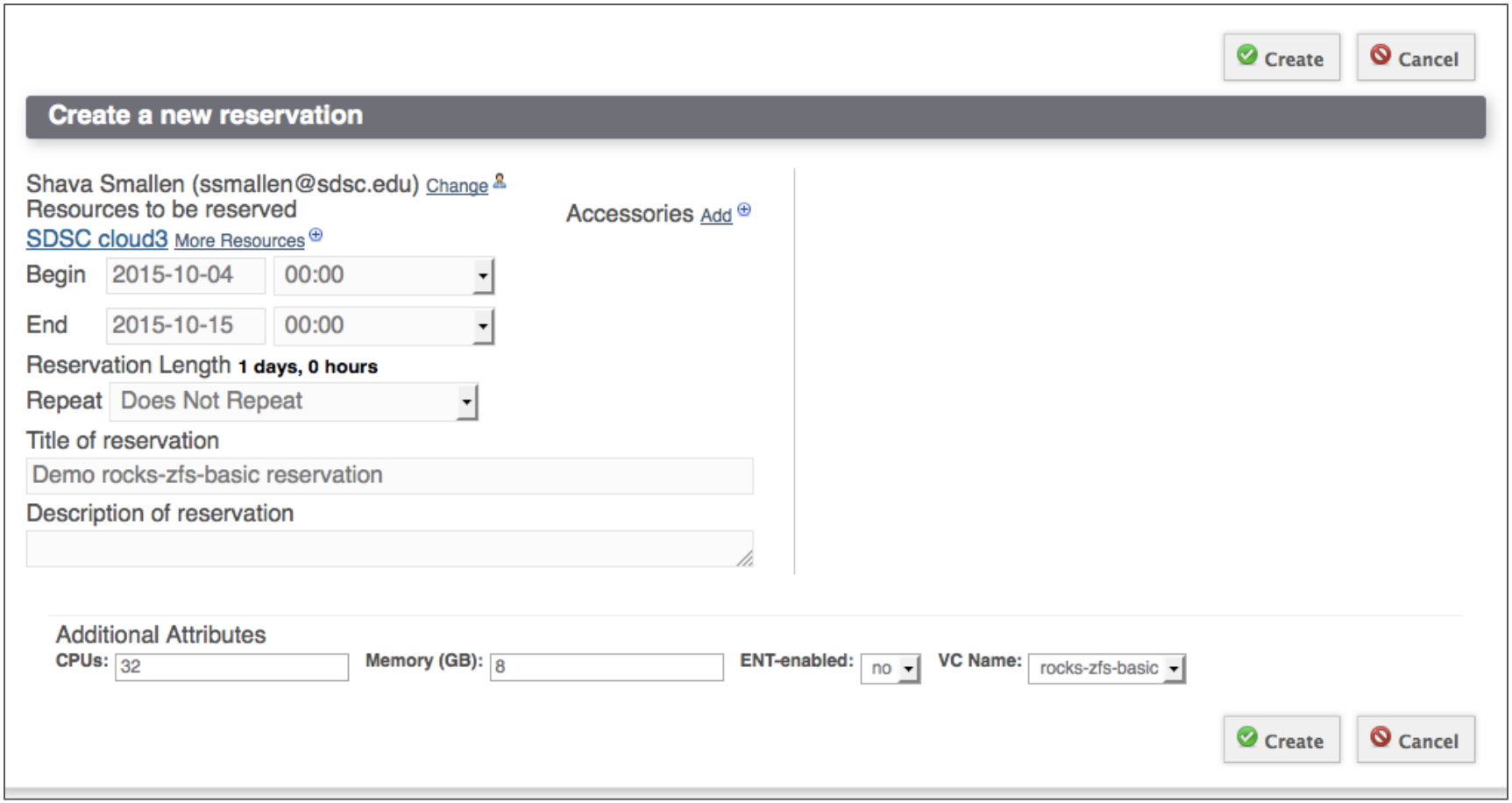
Log into Cloud Scheduler as ssmallen
Show NCHC resource



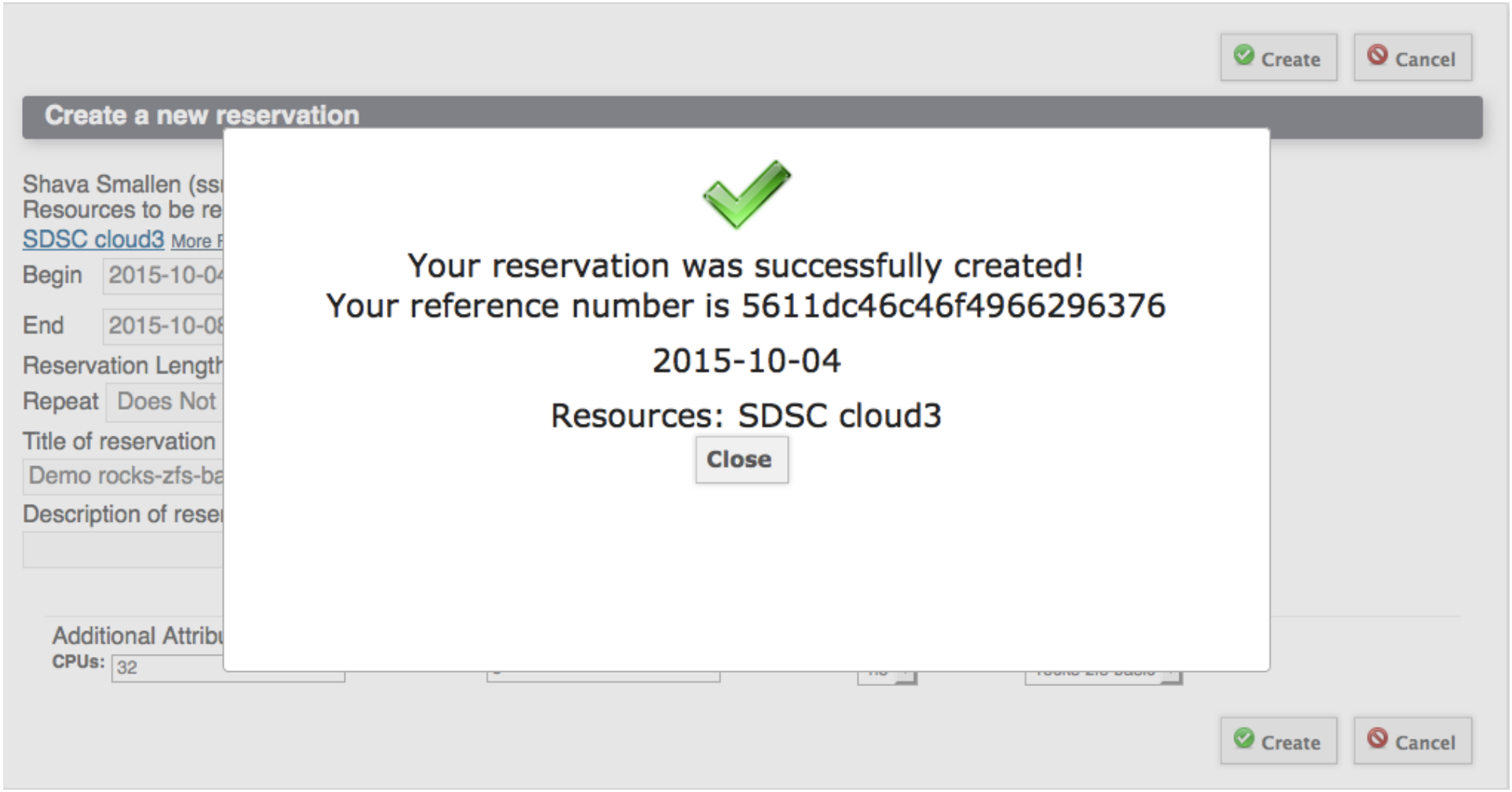
Show UCSD resource



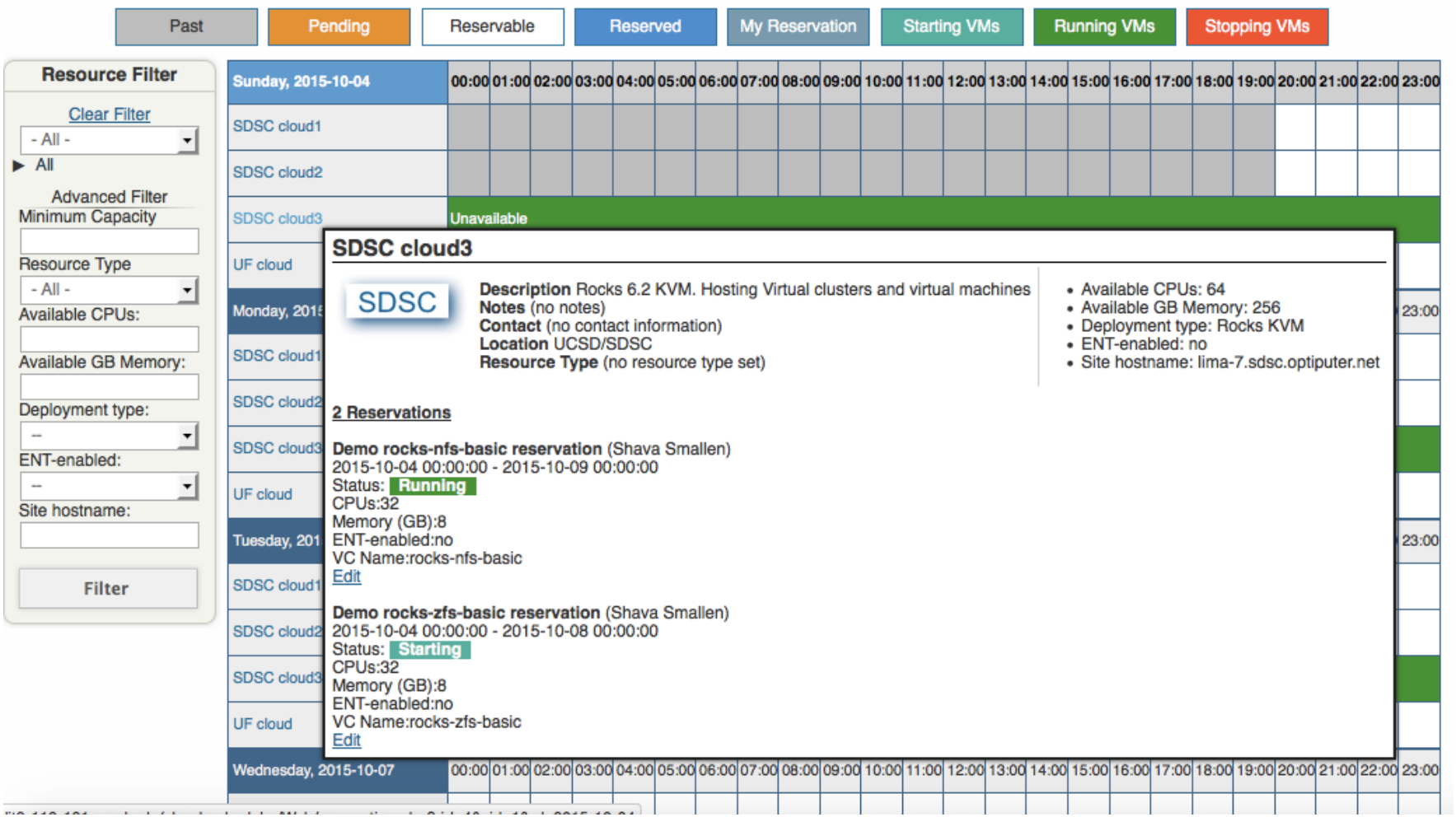
Start ZFS Reservation



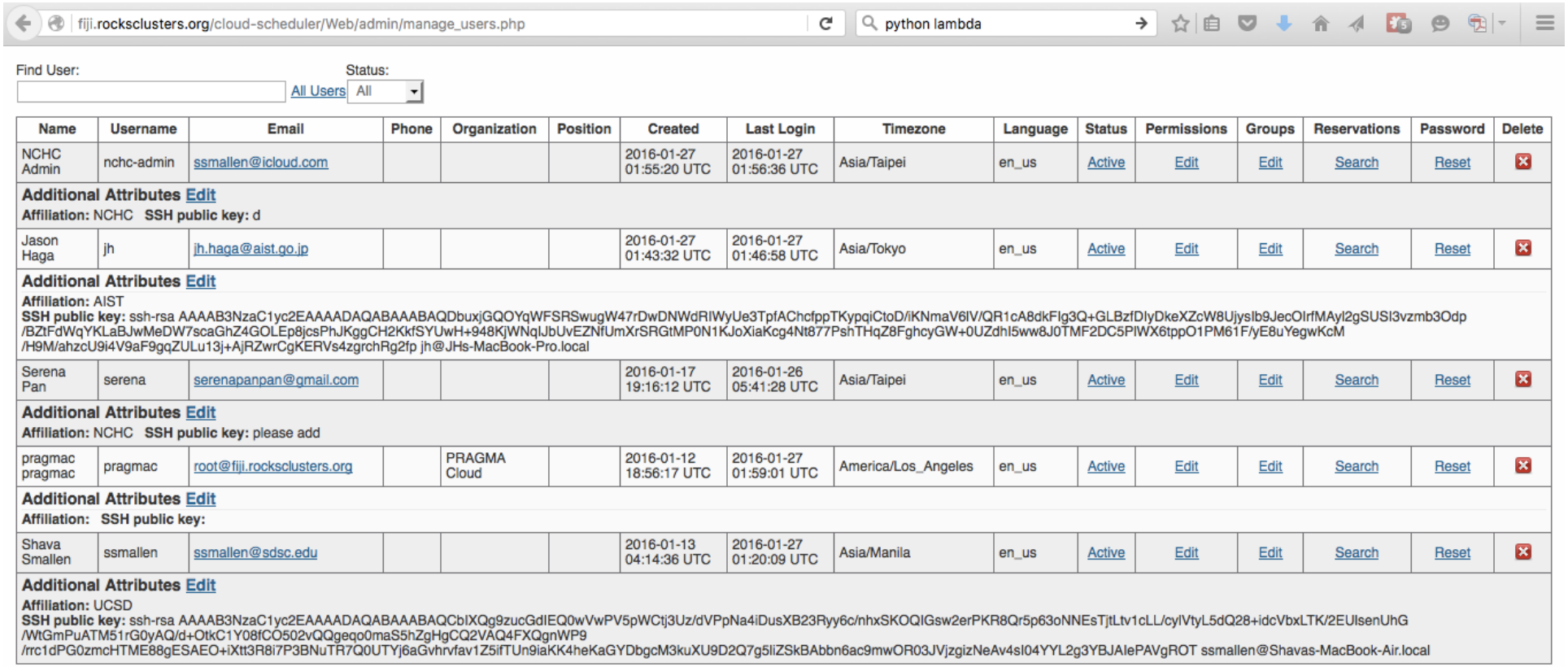
ZFS Reservation Confirmed



Confirm ZFS Reservation



Login in as pragmac
Show users



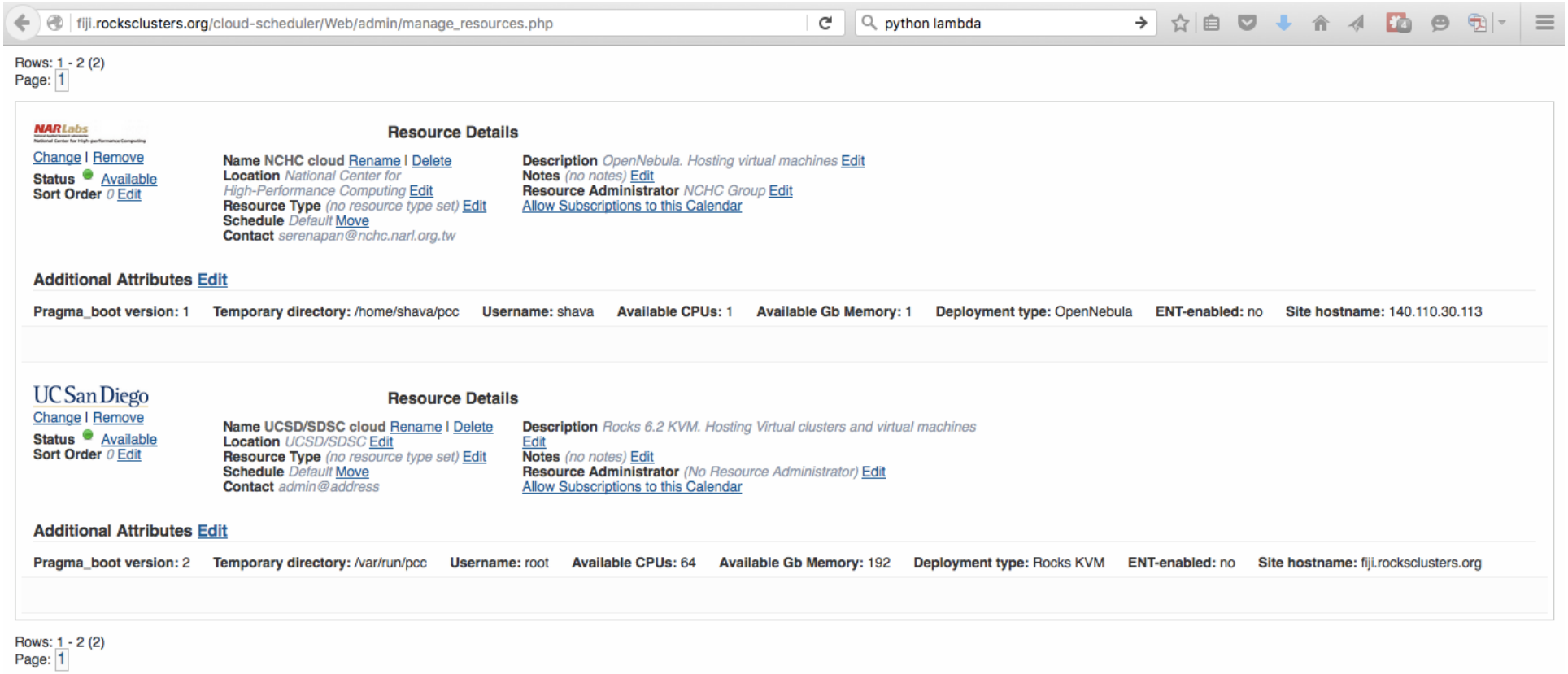
Find User: Status:

All Users All

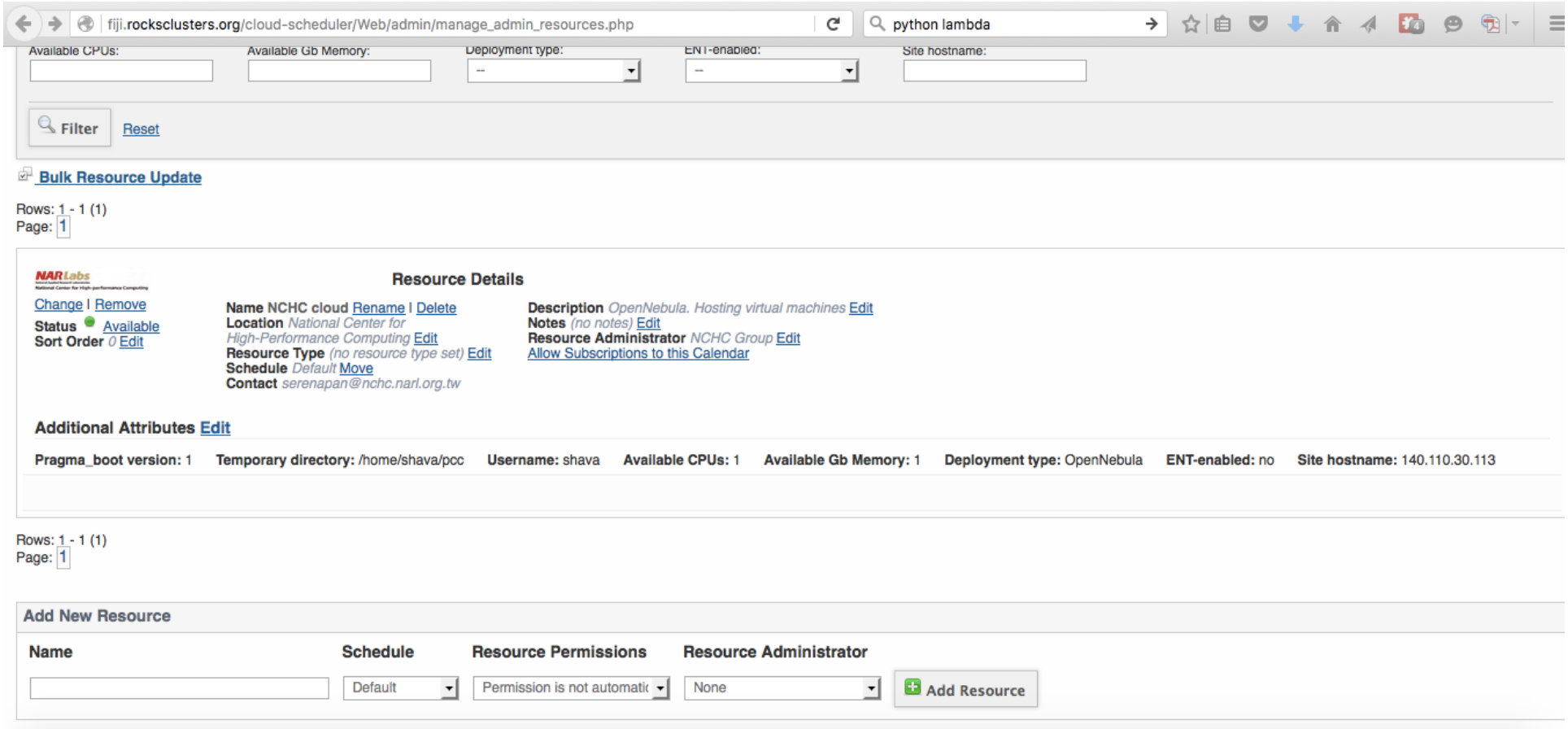
Name	Username	Email	Phone	Organization	Position	Created	Last Login	Timezone	Language	Status	Permissions	Groups	Reservations	Password	Delete
NCHC Admin	nchc-admin	ssmallen@icloud.com				2016-01-27 01:55:20 UTC	2016-01-27 01:56:36 UTC	Asia/Taipei	en_us	Active	Edit	Edit	Search	Reset	
Additional Attributes Edit **Affiliation:** NCHC **SSH public key:** d															
Jason Haga	jh	jh.haga@aist.go.jp				2016-01-27 01:43:32 UTC	2016-01-27 01:46:58 UTC	Asia/Tokyo	en_us	Active	Edit	Edit	Search	Reset	
Additional Attributes Edit **Affiliation:** AIST **SSH public key:** ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQDbuxjGQOYqWFSRSwugW47rDwDNWdRIWyUe3TpfAChcfppTKypqiCtoD/iKNmaV6IV/QR1cA8dkFlg3Q+GLBzfDlyDkeXZcW8UjysIb9JecOlrfMAyl2gSUSl3vzmb3Odp /BZtFdWqYKLaBJwMeDW7scaGhZ4GOLEp8jcsPhJKggCH2KkfSYUwH+948KjWNqlJbUvEZNfUmXrSRGtMP0N1KJoXiaKcg4Nt877PshTHqZ8FghcyGW+0UZdhI5ww8J0TMF2DC5PlWX6tppO1PM61F/yE8uYegwKcM /H9M/ahzcU9i4V9aF9gqZULu13j+AjRZwrCgKERVs4zgrchRg2fp jh@JHs-MacBook-Pro.local														
Serena Pan	serena	serenapanpan@gmail.com				2016-01-17 19:16:12 UTC	2016-01-26 05:41:28 UTC	Asia/Taipei	en_us	Active	Edit	Edit	Search	Reset	
Additional Attributes Edit **Affiliation:** NCHC **SSH public key:** please add															
pragmac pragmac	pragmac	root@fiji.rocksclusters.org		PRAGMA Cloud		2016-01-12 18:56:17 UTC	2016-01-27 01:59:01 UTC	America/Los_Angeles	en_us	Active	Edit	Edit	Search	Reset	
Additional Attributes Edit **Affiliation:** **SSH public key:**															
Shava Smallen	ssmallen	ssmallen@sdsc.edu				2016-01-13 04:14:36 UTC	2016-01-27 01:20:09 UTC	Asia/Manila	en_us	Active	Edit	Edit	Search	Reset	
Additional Attributes Edit **Affiliation:** UCSD **SSH public key:** ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQCbIXQg9zucGdIEQ0wVwPV5pWCtj3Uz/dVPpNa4iDusXB23Ryy6c/nhxSKOQIGsw2erPKR8Qr5p63oNNEsTjtLtv1cLL/cylVtyL5dQ28+idcVbxLTK/2EUlsenUhG /WtGmPuATM51rG0yAQ/d+OtkC1Y08fCO502vQQgeqo0maS5hZgHgCQ2VAQ4FXQgnWP9 /rrc1dPG0zmcHTME88gESAEO+iXtt3R8i7P3BNuTR7Q0UTYj6aGvhrvfav1Z5ifTUn9iaKK4heKaGYDbgcM3kuXU9D2Q7g5liZSkBAbbn6ac9mwOR03JVjzgizNeAv4sI04YYL2g3YBJAlePAVgROT ssmallen@Shavas-MacBook-Air.local														

Rows: 1 - 5 (5)
Page: 1

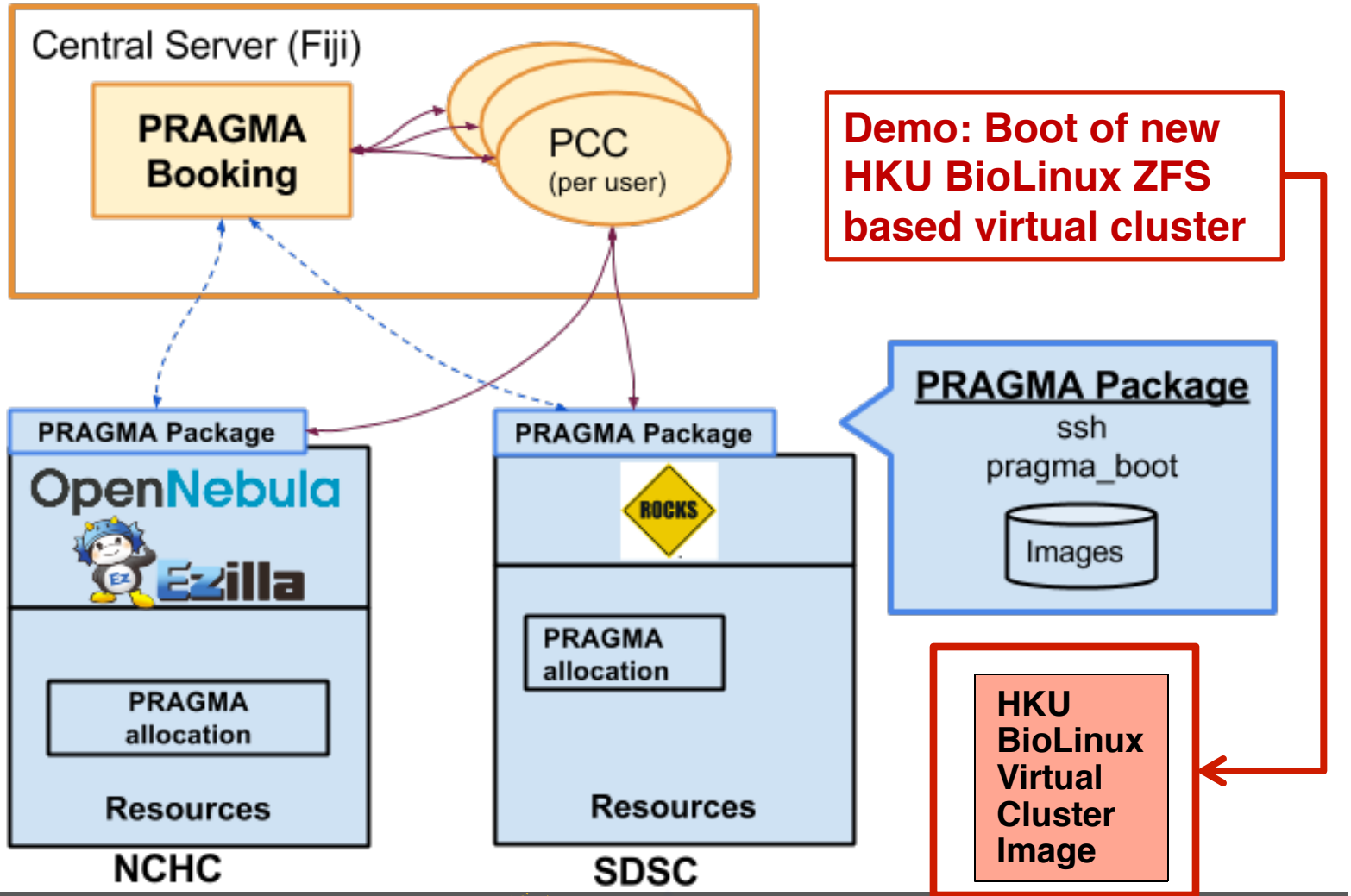
Show resources



Login as nchc-hadmin



Architecture Overview



Creating pragma_boot virtual cluster images

http://sdlab.naist.jp/~ichikawa/data/hku_biolinux.img.tar.bz2

- **Step1: Make image reconfigurable**
  - Installed vc-out-parser (Redhat) and modified to work on Ubuntu
- **Step 2: Make frontend image**
  - Setup IP forwarding and Iptables NAT rules
- **Step 3: Create vc-in.xml file**

Step 4: Add to Pragma_boot repository

- **Repository directory**
  - vcdb.txt
  - hku_biolinux
    - vc-in.xml (based on libvirt XML syntax)
    - hku_biolinux_frontend.img
    - hku_biolinux_compute.img

```
<vc type='Local Beowulf'>
  …
  <frontend> …
      <devices> …
        <disk type='file' device='disk'>
          <source file='hku_biolinux_frontend.img'/>
          <target dev='vda' bus='virtio'/>
        </disk> …
      </devices>
    </domain>
  </frontend>
  <compute> …
      <devices> …
        <disk type='file' device='disk'>
          <source file='hku_biolinux_compute.img'/>
        </disk> ….
      </devices>
    </domain>
 </compute>
 …
</vc>
```

(Takes about 40 minutes to launch a 2-node cluster)

```
$ bin/pragma boot hku_biolinux 16
```

Leveraging ZFS and img-storage roll

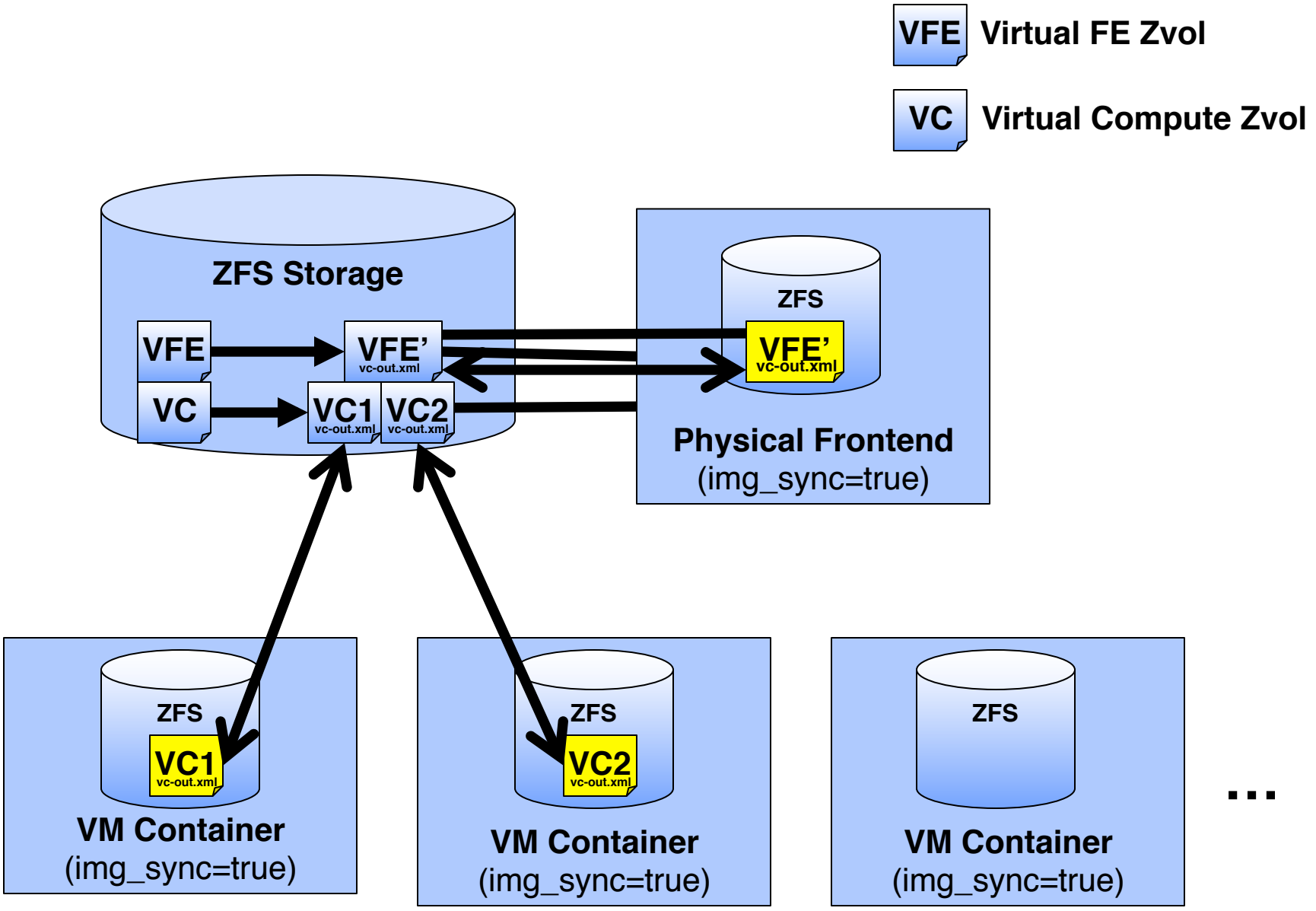
- **Step 5: Create a ZFS volume and dd image into volume (39.5 GB)**
  - Boots within 6 mins but takes 2.5 hours to sync images behind the scenes

OR

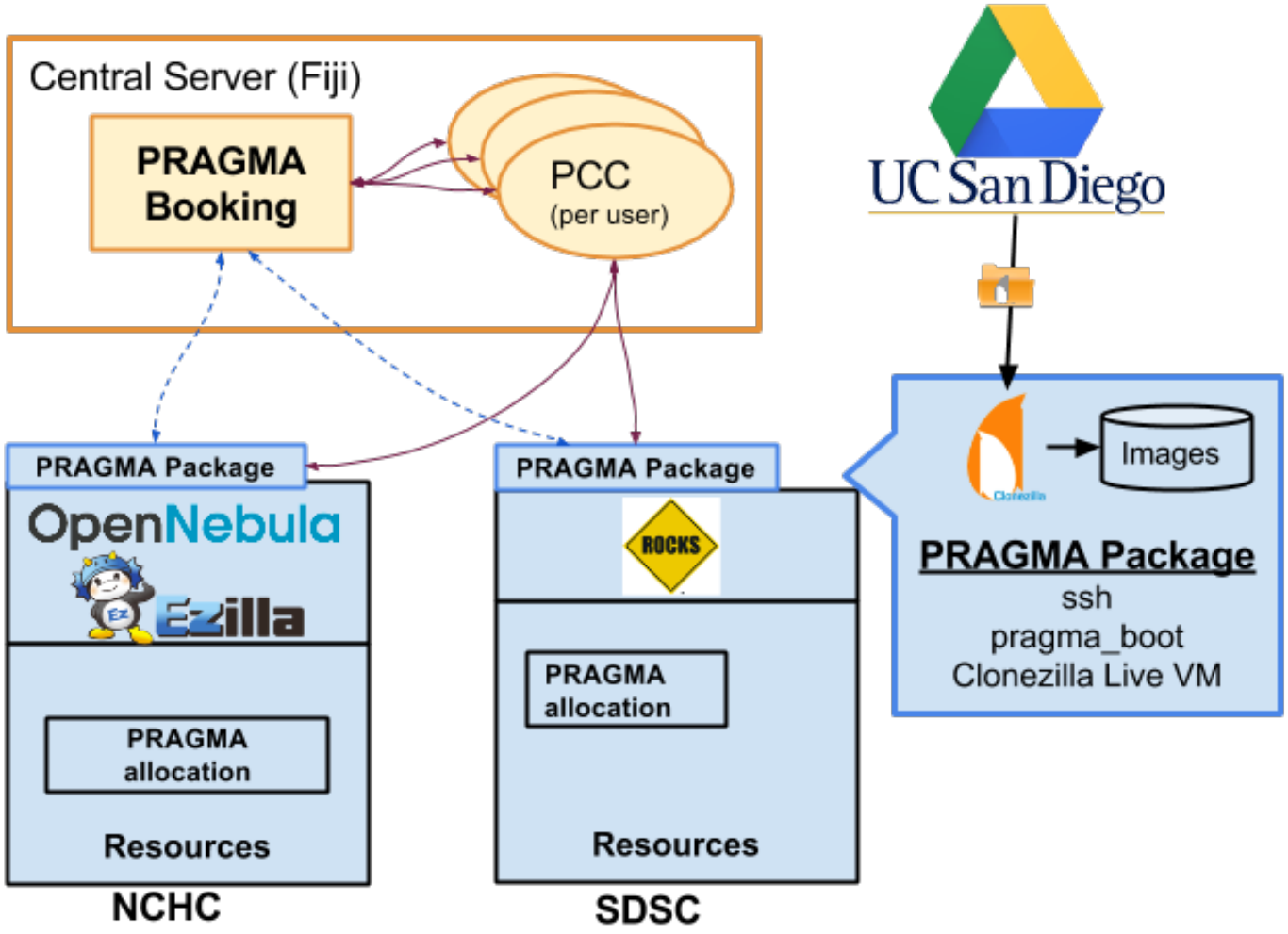
- **Step 5: Use Clonezilla to create ZFS volume (10.7GB)**
  - Boots within 6 mins but takes 25 mins to sync images

Step 6: Modify vc-in.xml

```
<vc type='Local Beowulf'>
  …
  <frontend> …
      <devices> …
        <disk type='volume' device='disk'>
          <source volume='state/hku_biolinux_fe'
                  pool='state' host='nas-0-0'/>
          <target dev='vda' bus='virtio'/>
        </disk> …
      </devices>
    </domain>
  </frontend>
  <compute> …
      <devices> …
        <disk type='volume' device='disk'>
          <source volume='state/hku_biolinux_compute'
                  pool='pragma' host='nas-0-0'/>
          <target dev='vda' bus='virtio'/>
        </disk> ….
      </devices>
    </domain>
 </compute>
 …
</vc>
```


VFE
Virtual FE Zvol
VC
Virtual Compute Zvol
ZFS Storage
VFE
VFE'
vc-out.xml
VC
VC1
vc-out.xml
VC2
vc-out.xml
ZFS
VFE'
vc-out.xml
Physical Frontend
(img_sync=true)
ZFS
VC1
vc-out.xml
VM Container
(img_sync=true)
ZFS
VC2
vc-out.xml
VM Container
(img_sync=true)
ZFS
VM Container
(img_sync=true)
...

Revised Cloud Scheduler Architecture



Goals for PRAGMA 31

- Integrate 6 additional sites: AIST, NAIST, Thammasat / Kasetsart, IU, UF, CNIC
- Integrate PRAGMA-ENT
- Integrate more virtual cluster images and get feedback from early users (Jason, others?)
- Integrate image management with Google drive and Clonezilla
- Integrate Cloud Init and boto with pragma_boot for greater portability
- Package and document software

More Information

- **Contact email**
  - ssmallen@ucsd.edu

- **Website**
  - http://fiji.rocksclusters.org/cloud-scheduler
  - https://github.com/pragmagrid/pragma_boot

We welcome early users ☺

Thank you!