Virtual cluster image sharing on the PRAGMA cloud testbed

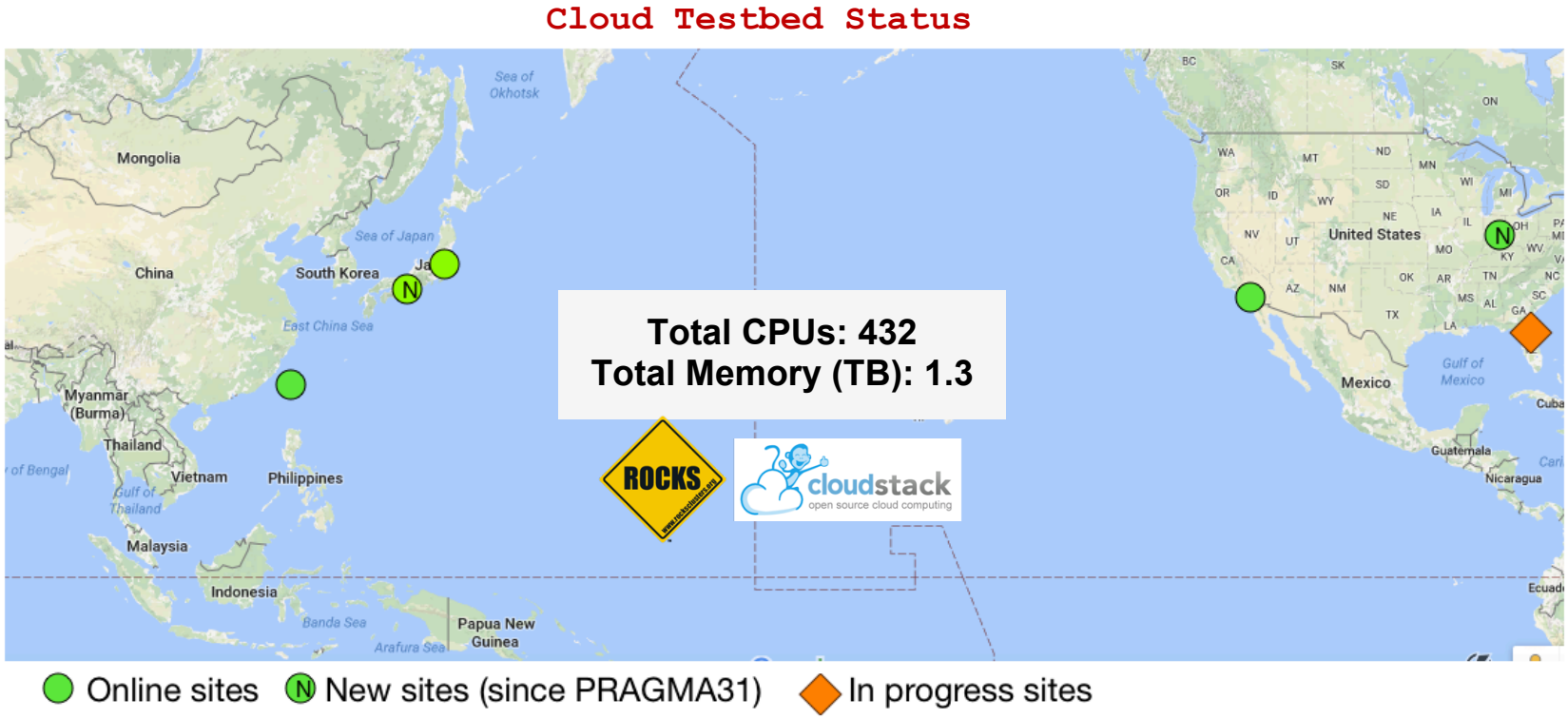
Shava Smallen
Steven Shiau, Nadya Williams,
Weicheng Huang, Philip Papadopoulos

ssmallen@sdsc.edu

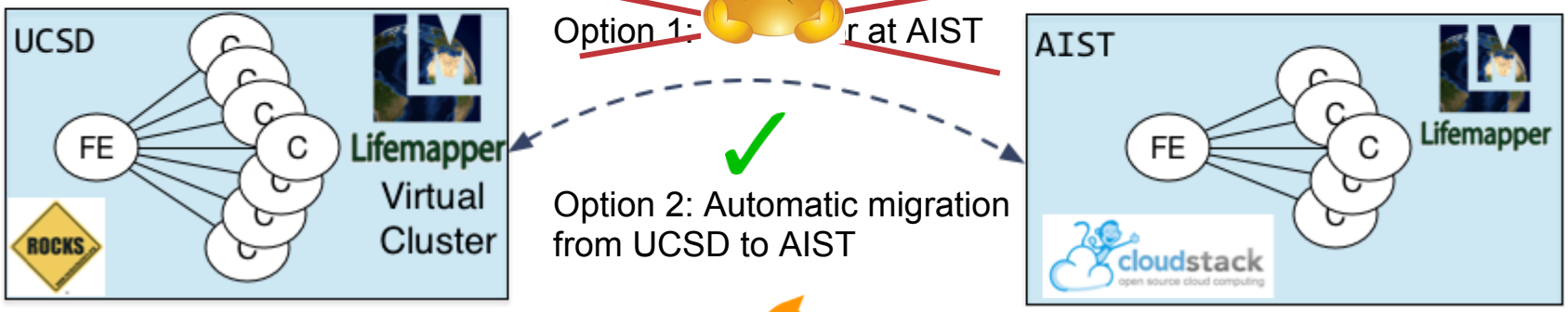
April 13, 2017
PRAGMA32

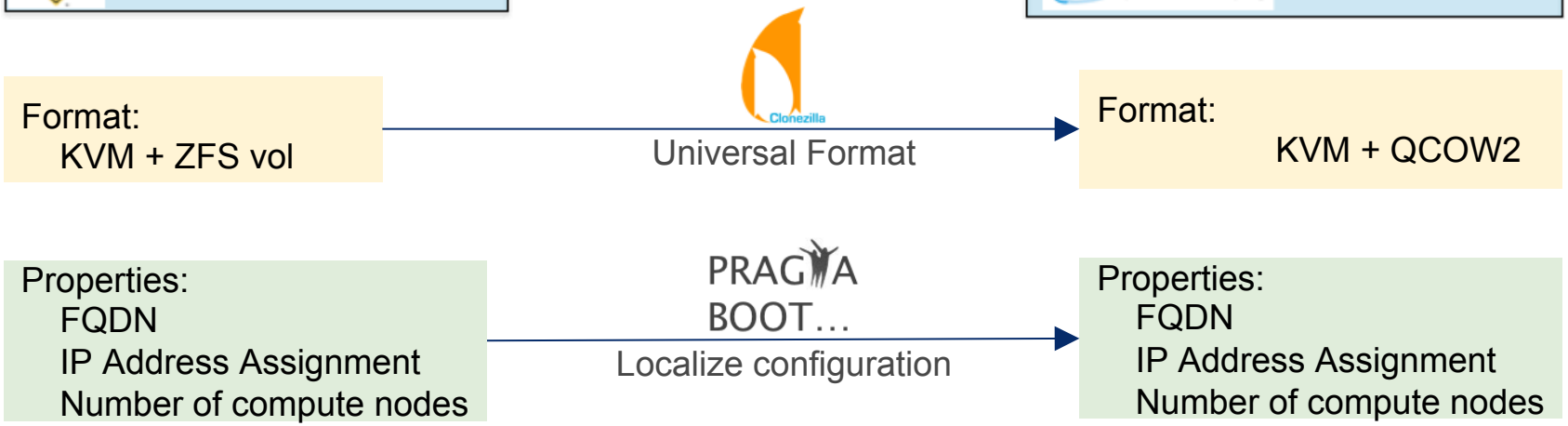
PRAGMA Cloud Testbed

- **Goal**: A persistent Cloud testbed for Biosciences and other PRAGMA working group members to run application experiments.



Motivation



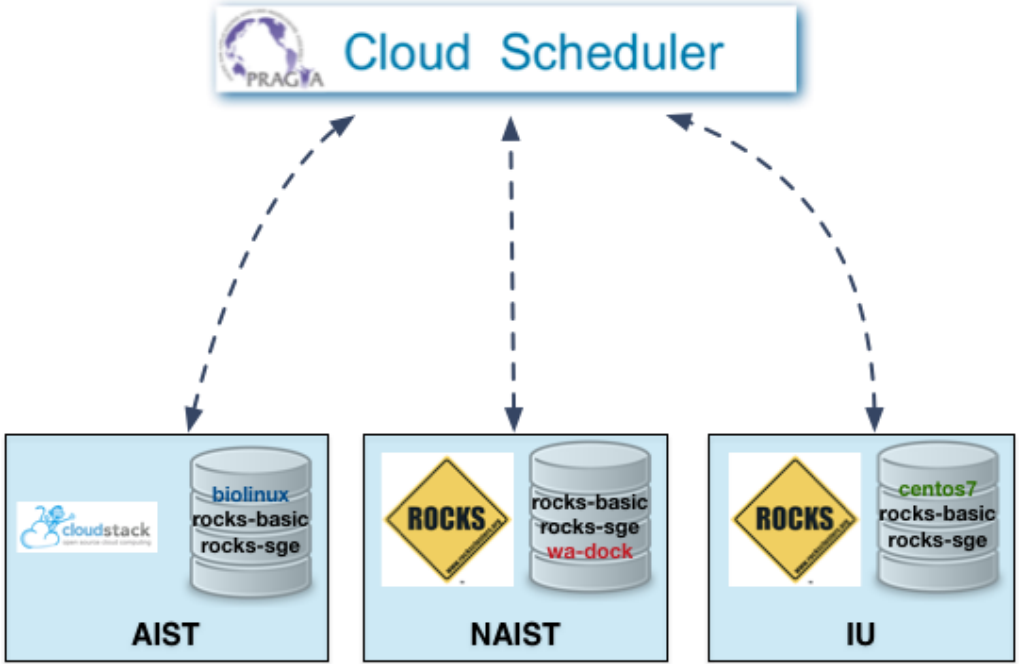


When (and where else) can I run my virtual cluster? cloud scheduler.

PRAGMA Cloud Scheduler

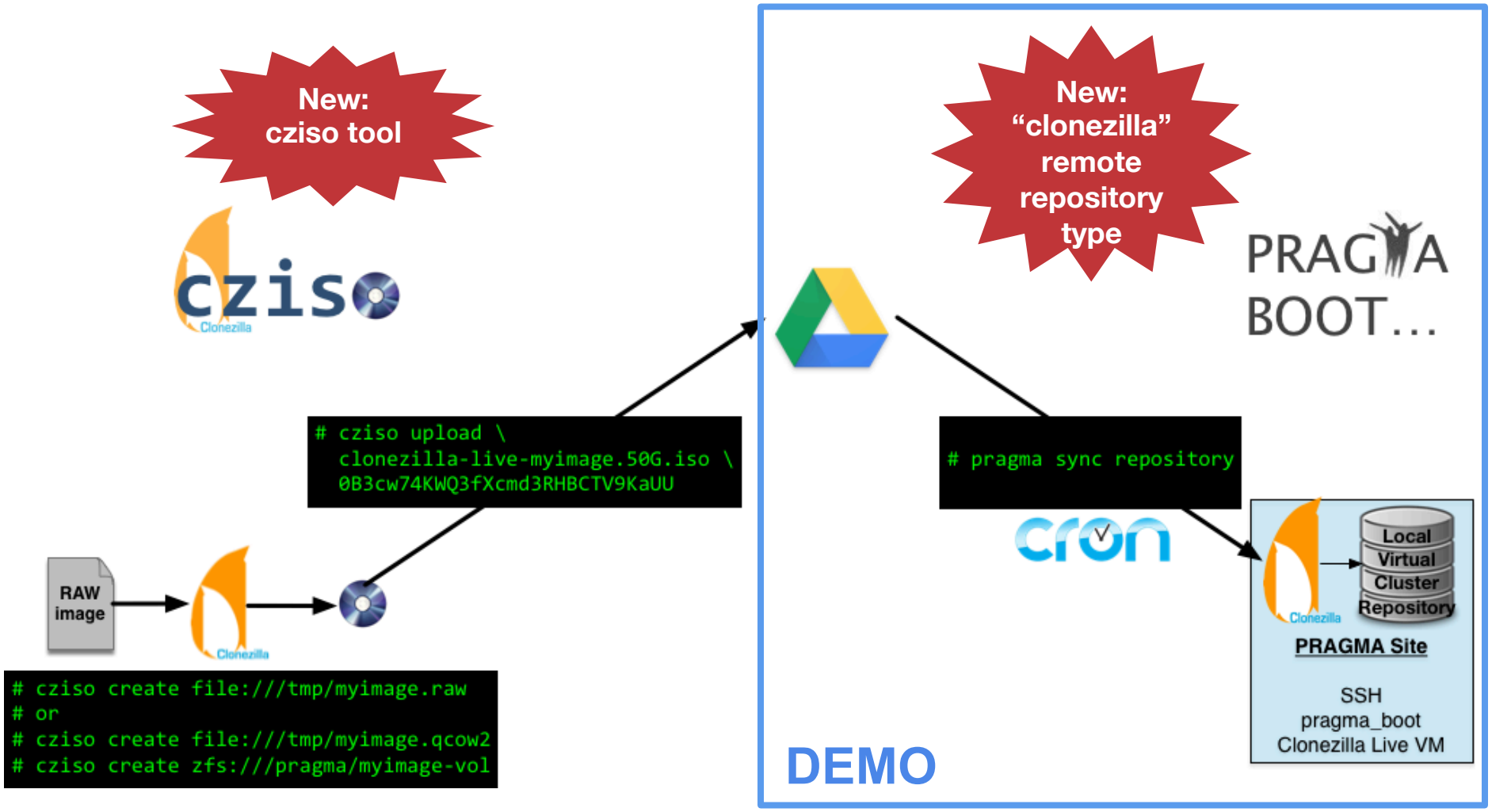
- **Goal: Low participation overhead and easy to use**
  - Sites only have to install a small package (SSH and pragma_boot) to participate
  - Users have convenient web interface to start up and manage their virtual clusters
- **Currently leverages the following tools:**
  - **Booked**: Open source room reservation software from Twinkle Toes (will be replaced by new GUI -- next demo)
  - **pragma_boot**: Boots virtual clusters for users across PRAGMA institutions using local VM provisioning system. Currently supports Rocks and Cloudstack. Openstack coming.

Sharing local virtual cluster images

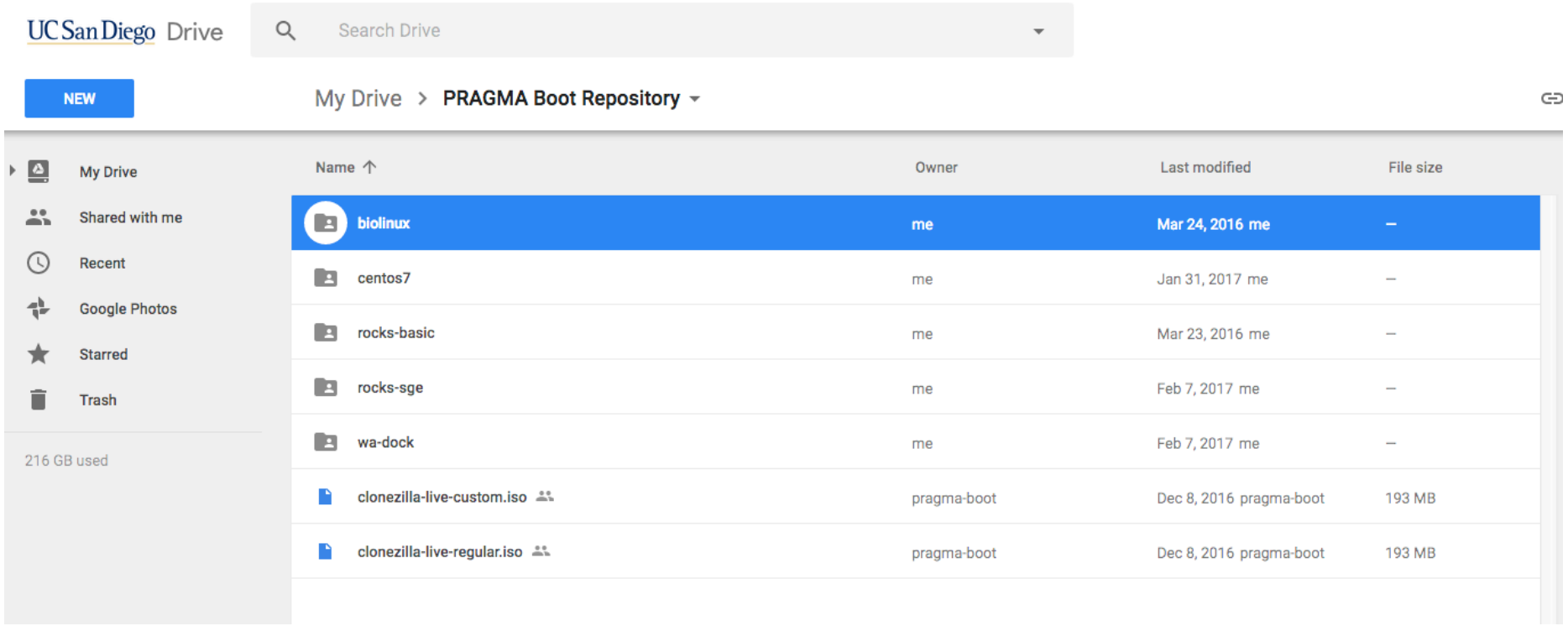


- Goal: Leverage UCSD's unlimited Google storage and Clonezilla to convert images to different formats (raw, qcow, zvol)

Virtual cluster image sharing



Google Drive Repository



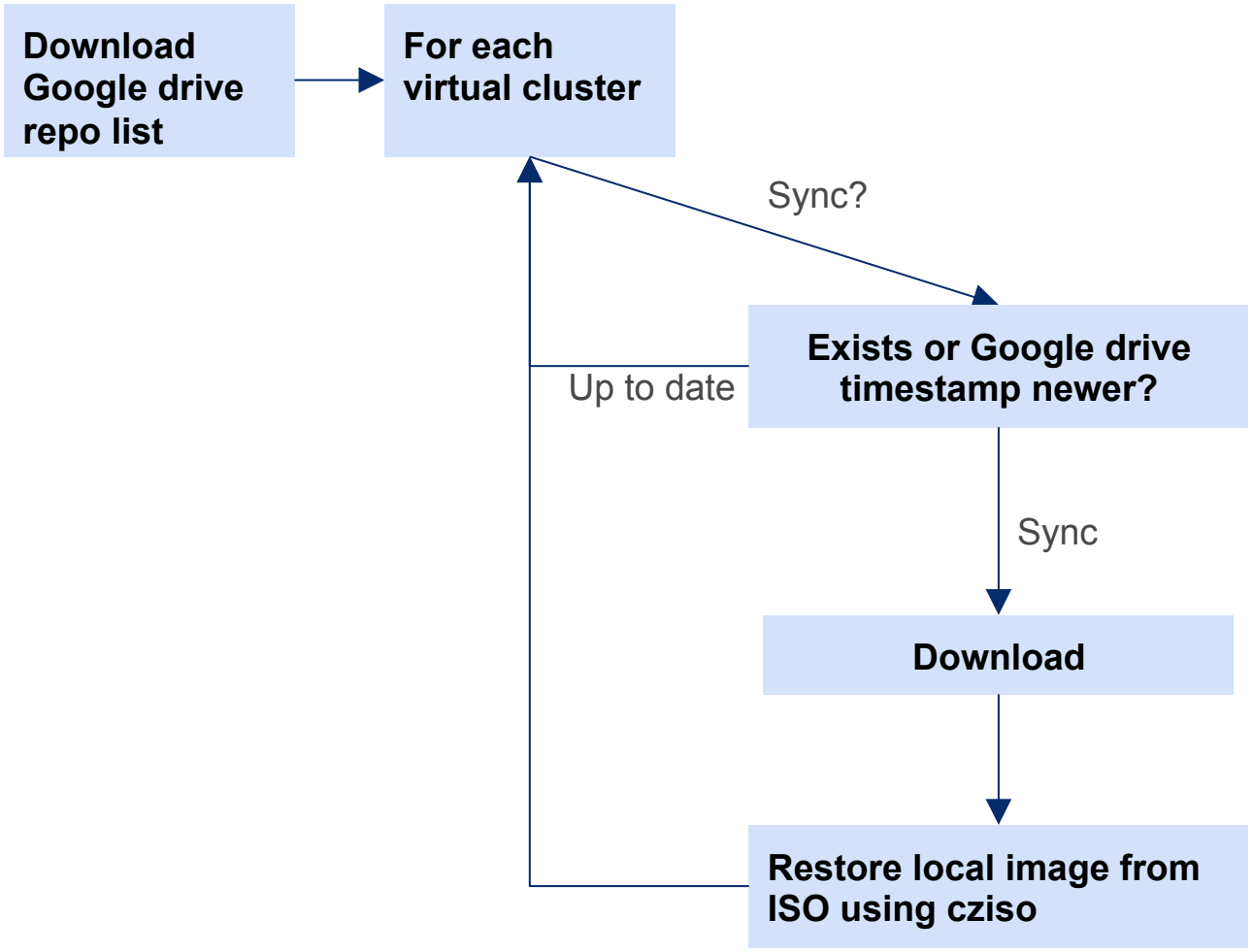
Local repository list on Fiji

```
[root@fiji pragma_boot]# ./bin/pragma list
  repository
VIRTUAL IMAGE
hku_biolinux
rocks-basic
rocks-sge
wa-dock
```

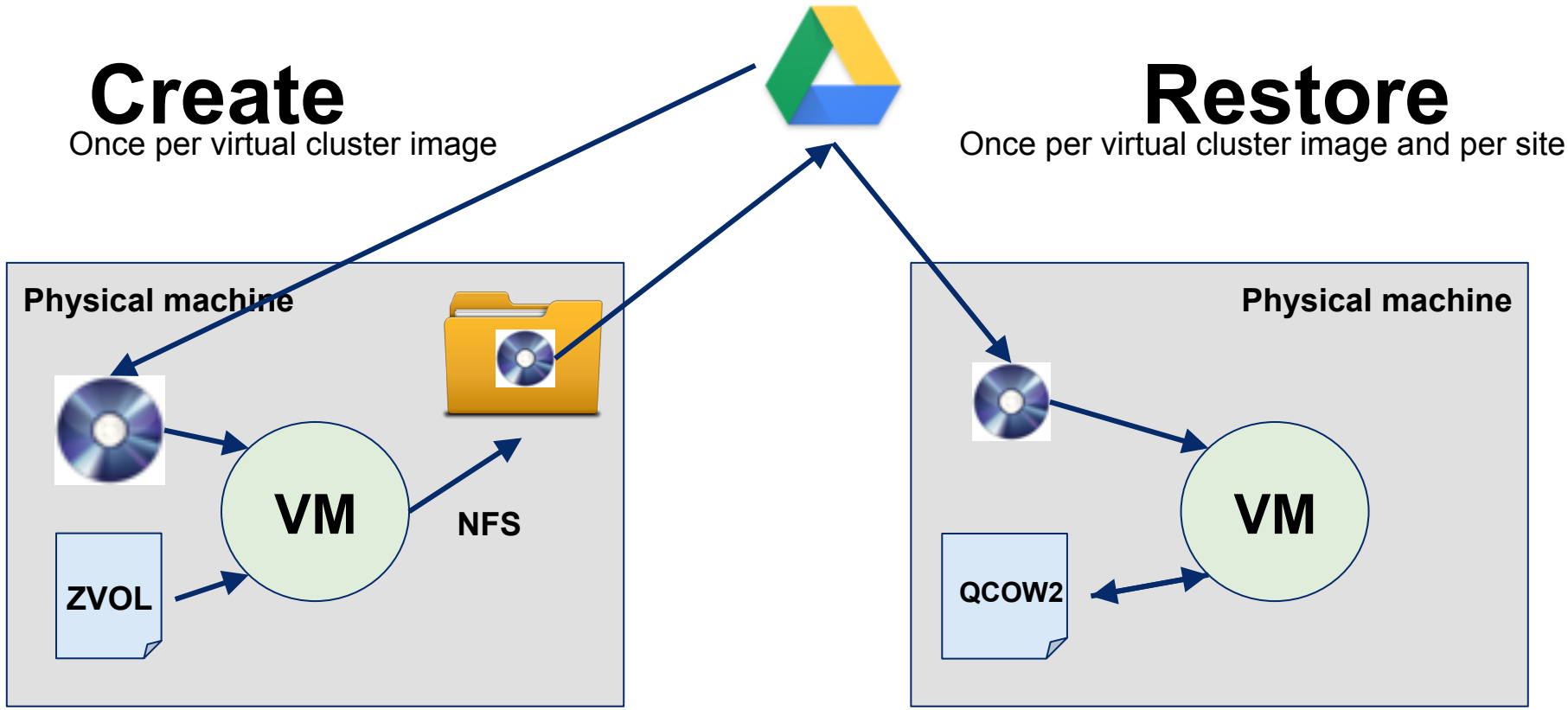
Run PRAGMA Sync repository

Takes ~5 mins to sync

Clonezilla Remote Repository Details



cziso

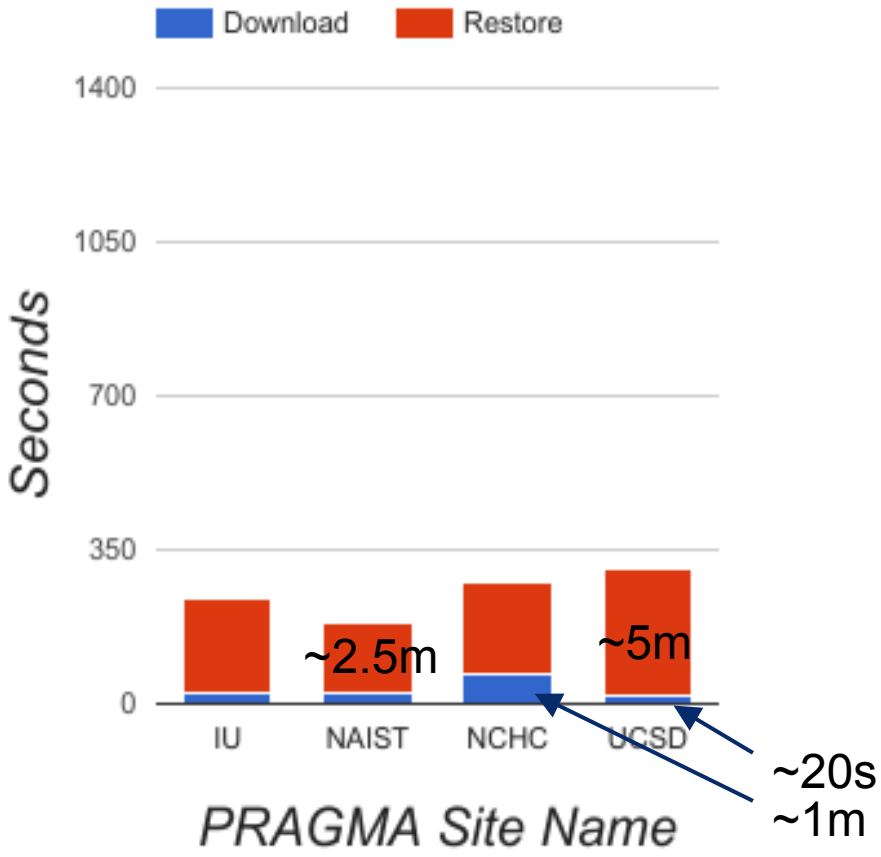


7. Uploads cziso to Google drive

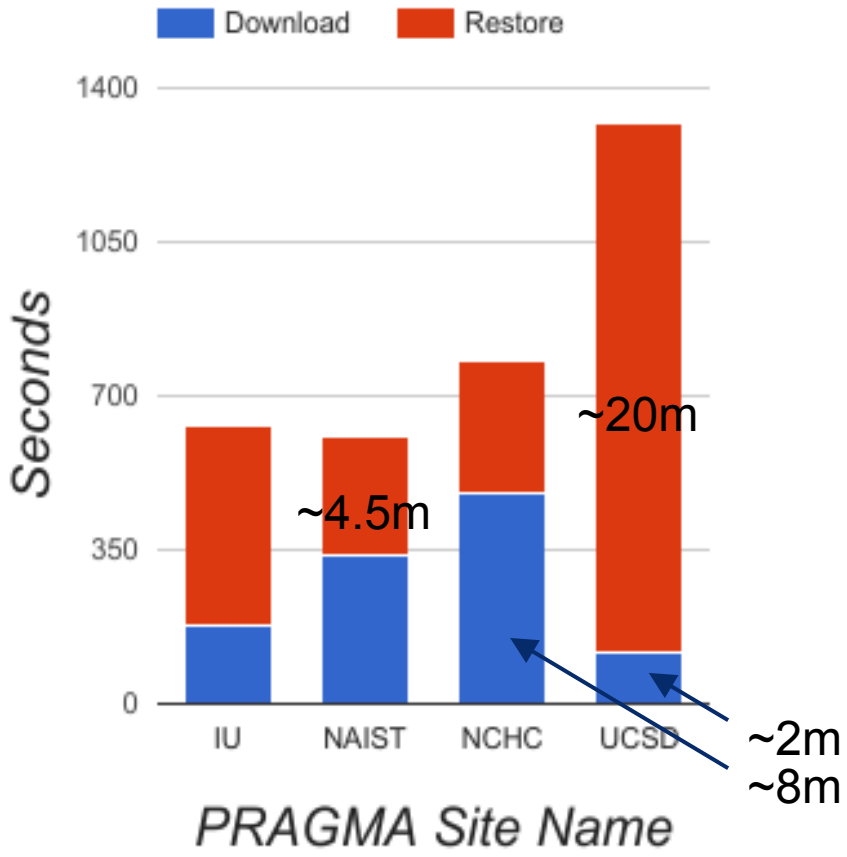
5. cziso VM automatically shuts down and restored image is ready to be used.

Examples of Download/Restore Times

centos7 frontend image (733 MB)



biolinux frontend image (5 GB)



Launch CentOS 7 cluster

Goals for PRAGMA 33

- Integrate Thammasat Cloud Scheduler GUI
- Finish Clonezilla repository integration with different drivers (e.g., Cloudstack)
- Finish PRAGMA-ENT integration
- Finish integration of University of Florida
- Add UCSD Rockstar resource
- Add Openstack PRAGMA Boot driver
- Add Airbox virtual cluster image (CENTRA)
-

More information

- Thank you to our collaborators at Thammasat University, NAIST, and Indiana University

- Email: pragma-cloud-admin@googlegroups.com
- Websites:
  - https://fiji.rocksclusters.org/cloud-scheduler
  - http://pragma-grid.net/site-setup

PRAGMA BOOT…

Clonezilla Remote Repository Type

```
repository_settings = {
   'repository_class' : 'clonezilla',
   'repository_dir' :  '/state/kvmdisks/repository',
   'vcdb_filename' : 'vcdb.txt',
   'repository_url': 'https://drive.google.com/drive/u/1/folders/0B3cw7uKWQ3fXcmdfRHBCTV9KaUU',

   # The following settings are Required for clonezilla repository
   'cziso': '/opt/cziso/bin/cziso',
   'local_image_url': 'zfs://nas-0-0/state/$imagename-vol',
   #'local_image_url': 'file://$repository_dir/$imagename.raw',
   # 'local_image_url': 'file://$repository_dir/$imagename.qcow2',

   # Only sync images from remote repository matching the below pattern
   #'include_images': 'biolinux',

   # Sync all images from remote repository except those matching below pattern
   # 'exclude_images': 'lifemapper'
}
```

Clonezilla ISO tool

https://github.com/pragmagrid/cziso

- **create <image>** - create a cziso file from a RAW, QCOW2, or ZVOL VM image
- **restore <iso> <image>** - restore a RAW, QCOW2, or ZVOL VM image from a cziso file
- **test <image>** - test a generated a RAW, QCOW2, or ZVOL VM image
- **update <zip>** - create a new custom Clonezilla ISO from latest Clonezilla distro
- **upload <file> <gdrive_folder>** - upload file to Google drive

PRAGMA Cloud Scheduler History

- **PRAGMA28**: Prototype of Lightweight Cloud Scheduler using Booked
  - **PRAGMA29**: Enabled multiple reservations per resource
  - **PRAGMA30**: Introduced pragma_boot version 2 and Rocks ZFS enabled virtual clusters
  - **PRAGMA31**: Cloudstack driver, Clonezilla experimentation, and student hackathon