PENNSTATE
1855

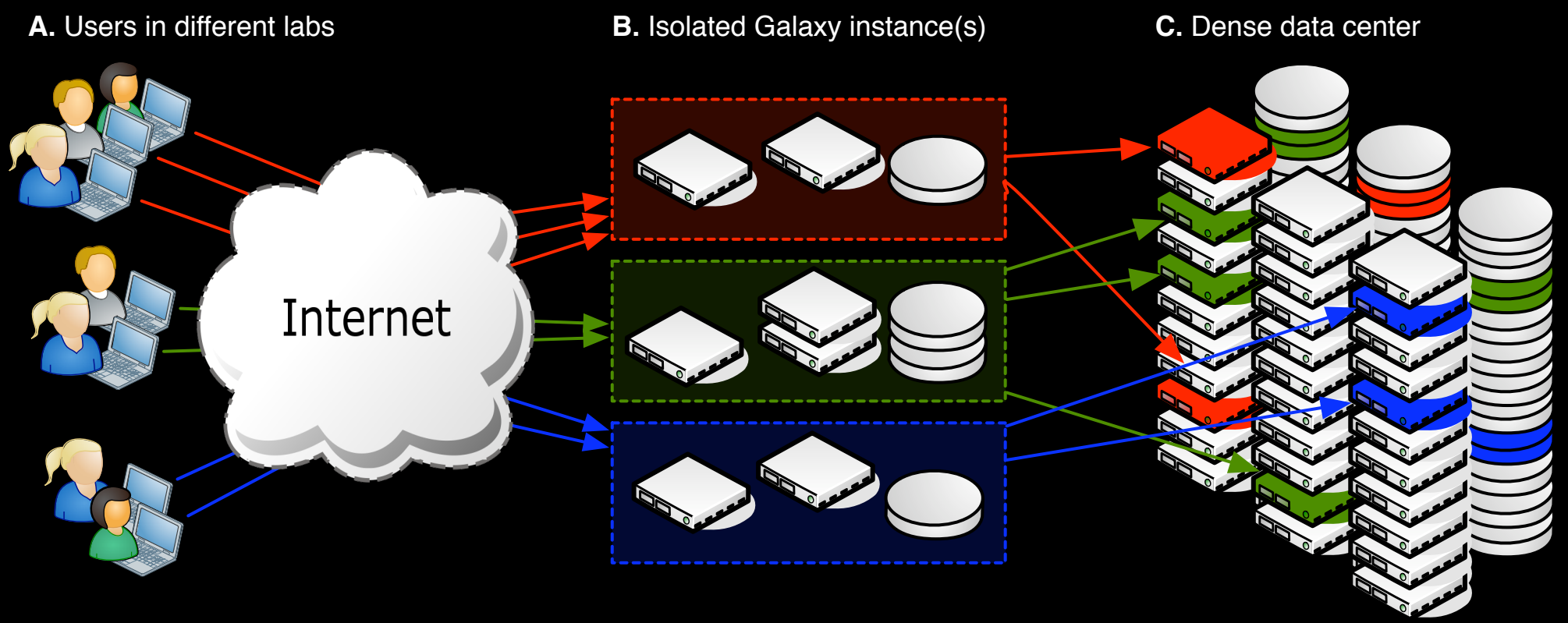
Enabling NGS analysis with(out) the infrastructure

Enis Afgan, the Galaxy Team, Anton Nekrutenko, James Taylor

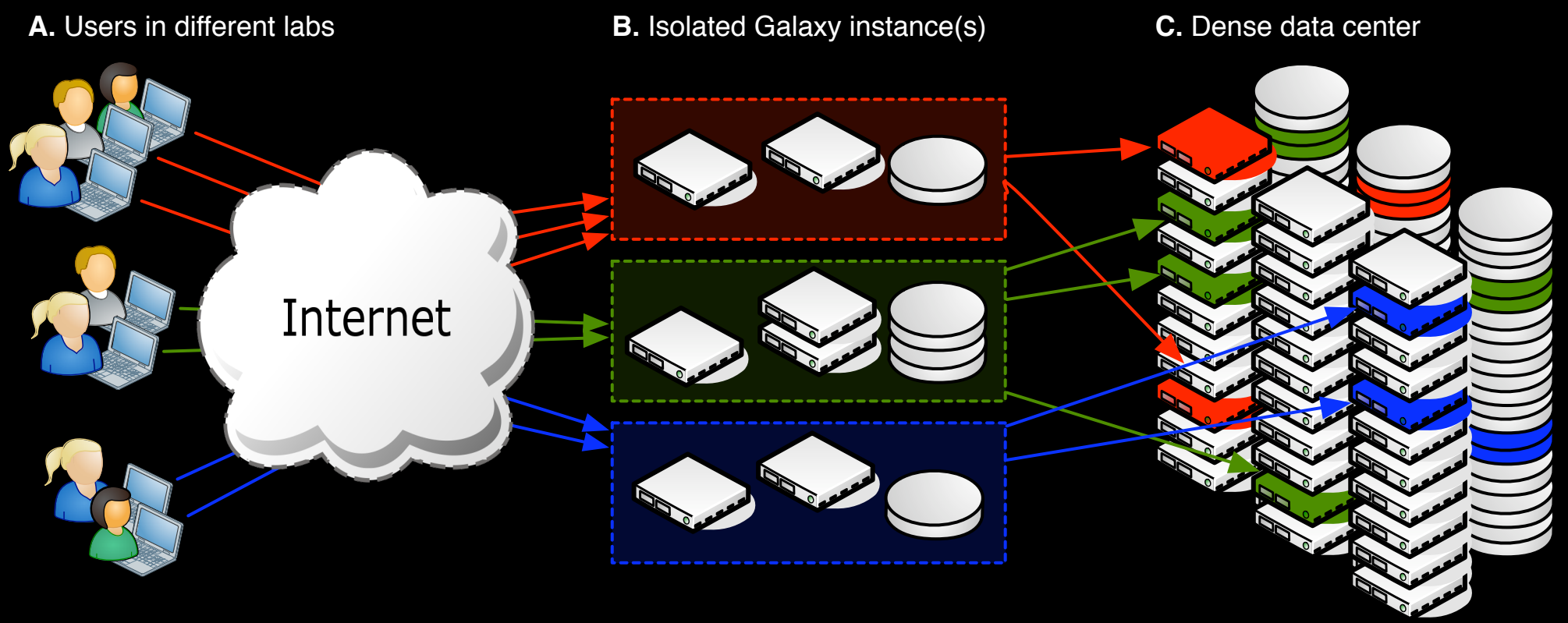
Bioinformatics Open Source Conference, July 16, 2011, Vienna

EMORY

CloudMan-as-a-Bridge

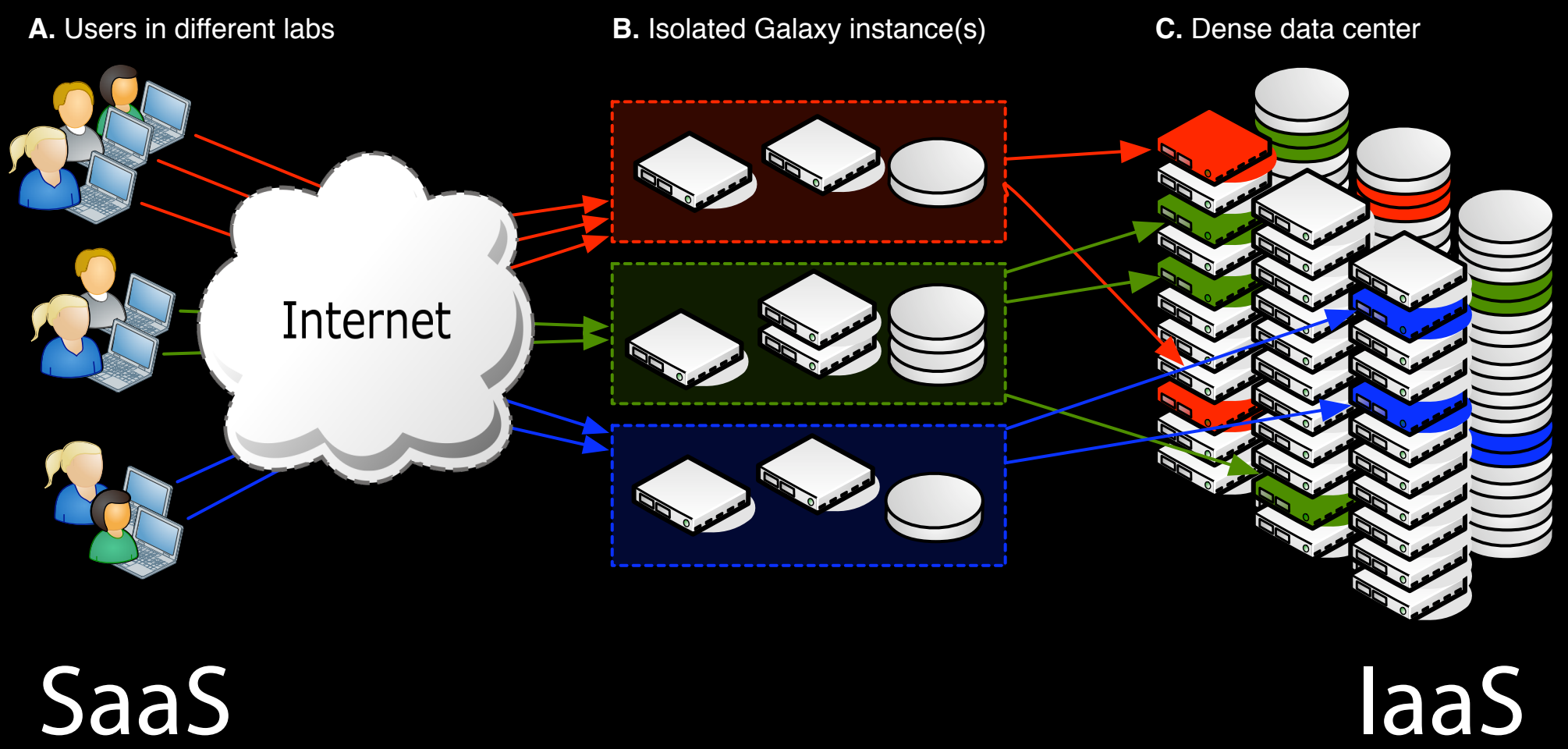


CloudMan-as-a-Bridge

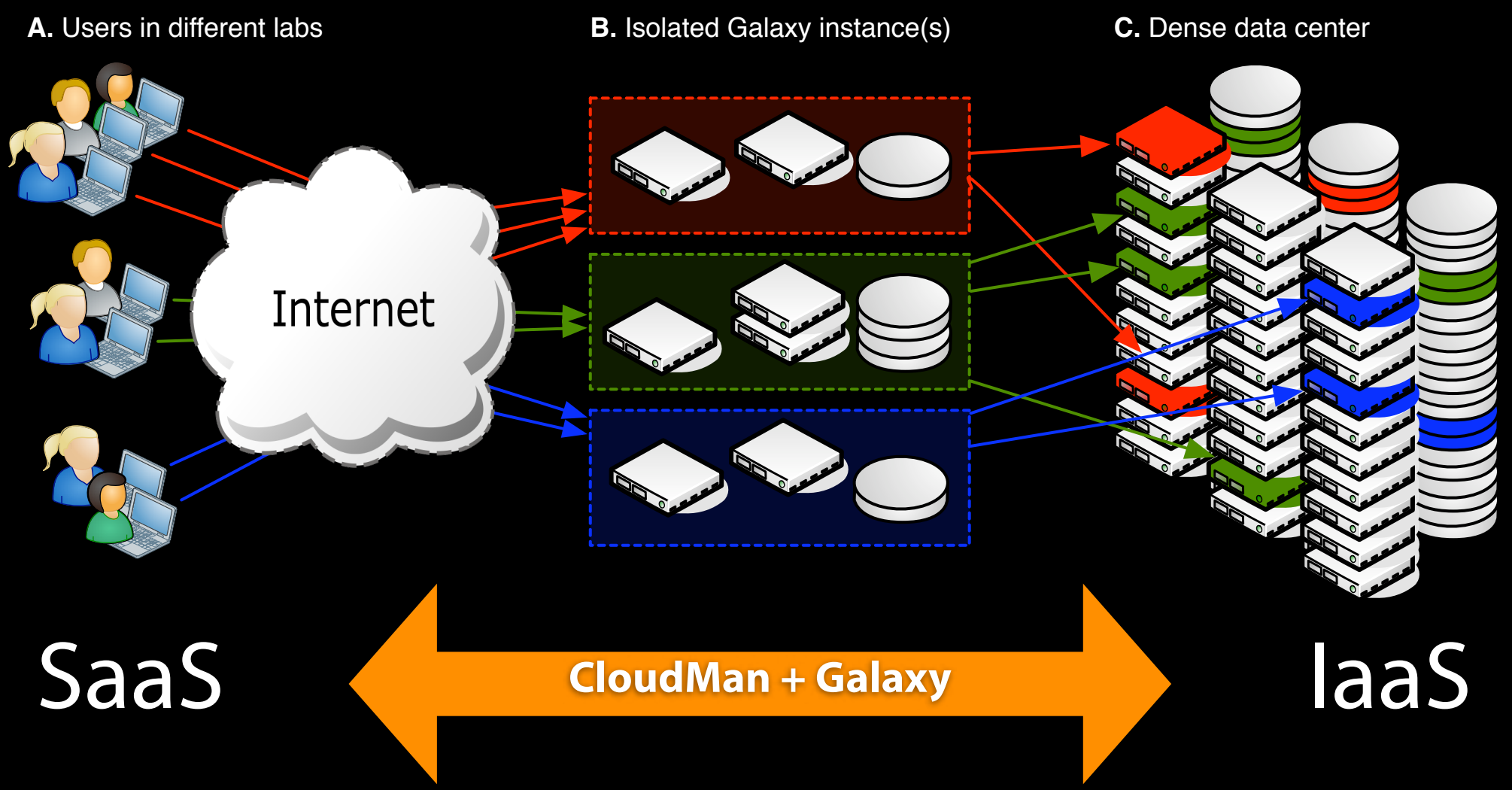


SaaS

CloudMan-as-a-Bridge



CloudMan-as-a-Bridge



CloudMan Platform

A complete solution for instantiating and managing cloud resources

- With automatically configured Galaxy (if desired)
- Scope of tools and reference datasets exceed Galaxy Main

Deploy a (Galaxy) cluster in minutes!

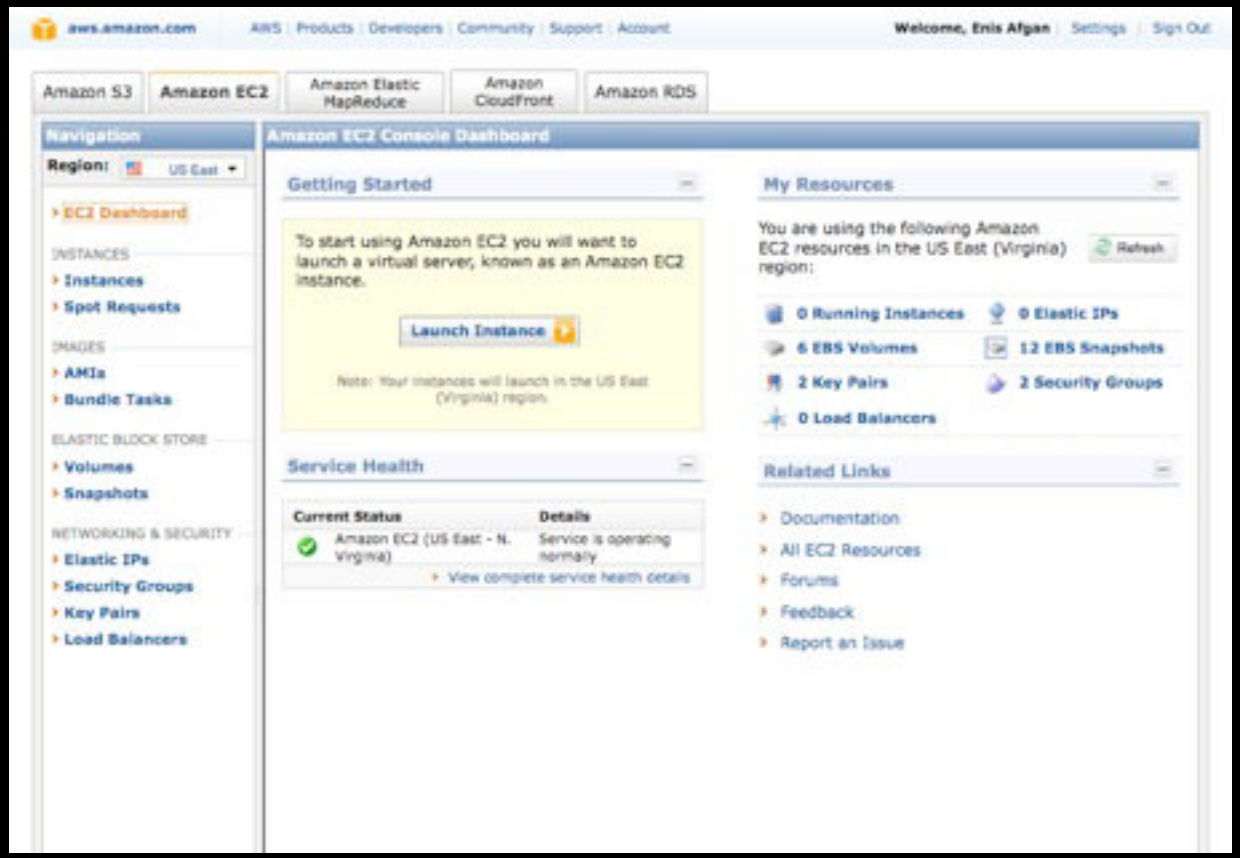
CloudMan features

- Deployment on Amazon Web Services Cloud
  - Wizard-guided setup: requires no computational expertise, no infrastructure, no software
- Automated (thus reproducible) configuration for machine image, tools, and data
- Four modes of cluster type setup
- Dynamic persistent storage
- Elastic resource scaling: manual or automatic based on workload
- Standalone deployment, requiring no external dependencies or services
- Customizable by individual users
- Sharing of derived cluster instances -> even the customized ones!

CloudMan features

- Deployment on Amazon Web Services Cloud
  - Wizard-guided setup: requires no computational expertise, no infrastructure, no software
- Automated (thus reproducible) configuration for machine image, tools, and data
- Four modes of cluster type setup
- Dynamic persistent storage
- Elastic resource scaling: manual or automatic based on workload
- Standalone deployment, requiring no external dependencies or services
- Customizable by individual users
- Sharing of derived cluster instances -> even the customized ones!

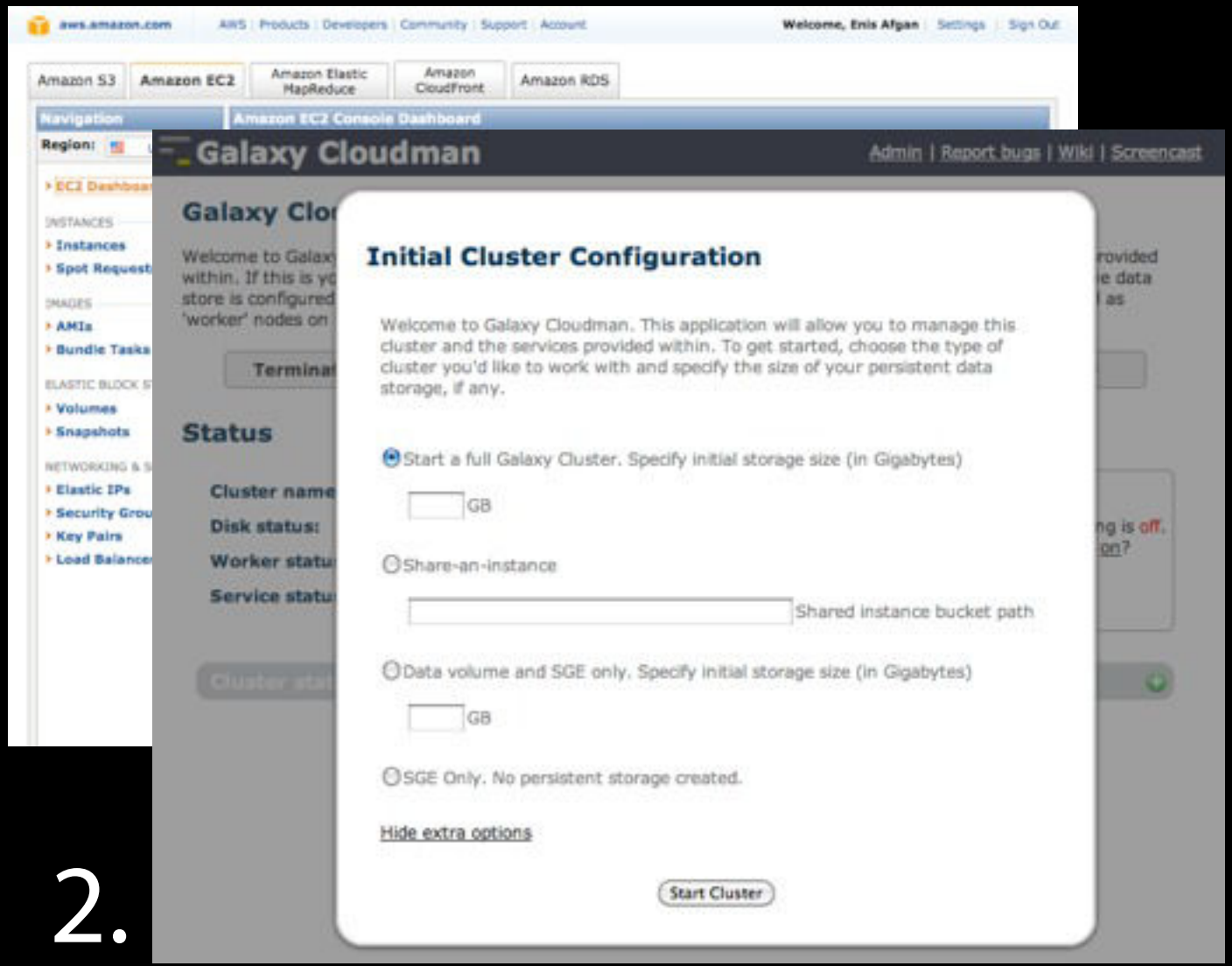
Deploying a cluster on AWS

1. 

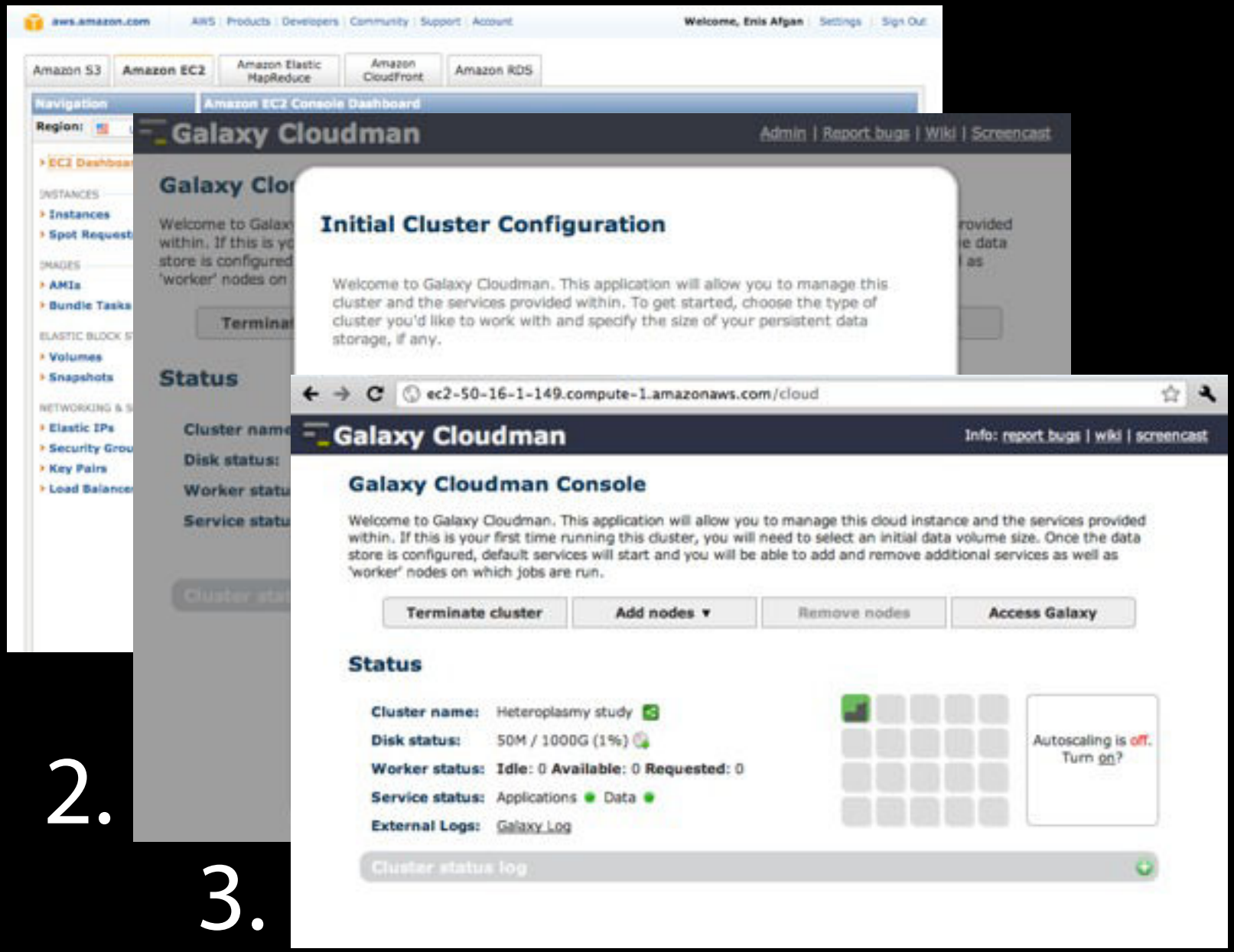
Deploying a cluster on AWS

1.

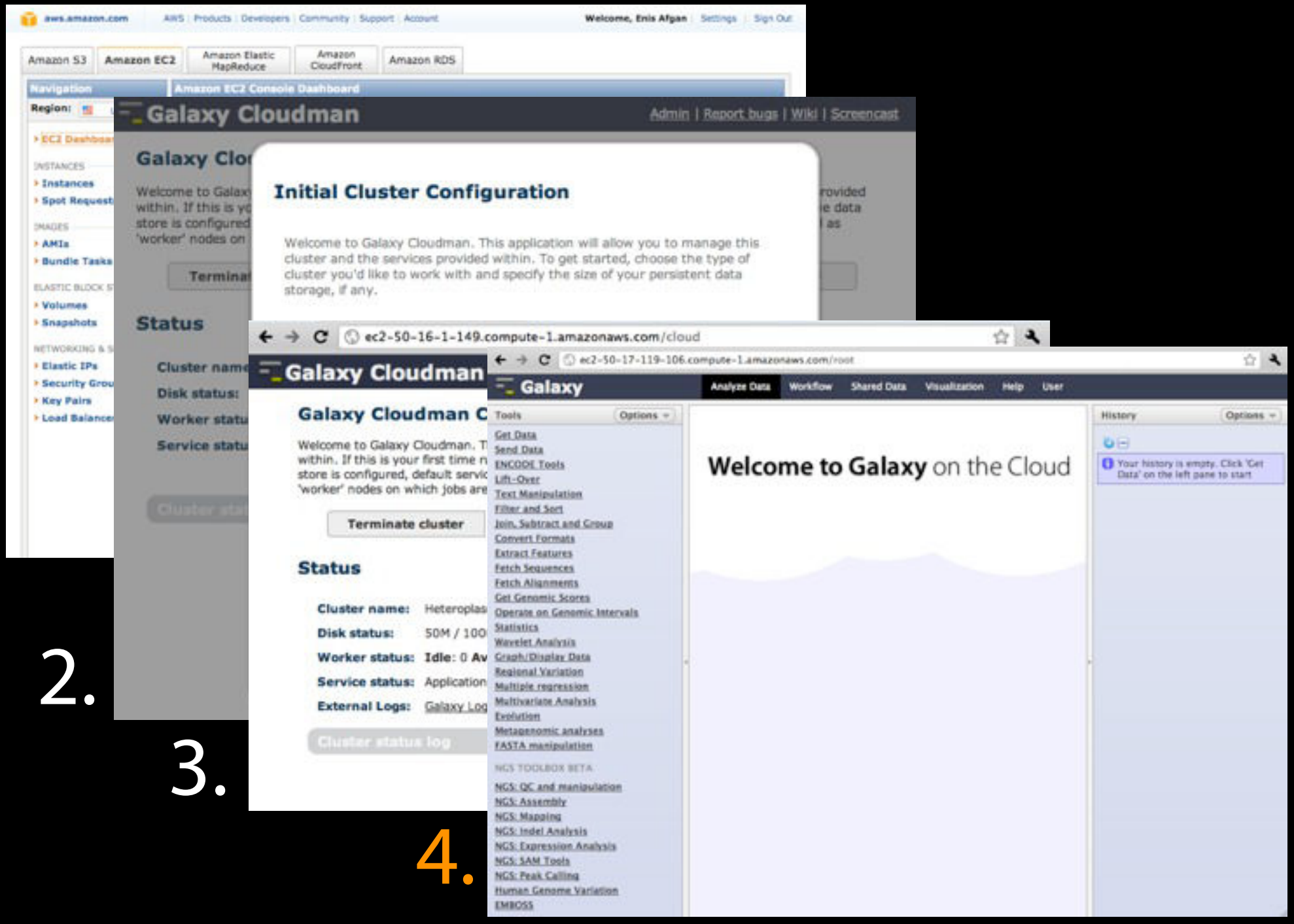
2.



Deploying a cluster on AWS



Deploying a cluster on AWS



CloudBioLinux + Galaxy + CloudMan =

- A lot of (NGS) tools immediately available and easily accessible
- 700GB of reference genome data

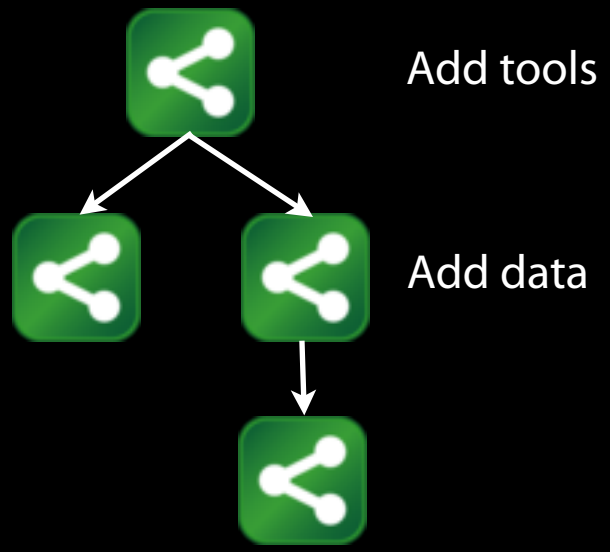
Bowtie, BWA, Samtools, MAQ, BFAST, ABySS, Velvet, MACS, Tophat, Cufflinks, MegaBLAST, BLAST, Sputnik, Taxonomy, HyPhy, Lastz, Perm, GATK, Srma, Beam, Pass, LPS, Plink, Haploview, Freebayes, Mosaik, Picard, ...

But what if your tool (or data) is missing?

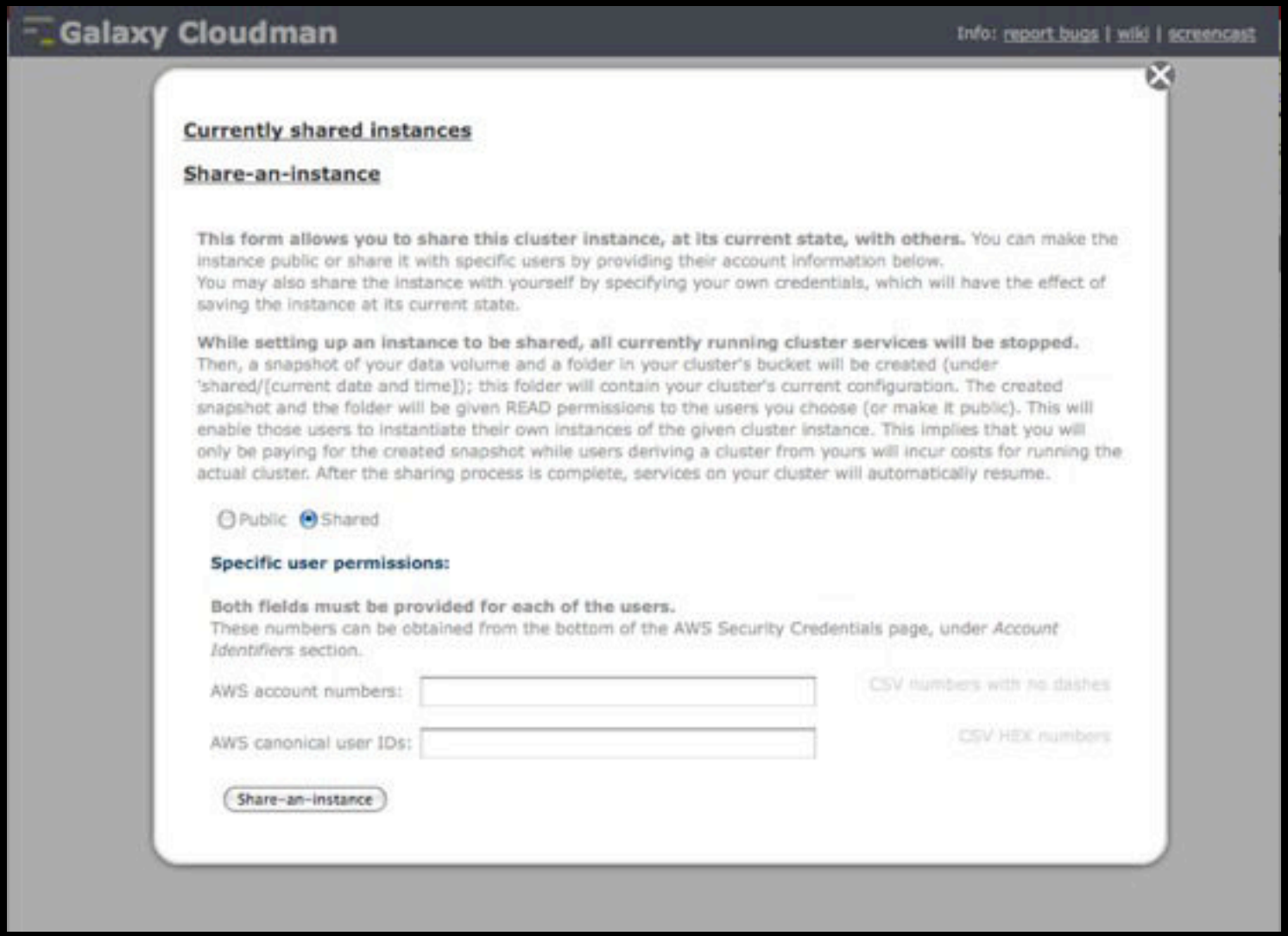
1. Add it! (via automation)
   - CloudMan instances are self-contained

```
def _install_picard():
    version = '1.45'
    mirror_info = "?use_mirror=voxel"
    url = 'http://downloads.sourceforge.net/project/picard/picard-tools/%s/picard-tools-%s.zip' \
          % (version, version)
    pkg_name = 'picard'
    install_dir = os.path.join(env.install_dir, pkg_name, version)
    install_cmd = sudo if env.use_sudo else run
    if not exists(install_dir):
        install_cmd("mkdir -p %s" % install_dir)
    with _make_tmp_dir() as work_dir:
        with cd(work_dir):
            run("wget %s%s -O %s" % (url, mirror_info, os.path.split(url)[-1]))
            run("unzip %s" % (os.path.split(url)[-1]))
            install_cmd("mv picard-tools-%s/*.jar %s" % (version, install_dir))
    sudo("touch %s/env.sh" % install_dir)
    sudo("chmod +x %s/env.sh" % install_dir)
    install_dir_root = os.path.join(env.install_dir, pkg_name)
    sudo('if [ ! -d %s/default ]; then ln -s %s %s/default; fi' % (install_dir_root, install_dir, install_
    print(green("------ Picard %s installed to %s ------" % (version, install_dir)))
```

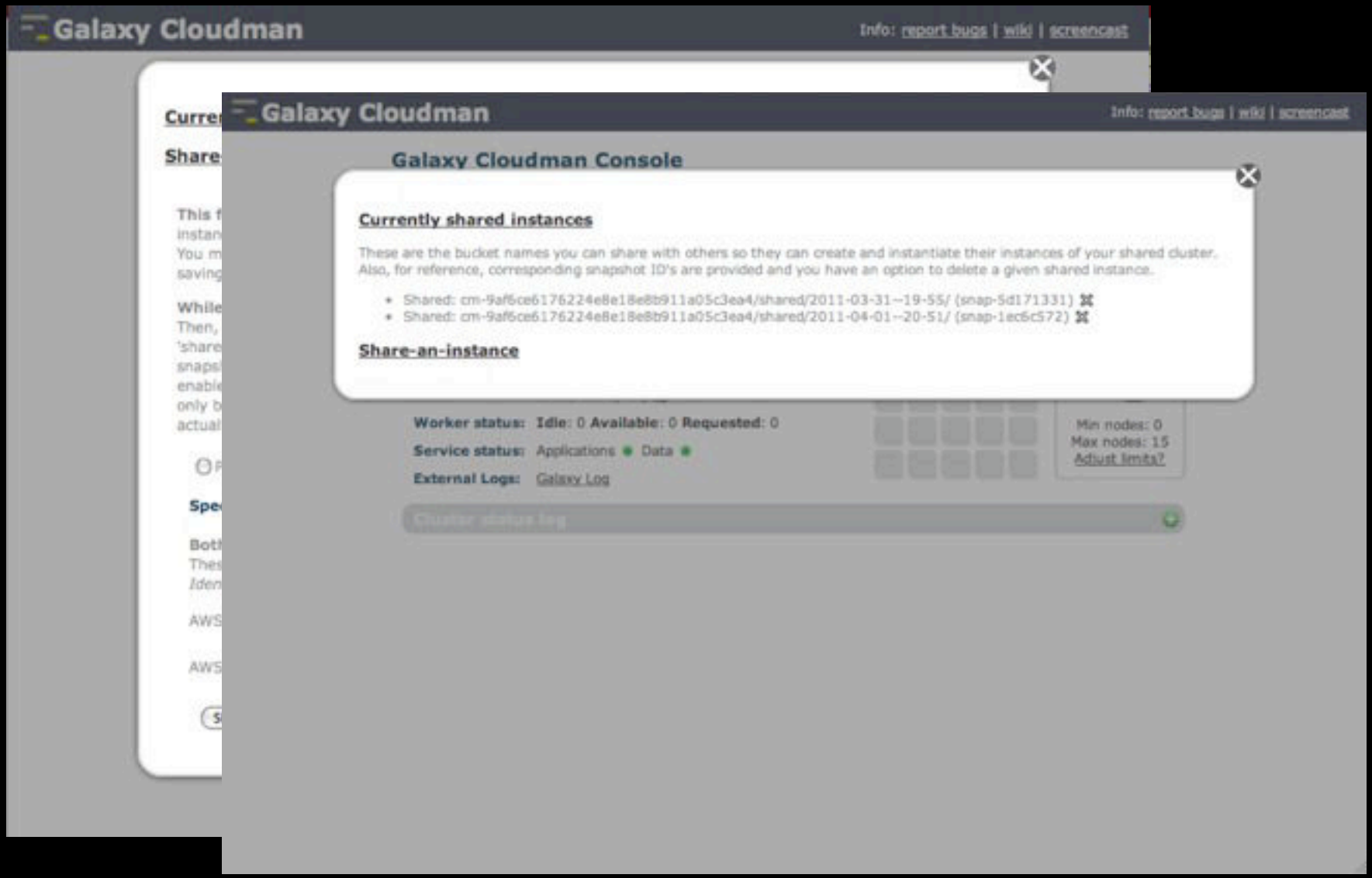
2. Save & share
   - With individual users or make it public



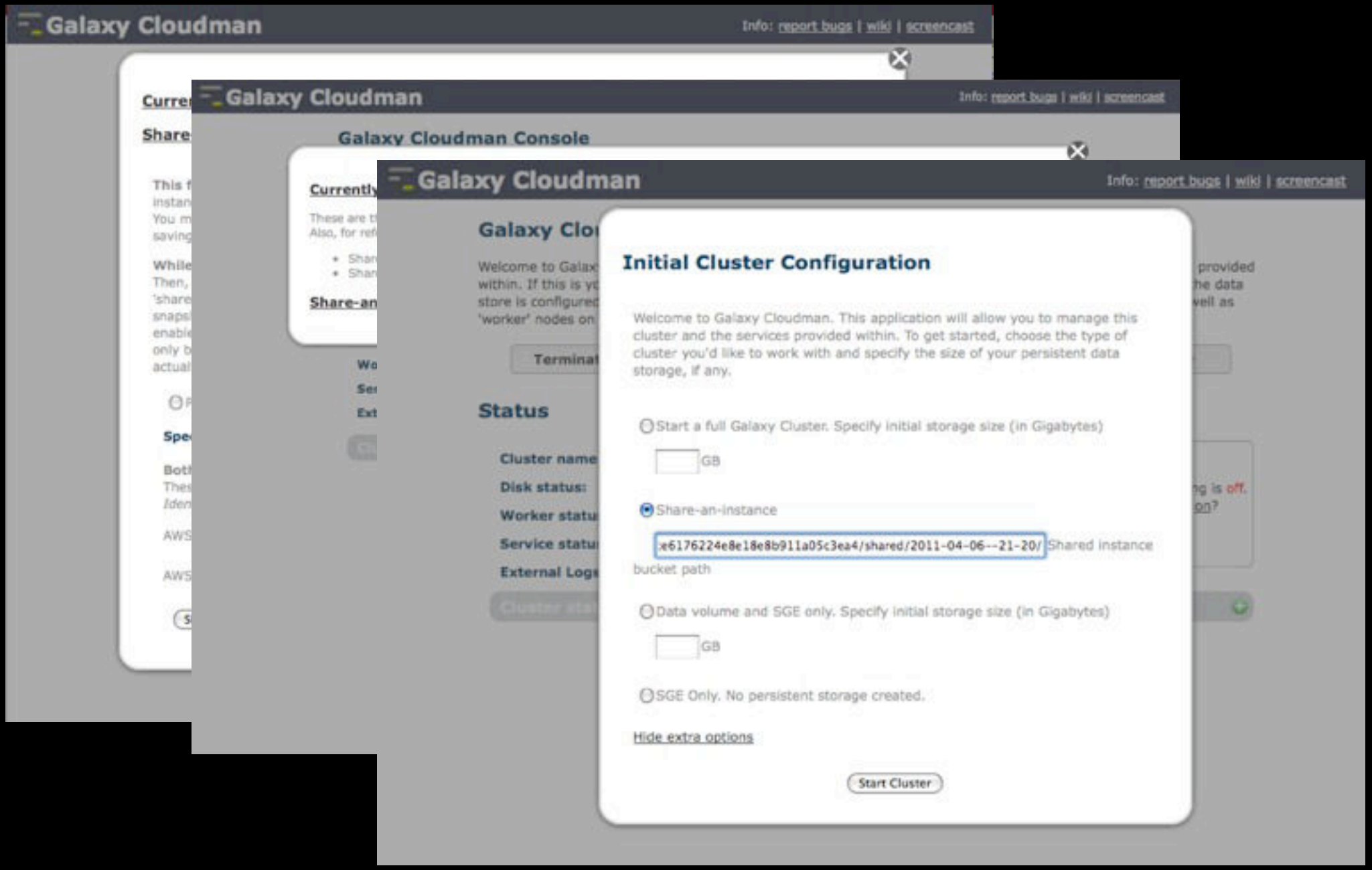
Deployment sharing



Deployment sharing



Deployment sharing



Use CloudMan as SaaS

Use CloudMan as PaaS

It’s automated, reproducible, extensible, and transparent.

EMORY

Enis Afgan

Dannon Baker

Dave Clements

Jeremy Goecks

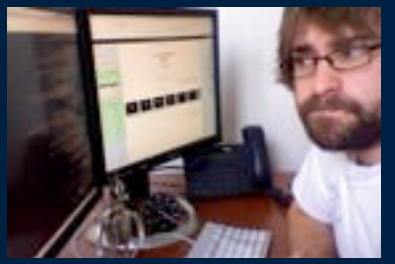
Kanwei Li

James Taylor

PENNSTATE

IBSS

Dan Blankenberg



Nate Coraor

Jennifer Jackson

Greg von Kuster

Kelly Vincent

Anton Nekrutenko

Supported by the **NHGRI** (HG005542, HG004909, HG005133), **NSF** (DBI-0850103), Penn State University, Emory University, and the Pennsylvania Department of Public Health