Transforming Science Through Data-driven Discovery

Atmosphere Overview

Intro Cloud Computing

Welcome to Atmosphere

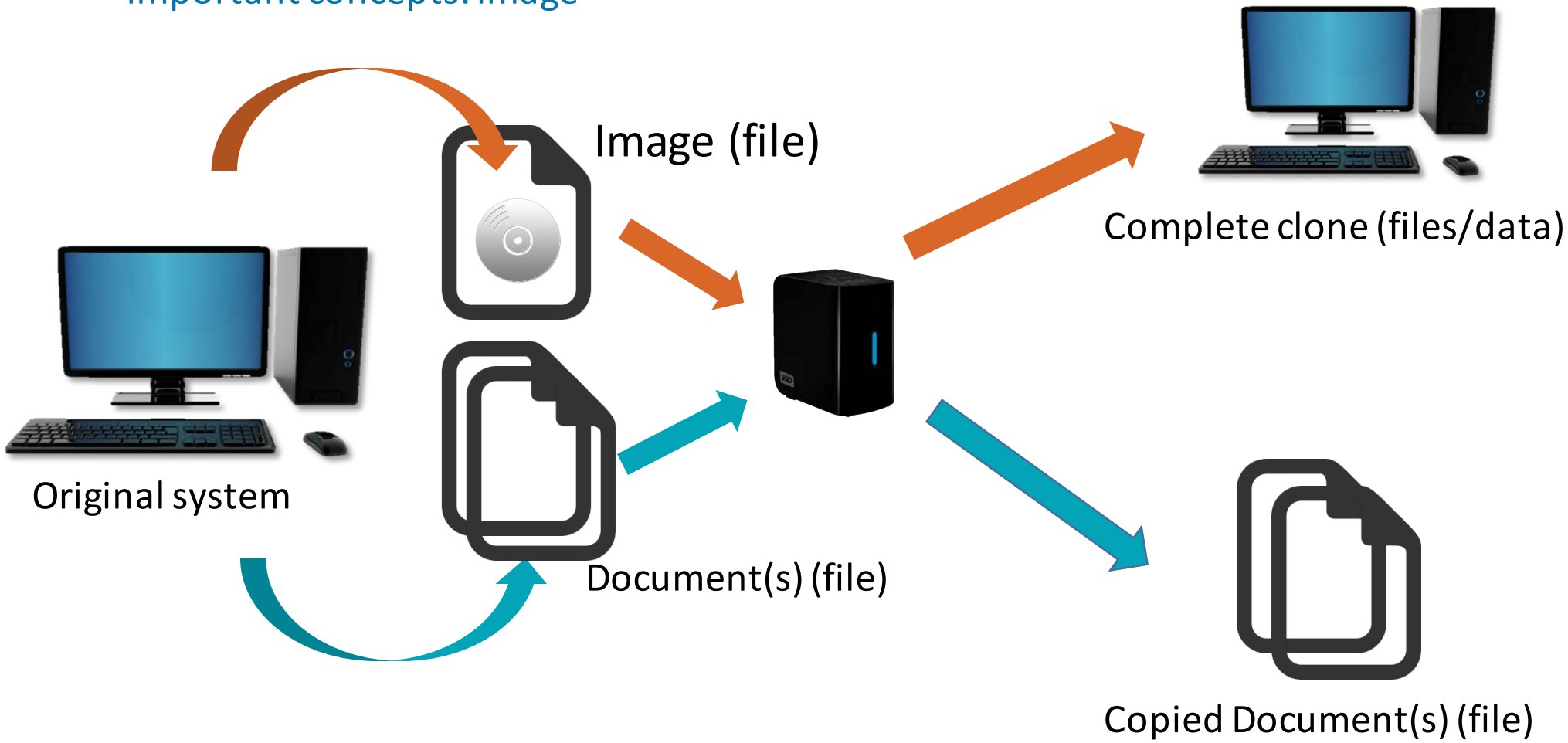
Custom Cloud Computing for Life Sciences

What is Cloud Computing?

Yet another round of jargon

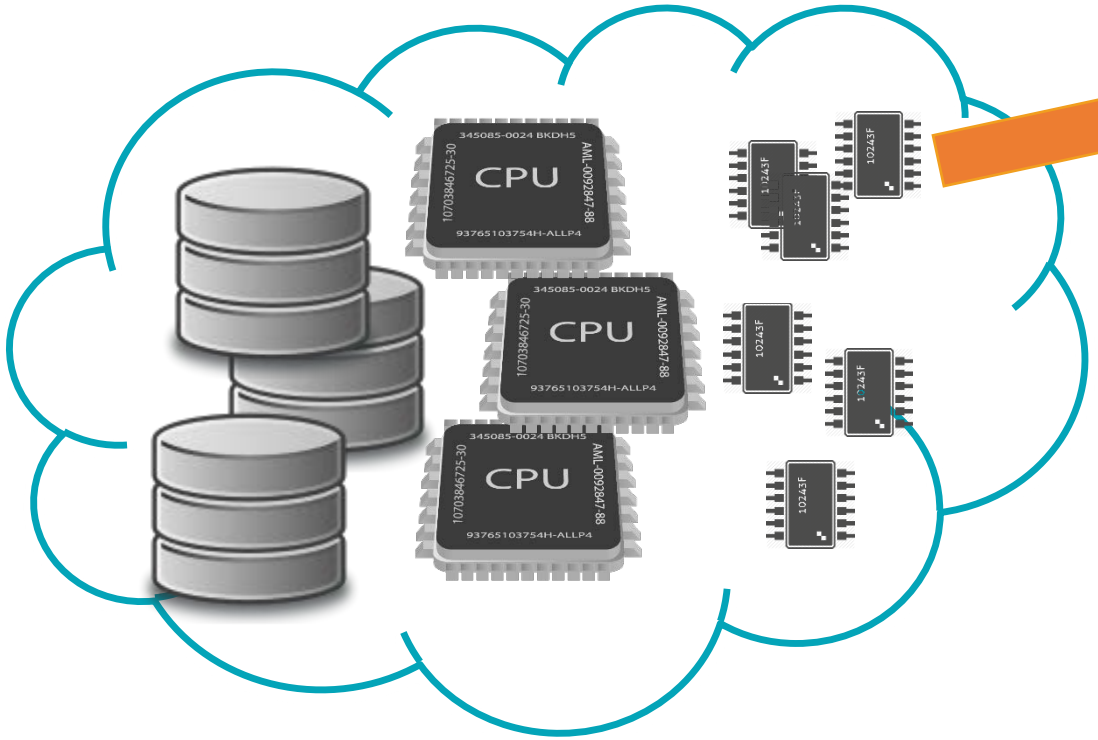
What is Cloud Computing?

Important concepts: Image

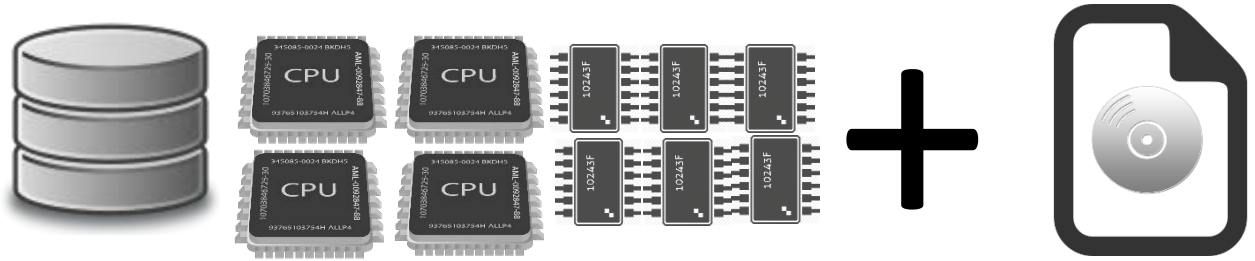


What is Cloud Computing?

Important concepts: Instance



CyVerse Cloud



(Disk + CPU + Memory) + (Image)

128.196.34.158

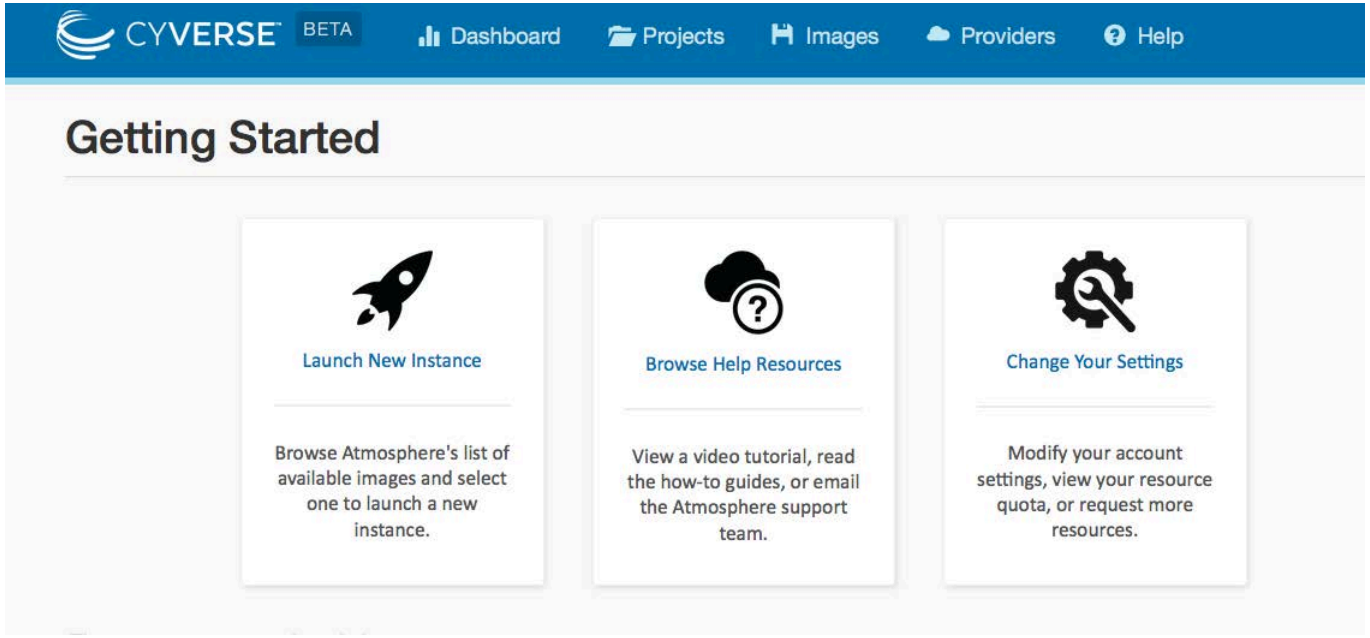
Atmosphere Instance (virtual machine)

What is Cloud Computing?

Yet another round of jargon

Atmosphere Overview

Largest, easiest to use cloud for Life Sciences



- Choose an existing image or customize
- Instances up to 16-Core / 128 GB RAM
- Access via shell or VNC
- Share you image with selected users, or make them public

Atmosphere Overview

Benefits

Get Science Done

- Work in an on-demand Linux environment (most bioinformatics)
- Collaborate with students and colleagues on the same instance

Reproducibility

- Make data, workflows, and analyses available in a public image
- Access previous software version and images

Productivity

- Multicore high memory images to run multithreading applications
- Move your analyses from your laptop to the cloud

Atmosphere Overview

Connecting to your instance

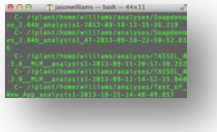
Windows

VNC Viewer

PuTTY

Mac

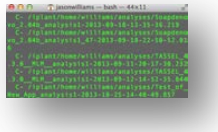
VNC Viewer



Shell/terminal

Linux

VNC Viewer



Shell/terminal

VNC Viewer: www.realvnc.com/download/viewer

PuTTy: www.putty.org

Atmosphere Overview

User perspectives and possible applications

Bench Scientist

- Learned how to use the shell and how to work with Linux
- Mastered using R to develop plots for his manuscript

Bioinformatician

- Launches an image and has full SUDO access to customize
- Developed a software with numerous R and Python library dependencies
- She updates it regularly by making a new image

Core Facilities

- Linked several atmosphere instances with Apache Hadoop
- Worked with CyVerse to import existing Amazon image

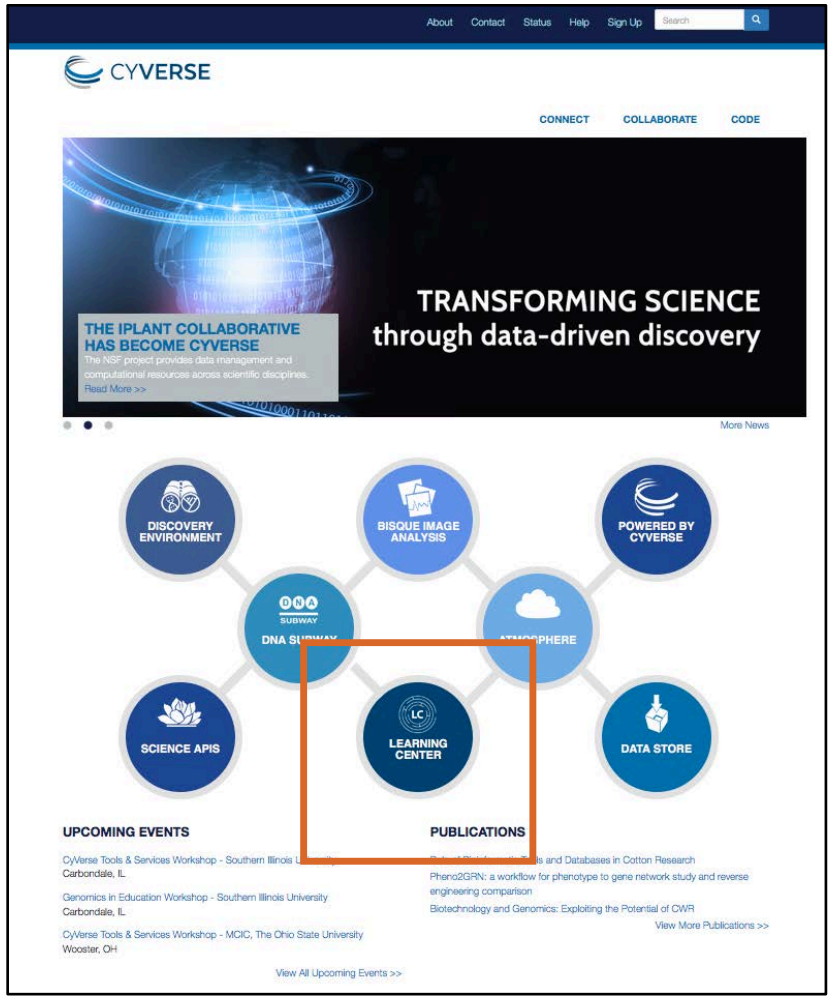
Images from personas based on: Bioinformatics Curriculum Guidelines: Toward a Definition of Core Competencies
PLOS Biology DOI: 10.1371/journal.pcbi.1003496

Atmosphere Overview

Key things to remember when you try this yourself

- Images do not have automatic access to your Data Store
  - Use Cyberduck to access the Data Store
  - Use iCommands
- Users have monthly allocation limits
  - Terminate, or stop instances not in use
  - If a larger allocation is needed, contact support
- All data on terminated instances will be destroyed
  - Use Cyberduck or iCommands to transfer data off the instance
  - You may also create an EBS Volume (see documentation)

Where to go from here:



Learning Center

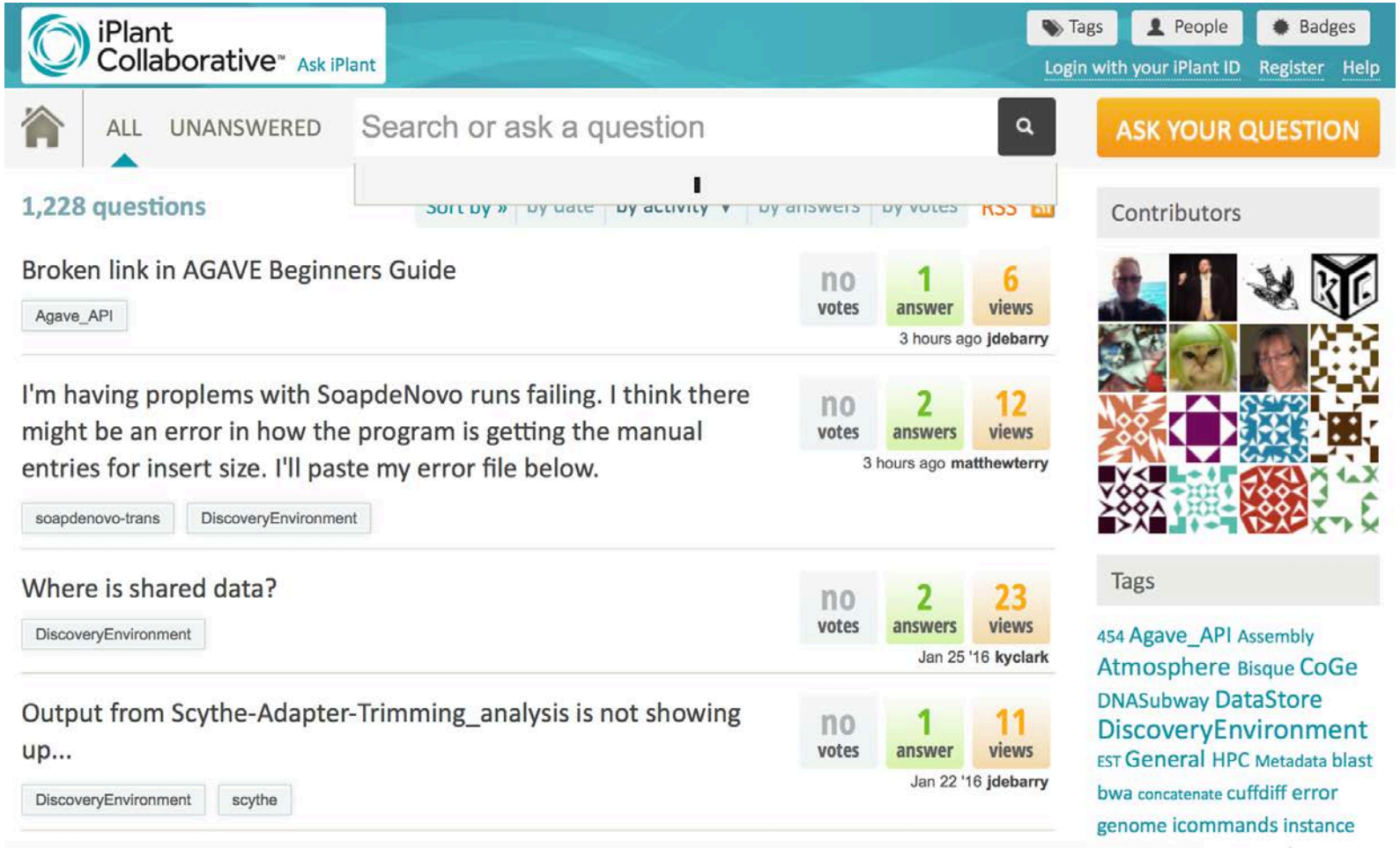
- Get Started Guide
- Tutorials and Videos
- Documentation

Upcoming Events

- Workshops
- Webinars

Keep asking: ask.iplantcollaborative.org

Detailed instructions with videos, manuals, documentation in Learning Center



CYVERSE

Transforming Science Through Data-driven Discovery

THE UNIVERSITY OF ARIZONA

Parker Antin
Nirav Merchant
Eric Lyons

TACC
TEXAS ADVANCED COMPUTING CENTER

Matt Vaughn

CSH Cold Spring Harbor Laboratory

Doreen Ware
Dave Micklos

NSF

CyVerse is supported by the National Science Foundation under Grant No. DBI-0735191 and DBI-1265383.