CYVERSE™

Transforming Science Through Data-driven Discovery

Cyverse Science APIs

John Fonner, @johnfonner
Matt Vaughn, @mattdotvaughn
#cyverse #agaveapi #usetacc

THE UNIVERSITY OF ARIZONA.

TEXAS ADVANCED COMPUTING CENTER

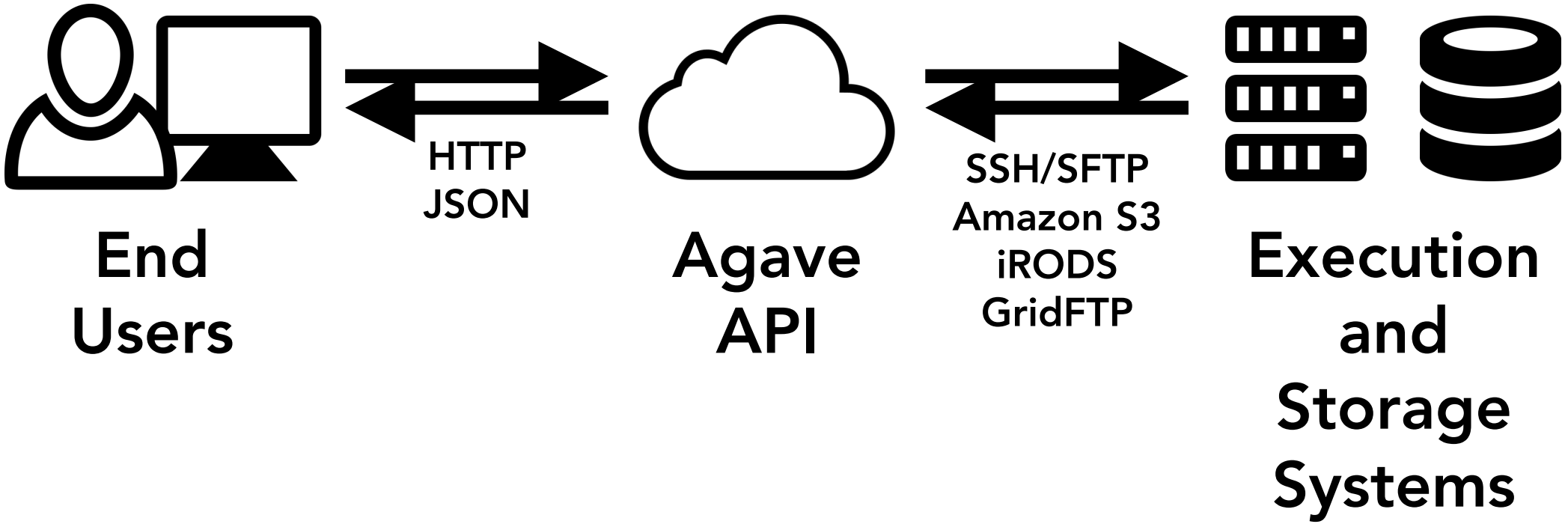
Cold Spring Harbor Laboratory

Rapid discussion questions

- For a typical workflow, how many steps are involved? How many compute systems do you use to run it?
- APIs let systems and services talk to each other. Can you think of examples of websites where content/capabilities of one site appear on a different site?
- What does it look like to execute a workflow that runs on multiple compute systems?

What is an API?

Application programming interface



What is an API?

Important concepts: endpoints

clients	Create and manage API keys (WSO2)
apps	Register and manage apps
files	Move and manage data
jobs	Run and manage jobs
meta	Create and manage metadata
monitors	Create and manage system monitors
notifications	Subscribe to and manage notifications
postits	Create pre-authenticated, disposable URLs
profiles	Create and manage application users
systems	Register and manage systems
tenants	List available tenants
transforms	Transform and stage data
usage	Query for usage across APIs

Agave API Overview

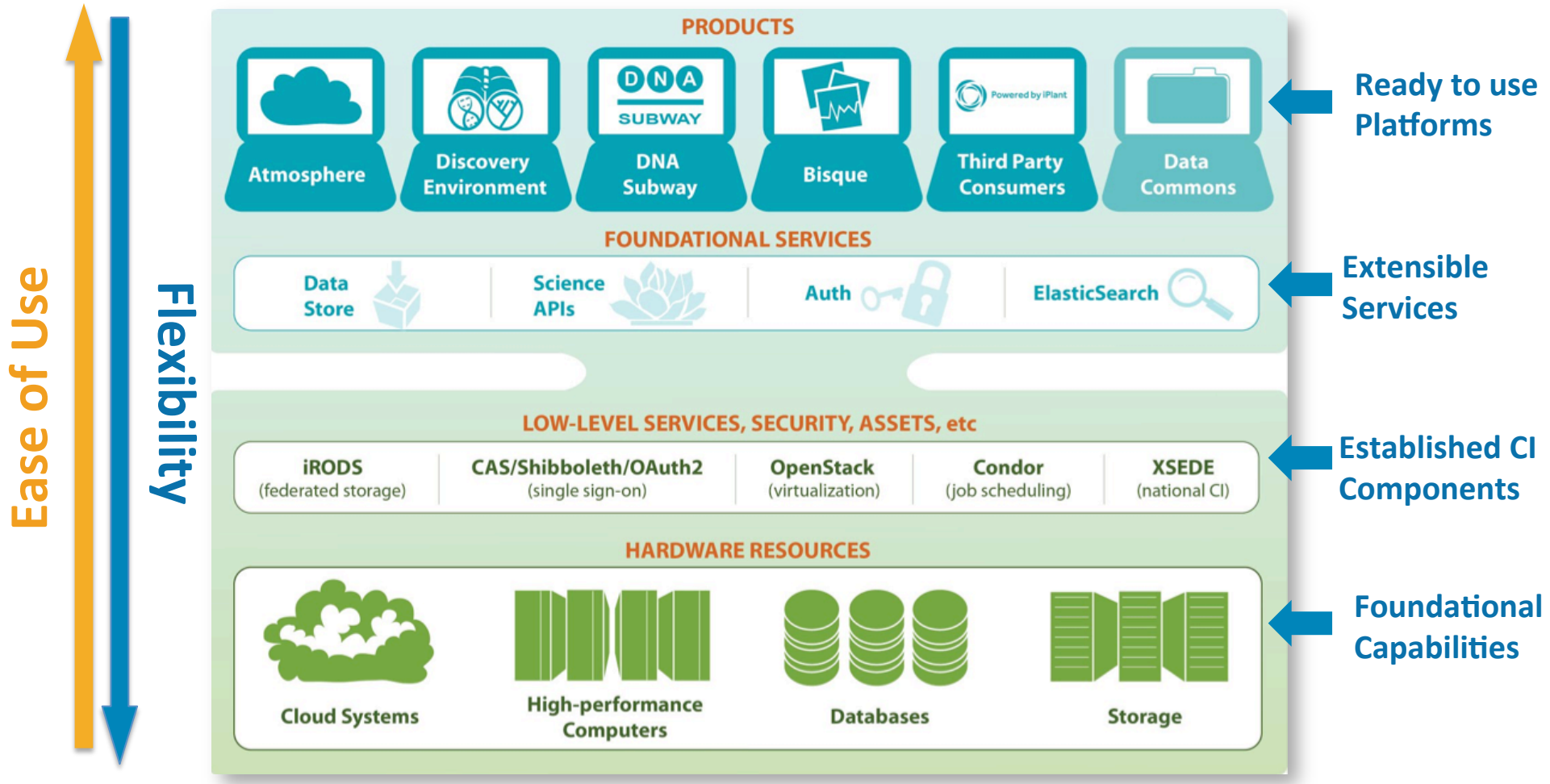
Reproducible scientific computing infrastructure

- Science-as-a-service platform
- Use your own compute, and storage resources (or iPlant's)
- Clone existing apps and register your own
- Share your apps and access them through the DE

Agave API Overview
Benefits

Get Science Done 	• Securely make your cluster and apps available on the web • Integrate the iPlant datastore and apps into command line scripts
Reproducibility 	• All files, jobs, and apps are tracked with unique IDs • Programmatically associate metadata with your files/jobs
Productivity 	• Automate workflows across sites and architectures • Create powerful new web portals quickly and cheaply

CyVerse product stack



Agave API Overview

Key things to remember when you try this yourself

- The command line interface (CLI) tools are great for users with basic Linux knowledge
  - http://agaveapi.co/tools/command-line-interface/
- Computers love reading JSON, humans may not
  - Catch syntax errors with a lint tool: http://jsonlint.com/
  - You can paste JSON into the Agave ToGo web app (or use the wizard): http://togo.agaveapi.co/app/#/apps/new
- Data on storage systems, apps on execution systems
  - Even the files associated with an app reside on storage systems
  - App inputs can be URLs or files on storage systems

Agave API Overview

User perspectives and possible applications

Bench Scientist

- Learned how APIs work
- Uses the Agave CLI Tools in the shell

Bioinformatician

- Integrated API calls into analysis scripts to handle data flow
- Made tools on our lab cluster available through the Discovery Environment

Core Facilities

- Used the “files” and “postits” to share data with collaborators
- Built a custom web interface for browsing data and metadata

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

Key take home knowledge

- APIs are tools that let developers and power users wire together web services to provide custom functionality.
- iPlant lets **any** user integrate their own compute systems and apps into the Discovery Environment through the Agave API.
- The Powered by iPlant initiative actively engages with developers to help them build community facing web portals using iPlant API services.

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