Free Beer Today

How iPlant + Agave + Docker are changing our assumptions about reproducible science

John Fonner (@johnfonner)
Texas Advanced Computing Center
BOSC 2015

This is a whale free conference.

It has been

1

presentations since the last incident

This is a whale free conference.

It has been

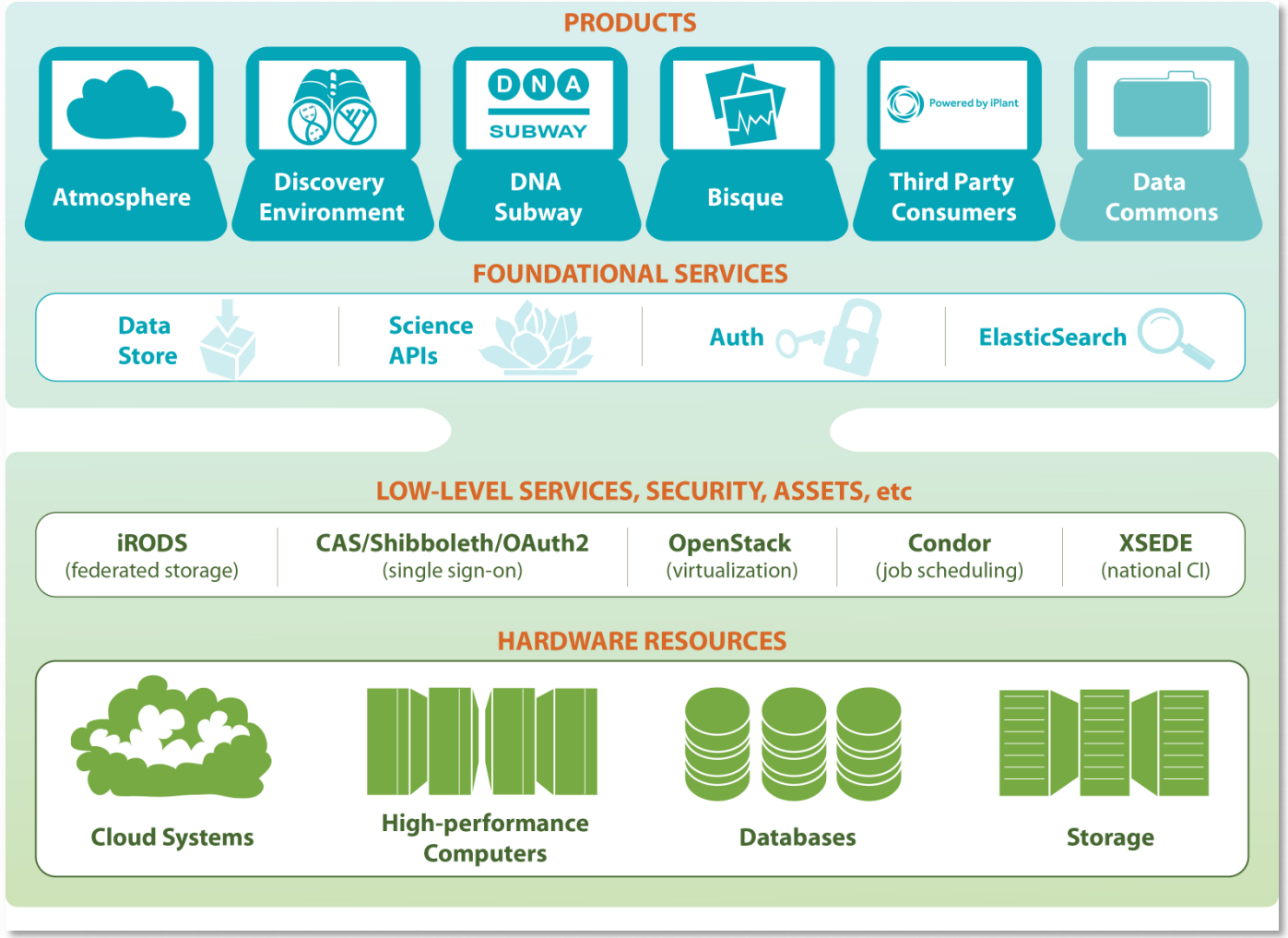
presentations since the last incident

What is iPlant?

www.iplantcollaborative.org

- iPlant's mission is to build the Cyberinfrastructure to support ~~plant~~ biology's Grand Challenge solutions
  - Funded by the National Science Foundation as a $50M "five year mission" to build a virtual organization of experts in ~~plant~~ biology, computing, statistics, etc.
  - Long term focus on sustainable food supply, climate change, biofuels, ecological stability, etc.
  - Renewed in September 2013
- Led by U. of Arizona, TACC, and Cold Spring Harbor Labs; Hundreds of participants globally - Working group members at >50 US institutions, USDA, DOE, etc.

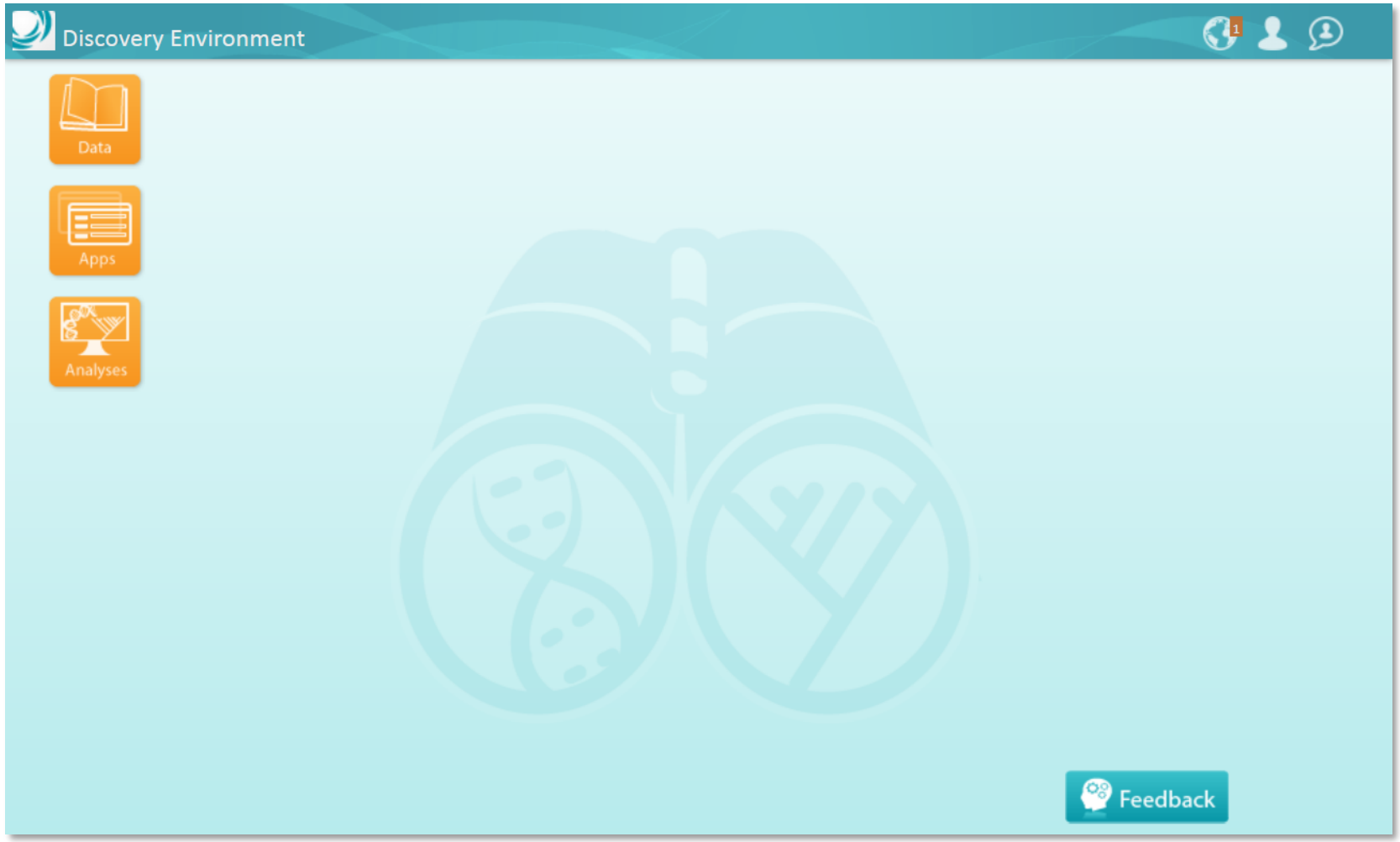
iPlant Cyberinfrastructure



- Web-based user interfaces for accessibility
- API level services for CLI and third party services
- Hardware resources built on the national cyberinfrastructure

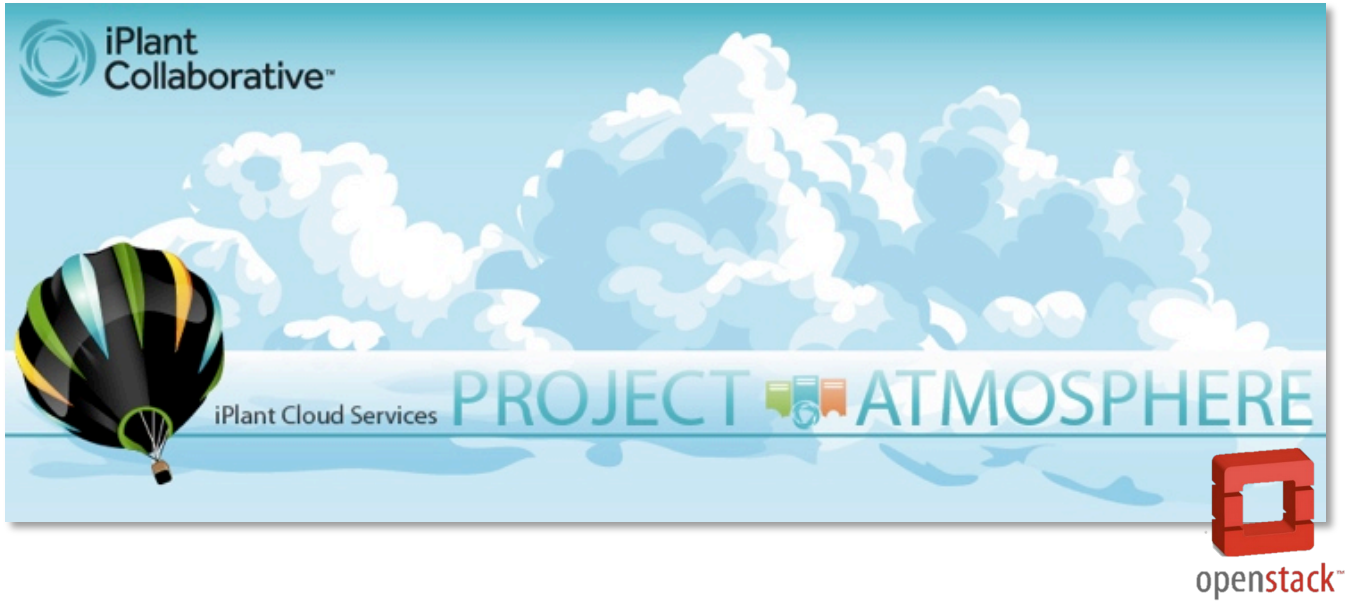
Discovery Environment

de.iplantc.org



Atmosphere

atmo.iplantc.org



- Largest, easiest to use open cloud for Life Science
- Choose an existing image or customize
- Instances up to 16-Core/128 GB RAM
- Share your image with selected users, or make them public

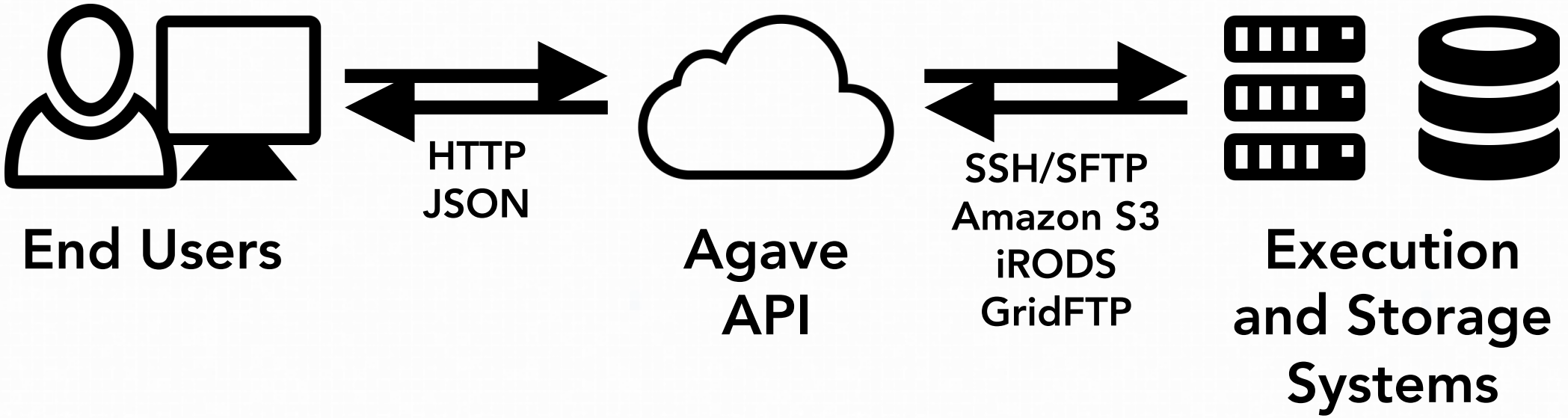
What is the Agave API?

agaveapi.co

- 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

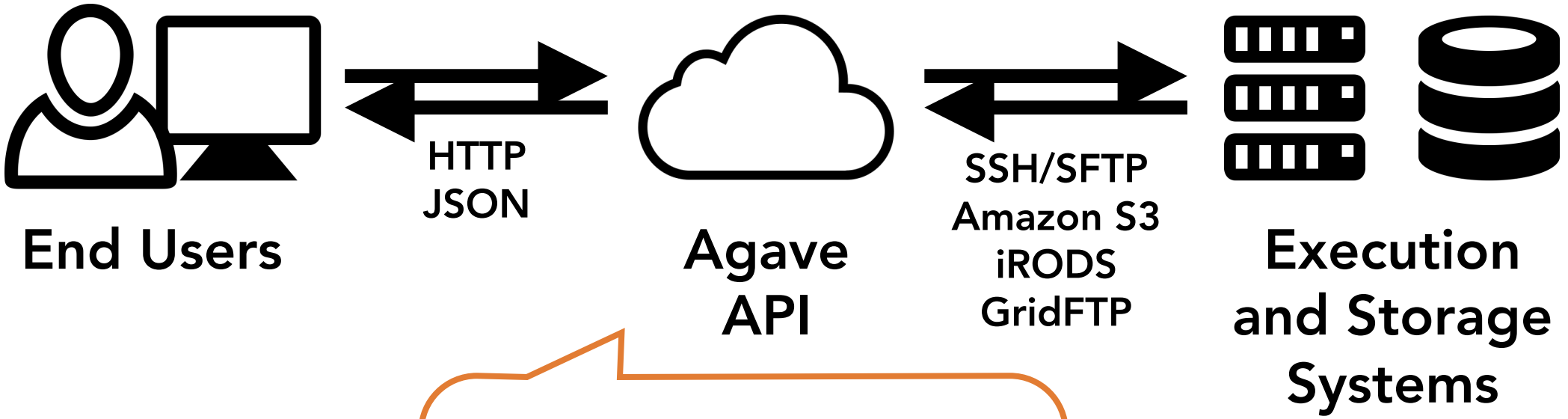
What is an API?

Application programming interface



What is an API?

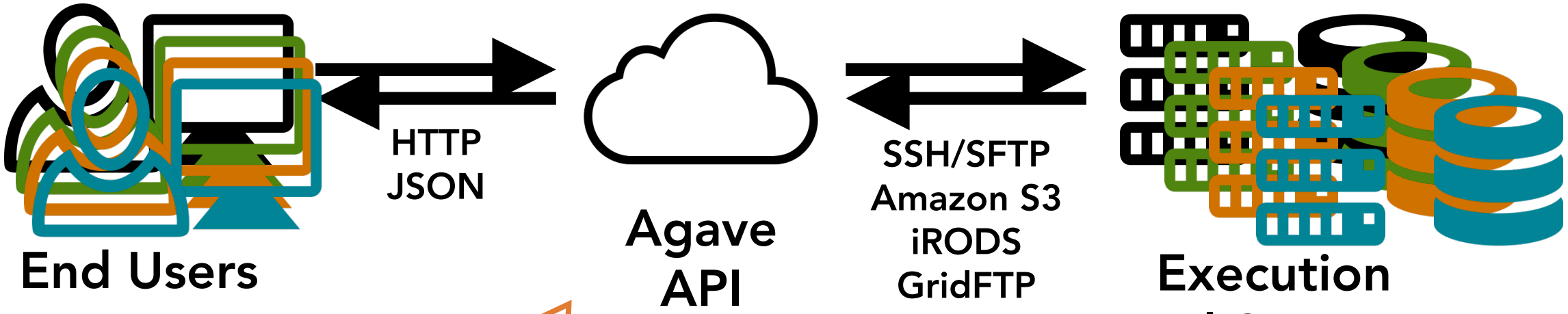
Application programming interface



authentication systems
apps jobs files
notifications metadata
postits

What is an API?

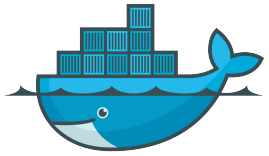
Application programming interface



authentication systems
apps jobs files
notifications metadata
postits

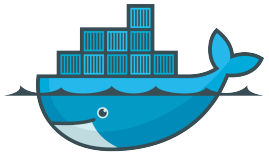
bit.ly/agave-togo

bitbucket.org/
taccaci/foundation-cli



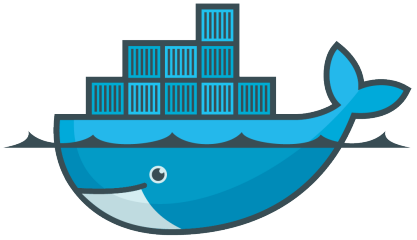
Docker

Friend of developers… and users



Docker

Friend of developers… and users

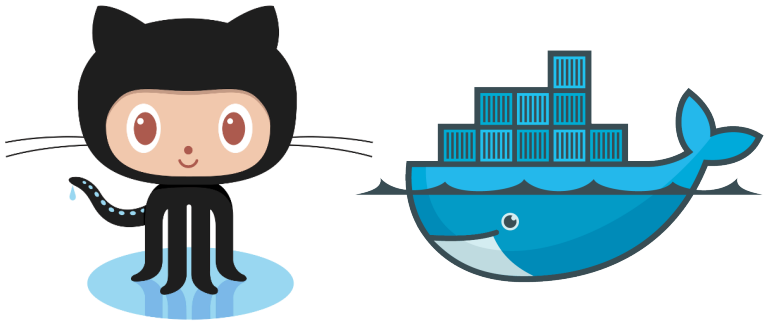


WHAT'S YOUR
POINT?

DID

Reproducibility

technology + convention



- open source
- version control
- app distribution
- runnable code
- sample data
- automated testing

- version control
- job provenance
- sample data
- platform integration

- fine-grain sharing
- cloud/cluster execution
- graphical user interface
- platform integration

Reproducibility

Is this is good as it gets?

…Nope

- More Docker support on advanced computing resources
- iPlant Data Commons
- Single click workflow re-execution
- Better OS support in Docker
- https://docs.docker.com/articles/dockerfile_best-practices/

Thanks!

Slides available at github.com/johnfonner/presentations

Email: jfonner@tacc.utexas.edu

Twitter: @johnfonner

iPlant Collaborative™

The iPlant Collaborative is funded by a grant from the National Science Foundation Plant Cyberinfrastructure Program (#DBI-0735191).