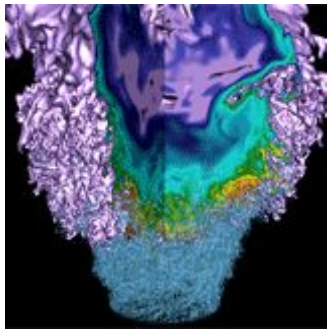
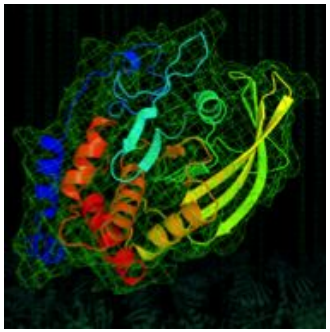
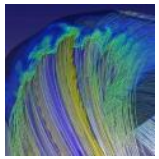
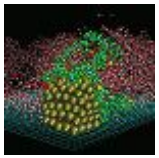
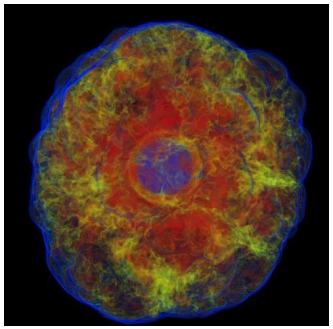
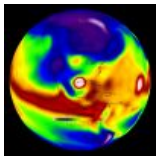
Introduction to Spin


Tony Wildish

On behalf of the Spin working group

What problem are we solving?

People need a place to run all sorts of services, not just batch jobs.

- Workflow managers that manage lots of batch jobs
- Database servers to record results
- Web servers to publish information
- Other services that integrate with a scientific collaboration across the WAN

Previously, these services would be on dedicated hardware, with custom installations, or installed ad-hoc in user-space

Spin gives these services a proper home

What is Spin?

Spin* is a system to **deploy Docker containers** in a manageable environment.

A **flexible, scalable**, platform, **tightly integrated** with NERSC resources:

- Develop on your laptop; **deploy in minutes**
- **Scale out** for performance
- Access **HPC networks and file systems**
- NERSC manages everything under the hood

* Scalable Platform Infrastructure at NERSC

What do you get out of this?

- Focus on developing your app, not on managing services
- Create a new database, web portal or other app on demand
- Ignore the details of the underlying infrastructure
  - Everyone cares about **services** not **servers**
  - System packages won't interfere with your packages
- Get scalability and fault tolerance for free
- Standard application images. Easy application updates

The benefits of VMs without the overhead.

The benefits of cloud with tight integration to NERSC.

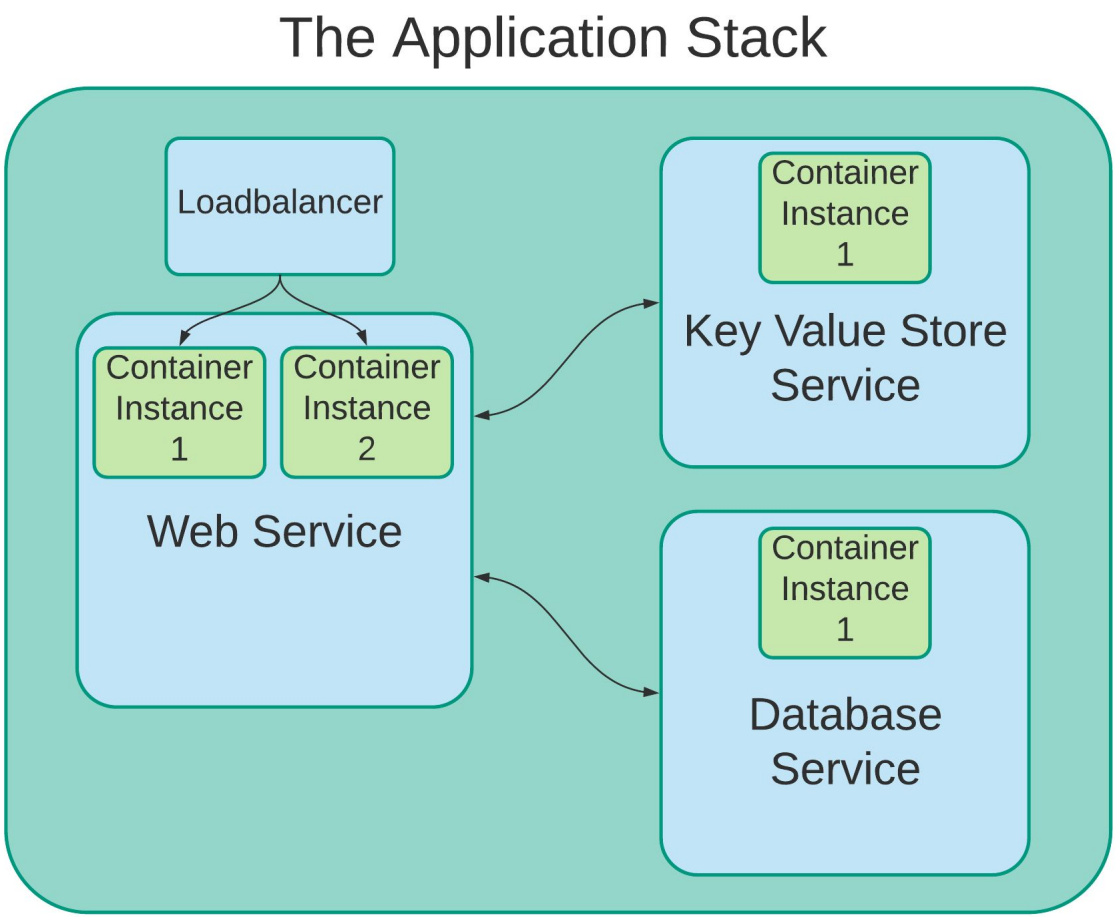
What is Docker?

- Docker:
  - A tool for building and running **containers**
  - Easy to learn the basics, good documentation
  - Large ecosystem, many good examples to borrow from
- Container:
  - lightweight, portable package for software
  - Isolates you from the operating system
  - Build on your laptop, run identically on NERSC
- More flexible than virtual machines
- Very efficient use of resources

Container, image, service, stack

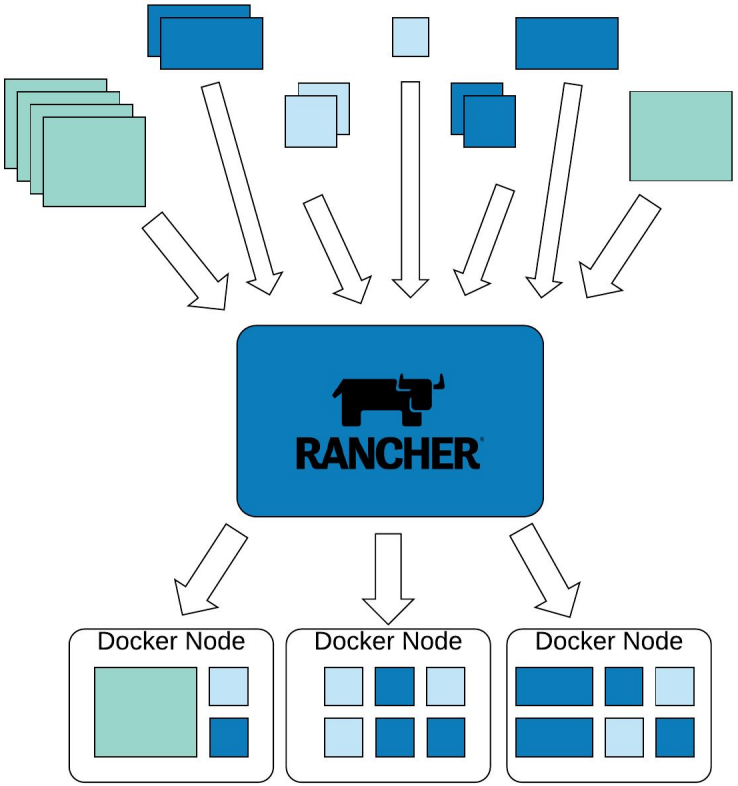
- An **Image** is a lightweight, stand-alone package that includes everything needed to run a piece of software
- A **Container** is a runnable instance of an image and provides one service
- A **Service** is comprised of one or more containers, each of which provide a single capability. For example:
  - A web service with two containers
  - A database service with a single container
- One or more services form an **Application Stack**

Container, image, service, stack



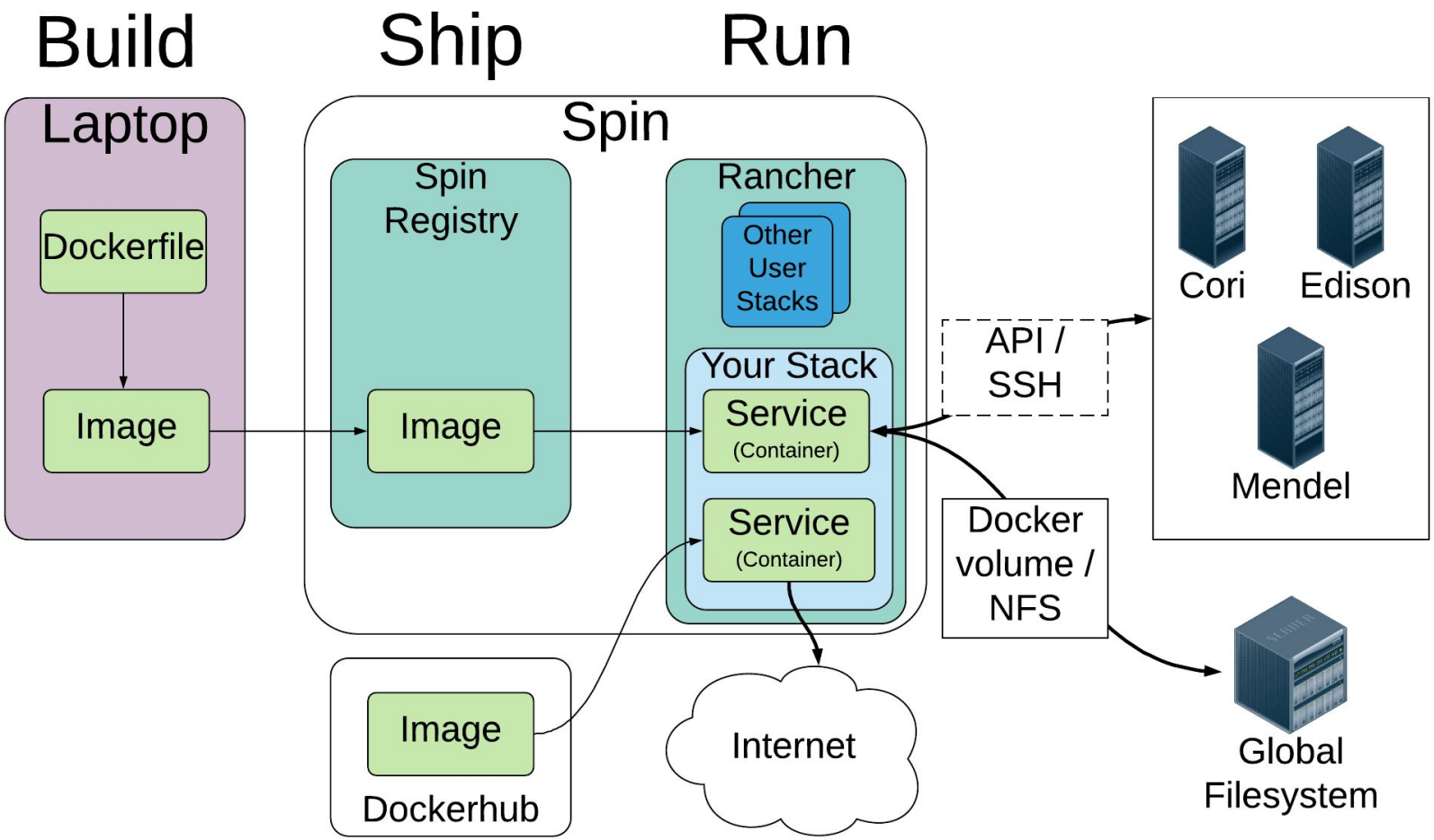
Managing multiple stacks

Without orchestration, a pool of servers and no coordination for users



Managed and assigned to Docker nodes, enabling holistic management, failover, service ownership.

Docker Philosophy: Build an Image, Ship it to a registry & Run it in Spin



Current users

- **Several Early Adopter projects are running on Spin**
  - CRD DST sciencesearch-dev.nersc.gov (Prod coming soon)
  - ESS-DIVE Data Archive collab between NERSC & CRD
  - Joint Genome Institute, several portals, DBs
- **NERSC services**
  - License servers (Cray Intel, Cray CCE, PGI, etc)
  - NERSC Jupyter, RStudio, Data Portals
  - Prototypes: pytokio, User MFA
  - Edge Services: Frontier Cache & CVMFS
  - CSGDOC, consult-history support archive
  - Shifter registry

Current status

- Spin is currently in a pilot mode
  - Internal projects at NERSC
  - Several users working directly with NERSC staff
  - Will soon be open to trained users

- Our first "SpinUp" hands-on training session is planned for mid-May
  - Capacity will be limited at first
  - The application will be available at the web site
  - Watch the weekly newsletter for announcements

Early-adopter pilot program

- Do you have a service that looks like it fits the bill?
  - Talk to us about your needs
  - nersc-spin@lbl.gov

- Getting started:
  - Learn Docker (several tutorials online)
  - Read the NERSC Spin documentation at www.nersc.gov