JGI JOINT GENOME INSTITUTE
A DOE OFFICE OF SCIENCE USER FACILITY

JGI Data Services at NERSC

Alicia Clum & Georg Rath
2019-07-19

What is the Joint Genome Institute?

- **A U.S. Department of Energy Office of Science User Facility**
  - Walnut Creek, CA facility opened in 1999
  - ~280 staff
  - ~$70M annual funding
  - Services used by 1,598 DOE affiliated researchers in 2017
- **An Experimental and Observational Data Facility**
  - DNA sequencing and other advanced genomic technologies
  - Computational Analysis
  - 75 Million core hours in 2018

Facility Software

- **User-Facing Web Portals and APIs - “Science Gateways”**
  - Complex applications
  - Integrate with compute and storage infrastructure (eg scheduler, parallel FS)
  - Need supporting infrastructure (eg databases, virtual machines)
- **Workflow Managers**
  - Used for internal management of workflows
  - Automatically rerun job steps and perform job “packing”
- **Instrument support**
  - Drives instruments
  - Pre- and post-processing of data

The Mendel Years

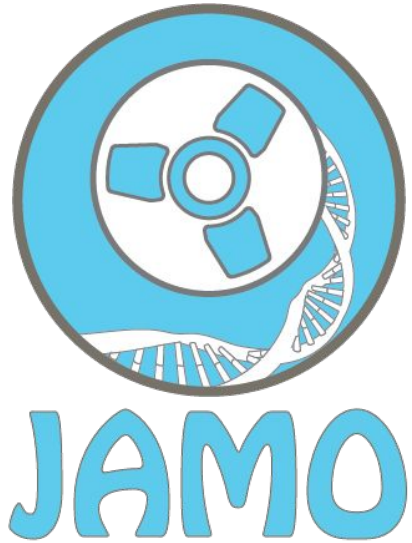
- **Mendel is a meta-commodity-cluster containing**
  - PDSF
  - Matgen
  - Genepool and Denovo
- **As well as service nodes**
  - interactive
  - web
  - database
  - and everything that did not fit elsewhere
- **Not part of Cori maintenance schedule**
- **Mendel will retire on July 26th**

A New World Order

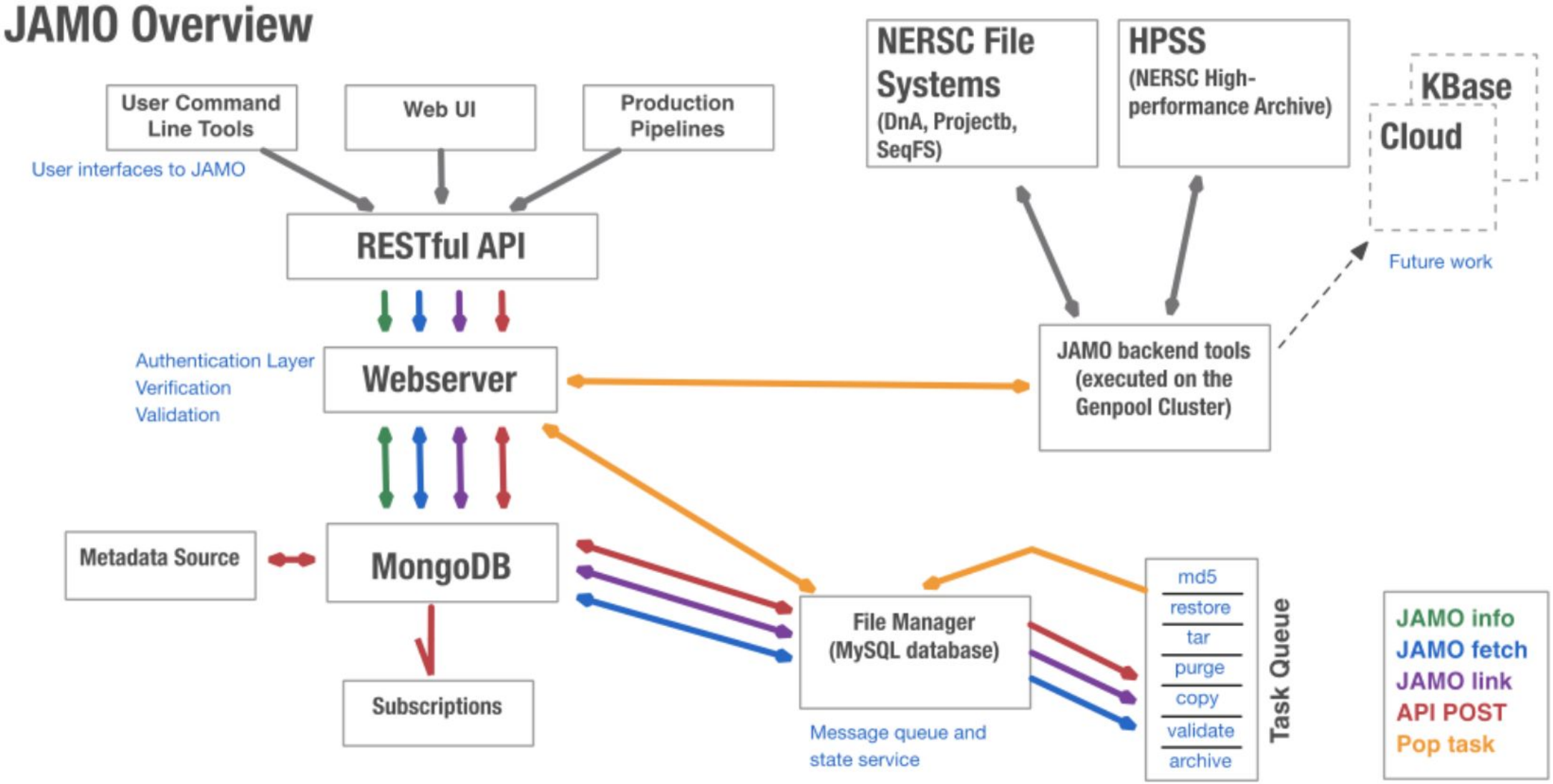
- **One batch system to rule them all**
  - Cori only compute system at NERSC
  - Usual HPC machine lifecycle
- **Migration to replacement infrastructure**
  - Compute: Denovo to Cori
  - Storage: GPFS/local disk to Lustre/GPFS/NFS/DVS/DataWarp
  - Services: Spin (Container CaaS), VMware (VM IaaS)
  - Interactive: Login Nodes

JAMO - JGI Archive and Metadata Organizer

- JGI data management middleware
- Takes care of data movement
- Stores Metadata
- Interfaces with all NERSC storage systems



JAMO - Architecture



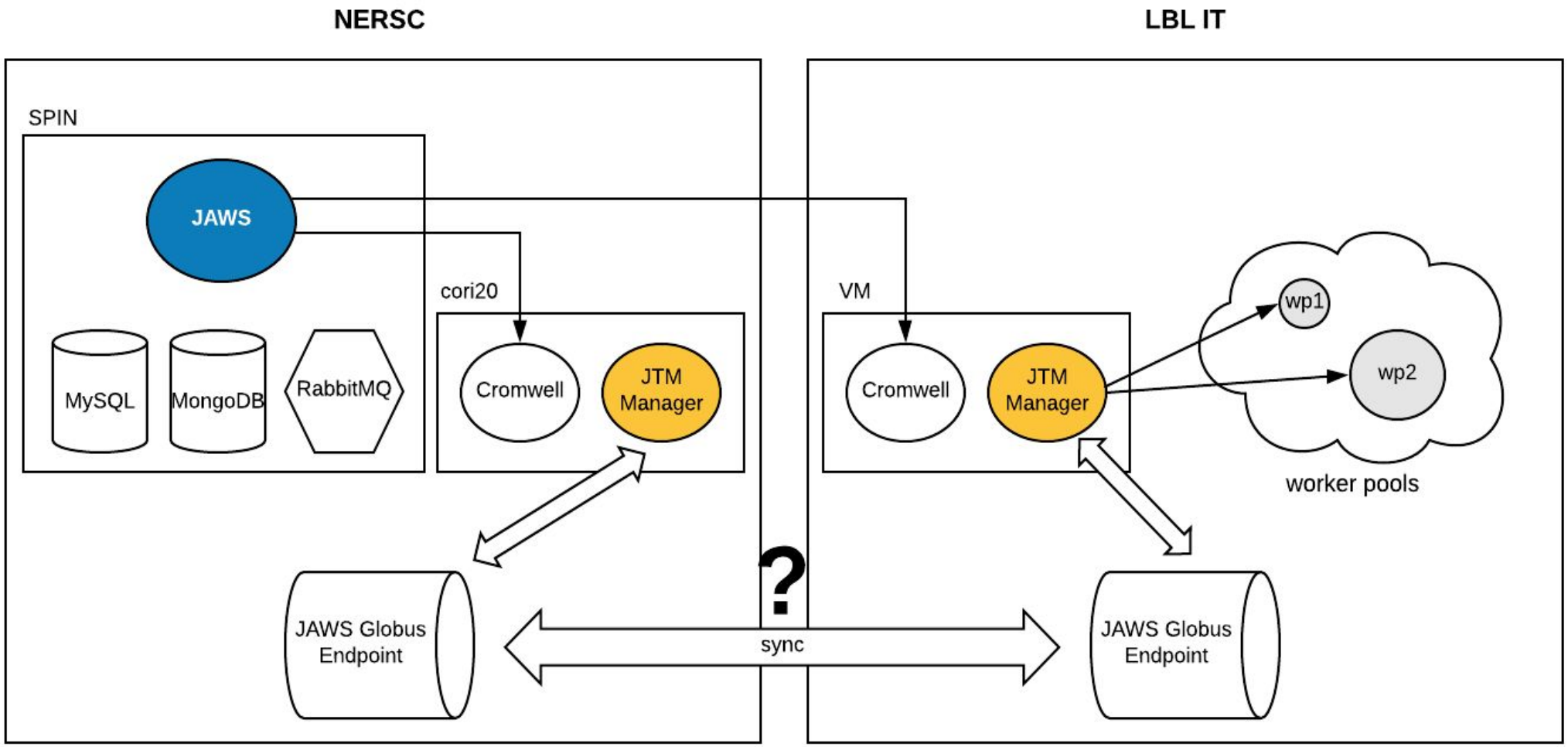
JAMO - JGI Archive and Metadata Organizer

- **Collaboration between infrastructure group and developers**
- **Migrated to virtual machines**
  - effort to port to container paradigm prohibitive
  - footprint of database servers too big
- **Move services off parallel filesystems where possible**
  - database corruption
  - general instability
- **Access to filesystem via DTN agent**
- **Effort: several months**

JAWS - JGI Analysis Workflow System

- **Workflow middleware**
- **Abstraction between compute resources and user**
  - Accepts workflow definition language (WDL) workflows
  - Enables efficient workflows
  - Packs many small tasks into bigger jobs (“pilot jobs”)
  - Enables sharing of analysis pipelines

JAWS - Architecture



JAWS - Deployment

- Deployed largely in Spin
- Access to scratch filesystem through cori20
- Access to scheduler through cori20
- Uses provided RabbitMQ service
- Needs to submit jobs as user

Experiences - Spin

- Learning curve is medium to high
- Documentation is extensive, consultants very responsive
- Required three day training session to get access is excessive
- No UI
- Off-the-shelf containers sometimes do not work
  - security restrictions
- Debuggability lacking
  - hard to access logs
- No resource monitoring (cpu, memory, network, disk)
- No APIs accessible
- Not possible to integrate with CI/CD systems
- Still quite a few humans in the loop
- Scheduler access needs workarounds
- Filesystem access* needs workarounds

Experiences - Virtual Machines

- **No learning curve**
- **Not an official service**
- **Human in the loop**
  - create/start/stop/snapshot machines requires ticket
  - installing software requires ticket
- **No resource monitoring (cpu, memory, network, disk)**
- **No APIs accessible**
- **Scheduler access needs workarounds**
- **Filesystem access* needs workarounds**

Organizational Impact

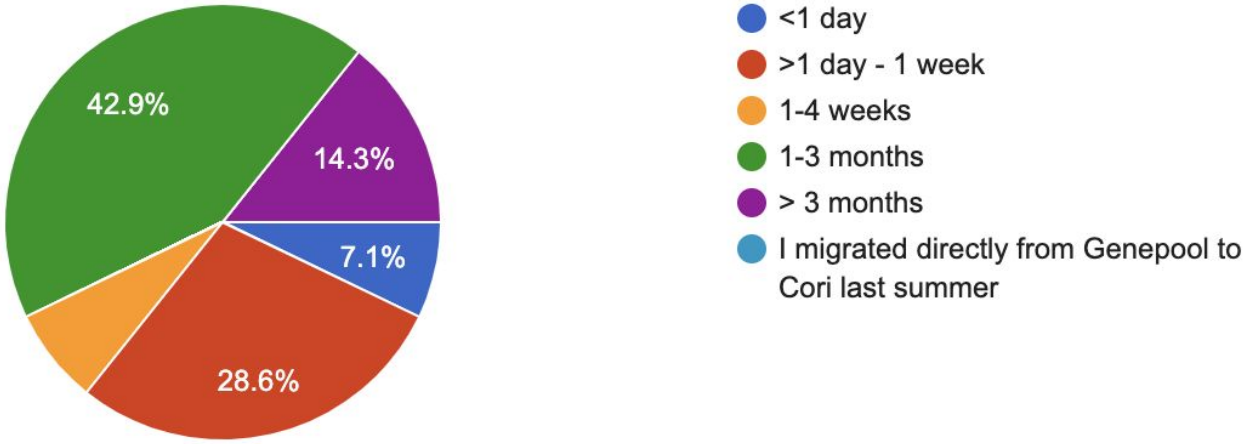
- **Integration effort is non-trivial**
  - Development of APIs necessary
- **Adaptation of existing services to new infrastructure has substantial impact**
- **Stable* infrastructure key to high scientific output**
- **Reliability matters**

Organizational Impact - cont.

- **10 percent of JGI institutional milestones this year were around migration or application readiness**

How much time did you spend migrating pipelines and/or services or exploring alternative software if your code wasn't able to run on Cori?

14 responses



Recommendations/Wish List

- **Let users control their infrastructure**
  - self-service, no humans in the loop
  - provide a flexible substrate
- **Let users see their infrastructure**
  - telemetry and logs accessible to users
- **Provide APIs everywhere**
  - as close to “industry standard” as possible
  - make integration easier (scheduler, data movement, VMs,...)
  - integration does not end at the NERSC border
- **Provide PaaS offerings (eg databases, message queues, etc)**
- **Explore filesystems with scalable access mechanisms (eg object stores)**
- **Consistent experience across the organization**
  - provide the same semantics everywhere (eg filesystems)

JGI JOINT GENOME INSTITUTE
A DOE OFFICE OF SCIENCE USER FACILITY

Questions?