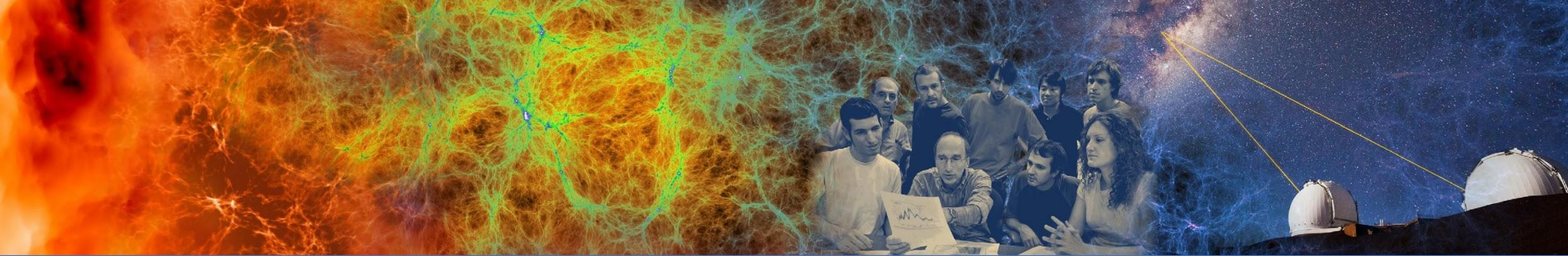
Prototyping a Cromwell Parsl Backend

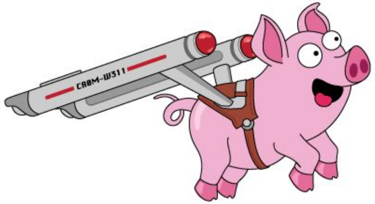
Kelly L. Rowland, Ph.D.
NERSC Data Science Engagement Group
ParslFest 2020
October 7, 2020



Cromwell

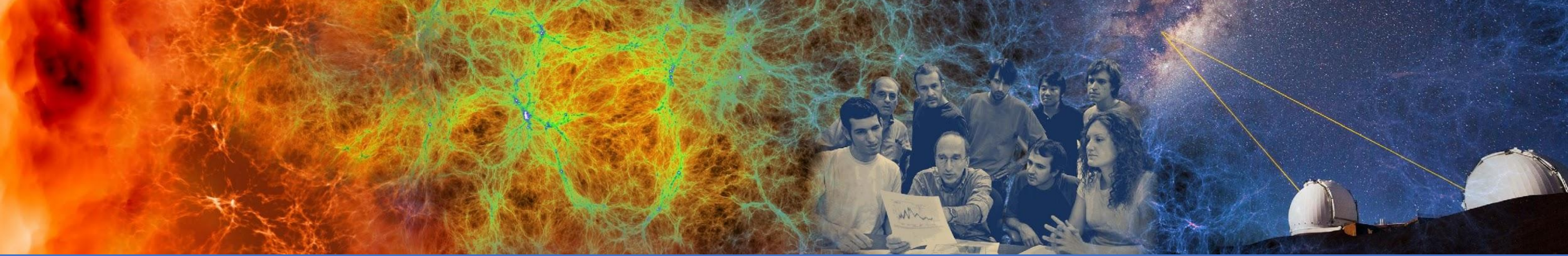
Cromwell

- “Workflow Management System geared towards scientific workflows”
- Developed by Broad Institute of MIT and Harvard
    - Biomedical and genomic research center
- Workflow input defined using one of:
    - Workflow Description Language (WDL)
    - Common Workflow Language (CWL)
- Runs on a multitude of platforms
    - Local instances
    - HPC schedulers - Slurm, SGE
    - Cloud providers - AWS, GCP
    - How? **Scripts are generated and sent to backends.**



Cromwell Backends

```
backend {
  providers {
    example_backend {
      config {
        <options>
        submit = "/usr/bin/env bash ${script}"
      }
    }
  }
}
```



Parsl Backend Prototype

Cromwell Backend

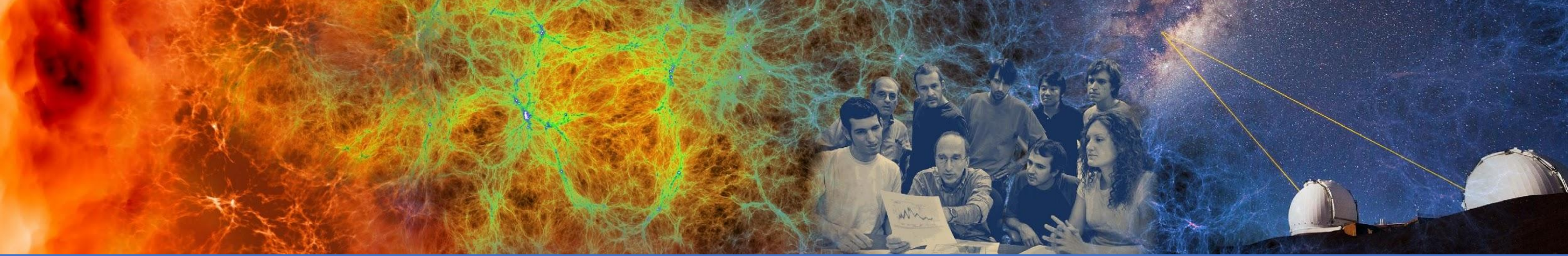
```
backend {
  providers {
    example_backend {
      config {
        <options>
        submit = "/usr/bin/env bash ${script}"
      }
    }
  }
}
```

Cromwell Backend, but Make it Parsl

```
backend {
  providers {
    parsl_backend {
      config {
        <options>
        submit = "python parsl_test.py ${script}"
      }
    }
  }
}
```

parsl_test.py

```
<Parsl imports>
import sys

script = sys.argv[1]

config = Config( <htex, label='cori', SlurmProvider> )

load(config)

@bash_app(executors=['cori'])
def run_script(script):
    cmd = '/usr/bin/env bash ' + script
    return cmd

script_out = run_script(script)
```



JAWS (or: “Why are you doing this?”)

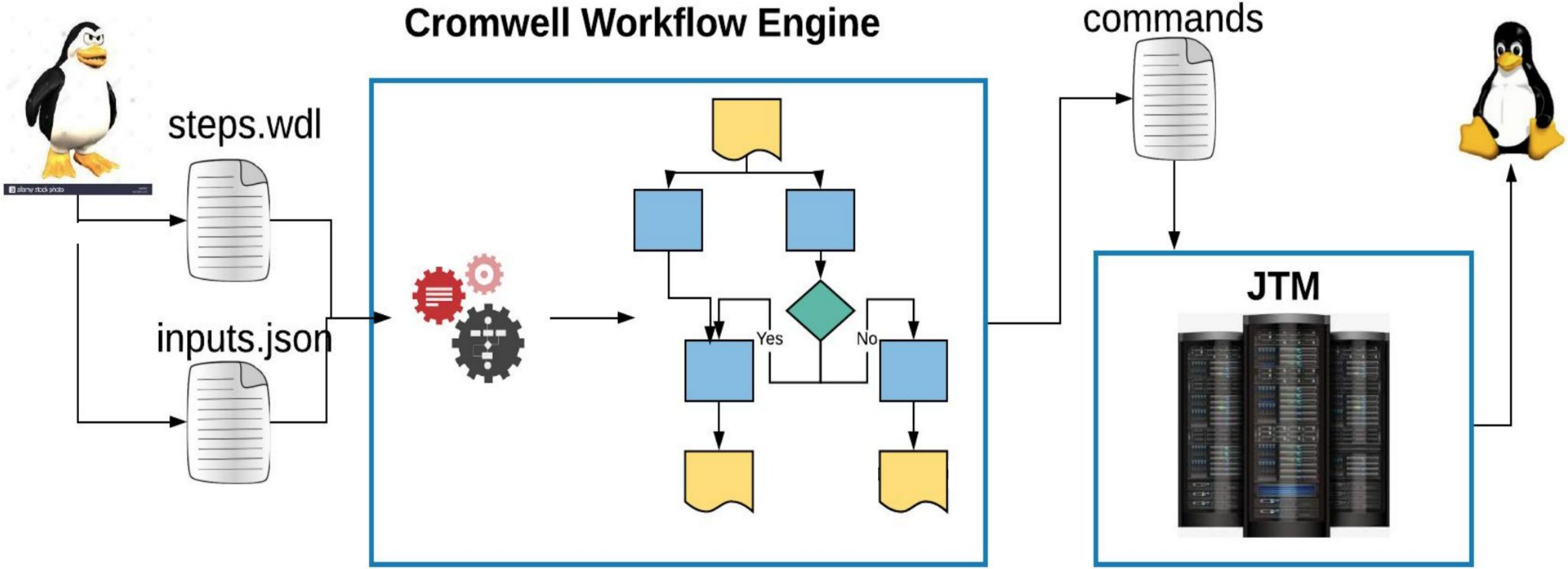
JAWS - JGI Analysis and Workflow System

- Developed by Joint Genome Institute
  - U.S. Dept. of Energy national user facility
  - DNA sequencing and other advanced genomic technologies
- Framework to run computational workflows
- Intended to improve re-usability and robustness of workflows in HPC environments
- Recently cleared for source code release!
  - Stay tuned

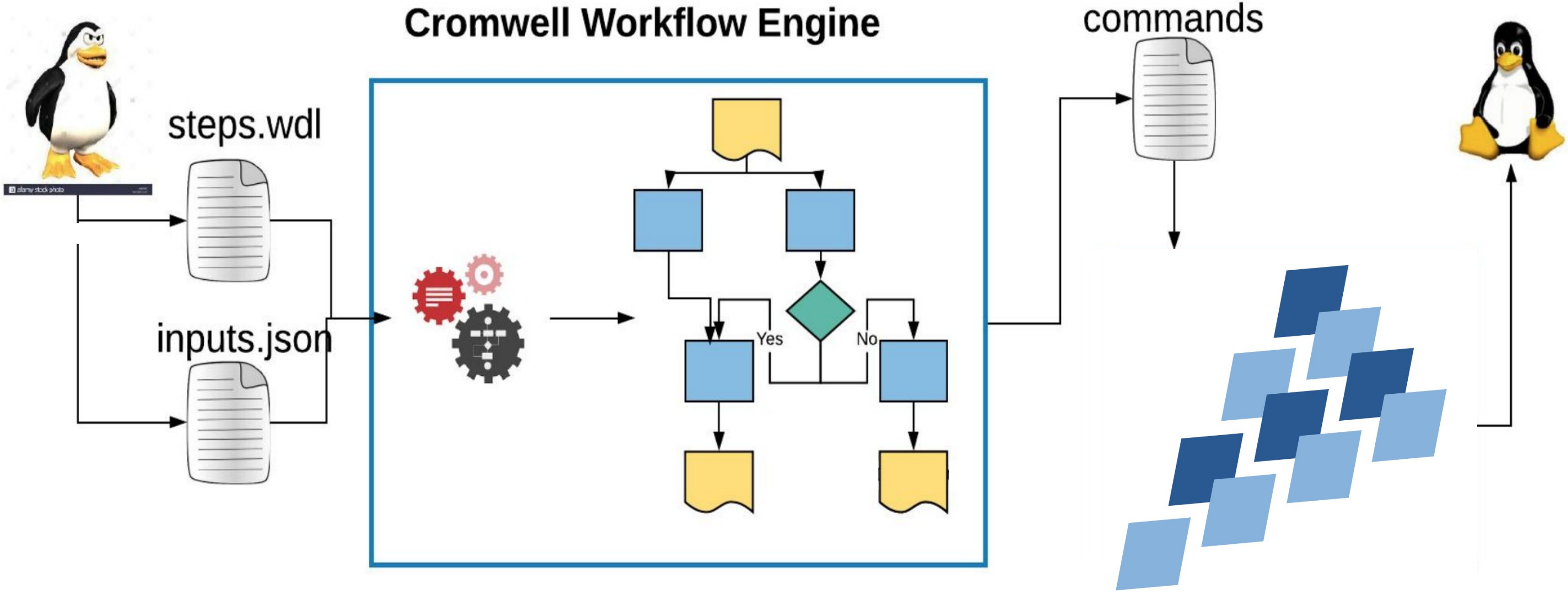
JAWS

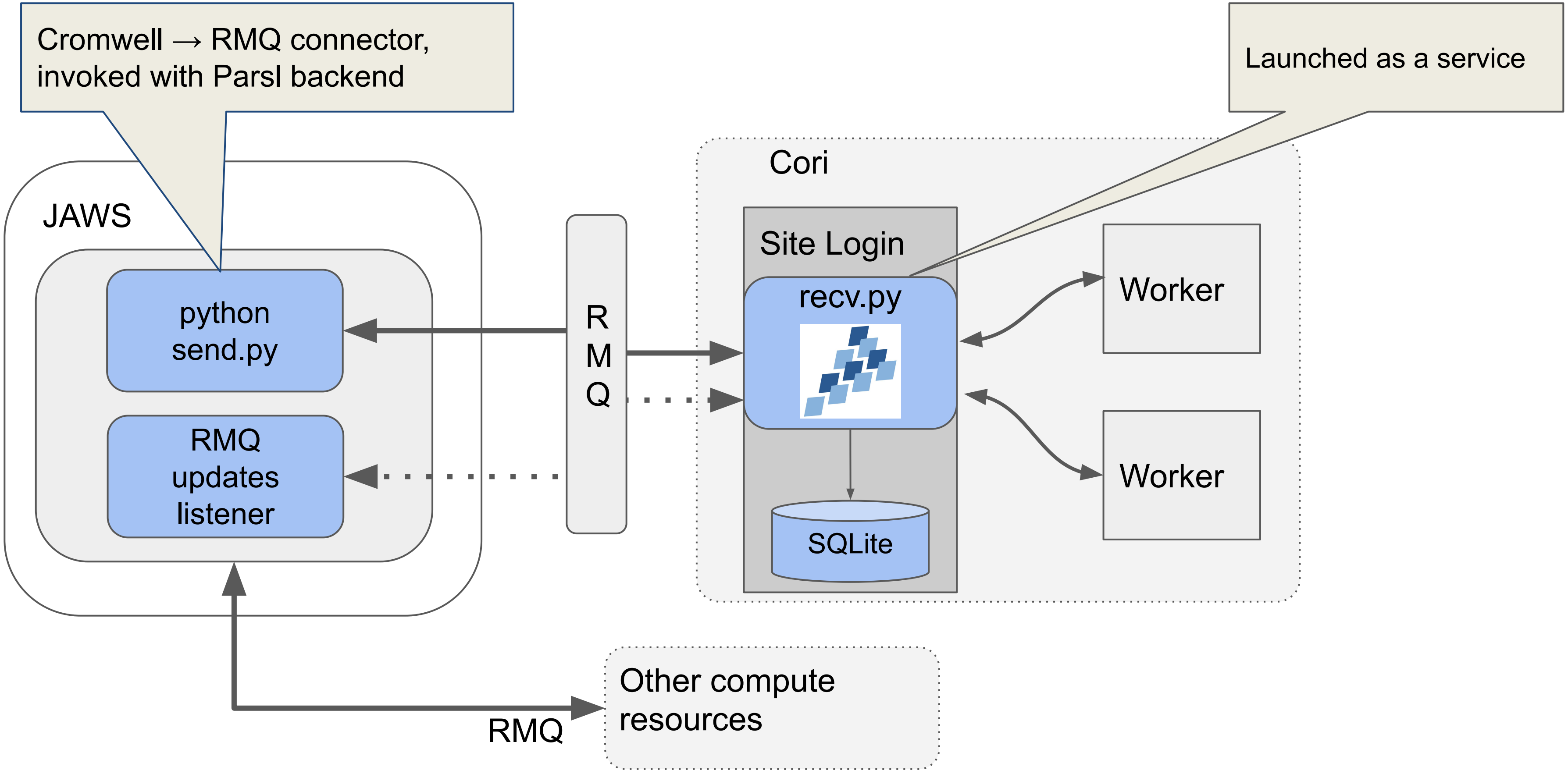
Workflows in the Data Ocean

JAWS - JGI Analysis and Workflow System



JAWS - JGI Analysis and Workflow System




Cromwell → RMQ connector, invoked with Parsl backend
Launched as a service
JAWS
python send.py
RMQ updates listener
RMQ
Cori
Site Login
recv.py
SQLite
Worker
Worker
Other compute resources
RMQ

Parsl Wish Item

- More flexibility in executors
    - Caveat to this: investigating Work Queue is on the to-do list

Special thanks to...

- Georg Rath
- Mario Melara
- Kyle Chard
- Yadu Babuji