# Running Flux in Slurm

Ryan Day
LC Resource Management

Flux development team: Al Chu, James Corbett, Jim Garlick, Mark Grondona, Dan Milroy, Chris Moussa, Tom Scogland, Vanessa Sochat, Jae-Seung Yeom

Sept 13, 2023

LLNL-PRES-853734



Lawrence Livermore National Laboratory

# What is Flux?

- Flux is a next generation open source resource manager being developed at LLNL.
- Flux is hierarchical. Every flux batch job is a full flux instance with the ability to schedule more jobs on its resources.
- Flux python API gives users powerful tools for running complex workflows.
- Flux can be easily run by users in their Slurm allocations.

# What is Flux?

- Flux is a next generation open source resource manager being developed at LLNL.*
- Flux is hierarchical. Every flux batch job is a full flux instance with the ability to schedule more jobs on its resources.
- Flux python API gives users powerful tools for running complex workflows.
- Flux can be easily run by users in their Slurm allocations.

*Flux is NOT a replacement for Slurm.**

** yet

# Flux on Github

## flux-framework

52 followers  http://flux-framework.github.io  @FluxFramework

Overview  Repositories 43  Projects 1  Packages  People 9

### Pinned

**flux-core** Public
core services for the Flux resource management framework
C  136  47

**flux-sched** Public
Fluxion Graph-based Scheduler
C++  76  36

**flux-coral2** Public
Plugins and services for Flux on CORAL2 systems
Shell  6  5

**flux-docs** Public
Documentation for the Flux-Framework
CSS  7  20

**flux-k8s** Public
Project to manage Flux tasks needed to standardize kubernetes HPC scheduling interfaces
Go  16  10

**PerfFlowAspect** Public
An Aspect Oriented Programming (AOP)-based tool to analyze cross-cutting performance concerns of composite science workflows.
C++  2  13

### People

### Top languages

C  Python  Shell  Rust  C++

### Most used topics

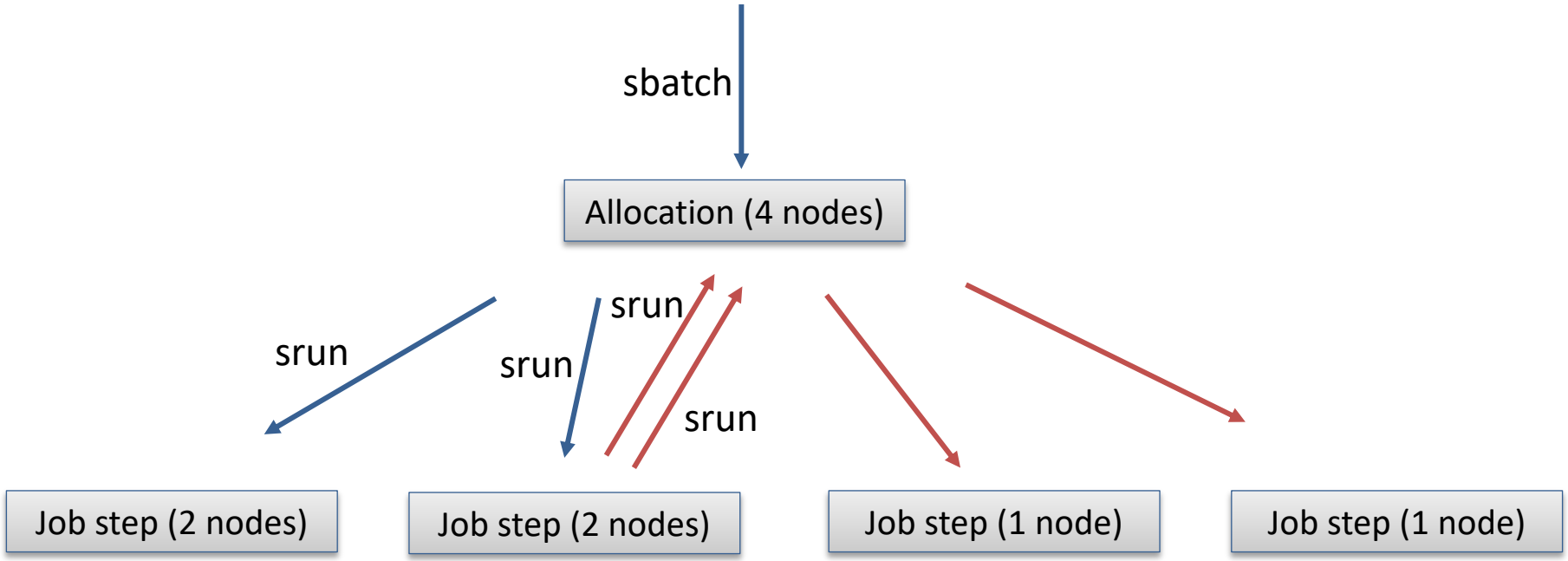
hpc  flux-framework  workflows  radiuss  resource-manager

https://github.com/flux-framework

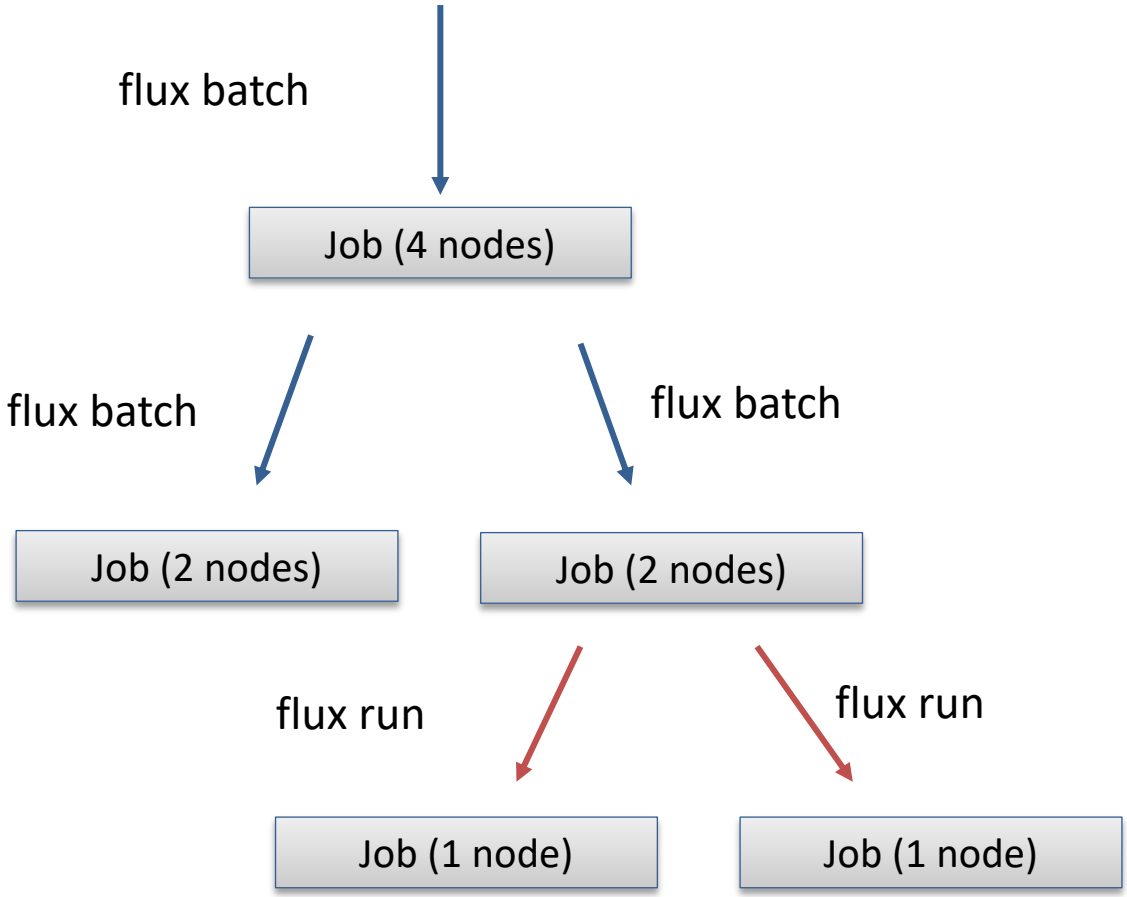
# What is Flux?

- Flux is a next generation open source resource manager being developed at LLNL.
- Flux is hierarchical. Every flux batch job is a full flux instance with the ability to schedule more jobs on its resources.
- Flux python API gives users powerful tools for running complex workflows.
- Flux can be easily run by users in their Slurm allocations.

# Flux is hierarchical: Launching steps in Slurm

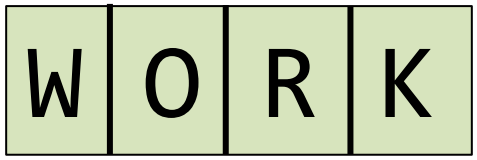


# Flux is hierarchical: Launching jobs in Flux



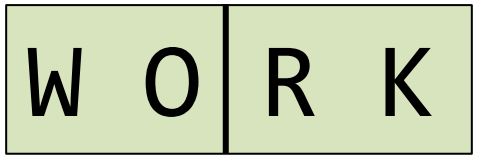
# Digression: a personal timeline of simulation

- 2000: 1 simulation takes many jobs.
- 2002: still waiting for those jobs.
- 2004: run 100 copies of 1 simulation across many jobs.
- 2006: many simulations across many jobs. Exchange between simulations.
- 2008: build statistical models instead of running simulations.
- 2010: have kids and go work in HPC support



# Digression: a personal timeline of simulation

- 2000: 1 simulation takes many jobs.
- 2002: still waiting for those jobs.
- 2004: run 100 copies of 1 simulation across many jobs.
- 2006: many simulations across many jobs. Exchange between simulations.
- 2008: build statistical models instead of running simulations.
- 2010: have kids and go work in HPC support



# Digression: a personal timeline of simulation

2000: 1 simulation takes many jobs.

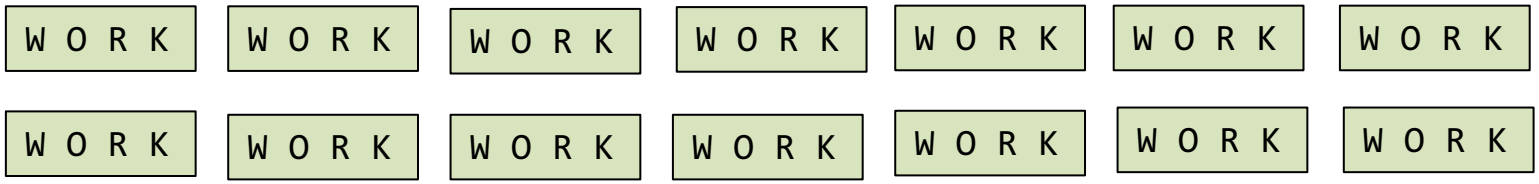
2002: still waiting for those jobs.

2004: run 100 copies of 1 simulation across many jobs.

2006: many simulations across many jobs. Exchange between simulations.

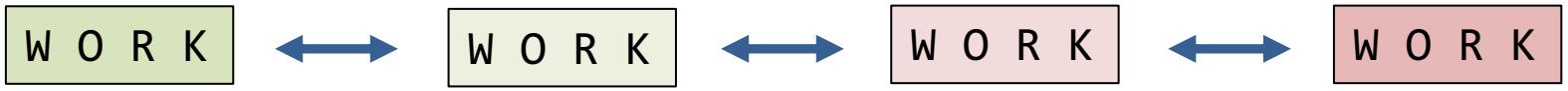
2008: build statistical models instead of running simulations.

2010: have kids and go work in HPC support



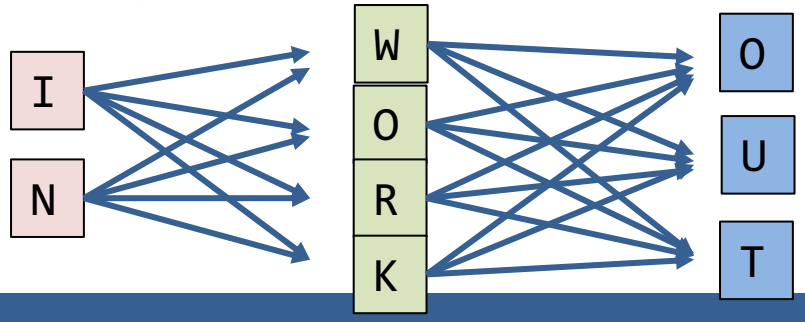
# Digression: a personal timeline of simulation

- 2000: 1 simulation takes many jobs.
- 2002: still waiting for those jobs.
- 2004: run 100 copies of 1 simulation across many jobs.
- 2006: many simulations across many job steps. Exchange between simulations.
- 2008: build statistical models instead of running simulations.
- 2010: have kids and go work in HPC support

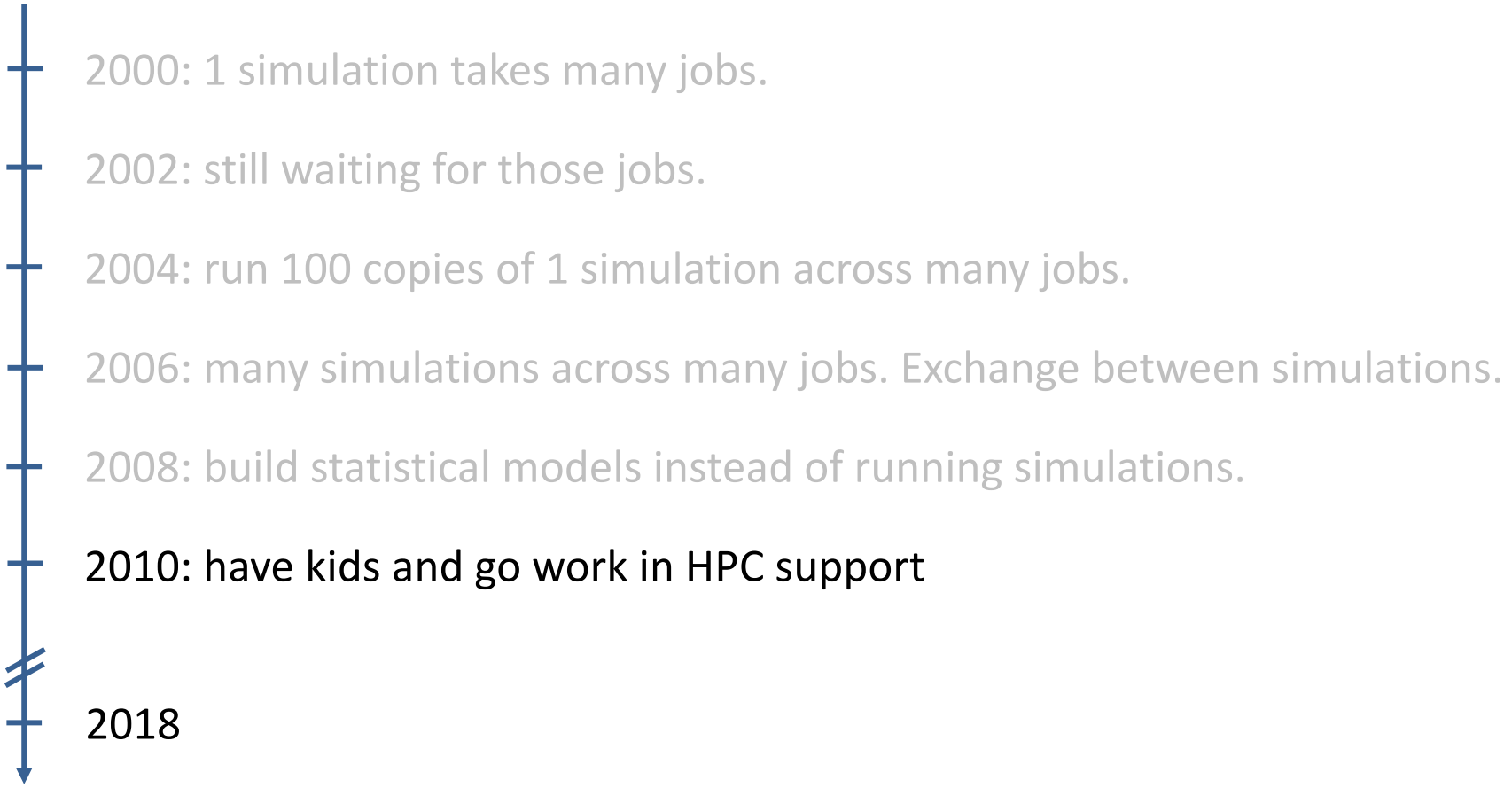


# Digression: a personal timeline of simulation

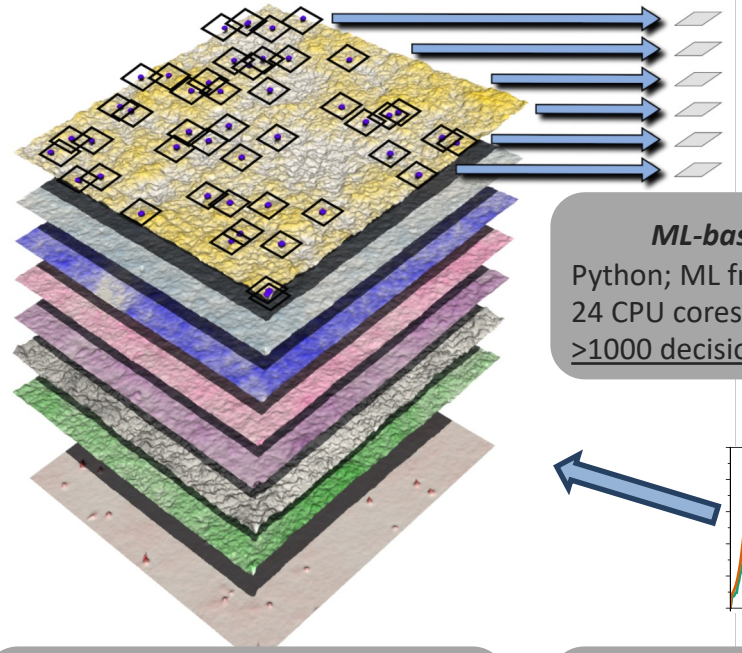
- 2000: 1 simulation takes many jobs.
- 2002: still waiting for those jobs.
- 2004: run 100 copies of 1 simulation across many jobs.
- 2006: many simulations across many jobs. Exchange between simulations.
- 2008: build statistical models instead of running simulations.
- 2010: have kids and go work in HPC support



# Digression: a personal timeline of simulation



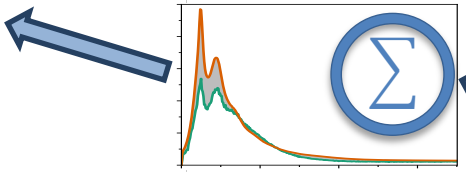
# MuMMI implements a complex and dynamic workflow



***CG Setup***
Python; Custom; GROMACS
24 CPU cores each
1.5 hr each

***ML-based Selection***
Python; ML frameworks; FAISS
24 CPU cores
>1000 decisions per minute

FIFO; real-time tracking & update



***Single Macro Simulation***
C++ (with MPI); MOOSE; ddcMD
24 CPU cores/node; 2400 MPI tasks
242 GB per day

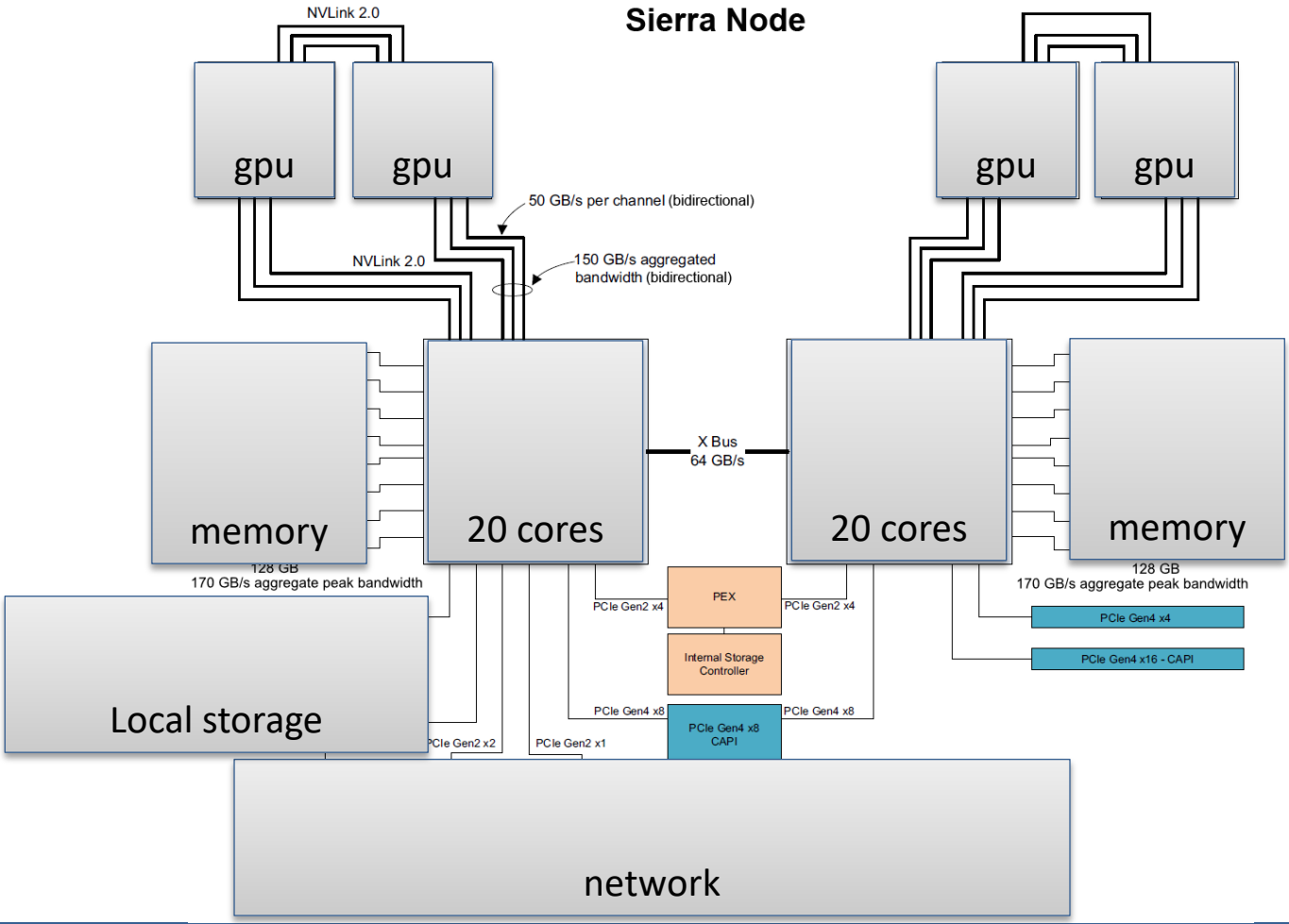
***Analysis Aggreg. & Feedback***
Python; Custom
24 CPU cores
120,000 reads per cycle

***CG Simulation***
C++ (with CUDA); ddcMD
1 GPU + 1 CPU core each
1.04 μs and ~6.5 GB per day

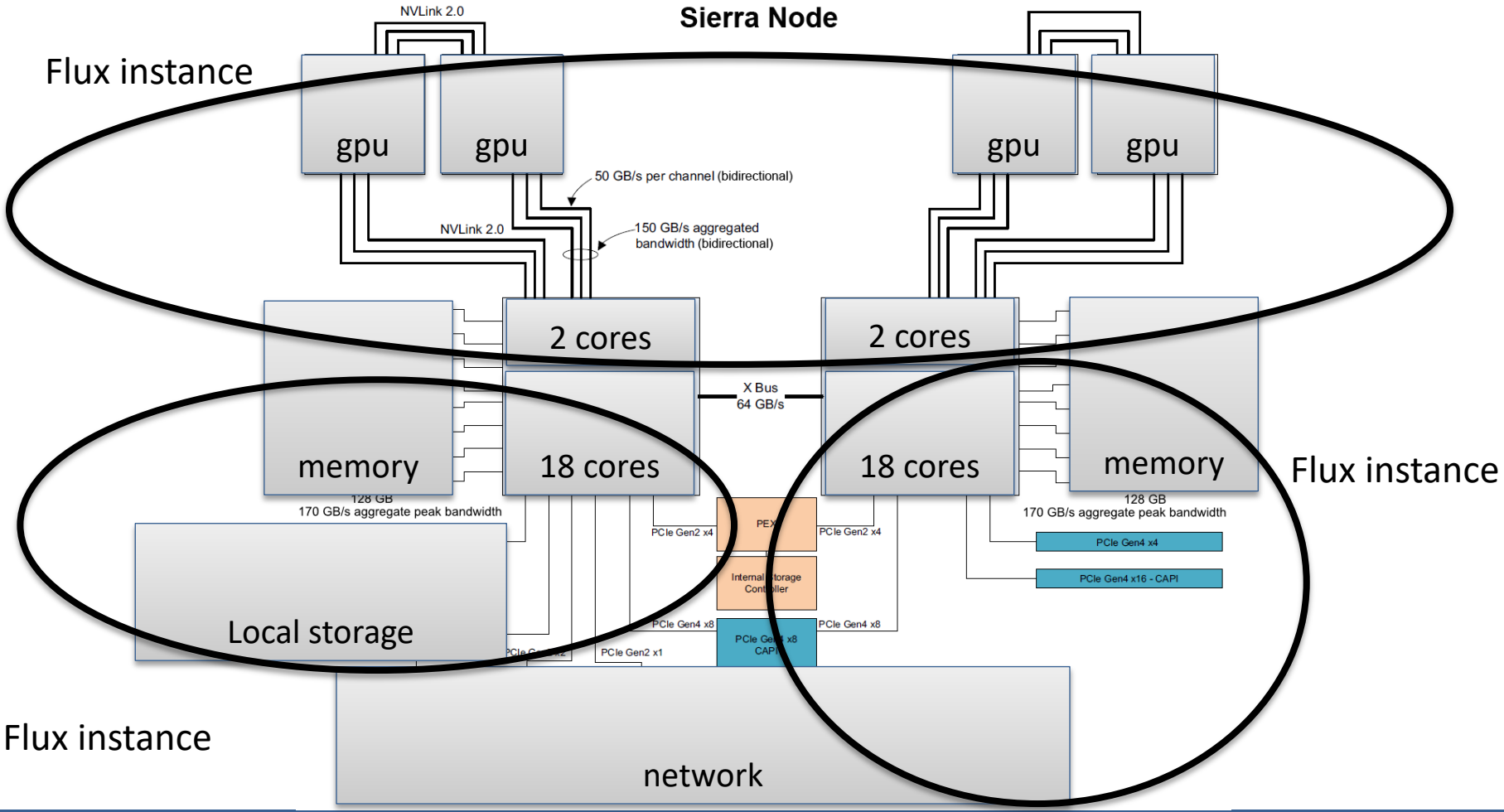
***In situ CG Analysis***
Python; Custom
3 CPU cores each
>2K frames per day

https://github.com/flux-framework/Tutorials/tree/master/2020-ECP (Di Natale)

# Flux is hierarchical: Sierra node diagram

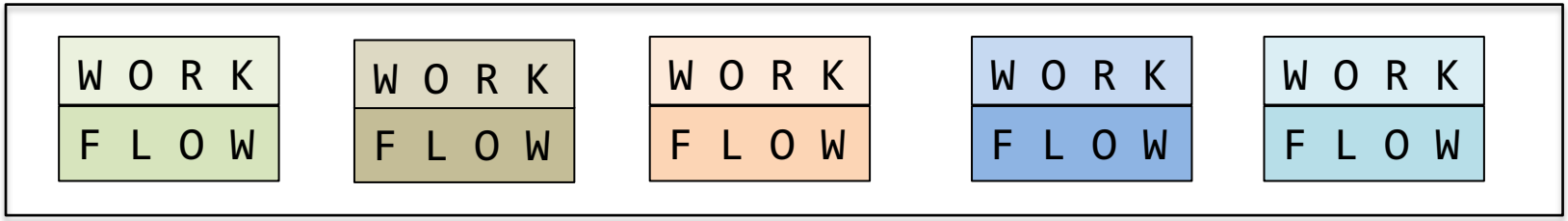


# Flux is hierarchical: ATS node diagram

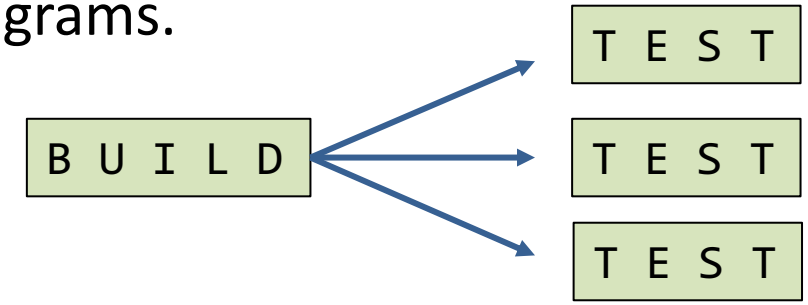


# Flux is hierarchical: other workflow examples

- Uncertainty quantification (UQ): run an ensemble of simulation pipelines with different starting conditions.




- Automated testing / CI: running arbitrary sets of unit tests for parallel programs.



# What is Flux?

- Flux is a next generation open source resource manager being developed at LLNL.
- Flux is hierarchical. Every flux batch job is a full flux instance with the ability to schedule more jobs on its resources.
- Flux python API gives users powerful tools for running complex workflows.
- Flux can be easily run by users in their Slurm allocations.

# Usability: Submitting a job

- Slurm

```
srun –N2 –n4 –t 2:00 sleep 120
```

- Flux CLI

```
flux run –N2 –n4 –t 2m sleep 120
```

- Flux API:

```
import flux
from flux.job import JobspecV1

f = flux.Flux()
j = JobspecV1.from_command(command=["sleep", "120"],
                           num_nodes=2,
                           num_tasks=4)
j.set_duration(120)
resp = flux.job.submit(f, j)
```

https://flux-framework.readthedocs.io/en/latest/auto_examples/example_job_submit_api.html

# Usability: Futures

```
from flux.job import JobspecV1, FluxExecutor

def event_callback(future, event):
    print(f"job {future.jobid()} triggered event {event.name!r}")

def main():
    # set up jobspecs
    …
    # submit jobs and register event callbacks for all events
    with FluxExecutor() as executor:
        futures = [executor.submit(compute_jobreq) for _ in range(args.njobs // 2)]
        …
        for fut in futures:
            # each event can have a different callback
            for event in executor.EVENTS:
                fut.add_event_callback(event, event_callback)
```

https://flux-framework.readthedocs.io/projects/flux-workflow-examples/en/latest/

# What is Flux?

- Flux is a next generation open source resource manager being developed at LLNL.
- Flux is hierarchical. Every flux batch job is a full flux instance with the ability to schedule more jobs on its resources.
- Flux python API gives users powerful tools for running complex workflows.
- Flux can be easily run by users in their Slurm allocations.

# Running Flux: starting Flux in a Slurm allocation

You can start flux in a Slurm allocation on any cluster:

```
[day36@quartz2306:~]$ salloc -N2 -p pdebug
salloc: Pending job allocation 1228472
salloc: job 1228472 queued and waiting for resources
salloc: job 1228472 has been allocated resources
salloc: Granted job allocation 1228472
salloc: Waiting for resource configuration
salloc: Nodes quartz[44-45] are ready for job

sh-4.4$ srun -N2 -n2 --pty flux start

sh-4.4$ flux resource list
     STATE NNODES   NCORES    NGPUS NODELIST
      free      2       72        0 quartz[44-45]
 allocated      0        0        0
      down      0        0        0

sh-4.4$ flux run -N2 -n2 hostname
quartz44
quartz45

sh-4.4$
```

https://flux-framework.readthedocs.io/en/latest/quickstart.html#starting-a-flux-instance

# Running Flux: running a batch script

```
sh-4.4$ cat quickexample.sh
#!/bin/sh
# flux : -N 2
# flux : -n 2
date
flux run -n 2 ~/hello/mpi_hello/hello_mpi

sh-4.4$ srun -N2 -n2 flux start \\
'flux batch quickexample.sh && flux queue drain'
f2SqkPL3
sh-4.4$ cat flux-f2SqkPL3.out
Fri Aug 25 13:37:05 PDT 2023
Hello from task 1 on quartz45!
Hello from task 0 on quartz45!

Number of MPI tasks is: 2

sh-4.4$
```

# Running Flux: running a python workflow

```
sh-4.4$ cat flux-workflow-examples/job-status-control/bookkeeper.py
#!/usr/bin/env python3
…
def event_callback(future, event):
    print(f"job {future.jobid()} triggered event {event.name!r}")
…
sh-4.4$ srun -N2 -n2 flux start \\
./flux-workflow-examples/job-status-control/bookkeeper.py 2
bookkeeper: all jobs submitted
bookkeeper: waiting until all jobs complete
job f2RSD4q5 triggered event 'submit'
job f2RSD4q6 triggered event 'submit'
job f2RSD4q5 triggered event 'depend'
job f2RSD4q5 triggered event 'priority'
job f2RSD4q5 triggered event 'exception'
job f2RSD4q5 triggered event 'clean'
job f2RSD4q6 triggered event 'depend'
job f2RSD4q6 triggered event 'priority'
job f2RSD4q6 triggered event 'exception'
job f2RSD4q6 triggered event 'clean'
bookkeeper: all jobs completed
sh-4.4$
```

https://flux-framework.readthedocs.io/projects/flux-workflow-examples/en/latest/

# Running Flux: getting Flux

- Install with spack:

```
spack install flux-sched
```

- Build from source:

```
git clone https://github.com/flux-framework/flux-core.git
configure, make, make install

git clone https://github.com/flux-framework/flux-sched.git
configure, make, make install
```

- Docker (quick single node deployment):

```
docker run -ti fluxrm/flux-sched:latest
```

https://flux-framework.readthedocs.io/en/latest/quickstart.html#building-the-code

# Where to find out more

- https://github.com/flux-framework/
- https://flux-framework.readthedocs.io/en/latest/
- https://github.com/flux-framework/Tutorials
- https://hpc-tutorials.llnl.gov/flux/
- https://hpc.llnl.gov/banks-jobs/running-jobs/batch-system-cross-reference-guides
- `flux help`, `flux help run`, `flux help jobs`, …

Lawrence Livermore
National Laboratory