Towards Diversified Exascale NWP Workflows

Christopher W. Harrop, Stefan F. Gary, Alvaro Vidal Torreira, Christina Holt, Naureen Bharwani, Venita Hagerty, Isidora Jankov, Kyle Chard, Michael Wilde

ParslFest 2023 - Oct 20, 2023, Chicago

Acknowledgements

This work is funded by the NOAA Software Engineering for Novel Architectures (SENA) Program

In collaboration with U.S. Department of Energy Exascale Computing Project & Parallel Works (funded by DOE Small Business Innovation Research)

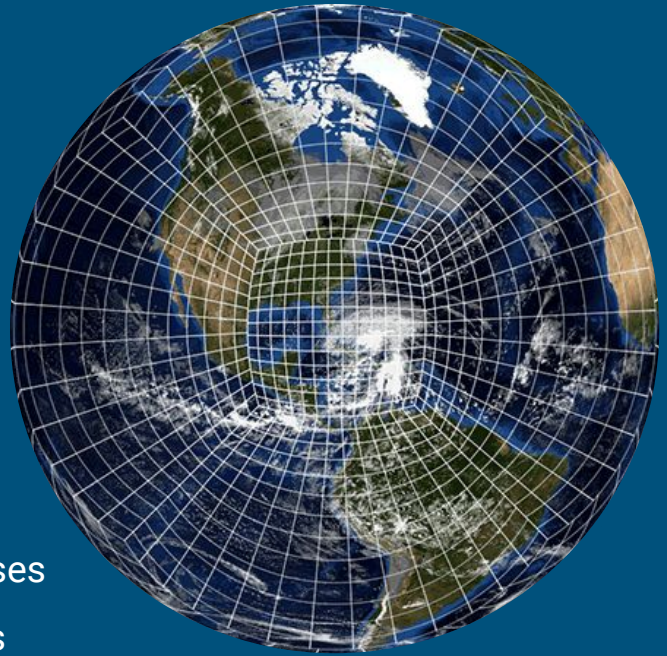
Supported by NOAA Research and Development HPC (RDHPCS) Program

Special thanks to: Todd Munson, Eric Schnepp, James Corbett, Jim Garlick

NWP Grand Challenge

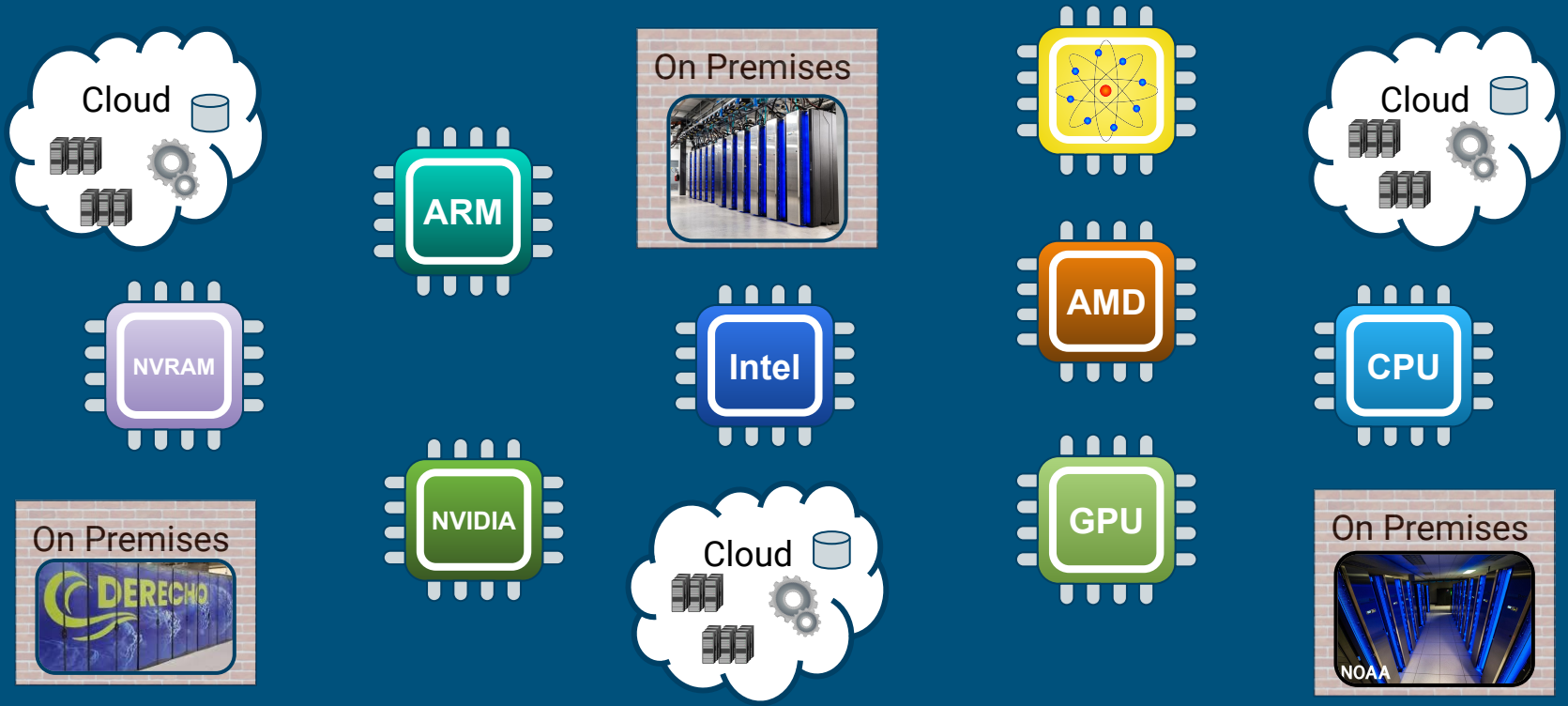
Numerical Weather Prediction at Exascale

Goal: Global <3km cloud-resolving model



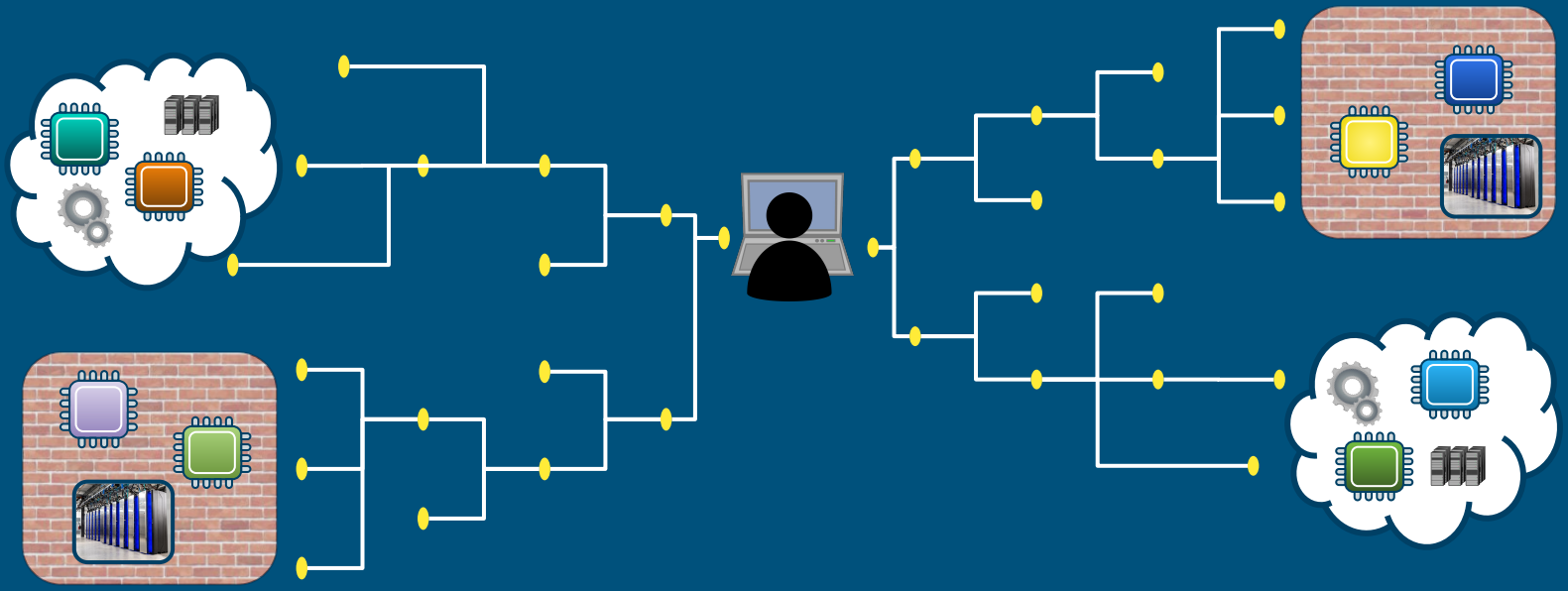
- Large ensembles / Digital twin
- ML models to replace expensive calculations
  - radiation, microphysics, data assimilation, etc
- Increased fidelity for representation of physical processes
- New programming models and computational methods

A Diversification Disruption

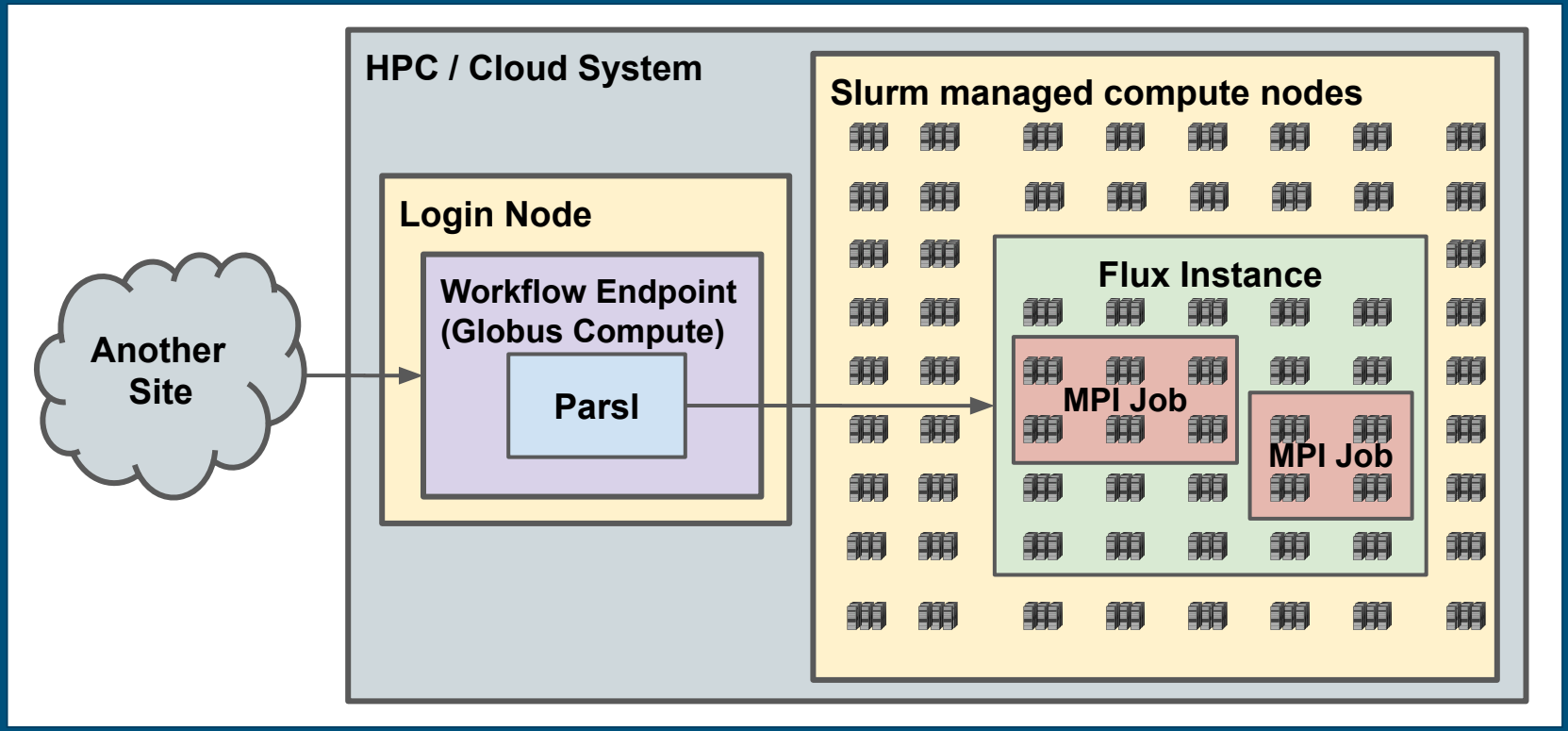


Our Vision

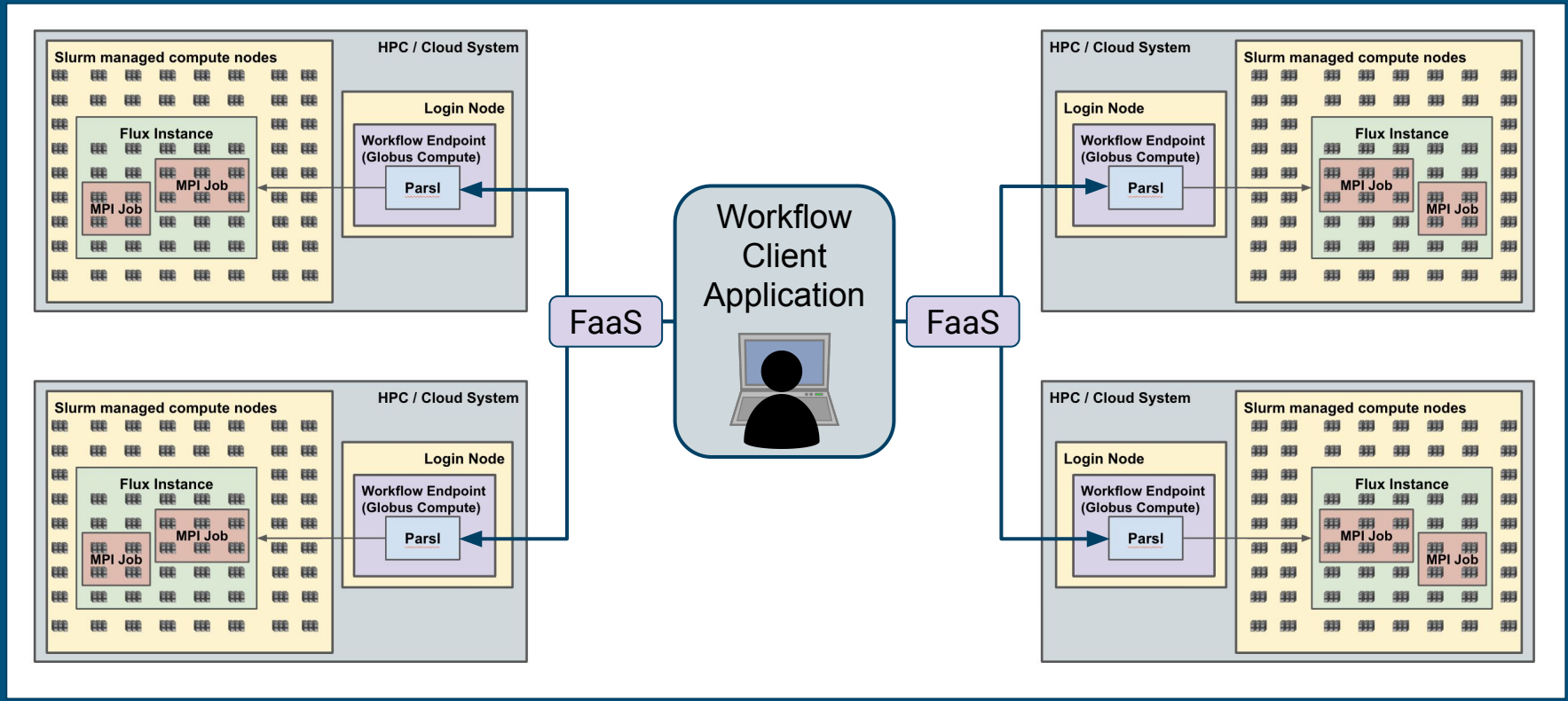
Diversified distributed workflows to enable NWP HPC Research



The Current Architecture



The Current Architecture



Outstanding Questions

- Should workflows be expressed as programs or as configurations?
  - Python or YAML?
- How do we monitor and steer distributed workflows to diagnose problems?
  - If a workflow is a Python program, how do you interrogate and control it?
  - How do we measure and optimize performance of a distributed workflow?
- How to address advanced MPI + X requirements
  - Core / Socket affinity - Custom MPI vendor-specific settings
  - MPMD launch for coupled models - mpiexec -n 100 atmosphere.exe : -np 10 ocean.exe
  - Customized, non-uniform, layout of MPI ranks - I/O task groups vs model compute ranks

Outstanding Questions

- What about the data?
  - How and when should we move distributed data?
  - File transfers or streaming from one application to another?
- How best to manage complex mixture of HPC & HTC tasks?
- Best practices for Parsl application design / development
- How do we test while developing distributed workflow capabilities?
  - Requires large, complex supporting software stacks
  - CI / CD using containerized Slurm clusters?

Summary Remarks

- We have a vision, and many questions, but do not have all the answers
- We are testing Parsl + Flux + Globus Compute
  - Parsl → High throughput computing and powerful programming interface
  - Flux → MPI-aware scheduling within Parsl workflows
  - Globus Compute → Function as a service (FaaS) for secure distributed execution
- We are starting small for testing and exploration purposes
  - Simple Quasi-Geostrophic data assimilation workflow
- We are reusing existing workflow development where possible
- Demonstration with a “real” model once foundational pieces are settled

Questions / Discussion