Jumping the Queue: From NASA to the Commercial Cloud

Robert Hood
InuTeq – NASA Ames

The HECC Project at NASA Ames

High End Computing Capability

- Facilitating Science/Engineering at NASA
- 4 compute systems, including (as of 6/18):
  - Pleiades: #24 in Top500; #14 in HPCG
  - Electra: #43 & #37 (both should go up)
- 1500+ users from across Agency Directorates
  - Aeronautics, Human Exploration, Science
- ~900 MPI applications (+ non-MPI codes)

Goal: Maximize resource delivery to users

- Strive for >80% utilization of resources
- Embrace new facility technologies
- Continually evaluate procurement options

When Does It Make Sense for HECC to Use Clouds?

Conduct a trade study: on-premises costs vs in-cloud costs

Main Finding:

- Commercial clouds do not offer a viable, cost-effective approach for replacing on-premises HPC resources at NASA.

Additional Finding:

- Commercial clouds provide a variety of resources not available at HECC. Certain use cases may be cost-effective to run there.

Example Conditions That May Warrant Cloud Usage:

- When utilization would be low—such as when using new resource types
  - But acquire in-house resources when demand grows sufficiently
- When there are other costs to consider, such as opportunity costs associated with high utilization (longer queue waits)
- When there are real-time requirements, such as web services

Follow-up Work Identified by Trade Study

Gain a better understanding of potential benefits and costs of having a portion of the HECC workload in the cloud:

- Understand performance characteristics of jobs that might run there
- Define a comprehensive model that allows accurate comparisons of cost of running jobs depending on resources used

Prepare for a broadening of HECC services to include a portion of its workload running on commercial cloud resources:

- For HECC users
- For non-HECC users

Pilot Project: Enabling Cloud Usage

Move Jobs from HECC Resources to Amazon Web Services (AWS)

- HECC user logs into an AWS cloud front end to build executable
- User annotates batch scripts, indicating files that need to be staged to/from cloud
- User submits batch jobs to “cloud” queue
- PBS server moves jobs to server running at AWS; stages input files to AWS
- Server in cloud allocates resources, runs job; PBS server in HECC stages output files back
- Accounting is done manually; HECC pays
- Limited to non-export controlled codes and data (i.e. “low” security plan)

User-Defined Software Stacks

- Container technologies (Charliecloud, Singularity) are under evaluation

User testing of the “cloud bursting” started in September 2018

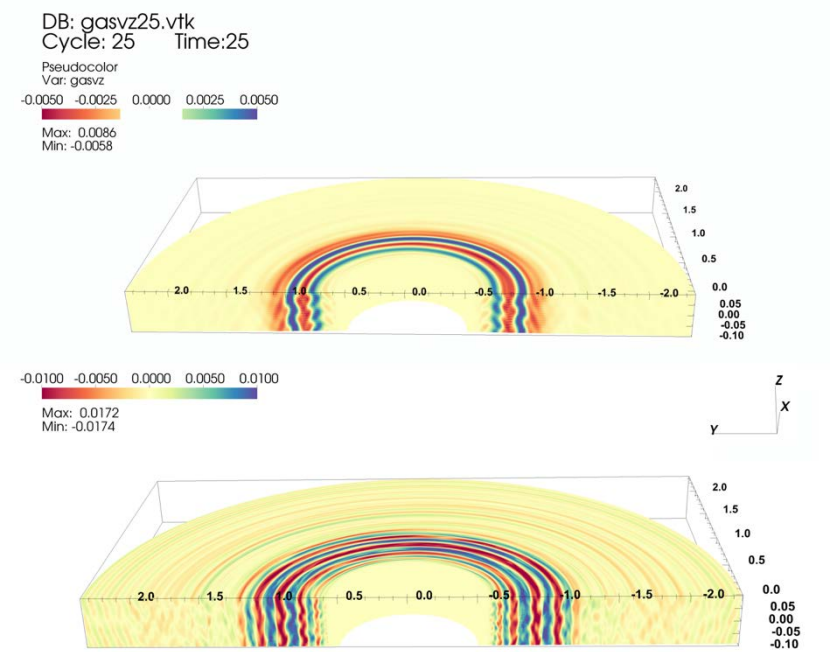
Modeling of Spiral Wave Instabilities (SWI)

Researchers increased resolution of spiral waves in protoplanetary disks by 4x and captured cascade of SWI-driven turbulence at smaller scales.

- Previous resolution was limited to 512✖128✖128 due to computational costs.
- Using NVIDIA V100 nodes at AWS, the resolution increased to 2048✖512✖512.

Higher-resolution computation took 100 hours on single AWS compute node with eight V100s.

The researchers plan to test the SWI case with more realistic models and higher resolutions.



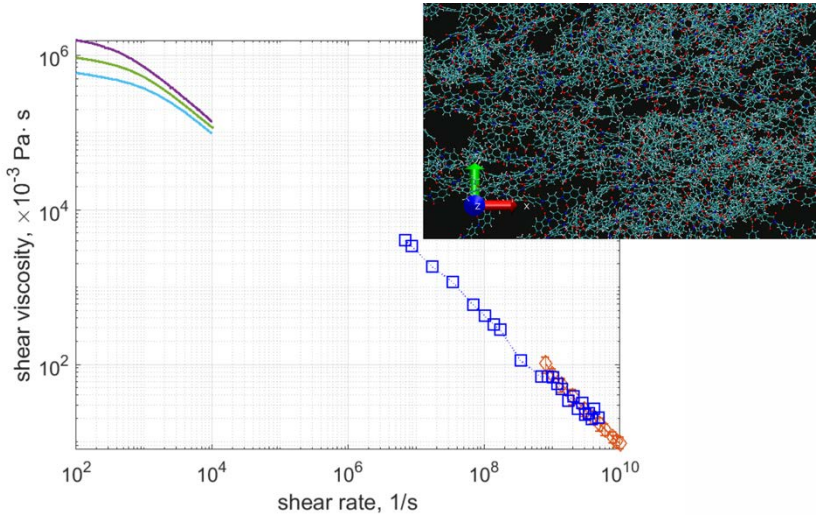
Top: Results with standard resolution (512x128x128). Bottom: 4x higher resolution (2048x512x512). While the spiral wave instability (SWI) develops with both resolutions, the SWI-induced turbulence is better resolved with a larger number of grid cells. *Jaehan Bae, Carnegie Institute of Washington.*

3D Aerospace Manufacturing Simulations

Material scientists narrowed the gap between experiment and simulation of polymer-polymer interfaces by more than two orders of magnitude.

- Previously, coarse-grained approximations used the LAMMPS molecular dynamics (MD) code running without GPUs (red squares).
- New fine-grained simulations used GROMACS MD package running on GPUs at AWS (blue squares).

The researchers will extend these results to predict other important aerospace manufacturing parameters, paving the way for exploration of novel frontiers of nanomanufacturing.

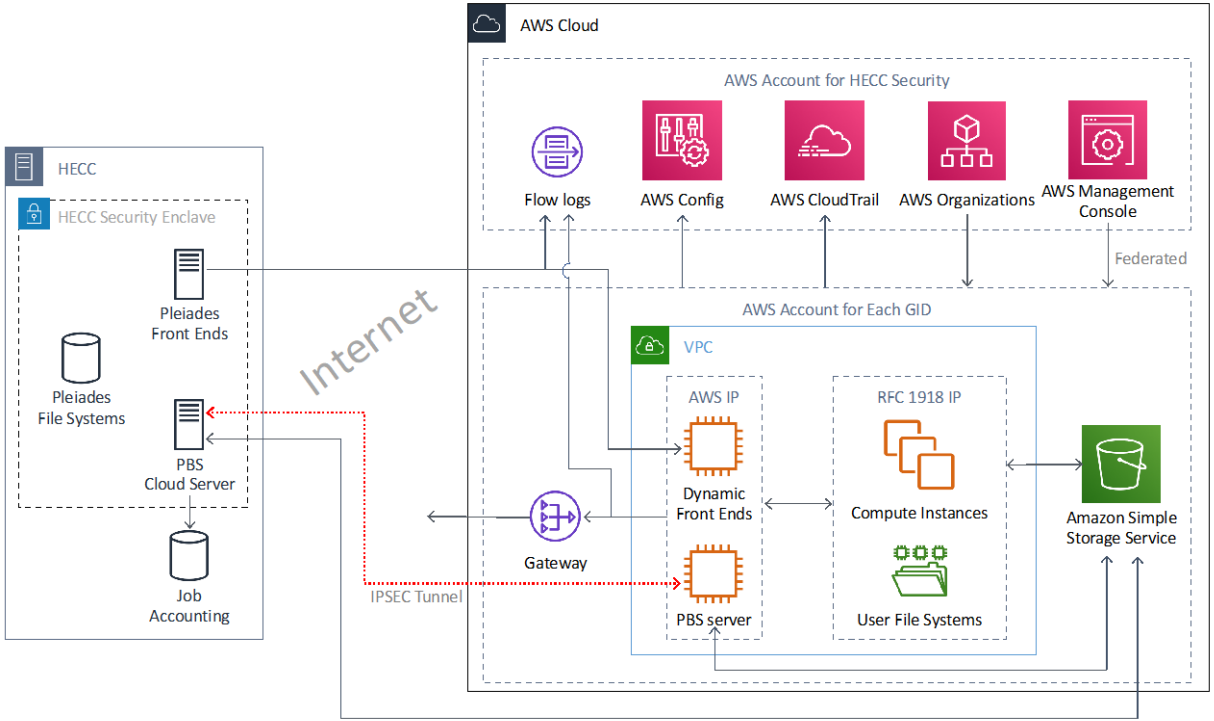


This graph shows atomistically resolved molecular dynamics (MD) predictions of the shear viscosity for flat polymer-polymer interfaces (squares at lower right). These can be compared with experimental data on bulk samples (lines at upper left). The inset shows MD predictions obtained on the Pleiades supercomputer. *Dmitry Luchinsky, SGT, Inc., Intelligent Systems Division, NASA Ames Research Center.*

Current Work

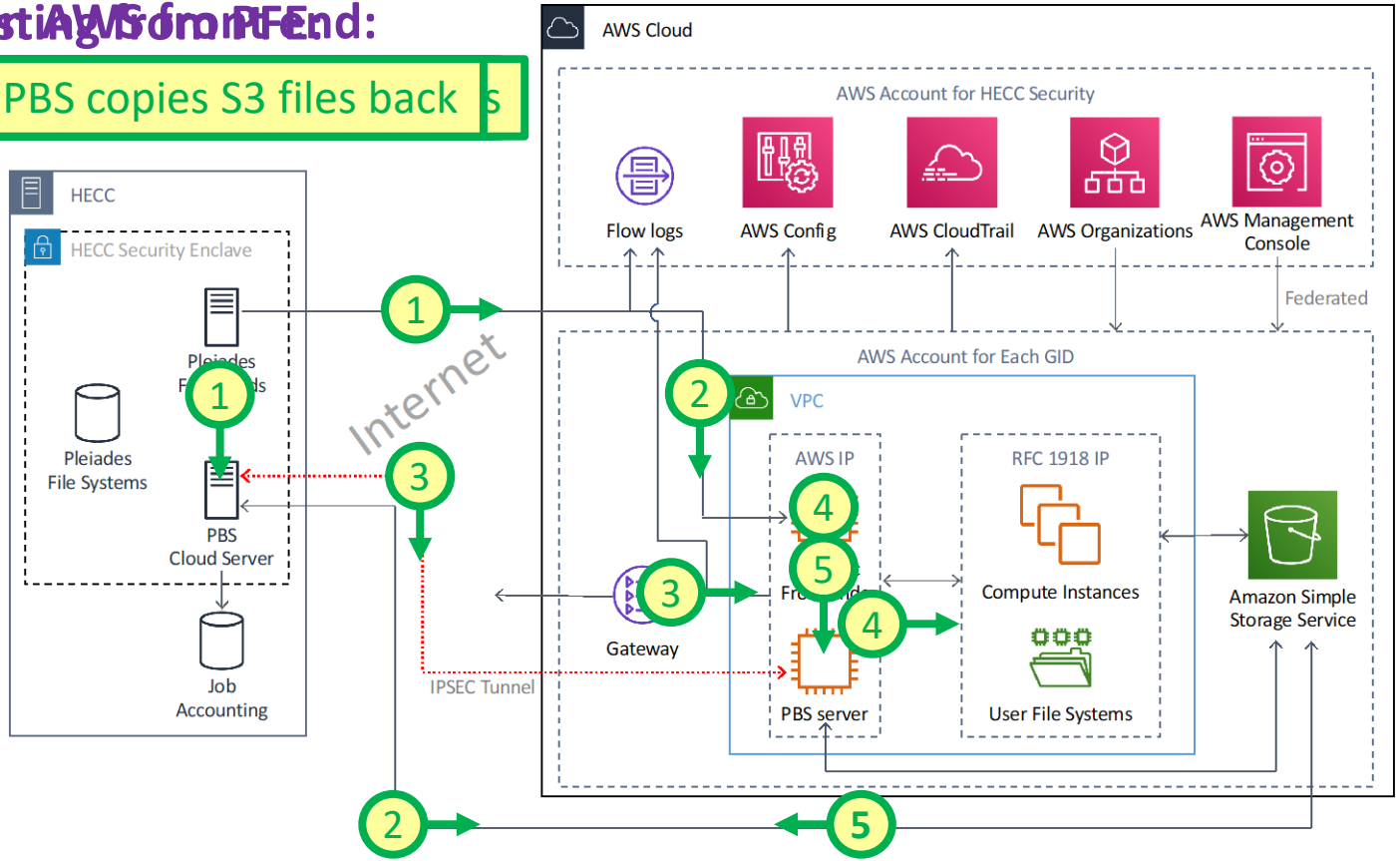
Pilot Extended to Include:

- Moderate security plan
- Full accounting, with account limits and automated tracking of consumption
  - Users bring funding
  - Accounts will be charged for:
    - Job resource usage
    - Ongoing storage usage
    - Data transfer bandwidth
    - Overhead



Usage Scenarios

5: HECC PBS copies S3 files back



Status

The HECC AWS Cloud Offering is Now Operational (Started August 7th):

- Full accounting, with account limits and automated tracking of consumption
- Processes for:
  - Account setup requests
  - Transfers from NASA WBS to HECC cloud account
  - Monthly statements

Future Work

Extend Services as Required by HECC Users

- New resource types at Amazon
- Other cloud providers (Google? Azure?)

Extend Services to Include Non-HECC Users from NASA

- They also need to bring their own funding
- Provide web-based user interface for defining and running jobs, moving data, etc.
- HECC would add cost-recovery fee to make this self-sustaining

Devise Cost Methodology

- Must be able to do meaningful cost comparisons between on-premises and in-cloud resources in order to determine which would be most cost effective
  - Include opportunity costs
  - Establish benchmark suite
- Adjust processes for acquisition and phase out of resources to include commercial cloud

Acknowledgements & Additional Information

The Trade Study Team:

- S. Chang, R. Hood, H. Jin, S. Heistand, J. Chang, S. Cheung, J. Djomehri, G. Jost, D. Kokron

The Trade Study: NAS Technical Report NAS-2018-001

- Posted on our website at:

  https://www.nas.nasa.gov/assets/pdf/papers/NAS_Technical_Report_NAS-2018-01.pdf

The Cloud Team:

- S. Chang, S. Heistand, H. Yeung, M-Y. Koo, J. Chang, S. Cheung, J. Djomehri, R. Hood, H. Jin

Knowledge Base:

- See the sixteen new articles published under Cloud Computing > AWS Cloud:

  https://www.nas.nasa.gov/hecc/support/kb/175/