USC University of Southern California

Security in High Performance Computing Environments

Taolue Wang, Zhekai Xie

USC

Background for High Performance Computing

- *High performance computing (HPC) is the ability to process data and perform complex calculations at high speeds.*
- To build a high-performance computing architecture, compute servers are networked together into a cluster. Software programs and algorithms are run simultaneously on the servers in the cluster.
- To operate at maximum performance, each component must keep pace with the others.

HPC use cases

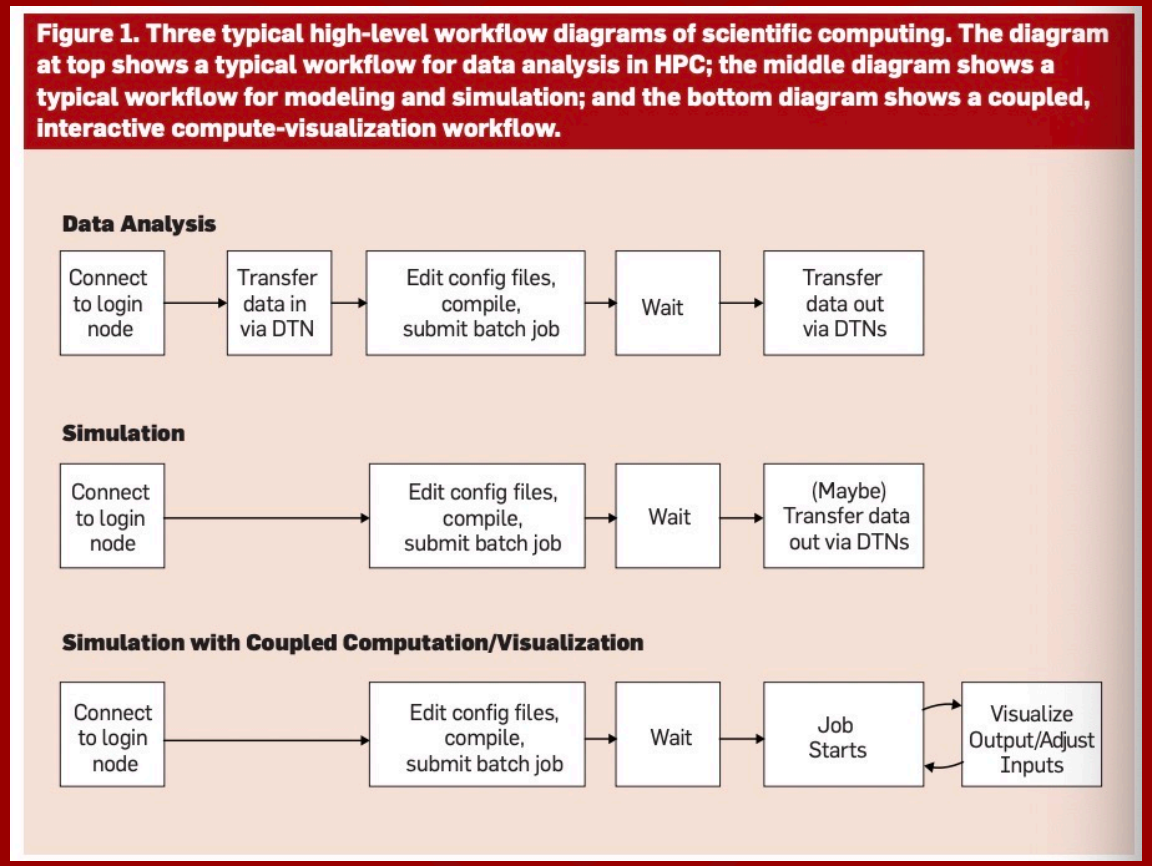
- *Research lab*. HPC is used to help scientists find sources of renewable energy, understand the evolution of our universe, predict and track storms, and create new materials.
- *Media and Film.* HPC is used to edit feature films, render mind-blowing special effects, and stream live events around the world.
- *Artificial Intelligence.* HPC is used to detect credit card fraud, provide self-guided technical support, teach self-driving vehicles, and improve cancer screening techniques.

Distinctiveness of security for HPC system

USC

Distinctive purposes

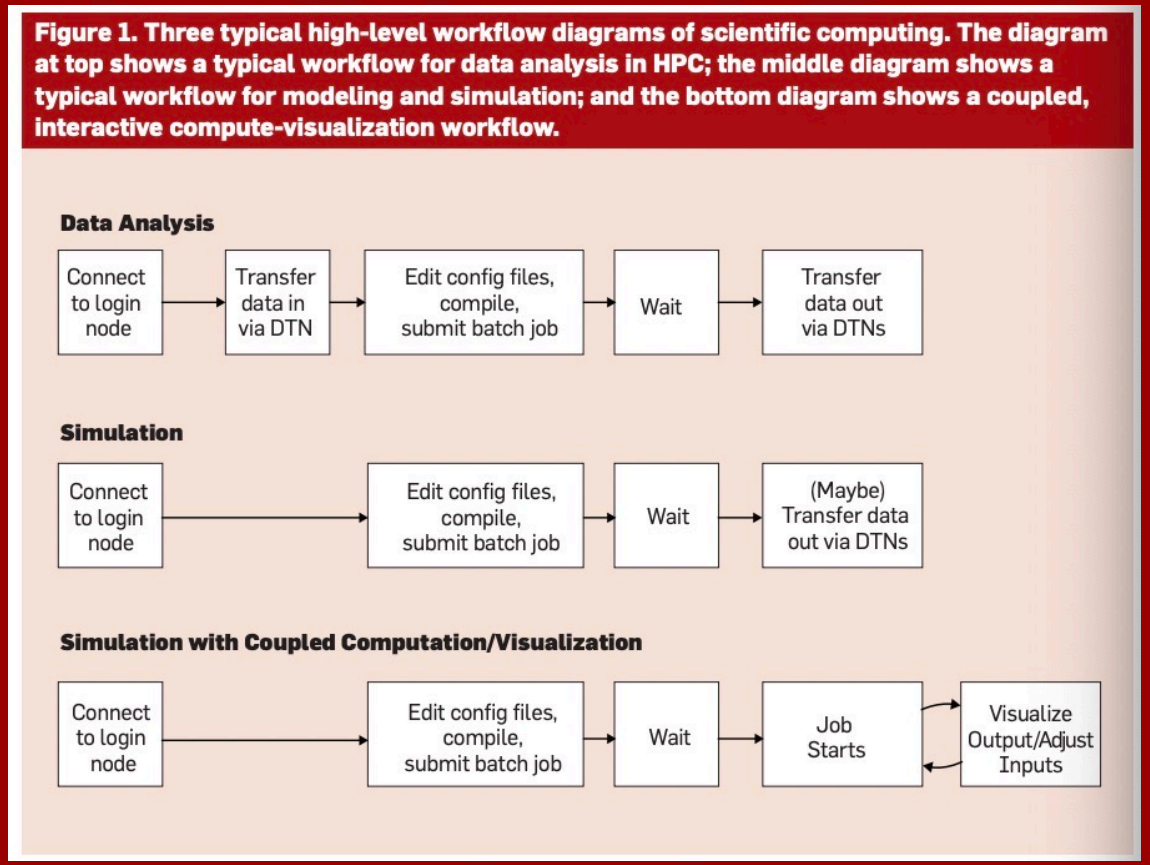
- High performance system are typically used for automated computation of some kind, typically performing some set of mathematical operations.
- Historically, this has often been for the purpose of modeling and simulation, and increasingly today, for data analysis as well.
- Users regard such a security solution as a waste of cycles at worst, and an unacceptable delay of scientific results at best.



USC

Distinctive modes of operation

- The typical mode of operation for using a scientific high-performance machine involves connecting through a login node of some kind.
- A batch scheduler that determines when jobs should run based on analyzing the batch scripts that have been submitted according to a given optimization policy.
- This may happen on the HPC system, or the output of the HPC computation may be downloaded to a non-HPC system for analysis in a separate environment such as using Jupyter/Ipython.



USC

Custom operating system stack

- HPC systems is that these systems often have highly exotic stacks. Current HPC environments represent a spectrum of hardware and software components, ranging from exotic and highly custom to fairly commodity.
- Aurora, the system scheduled to be installed at ALCF in 2019, will be constructed by a partnership between Cray and Intel
- Custom hardware and software components may have both positives and negatives.

Openness

- That is, scientists from all over the world whose identities have never been validated may use them.
- many such systems, such as those used by NSF or DOE ASCR, have no traditional firewalls between the data transfer nodes and the Internet

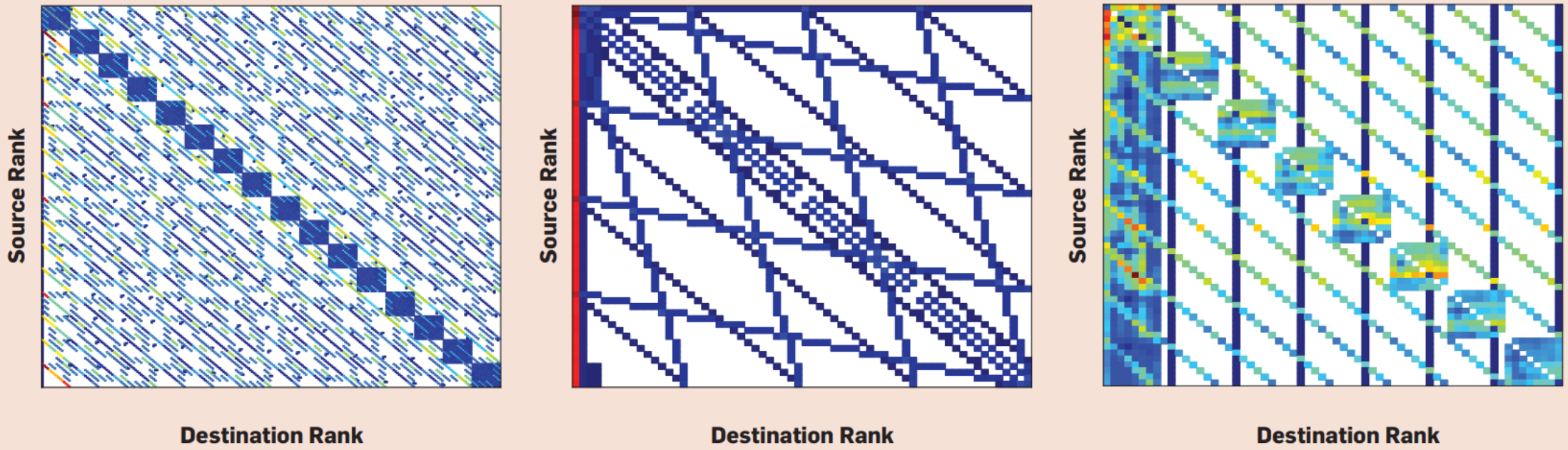
Security Mechanisms Overcome Constraints of HPC Environments

- Traditional IT security solutions don't work well in HPC
  - Even 0.0046% packet loss can cause a 90% loss through data transfers
- Science DMZ (Demilitarized Zone)
  - Defines a set of security policies, procedures, and mechanisms (HPC theme #1)
  - Moved to own enclave; transfer through single network ingress and egress point
  - Leverage packet filtering firewalls; reducing complexity
  - Less delay on transmission while inspection; less congestion lead to packet loss

Leveraging the Distinctiveness of HPC as an Opportunity

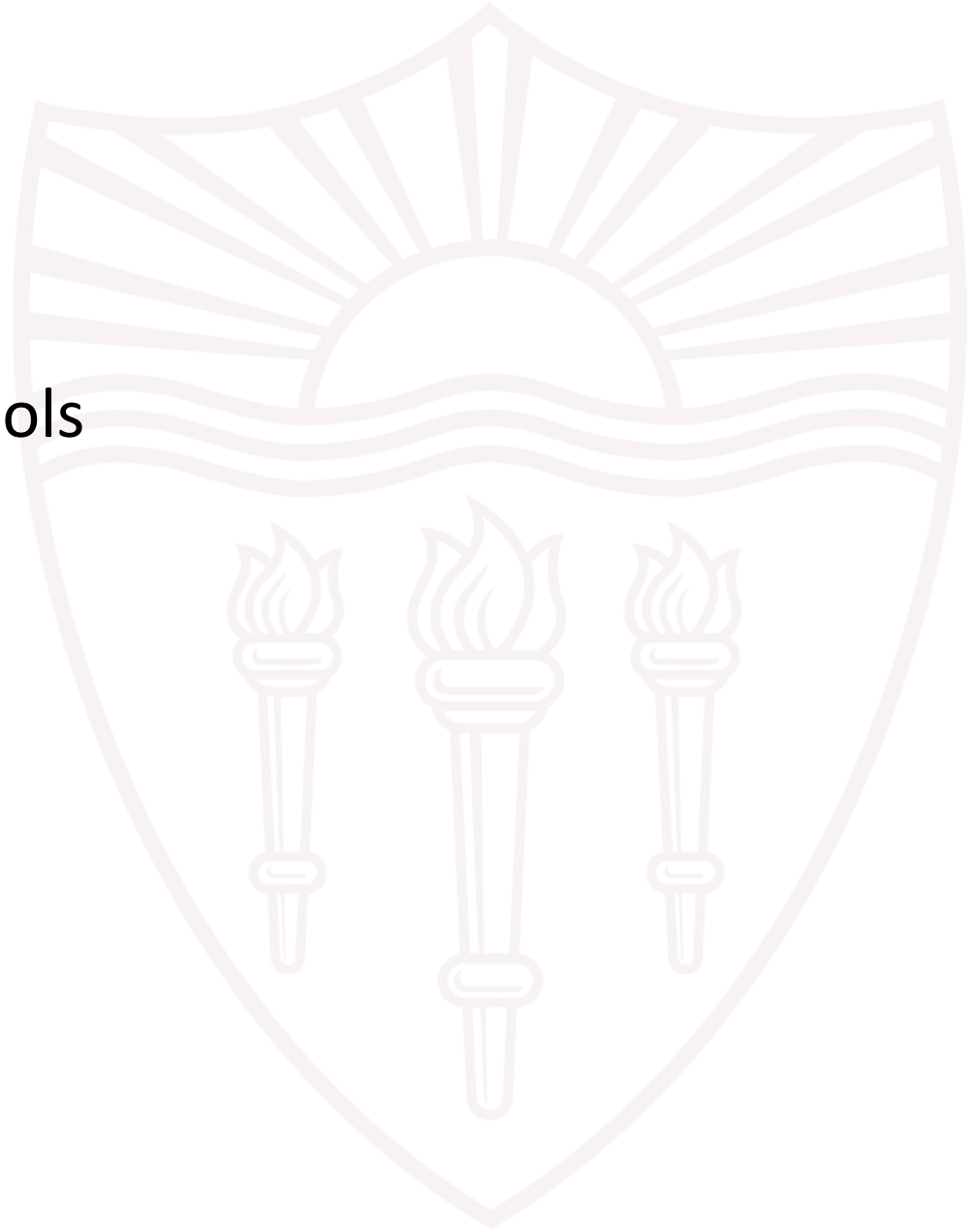
- Anomaly-based Detection in HPC
  - Network traffic is much more diverse than expected.
  - Patterns of behavior in HPC are more regular.
- Analyzing system behavior with Machine Learning.
  - Answering question: Are users running something illegal?
  - Transform patterns of communication into graphs by Message Passing Interface
  - Applied Bayesian-based techniques for classification and matching of graphs
  - 95%–99% accuracy

Figure 2. "Adjacency matrices" for individual runs of a performance benchmark, an atmospheric dynamics simulator, and a linear equation solver SUPERLU. Number of bytes sent between ranks is linearly mapped from dark blue (lowest) to red (highest), with white indicating an absence of communication.[47,48]



Looking to the Future

- Automated static/runtime analysis tools
- New virtualized environments
- “Science Gateways”
- Simulated homomorphic encryption
- Blockchains
- …



References

- https://dl.acm.org/doi/pdf/10.1145/3096742
- Dart, E., Rotman, L., Tierney, B., Hester, M. and Zurawski, J. The science DMZ: A network design pattern for data-intensive science.
- Sommer, R. and Paxson, V. Outside the closed world: On using machine learning for network intrusion detection.
- DeMasi, O., Samak, T. and Bailey, D.H. Identifying HPC codes via performance logs and machine learning.
- Peisert S. Fingerprinting Communication and Computation on HPC Machines. TR LBNL-3483E, Lawrence Berkeley National Laboratory, June 2010.
- Whalen, S., Peisert, S. and Bishop, M. Network-theoretic classification of parallel computation patterns.
- Whalen, S., Peisert, S. and Bishop, M. Multiclass Classification of Distributed Memory Parallel Computations.