HPC Security: Why the Time is Now

Albert Reuther and Andrew Prout
MIT Lincoln Laboratory Supercomputing Center

S-HPC @ SC24, Atlanta, GA

17 Nov 2024

This material is based upon work supported by the Department of the Air Force under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Department of the Air Force.

MIT LINCOLN LABORATORY
SUPERCOMPUTING CENTER

Outline

- **Introduction and Motivation**
- **Expanding Userbase**
- **Changes in Cybersecurity**
- **HPC is Different**

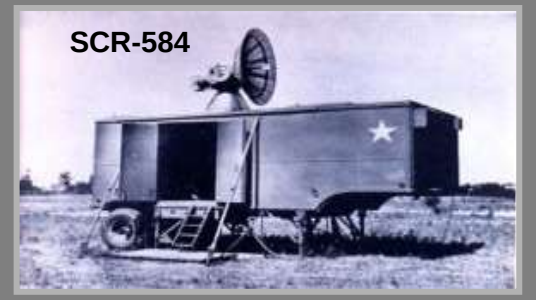
Who We Are – a Little History


MIT Building 20

Mission: ***Development of radar systems and technology***

Main projects: Surveillance radar
Fire control radar
Navigation systems

4000 employees
Designed half of all US WWII radars

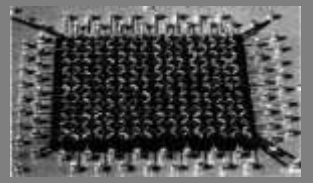


Est. 1951: ***Air defense and technology development***

Main projects: Semi-Automatic Ground Environment (SAGE)

Major Innovations:

Real-Time Computing


Magnetic-core Memory


Light-pen CRT Interface

https://www.ll.mit.edu/sites/default/files/other/doc/2018-04/MIT_Lincoln_Laboratory_history_book.pdf

MIT Lincoln Laboratory

Department of Defense Federally Funded Research and Development Center

Massachusetts Institute of Technology

MIT Lincoln Laboratory

Mission: Technology in Support of National Security

Key Roles:
System architecture engineering
Long-term technology development
System prototyping and demonstration

FY23 Employees: 4419

FY23 Funding: $1.39B

Facilities: 2.1M sq-ft

Mission Areas:

- Air and Missile Defense
- Homeland Protection
- Air Traffic Control
- Communication Systems
- Cyber Security
- Advanced Technology
- Space Control
- ISR Systems and Technology
- Tactical Systems
- Engineering
- Biotechnology & Human Systems

History of Supercomputing at Lincoln Laboratory

1951 Whirlwind

1963 Sketchpad

1956 TX-0

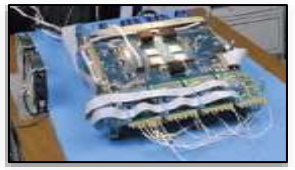
1970 Fast Digital Processor (FDP)

1977 : Lincoln Digital Signal Processor (LDSP)

Late 1970s High-speed FFT pipelined processor

1982 Compact LPC Vocoder

Early 1990s
- APT processor
- RAPTOR processor
- Space-Time Adaptive Processing Library (STAPL)



2002
- ISR Processing and Array Technology (IPAT) processor
- MatlabMPI

2004 Knowledge-Aided Sensor Signal Processing and expert Reasoning (KASSPER) processor

1950 **1960** **1970** **1980** **1990** **2000** **2010**

1953 Magnetic-Core Memory Array

1962 Lincoln Instrument Computer (LINC)

1970 GENESYS

1978 Micro-Processor Based LPC Vocoder (LPCM)

1992 Radar Surveillance Technology Experimental Radar (RSTER) processor

1958
- AN/FSQ-7 (Whirlwind II)
- Average Response Computer (ARC)
- CG-24
- TX-2

1974 Lincoln Digital Voice Terminal (LDVT)

1999 Parallel Vector Library (PVL)

2003 pMatlab

2004
- pMapper gridMatlab
- LLGrid TX-2500

2007
- Parallel Vector Tile Optimizing Library (PVTOL)
- Real-Time Communication Layer (RTCL)

2012 D4M

2014 LLGrid TX-Green

2015 BigDAWG

2016 Lincoln Laboratory Supercomputing Center (LLSC)

https://www.ll.mit.edu/r-d/cyber-security-and-information-sciences/lincoln-laboratory-supercomputing-center

Lincoln Laboratory Supercomputing Center (LLSC) Role

Mission Areas

- Air and Missile Defense
- Homeland Protection
- Air Traffic Control
- Communication Systems
- Cyber Security
- Advanced Technology
- Space Control
- ISR Systems and Technology
- Tactical Systems
- Engineering
- Biotechnology & Human Systems

Vast Data & Computation

OSINT


Weather


HUMINT

C2


Ground


Maritime


Air


Space


Cyber

LLSC develops & deploys unique, energy-efficient supercomputing that provides cross-mission

- **Data centers, hardware, software, user support, and pioneering research**
- **100x more productive than standard supercomputing[1]**
- **100x more performance than standard cloud[2]**

[1] *Interactive Grid Computing at Lincoln Laboratory*, Bliss et al., LL Journal, 2006
[2] *Scalability! But at what COST?*, McSherry et al., HotOS XV, 2015

Outline

- Introduction and Motivation
- Expanding Userbase
- Changes in Cybersecurity
- HPC is Different

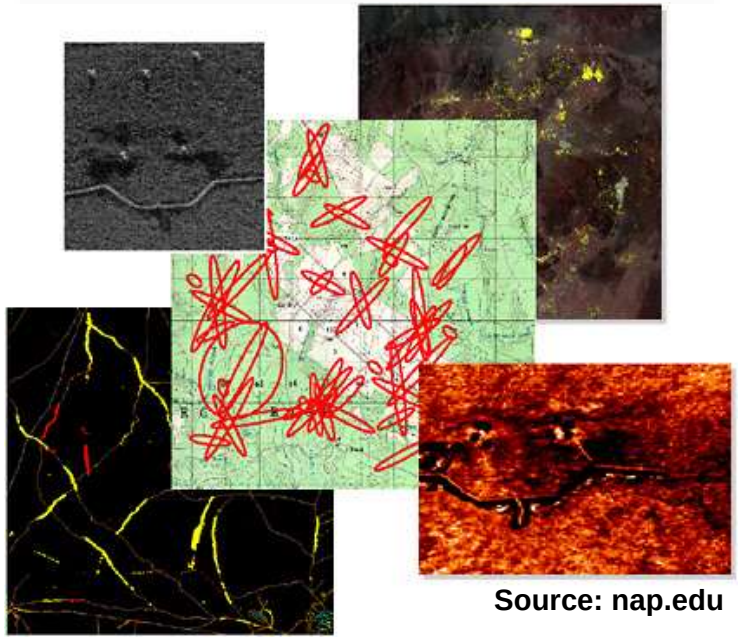
Supercomputing Application Types

Physical Modeling and Simulation

Source: olcf.ornl.gov

- Molecular Dynamics
- Finite Element Analysis
- Computational Fluid Dynamics
- Multi-Physics Modeling
- Weather Simulation
- Etc.

Signal and Image Processing



Source: nap.edu

- Radar (GMTI, SAR, ISAR)
- Sonar
- Wide Area Motion Tracking
- Signal Intelligence (SIGINT)
- Electronic Warfare
- Etc.

Machine Learning and Data Analysis

Source: cooperenvironmental.com

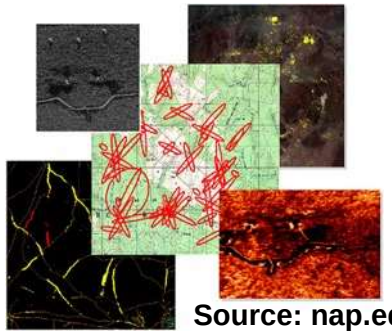
- Target Tracking
- Phenomenology Detection
- Visualization
- Data Clustering
- Graph/Network Analysis
- Etc.

Supercomputing Application Types

Physical Modeling and Simulation

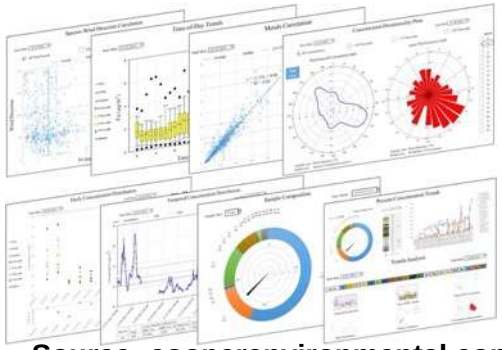
Source: olcf.ornl.gov

Signal and Image Processing



Source: nap.edu

Machine Learning and Data Analysis



Source: cooperenvironmental.com

1970s | 1980s | 1990s | 2000s | 2010s | 2020s | 2030s

Cray 2

Cray Y-MP

Intel Paragon

IBM Blue Gene

HPE Cray EX

STAP Processor

DARPA KASSPER

Cell Cluster AFRL HPCMP

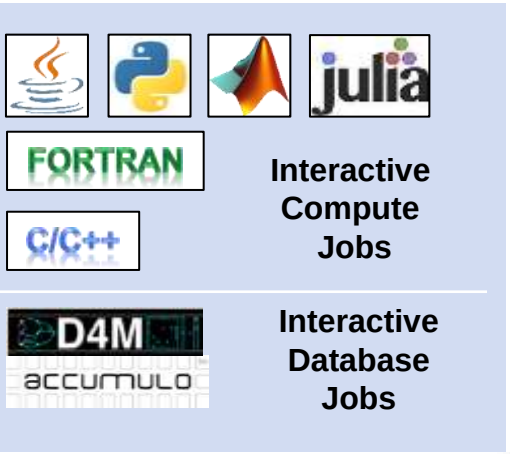
Radient Falcon

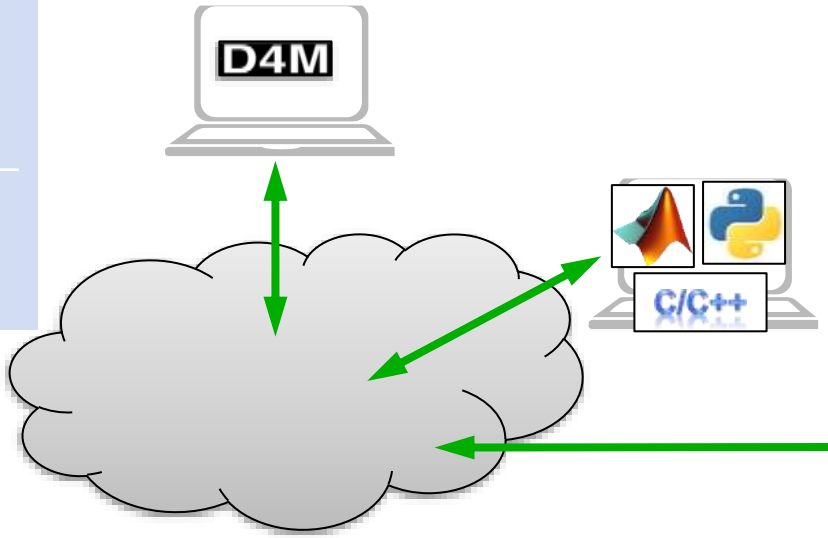
Google BigTable

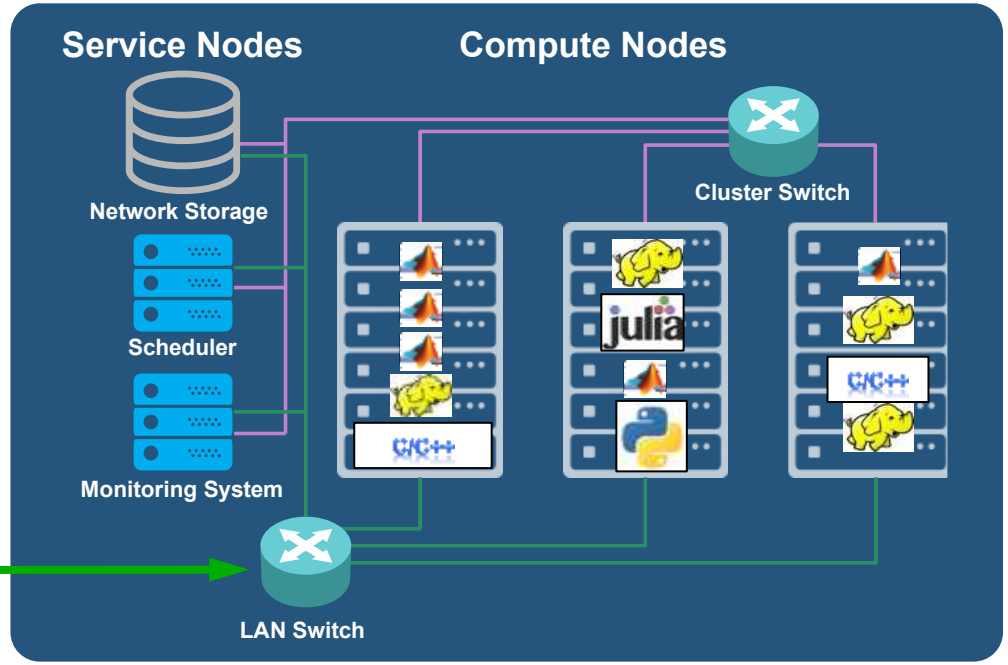
MATLAB PCT

STAP = Space-Time Adaptive Processing
PCT = MATLAB Parallel Computing Toolbox

Interactive HPC Enables Broader Userbase







- LLSC provides a software platform that allows users to
  - Launch interactive compute jobs from their desktop
  - Share large volumes of project data
- The LLSC experience provides
  - Reference datasets pre-positioned in databases
  - Software modules and training to reduce user ramp up time
- LLSC responsible for cybersecurity across entire system

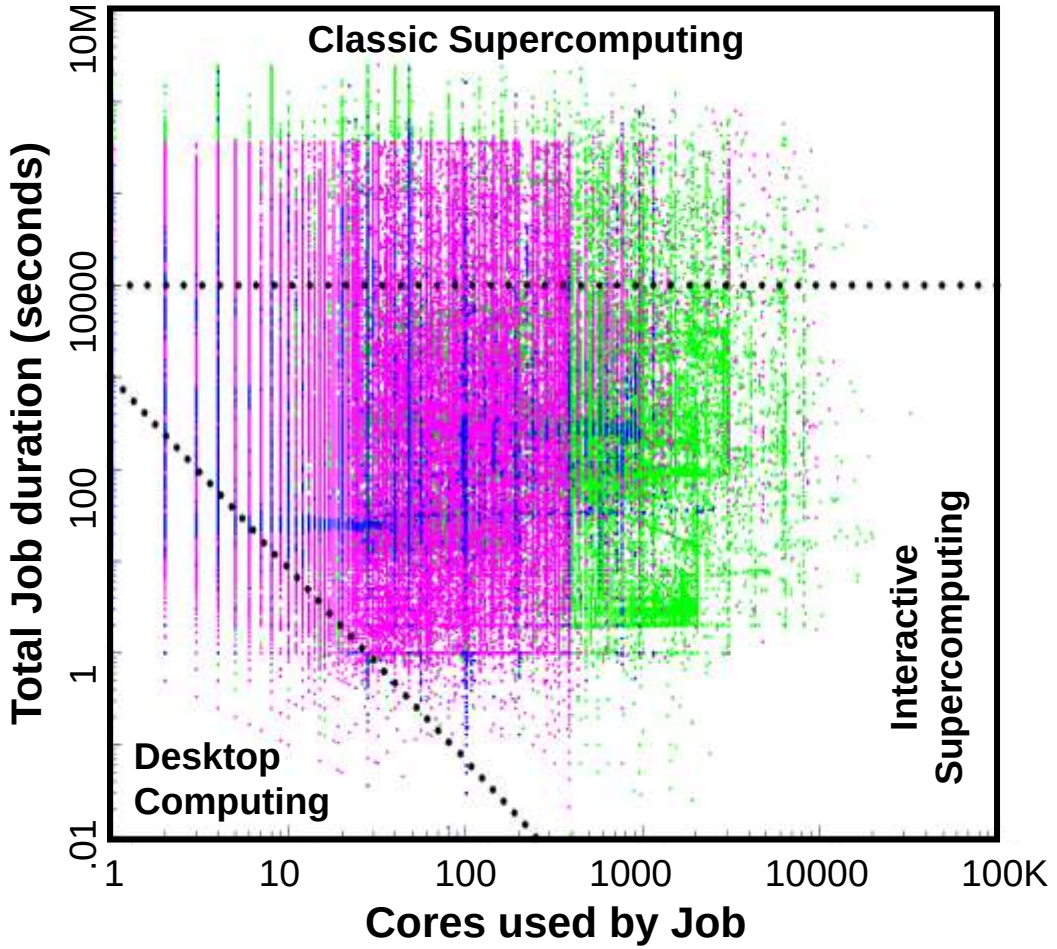
- Rapid Prototyping
  - Algorithm development
  - Data analysis
  - Machine learning training
- Application Steering
  - Real-time / streaming data analytics
  - Debugging/validation
- Visualization

N. T. Bliss, R. Bond, J. Kepner, H. Kim, and A. Reuther, "Interactive Grid Computing at Lincoln Laboratory," *Lincoln Lab. J.*, vol. 16, no. 1, p. 165, 2006.

Changes in Desktop Cybersecurity

- Desktop cybersecurity is now big business
- Scientific computing is generally not the customer
- Magnitude of problems that desktops/laptops can solve is shrinking

Userbase Enablement



Interactive HPC enables users to run jobs that don't fit on laptop/desktop

- **Desktop Computing**
  - **CPU-time <20 minutes**
- **Classic Supercomputing**
  - **Wall-clock time >3 hours**
- **Interactive Supercomputing**
  - **Between desktop and classic supercomputing**
  - **Shortens the "time to insight"**
  - **Ten development turns/day instead of one turn/week**

TX-Green (89020 Cores) TX-E1 (32000 Cores)
TX-C (6740 Cores)

Note: Profile of jobs executed between April 1st, 2023 and March 31st, 2024.

Diversified Software, Tools, and Workflows

High Productivity Languages & Environments

MATLAB
SIMULINK
GNU Octave
julia
python
OPEN OnDemand
jupyterlab
R

Machine Learning Frameworks

PYTORCH
DASK
TensorFlow
TensorBoard
RAY
flux
APACHE Spark
ANACONDA
Powered by Continuum Analytics

3rd Party / Commercial Software

XCrySDen
X-window CRYstalline Structures and DENsities
FEKO
STAR-CCM+
LS-DYNA
Cart3D
ANSYS FLUENT
NCBI BLAST
SIMULIA ABAQUS
ParaView
WRF Weather Research & Forecasting

High Performance Languages

Fortran
C
C++
MPI

Compilers, Debuggers, Performance Analysis

VS Code
oneAPI
intel oneAPI
GCC
NVIDIA HPC
NVIDIA Nsight Compute

High Performance Databases

SciDB
accumulo

Outline

- Introduction and Motivation
- Expanding Userbase
- Changes in Cybersecurity
- HPC is Different

Changes in Cybersecurity World

- The “thick outer shell” of HPC systems is no longer enough
  - Another way to say we need ZeroTrust inside the HPC
  - HPC ZeroTrust is very different than Enterprise IT ZeroTrust
- Coarse “buckets” to not enough separation
  - Userbase expansion, multi-tenancy: Less single-mission dedicated HPC
  - Insider threats: Edward Snowden, Harold Martin, Reality Winner
  - Relying on personnel security alone is no longer viable
- HPC is now a target
  - Artifact of Enterprise cybersecurity getting better, or just attacks getting more numerous?
  - Cryptocurrency mining: They may not be targeting the data, but we’ll need to analyze what they had access to in the after-action report

Fine-grained always-on need-to-know separation is now what’s expected

Why Worry About HPC Security?

Bloomberg

Subscribe

Technology | Cybersecurity

Hackers Target European Supercomputers Researching Covid-19

HPC wire

Search... Go

Hacking Streak Forces European Supercomputers Offline in Midst of COVID-19 Research Effort

By Oliver Peckham

May 18, 2020

Suche | Kontakt | Impressum | Datenschutz

lrz Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften

AKTUELLES FAQ SERVICES SERVICEDESK IDM-PORTAL

www.lrz.de / presse / ereignisse / 2020-06-03_InterviewSecurityIncident

Supercomputers offline across Europe

UC-NG at the LRZ - Photo: V. Hohenegger for LRZ

Security Boulevard

Home » Security Boulevard (Original) » News » 11-Plus Supercomputers Hacked With Cryptominers

11-Plus Supercomputers Hacked With Cryptominers

by Richi Jennings on May 18, 2020

Cryptocurrency mining: we have a bullseye on our back from that crowd

https://www.educv.de/blog/post-2021-02-17-analyzing-a-compromised-hpc-cluster/

Why Worry About HPC Security?

- **Potential for losses**
  - **Economic loss when HPC system has to be rebuilt and no jobs being run**
  - **Risk of not getting new grants and projects**
  - **Risk of no longer being able to trust the integrity of data and jobs/work on HPC system**
- **Will likely be written into future government funded project requirements**

Outline

- Introduction and Motivation
- Expanding Userbase
- Changes in Cybersecurity
- HPC is Different

Three Big Environment Differences Affect Security Enterprise vs. HPC

Enterprise Systems

- Huge number of files on data storage
  - Change slowly – human editing pace
  - Change is localized – many changes to same file in quick succession
- Users from same organization
  - Under same rules and requirements
  - Same organizational goals
- Homogeneous applications
  - Users run limited number of sanctioned applications on own computer
    - Office applications
    - Browsers
  - Business apps and databases centralized

HPC Systems

- Huge number of files on data storage
  - Many files changing quickly – mostly from executing parallel jobs
  - Change occurring in many places in file system corresponding to parallel jobs
- Users from multiple organizations
  - Collaboration
  - Differing organizational rules and goals
- Every user is software developer
  - Prototyping code and simulations
  - Little thought to designing with security
  - Open source software and libraries

HPC is Different

- Enterprise cybersecurity tools expect a large “budget” to work with
  - Measured against “human will notice” standard
  - Can use lots of CPU, can delay kernel syscalls, can limit bandwidth
- Assume each “node” is the system boundary
  - Does not view the HPC system as a whole
- Chicken and the egg problem:
  - HPC keeps getting waivers, there's no market ROI for a vendor to tune their tools for HPC use
  - There's no market for HPC security tools, HPC will continue to need waivers
- People here in this room need to chart path to improvement
  - Specialized HPC cybersecurity solutions
  - May be investments in vendors, may be investments in enhancing open source projects

HPC Security Path Forward

- Early cybersecurity research focused on threat modeling, theoretically sound solutions
  - Bell–LaPadula model (1973)
  - DoD Orange Book (1983)
- Modern guidance has gotten more tactical
  - DoD STIGs: “ensure /proc/x/y/z is set to 1”
- As the guidance got more specific, more assumptions were baked in
  - The system’s use case was assumed, and specific to enterprise IT applications
  - Need to acknowledge the shortcomings of this guidance when applied to HPC: the guidance was never written for us in the first place!
- Not only need to “exempt” HPC from wrong guidance, also need to identify HPC-specific gaps and develop new solutions

Need to go back to the basics of threat modeling and review all guidance

Conclusion

- **Time is now for HPC security**
- **Different from Enterprise**
- **We as community need to set HPC cybersecurity standards**
  - **Before someone who doesn't know HPC does it for us!**
- **Work has begun – NIST SP 800-223**
  - **Enterprise cybersecurity is culmination of learned experiences over last 40-50 years**
  - **HPC cybersecurity needs to follow same process of evolving – just quicker**

NIST Special Publication 800
NIST SP 800-223

High-Performance Computing Security

Architecture, Threat Analysis, and Security Posture

Yang Guo
Ramaswamy Chandramouli
Computer Security Division
Information Technology Laboratory

Lowell Wofford
Amazon.com, Inc.

Rickey Gregg
Gary Key
HPCMP

Antwan Clark
Laboratory for Physical Sciences

Catherine Hinton
Los Alamos National Laboratory

Andrew Prout
Albert Reuther
MIT Lincoln Laboratory

Ryan Adamson
Oak Ridge National Laboratory

Aron Warren
Sandia National Laboratories

Purushotham Bangalore
University of Alabama

Erik Deumens
University of Florida

Csilla Farkas
University of South Carolina

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-223

February 2024

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

https://csrc.nist.gov/pubs/sp/800/223/final
https://csrc.nist.gov/Projects/high-performance-computing-security

Acknowledgements

- LaToya Anderson
- William Arcand
- William Bergeron
- David Bestor
- Alex Bonn
- Daniel Burrill
- Chansup Byun
- Vijay Gadepally
- Michael Houle
- Matthew Hubbell
- Hayden Jananthan
- Michael Jones
- Jeremy Kepner
- Piotr Luszczek
- Peter Michaleas
- Lauren Milechin
- Guillermo Morales
- Julie Mullen
- Antonio Rosa
- Siddharth Samsi
- Jason Williams
- Charles Yee

Contact


reuther@ll.mit.edu
aprout@ll.mit.edu
MIT LINCOLN LABORATORY
SUPERCOMPUTING CENTER