OAK RIDGE National Laboratory

# The Role of Precise Time in Supercomputers

Terry Jones, Oak Ridge National Laboratory

March 16, 2023

2023 WSTS

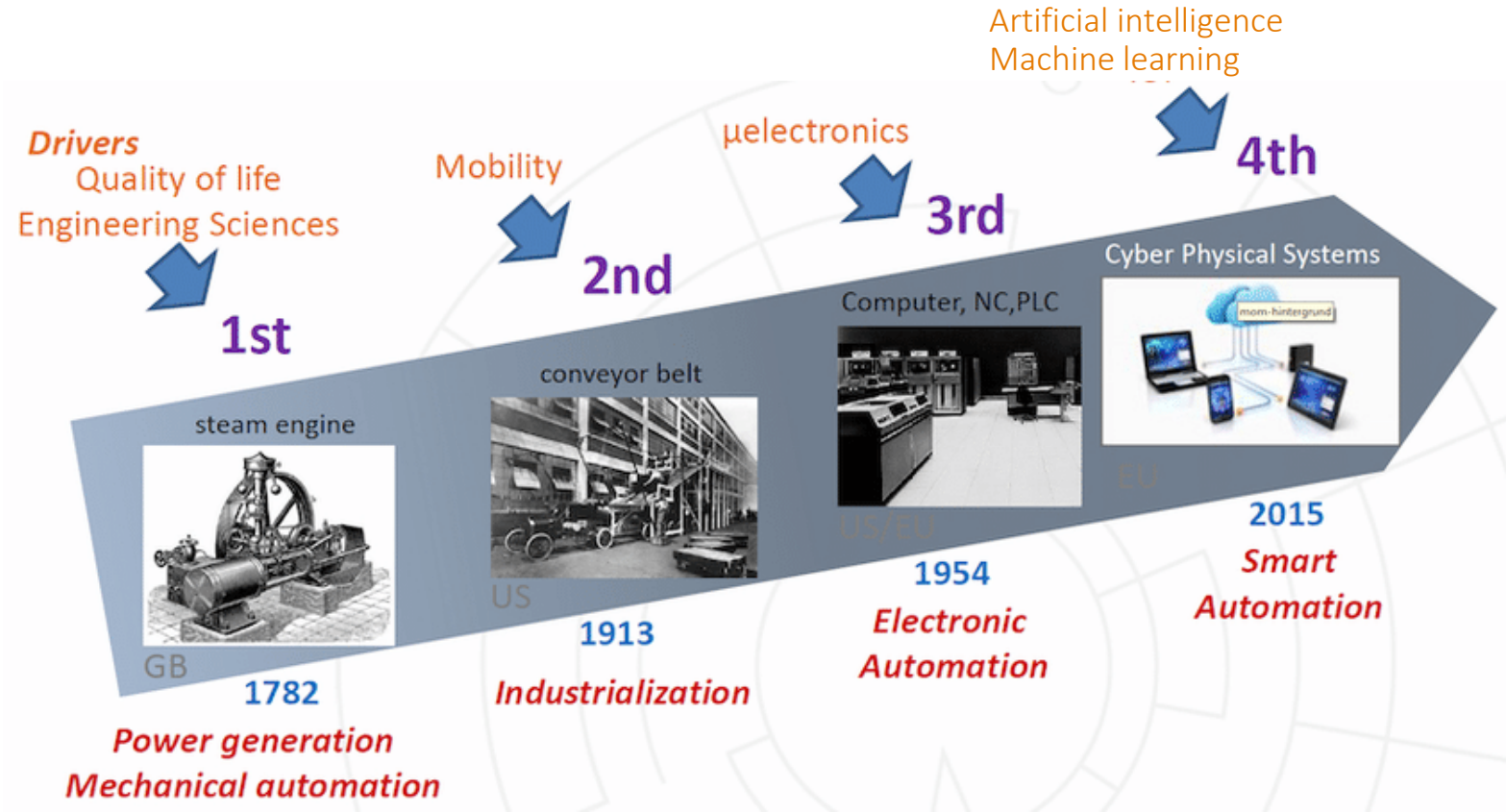
ORNL is managed by UT-Battelle, LLC for the US Department of Energy

U.S. DEPARTMENT OF ENERGY

# Outline

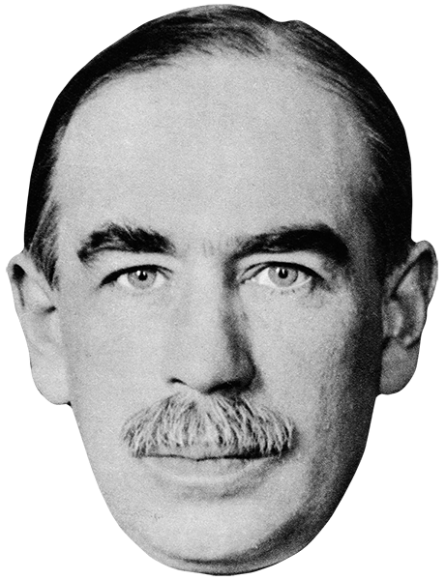
- What?
  - Why should the average person care about the successes of supercomputers?
- So What?
  - How do supercomputers utilize precise timing?
  - How precise are we talking about?
- Now What?
  - Where is supercomputing headed and are there any timing ramifications?

# The March of Science




"Equipped with his five senses, man explores the universe around him and calls the adventure Science."

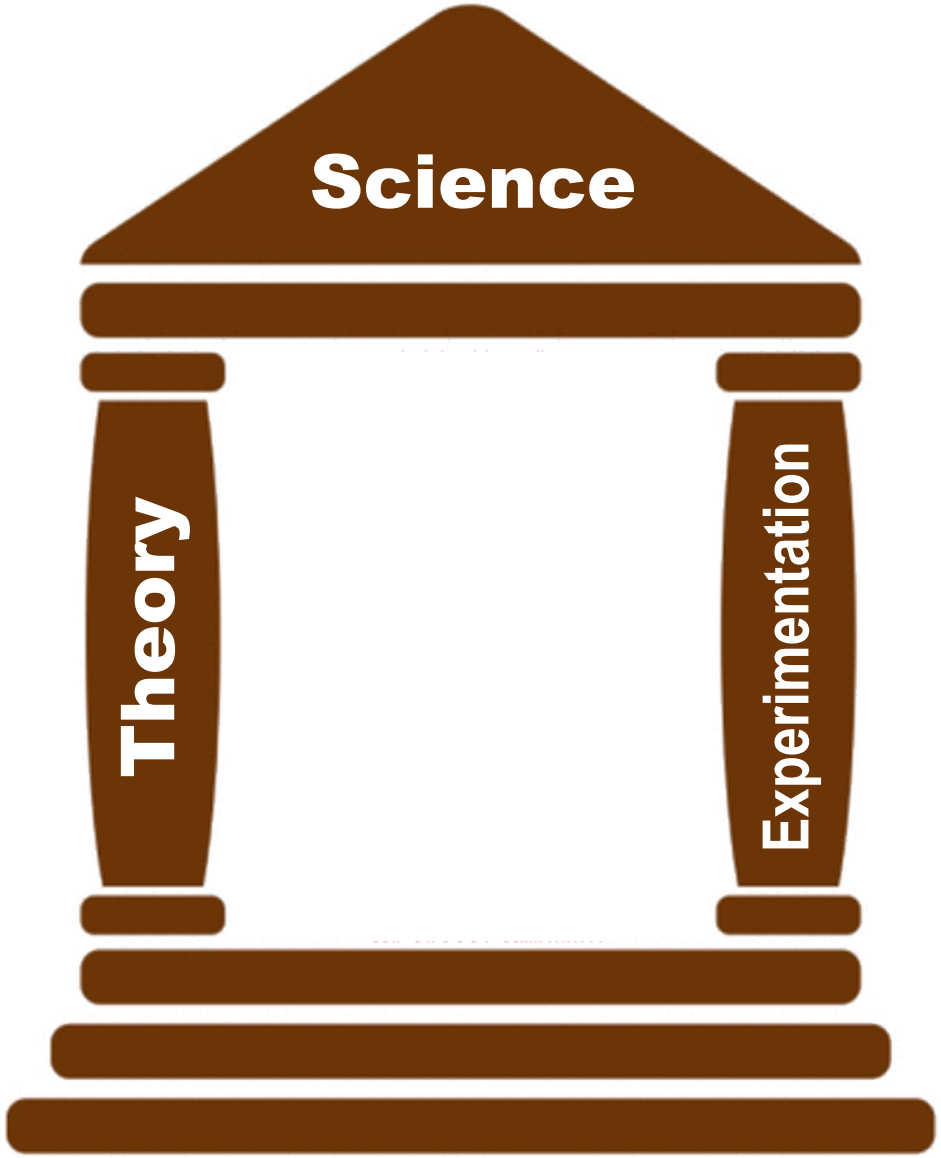
- Edwin Hubble



"Ideas shape the course of history."

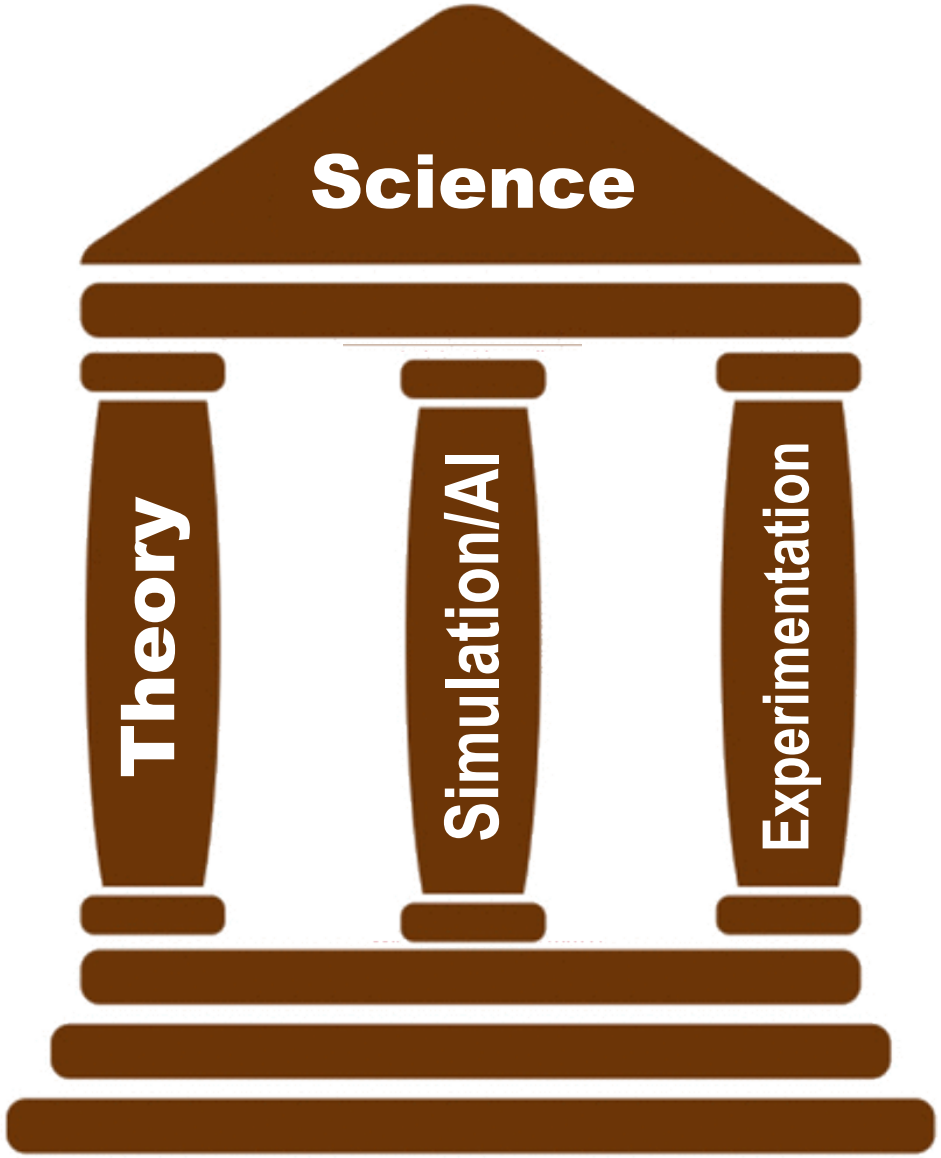
- John Maynard Keynes

# The Two Pillars of Science



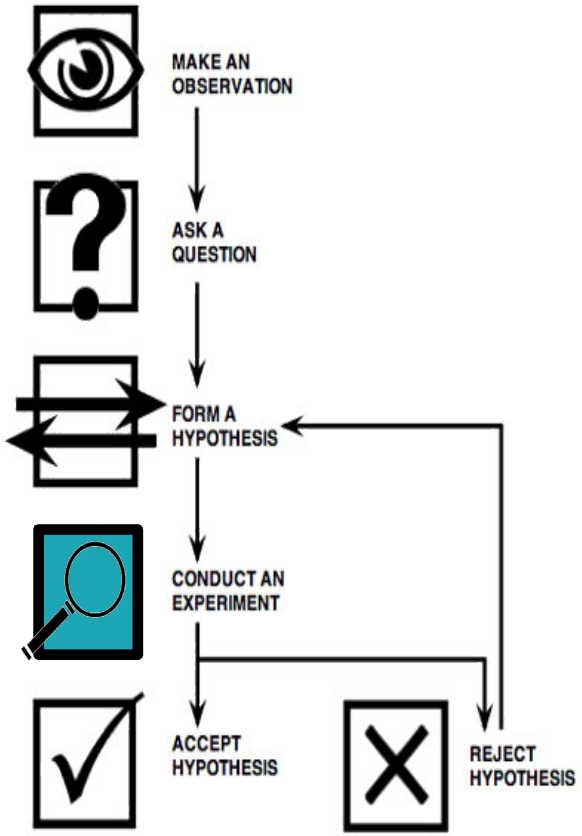
# The ~~Two~~ Three Pillars of Science

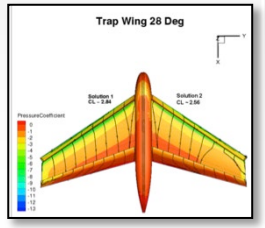
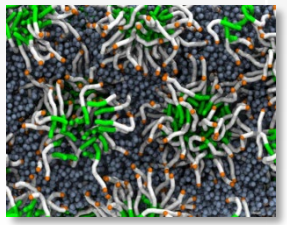
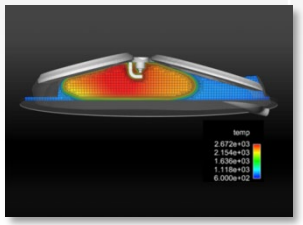
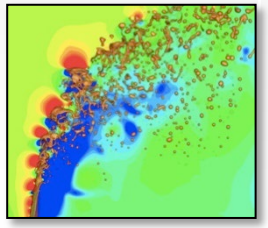
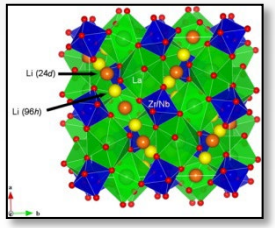
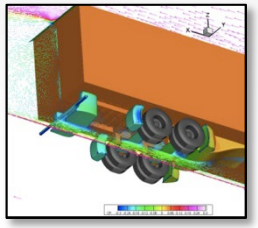
Computers have changed the way we conduct experiments. Given enough computer power, we can perform accurate experiments more quickly, more cheaply, and often with greater control.



# The Impact of High Performance Computing

- Over a 5 year span, 38% of the international innovation “R&D 100” awards went to US National Labs




| BOEING | P&G | Ford | United Technologies Research Center | BOSCH | SMART TRUCK |
|---|---|---|---|---|---|
| **Aircraft design** | **Consumer products** | **Engine cycle-to-cycle variation** | **Jet engine efficiency** | **Li-ion batteries** | **Long-haul truck fuel efficiency** |
| Simulating takeoff and landing scenarios improved a critical code for estimating characteristics of commercial aircraft, including lift, drag, and controllability | Leadership computing and molecular dynamics software advanced understanding of chemical processes that can limit product shelf life | Emerging model of engine cyclic variation will apply thousands of processors to a challenging problem | Accurate predictions of atomization of liquid fuel by aerodynamic forces enhance combustion stability, improve efficiency, and reduce emissions | New classes of solid inorganic Li-ion electrolytes could deliver high ionic and low electronic conductivity and good electrochemical stability | Simulations reduced by 50% the time to develop a unique system of add-on parts that increases fuel efficiency by 7−12% |
|  |  |  |  |  |  |





OAK RIDGE National Laboratory

# A Brief History of High Performance Computing


| Gigaflops | Teraflops | Petaflops | Exaflops |
|---|---|---|---|
| $10^9$ | $10^{12}$ | $10^{15}$ | $10^{18}$ |
| 1985 | 1998 | 2008 | 2022 |

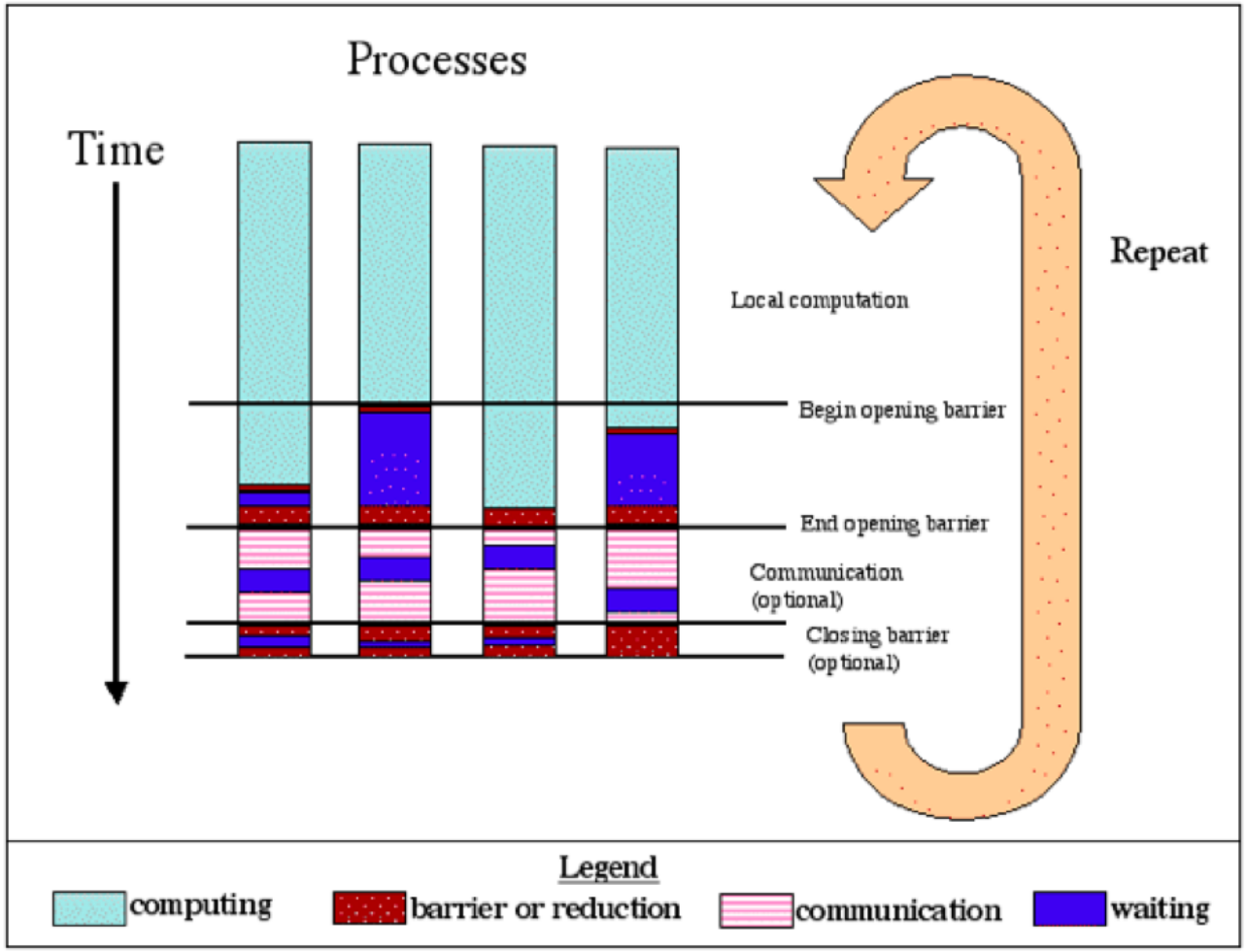
Calendar Year



OAK RIDGE National Laboratory

# The *Tightly Coupled** Case – Bulk Synchronous Parallel Programs[1]

- Consider a program with 4 threads
- Each thread cycles between several states:
  - Computing
  - Barrier
  - Communication
  - Waiting



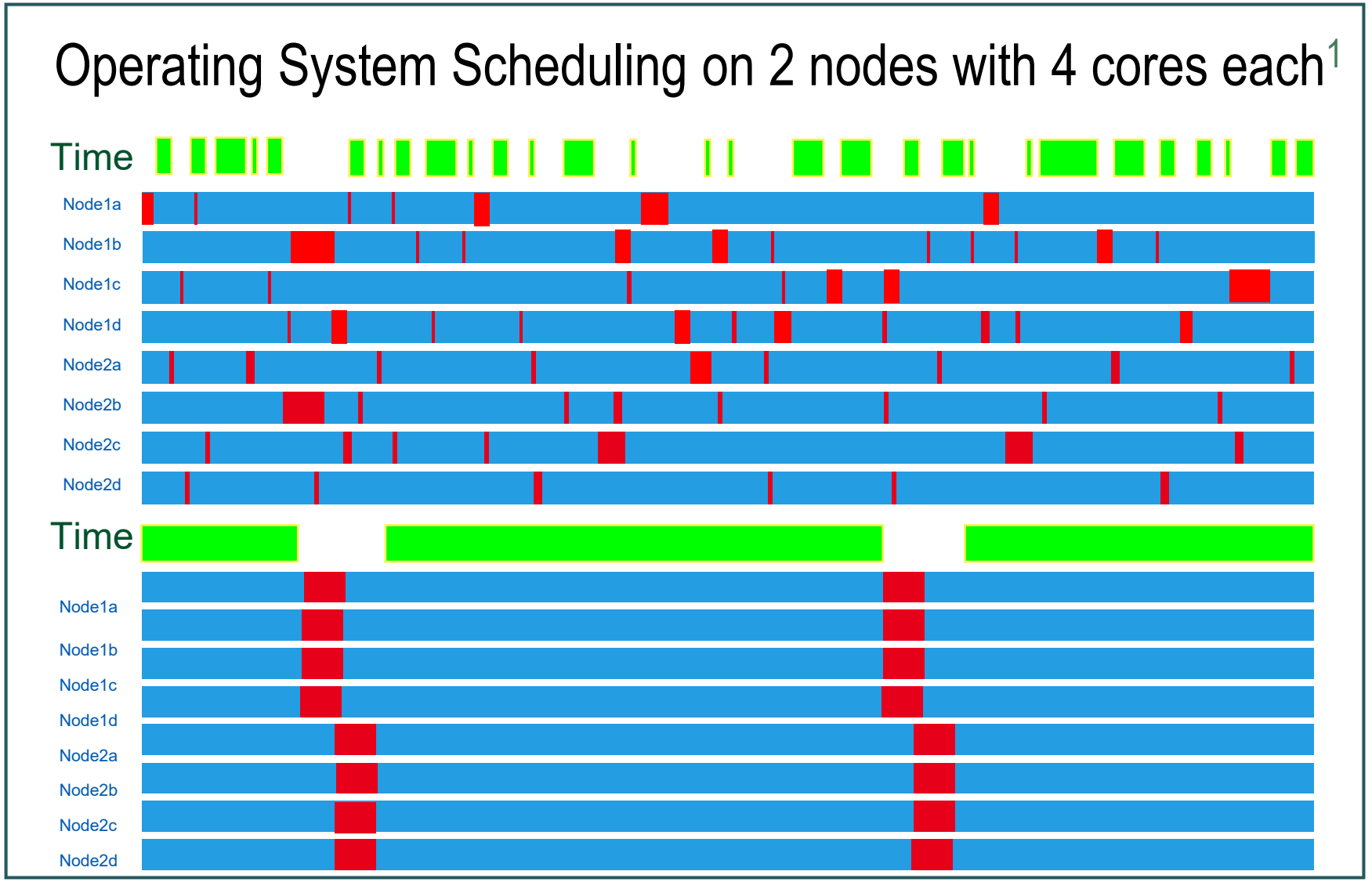

* Tightly Coupled = Within one machine

1. Leslie G. Valiant, A bridging model for parallel computation, Communications of the ACM, Volume 33 Issue 8, Aug. 1990

# The *Tightly Coupled* Case (continued)

- Goal is to eliminate waiting
- Have each core/thread complete at same time
- Q: But what if a thread gets delayed by something external?

  A: We need precise time agreement!

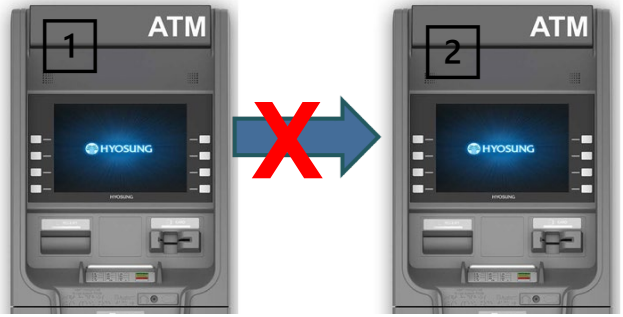



1. Terry Jones, Shawn Dawson, Rob Neely, William Tuel, Larry Brenner, Jeff Fier, Robert Blackmore, Pat Caffrey, Brian Maskell, Paul Tomlinson, and Mark Roberts, "Improving the Scalability of Parallel Jobs by adding Parallel Awareness to the Operating System." In The 16th International Conference for High Performance Computing, Networking, Storage and Analysis (SC'03). Phoenix, AZ, November 2003. https://doi.org/10.1109/SC.2003.10024

# The *Loosely Coupled** Case – Consistency of Shared Data

Step 1: Deposit $100 initiated in ATM [1]


Step 2: ATM [1] fails to update ATM [2]




Step 3: What happens next depends on your Consistency protocol:
(a) Deny the deposit
(b) Accept the deposit & reconcile later

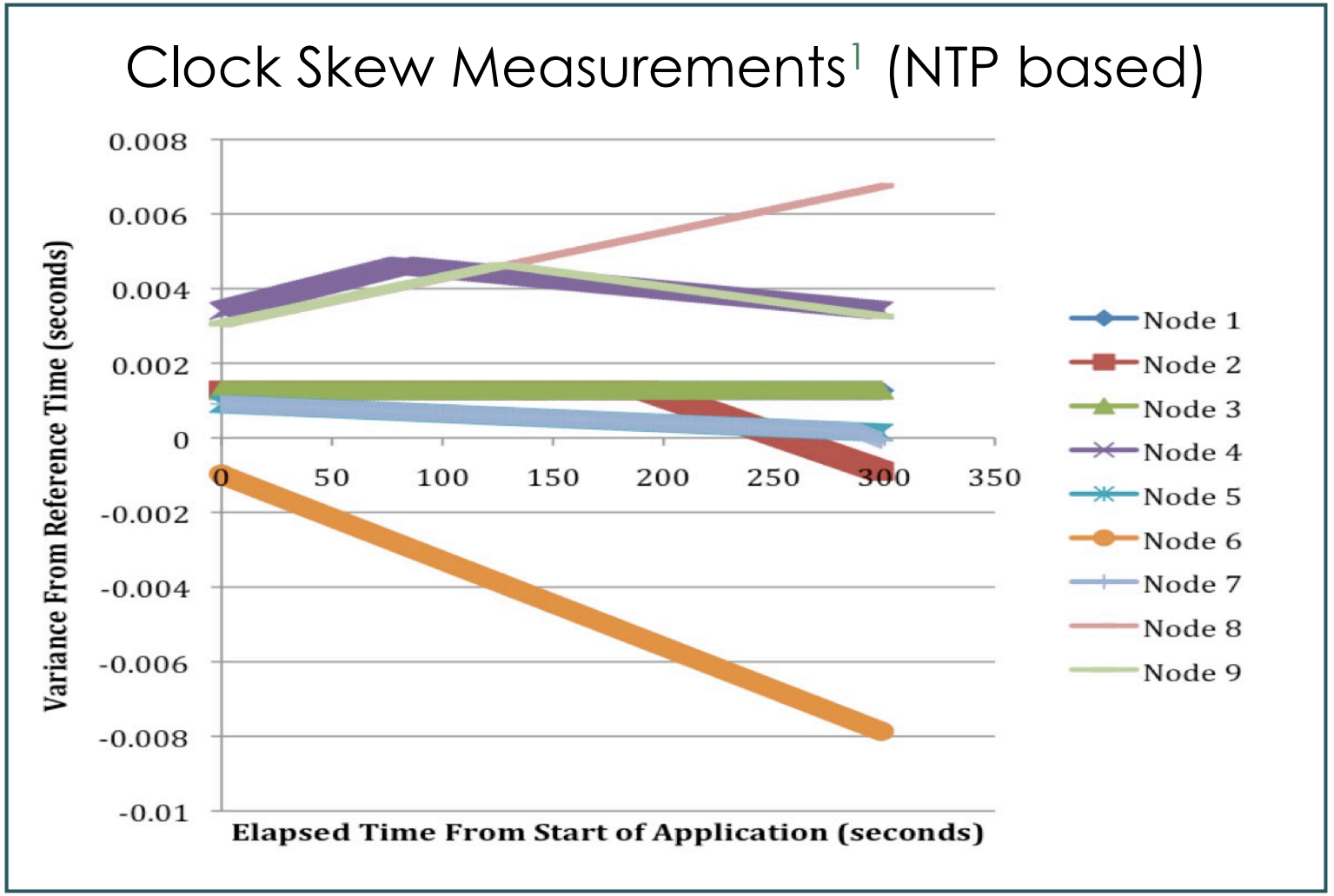

A primary characteristic of a Consistency protocols Is how many messages are required to maintain a well-known state.

If both machines have a common notion of time, there is a nice optimization called "***leases***" which Results in fewer messages to maintain well-known state.
We need precise time agreement!

* Loosely Coupled = Spanning multiple machines

# …So We Want Precise Time Agreement In These Cases

- In the early days of massively parallel systems
- Each node was an NTP client
- And NTP wasn't particularly accurate on those systems




1. Terry Jones, and Gregory Koenig. "Clock Synchronization in High-end Computing Environments: A Strategy for Minimizing Clock Variance at Runtime." Journal of Concurrency and Computation: Practice & Experience, Volume 25, Issue 6, doi: 10.1002/cpe.2868, pages 881-897, April 25, 2013. https://doi.org/10.1002/cpe.2868

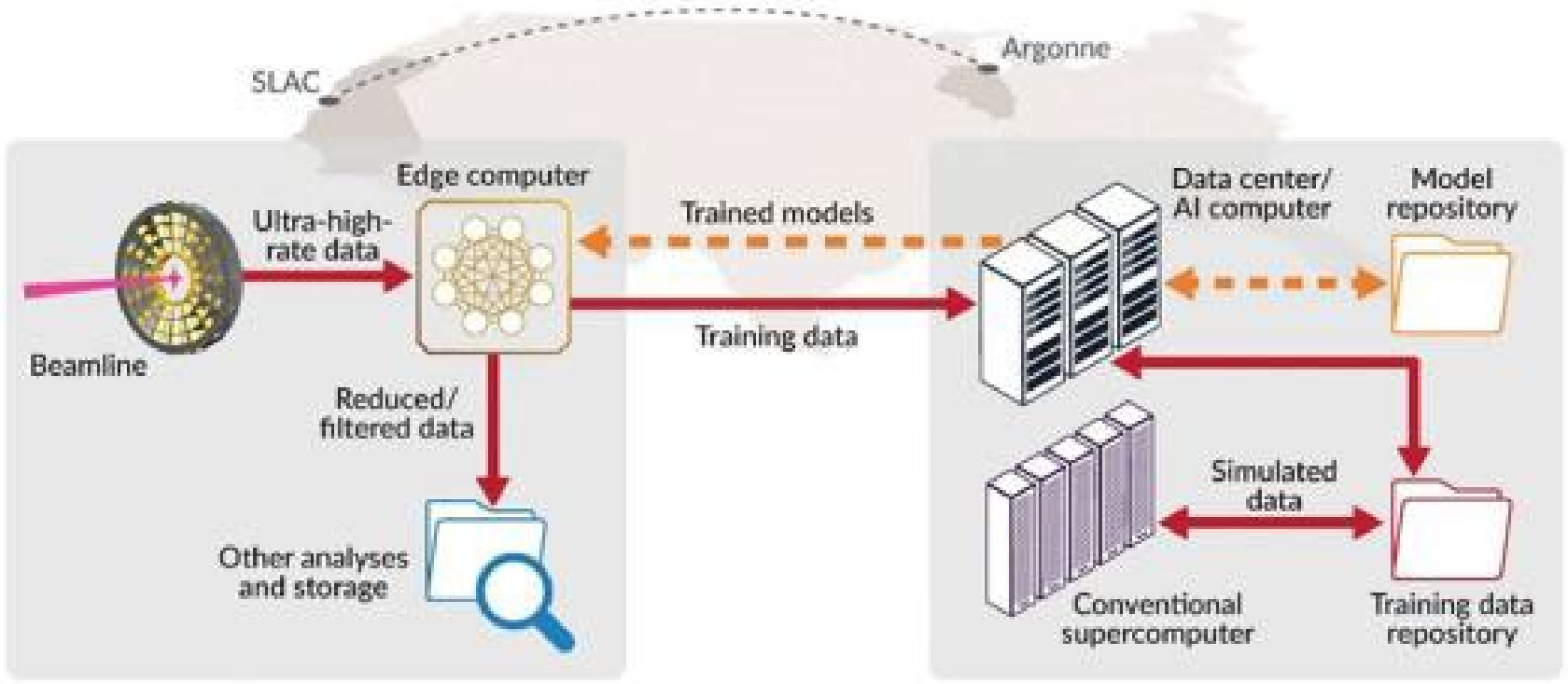
# Today, Each HPC Machine Is Designed With Jitter In Mind

- Heartbeat available from custom interconnect
  - Accuracies in nano-seconds
- Cores set aside for OS interruptions

- But getting every thread to complete at the same time is still a challenging problem!

# Tomorrow, the Integrated Supercomputer Continuum

- Workflows that extend the supercomputer to concurrent use with other supercomputers and machines



# Summary

- Simulation and high performance computing are critical elements to scientific advancement
- Today's supercomputers are massively parallel machines built with high-end commodity (off-the-shelf) parts and specific customization
- Bulk Synchronous algorithms require anti-jitter measures which rely on coordination and available cycles (must solve the tightly-coupled problem)
- Tomorrow's workflows are projected to extend past machines, new levels of coordination will need to be explored (must solve loosely-coupled)

Questions & Comments Welcome: trj@ornl.gov

# Acknowledgements


This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.