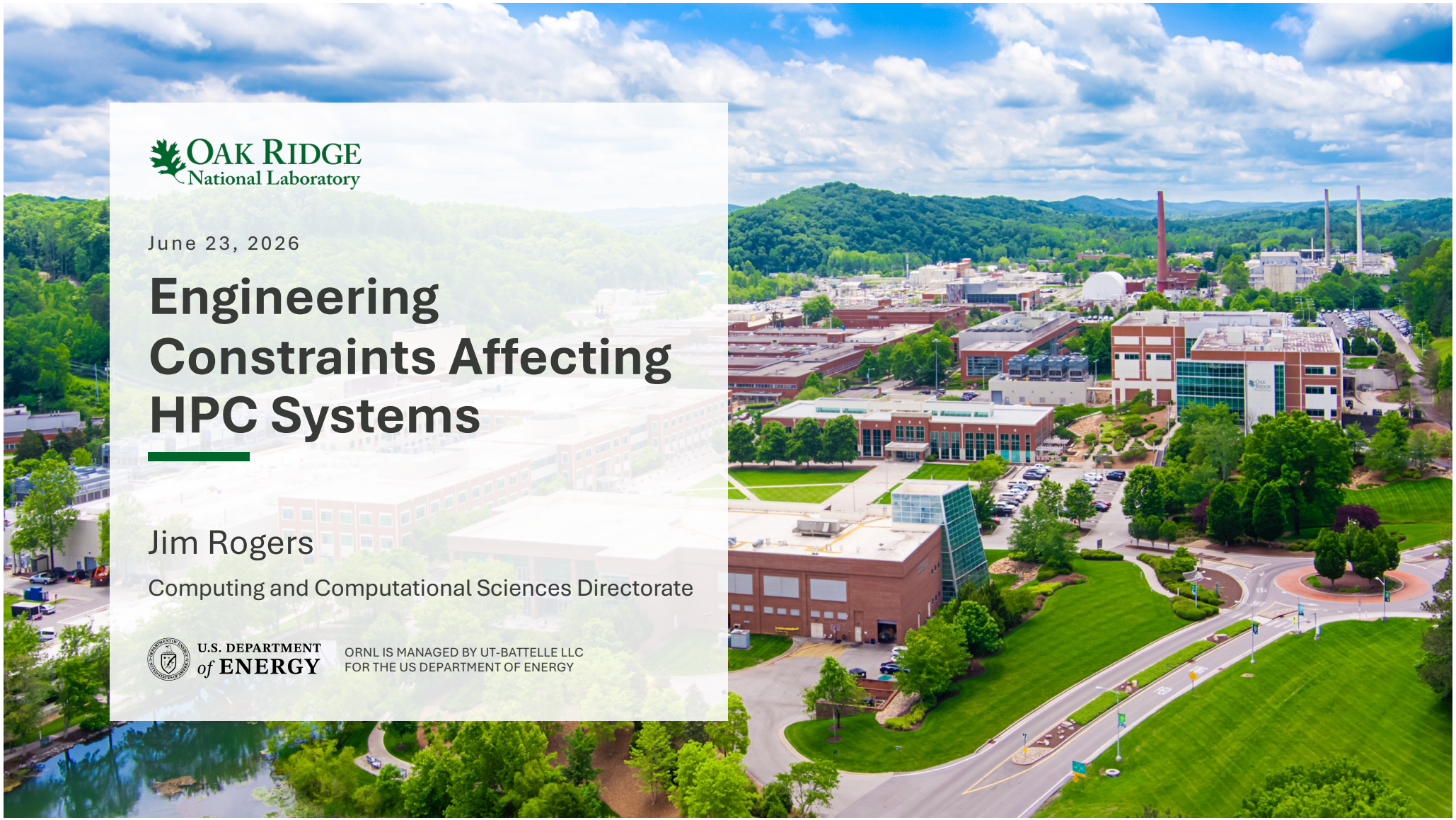


OAK RIDGE National Laboratory

June 23, 2026

Engineering Constraints Affecting HPC Systems

Jim Rogers

Computing and Computational Sciences Directorate

U.S. DEPARTMENT of ENERGY

ORNL IS MANAGED BY UT-BATTELLE LLC
FOR THE US DEPARTMENT OF ENERGY

Balancing Hard Constraints with Co-design

Titan (2012 – 2019)

Constraints

- Contract executed as a sole-source upgrade of an existing computer (Jaguar).
  - 2x6-core AMD -> 16-core CPU + NVIDIA Kepler GPU (1C:1G)
  - No specific opportunity to compete
- Immense schedule pressure to deliver to production as a rolling upgrade of an operational system
  - Cascading upgrade strategy changed the system size continuously for more than a year.
  - Impact on global bandwidth during upgrade
- Electrical power demand jumped from 350W TDP to 515 TDP per node.
  - Required 480VAC 100A connection/cabinet

Wins

- ECOPhlex phase-change cooling solution ejected waste heat to refrigerant, and then to chilled water. ITUE of 1.29 vs. typical air-cooled solution at ~1.8.
- Dramatic improvement in both Time to Solution and Energy to Solution v. CPU-only solutions

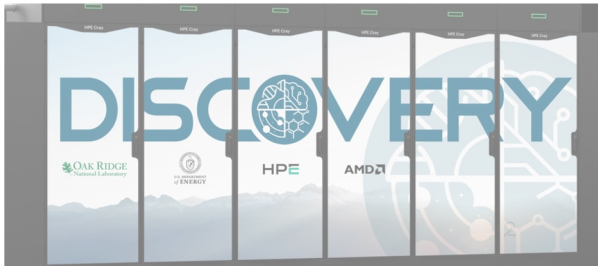
OAK RIDGE National Laboratory

2012

2018

2023

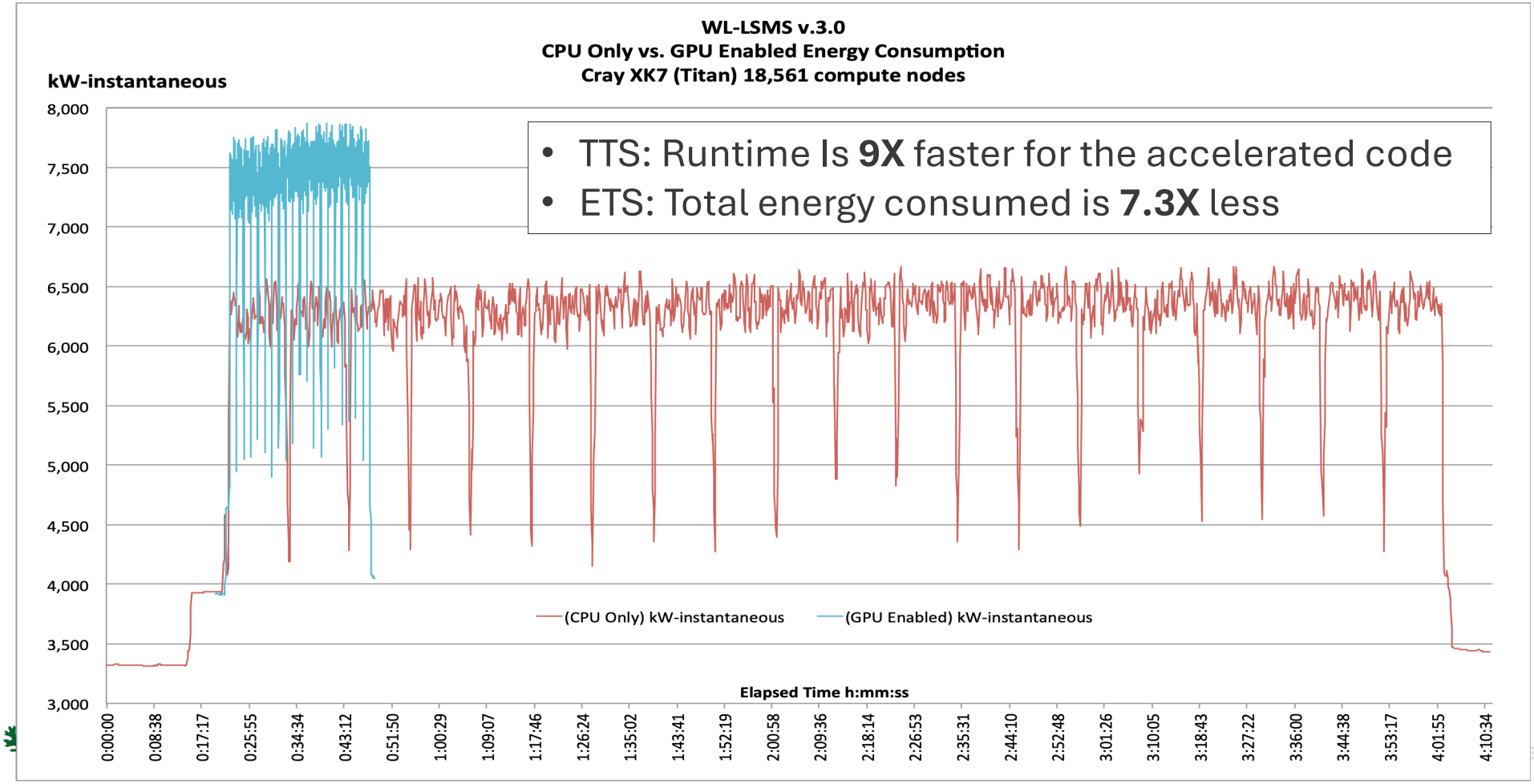
2028



Application Power Efficiency on the Cray XK7

Comparing CPU-Only and GPU-Enabled WL-LSMS

Slide **from 2012** introduces the impact of GPUs on Time- and Energy- to Solution



Balancing Hard Constraints with Co-design

Summit (2018 – 2024)

Contract Basis

- Competitive acquisition awarded to IBM for 4,608 compute nodes, each with 2 POWER9 CPU : 6 NVIDIA GV100.

Constraints

- Multi-lab CORAL acquisition process – shared requirements set a low bar
- Aggregate power budget for compute + file system: **20MW**

Wins

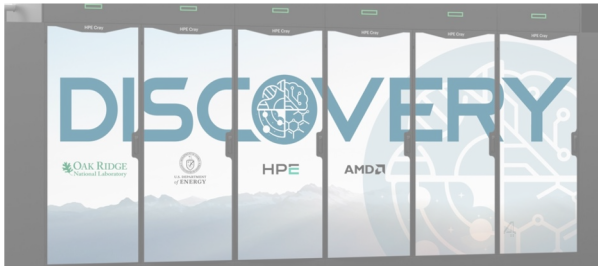
- Facility Central Energy Plant (CEP) design used **20C** supply directly to the rack
- **Non-recurring Engineering** (NRE) for the compute node, using 20C facility water, produced ~100% waste-heat capture.
  - Direct-liquid-cooling (DLC) design for CPUs and GPUs, with RDHX for memory, NICS, power supplies.
  - Annualized **ITUE of 1.1** contributes to improved OPEX
- 10x sustained performance from NVIDIA Kepler to NVIDIA Volta
  - Dramatic advances in exascale AI and many scientific domains
  - Introduction of mixed precision computing and tensor cores

2012

2018

2023

2028



Balancing Hard Constraints with Co-design

Frontier (2022 –)

Contract Basis

- Competitive acquisition awarded to HPE for 9,856 compute nodes, each with 1 AMD Gen 3 EPYC (Trento) CPU: 3 AMD MI250X GPU

Constraints/Requirements

- Multi-lab CORAL-2 acquisition process – shared requirements set a low bar
- Peak Performance Target: 1.3EF FP64
- 5x scientific throughput improvement
- Aggregate power budget for compute + file system: **40MW**

Wins

- Facility CEP design delivered 32C supply directly to the rack
  - 100% chiller-free
- Non-recurring Engineering (NRE) for the compute node produced ~100% waste-heat capture.
  - HPE Cray EX design is fan-less. Waste heat directly to water
  - ITUE as low as 1.03 (depending on load)
- 5x sustained performance from Summit to Frontier supports advanced fusion, turbulence modeling, mixed-precision AI

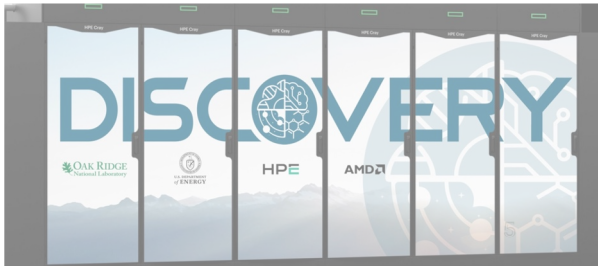
OAK RIDGE National Laboratory

2012

2018

2022

2028



Balancing Hard Constraints with Co-design

Discovery (2028 –)

Contract Basis

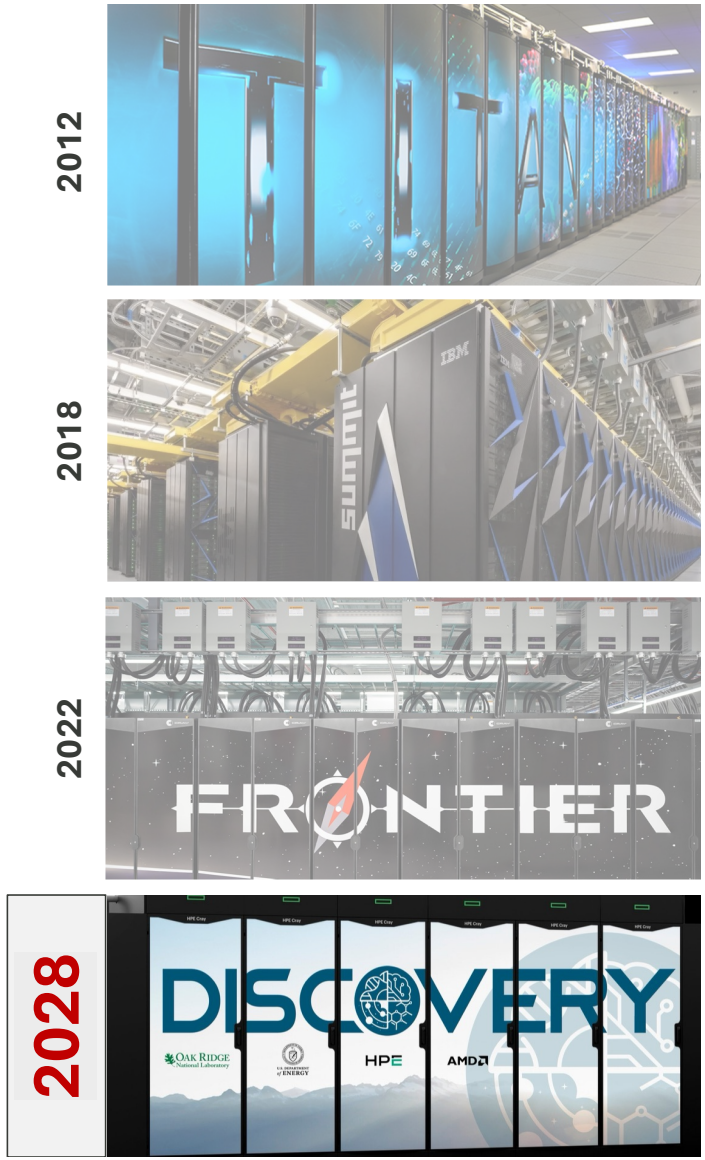
- Competitive acquisition awarded to HPE for a GX5000-based system, with each node configured as 1 AMD EPYC CPU: 4 AMD MI430X GPU.
- Pending GO/NO-GO will define the final system size based on funding, component pricing

Constraints/Requirements

- "Significant scientific throughput improvement versus Frontier"
- Aggregate power budget for compute + file system: 40MW
- Must use 32C supply CEP (from Frontier)
- **Stress on facilities to accommodate Frontier+Discovery for ~1-year.**

Wins

- Facility-
  - W32 CEP upgraded in-place to support both Frontier and Discovery
- System-
  - Architectural focus on improved memory and network bandwidth
  - Significant leverage of the OCP ORV3 HPR ecosystem
  - User-level access to runtime power management features



OAK RIDGE National Laboratory

Volume of Work as a Key Source Selection Criteria

Basis –

ORNL frequently issues request for RFPs that describe a resulting Firm Fixed Price Contract at a specific value.

- Drives system size in accordance with the target requirements for the system.
- Eliminates Cost as an evaluation factor
- Seeks solutions that maximize the Volume of Work (performance-weighted)
- Does not consider energy consumption as a Bidder/Seller responsibility

Volume of Work –

- A weighted assessment of an RFP's benchmarks that describe ***an increase in the total delivered work versus the baseline system on the target architecture.***
- Focuses on Time to Solution for principal applications

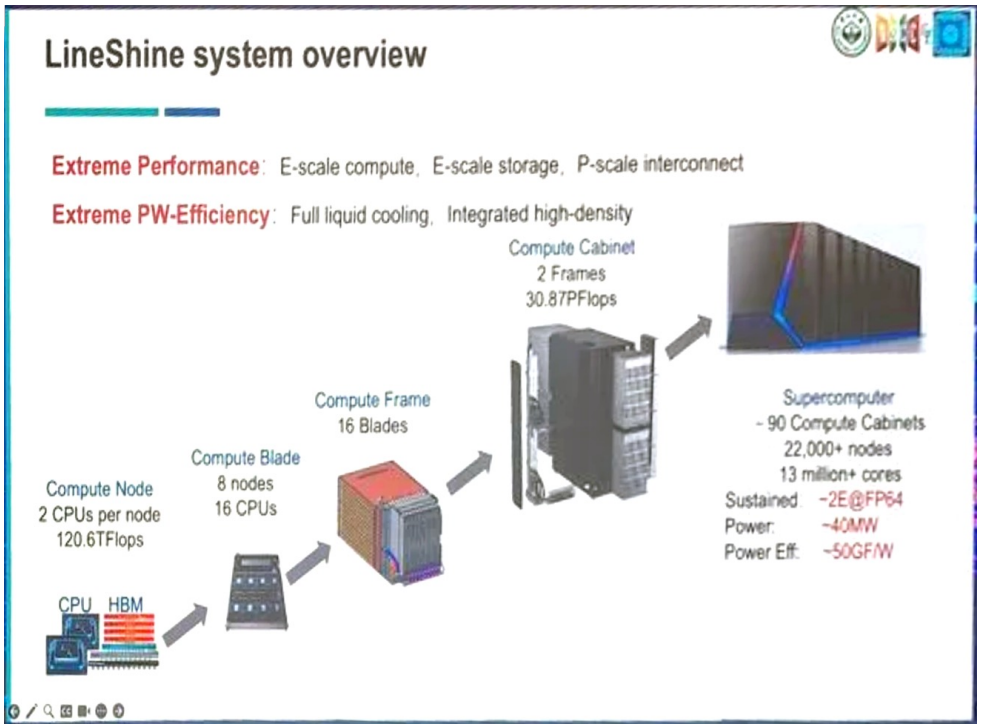
Consider: Supplement Volume of Work with a similar calculation of Energy to Solution to support the calculation of TCO

Proposed					C7 System Size				
						nodes			
Benchmark	T^i	C7 Max Ti	Total Runtime	Cores Used	Nodes Alloc.	C^i	Weight W^i	$V^i = \frac{T_b^i C^i}{T^i C_b^i}$	$V = \frac{1}{\sum \frac{W^i}{V^i}}$
	secs	secs	secs			copies			
ESM4 (production)		10790					0.35		
CM5 (production)		8993					0.55		
MPAS (QU20km)		2633					0.05		
MPAS (VR20km)		3009					0.05		
Total							**1.00**	**Calculated V**	

A recent ORNL RFP example, where just three benchmarks, one with two input data sets, defined the increase in performance of a proposed system.

Worldwide HPC+AI Procurement Demonstrates Power, Space, and Cooling Demand

China's LineShine
40k CPUs, 0 GPUs, 2EF FP64, 40MW, 50GF/W



Source: HACI 2026 Conference, May 22-25, 2026, Shenzen China

- **Facility investment in the Shenzhen Guangming District National Supercomputing Project: RMB 6.45 billion (~ € 0.8B)**
  - **Principle Energy Supply- Guangming District including the Shenzhen Energy Guangming Natural Gas Power Station**
- **LineShine CAPEX: Undisclosed, but €100M's**
- **LineShine OPEX: Estimated at ~€0.08 per kWh**

"The LineShine supercomputer, developed by the National Supercomputing Center in Shenzhen (NSCC-SZ), is an exascale system consisting of 20,480 computing nodes, denoted the full machine. Each node is equipped with two ARMv9- based LX2 processors. Each LX2 integrates two compute dies (304 cores total) and eight on-package HBM stacks (32 GB, 4 TB/s aggregate bandwidth). Each compute die contains 152 cores and 128 GB of off-package DDR memory organized into four NUMA domains."

Source: Breaking the Training Barrier of Billion-Parameter Universal Machine Learning Interatomic Potentials. https://arxiv.org/abs/2604.15821