The Changing Landscape for Scientific Computing at NSF and DOE

Mike Norman
Director, SDSC
Distinguished Professor of Physics

Tale of 2 Agencies: Different priorities and strategies

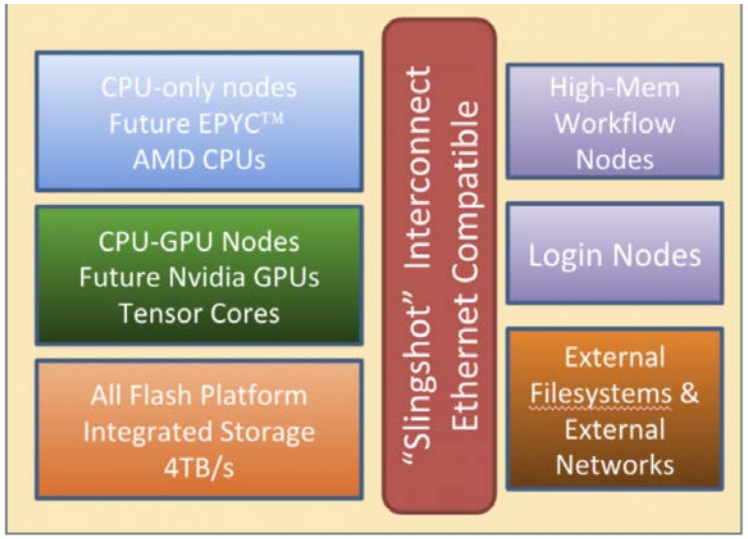
- **DOE**
  - Goal: International leadership via deployed exascale systems
  - Vision: Converged analytics, AI, and simulation at extreme scale
  - Strategy: HPC systems deployed through lab procurements
  - Strategy: Massive investment in exascale software teams (ECP)

- **NSF**
  - Goal: Broaden access to advanced computing for an increasingly diverse set of academic researchers
  - Vision: Innovative HPC systems integrated into a national cyberinfrastructure including cloud and campus resources
  - Strategy: Cost-effective leadership and capacity systems deployed by winning proposer
  - Strategy: Proportionate investment in diverse PI-led software projects (CSSI)

DOE/NNSA exascale roadmap

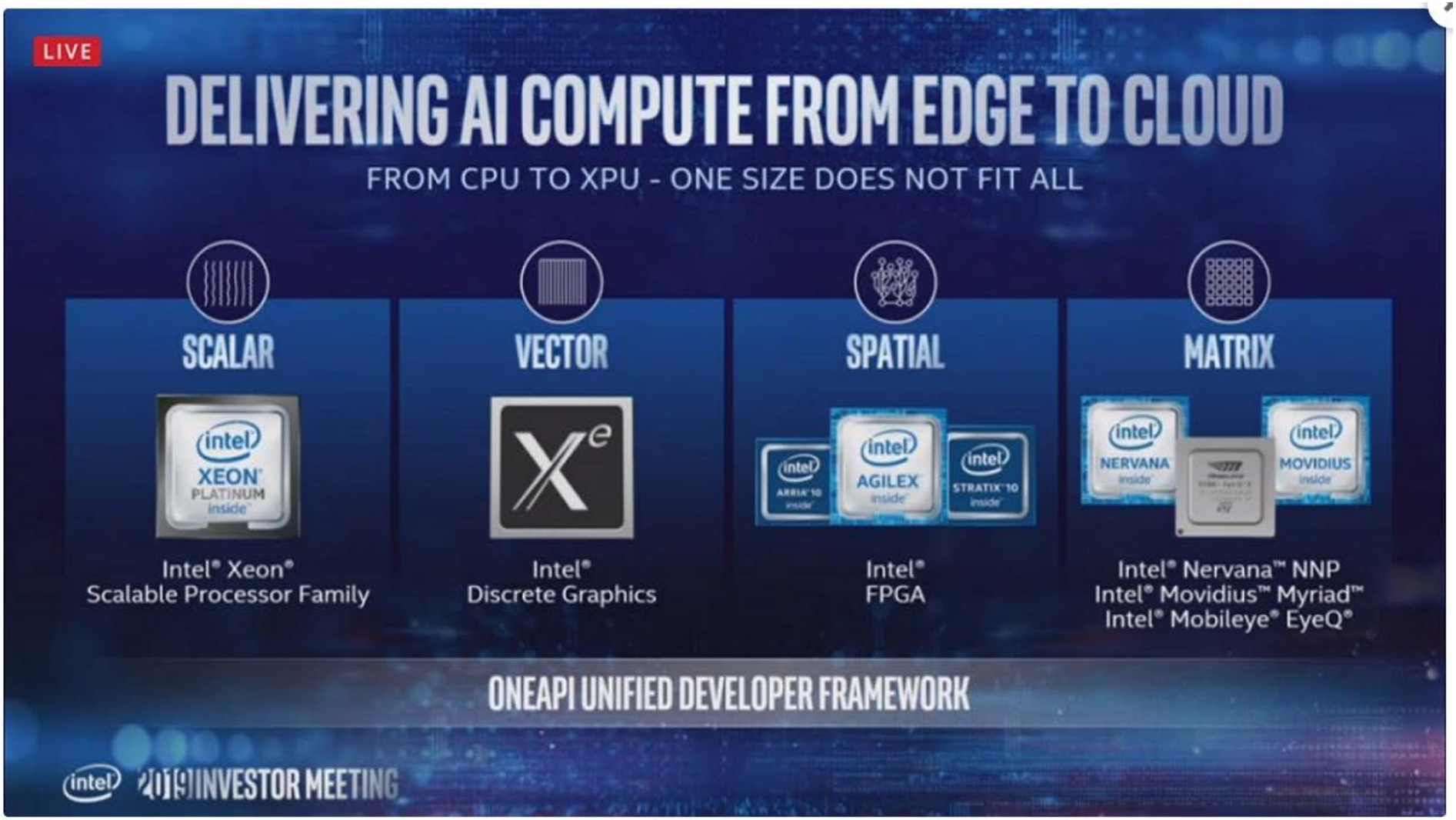
NERSC Perlmutter

- Cray Shasta platform
- Slingshot interconnect
- AMD CPUs
- Nvidia GPUs
- Diverse MPI+X programming models supported
- All-flash PFS



ANL Aurora

- Cray Shasta platform
- Slingshot interconnect
- Intel CPUs & GPUs
- Other Intel XPUs?
- Diverse MPI+X programming models supported
- Unified node-level programming via One API


LIVE
DELIVERING AI COMPUTE FROM EDGE TO CLOUD
FROM CPU TO XPU - ONE SIZE DOES NOT FIT ALL
SCALAR
Intel® Xeon® Scalable Processor Family
VECTOR
Intel® Discrete Graphics
SPATIAL
Intel® FPGA
MATRIX
Intel® Nervana™ NNP
Intel® Movidius™ Myriad™
Intel® Mobileye® EyeQ®
ONEAPI UNIFIED DEVELOPER FRAMEWORK
2019 INVESTOR MEETING

ORNL Frontier

- Cray Shasta platform
- Slingshot interconnect
- AMD CPUs and GPUs
- Diverse MPI+X programming models supported
- Unified node-level programming via ROCm
- $600M / 40 MW

ROCm, a New Era in Open GPU Computing

Platform for GPU-Enabled HPC and Ultrascale Computing

Overview | Getting Started | Documentation | ROCm Ecosystem | Deep Learning | Tutorials | Community | Contribute

ROCm Tutorials

ROCm comes with a set of tutorials designed to help you understand how you can use and extend the platform.

- hipCAFFEHow use Caffe on ROCm
- Vector-Add Vector Add example ussing the HIP Programing Language
- Mini N-body This sample demonstrates the use of the HIP API for a mini n-body problem.
- Assembly Sample The Art of AMDGCN Assembly:How to Bend the Machine to Your Will. This tutorial demonstrates GCN assembly with ROCm application development.
- Optimizing Dispatches ROCm With Rapid Harmony: Optimizing HSA Dispatch: This tutorial shows how to optimize HSA dispatch performance for ROCm application development.
- CLOC Offline Compiler ROCm With Harmony: Combining OpenCL Kernels, HCC and HSA in a Single Program. This tutorial demonstrates how to compile OpenCL kernels using the CL offline compiler (CLOC) and integrate them with HCC C++ compiled ROCm applications.

Upshot: DOE is “all in” on dense GPU clusters

- **3 systems, 3 different kinds of GPUs**
- **Application porting is a major issue**
- **Open platforms Perlmutter, Aurora and Frontier require different node-level programming approaches (CUDA, One API, ROCm), creating consternation amongst application teams**
  - Only CUDA is mature
- **$1.8B being spent on ECP, which is an ambitious software readiness program**
  - Big effort in “performance portability” SW: Kokkos, RAJA, hipify
  - Big effort in math libraries too
  - All exascale software will be open source and available via SDKs

Tale of 2 Agencies: Different priorities and strategies

- **DOE**
  - Goal: International leadership via deployed exascale systems
  - Vision: Converged analytics, AI, and simulation at extreme scale
  - Strategy: HPC systems deployed through lab procurements
  - Strategy: Massive investment in exascale software teams (ECP)

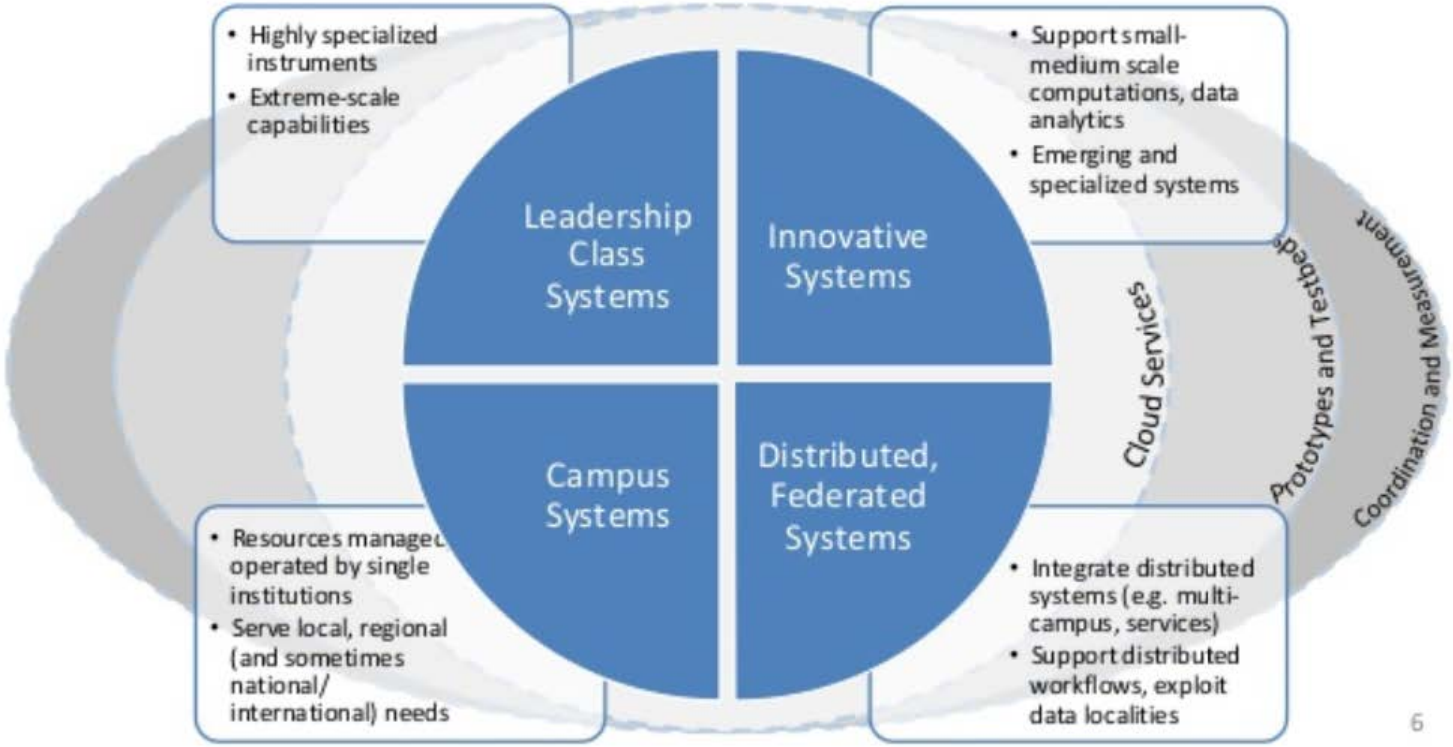
- **NSF**
  - Goal: Broaden access to advanced computing for an increasingly diverse set of academic researchers
  - Vision: Innovative HPC systems integrated into a national cyberinfrastructure including cloud and campus resources
  - Strategy: Cost-effective leadership and capacity systems deployed by winning proposer
  - Strategy: Proportionate investment in diverse PI-led software projects (CSSI)

NSF OAC's philosophy

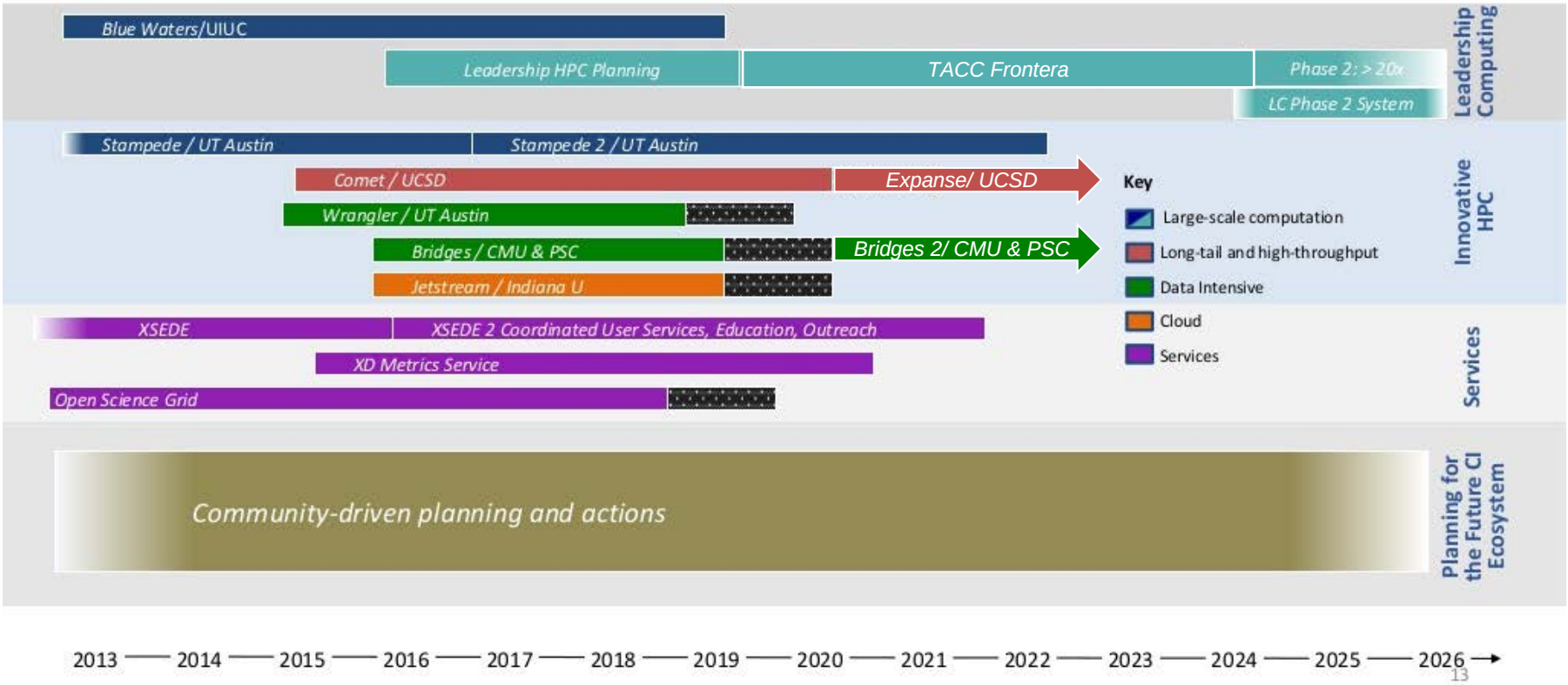
Conclusion

- Science and society are being transformed by compute and data – an agile cyberinfrastructure ecosystem is essential
- Rapidly changing application requirements; resource and technology landscapes
  - Our cyberinfrastructure ecosystem must efficiently evolve
- Forward-looking approach to cyberinfrastructure ecosystem, aimed at transforming science

Leadership Computing, part of a Balanced Portfolio

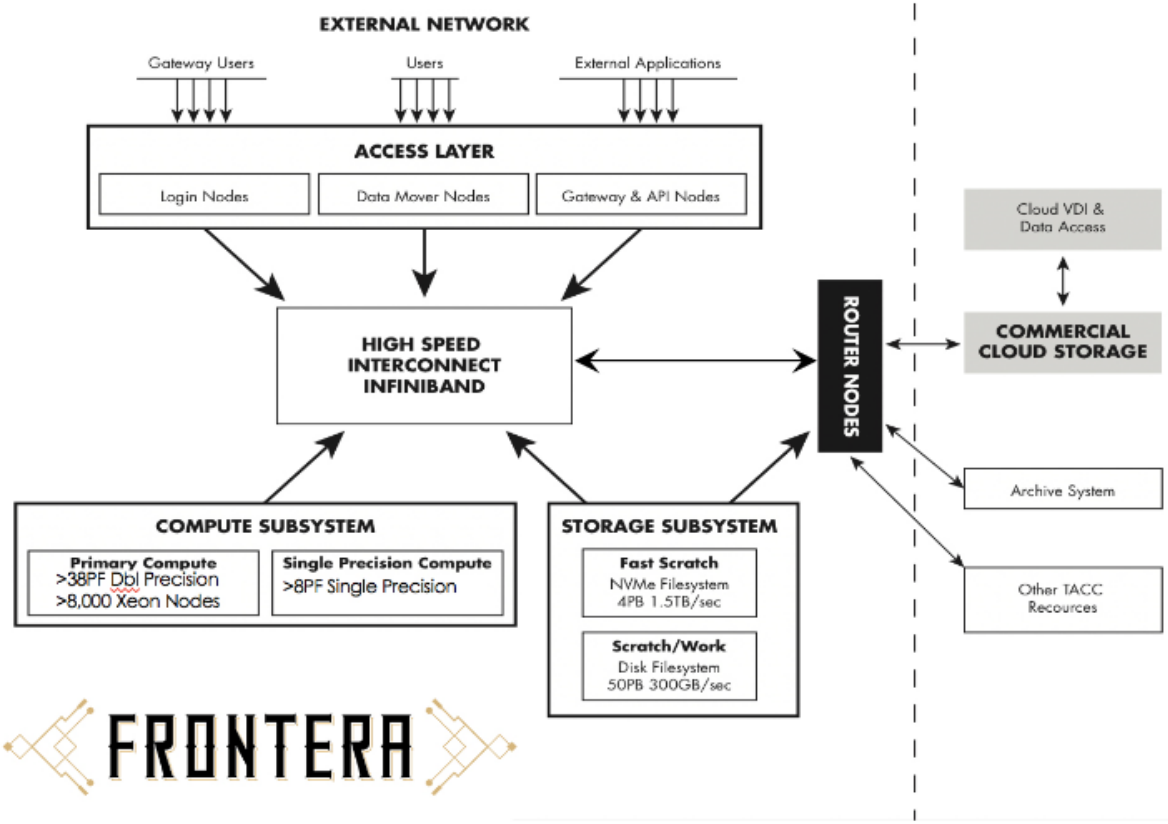


Planning for the Future: OAC-Funded Computing Ecosystem



System Hardware and Software Overview

- "Vanilla" 8000 node Dell/EMC Intel CPU cluster
- 56 core dual socket Cascade Lake nodes
- HDR Infiniband fabric
- CPU applications "compile and go"
- Standalone mini-Summit cluster (IBM P9/Nvidia V100)
- Shared PFS



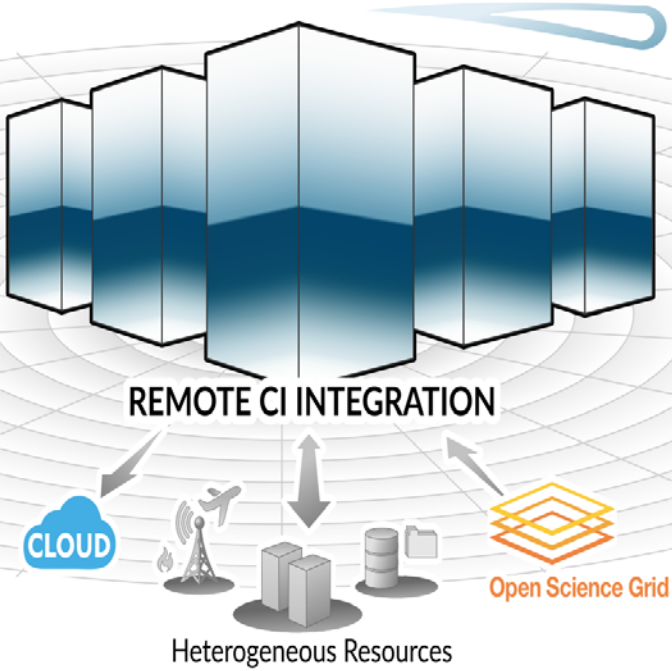

EXPANSE
COMPUTING WITHOUT BOUNDARIES

EXPANSE

COMPUTING WITHOUT BOUNDARIES
5 PETAFLOP/S HPC and DATA RESOURCE

HPC RESOURCE
13 Scalable Compute Units
728 Standard Compute Nodes
52 GPU Nodes: 208 GPUs
4 Large Memory Nodes

DATA CENTRIC ARCHITECTURE
12PB Perf. Storage: 140GB/s, 200k IOPS
Fast I/O Node-Local NVMe Storage
7PB Ceph Object Storage
High-Performance R&E Networking



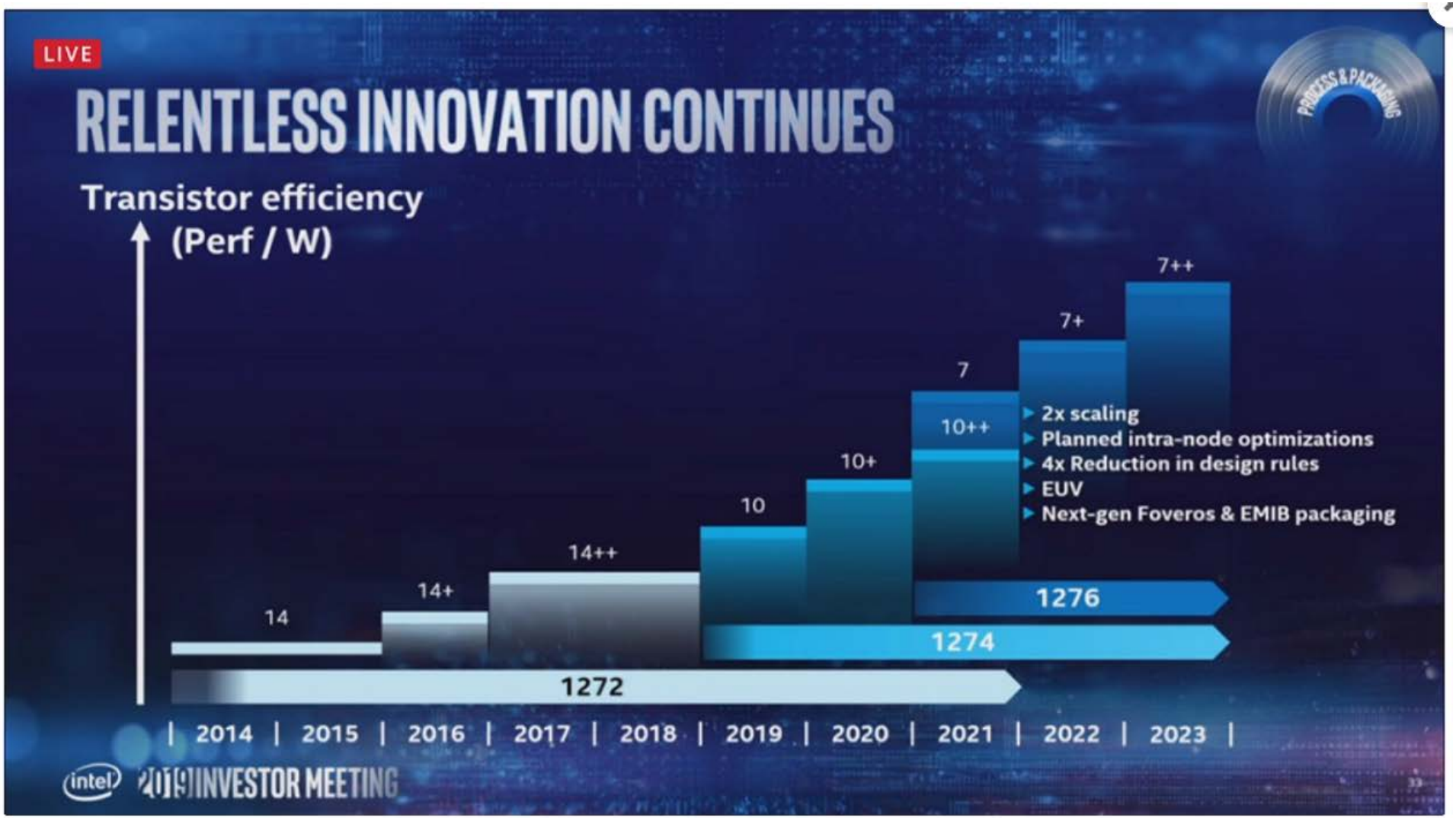
LONG-TAIL SCIENCE
Multi-Messenger Astronomy
Genomics
Earth Science
Social Science

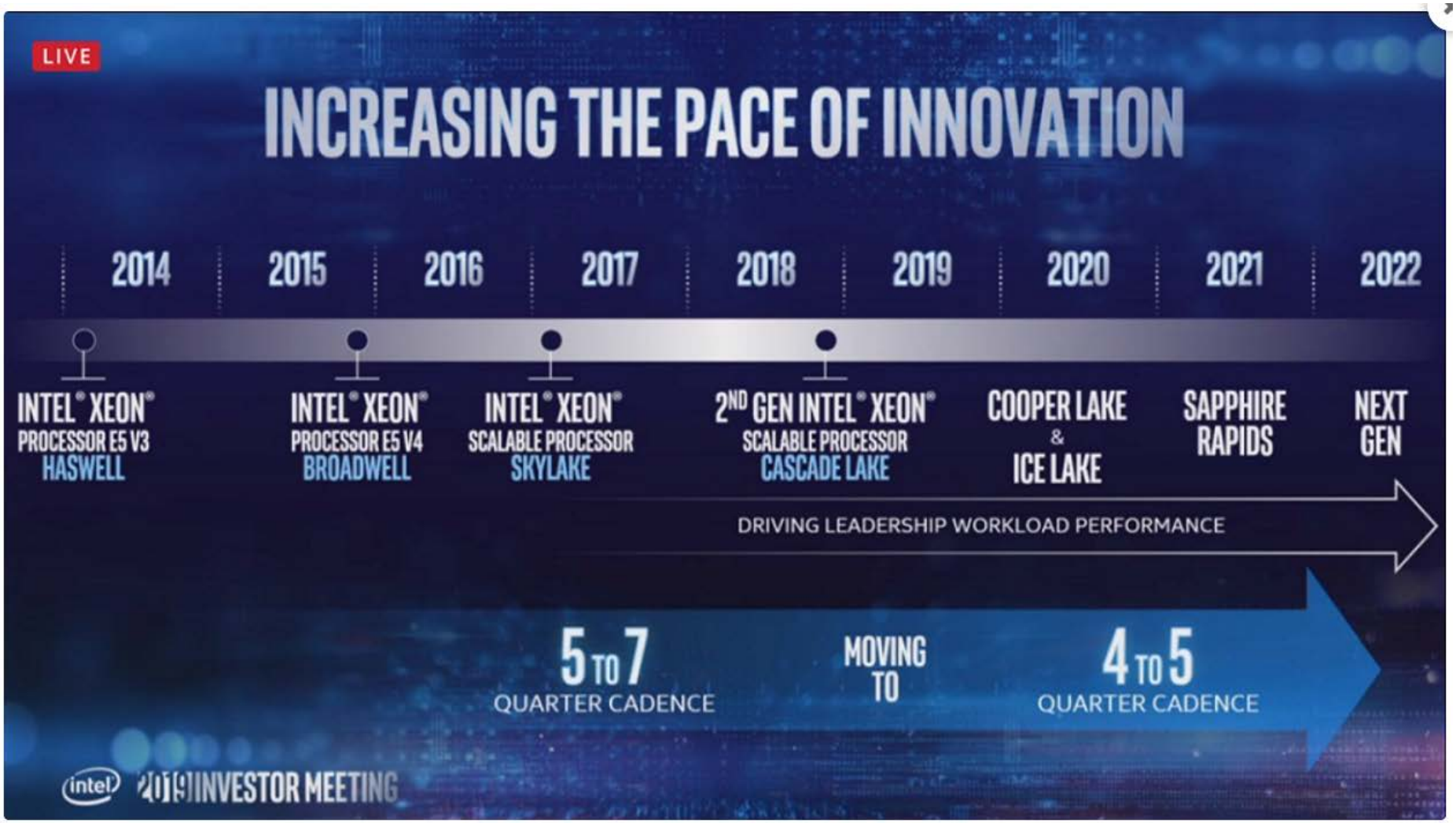
INNOVATIVE OPERATIONS
Composable Systems
High-Throughput Computing
Science Gateways
Interactive Computing
Containerized Computing
Cloud Bursting

Closing thoughts/suggestions

- **SCEC can build on a diversity of compute and data resources available from NSF and DOE**
  - High end simulations can be accommodated via PRAC and INCITE programs
  - Some architectures will require substantial code development to exploit (Aurora, Frontier), others no so much (Frontera)
  - Need to chose your HPC path forward taking software development into account
  - Other NSF resources (Expanse, Bridges 2) may be useful for ensemble runs, data analysis, and data sharing
- **SCEC should strongly advocate for its needs in multiple venues**
  - LCCF Phase 2 requirements gathering (next 12 months)
  - SCEC should advocate NSF OAC for something like PetaApps program for coming leadership computing systems

RESERVE


LIVE
RELENTLESS INNOVATION CONTINUES
PROCESS & PACKAGING
Transistor efficiency
(Perf / W)
14
14+
14++
10
10+
10++
7
7+
7++
2x scaling
Planned intra-node optimizations
4x Reduction in design rules
EUV
Next-gen Foveros & EMIB packaging
1276
1274
1272
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
intel
2019 INVESTOR MEETING


LIVE
INCREASING THE PACE OF INNOVATION
2014
2015
2016
2017
2018
2019
2020
2021
2022
INTEL® XEON® PROCESSOR E5 V3 HASWELL
INTEL® XEON® PROCESSOR E5 V4 BROADWELL
INTEL® XEON® SCALABLE PROCESSOR SKYLAKE
2ND GEN INTEL® XEON® SCALABLE PROCESSOR CASCADE LAKE
COOPER LAKE & ICE LAKE
SAPPHIRE RAPIDS
NEXT GEN
DRIVING LEADERSHIP WORKLOAD PERFORMANCE
5 TO 7 QUARTER CADENCE
MOVING TO
4 TO 5 QUARTER CADENCE
intel 2019 INVESTOR MEETING