Big memory and unified memory supercomputers in CCS

Osamu Tatebe

University of Tsukuba

University of Tsukuba
Center for Computational Sciences

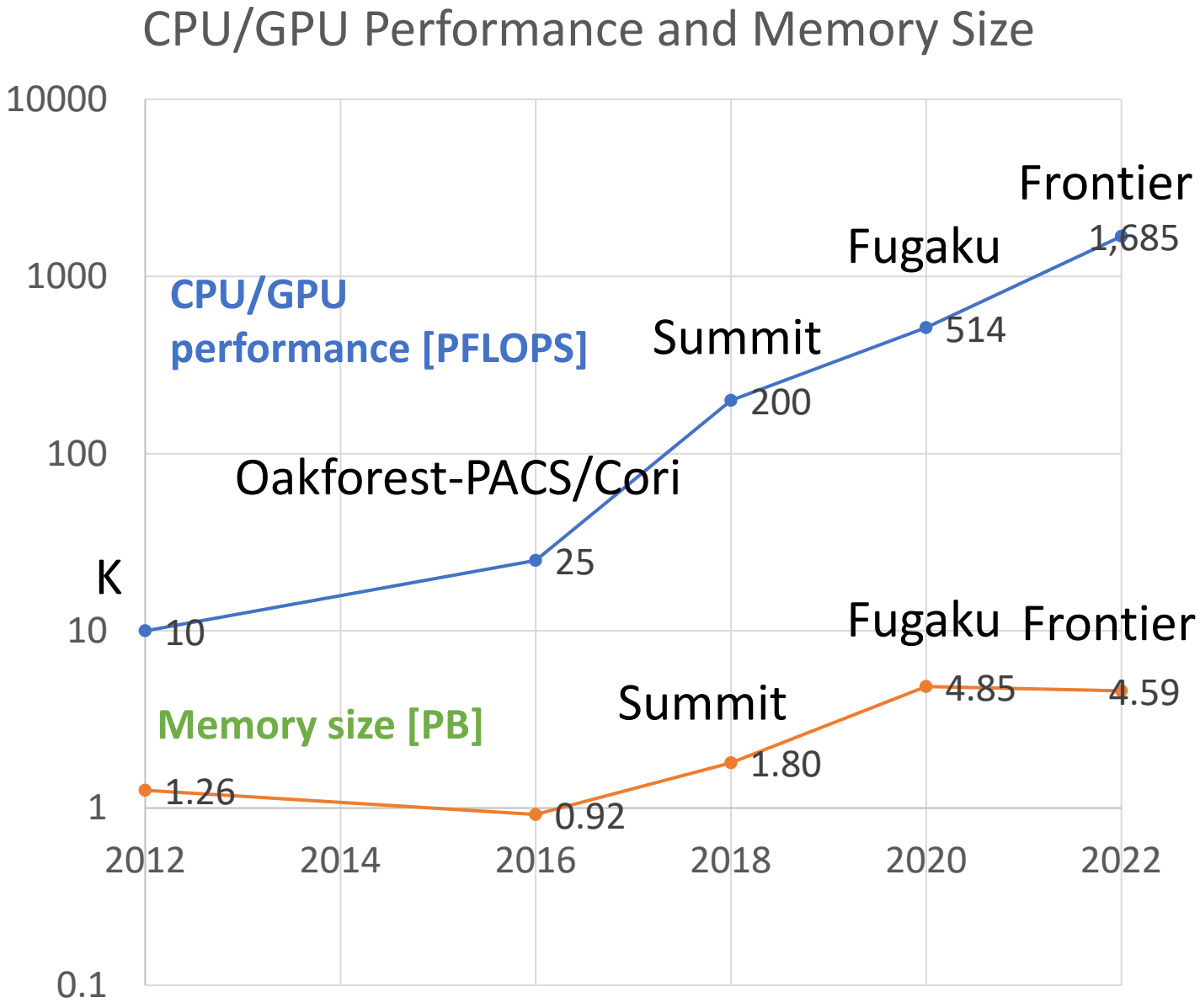
Supercomputers in CCS and JCAHPC

- CCS
  - Cygnus – 2.3PFlops GPU/FPGA multi-hybrid accelerated supercomputer (2019.4 – 2025.3)
  - Pegasus – 8.1PFlops big memory supercomputer (2023.4 – 2028.1)
  - Post Cygnus (tentative) under procurement – unified memory supercomputer (2025.10 –)
- JCAHPC
  - Oakforest-PACS – 25PFlops manycore supercomputer (2017.4 – 2022.3)
  - Miyabi – 80PFlops GH and Xeon CPU Max supercomputer (2025.1 –)

Pegasus background

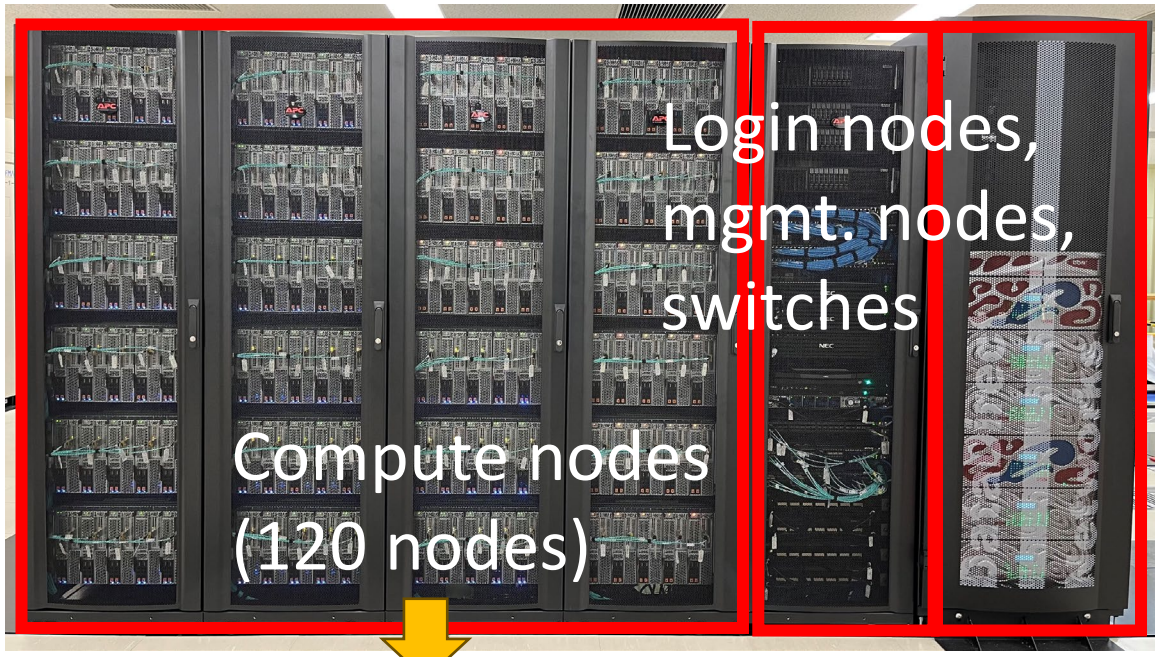
- CPU performance 170x, but memory size 3.6x in 10 years
- It matters for Data-driven and AI-driven Science
  - Memory size and Storage performance are important

- Introduce Persistent Memory
  - Low power and cost effective
  - Big memory space and high-performance storage



Pegasus Highlights

- Build with 4th Gen Intel Xeon (SPR), NVIDIA H100 Tensor Core PCIe GPU, and Intel Optane persistent memory 300, which will strongly drive Big Data and AI



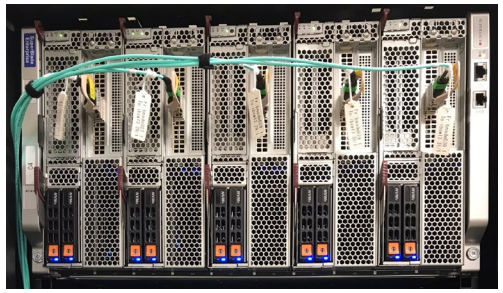
150 nodes

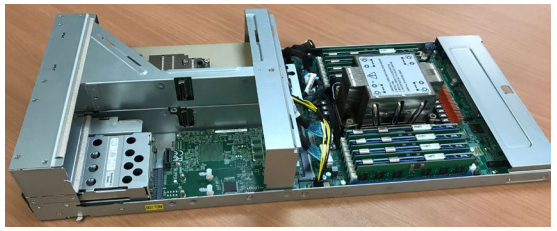
Parallel File System

Pegasus Specification

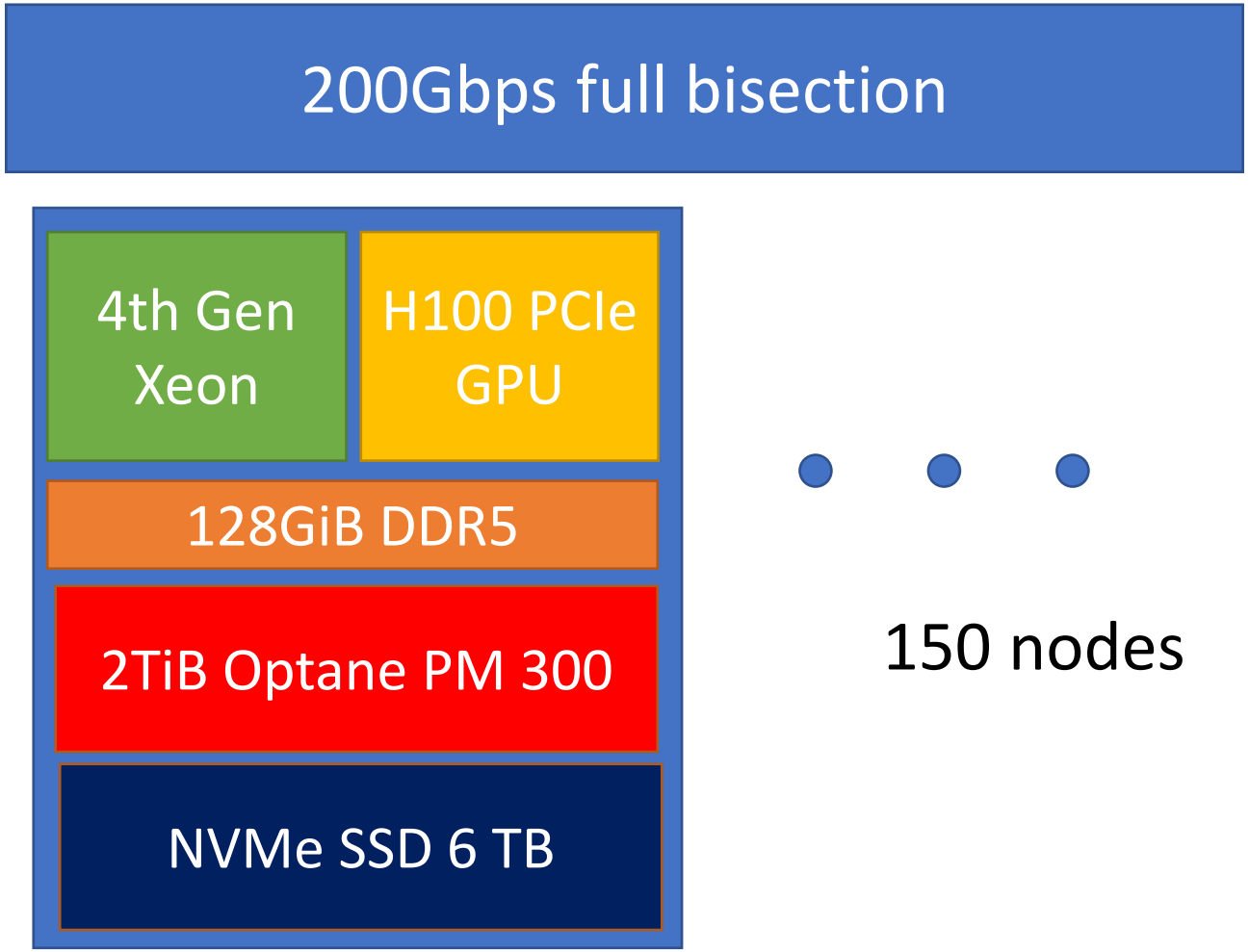
- Installed in Q4 2022
- Total Performance
  - 150 nodes, 8.1 PFlops, 300 TiB Pmem
- Node specification
  - 3.2 TFlops Intel Platinum 8468 (Sapphire Rapids)
  - 51 TFlops NVIDIA H100 PCIe GPU
  - 128 GiB DDR5 DRAM (282 GB/s)
  - 2 TiB Optane PM 300 series (Crow Pass)
  - 6 TB NVMe SSD (7 GB/s)
- Interconnection Network
  - NVIDIA Quantum-2 InfiniBand platform (200 Gbps) full bisection (InfinBand NDR200)
- Parallel File System
  - 7.1 PByte DDN EXAScaler (40 GB/s)

NEC LX B1000E Blade Enclosure

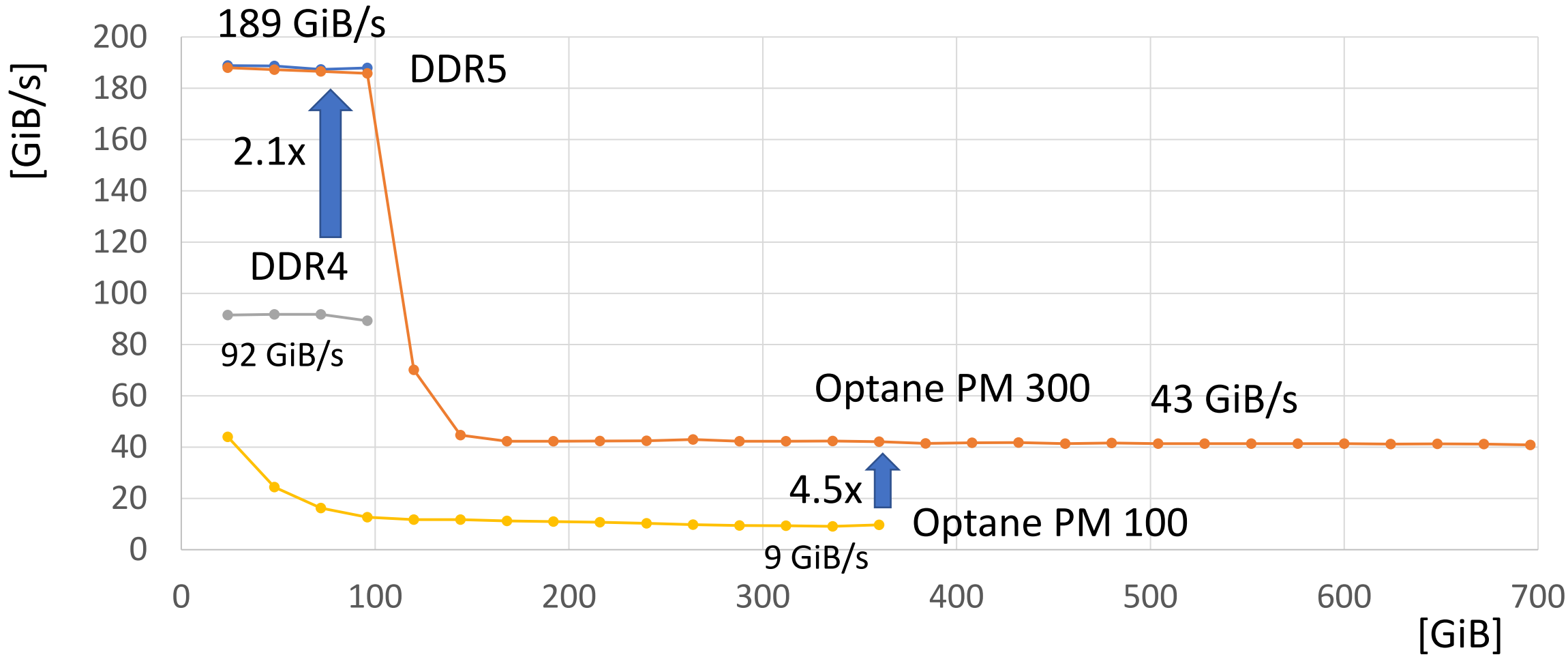




NEC LX 102Bk-6



Stream Benchmark (Triad; a[] = b[] + s * c[])

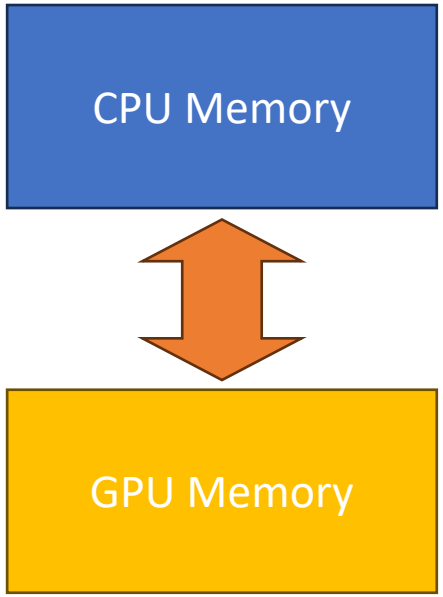


Special thanks to Akira Nukada

Post Cygnus

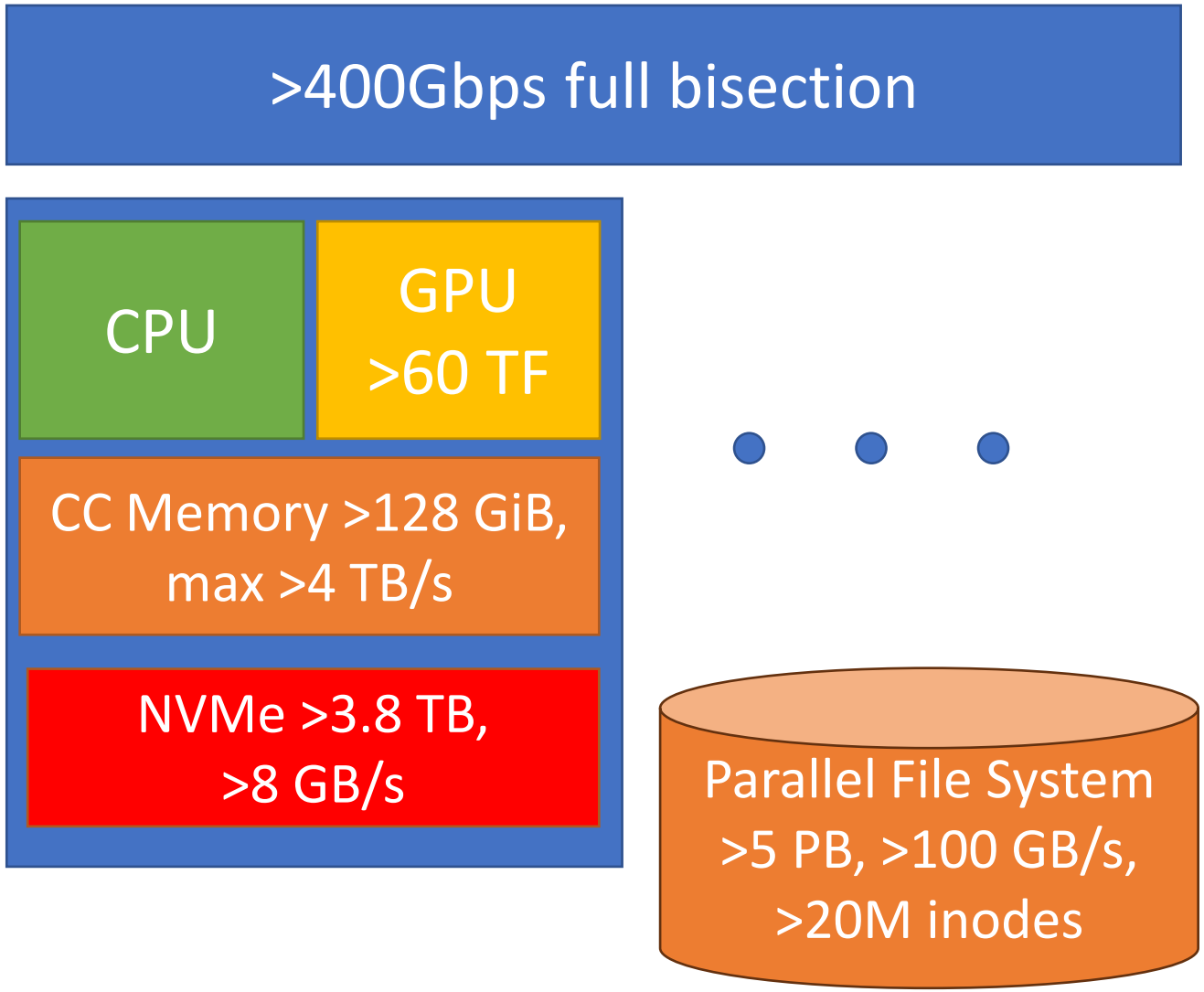
- Background
  - Large memory issue
    - Optane PM discontinued
    - Node-level memory disaggregation (CXL 3.0) is not available yet
  - Discrete memory issue between CPU and GPU
    - Copy overhead between CPU and GPU memories
    - New tightly coupled CPU/GPU system architecture
      - Miyabi, JUPITER, El Capitan

Cache-coherent unified memory for CPU and GPU



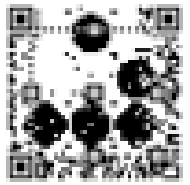
Requested Specification (1)

- Total DP Performance
  - 6 to 16 PFlops
- Compute node
  - GPU (>60 TFlops)
  - **Cache-coherent memory** (>128 GiB, max >4 TB/s) per GPU
  - NVMe SSD (>3.8 TB, > 8 GB/s) per GPU
- Interconnection Network
  - >400 Gbps per GPU full bisection
- Parallel File System
  - >5 PB, >100 GB/s, >20M inodes



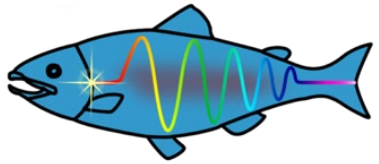
Requested Specification (2)

- Benchmarks
  - BabelStream
    - >200 GB/s (CPU)
    - >2,000 GB/s (GPU)
  - MDtest (empty file creation performance for a shared directory)
    - >100K IOP/s
  - IOR (striped single shared file access)
    - >35 GiB/s (Write, Read)
  - HPL
    - >3 PFlops
  - Applications
    - CCS QCD (F90 + Cuda C)
    - Gromacs
    - Collision (Cuda C)
    - SALMON – scalable ab-initio light-matter simulator for optics and nanoscience (OpenACC Fortran)



Innovative Computing Laboratory
UNIVERSITY OF TENNESSEE
COMPUTER SCIENCE DEPARTMENT

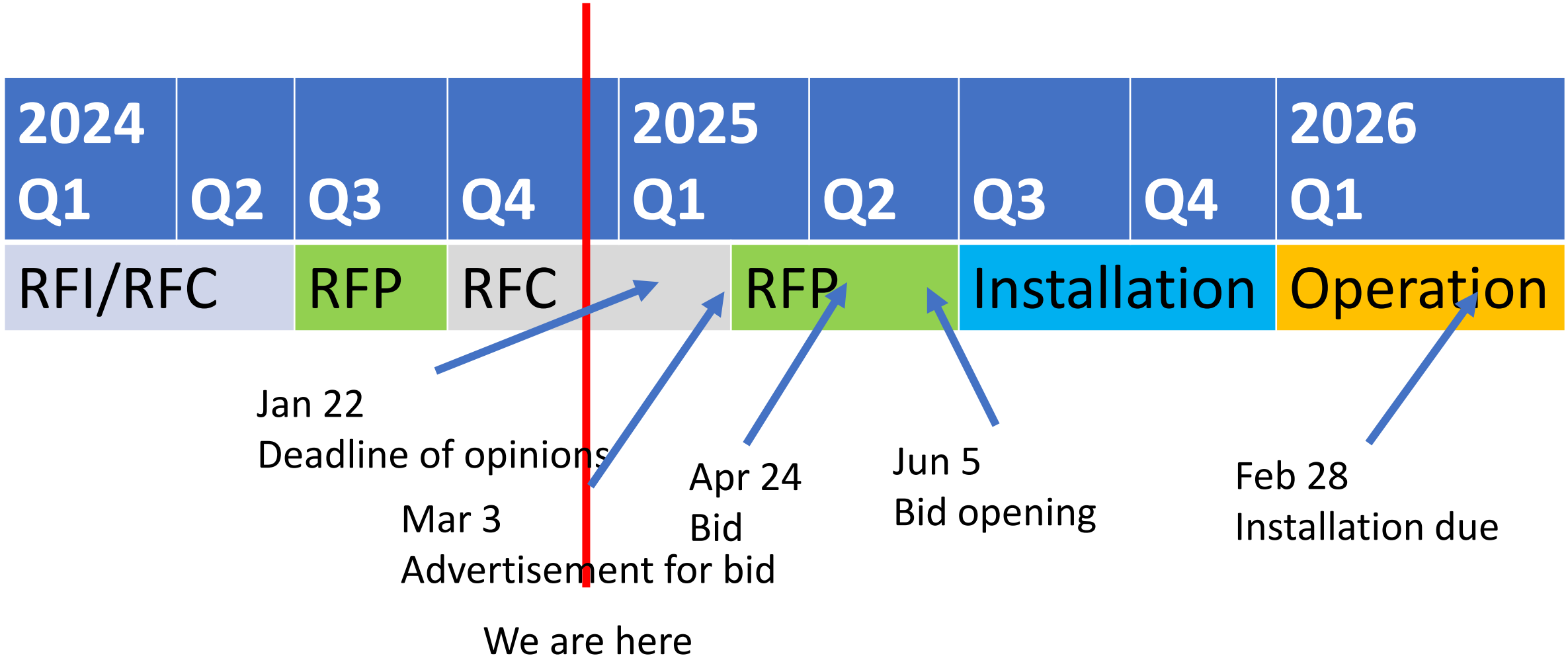
筑波大学
計算科学研究センター
Center for Computational Sciences



SALMON

GROMACS
FAST. FLEXIBLE. FREE.

Procurement Schedule (tentative)



学際ハブ拠点スパコンお試し利用

- 事業の目的
  - 新たな需要の開拓、企業での計算科学的手法の導入の裾野拡大
- 対象者
  - CCSと共同研究を継続的に実施する企業
- 制度
  - センター教員との共同研究を前提としたスパコンお試し企業利用
- 申込
  - https://www.ccs.tsukuba.ac.jp/kyodoriyou/hub-trial/

まとめ

- Pegasusは2024年12月より8.1 PFlops（150ノード）に増強
  - SPR, H100, Optane PM 300, NDR200
- Miyabiは2025年1月より運用開始
  - 80PFlops GH and Xeon Max CPU supercomputer
- Post Cygnusは2026年2月に導入予定
  - ユニファイドメモリー型スーパーコンピュータ
- 学際ハブ拠点スパコンお試し利用により企業利用を開始