

CS 470 Spring 2024

Mike Lam, Professor



Frontier supercomputer, Oak Ridge

Novel Systems

System architectures

- Shared memory
 - Primary goal: **make faster computers**
 - Programming paradigm: threads
 - Technologies: [Pthreads](#), [OpenMP](#), [CUDA](#)
- Distributed memory
 - Primary goal: **add more computers**
 - Programming paradigm: message passing
 - Technologies: [MPI](#), [SLURM](#)

Where do we go from here?

Hybrid HPC architectures

- Massive parallelism on the node
 - Hardware: CPU w/ accelerators
 - GPUs or manycore processors
 - Technologies: [OpenMP](#), [CUDA](#), [OpenACC](#), [OpenCL](#)
- Distributed between massive number of nodes
 - Hardware: fast interconnect and distributed FS
 - Technologies: [MPI](#), [Infiniband](#), [Lustre](#), [HDFS](#)

Summit, Oak Ridge National Lab, USA



Fugaku (富岳), Riken Center, Japan



Sunway TaihuLight (神威·太湖之光),
National Supercomputing Center in Wuxi, China



Top10 systems (Spring 2016)

| RANK | SITE | SYSTEM | CORES | RMAX (TFLOP/S) | RPEAK (TFLOP/S) | POWER (KW) |
|------|-----------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|-----------|-------------------|--------------------|---------------|
| 1 | National Super Computer Center in Guangzhou China | Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT | 3,120,000 | 33,862.7 | 54,902.4 | 17,808 |
| 2 | DOE/SC/Oak Ridge National Laboratory United States | Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc. | 560,640 | 17,590.0 | 27,112.5 | 8,209 |
| 3 | DOE/NNSA/LLNL United States | Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM | 1,572,864 | 17,173.2 | 20,132.7 | 7,890 |
| 4 | RIKEN Advanced Institute for Computational Science (AICS) Japan | K computer , SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu | 705,024 | 10,510.0 | 11,280.4 | 12,660 |
| 5 | DOE/SC/Argonne National Laboratory United States | Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM | 786,432 | 8,586.6 | 10,066.3 | 3,945 |
| 6 | DOE/NNSA/LANL/SNL United States | Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc. | 301,056 | 8,100.9 | 11,078.9 | |
| 7 | Swiss National Supercomputing Centre (CSCS) Switzerland | Piz Daint - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x Cray Inc. | 115,984 | 6,271.0 | 7,788.9 | 2,325 |
| 8 | HLRS - Höchstleistungsrechenzentrum Stuttgart Germany | Hazel Hen - Cray XC40, Xeon E5-2680v3 12C 2.5GHz, Aries interconnect Cray Inc. | 185,088 | 5,640.2 | 7,403.5 | |
| 9 | King Abdullah University of Science and Technology Saudi Arabia | Shaheen II - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc. | 196,608 | 5,537.0 | 7,235.2 | 2,834 |
| 10 | Texas Advanced Computing Center/Univ. of Texas United States | Stampede - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell | 462,462 | 5,168.1 | 8,520.1 | 4,510 |

4 / 10

Top10 systems (Spring 2017)

| Rank | Site | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|--------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|------------|----------------|-----------------|------------|
| 1 | National Supercomputing Center in Wuxi China | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCCPC | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 2 | National Super Computer Center in Guangzhou China | Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2 Intel Xeon Phi 31S1P NUDT | 3,120,000 | 33,862.7 | 54,902.4 | 17,808 |
| 3 | DOE/SC/Oak Ridge National Laboratory United States | Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc. | 560,640 | 17,590.0 | 27,112.5 | 8,209 |
| 4 | DOE/NNSA/LLNL United States | Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM | 1,572,864 | 17,173.2 | 20,132.7 | 7,890 |
| 5 | DOE/SC/LBNL/NERSC United States | Cori - Cray XC40, Intel Xeon Phi 7250 80C 1.4GHz, Aries interconnect Cray Inc. | 622,336 | 14,014.7 | 27,880.7 | 3,939 |
| 6 | Joint Center for Advanced High Performance Computing Japan | Oakforest PACS - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 80C 1.4GHz, Intel Omni-Path Fujitsu | 556,104 | 13,554.6 | 24,913.5 | 2,719 |
| 7 | RIKEN Advanced Institute for Computational Science (AICS) Japan | K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu | 705,024 | 10,510.0 | 11,280.4 | 12,660 |
| 8 | Swiss National Supercomputing Centre (CSCS) Switzerland | Piz Daint - Cray XC50, Xeon E5-2690-2 12C 2.6GHz, Aries interconnect, NVIDIA Tesla P100 Cray Inc. | 206,720 | 9,779.0 | 15,988.0 | 1,312 |
| 9 | DOE/SC/Argonne National Laboratory United States | Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM | 786,432 | 8,586.6 | 10,066.3 | 3,945 |
| 10 | DOE/NNSA/LANL/SNL United States | Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc. | 301,056 | 8,100.9 | 11,078.9 | 4,233 |

Top10 systems (Spring 2018)

| Rank | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|-------------------|--------------------|---------------|
| 1 | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 2 | Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-7, Intel Xeon Phi 31S1P, NJDT National Super Computer Center in Guangzhou China | 3,120,000 | 33,862.7 | 54,902.4 | 17,808 |
| 3 | Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect, NVIDIA Tesla P100, Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland | 361,760 | 19,590.0 | 25,326.3 | 2,272 |
| 4 | Gyokou - ZettaScaler-2.2 HPC system, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 700Mhz, ExaScaler Japan Agency for Marine-Earth Science and Technology Japan | 19,860,000 | 19,135.8 | 28,192.0 | 1,350 |
| 5 | Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x, Cray Inc. DOE/SC/Oak Ridge National Laboratory United States | 560,640 | 17,590.0 | 27,112.5 | 8,209 |
| 6 | Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom, IBM DOE/NNSA/LLNL United States | 1,572,864 | 17,173.2 | 20,132.7 | 7,890 |
| 7 | Trinity - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect, Cray Inc. DOE/NNSA/LANL/SNL United States | 979,968 | 14,137.3 | 43,902.6 | 3,844 |
| 8 | Cori - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect, Cray Inc. DOE/SC/LBNL/NERSC United States | 622,336 | 14,014.7 | 27,880.7 | 3,939 |
| 9 | Oakforest-PACS - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path, Fujitsu Joint Center for Advanced High Performance Computing Japan | 556,104 | 13,554.6 | 24,913.5 | 2,719 |
| 10 | K computer , SPARC64 VIIIfx 2.0GHz, Tofu interconnect, Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan | 705,024 | 10,510.0 | 11,280.4 | 12,660 |

7 / 10

Top10 systems (Spring 2019)

| Rank | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|-------------------|--------------------|---------------|
| 1 | Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States | 2,397,824 | 143,500.0 | 200,794.9 | 9,783 |
| 2 | Sierra - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 1,572,480 | 94,640.0 | 125,712.0 | 7,438 |
| 3 | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.75GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 4 | Tianhe-2A - TH-MB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NJDT National Super Computer Center in Guangzhou China | 4,981,760 | 61,444.5 | 100,678.7 | 18,482 |
| 5 | Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland | 387,872 | 21,230.0 | 27,154.3 | 2,384 |
| 6 | Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States | 979,072 | 20,158.7 | 41,461.2 | 7,578 |
| 7 | AI Bridging Cloud Infrastructure (ABCI) - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan | 391,680 | 19,880.0 | 32,576.6 | 1,649 |
| 8 | SuperMUC-NG - ThinkSystem SD530, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path , Lenovo Leibniz Rechenzentrum Germany | 305,856 | 19,476.6 | 26,873.9 | |
| 9 | Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States | 560,640 | 17,590.0 | 27,112.5 | 8,209 |
| 10 | Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States | 1,572,864 | 17,173.2 | 20,132.7 | 7,890 |

Top10 systems (Spring 2021)

| Rank | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|----------------|-----------------|------------|
| 1 | Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan | 7,630,848 | 442,010.0 | 537,212.0 | 29,899 |
| 2 | Summit - IBM Power system, AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States | 2,414,592 | 148,600.0 | 200,794.9 | 10,096 |
| 3 | Sierra - IBM Power System, AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 1,572,480 | 94,640.0 | 125,712.0 | 7,438 |
| 4 | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCP National Supercomputing Center in Wuxi China | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 5 | Stetson - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States | 555,520 | 63,460.0 | 79,215.0 | 2,646 |
| 6 | Tianhe-2A - TH-IVB-EPY Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China | 4,981,760 | 61,444.5 | 100,678.7 | 18,482 |
| 7 | JUWELS Booster Module - Bull Sequoia XH2000, AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec Parallel Station Cluster Suite, Atos Forschungszentrum Juelich [FZJ] Germany | 449,280 | 44,120.0 | 70,980.0 | 1,764 |
| 8 | HPCs - PowerEdge C6140, Xeon Gold 6252 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR Infiniband, Dell EMC Eco S.p.A. Italy | 669,760 | 35,450.0 | 51,720.8 | 2,252 |
| 9 | Frontera - Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox InfiniBand HDR, Dell EMC Texas Advanced Computing Center/Univ. of Texas United States | 448,448 | 23,516.4 | 38,745.9 | |
| 10 | Dammam 7 - Dell C6420, Xeon Gold 6248 20C 2.5GHz, NVIDIA Tesla V100 SXM2, InfiniBand HDR 100, HPE Saudi Aramco Saudi Arabia | 672,520 | 22,400.0 | 55,423.6 | |

Top10 systems (Spring 2022)

| Rank | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|----------------|-----------------|------------|
| 1 | Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan | 7,630,848 | 442,010.0 | 537,212.0 | 29,899 |
| 2 | Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States | 2,414,592 | 148,600.0 | 200,794.9 | 10,096 |
| 3 | Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 1,572,480 | 94,640.0 | 125,712.0 | 7,438 |
| 4 | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 5 | Perlmutter - HPE Cray EX2350n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Hingshot-10, HPE DOE/SC/LBNL/NERSC United States | 761,856 | 70,870.0 | 93,750.0 | 2,589 |
| 6 | Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States | 555,520 | 63,460.0 | 79,215.0 | 2,646 |
| 7 | Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China | 4,981,760 | 61,444.5 | 100,678.7 | 18,482 |
| 8 | JUWELS Booster Module - Bull Sequana XH2000, AMD EPYC 7402 24C 2.3GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation Cluster Suite, Atos Forschungszentrum Juelich [FZJ] Germany | 449,280 | 44,120.0 | 70,980.0 | 1,764 |
| 9 | HP5 - PowerEdge C4140, Xeon Gold 6252 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR Infiniband, DELL EMC EMI S.p.A Italy | 669,760 | 35,450.0 | 51,720.8 | 2,252 |
| 10 | Voyager-EUS2 - ND7000n1_A100_v4, AMD EPYC 7V12 48C 2.35GHz, NVIDIA A100 80GB, Mellanox HDR Infiniband, Microsoft Azure Azure East US 2 United States | 253,440 | 30,050.0 | 39,531.2 | |

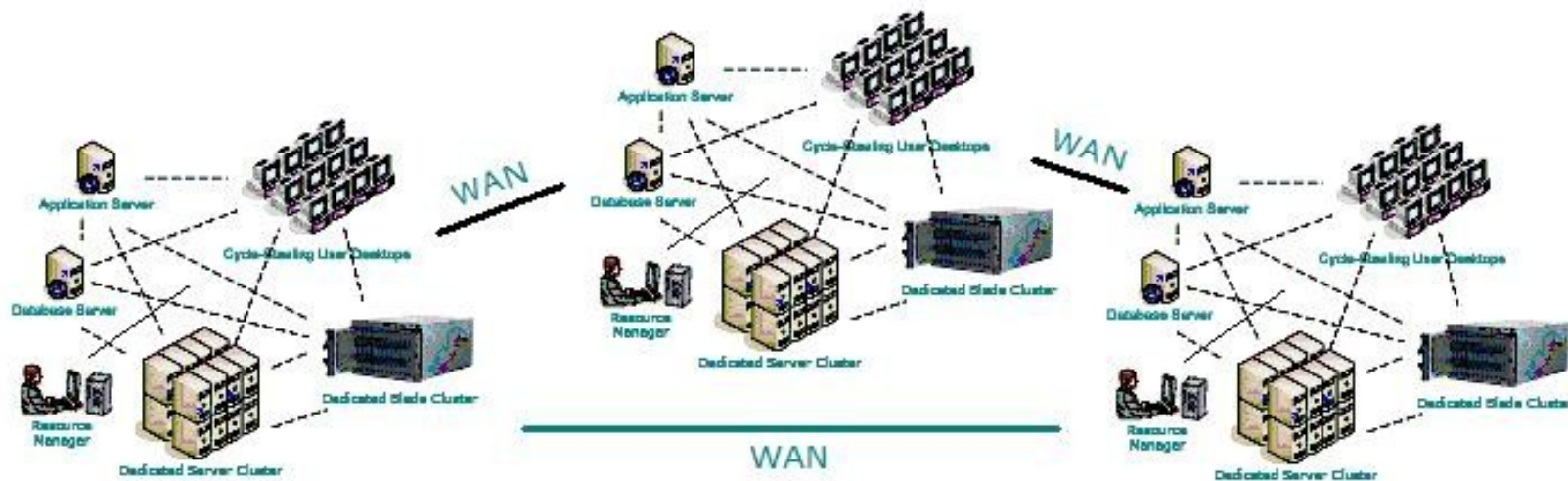
10 / 10

What's next?

- What's even **more** parallel and/or distributed than these hybrid systems?

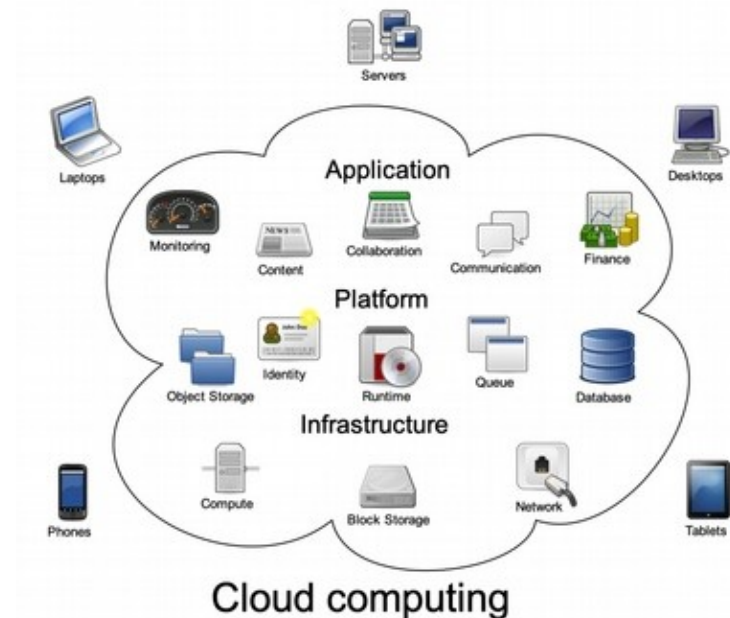
Grid computing

- Heterogenous nodes in disparate physical locations
 - Solving problems or performing tasks of interest to a large number of diverse groups
 - Hardware: different CPUs, GPUs, memory layouts, etc.
 - Software: different OSes, [Folding@Home](#), [Condor](#), [GIMPs](#), etc.



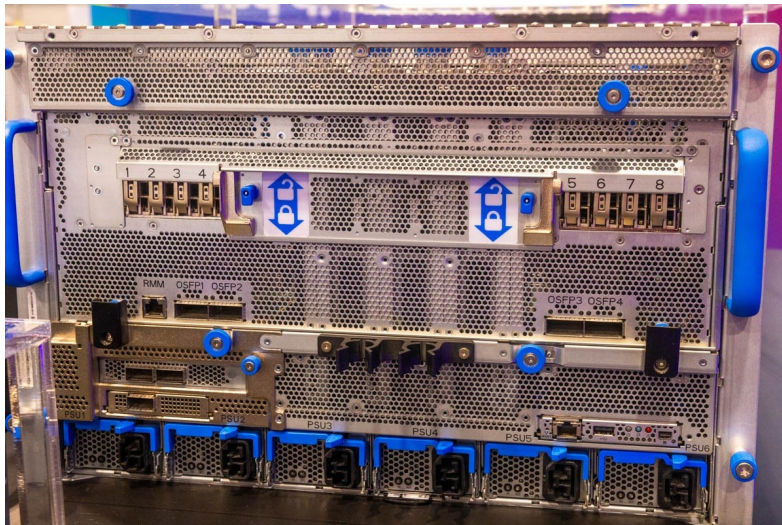
Cloud computing

- Homogenous centralized nodes
 - **Infrastructure as a Service** (IaaS) and **Software as a Service** (SaaS)
 - Hardware: large data centers with thousands of servers and a high-speed internet connection
 - Software: virtualized OS and custom software (**Docker**, etc.)



Cloud computing

- Surprise #3 ranked result in Top500 list revealed at SC23 in November 2023
 - Intel Xeon Platinum 8480C CPUs and Nvidia H100 GPUs
 - 1.1M cores total, sustaining 561 petaflops on HPL
 - Infiniband interconnect w/ direct GPU-to-GPU links



Dulles Technology Corridor

- Business cluster in Northern Virginia
 - Ashburn to Tysons Corner, along VA 7 and VA 267
 - In 2009, over 50% of all US-based Internet traffic was routed through data centers in this region
 - Home of AWS “US East” region, hosting ~70% of AWS IP addresses
 - <https://www.datacentermap.com/usa/virginia/>

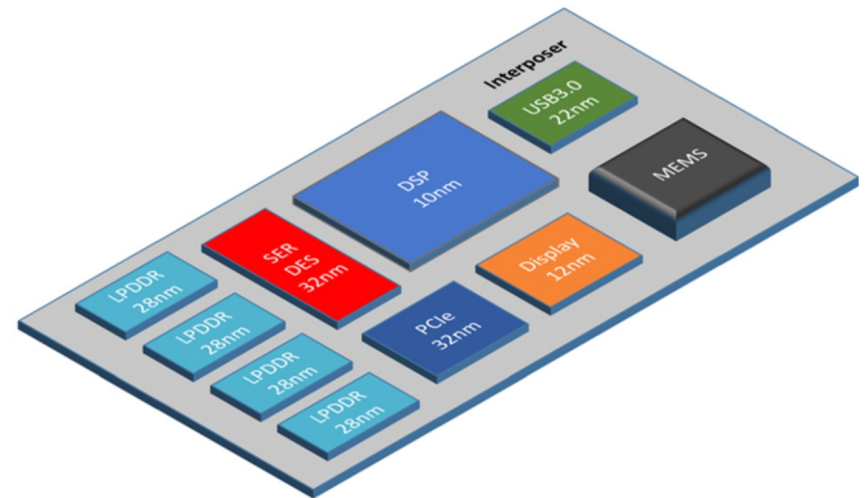
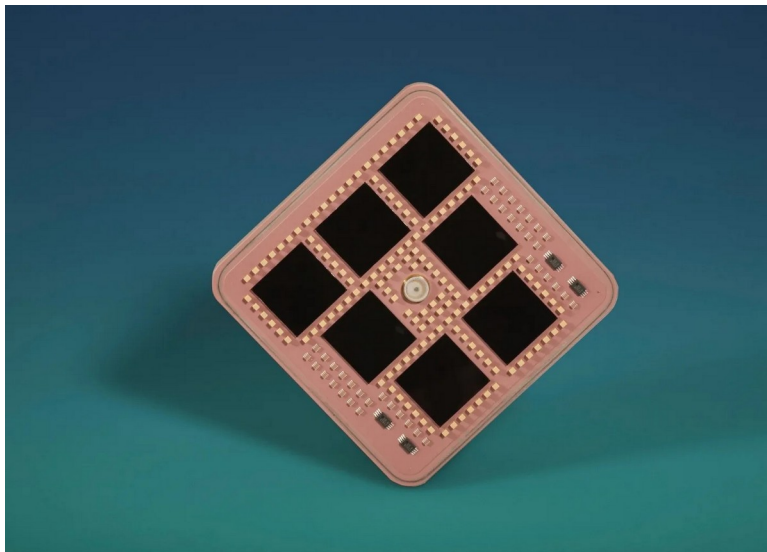


What's next?

- What's on the horizon?

Novel technologies

- **Chiplet**: small integrated circuit
 - Optimized for a specific functional purpose
 - Combined with other chiplets in a single **multi-chip module**
 - Easier to replace faulty chips during manufacturing



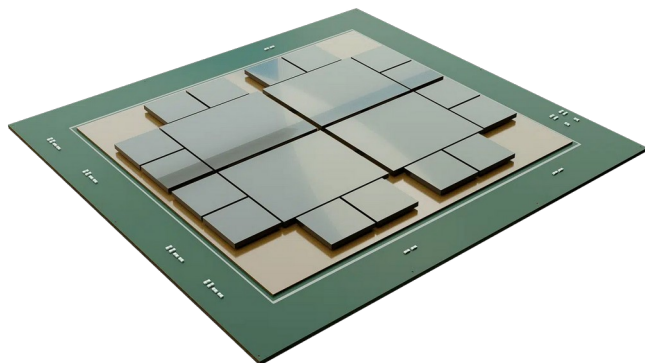
<https://en.wikipedia.org/wiki/Chiplet>

<https://www.eetimes.com/chiplets-a-short-history/>

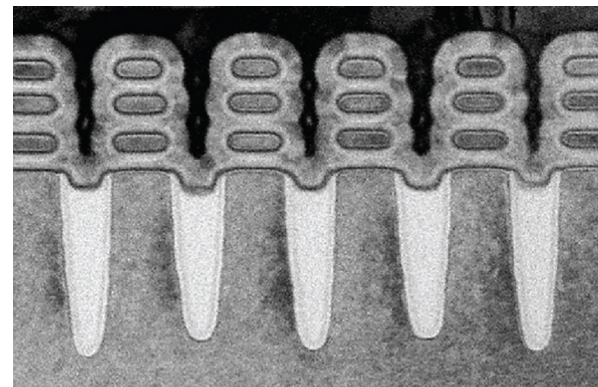
https://semiengineering.com/knowledge_centers/packaging/advanced-packaging/chiplets/

Novel technologies

- **Optical**
 - Use photon waves instead of electrons for visual AI, pattern recognition, and cryptography (recent emphasis on fully-homomorphic encryption)
 - Example: **Optalysis Enable** etile
- **Nanosheet transistors**
 - Vertical stacking technology that promises 40% performance boost and 75% power reduction over traditional transistors
 - Samsung working on commercial production

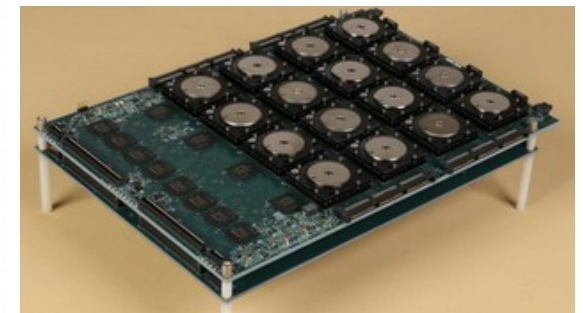
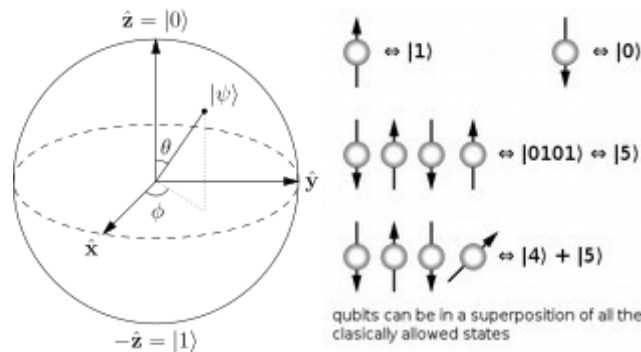


Optalysis Enable



Novel architectures

- **Memory-centric**
 - Fast memory fabrics w/ in-chip processing
 - Example: [HPE The Machine](#) (announced in 2014)
- **Quantum**
 - Leverage quantum superposition and entanglement (qubits)
 - Example: [D-Wave 2000Q](#) (2048 qubits) and [IBM QX](#) (5 and 16 qubits)
- **Neuromorphic**
 - Specialized, low-power hardware that emulates neural networks
 - Example: [IBM TrueNorth](#) released in 2014 (4096 cores, 1 million neurons)



Quantum computing

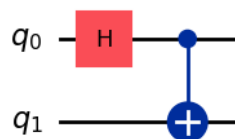
- IBM Quantum System Two

- Three 133-qubit Heron processors
- Similar capabilities available on IBM Cloud for free (10 minutes per month) or \$1.60/second

```
# Create a new circuit with two qubits  
qc = QuantumCircuit(2)
```

```
# Add a Hadamard gate to qubit 0  
qc.h(0)
```

```
# Perform a controlled-X gate on qubit 1,  
# controlled by qubit 0  
qc.cx(0, 1)
```



<https://www.youtube.com/watch?v=Qndz54SGCAs>

Neuromorphic computing

- April 2024: Intel Hala Point (Sandia)

Hala Point, the world's largest neuromorphic system

Largest capacity: 1.15 billion neurons and 128 billion synapses

Scalable: 1,152 Loihi 2 processors with 140,544 neuromorphic cores and 2,304 x86 cores



Fast: 380 trillion synaptic operations per second, 240 trillion neuron operations per second, 16 petabytes per second of memory bandwidth

Efficient: 15 TOPS/W executing sparse deep neural networks with an equivalent 30 quadrillion 8-bit operations per second*

Introducing Intel's most advanced neuromorphic system to date, code-named Hala Point. This system demonstrates state-of-the-art computational efficiencies for more efficient and sustainable AI.

Hala Point could enable advancements in future real-time continual learning, inference, and optimization for applications like science and engineering problem-solving, logistics, smart city infrastructure management, large language models (LLMs) and AI agents.

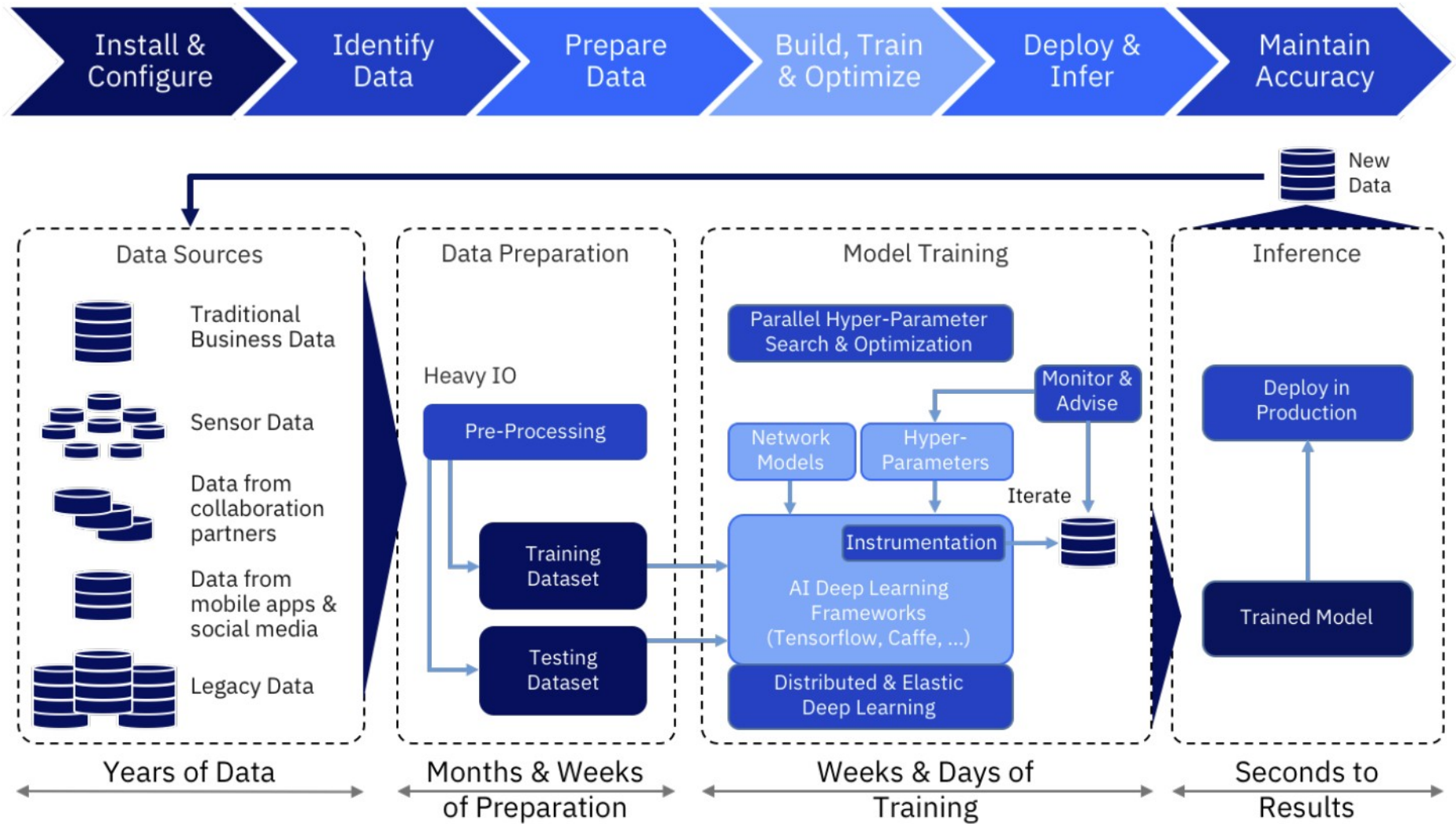
*Characterization performed with a multi-layer perceptron (MLP) network with 14,784 layers, 2048 neurons per layer, 8-bit weights stimulated with random noise. The Hala Point implementation of the MLP network is pruned to 10:1 sparsity with sigma-delta neuron models providing 10% activation rate. Results as of testing in April 2024. Results may vary.

© Copyright Intel Corporation. All rights reserved. Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries. Other names and brands may be claimed as the property of others.

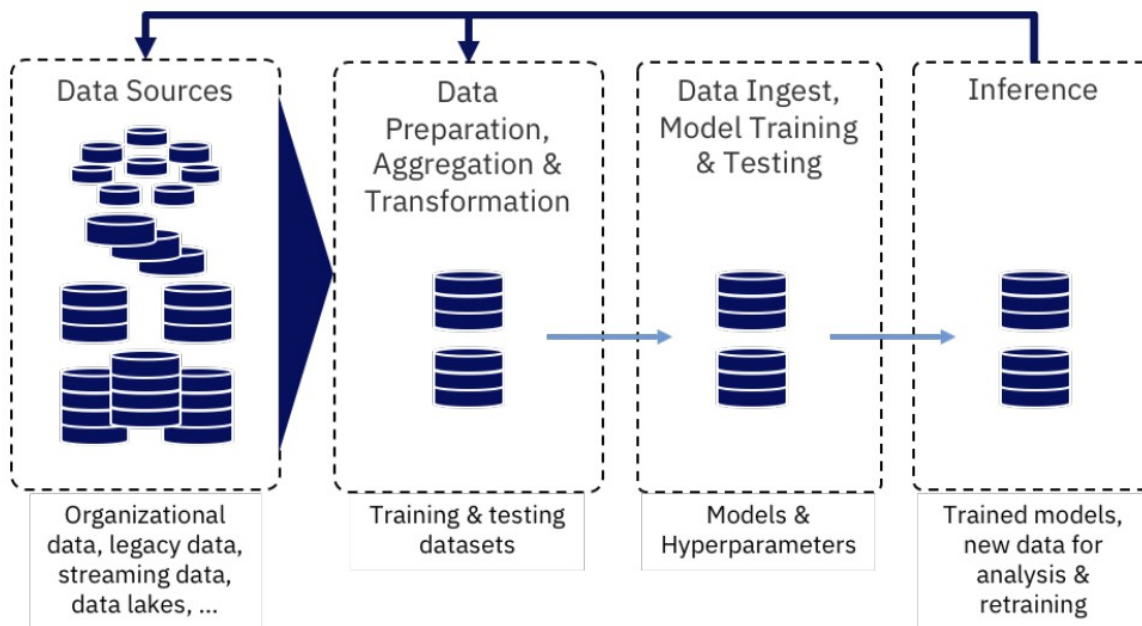
intel.



Deep Learning ML



Deep Learning ML



High bandwidth and low latency between the storage and compute nodes is absolutely critical, and sufficient bandwidth between the nodes needs to also be considered for data ingest and transformation phase of the workflow. Performance is key when training models to make sure sufficient data is delivered to the systems to keep the GPUs running at capacity, so a high-speed network subsystem is needed for the training cluster (i.e., Fast ethernet and InfiniBand).

Novel HPC Architectures for AI

NVIDIA DGX SuperPOD

Purpose-built for the unique demands of AI.



The World's First Turnkey AI Data Center

NVIDIA DGX SuperPOD™ is AI data center infrastructure that enables IT to deliver performance—without compromise—for every user and workload. As part of the **NVIDIA DGX™ platform**, DGX SuperPOD supports hybrid deployments and offers leadership-class accelerated infrastructure and scalable performance for the most challenging AI workloads, with industry-proven results.

Top10 systems (Spring 2024)

Microsoft

| Rank | System | Cores | Rmax (PFlop/s) | Rpeak (PFlop/s) | Power (kW) |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|----------------|-----------------|------------|
| 1 | Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 8,699,904 | 1,194.00 | 1,679.82 | 22,703 |
| 2 | Aurora - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States | 4,742,808 | 585.34 | 1,059.33 | 24,687 |
| 3 | Eagle - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Microsoft Azure United States | 1,123,200 | 561.20 | 846.84 | |
| 4 | Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu Interconnect D, Fujitsu RIKEN Center for Computational Science Japan | 7,630,848 | 442.01 | 537.21 | 29,899 |
| 5 | LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland | 2,752,704 | 379.70 | 531.51 | 7,107 |
| 6 | Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, EVIDEN EuroHPC/CINECA Italy | 1,824,768 | 238.70 | 304.47 | 7,404 |
| 7 | Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States | 2,414,592 | 148.60 | 200.79 | 10,096 |
| 8 | MareNostrum 5 ACC - BullSequana XH3000, Xeon Platinum 8460Y+ 40C 2.3GHz, NVIDIA H100 64GB, Infiniband NDR200, EVIDEN EuroHPC/BSC Spain | 680,960 | 138.20 | 265.57 | 2,560 |
| 9 | Eos NVIDIA DGX SuperPOD - NVIDIA DGX H100, Xeon Platinum 8480C 56C 3.8GHz, NVIDIA H100, Infiniband NDR400, Nvidia NVIDIA Corporation United States | 485,888 | 121.40 | 188.65 | |
| 10 | Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 1,572,480 | 94.64 | 125.71 | 7,438 |

NVIDIA

The Future™

WHO WON?

WHO'S NEXT?

YOU DECIDE

