Parallel and Distributed Programming

Imad Kissami[1]

[1]Mohammed VI Polytechnic University, Benguerir, Morocco

October 27, 2025

OUTLINE

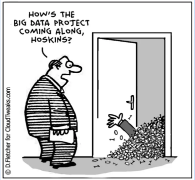
- The flood of Data
- What's HPC ?
- A brief introduction on hardware
- Modern supercomputers

THE FLOOD OF DATA

- In 2023: Internet user ~ 15.87 TB of data daily.
- By 2024: TikTok videos ~ 7.35 TB per day.
- By 2024: Self-driving car ~ 20 TB per day
- By 2025: cloud storage ~ 200+ zettabytes.

- 1 TB => 10^{12} Bytes
- 1 Zettabyte => 10^{21} Bytes

Enough memory?

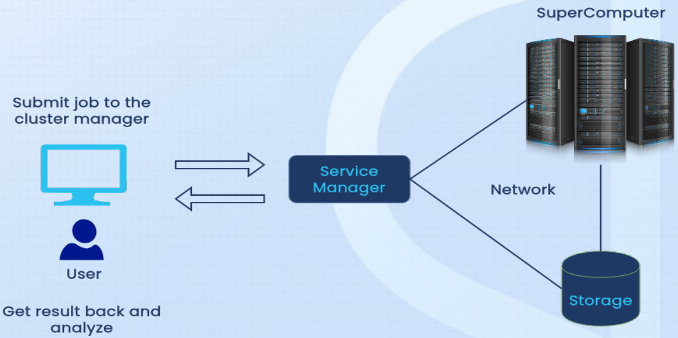


Enough compute?

WHAT'S HPC?

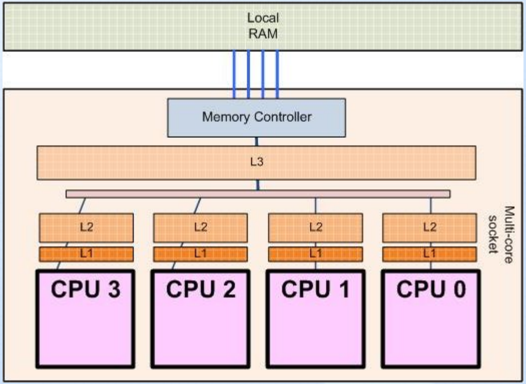
Leveraging distributed compute resources to solve complex problems

- Terabytes ⟶ Petabytes ⟶ Zetabytes of data
- Results in minutes to hours instead of days or weeks



A BRIEF INTRODUCTION ON HARDWARE

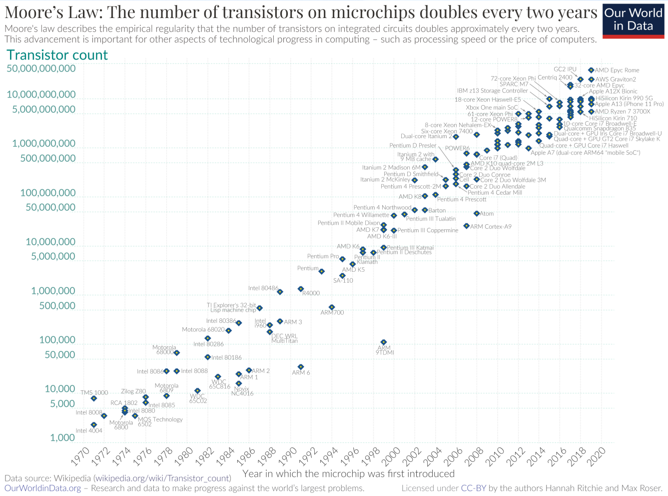
Modern architecture (CPU)



A BRIEF INTRODUCTION ON HARDWARE

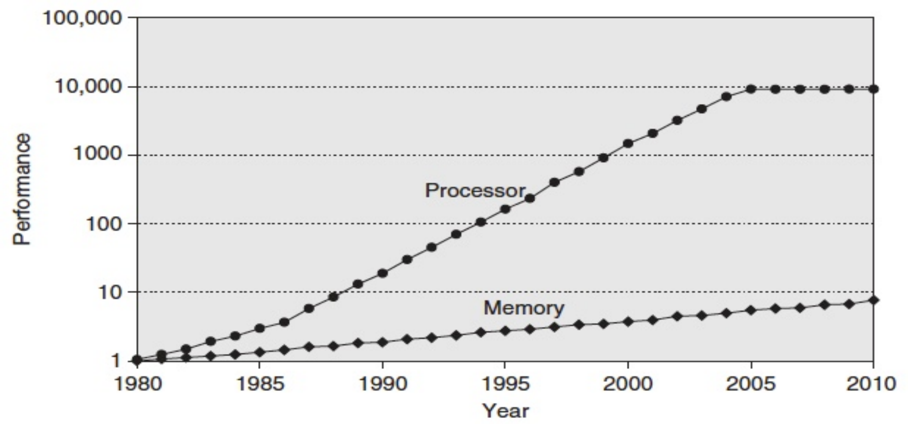
Moore's Law

- Number of transistors: from 37.5 million(2000) to 100 billion(2022)
- Cpu speed: from 1.3GHz to 6.5GHz
- By 2030 => 1 trillion transistors!



A BRIEF INTRODUCTION ON HARDWARE

CPU vs RAM speeds



Common Processors

Processor	L3 cache (MB)	Nb. of Cores	Freq. (Ghz)
Xeon 6766E (formerly Sierra Forest)	108	144	1.9-2.7
Xeon 6776P-B (formerly Granite Rapids-D)	288	72	2.3-3.5
i9-285K (Desktop, formerly Arrow Lake)	30	24	3.2-5.7
i7-251E (Mobile, formerly Bartlett Lake)	26	24	2.1-5.6

Table: Some Intel processors

Processor	L3 cache (MB)	Nb. of Cores	Freq. (Ghz)
AMD EPYC 9965 (server)	384	192	2.25-3.7
AMD EPYC 9845 (server)	320	160	2.1-3.7
AMD Ryzen 9 9950X3D (Desktop)	72 MB	16	4.3-5.7

Table: Some AMD processors

HPC MYTHS (AND REALITIES)

- Myth: A niche for researchers, geeks, and "eggheads."
  Reality: Used widely in oil & gas, automotive, aerospace, manufacturing, pharma, finance, and more.
- Myth: It's not for the cloud / not needed in the cloud.
  Reality: Cloud offers HPC instances and high-speed interconnects.
- Myth: HPC means one giant mainframe/supercomputer only.
  Reality: Modern HPC spans clusters, accelerators, and even edge devices.
- Myth: HPC is only MPI / Fortran.
  Reality: The ecosystem includes Python, R, C++, CUDA/HIP, OpenMP, SYCL, and task-based runtimes/workflows.

HPC MYTHS (AND REALITIES)

- Myth: HPC is only about FLOPS.
  Reality: Memory bandwidth, storage I/O, and latency are often the real bottlenecks.
- Myth: HPC adoption is too costly.
  Reality: Shared facilities and cloud reduce costs; pay-as-you-go models are available.
- Myth: HPC isn't reproducible.
  Reality: Containers, modules, and workflow managers enable portability and reproducibility.

CLOUD COMPUTING

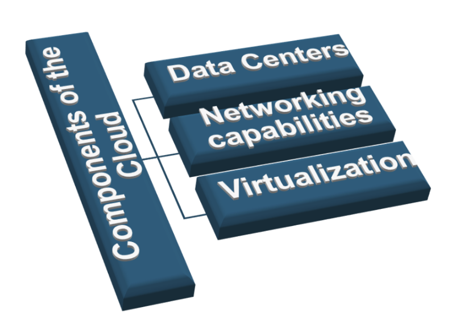
Accessing powerful computing resources via the internet from remote data centers.

- Enterprises – Store and process large-scale business data offsite.
- AI Research – Train deep learning models using high-performance cloud GPUs.
- Application Deployment – Seamless deployment and scaling of apps across regions and devices.

–> Elastic computing resources

–> Reduces infrastructure and maintenance costs

–> Enables ubiquitous access to data and services from anywhere



- AWS, Microsoft Azure, GCP
- ACS (UM6P)

EDGE COMPUTING

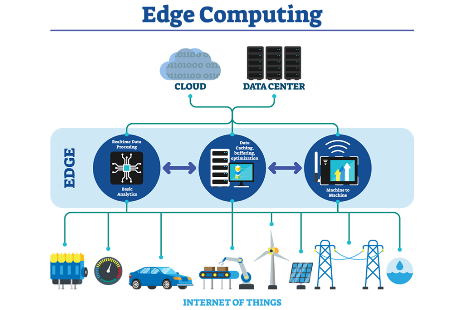
Local Processing: Processing data as close to their source as possible.

- Smartphones – Local face recognition for fast, secure authentication.
- Autonomous Vehicles – Real-time sensor processing for instant decisions.
- Industrial IoT – Edge servers monitor and control machines efficiently.

-> Minimizes latency

-> Reduces data transit costs

-> Enables real-time feedback and decision-making



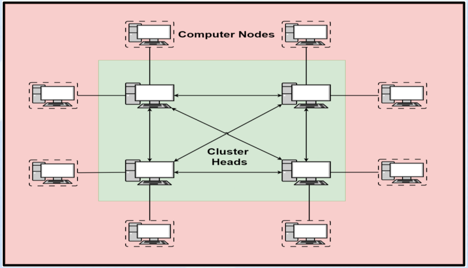
SUPERCOMPUTER

Use of powerful computing systems and parallel architectures to solve large-scale problems efficiently.

- Scientific Research – Simulate climate models, particle interactions, and astrophysics.
- Engineering – Run complex fluid dynamics simulations, structural analysis, and optimization.
- Business & Finance – Perform large-scale risk analysis, forecasting, and market modeling.

-> Fast massive data processing

-> Accelerated simulations



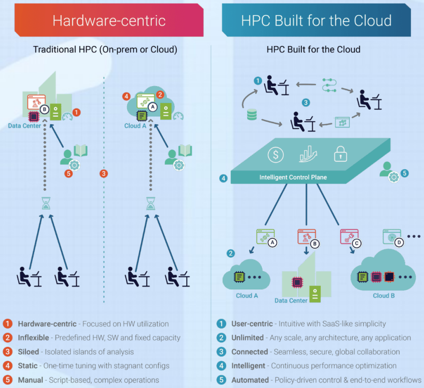
- El Capitan, Frontier, Aurora, Jupiter
- TOUBKAL (UM6P)

HPC CLOUD

Combines the scalability of cloud computing with the power of high-performance computing.

- Researchers – Run large-scale simulations on demand without owning infrastructure.
- Industry – Perform engineering, AI, and data analytics tasks in a flexible cloud environment.

- Elastic HPC resources
- On-demand large-scale simulations



- AWS, Azure, GCP

FLOPS: THE SPEED OF HPC (KILO $\rightarrow$ YOTTA)

HPC term often applies to systems that function above a TFLOPS or $O(10^{12})$ floating-point operations per second (FLOPS).

Name	Unit	Value
kiloFLOPS	kFLOPS	10^3
megaFLOPS	MFLOPS	10^6
gigaFLOPS	GFLOPS	10^9
teraFLOPS	TFLOPS	10^{12}
petaFLOPS	PFLOPS	10^{15}
exaFLOPS	EFLOPS	10^{18}
zettaFLOPS	ZFLOPS	10^{21}
yottaFLOPS	YFLOPS	10^{24}

Table: Scales of Floating-Point Operations per Second (FLOPS)

TOP500 AT A GLANCE

Biannual ranking of the world's fastest supercomputers (released at ISC in June and SC in November).

Benchmark: HPL (LINPACK) in double precision; key metrics include:

- R_{max} – Measured performance (TFLOPS / PFLOPS)
- R_{peak} – Theoretical peak performance
- System Power – Total power consumption (MW)
- Efficiency – Performance per Watt (GFLOPS/W)

Companion views:

- Green500 – Focus on energy efficiency
- HPCG – Memory and communication-intensive performance

TOP500 AT A GLANCE

Trends:

- Heterogeneous architectures (CPU + GPU) dominate
- High-speed interconnects (e.g., InfiniBand, Slingshot)
- Increasing focus on performance per watt (Perf/W)

Caveat:

- HPL is compute-bound and may overestimate performance for memory- or communication-bound workloads.
- Use HPCG or application-level benchmarks for balanced performance assessment.

LINPACK (HPL): TOP500 PERFORMANCE BENCHMARK

Solves a dense FP64 linear system via LU factorization; reports sustained PFLOP/s (R_{max}). Basis of the TOP500.

How it runs?

- Uses MPI + threads (often OpenMP)
- 2D block-cyclic data layout
- Tuned by:
  - Problem size N
  - Process grid $P \times Q$
  - Block size, panel factorization/lookahead
  - GPU BLAS, pinned memory

Strengths:

- Portable and comparable across systems
- Good proxy for peak floating-point throughput
- Exposes node/GPU compute and interconnect broadcast performance

LINPACK (HPL): TOP500 PERFORMANCE BENCHMARK

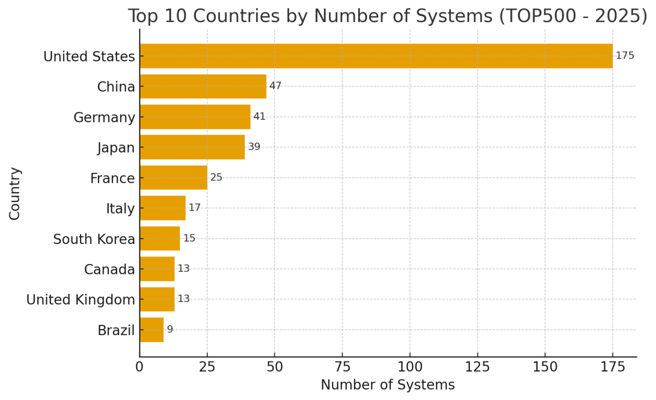
Caveats:

- Not representative of memory-bound or irregular applications
- Other benchmarks (e.g., HPCG) complement HPL

Related:

- HPL-AI / HPL-MxP – mixed precision, tensor-core acceleration
- HPCG – memory/communication intensive performance
- Energy efficiency tracked via GFLOPS/W (Green500)

TOP 10 COUNTRIES BY NUMBER OF SYSTEMS



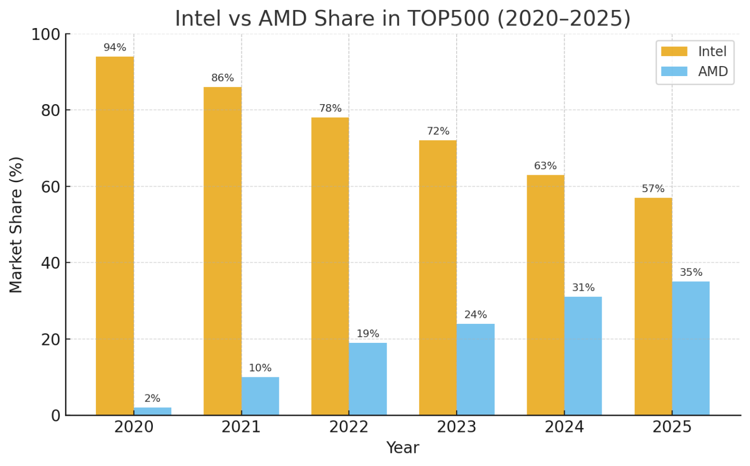
TOP500 – 2025

HPL TOP 10 SYSTEMS -- JUNE 2025

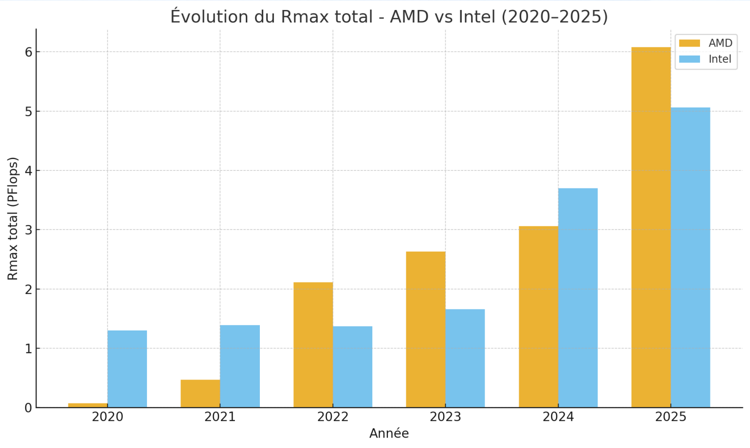
Site	Manufacturer	Computer	Country	Cores	R_{max} (PFLOP/s)	R_{peak} (PFLOP/s)	Power (kW)
Lawrence Livermore National Lab	HPE	El Capitan	USA	11,039,616	1,742.00	2,746.38	29,581
Oak Ridge National Laboratory	HPE	Frontier	USA	9,066,176	1,353.00	2,055.72	24,607
Argonne National Laboratory	Intel	Aurora	USA	9,264,128	1,012.00	1,980.01	38,698
Jülich Supercomputer Center	Eviden	JUPITER Booster	Germany	4,801,344	793.40	930.00	13,088
Microsoft Azure	Microsoft	Eagle	USA	2,073,600	561.20	846.84	N/A
Eni S.p.A.	HPE	HPC6	Italy	3,143,520	477.90	606.97	8,461
RIKEN R-CCS	Fujitsu	Fugaku	Japan	7,630,848	442.01	537.21	29,899
Swiss National Super-computing Centre	HPE	Alps	Switzerland	2,121,600	434.90	574.84	7,124
EuroHPC/CSC	HPE	LUMI	Finland	2,752,704	379.70	531.51	7,107
EuroHPC/CINECA	Eviden	Leonardo	Italy	1,824,768	241.20	306.31	7,494
UM6P	Intel	Toubkal	Morocco	72,000	3.16	5.01	754,92

Table: Top 10 + Toubkal Supercomputers (TOP500 - JUNE 2025)

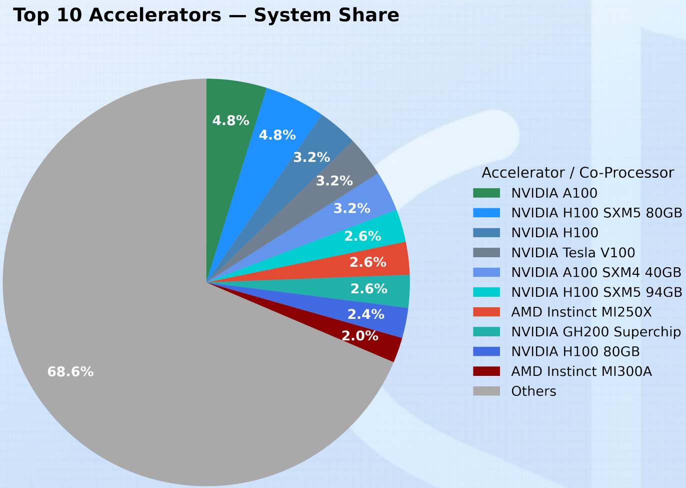
INTEL Vs AMD Share in TOP500 (2020–2025)



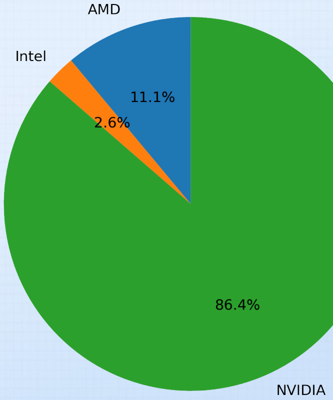
INTEL Vs AMD Performance in TOP500 (2020–2025)

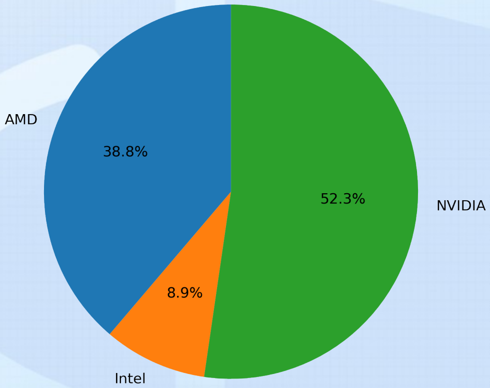


TOP500 ACCELERATORS -- JUNE 2025

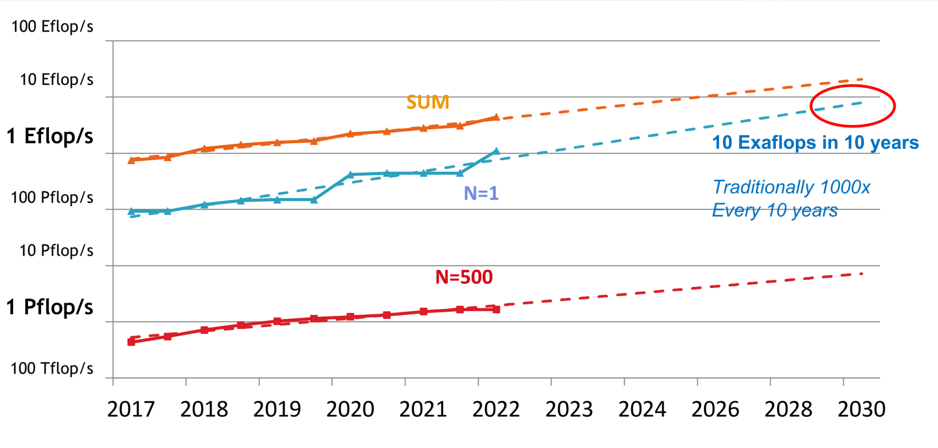


NVIDIA VS AMD VS INTEL ACCELERATORS -- JUNE 2025





PROJECTED PERFORMANCE



HPCG: BENCHMARK FOR SCIENTIFIC APPLICATIONS

hpcg-benchmark.org J. Dongarra, P. Luszczek, M. Heroux

- High Performance Conjugate Gradients (HPCG).
- Solves $Ax = b$, with A large and sparse, b known, and x computed.
- An optimized implementation of PCG contains essential computational and communication patterns that are prevalent in a variety of methods for discretization and numerical solution of PDEs.

Patterns:

- Dense and sparse computations.
- Dense and sparse collectives.
- Multi-scale execution of kernels via MG (truncated) V-cycle.
- Data-driven parallelism (unstructured sparse triangular solves).

Strong verification: via spectral properties of PCG.

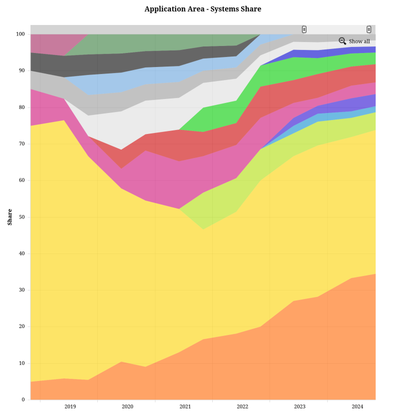
HPCG TOP 10 SYSTEMS -- JUNE 2025

#	TOP500 Rank	Manufacturer	Computer	Country	Rmax [TFlop/s]	HPCG [PFlop/s	HPCG/ Peak	HPCG / HPL
1	1	HPE	El Capitan – Cray EX255a, AMD EPYC 4th Gen 24C 1.8GHz, Instinct MI300A, Slingshot-11	USA	1,742.0	17.407	1.0%	1.0%
2	7	Fujitsu	Fugaku – A64FX 48C 2.2GHz, Tofu interconnect D	Japan	442.0	16.005	3.0%	3.6%
3	2	HPE	Frontier – Cray EX235a, AMD EPYC 3rd Gen 64C 2GHz, Instinct MI250X, Slingshot-11	USA	1,353.0	14.054	0.8%	1.2%
4	3	Intel	Aurora – Cray EX, Xeon Max 9470 52C 2.4GHz, Intel GPU Max, Slingshot-11	USA	1,012.0	5.613	0.6%	1.0%
5	9	HPE	LUMI – Cray EX235a, AMD EPYC 3rd Gen 64C 2GHz, Instinct MI250X, Slingshot-11	Finland	379.7	4.587	0.8%	1.1%
6	8	HPE	Alps – Cray EX254n, NVIDIA Grace 72C 3.1GHz, GH200 Superchip, Slingshot-11	Switzerland	434.9	3.671	0.9%	0.8%
7	10	Atos	Leonardo – BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100, HDR InfiniBand	Italy	241.2	3.114	1.0%	1.3%
8	15	HPE	ABCI 3.0 – Cray XD670, Xeon Platinum 8558 48C 2.1GHz, NVIDIA H200, Infiniband NDR200	Japan	145.1	2.446	1.2%	1.7%
9	25	HPE	Perlmutter – Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100, Slingshot-11	USA	79.23	1.905	2.0%	2.7%
10	20	IBM	Sierra – Power System AC922, POWER9 22C 3.1GHz, NVIDIA Volta GV100, Mellanox EDR	USA	94.64	1.796	1.4%	1.9%

Green500 TOP 10 SYSTEMS -- JUNE 2025

Green Rank	TOP500 Rank	Computer	Accelerator	Interconnect	Rmax/Power (GFlops/W)
1	259	JEDI – BullSequana XH3000 – NVIDIA Grace Hopper Superchip 72C 3GHz	NVIDIA GH200 Superchip	NVIDIA InfiniBand NDR200	72.733
2	148	ROMEO-2025 – BullSequana XH3000 – NVIDIA Grace Hopper Superchip 72C 3GHz	NVIDIA GH200 Superchip	NVIDIA InfiniBand NDR200	70.912
3	484	Adastra 2 – HPE Cray EX255a – AMD 4th Gen EPYC 24C 1.8GHz	AMD Instinct MI300A	HPE Slingshot-11	69.098
4	183	Isambard-AI phase 1 – HPE Cray EX254n – NVIDIA Grace 72C 3.1GHz	NVIDIA GH200 Superchip	HPE Slingshot-11	68.835
5	255	Otus (GPU only) – Lenovo ThinkSystem SD665-N V3 – AMD EPYC 9655 96C 2.6GHz	NVIDIA H100 SXM5 80GB	InfiniBand NDR	68.177
6	66	Capella – Lenovo ThinkSystem SD665-N V3 – AMD EPYC 9334 32C 2.7GHz	NVIDIA H100 SXM5 94GB	InfiniBand NDR200	68.053
7	304	SSC-24 Energy Module – HPE Cray XD670 – Intel Xeon Gold 6430 32C 2.1GHz	NVIDIA H100 SXM5 80GB	InfiniBand NDR400	67.251
8	85	Helios GPU – HPE Cray EX254n – NVIDIA Grace 72C 3.1GHz	NVIDIA GH200 Superchip	HPE Slingshot-11	66.948
9	399	AMD Ouranos – BullSequana XH3000 – AMD 4th Gen EPYC 24C 1.8GHz	AMD Instinct MI300A	InfiniBand NDR200	66.464
10	412	Henri – Lenovo ThinkSystem SR670 V2 – Intel Xeon Platinum 8362 32C 2.8GHz	NVIDIA H100 80GB PCIe	InfiniBand HDR	65.396

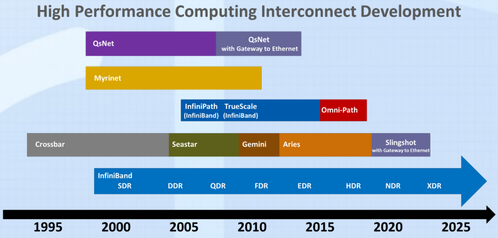
APPLICATIONS AREA



- Others
- Research
- IT Services
- Finance
- Chemistry
- Energy
- Weather and Climate Research
- Software
- Geophysics
- Information Service
- Aerospace
- Logistic Services
- Electronics
- Services
- Web Services
- Information Processing Service
- Semiconductor
- Automotive
- Telecommunication
- Database
- Defense

INTERCONNECTS IN SUPERCOMPUTERS

- Definition: The network fabric that links CPUs, GPUs, and compute nodes, allowing them to exchange data quickly and in parallel.
- Key Performance Metrics:
  - Latency: measured in microseconds (μs)
  - Per-link bandwidth: measured in GB/s
  - Message rate: measured in million messages per second (Mmsg/s)
- Common Network Fabrics:
  - InfiniBand (HDR, NDR generations)
  - HPE Slingshot (used in Frontier, El Capitan, Aurora)
  - High-speed Ethernet / RoCE
  - NVLink / NVSwitch for in-node GPU communication



- Role: Critical for scaling performance across thousands of nodes; determines overall efficiency and parallel communication overhead.

TOP PERFORMING INTERCONNECT -- TOP 5

Rank	Interconnect	Link Rate (Gb/s)	Why it leads
1	InfiniBand NDR400	400	Very low latency, high message rate; SHARP offload; GPUDirect RDMA — common in exascale/AI systems.
2	InfiniBand NDR200	200	NDR features with lower per-port rate; mature ecosystem and toolchain.
3	HPE Slingshot-11	200	Adaptive routing on Dragonfly+ topology, congestion control; backbone of HPE Cray exascale systems.
4	InfiniBand HDR	200	Proven across many TOP500 systems; strong collectives and RDMA offloads.
5	InfiniBand EDR	100	Lower latency and jitter than 100G Ethernet/RoCE for HPC collectives.

Note: “Best” depends on workload and network design (topology/rails), but these five typically deliver the highest sustained HPC performance.

TOP 500 INTERCONNECT SYSTEM SHARE -- JUNE 2025