International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Overview of High-Performance Computing and It's Current Trend in India

Prafull Siddharth Barathe

Dept. of AI&DS, SRTTC's Faculty of Engineering, Kamshet, Pune

Abstract

This is an era of High-Performance Computing abbreviated as HPC. HPC evolved due to increase in demand of Processing speed. It is a hybrid technology that continuously changes its colors, so it is hard to define it in static words. However, if we still persist to define it the effort leads to a definition as below – "HPC is a branch of computer science that concentrates on combining the power of computing devices, networking infrastructure and storage systems using parallel hardware, networking and software technology that utilizes human expertise to gain optimized and sustained performance for solving scientific, engineering, big data and research problems including grand challenges like computer modeling, simulations and analysis of data.

Keywords: HPC, High-Performance Computing

Introduction:

High-Performance Computing (HPC) refers to the practice of aggregating computing power to deliver far greater performance than typical desktop computers or workstations. HPC systems allow organizations to process large amounts of data quickly and solve complex problems that require significant computational resources. HPC architecture involves using parallel processing to run advanced applications efficiently and reliably.

High-performance computing (HPC) involves multiple interconnected robust computers operating in parallel to process and analyze data at high speeds. HPC architecture refers to the HPC design and structure that enable HPC clusters to handle tasks involving large datasets and complex calculations.

History of High-Performance Computing (Indian Perspective): India's history of high-performance computing (HPC) includes:

In the 1980s, India faced difficulties in purchasing supercomputers for academic and weather forecasting purposes.

In 1987, the Indian government requested a Cray X-MP supercomputer, but the US government denied the request due to concerns about India using it for weapons development. In response, India began developing its own supercomputer.

The PARAM 8000 was released in 1991 and was considered a success for C-DAC. It demonstrated that India had the second most powerful, publicly demonstrated, supercomputer in the world after the United States.

In April 2015, the Government of India initiated the NSM to make India a world leader in HPC. The NSM is being implemented jointly by the Ministry of Electronics and Information Technology (MeitY) and

Department of Science and Technology (DST).
The NSM has resulted in the creation of the Rudra server platform, which is manufactured in India. The Rudra server platform is cost efficient and power efficient in operation.
The NSM has also resulted in the creation of Trinetra high-speed interconnects.

Components of HPC: In an HPC architecture, a group of computers (nodes) collaborates on shared tasks. Each node in this structure accepts and processes tasks and computations independently. The nodes coordinate and synchronize execution tasks, ultimately producing a combined result.The HPC architecture has mandatory and optional components.

Mandatory Components
The compute, storage, and network components are the basis of an HPC architecture. The following sections elaborate on each component.

Compute: The compute component is dedicated to processing data, executing software or algorithms, and solving problems. Compute consists of computer clusters called nodes. Each node has processors, local memory, and other storage that collaboratively perform computations. The common types include:

- **Headnode or login nodes.** Entry points where users log in to access the cluster.
- **Regular compute nodes.** Locations for executing computational tasks.
- **Specialized data management nodes.** Methods for efficient data transfer within the cluster.
- **Fat compute nodes**. Handlers for memory-intensive tasks, as they have large memory capacity, typically exceeding 1TB.

Storage: The high-performance computing storage component stores and retrieves data generated and processed by the computing component.
HPC storage types are:

- **Physical:** Traditional HPC systems often use physical, on-premises storage. On-premise storage enables the inclusion of high-performance parallel file systems, storage area networks (SANs), or network-attached storage (NAS) systems. Physical storage is directly connected to the HPC infrastructure, providing low-latency access to data for compute nodes within the local environment.
- **Cloud Storage:** Cloud-based HPC storage solutions offer scalability and flexibility. In contrast to traditional external storage, typically the slowest computer system component, cloud storage within an HPC system operates at a high speed.
- **Hybrid**: A hybrid HPC storage solution combines both on-premises physical storage and cloud storage to create a flexible and scalable infrastructure. This approach allows organizations to address specific requirements, optimize costs, and achieve a balance between on-site control and the scalability offered by the cloud.

Network: The HPC network component enables communication and data exchange among the various nodes within the HPC system.
HPC networks focus on achieving high bandwidth and low latency. Different technologies, topologies, and optimization strategies are utilized to support the rapid transfer of large volumes of data between nodes.

HPC Scheduler: Task requests from the headnode are directed to the scheduler.
A scheduler is a vital HPC component. This utility monitors available resources and allocates requests across the nodes to optimize throughput and efficiency.
The job scheduler balances workload distribution and ensures nodes are not overloaded or underutilized.

Optional Components

Optional components in HPC environments are based on specific requirements, applications, and budget considerations. Optional components organizations choose to include in their HPC setups are:

- **GPU-Accelerated systems:** Boost computations for tasks that can be parallelized on both CPU cores and Graphics Processing Units (GPUs), such as simulations, machine learning, and scientific modeling. GPU acceleration operates in the background, facilitating large-scale processing within the broader system.
- Data management **software:** Systems that handle data storage, retrieval, organization, and movement within an HPC environment. Data management programs optimize system resource management according to specific needs.
- **InfiniBand** switch**:** Connects and facilitates communication between all nodes in the cluster.
- **Facilities and power:** Physical space required to accommodate HPC.
- **FPGAs (Field-Programmable Gate Arrays):** Customizable hardware acceleration is used in environments where highly efficient and low-latency processing is essential.
- **High-performance storage accelerators:** Parallel file systems or high-speed storage controllers enhance data access and transfer speeds.
- **Remote visualization nodes:** Help maintain computational efficiency when data visualization is critical to HPC workflows. The nodes offload visualization tasks from the main compute nodes.
- **Energy-efficient components:** Energy-efficient processors, memory, and power supplies minimize the environmental impact and operational costs and improve data center sustainability.
- **Scalable and flexible network fabric:** Enhances node communication, improving overall system performance.
- **Advanced security mechanisms:** Include hardware-based encryption, secure boot processes, and intrusion detection systems.

Types of HPC Architecture: The main hardware and system types for processing demanding computational tasks in an HPC system include parallel computing, cluster computing, and grid computing.

Parallel Computing: In HPC architecture, parallel computing involves organizing multiple nodes to work simultaneously on the same or similar computations. The primary goal is to increase computational speed and efficiency by dividing tasks into smaller subtasks that can be processed concurrently.
There are different parallelism forms, including:

- **Task parallelism.** A complex computational task is split into smaller, independent subtasks when different task parts are performed concurrently without significant inter-task dependencies.
- **Data parallelism.** Data distributed across multiple nodes allows each unit to execute the same operation on different data subsets simultaneously. Effective when the same operation needs to be performed on different segments concurrently.
- **Instruction-level parallelism.** Multiple instructions are executed concurrently within a single

processor. The structure can exploit the parallel nature of instruction-level operations to enhance computation's overall speed and efficiency.

This architecture type is well-suited for handling large-scale computations, such as simulations, scientific modeling, and data-intensive tasks, as it enables simultaneous calculations by harnessing the power of hundreds of processors.

Cluster Computing: In HPC architecture, cluster computing involves connecting multiple individual computers (nodes) that function as a single resource. Cluster computing enables nodes to collaborate, handle computational tasks, and share resources. A scheduler typically oversees the task coordination and resource management within a cluster.

This architecture is used in HPC for its ability to provide substantial processing power by leveraging the collective capabilities of interconnected nodes. Cluster computing is cost-effective because standard, off-the-shelf hardware can be scaled according to budget and requirements.

Grid and Distributed Computing: Grid or distributed computing connects geographically dispersed computing resources to form a single, virtual HPC system. Unlike cluster computing, which typically involves a localized node group, the grid reaches multiple locations within a single organization or across various institutions.

These nodes aren't necessarily working on the same or similar computations but are parts of a more complex problem. This architecture allows the pooling of computational power from diverse sources, enabling resource sharing and collaborative problem-solving over large distances.

General Design Principles: Despite HPC's different architecture types, several overarching principles apply when designing an HPC environment. General architecture principles are:

- **Dynamic architecture:** Avoid rigid, static architectures and cost estimates dependent on a steady-state model.
- **Scalability:** Create a scalable architecture that allows you to seamlessly add or remove computational resources, like nodes or processors.
- **Data management:** Devise an effective data management strategy to accommodate different data types, sizes, and access patterns. How data is stored, processed, and accessed impacts system performance.
- **Heterogeneous computing:** Combine hardware and software components, like GPUs, FPGAs, and accelerators, to streamline specific computational tasks.
- **Automation:** Use automated solutions for managing and provisioning system resources. Manual input is the least efficient component in an HPC environment.
- **Collaboration:** When working on a project that spans multiple organizations, it is vital to preselect collaborative tools, scripts, and data-sharing protocols during the design phase to avoid future incompatibility issues.
- **Workload testing:** HPC applications are complex and require comprehensive testing. This is the only way to measure HPC application performance in various environments.
- **Cost vs. Time:** For time-critical tasks, prioritize performance over cost. Focus on cost optimization for non-time-sensitive workloads.

Current Trends and Challenges: HPC continues to evolve, with trends focusing on exascale computing, cloud integration, and energy efficiency. However, there are challenges to widespread adoption, including:

Power Consumption: HPC systems consume significant amounts of energy, making efficiency a key concern in the design and deployment of these systems.

Scalability: As data volumes increase, ensuring that HPC systems can scale efficiently is a major challenge.

Cost: Building and maintaining HPC infrastructure is costly, often limiting access to large organizations or research institutions. Setting up an on-site HPC system requires a substantial budget, involving significant upfront costs for hardware, a skilled technician team, and establishing an on-premises data center. Maintaining HPC also entails high ongoing costs related to cooling and power bills.
Apart from common challenges, each architecture type also has specific issues:

- **Cluster computing:** Managing many nodes in a cluster is complex, and inefficient networking or poorly optimized parallelization leads to bottlenecks, limiting overall performance.
- **Grid or Distributed Computing:** Communication between geographically distributed resources introduces latency and communication overhead. Also, integrating resources across different administrative domains raises security and data privacy concerns.
- **Parallel computing:** Coordinating and synchronizing parallel tasks leads to potential synchronization issues. Parallel computing also comes with limited scalability.

Working systems in India : The **"National Supercomputing Mission (NSM)**: Building Capacity & Capability", has been initiated by Government of India in April 2015. The Mission envisages empowering our national academic and R&D institutions spread over the country by installing supercomputers of various capacities. Access to these supercomputers is provided through the National Knowledge Network (NKN). The NKN is another programme of the government which connects academic institutions and R&D labs over a high-speed network. Academic and R&D institutions as well as key user departments/ministries would participate by using these facilities and develop applications of national relevance. The Mission also includes development of highly professional High-Performance Computing (HPC) aware human resource for meeting challenges of development of these applications. The Mission implementation would bring supercomputing within the reach of the large Scientific & Technology community in the country and enable the country with a capacity of solving multi-disciplinary grand challenge problems.
The Mission would be implemented and steered jointly by the Department of Science and Technology (DST) and Department of Electronics and Information Technology (DeitY) at an estimated cost of Rs.4500 crore over a period of seven years.

I - NSM Systems

Sr. No.	Institute Name	HPC System Name	Computing Power
1. 1	IIT(BHU), Varanasi	PARAM Shivay	838TF

1. 2	IISER, Pune	PARAM Brahma	1.70PF
1. 3	IIT, Kharagpur	PARAM Shakti	1.66PF
1. 4	JNCASR, Bangalore	PARAM Yukti	1.8PF
1. 5	IIT, Kanpur	PARAM Sanganak	1.66PF
1. 6	C-DAC, Pune	PARAM Siddhi-AI	5.2PF/210PF (AI)
1. 7	IIT, Hyderabad	PARAM Seva	838TF
1. 8	NABI, Mohali	PARAM Smriti	838TF
1. 9	IISc, Bangalore	PARAM Pravega	3.3PF
1. 10	C-DAC, Bangalore	PARAM Utkarsh	838TF
1. 11	IIT, Roorkee	PARAM Ganga	1.66PF
1. 12	IIT, Gandhinagar	PARAM Ananta	838TF
1. 13	NIT, Trichy	PARAM Porul	838TF
1. 14	IIT, Guwahati	PARAM Kamrupa	838TF
1. 15	IIT, Mandi	PARAM Himalaya	838TF

II - R&D and Application development systems

Sr. No.	**Institute Name**	**HPC System Name**	**Computing Power**	**Year of Commissioning**
1. 16	C-DAC, Pune	SANGAM Testbed	150 TF	2017
1. 17	C-DAC, Pune	PARAM Shrestha	100 TF	2018
1. 18	C-DAC, Pune	PARAM Embryo	100 TF	2020

1. 19	C-DAC, Pune	PARAM Neel	100 TF	2020
1. 20	SETS, Chennai	PARAM Spoorthi	100 TF	2020
1. 21	C-DAC, Pune	Bioinformatics R&D Facility	230 TF	2021
1. 22	C-DAC, Bangalore	System Software lab	82TF	2020
1. 23	C-DAC, Pune	PARAM Sampooran	27 TF	2020

III - PARAM Vidya under NSM HRD (For education and training)

Sr. No.	**Institute Name**	**HPC System Name**	**Computing Power**	**Year of Commissioning**
1. 24	C-DAC, Pune	PARAM Vidya1	52.3 TF	2022
1. 25	IIT, Kharagpur	PARAM Vidya2	52.3 TF	2022
1. 26	IIT, Palakkad	PARAM Vidya3	52.3 TF	2022
1. 27	IIT, Chennai	PARAM Vidya4	52.3 TF	2022
1. 28	IIT, Goa	PARAM Vidya5	52.3 TF	2022

TOP 500: Top500 lists computers ranked by their performance on the LINPACK Benchmark. While top500 make every attempt to verify the results obtained from users and vendors, errors are bound to exist and should be brought to our attention. Top500 intend to continue to update this list half-yearly and, in this way, to keep track with the evolution of computers.

The main objective of the TOP500 list is to provide a ranked list of general purpose systems that are in common use for high end applications.

The 63rd edition of the TOP500 reveals that Frontier has once again claimed the top spot, despite no longer being the only exascale machine on the list. Additionally, a new system has found its way into the Top 10. The Frontier system at Oak Ridge National Laboratory in Tennessee, USA remains the most powerful system on the list with an HPL score of 1.206 EFlop/s. The system has a total of 8,699,904 combined CPU and GPU cores, an HPE Cray EX architecture that combines 3rd Gen AMD EPYC CPUs optimized for HPC and AI with AMD Instinct MI250X accelerators, and it relies on Cray's Slingshot 11 network for data transfer. On top of that, this machine has an impressive power efficiency rating of 52.93 GFlops/Watt – putting Frontier at the No. 13 spot on the GREEN500.

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	1,206.00	1,714.81	22,786
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	9,264,128	1,012.00	1,980.01	38,698
3	**Eagle** - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Azure Microsoft Azure United States	2,073,600	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	**Alps** - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200	1,305,600	270.00	353.75	5,194

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
	Superchip, Slingshot-11, HPE Swiss National Supercomputing Centre (CSCS) Switzerland				
7	**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	241.20	306.31	7,494
8	**MareNostrum 5 ACC** - BullSequana XH3000, Xeon Platinum 8460Y+ 32C 2.3GHz, NVIDIA H100 64GB, Infiniband NDR, EVIDEN EuroHPC/BSC Spain	663,040	175.30	249.44	4,159
9	**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
10	**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	

APPLICATIONS OF HPC:

HPC applications of national relevance are going to be developed and deployed. The areas chosen for developing applications include,

- Computational biology
- Climate modelling, weather prediction
- Engineering including CFD, CSM, CEM
- Disaster simulations and management
- Computational chemistry and material science
- Discoveries beyond Earth (Astrophysics)
- Big data Analytics

CONCLUSION:

HPC architecture plays a critical role in solving complex computational problems that cannot be handled by traditional computing systems. As HPC technology continues to advance, we can expect to see improvements in computational power, energy efficiency, and accessibility, allowing a broader range of industries to benefit from high-performance computing.

REFERENCES

1. https://dst.gov.in/national-super-computing mission#:~:text=The%20salient%20achievements%20of%20the,(cloud%2C%20HPC)%20applications.
2. https://www.meity.gov.in/content/high-performance-computinghpc
3. https://phoenixnap.com/kb/hpc-architecture
4. https://top500.org/
5. https://nsmindia.in/Apps