

Supercomputation @ IOPB

SAMKHYA (सांख्य)

Terascale Supercomputing Facility

Makrand Siddhabhatti
Systems Manager

Content

- ▶ What is High Performance Computing and computing power of a system
- ▶ Need of HPC system and Application
- ▶ Our System Components
 - * Compute Nodes (CPU Only, Nvidia and Intel Xeon Phi)
 - * Job Scheduler
 - * Storage
 - * Network Interconnect (Ethernet & Infiniband)
 - * Software and Tools
 - * Message Passing Interface (MPI)
- ▶ Infrastructure Detail
- ▶ Super computing in World & India

High Performance Computing

- ▶ A compute and data intensive dedicated infrastructure to run serial/parallel jobs(tasks) uninterruptedly for significant duration of time over high speed network.
- ▶ Evolution and HPC terminology
 - ▶ Cluster Computing
 - ✓ Homogeneous environment with limited number of resource confined at one place.
 - ▶ Grid Computing
 - ✓ Heterogeneous environment and geographically distributed in nature.
 - ▶ Super Computer
 - ✓ Homogeneous environment with exceptionally high number of compute at one place.

Need of HPC system and Application

- ▶ Computational material science.
- ▶ Condensed matter and complex systems.
- ▶ Nuclear and high energy physics.
- ▶ Physics @ LHC
 - ✓ Only 1 Higgs boson per 10 billion collisions in the Large Hadron Collider.
 - ✓ Collision produced 25 Petabytes of data / year.
- ▶ Weather forecast and Earth Science
 - ✓ Weather Forecast
 - ✓ Monsoon Prediction

Performance of HPC (Super Computer System)

“Samkhya (सांख्य), also referred to as Sankhya, Sāṃkhya, or Sāñkhya, is a Sanskrit word that, depending on the context, means "to reckon, count, enumerate, calculate, deliberate, reason, reasoning by numeric enumeration, relating to number, rational.”

- ▶ Performance of supercomputer system is measured in Floating point Operations Per Seconds (FLOPS), which is calculate as,

$$\text{FLOPS} = \text{No. of Cores} \times \text{Frequency} \times \text{Floating point number}$$

Floating point number = provided by Original Equipment Manufacturer (OEM)

Our System Components

▶ Head Node (Login) :

- ▶ A master node and user interface to system
- ▶ Responsible for job submission, accounting and final outcome.
- ▶ Runs Job Schedulers master daemon.
- ▶ 2 Login Node.

▶ Compute Nodes :

- ▶ Slave Node to master (head node).
- ▶ Compute intensive and involved in computation.
- ▶ Runs Job schedulers client daemon.

* 20 Nodes (CPU Only)

* 20 Nodes (CPU + Nvidia GPU)

* 20 Nodes (CPU + Intel Xeon Phi)

▶ I/O Nodes :

- ▶ Managing storage and its configuration.
- ▶ Manage meta data generated between transactions.
- ▶ 4 I/O Nodes.

Our System Components : contd....

Node Configuration : At a glance

Nodes	Cores / Cards	CPU Frequency	RAM	Processor / Cards
Login Node and I/O Node	16 Core/Node	2.40 GHz	64 GB	Intel Xeon E5-2630v3
Compute Node (20 nos.) (CPU Only)	24 Core/Node	2.50 GHz	128 GB	Intel Xeon E5-2680v3
Compute Node (20 nos.) (CPU + Nvidia GPU)	24 Core/Node 2 Nvidia Cards/Node	2.50 GHz	128 GB	Intel Xeon E5-2680v3 Nvidia Tesla K80 GPU (4992 CUDA Cores)
Compute Node (20 nos.) (CPU + Intel Xeon Phi)	24 Core/Node 2 Intel Xeon Phi Cards/Node	2.50 GHz	128 GB	Intel Xeon E5-2680v3 Intel Xeon Phi 7120P (61 Cores)

1440 CPU Cores, 40 Nvidia GPU Cards and 40 Intel Xeon Phi Cards delivers peak performance (theoretical) of ~ 222 Tera FLOPS.

Our System Components : contd....

▶ Job Scheduler:

- Responsible for job scheduling, tracking across compute nodes.
- Scheduler used in our HPC is Torque (Tera scale Open source Resource and Queue Manager).
- Other Scheduler : PBS, Moab, LoadLeveler, condor, SLURM.

▶ Storage:

- Lustre file system.
- Raid 6 (8+2) configuration.
- Usable storage space of ~ 50 TB.

Our System Components : contd

► Interconnect:

□ Ethernet

- Used for local network communication between compute and head node.
- CPU overhead and bottleneck in case of high speed transfer.

□ Infiniband

- Uses switched dedicated fabric topology widely used in high performance computing system
- High bandwidth & Low latency
- Direct memory Access (DMA), low CPU overhead
- Technologies :

Single Data Rate(2.5 Gbps), Double Data Rate (5.0 Gbps)

Quad data rate (10 Gbps) : 4X is 40 Gbps used in our HPC facility.

Fourteen Data Rate 10 (10 Gbps) & FDR14 (14 Gbps)

Enhance Data Rate (EDR) (25 Gbps)

Our System Components : contd....

► Software and Tools

- Intel Parallel Suite Studio XE Cluster Edition
- GNU compilers
- Openmpi

Compiler	Executable	Version
GNU C/C++	gcc / g++	4.4.7
GNU Fortran	gfortran / f95	4.4.7
Intel C / C++	icc / icpc	17.0.2
Intel fortran	ifort	17.0.2
Cuda Compiler	nvcc	Cuda 8
Intel parallel Compiler Suite	mpiicc, mpiifort, mpirun, mpiexec	17.0.2

Message Passing Interface (MPI)

- ▶ It is a communication protocol used in parallel computing.
- ▶ It create parallel environment through its standard library which helps to port application parallel in distributed environment.
- ▶ Various implementations,
 - ▶ OpenMPI
 - ▶ Intel MPI
 - ▶ MVAPICH
 - ▶ IBM MPI

Infrastructure Detail

▶ Air Conditioning:

- Two precision AC.
- Capacity of 10 TON each.

▶ Uninterrupted Power Supply:

- One 60 KVA UPS cater to precision AC.
- Three 40 KVA UPS with redundancy cater to HPC system.

Super computing in World & India

► World:

- List is published in twice in a year containing top500 supercomputer in world.
- URL is top500.org.

top fastest supercomputer in world as on November, 2017 :

- ❖ Sunway TaihuLight in National Supercomputer Center, Wuxi, Jiangsu, China with peak performance of ~ 93 PetaFLOP.
- ❖ Tianhe-2 (MilkyWay-2) in National Supercomputer Center, Guangzhou, China with peak performance of ~ 33 PetaFLOP.
- ❖ Piz Daint in Swiss National Supercomputing Centre, Switzerland with peak performance of ~ 25 PetaFLOP.

► India:

- List is published by Supercomputer Education and Research Centre (SERC), IISc, Bangalore.

top fastest supercomputer in India as on November, 2017 :

- ❖ SahasraT (SERC - Cray XC40) in SERC, IISc, Bangalore with peak performance of ~ 1.2 PetaFLOP (World Rank : 228).
- ❖ Aaditya (iDataPlex DX360M4) in Indian Institute of Tropical Meteorology, Pune with peak performance of ~ 790 TeraFLOP (World Rank : 368).

Questions

Contact us :

Computer Center

Institute of Physics, Bhubaneswar

Email : hpcsupport@iopb.res.in

Thank You.