Supercomputing in India

Ashish P. Kuvelkar
Senior Director (HPC- Tech)
C-DAC, Pune

ACM Winter School 2019 on High Performance Computing (HPC)

Agenda

The past

The present

The Future

The Past

The Beginning of C-DAC

Mid 1980s: India plans to buy a supercomputer from USA, to be used for predicting weather

Sale of Supercomputer to India, is denied

Government of India sets up C-DAC in 1988, thus launching the country's initiative in supercomputing

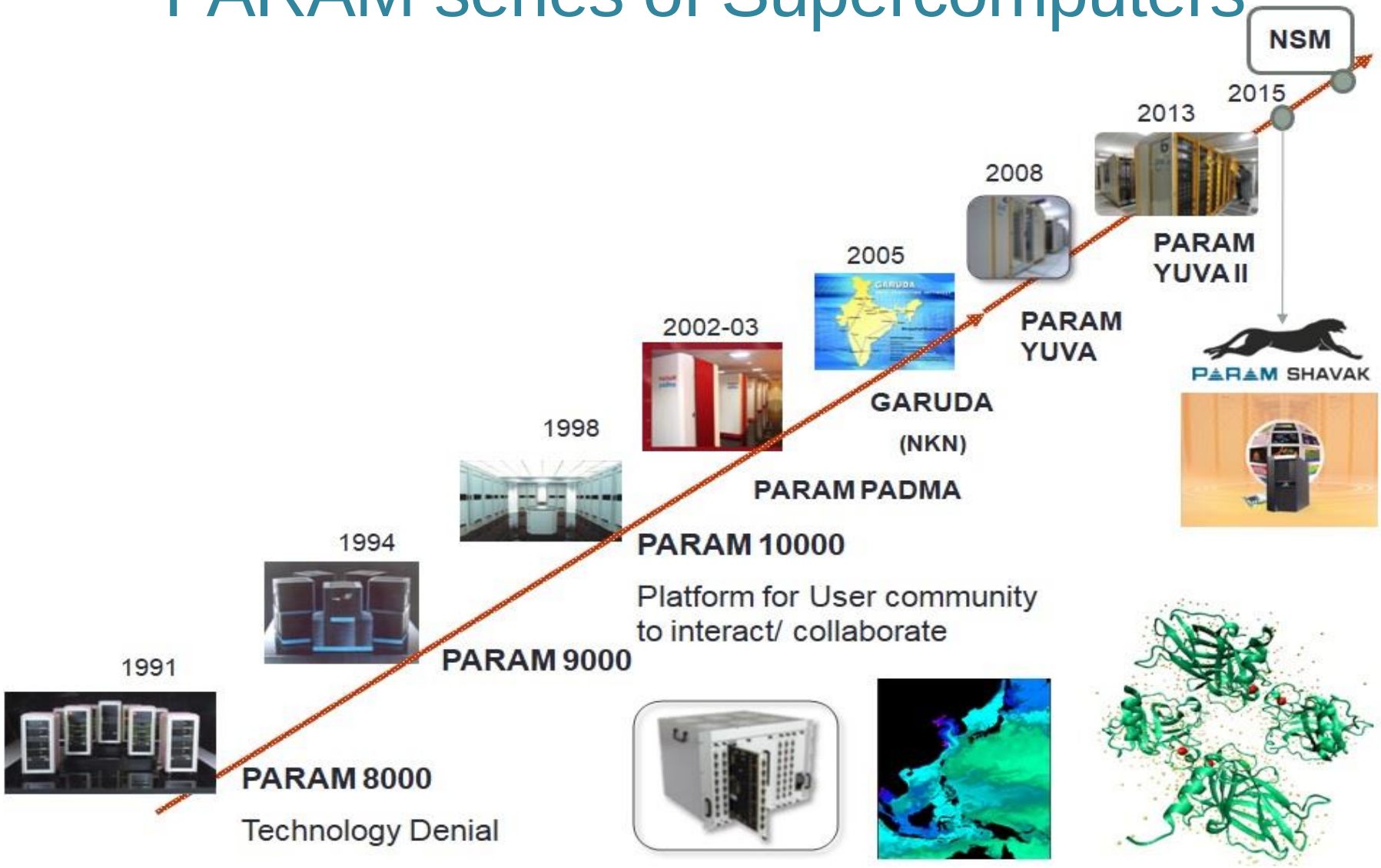
Goal of first mission

- To deliver parallel computer with peak computing power exceeding 1,000 Mflops (1 GigaFlops)
- Demonstrate applications of national importance

In July 1991, with mission effort of 300 man-years, PARAM8000 was delivered

Development of PARAM over the years

PARAM series of Supercomputers



Goals for the first Mission

In 3 years time, using 30 Crore budget delivering a Supercomputer with 1 GigaFlops peak performance

- Proportionate primary memory
- I/O bandwidth for connecting
  - Secondary storage
  - High performance Graphics
  - Standard network interfaces
- Parallel programming environments
- Plausible sustained performance for range of applications in science and engineering
- Make operational 3 to 4 large scale applications

PARAM 8000

Architecture

Uses INMOS Transputer

Distributed Memory Architecture

Message passing among nodes

Works as a back-end compute engine to hosts

- PC/AT, Sun Workstations or MicroVAx machine

Multiuser facility through

- Partitioning of nodes into logically disjoint fragments
- Assigning each partition to one user

Accommodates heterogeneous nodes

- Disk I/O nodes, Graphics nodes, Vector nodes, DSP nodes

PARAM Yuva

Nodes based on Intel Tigerton (Xeon 73XX) CPU – 54 TF compute power

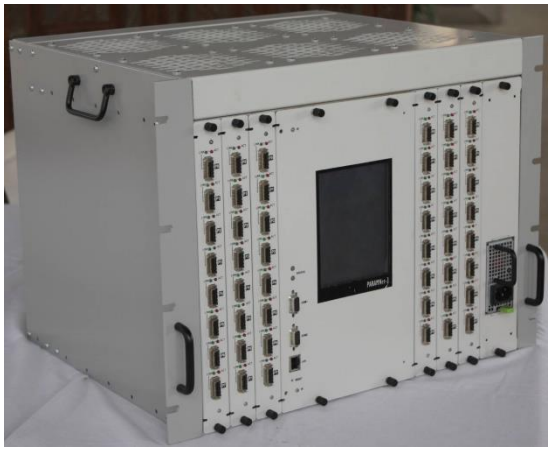
PARAMNet – III as the clos System Area Network running at 10 Gbits/sec

C-DAC's Communication Processor Gemini (FPGA based)

PARAMNet – III 48 port Switch and Network Interface Card

- DAPL protocol over copper cable CX-4

Components of PARAM – Yuva



PARAM Yuva II

PARAM Yuva II

Upgraded to 524 Tera Flops in the same electrical power envelope using Intel's SandyBridge (Xeon - E5-2670) processor

Intel Knights Landing (KNL) Many Integrated Core (MIC) as the accelarator

Total number of cores, in 288 nodes, 30,056

Connected to NKN network

The Present

Top 10 Systems in India

Rank	System	Rpeak (TF/S)	Rmax(TF/S)
1	Indian Institute of Trop Met's Prathyush	3763.9	4006.19
2	NCMRWF's Mihir	2570.4	2808.7
3	SERC IISc's Sahasrat	901.51	1244
4	Indian Institute of Trop Met's Aaditya	719.2	790.7
5	TIFR Indian Lattice Gauge Theory Initiative's ColourBoson	558.7	730
6	IIT Delhi	524.4	861.74
7	CDAC's PARAM Yuva II	388.44	520.4
8	IIT Bombay's SpaceTime	384.83	620.544
9	CSIR 4th Paradigm Institute	334.38	362.09
10	NCMRWF	318.4	350.1

Source: topsc.cdacb.in

System Description

#1 Prathyush at IITM, Pune

- Used for Ocean-Atmosphere Climate System Research
- Cray XC-40 class system with 3315 CPU nodes
- Cray Aries interconnect
- 10.686 PB Lustre Parallel file system, 30 PB archive

#2 Mihir at NCMRWF, Noida

- Used for Monsoon Mission, Climate Change Research
- Cray XC-40 class system with 2322 CPU nodes
- Cray Aries interconnect
- 7.639 PB Lustre Parallel file system, 21 PB archive

System Description (Contd.)

#3 Sahasrat at SERC, IISc, Bangalore

- Cray XC-40 class system with 1468 CPU nodes, 48 Intel Xeon Phi nodes and 44 K-40 GPU nodes
- Cray Aries interconnect
- 12.1PB Lustre Parallel file system

#4 Aaditya at IITM, Pune

- IBM iDataPlex system with 2384 CPU nodes
- Infiniband FDR10 interconnect
- 6 PB GPFS Parallel file system, 1 PB archive

System Description (Contd.)

#5 ColourBosom at ILGTI, TIFR, Hyderabad

- Cray XC-30 class system with 467 CPU nodes with one K-20x GPU
- Cray Aries interconnect
- 1.1PB Lustre Parallel file system

#6 PARAM Yuva - II at CDAC, Pune

- SuperMicro system with 288 CPU nodes
- PARAMNet and Infiniband HDR interconnect
- 100 TB Parallel file system, 100 TB home area

BRAF at C-DAC, Pune

- Sequence Analysis
- Molecular Modelling
- Ab-initio methods
- Microarray Data Analysis
- Genome Alignment
- Repeat finding tools
- RNA analysis algorithm

Biogene

Biocrome System

https://www.cdac.in/index.aspx?id=hpc_sa_braf_access

PARAM Shavak

PARAM Shavak - Supercomputing Solution

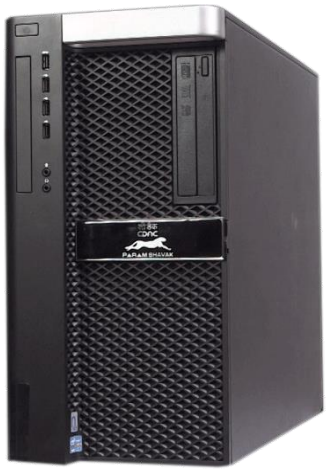
- Supercomputing in a box solution - affordable and pre-installed supercomputing systems which can be readily put to utilization among users with moderate needs

- Over 3TF from 36 Intel Xeon CPU cores, 7TF (DP* 1xGV100 for HPC) and 9TF (SP* 1xP5000 for DL)

- Is an initiative to empower scientists and researchers with state-of-the-art facilities for their cutting-edge research

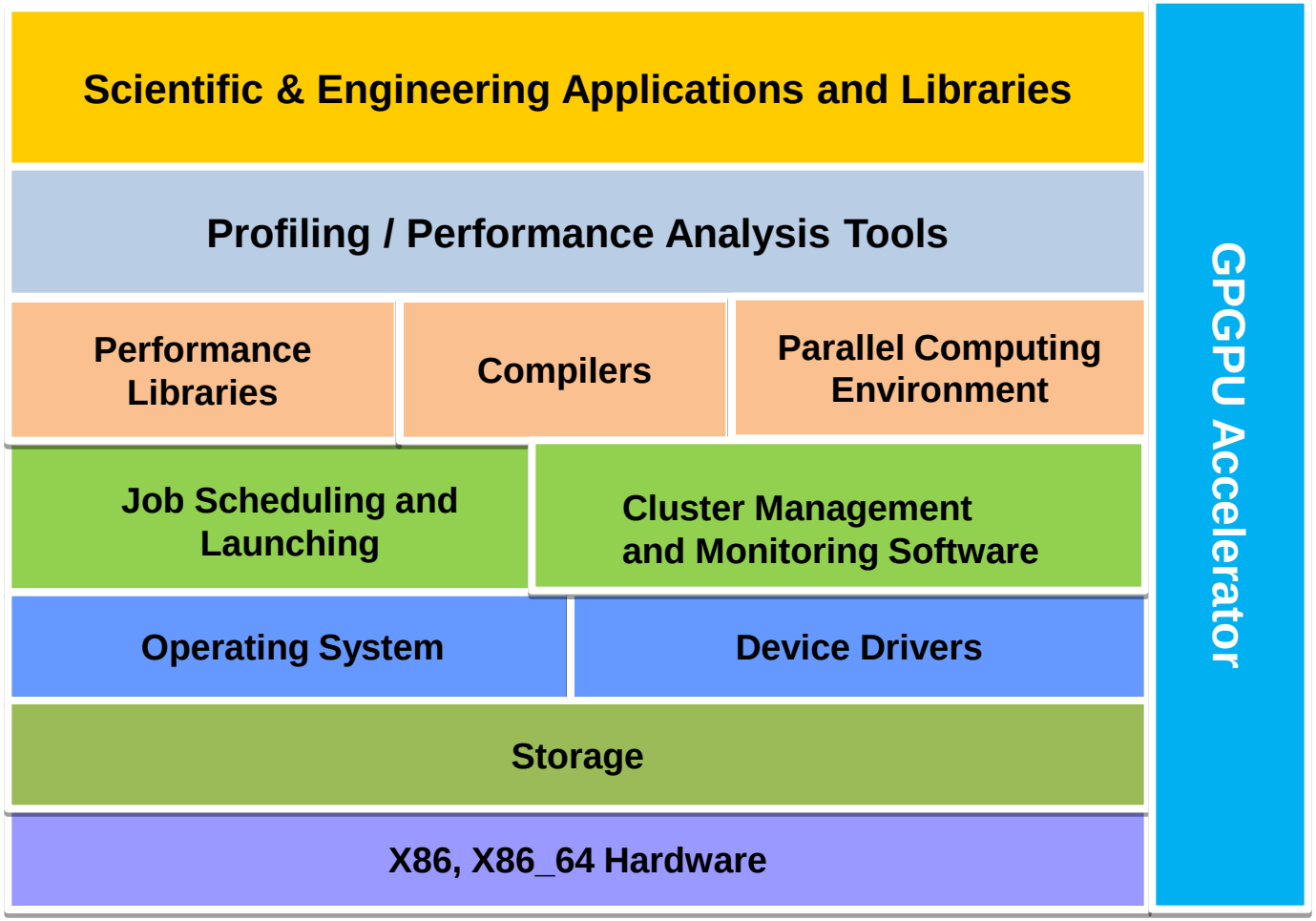
- Provides a skill development platform for academia

- Solutions for three type of technological challenges
  - High Performance Computing (HPC)
  - Deep Learning (DL)
  - Virtual Reality (VR)



*DP/SP : Double/Single Precision

Hardware and Software Stack



Some Applications from various domains

Bio-informatics : mpiBLAST

Molecular Dynamics : GROMACS, NAMD, LAMMPS

Materials Science : Quantum Espresso

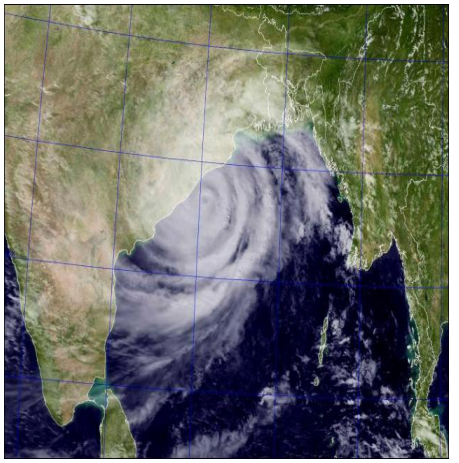
Quantum Chemistry : NWChem, ABINIT

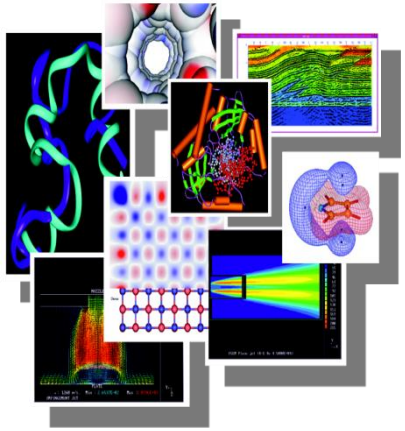
Atmospheric & Ocean Modeling : WRF, MOM

CFD : OpenFOAM

Libraries & Visualization Tools

Compilers, Libraries and Performance Analysis Tools: open source scientific & mathematical libraries, tools and Intel Parallel Studio





Indigenous Technologies in PARAM

- PARAMNet, a low latency high bandwidth network

- Reconfigurable Computing, FPGA based accelarator

PARAMNet

PARAMNet Over Years

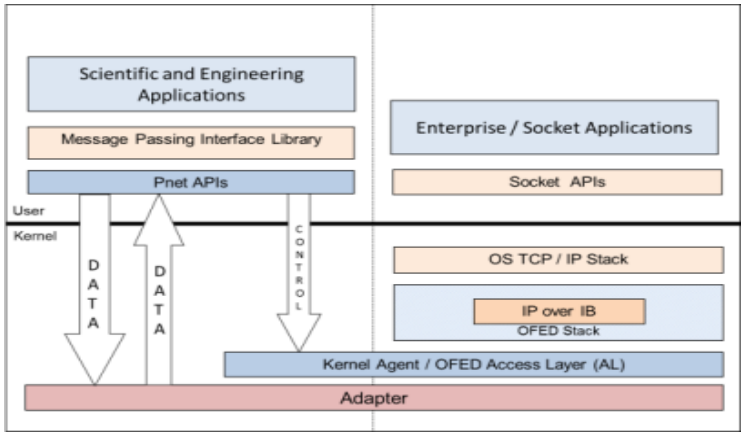
Network	Data Speed	Co-Processor	Switch	Software Stack	Primary Deployment
PARAMNet-I (beta)	100 Mbits/sec	CCP-I	8-port, 100 MBits/sec, Copper (DS Link)	CDAC Proprietary	PARAM 9000 (1997)
PARAMNet-I	400 Mbits/sec	CCP-II	8-port 400 Mbits/sec Copper (DS Link)	CDAC Proprietary Fast Sockets	PARAM 10000 (1998)
PARAMNet-II	2.5 Gbits/sec	CCP-III “Neo”	16-port 2.5 Gbps, full duplex, Fiber	VIPL compliant	PARAM Padma (2003)
PARAMNet-3	10 Gbits/sec	Gemini	48-port 10 Gbps/sec, full duplex Copper (CX-4)	DAPL SDP TCP/IP compliant	PARAM Yuva/Yuva-II (2008-2013)

Network Ecosystem Development

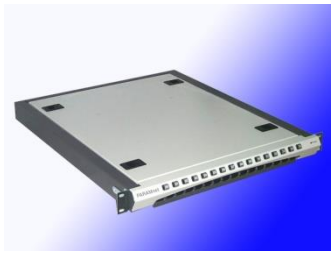
Silicon

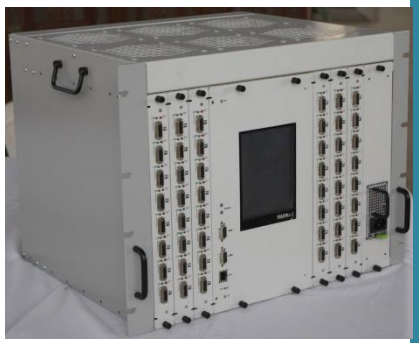
Interface Cards

LWP Stack



Switches





Reconfigurable Computing

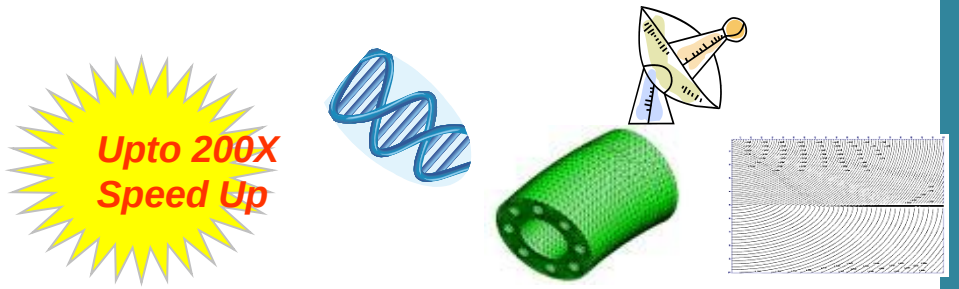
Reconfigurable Computing System

"Indigenous FPGA based *Energy and Space efficient Application Accelerators*"

- **RC Hardware**
  - Multi million gate FPGA for application
  - Small form factor
  - Low power consumption of 25W
- **"Avatars"- Hardware library**
  - Hardware library developed for the critical portions of the application
- **"Varada"- Programming environment**

Application accelerated in domains:

- **Life science**
- **Fracture mechanics**
- **Astrophysics**



Reconfigurable Computing

The Future

National Supercomputing Mission

About the NSM

- Government of India initiated 7 year project with an outlay of Rs. 4500 Crores
- Implementing agencies : C-DAC and IISc
- Activities
  - Infrastructure Development
  - Application Development
  - R&D for Exascale Computing
  - HR Development

Scope of the Mission

Number of supercomputing systems of different sizes and scales with total computer power to the tune of about 50 Petaflops

- Entry Level Supercomputers
- Midrange Supercomputers
- Large Supercomputers

Building National Supercomputing Grid by interconnecting various HPC systems over NKN

Developing supercomputing applications through collaborations

HPC manpower development

HPC Research and Development leading to Exascale computing readiness

NSM: Make In India

Build Approach

- Optimized architecture and system design
- Collaboration with technology partners to design subsystems
- Most of the subsystems are envisaged to be manufactured in India
- Customization of open source software

Goals

- Develop and enhance Design capability
- Increase manufacturing capabilities
- Catalyze the high end and complex electronics product design and manufacturing in the country

NSM Application Development

At least 5 applications of national relevance to be developed and deployed

Areas include but not limited to

- 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

NSM: R&D

Developments for building state-of-the-art HPC machines

Programming paradigm for Exascale

Scalable Algorithm developments

Power optimization technologies

Exploring next generation Data centre technologies

R&D for high speed network scalable for Exascale supercomputers

Special purpose machines – e.g. for Bio-informatics

HR Development Activities

- Generate enough human resources that can take-up and spearhead supercomputing activities in the country
- Development of highly professional HPC-aware human resource pool
  - For HPC applications development
  - For managing, monitoring and running complex HPC systems
- Manpower needs to be trained at all levels
  - Undergraduate
  - PhD and Masters
  - PG Diploma
- 20,000 strong manpower to be developed in 7 years

Needs to be addressed

- A large number of Universities with good representation and distribution across the country must adopt Parallel Computing as a core course.
- A good number of the students should also obtain working experience by working in the NSM projects
- NSM should also help facilitate job opportunities in India for parallel and supercomputing.
- Manpower needs to trained in
  - HPC related concepts
  - Languages, Tools
  - Hardware and network development

Systems under NSM

Installed

- PARAM Shivay at IIT (BHU) (~ 0.8PF)
- PARAM Shakti at IIT Kharagpur (~ 1.6PF)
- PARAM Brahma at IISER Pune (~0.8PF)

Upcoming

- IIT Kanpur (~ 1.6PF)
- IIT Hyderabad (~0.8PF)
- JNCASR, Bangalore (~0.8PF)

Highlights from Nov'19 Top500 List

- The top 10 systems are unchanged from the previous list.
- The aggregate performance of the 500 systems now is at 1.65 exaflops.
- The entry level is 1.14 petaflops, up from 1.02 petaflops
- Summit at Oak Ridge National Laboratory with HPL result of 148.6 PFs and Sierra with 94.6 PFs at Lawrence Livermore National Laboratory occupy top 2 positions
  - Both use IBM's Power9 CPUs, NVIDIA Tesla V100 GPUs
- No. 3 is the Sunway TaihuLight supercomputer, with 93.0 PFs.
- Tianhe-2A (Milky Way-2A), at No. 4 with 61.4 petaflops
- Frontera, a Dell C6420 system at TACC is at #5 with 23.5 PFs

Highlights from Nov'19 Top500 List

- The most powerful new supercomputer on the list is AiMOS, at no. 24, with 8.0 petaflops. Uses Power9 CPUs and V100 GPUs.
- The number of TOP500 installations in China continues to rise and now sits at 227, up from 219 six months ago.
- Share of US-based system at 118.
- Japan remains in third place with 29, followed by France with 18, Germany with 16, the Netherlands with 15, Ireland with 14, and the United Kingdom with 11.
- All other countries are in the single digits.

Highlights from Nov'19 Top500 List

- At the CPU level, Intel's processors are present in 470 of the 500 systems, split between multiple generations of Xeon and Xeon Phi hardware.
- IBM is second with 14 systems – 10 with Power CPUs and four with Blue Gene/PowerPC CPUs.
- AMD claims just three systems on the current list.
- Two Arm-based supercomputers on the list: the Astra system
- NVIDIA is the dominant vendor for accelerators. Its GPUs are present in 136 of the 145 accelerated systems.

Highlights from Nov'19 Top500 List

- Ethernet is used in 52 percent (258) of the TOP500 systems,
- InfiniBand is the network-of-choice in 28 percent (140) of systems
- Greenest computer, A64FX prototype delivered 16.9 gigaflops/watt.
- NA-1, a Zettascaler machine with PEZY-SC2 processors and delivers 16.3 gigaflops/watt

- (source: www.top500.org)

Thank You

ashishk@cdac.in