Azadi Ka Digital Mahotsav

Exa-scale Vision of India

Dr. Hemant Darbari
Ex-Director General, C-DAC
darbari@cdac.in

Organization of the Talk

- Exascale Computing Vision
- Why Exascale?
- Importance for India (Grand Challenges)
- Technological Roadmap
- PARAM SHANKH-
  ‘Exascale Supercomputing through Harmony of Innovations’
- Looking Ahead…
- Conclusions

Acknowledgments:

- C-DAC
- Intel, NVIDIA, ARM, IBM, AMD
- Atos, Mellanox

Exa-scale Computing Vision

Our Vision is to Establish Dependable and Secure Exa-Scale Eco-system with innovative designs, disruptive technologies and Expert Human resource.

Our Mission:

- to develop Hardware encompassing Exascale Chip Design, Design and Manufacture of Exascale Server Boards, Exascale Interconnects and Storage including Silicon-Photonics;

- to Exploit huge opportunities in System Software and High-end Software development to make India a Global Leader.

- To develop solutions for Grand Challenge Problems for the National/International Level

Why Exascale ?

Building Exascale Capable System is the key enabler for accelerating scientific discovery in the world. There is a strong co-relation between research output and availability of HPC systems.

Scaling from Tera(10^{12}) to Peta (10^{15}) to Exascale (10^{18}) computing is the key to futuristic research problems and providing state-of-the-art solutions to grand challenge problems. Each of these solutions will have a wide spread societal impact.

Investment in Exascale can give ROI of 1:500 as per authentic studies

Why Exascale ?

India, through our major initiatives in National Supercomputing Mission and Microprocessor Development Program has made significant progress towards indigenous realization of supercomputers.

Be it:

- **Servers,**
- **High Speed Network Interconnects,**
- **Storage Systems and**
- **HPC Microprocessors**

And now we are geared up to realize totally indigenous Exascale Supercomputer called PARAM SHANKH under Atmanirbhar Bharat Mission.

This will spawn 300-400 MSMEs, Industries and Start-ups

Grand Challenge Problems requiring Exascale Computing

- Numerical Weather Prediction (Monsoon Mission & Climate Modeling)
- Materials research (Design of new materials)
- New periodic table (cluster of atoms?)

- Faster than real simulation (Electrical Grid Failure & Rail Accidents)
- National security (Real Time analysis of satellite images)
- Grand challenge problems in mathematics (crypto – analysis)

- Creating energy efficient enzymes (Conversion of solar to useful energy)
- CFD (New insights into Internal Combustion Engine, Space-crafts)
- Ab-INITIO Molecular Dynamics

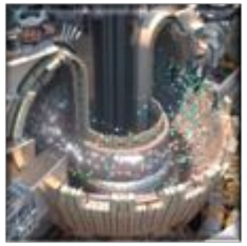
AND …….. A LOT MORE !!!

This will lead India towards Technological Supremacy in the Global arena

CONVERGED MODELS APPLIED TO GRAND CHALLENGES

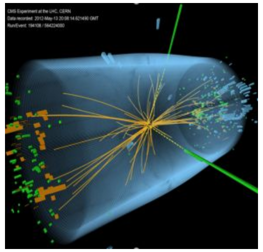
Commercially viable fusion energy

>95% Prediction Accuracy for Fusion Disruptions with 60ms warning



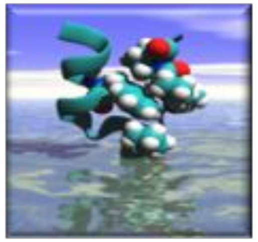
Improve/Validate the Standard Model of Physics

Improve FastSim by 1,000,000x
CERN Large Hadron Collider



Clinically viable precision medicine

300,000X Faster Predict Molecular Energetics for Drug Discovery

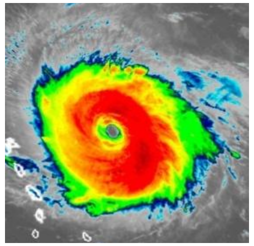


Understanding cosmological dark energy and matter

Model the origins of the universe in 30ms

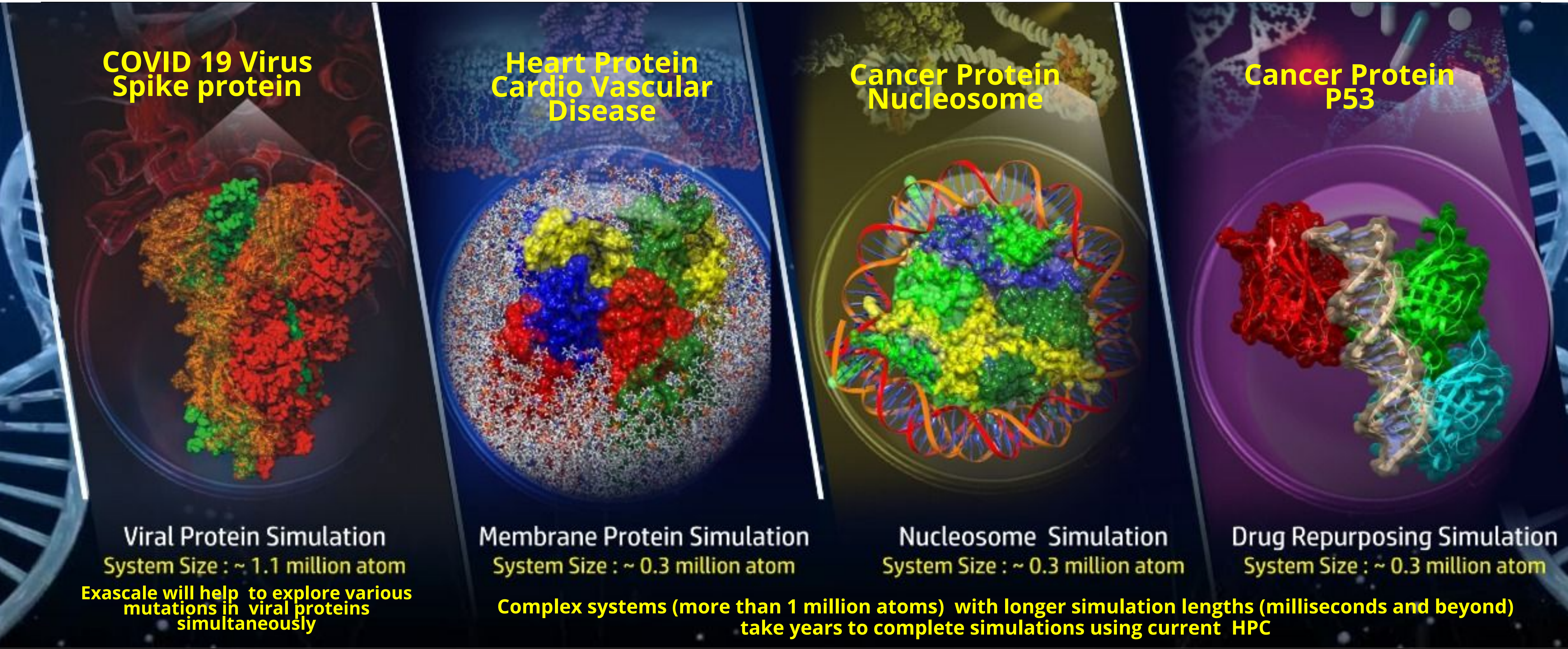
Climate/weather forecasts with ultra- high fidelity

Improve Climate and Weather models to sub 1km resolution



India is a Major Contributor in all these Mega Science Projects

Typical Application of Exascale Computing

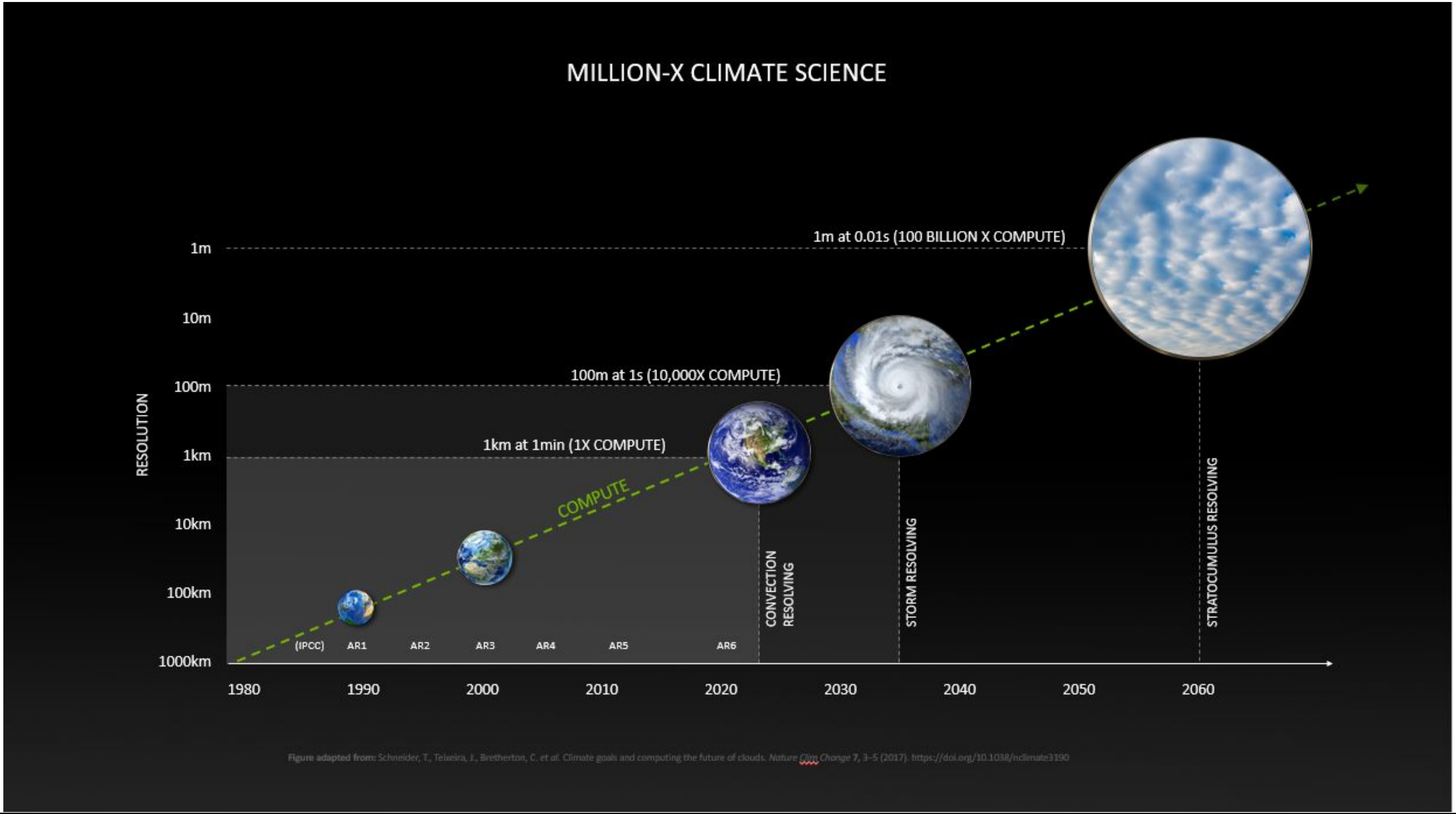


<u>1 Microsecond Simulation Length for a system of 1 million atoms</u>

100 TF: 3 months

Exascale: approx. 15 days

Typical Application of Exascale Computing



MODELING THE ORIGINS OF THE UNIVERSE IN LESS THAN 1 sec

Challenge

Simulation of the origins of the universe and large scale galaxy simulation is typically done with N-Body gravitational simulations that require 10**6 Node hours to run **(India is contributing in LIGO-Mega Science Project)**

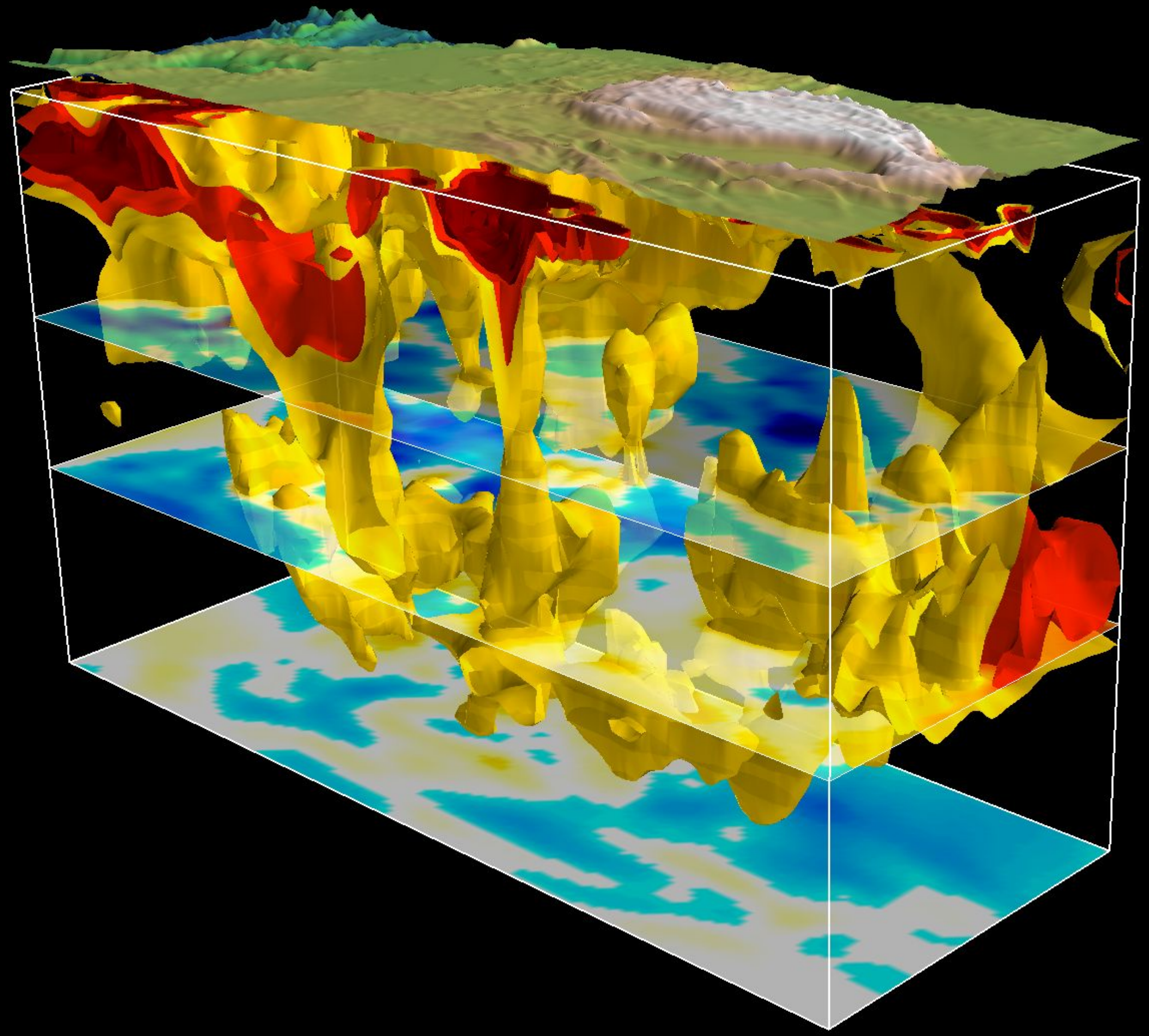
Solution

A custom DNN (D3M) based on the U:NET CNN, originally developed for biomedical microscopy

Impact

D3M was shown to generate High Res images in 30 ms

Early results show that the model correctly generated images with Dark Matter content greater than original training data based on prior science thinking



PLANNING AND WARNING MODELS FOR EARTHQUAKES

Challenge

The models to allow city planning to minimize the damage potential that can result from earthquakes can take months on supercomputers

Accurate and timely models to be used for prediction and early detection are currently intractable

Solution

2019 Gordon Bell Nominees developed a custom RNN trained with simulation data, and then used to replace the pre-conditioner in the solver

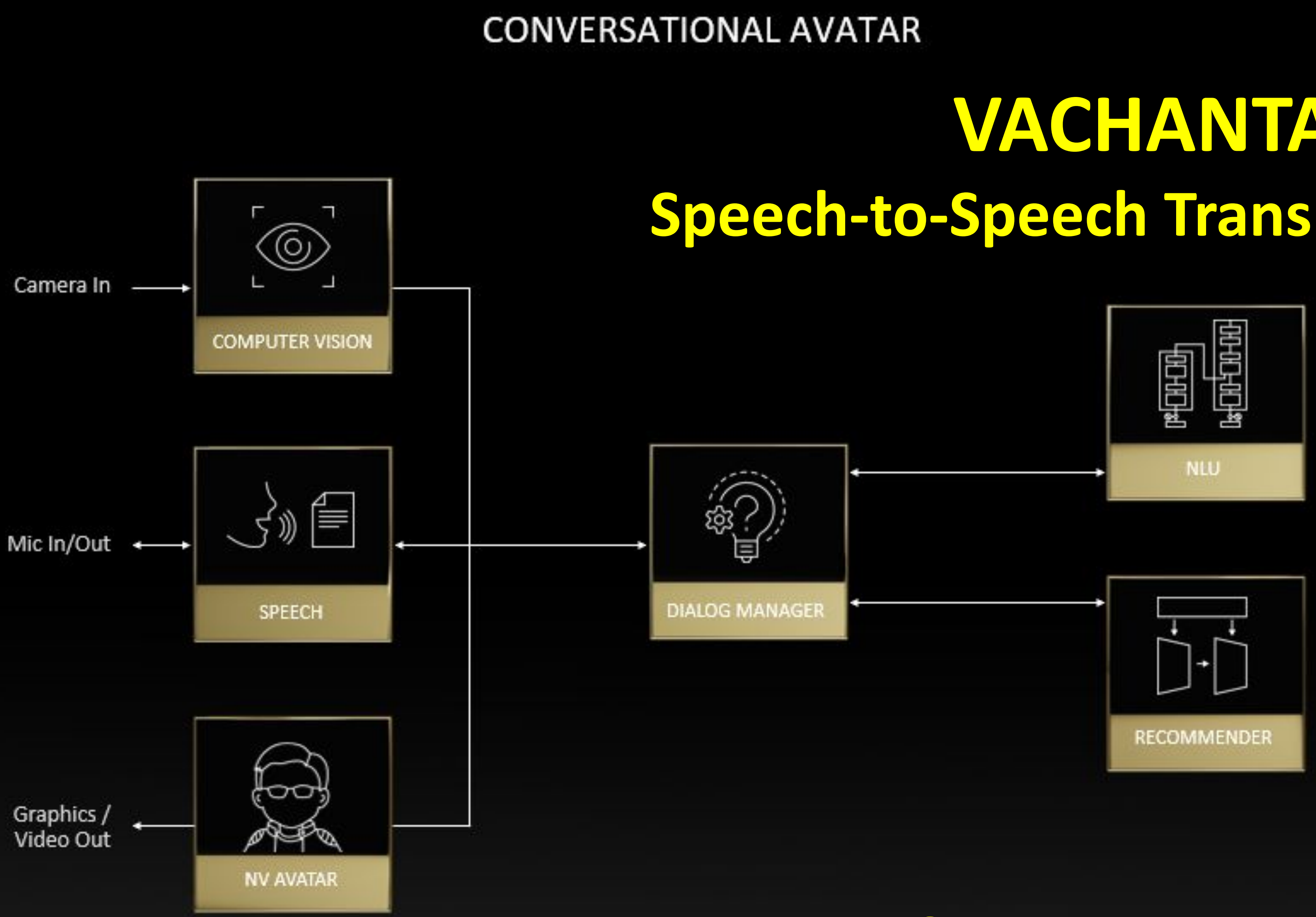
Mixed precision algorithms were carefully chosen to replace double precision in selected parts of the iterative solver

Impact

The overall speed up of 25X at super high resolution with a model size of 75Bn elements and 302Bn degrees of freedom on the SUMMIT system at ORNL

VACHANTAR

Speech-to-Speech Translation System



National Language Translation Mission

Department of
Science & Technology (DST),
Government of India

Ministry of Electronics &
Information Technology (MeitY),
Government of India

Indian Institute of Science,
Bengaluru

सी-डैक
CDAC
Center for Development of
Advanced Computing

National Supercomputing Mission (NSM)

One Vision One Goal...Advanced Computing for Human advancement...

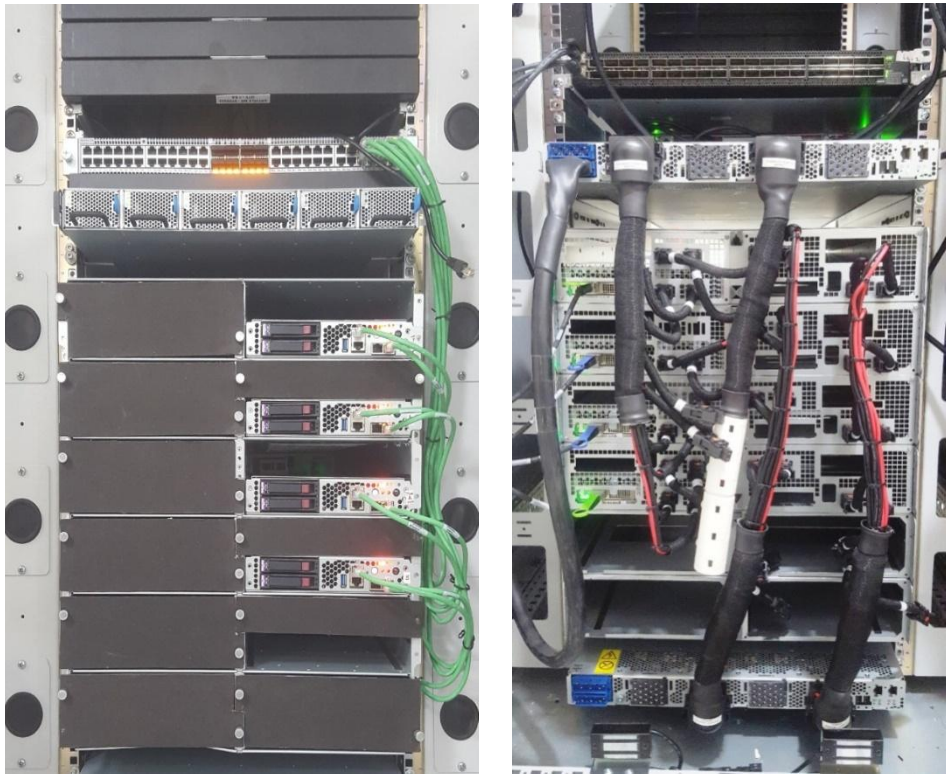
Indigenously designed and manufactured Rudra Board and 4 Node Cluster

Rudra : Mother board

- Open 19 form factor
- Centralized redundant power supply
- 4 Compute Nodes
  Each Node
  - 2 x 6240R (24 cores 2.4 Ghz 165W) CPUs
  - 192 GB DDR4
  - 10 Gbps Ethernet
  - 2 x 800GB NVME SSD
  - HDR 100 Mellanox NIC
  - 3.675 TF Peak (68% Rmax)
- 40 port HDR 200 Switch
- 14.75 TF Peak (64.1% Rmax)

Indigenous HPC interconnect (Trinetra)

40 Gbps x 6 = 240 Gbps and 100 Gbps x 6 = 600 Gbps



Rudra 4 Node Cluster

Supercomputers across the Nation

Overall 28,11,535 HPC Jobs Executed on NSM Systems

- PARAM Shivay Supercomputing facility at IIT (BHU), Varanasi (Dedicated to nation by PM on February 19, 2019)
- PARAM Brahma Supercomputing facility at IISER Pune (Visit by PM on December 08, 2019)
- PARAM Shakti Supercomputing facility IIT Kharagpur (Fully Operational)

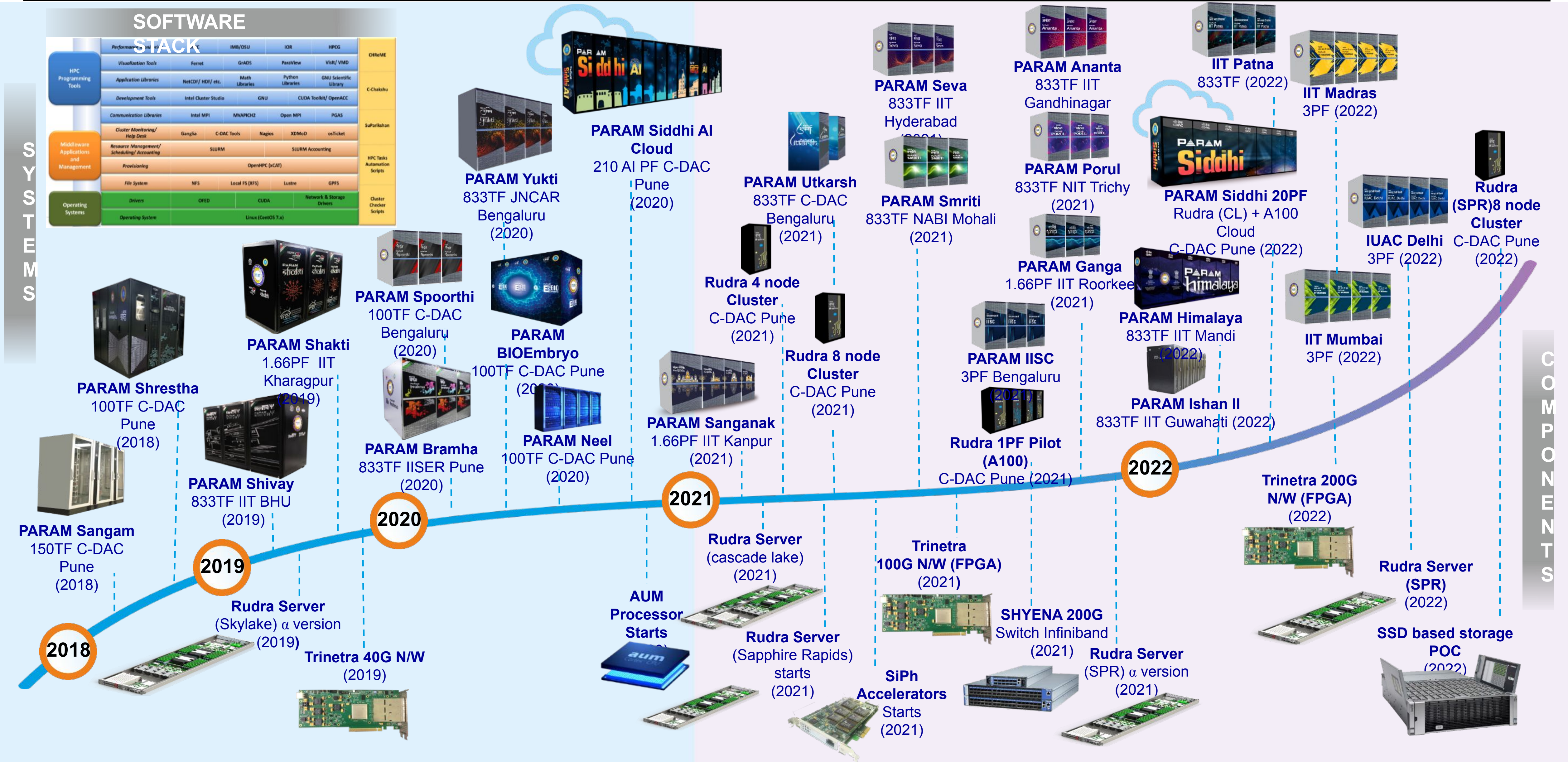
3,80,851 HPC Jobs Executed
1.3 PF (1.66 PF)

6,26,231 HPC Jobs Executed
650 TF (833 TF)

17,50,798 HPC Jobs Executed
650 TF (797 TF)

More than 11000 Supercomputing manpower/faculty trained

National Supercomputing Mission (NSM) Indigenous Components and Systems



One Vision. One Goal... Advanced Computing for Human Advancement...

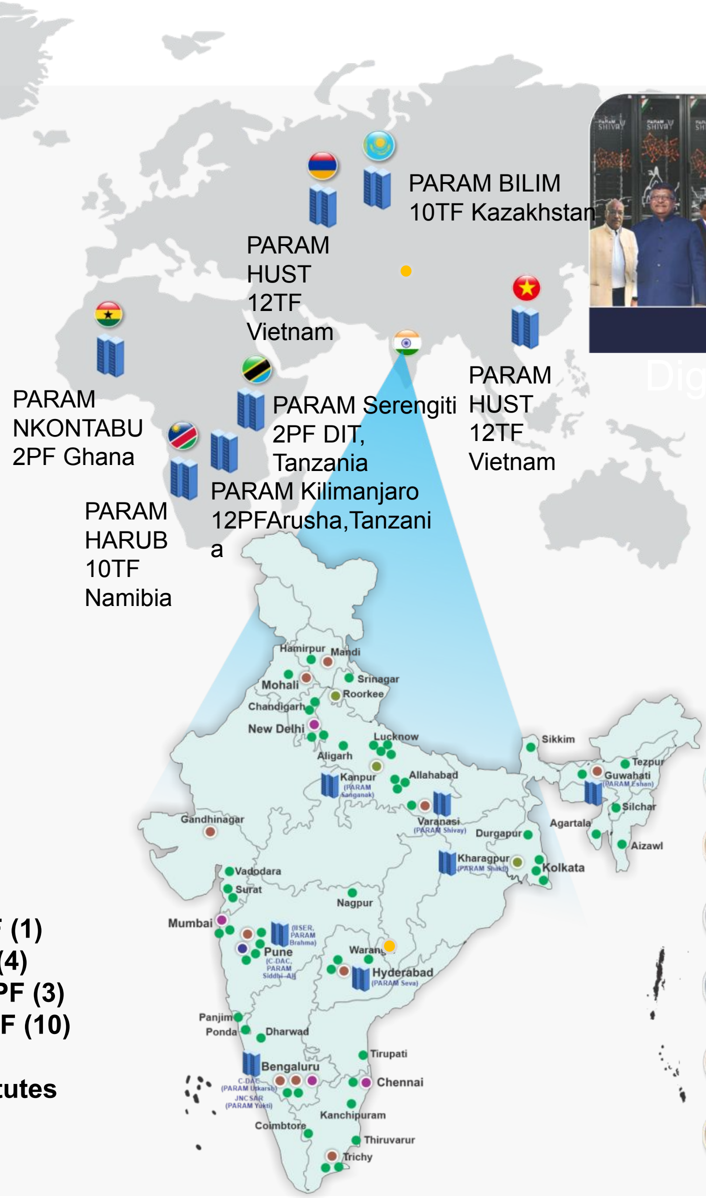
International and National Super Computing System

collaborations and consortia

Research Institutions: 30+

Academia: 70+

Industries & Start-ups: 25+



- **20PF**
  Pune
- **3PF**
  IISc, Bangalore
  IIT, Bombay
  IIT, Madras
  IUAC, Delhi
- **1.66PF**
  IIT, Kharagpur
  IIT Kanpur
  IIT, Roorkee
- **833TF**
  IIT BHU, Varanasi
  IISER, Pune
  IIT, Hyderabad
  JNCSAR, Bangalore
  IIT, Guwahati
  NABI, Mohali
  IIT, Gandhinagar
  NIT, Trichy
  C-DAC, Banaglore
  IIT, Mandi

- **User Institutes**

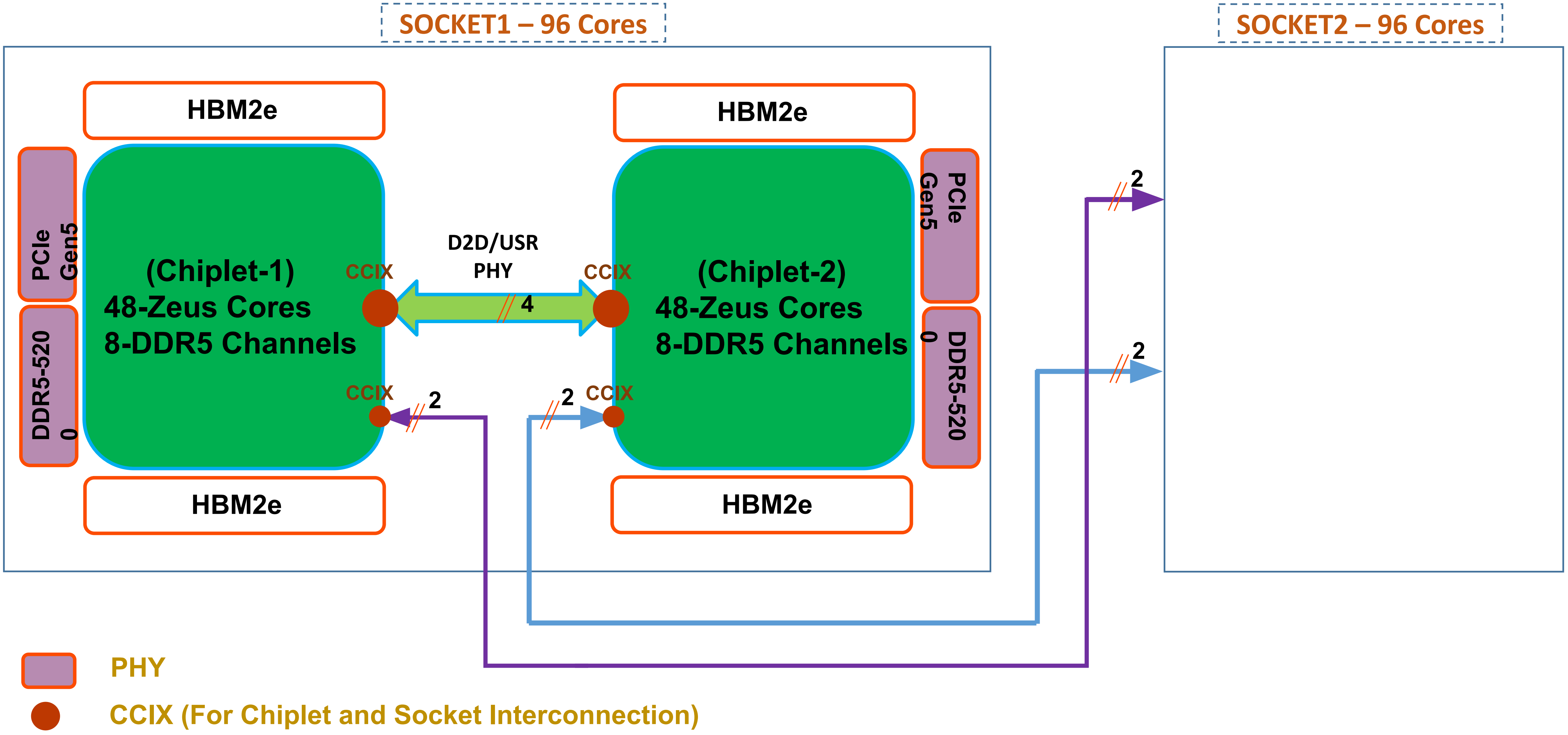
IIT-Dharwad
Goa University
4PI Bangalore
VJTI-Mumbai
Mumbai University
IIT-Goa
NIT-Goa
MNNIT-Allahabad
Banaras Hindu University-Varanasi
IIIT-Allahabad
IIIT-Lucknow
University of Allahabad
NCL-Pune
COEP
NCCS
NCRA
IUCAA

IIIT-Guwahati
NIT-Silchar
Tezpur University
NIT-Agartala
NIT-Sikkim
NIT- Meghalaya
NIT-Mizoram
NIT-Manipur
SVNIT-Surat
IIIT-Vadodara
IIIT-Surat
Lucknow University
NBRI-Lucknow
Aligargh Muslim University
IIIT-Lucknow
NIT-Hamirpur
NIT-Uttarakhand
Pune University
VNIT-Nagpur

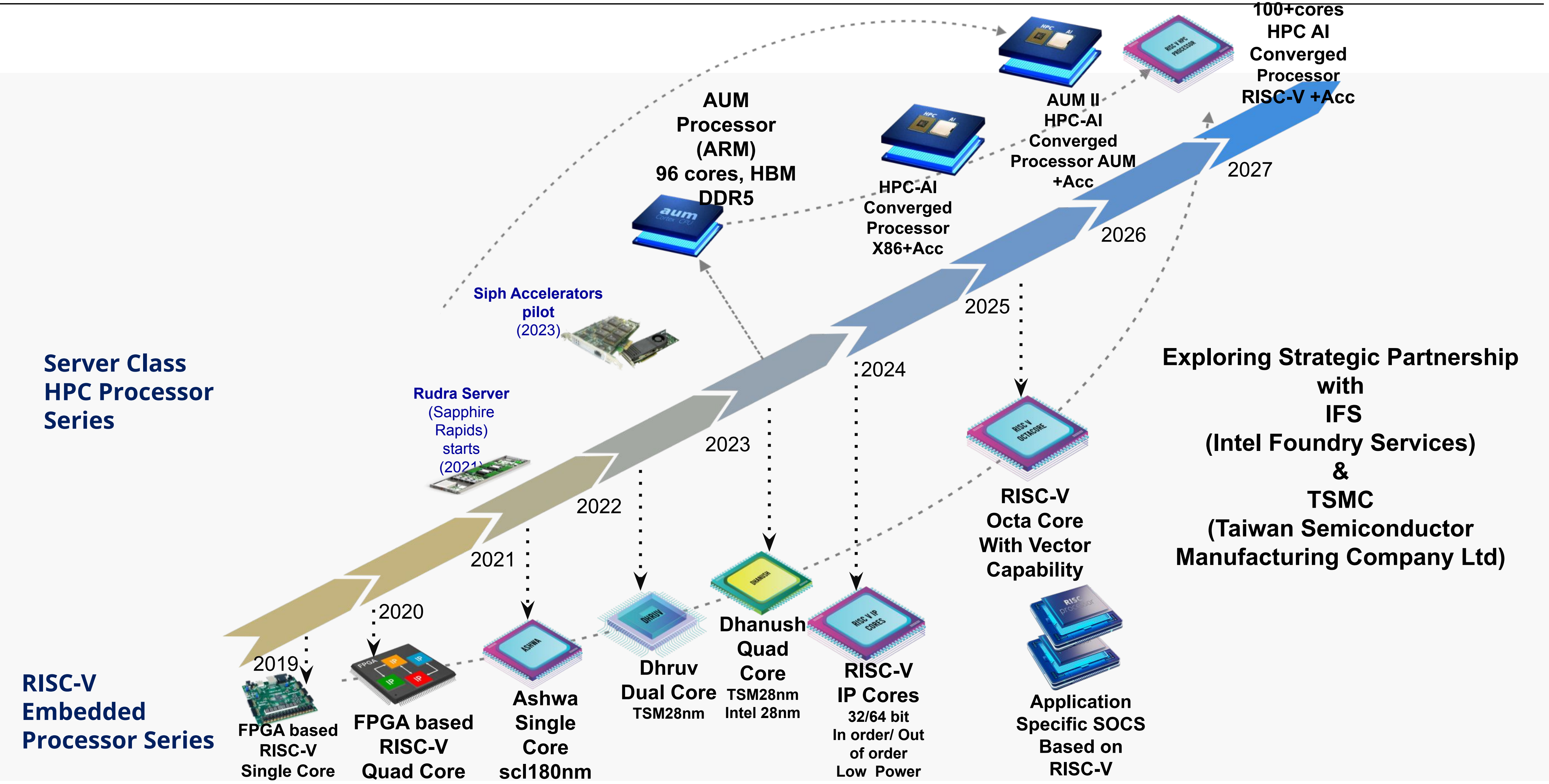
- **20PF (1)**
- **3PF (4)**
- **1.66PF (3)**
- **833TF (10)**
- **User Institutes**

Basic Science
Earth Science
Engineering Applications
Space Engineering
Defense & Atomic Energy Applications
Data Sciences

Indigenous HPC System on Chip (SoC) Configuration



Roadmap of C-DAC Microprocessors



C-DAC -Exascale Computing mission

परम शंख

PARAM Shankh

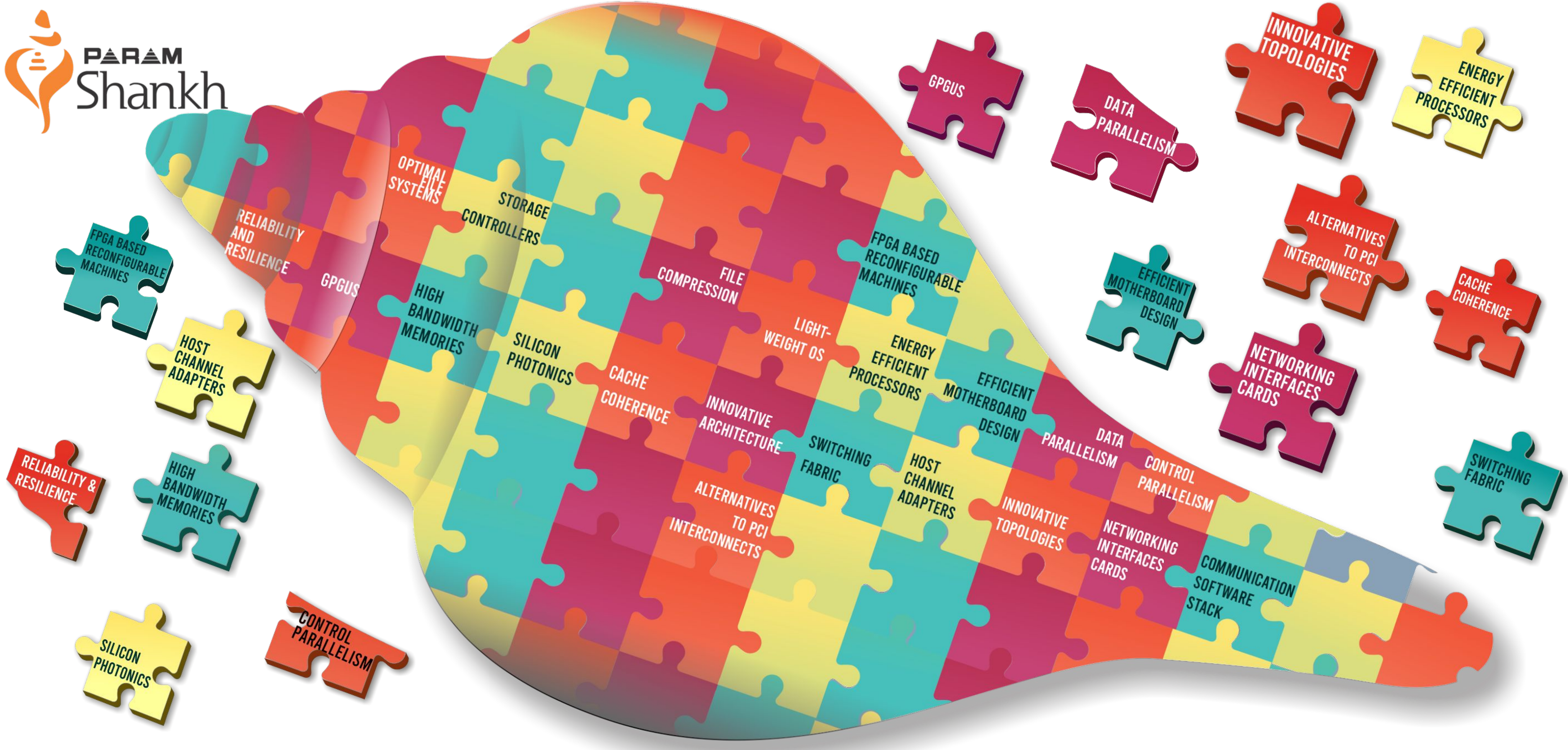
Sl. No.	Number	In Power (10 raise to)	Name in Ancient Indian System	English Equvalent
1	10	1	Das	Ten
2	100	2	Sau	Hundred
3	1000	3	Hazar (Sahastra)	Thousand
4	10000	4	Das Hazar(Sahastra)	Ten Thousand
5	100000	5	Lakh	Hundred Thousand
6	1000000	6	Das Lakh	Million
7	10000000	7	Crore	Ten Million
8	100000000	8	Das Crore	Hundred Million
9	1000000000	9	Arab	Billion
10	10000000000	10	Das Arab	
11	100000000000	11	Kharab	
12	1000000000000	12	Das Kharab	Trillion
13	10000000000000	13	Neel	
14	100000000000000	14	Das Neel	
15	1000000000000000	15	Padm	Quadrillion
16	10000000000000000	16	Maha Padm	
17	**100000000000000000**	**17**	**Shankh**	
18	1000000000000000000	18	Maha Shankh	Quintillion
19	10000000000000000000	19	Antya	
20	100000000000000000000	20	Maha Antya	
21	1000000000000000000000	21	Madhya	Sextillion
22	10000000000000000000000	22	Maha Madhya	
23	100000000000000000000000	23	Paradh	
24	1000000000000000000000000	24	Maha Parad	Septillion
25	10000000000000000000000000	25	Dhoon	
26	100000000000000000000000000	26	Maha Dhoon	
27	1000000000000000000000000000	27	Ashohini	Octillion
28	10000000000000000000000000000	28	Maha Ashohini	

"Exascale Supercomputing through Harmony of Innovations"

Description of Shankh

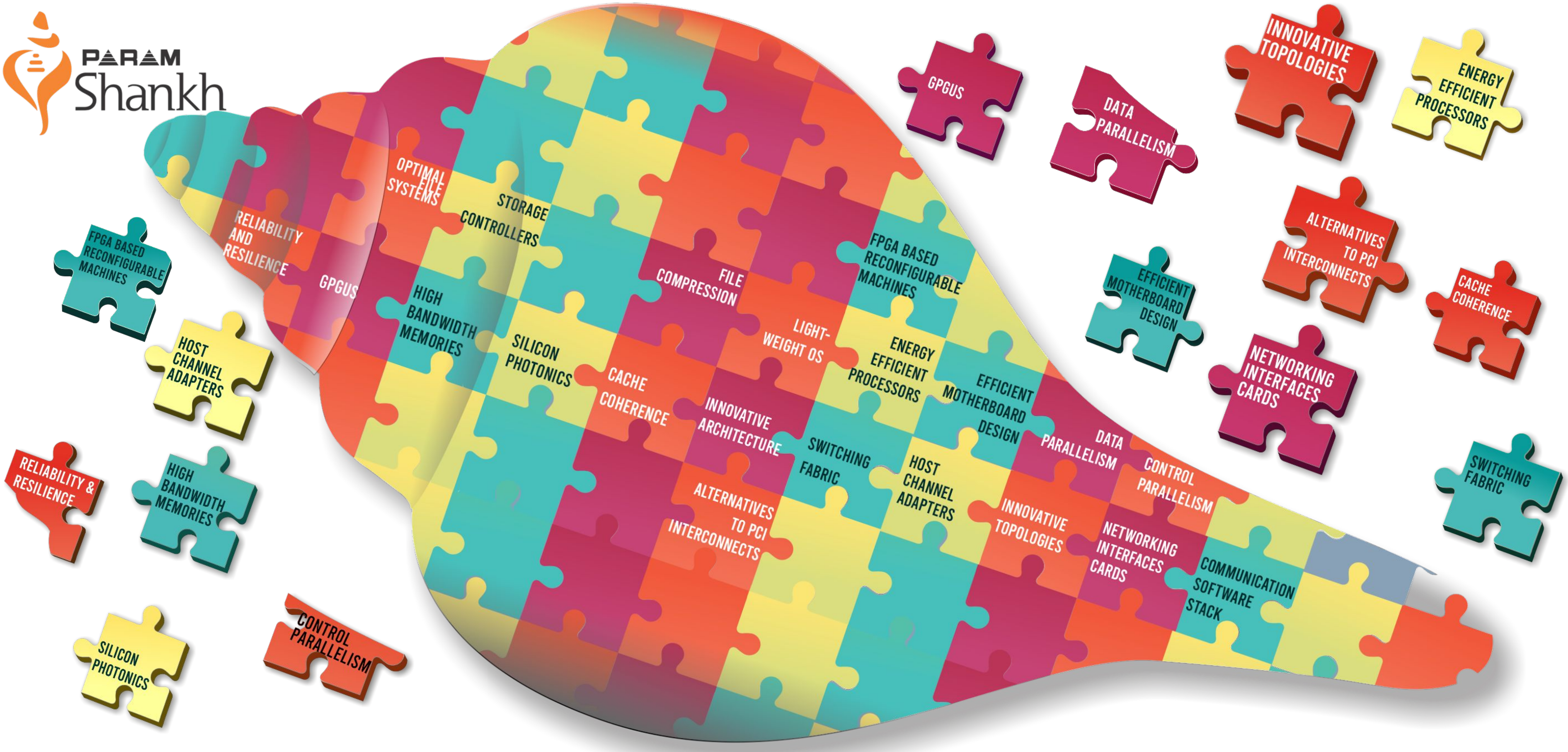
Sl. No.	Number	In Power (10 raise to)	Name in Ancient Indian System	English Equvalent
1	10	1	Das	Ten
2	100	2	Sau	Hundred
3	1000	3	Hazar (Sahastra)	Thousand
4	10000	4	Das Hazar(Sahastra)	Ten Thousand
5	100000	5	Lakh	Hundred Thousand
6	1000000	6	Das Lakh	Million
7	10000000	7	Crore	Ten Million
8	100000000	8	Das Crore	Hundred Million
9	1000000000	9	Arab	Billion
10	10000000000	10	Das Arab	
11	100000000000	11	Kharab	
12	1000000000000	12	Das Kharab	Trillion
13	10000000000000	13	Neel	
14	100000000000000	14	Das Neel	
15	1000000000000000	15	Padm	Quadrillion
16	10000000000000000	16	Maha Padm	
17	**100000000000000000**	**17**	**Shankh**	
18	1000000000000000000	18	Maha Shankh	Quintillion
19	10000000000000000000	19	Antya	
20	100000000000000000000	20	Maha Antya	
21	1000000000000000000000	21	Madhya	Sextillion
22	10000000000000000000000	22	Maha Madhya	
23	100000000000000000000000	23	Paradh	
24	1000000000000000000000000	24	Maha Parad	Septillion
25	10000000000000000000000000	25	Dhoon	
26	100000000000000000000000000	26	Maha Dhoon	
27	1000000000000000000000000000	27	Ashohini	Octillion
28	10000000000000000000000000000	28	Maha Ashohini	

Solving the Complex jigsaw of Exascale



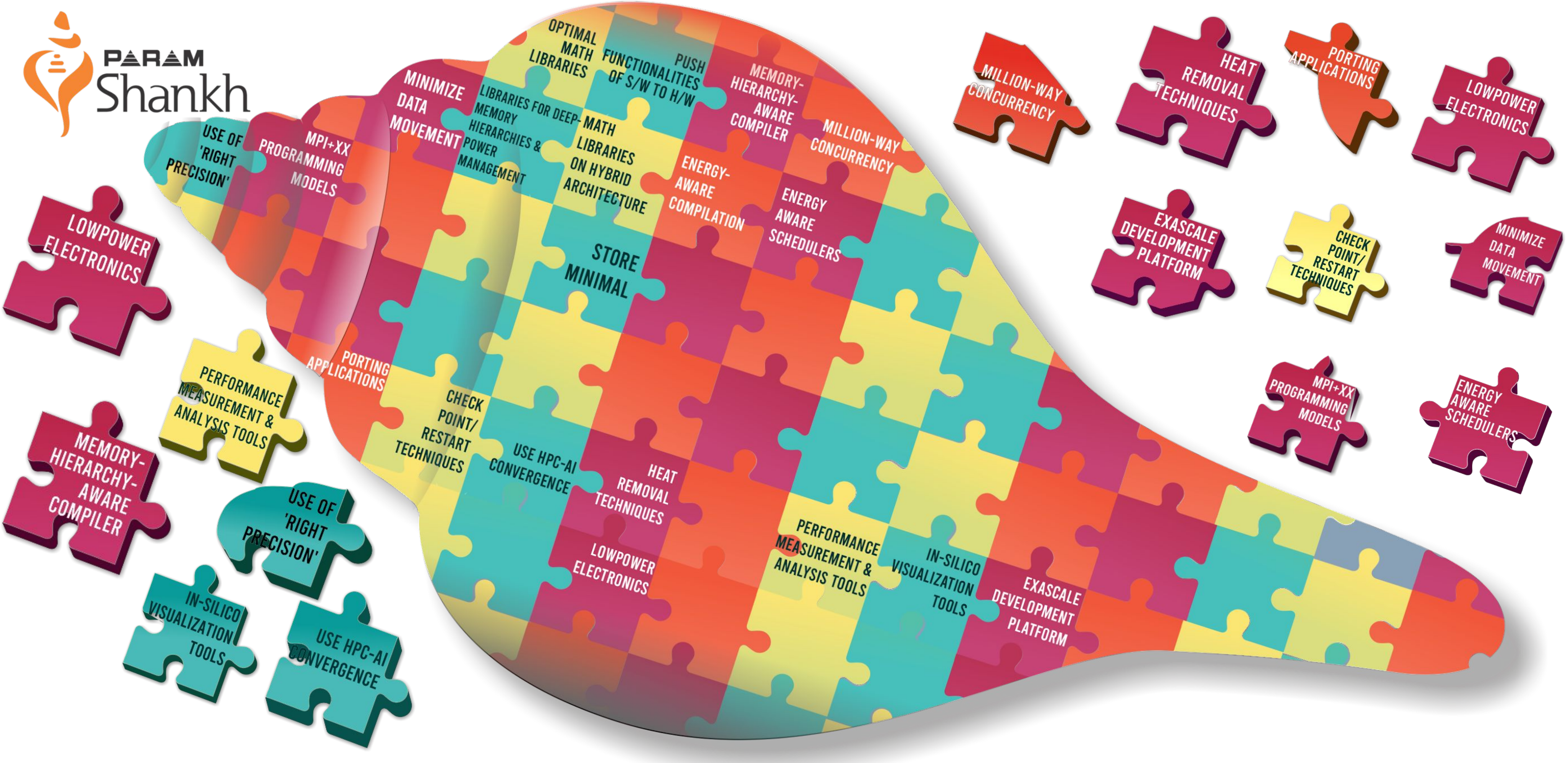
Solving the Complex jigsaw of Exascale

PARAM Shankh



Solving the Complex jigsaw of Exascale

PARAM Shankh



Research opportunities to be exploited:

Generic areas:

1.Cache Coherence:

- Cache Coherence is a way of providing architectural support for efficient shared memory.
- Idea is to maintain uniformity of data in multiple local caches.
- Very interesting problem is to find out what is “right data” in the context of parallel processing.
- Innovations are needed in designing new cache-coherent protocols and exploiting Cache Coherence for extremely parallel architecture.

2.Silicon Photonics:

- Concept of manufacturing silicon integrated circuits embedded with photonic elements.
- Communication using electrons through copper medium has large limitations.
- Communication by photons uses very less energy and also has less issues to be addressed.
- What percentage of components can be photonics based?
- Increasing this percentage is indeed an active area of research.

Research opportunities to be exploited:

Generic areas:

3.High Bandwidth Memories:

- Data movement between CPU and Memory results in latency and power consumption issues
- Heat removal is another issue to be death with
- These issues cause a “Memory Wall”.
- Why not move logic and memory closer in an advanced packaging?
- How best this can be achieved?
- Innovation is to find the right mix of various kind of memories to achieve the desired results in a given budget.

4.GPGPUs:

- GPUs with Novel architectures.
- Innovate tools to facilitate mapping of legacy codes to GPGPU architecture.
- Optimal mix of CPUs and GPUs.

Research opportunities to be exploited:

Generic areas:

5.Reliability and Resilience:

- It is obvious that when millions of components are involved, there could be some failures.
- The trick is to build resilience against failure by providing redundancy.
- Explore innovative methods/sensors to detect failures and seamlessly migrate the processes on redundant resources.

Processor/s:

6.Energy efficient Processors:

- Reduce Thresholds of Voltage swings
- Use power gating.
- Innovate methods of reducing leakage currents.

Research opportunities to be exploited:

Processor/s:

7.FPGA based reconfigurable machines:

- These are used to get deterministic response in a given application/ process.
- Push software capabilities to hardware to make software faster.

8.Efficient Motherboard design:

- There is always a better design than the best design!! Go for it!!!

Research opportunities to be exploited:

Processor/s:

9.Data Parallelism:

- Exploit Vector Registers, Vector Instructions and Vector Processors.
- Simulate these if not available in hardware.
- Facilitate with right compilers to exploit data parallelism.

10.Control Parallelism:

- Develop algorithms to exploit control parallelism (rarely exploited till now)

11.Faster alternatives to PCI interconnects:

- PCI interconnects CPU with the rest of the peripherals
- Keeping pace with the increase in speeds of CPU and Networks is indeed important.
- Innovation is needed in finding newer materials to be used of PCBS, Novel connectors and ways & means of reducing cross talks.

Research opportunities to be exploited:

Interconnects:

12.Interconnect switching fabric:

- Detect and resolve congestion.
- Check possibility of over provisioning.
- Guarantee bandwidth to critical operations.
- Arrive at optimal message latency.

13.Host Channel Adapters:

- Use spare processing power to compress and De-compress data to improve bandwidth.

Research opportunities to be exploited:

Interconnects:

14.Innovative Topologies:

- Use right topology that is suited to the traffic flow in a given application. Lot of modelling/ Research work is needed in this area.

15.Network interfaces cards:

- Use processing power at Network Interface Card (NIC) end to compress and De-compress data to improve bandwidth!

Research opportunities to be exploited:

Interconnects:

16.Communication software stack:

- Innovate communication software to minimise inherent delays!
- Push the software stack functionalities into Hardware/ FPGA to make it fast.

I/O

17.File Compression:

- Application dependent file compression technique needs to be thought off.
- Store only what is needed for an application.
- Exploit patterns!

Research opportunities to be exploited:

I/O:

18.Design of storage technologies/ Controllers:

- Innovate Storage technologies that cater to both Volume and access speeds.
- Innovate robust controllers with redundancies!
- It is important to store provenance information.
- Can security of data at rest be embedded?

19.File systems (optimized I/Os):

- Innovate file systems that ensure faster response to a given application.
- Exploit workflows.
- Innovate new parallel file systems that can exploit processing power to compress and decompress data on the fly.

Research opportunities to be exploited:

Operating System:

20.Light Weight OS:

- As millions of cores run the O.S, the Operating system should as light in weight as possible. The traditional, bulky operating system that was processor centric should be replaced by an O.S. that does optimal data movement.

Architecture:

21.Innovative Architecture:

- It is important to develop an architecture that can seamlessly accommodate and interoperate with newer machines/ components by introduction of an interface. If scheduling of tasks can be done by an over-arching operating system, it will be good. Lot of innovations are needed.

Research opportunities to be exploited:

System Software:

22.Energy aware Schedulers:

- It is important to save on energy.

23.Handle million-way concurrency:

- Minimal data needs to be used for concurrency control.

24.Compiler to exploit hierarchy of Memory:

- Extremely important as memory with varying access speeds are being made available.

Research opportunities to be exploited:

System Software:

25.Compiler to generate energy-aware optimal code:

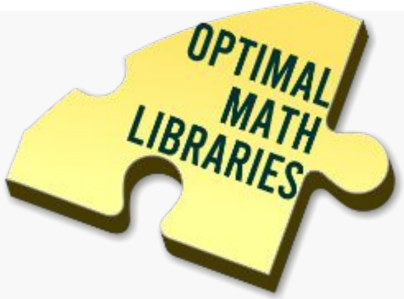
- The instructions generated must consume less energy. Alternate, equivalent instructions should be generated if a given in instruction is deemed to consume more energy.

26.Push functionalities of software to hardware:

- It is always a good idea to push the routinely used software into hardware to ensure order magnitudes of speedup.

Research opportunities to be exploited:

System Libraries:



27.Optimize Math Libraries for minimal data Movement:

- Hither to, most of the algorithms developed were CPU centric to ensure optimal utilization of CPU. With millions of cores, Processing power is not at premium, but data movement is. All functions will have to address this. Hence can be seen as a great challenge and huge opportunity.

28.Libraries for Deep-Memory Hierarchy & Power Management:

- Exploit the memory hierarchy and optimize on power that would be needed to execute the instructions generated by the Library.

29.Math libraries on hybrid architecture:

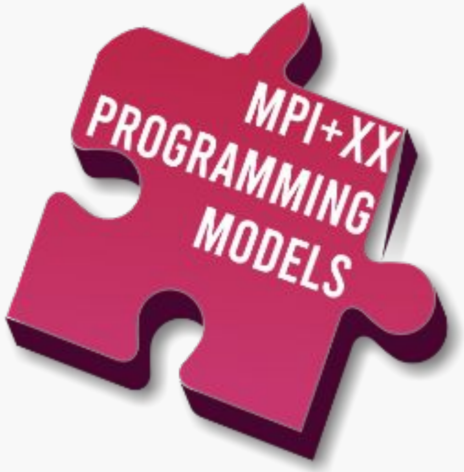
- Math libraries have been written to either exploit share memory architecture or distributed memory architecture. When the number of processors become too large, we will have to resort to Hybrid model to exploit locality and reduce on communication cost. This is yet another opportunity.

Research opportunities to be exploited:

Application Libraries:

30.Minimize data movement in Application S/W:

- Legacy algorithms and codes have been written to optimize CPU. Often this was done even if some data copying/ movement was involved. When these were ported to parallel architectures, they exhibited scalability issues. Now that CPU is available in abundance, it is worth re-writing the algorithms to optimize data movement and become scalable.



31.Enhanced MPI+xx programming models:

- MPI is compatible with Threading but heavy on memory.
- OPENMP is light on Memory.
- MPI+xx model promises scalability is Exa scale systems.
- The model can be more effective when interoperates with GPGPUs.

Research opportunities to be exploited:

Application Libraries:

32.Use of 'Right Precision' computation:

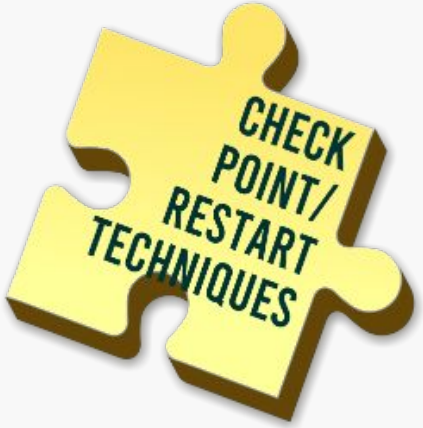
- Programming models used either single precision or double precision throughout a code based on the perceived requirement from an application. It is important to use right precision to various code segments which will execute faster and yet give results of desired accuracy. This will be more energy efficient as well.

33.Porting Applications to new architectures:

- Users are always reluctant to invest time in changing their codes to exploit newer architecture. Tools to facilitate porting to newer architectures should be developed with ease-of-use in mind.

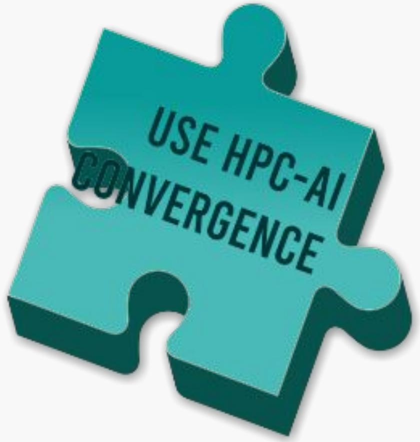
Research opportunities to be exploited:

Application Libraries:



34.Check point/ restart techniques:

- It is indeed important to consider the possibility of resources (core) gong bad and have built-in mechanisms to sense the failure and migrate the process to a redundant resource in a seamless manner and re-start.
- Further, the energy consumed by the Supercomputer is exceptionally large. In the event of a power-outage, one should be able to start from a point where one had left before the power outage.



35.USE HPC-AI convergence:

- It is a good idea to exploit patterns in each type of computation and to learn from data of previous executions.

Research opportunities to be exploited:

Monitoring Tools:

36.Parallel Performance Measurement & Analysis Tools:

- It is important to develop tools to facilitate performance measurement and analysis tools both at system level and application level. This will invariably enhance productivity.

37.In-situ visualization tools:

- These tools need to be developed to ensure fast response since the number of components in an exascale system is very huge.

Research opportunities to be exploited:

Development Platform:

38.Exascale development platform:

- User should be provided with a development platform where in, he/she can develop, debug, profile, execute and monitor his/her jobs.
- Perhaps an AI based tool can prompt him/her alternatives/ better methodologies.

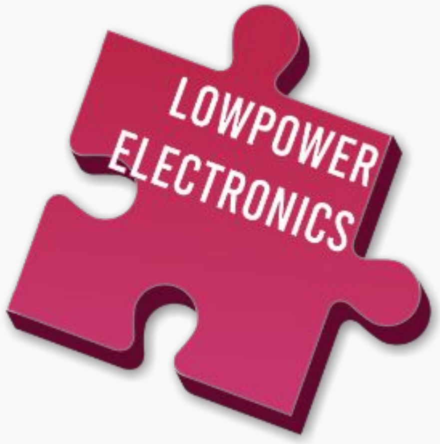
Heat removal research issues:

39.Innovations in heat removal techniques:

- Innovations in micro-fluidics.
- Innovations in Direct contact liquid cooling.
- Material research for efficient heat removal.
- Exploring Immersive Cooling Technologies.

Research opportunities to be exploited:

Heat removal research issues:



40.Adiabatic electronics:

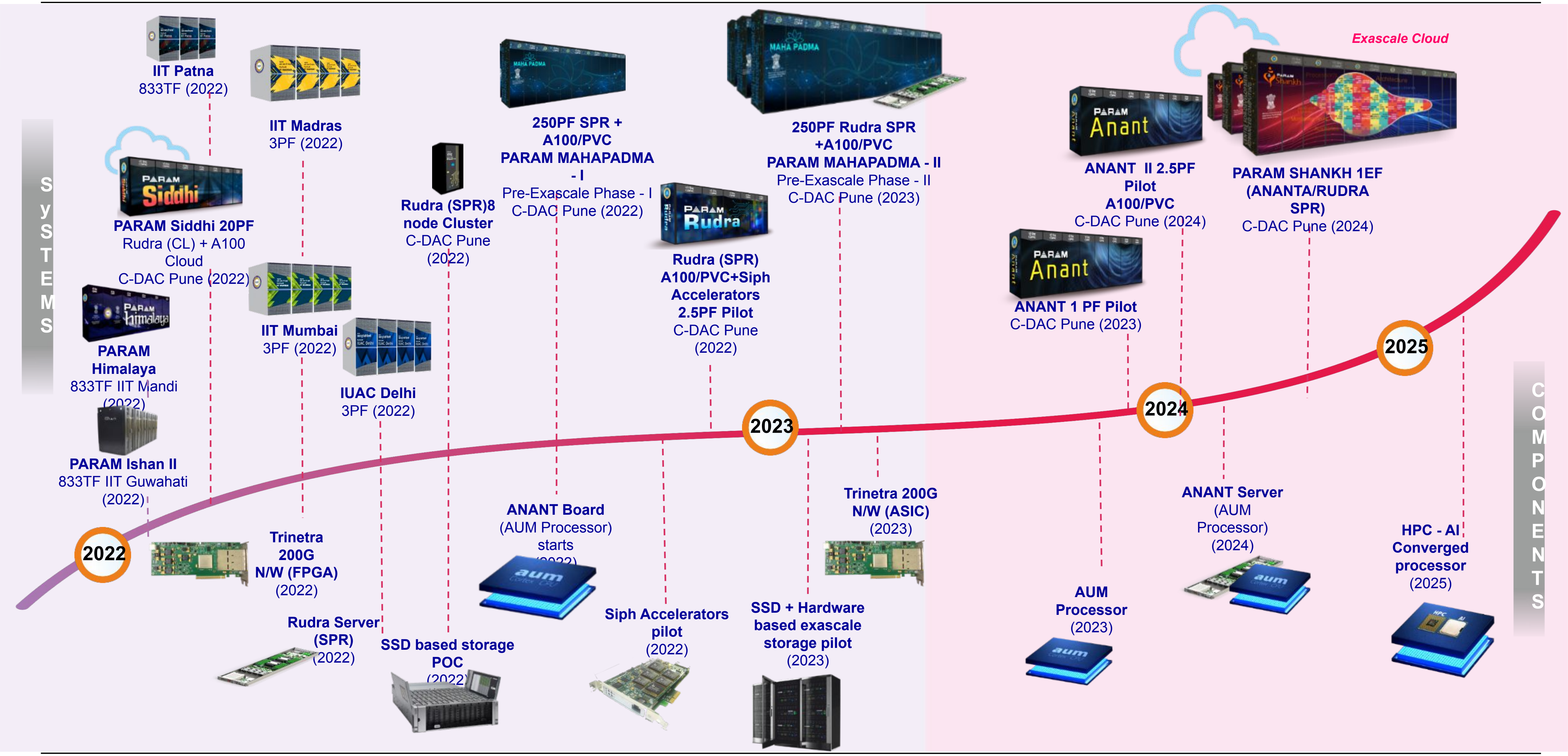
- Low power electronics that implement reversible logic
- Based on the concept that switching activities reduce power by giving back the stored energy to the supply
- Has huge potential for energy saving in switching and hence heat dissipation.
- Less heat dissipated, less will be the cost of heat removal.

C-DAC

Exascale Supercomputing through Harmony of Innovations

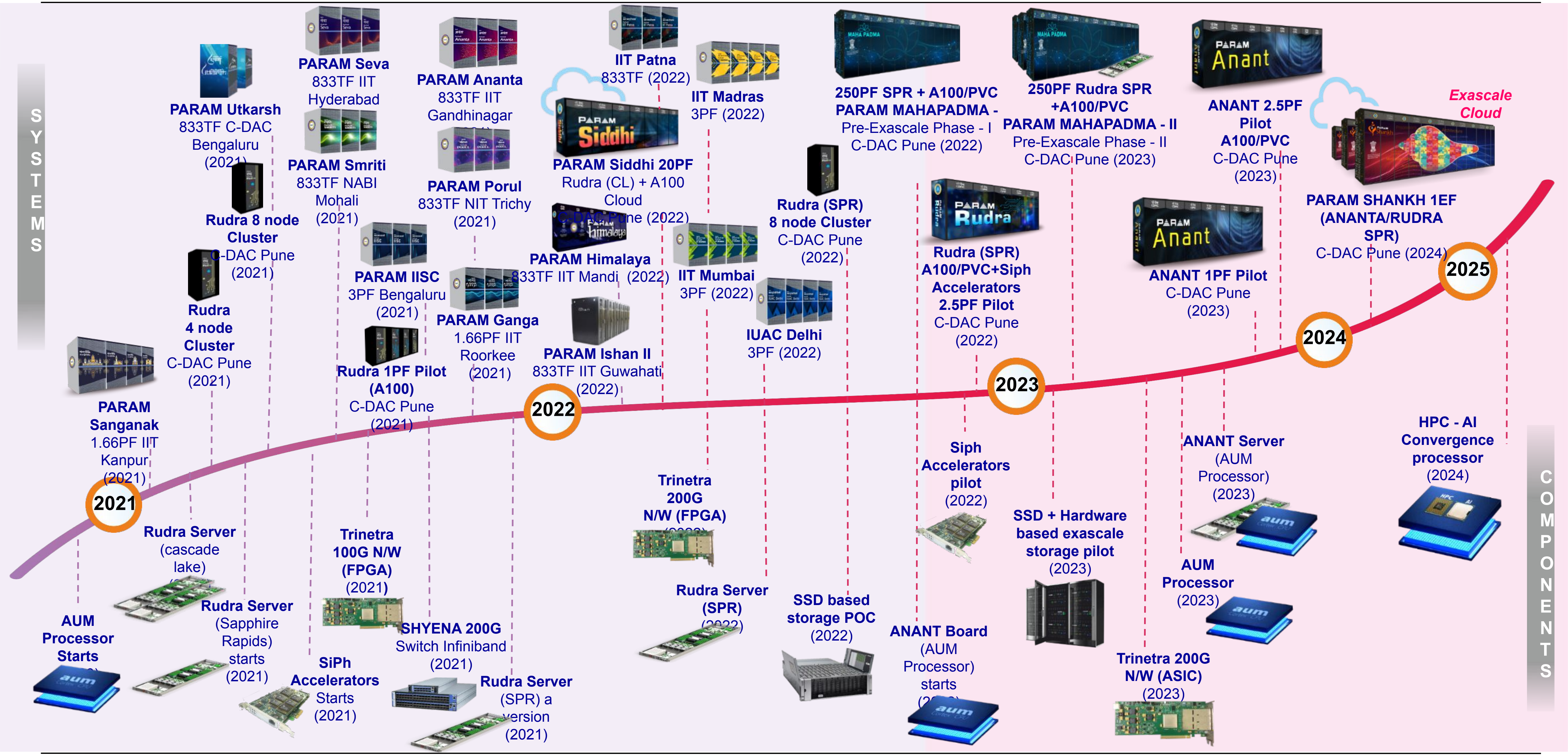
One Vision One Goal...Advanced Computing for Human advancement...

Roadmap Exascale Supercomputing Mission



One Vision. One Goal... Advanced Computing for Human Advancement...

Exascale Supercomputing through Harmony of Innovations



One Vision. One Goal... Advanced Computing for Human Advancement...

Global Initiatives

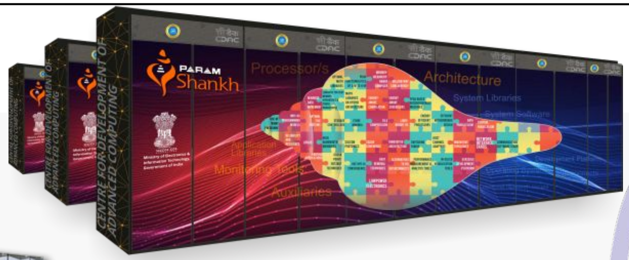
FRONTIER
Supercomputer,
1.0 Exaflops
ORNL, Tennessee, USA (2021)

AURORA
Supercomputer,
1.0 Exaflops
ANL, USA (2021)

Fugaku
Supercomputer
at Riken Lab, Japan (2021)

HPC5 - Eni Commercial HPC
Supercomputer,
1.0 Exaflops, EU (2023)

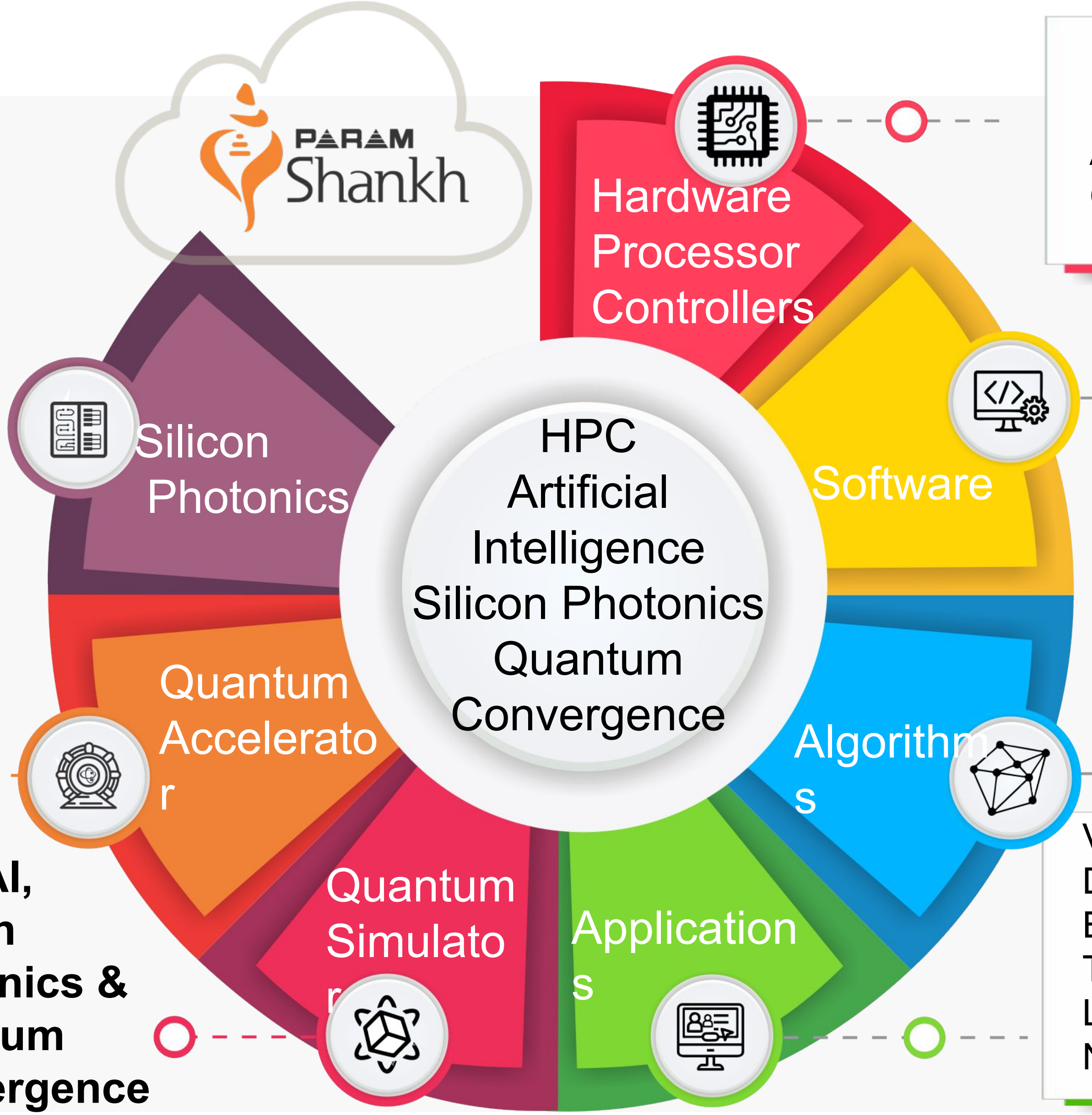
EL CAPITAN
Supercomputer,
1.0 Exaflops, LLNL, USA (2023)



PARAM SHANKH
Supercomputer,
1.0 Exaflops, C-DAC (2024)

One Vision. One Goal... Advanced Computing for Human Advancement...

Future Activities



Heterogeneous/ specialized Computer Architecture
Energy efficient, Fault tolerance Computer Architecture
C-DAC HPC-AI Converged Accelerator Development

OFED stack over Trinetra
Exscale Cluster Development Suite
Exscale Linux& Linux Distribution on Indigenous Exascale Processors

Programming Models for Exascale
Application Development Supplementing tradition MPI+X approaches
Multi-level adaptive routing algorithms for Hybrid Topologies

Visual Genomics for Molecular
Drug Repurposing
Ensemble based forecast techniques using HPC
Turbulent CFD Model of Transient flow for Engineering Apps
LSRTM applications for Advance Seismic Imaging
Next Generation Disaster Management Framework

HPC,AI, Silicon Photonics & Quantum Convergence

Conclusions

India will have full capability to Indigenously Design, Develop and Deploy Exascale Supercomputing Systems.

India will move from Service Economy to Knowledge Economy, where the gains are immeasurable in terms of Patents, Royalties and Innovations.

Finally, through National and International collaboration with Researchers, Academia, Industry and Start-Ups, India will emerge as a Leader rather than being a Follower.

Thank You !

Dr. Hemant Darbari
Ex-Director General, C-DAC
darbari@cdac.in