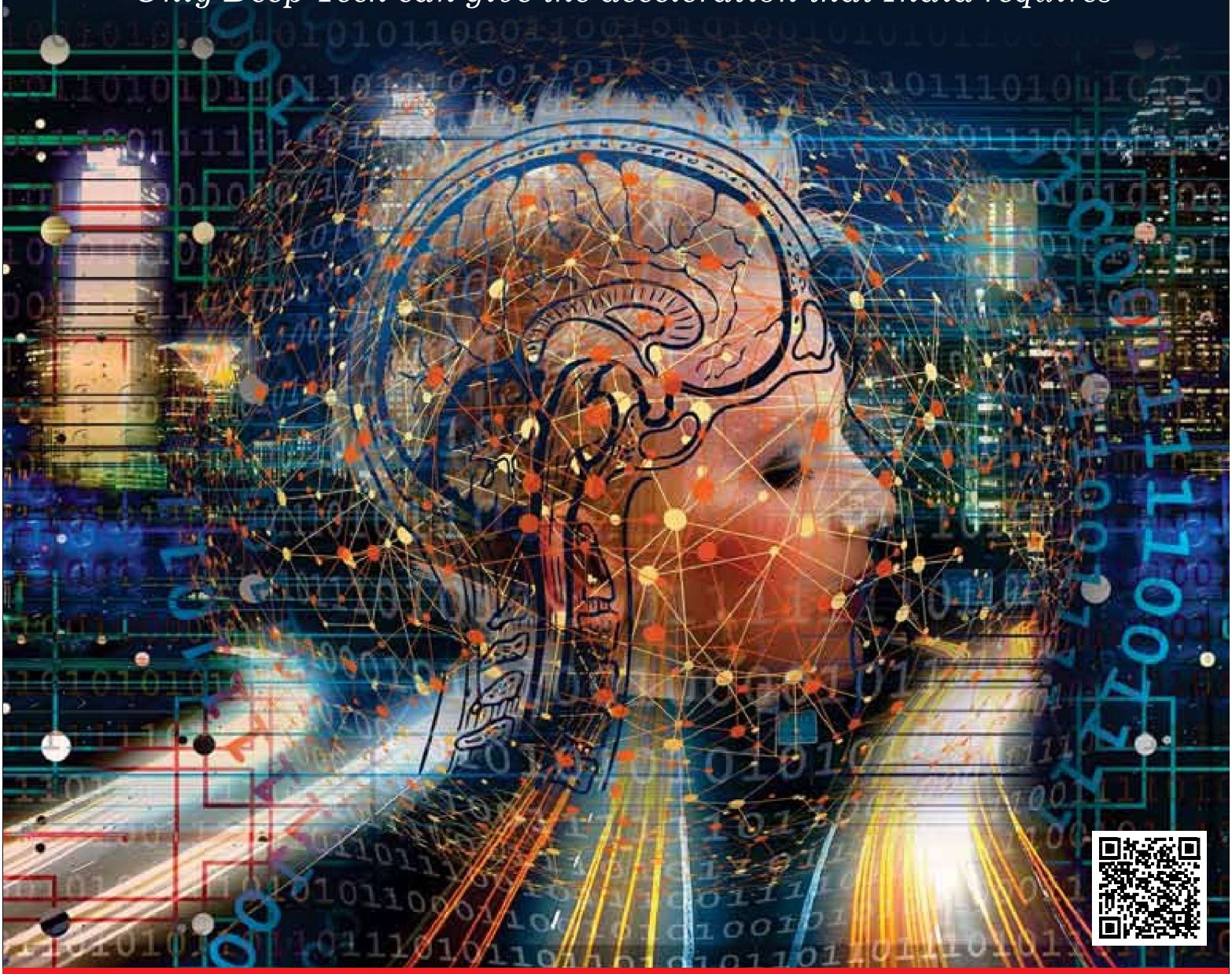
₹125
VOL. 34 | ISSUE 10 | October 2021
www.pcquest.com
PCQUEST
UNDERSTAND • CHOOSE • IMPLEMENT IT
CyberMedia
35TH YEAR
1987 - 2021

Time for Deep Techceleration

Only Deep Tech can give the acceleration that India requires



Special Subscription offer on page 54
76 pages including cover

Overcoming the hurdles in India's HPC path

Dr. Rupesh Nasre
pcquest@cybermedia.co.in

The National Supercomputing Mission (NSM) has given High Performance Computing (HPC) a much needed shot in the arm already. But there is still a long road ahead to democratize HPC

High Performance Computing (HPC) comes into play when individual machines fall short of solving a large problem. You can play games on a laptop, attend webinars from your mobile, and process spreadsheets on your desktops. But it is practically impossible to run a responsive gaming server on your laptop which is used by thousands of gamers worldwide, or to run the WhatsApp service with your mobile, or to host Microsoft or Google services on your desktop.

As users, you do not have to worry about these backend tasks. But if you are on the service-provider's side, bandwidth, throughput, latency, SLA, and performance become the key terms that guide your day-to-day job. If you run a company with thousands of customers, it is likely that you need to work with "high performance".

HPC is tightly linked with parallel computing. A typical computer program runs sequentially -- that is, it performs one task at a time. In contrast, in parallel computing, a program often performs multiple tasks at the same time, resulting in higher throughput (number of tasks completed in unit time). What is the scale of parallelism we are talking about here? A typical server may have a few tens of cores. Thus, this server can potentially execute a few tens of instructions simultaneously. If your server has an accelerator (graphics processing unit or GPU), the accelerator typically has thousands of cores.

Can I offer Google-like services if I have such a server with the associated accelerator? No. You need thousands of such servers to be able to address the scale. In HPC, at least a few hundreds of such servers work together to solve a given problem. If they have to work together, they need to share data. This is done by a fast interconnect between these servers -- whose top speed can be 100 times that of your best download speed on your desktop.

But are there problems which really require such a high computing power? In India, the government spends a large amount of money on weather modeling. This can be very helpful to predict monsoon, earthquakes, cyclones, etc. We receive a huge amount of data from our over 300 satellites (e.g., INSAT), which

DR. RUPESH NASRE, Faculty member, CSE Dept, IIT Madras | Head, National Supercomputing Mission Nodal Centre

quickly cross the memory limits of a single server computer. Therefore, not only that each server must have a large amount of RAM (over 100 gigabytes), but also a very large amount of hard-disk space (in petabytes; 1 petabyte = 1000 terabytes and 1 terabyte = 1000 gigabytes). Similarly, in healthcare, we need such a setup to find out the effects of various drugs, including vaccines. In the automobile industry, HPC helps in simulating various mechanical phenomena such as car collisions, which are immensely helpful in safety-critical applications such as airbags. HPC helps in simulating several natural phenomena used in materials engineering, aerospace, chemistry, physics, ocean engineering, mechanical sciences, nanomaterials, as well as biology.

As a specific example, several scientific problems can be modeled using a system of linear equations, to be solved using matrices. This often demands matrix operations such as matrix-vector product or matrix-matrix multiplication or matrix inversion. Depending on the domain, the size of this square matrix could be in a few thousands by a few thousand, resulting in a few million values in the matrix. Solving such a system

of equations with our high-school linear algebraic methods such as Gaussian elimination is possible, but demands a high computation power to get results in a reasonable time (which could be a few days).

So we are talking about a setup that has hundreds of high-end servers, each containing tens of cores, hundred GB RAM, a petabyte of hard-disk storage, and all connected with a high-speed network. Is this a fairy tale or such a setup really exists? India's fastest HPC computer, also called a supercomputer is in CDAC Pune, named *Param Siddhi-AI*. It has over 41,000 cores and a total of 84 TB of RAM spread across its servers. Among the whole world, it ranks 89 in terms of its computing power (as of June 2021). The current fastest supercomputer in the world is *Fugaku* from Japan which has over 7 million cores!

Clearly, our parents cannot afford such supercomputers, whose cost would be a few crore rupees for a small-sized supercomputer. Therefore, these are typically owned by private industries and government organizations. That is also one reason why we do not see them in day-to-day life. But we definitely use them almost everyday.

> PARAM Shivay, the first supercomputer assembled indigenously, was installed in IIT (BHU), followed by PARAM Shakti, PARAM Brahma, PARAM Yukti, PARAM Sanganak at IIT-Kharagpur IISER, Pune, JNCASR, Bengaluru and IIT Kanpur respectively. India's fastest HPC computer, also called a supercomputer is in CDAC Pune, named PARAM Siddhi-AI.

Challenges

Since supercomputers are not as easily accessible as TVs or smartphones, our overall understanding and appreciation of HPC is low. Due to the involved costs, not every academic institution can afford it. This has the ripple effect of not having HPC in the curriculum, not having enough human resource for teaching relevant courses, and lacking the overall HPC ambience in the country.

The supercomputer hardware itself becomes obsolete within a few years, exacerbating the problem. This poses serious

Image Source: NSM

challenges to the industry where it is difficult to find experts having the knowledge of both HPC and the domain problem. For instance, a physicist may not know how to parallelize a code, while a parallelization person may lack understanding of the Physics involved in the problem. Yet another perception issue is that ability to parallelize is sometimes not perceived to be a skill- since that is often equated with coding. Overall, our HPC expertise is scarce.

On the positive side, companies such as Amazon and Google provide limited, free online platforms where HPC codes can be executed. This makes use of the cloud based infrastructure and allows students from academic institutions to learn HPC. The government, under the aegis of the National Supercomputing Mission (NSM), has provided support to several institutions to purchase large HPC facilities to further the spread of HPC research and learning. Under NSM, there are nodal centres in different parts of the country wherein the HPC facilities would be made available to nearby institutions and users.

> Our HPC expertise is scarce. It is difficult to find experts having the knowledge of both HPC and the domain problem. An interdisciplinary curriculum encompassing computation and domain-specific problems would marry the two streams, resulting in student experts equipped to solve the scientific problems.

Ways to improve HPC ecosystem in India

These efforts are clearly in the right direction and would help improve the overall HPC ecosystem. To further the spread and visibility of HPC, we would need additional steps.

First, the private sector, the government, and academia need to work together to provide free and subsidized access to online HPC resources. This can be immensely beneficial to institutions with limited funds. Second, encouraging HPC personnel would foster growth of the related efforts. As an example, an award in the name of APJ Abdul Kalam is yearly given by Hewlett-Packard Enterprise to recognize efforts in the HPC space in research, teaching and infrastructure development. More such recognitions would encourage HPC spearheaders.

Third, curriculum courses in different HPC domains can be introduced in the physical schools as well as via platforms such as SWAYAM. Fourth, more interdisciplinary curricula need to be developed and encouraged. This is crucial because often computing students (such as from the Computer Science department) do not know enough about other domains (such as Materials Engineering or Biology), and on the other hand, the exposure of domain-specific students to computation could be limited.

An interdisciplinary curriculum encompassing computation and domain-specific problems would marry the two streams, resulting in student experts equipped to solve the scientific problems. Finally, one needs to have regular meetups of HPC experts and enthusiasts. For instance, HiPC conference is held in India in December which allows several researchers and practitioners to engage in HPC discussions and seed collaborations. Students can be nominated and supported to attend various sessions in this conference, without having a research paper to present. NSM nodal centres hold regular workshops and lecture series to satisfy the HPC hunger of the budding experts.

If one looks at the history of HPC in India, we observe that we have come a long way, and there is a longer way to go ahead. With the improvement in HPC ambience, we would build a taskforce ready to solve major scientific problems in India and beyond. ■

The author is faculty member, Dept. of Computer Science & Engineering, IIT Madras and Head-National Supercomputing Mission Nodal Centre