INSTITUTE FOR DEFENCE STUDIES & ANALYSES
रक्षा अध्ययन एवं विश्लेषण संस्थान

IDSA
Issue Brief

The Race for Leadership in Supercomputers – Does India Stand A Chance?

Prabha Rao

September 7, 2016

Summary

One outstanding feature of the supercomputer sector is that innovation is always taking place across the entire cycle, from new theories of computation to the design of chips and to new forms of software. Unlike other sectors which stabilize based on commercial considerations sooner or later, the innovation pot is always boiling over in the case of supercomputers. This is both a daunting barrier and an exciting opportunity for countries like India. There are several imperatives if India is to regain some measure of competitiveness in this strategically vital sector.

In June 2016, a significant and unusual event occurred in the world of supercomputing – the sector that specializes in very high speed computers that are used for applications such as weather forecasting and advanced weapons design. It was announced that the fastest supercomputer in the world was now the Sunway TaihuLight, a Chinese machine, which had performed at a speed of 93 petaflops – three times faster than the previous leader.[1]

Chinese supercomputers have been leading the field since 2011, but until now had depended to a large extent on key hardware components from American companies. What made the June 2016 event unusual was the announcement that, in a first for the industry, the Sunway TaihuLight was powered entirely by Chinese-designed and Chinese-manufactured processor chips. In other words, the new machine was evidence that China had mastered the entire computer engineering cycle, from conceptualization to detailed design and manufacture of individual semiconductor components. For the first time in the history of computing, the leadership at the cutting edge of a strategic technology – supercomputers – had passed from the United States to China.

Brief History of Supercomputing

To understand how this happened, and why countries like Japan, India, and many in the European Union have been overtaken by China, it is useful to understand the history of supercomputing, or High Performance Computing (HPC) as it is also referred to. The idea of HPC – specialized machines designed to operate at ever faster speeds to solve the most complex of real world problems – is universally credited to Seymour Cray, the legendary American computer designer. In 1964, the world's first supercomputer, the Control Data Corporation CDC 6600, was designed and manufactured under Cray's supervision and leadership. For almost the next 50 years, with a few exceptions, it was always a US-built supercomputer that set the trend.

Within that half century were contained two stages, or eras, in supercomputer development. The first era is usually referred to as the Monocomputer Era, and this lasted from around 1960 to 1995. The Monocomputer architecture utilized a single high speed processor accessing data stored in a single memory stack. Since this architecture was first developed by Seymour Cray and was used by all supercomputers in this era, the first era is also sometimes referred to as the Seymour Cray era of hardware.

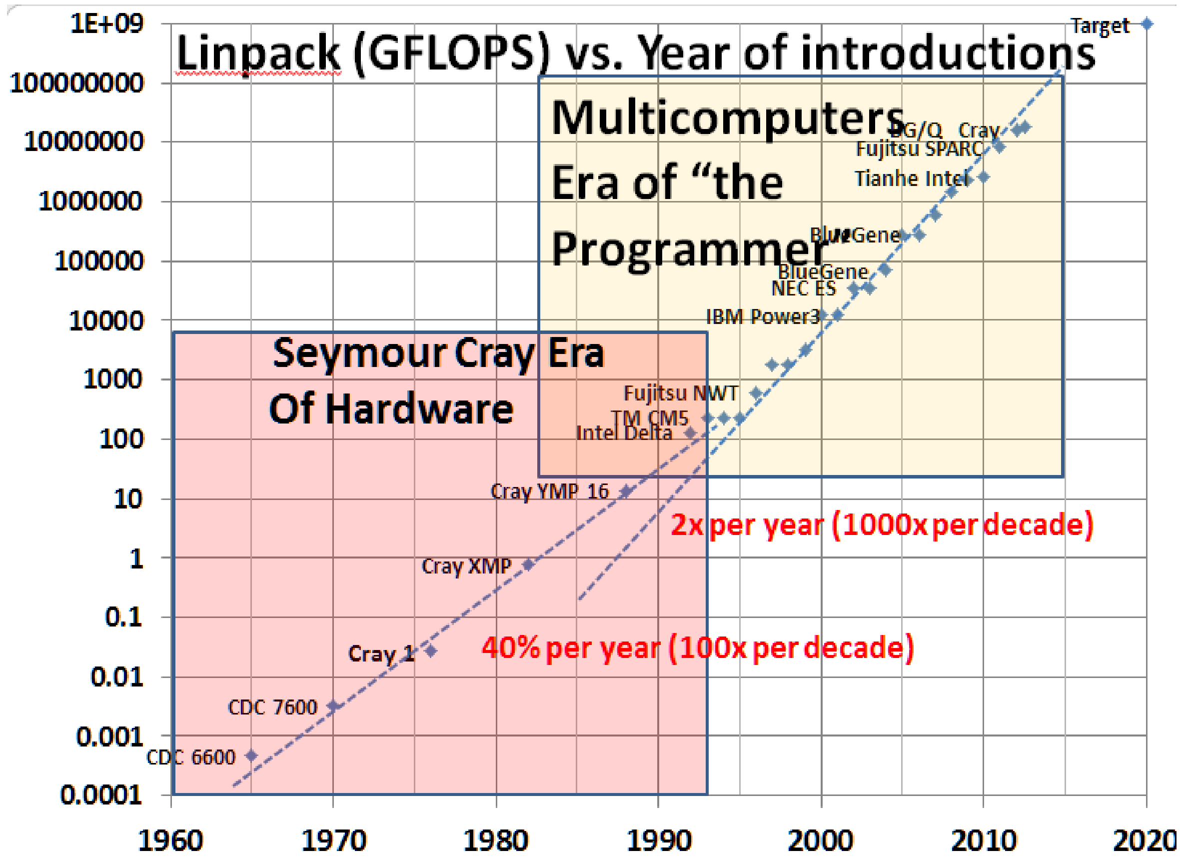
In the early 1980s, a radically different approach began to be adopted. This new approach, or architecture, used the idea that many computers or processors operating in parallel could do the job faster than a single computer using the single processor Cray architecture. Thus began the Multicomputer Era, which overlapped with the first era starting around 1985, and is continuing till date. The Multicomputer era places far greater emphasis on the software that distributes the

1 Top 500, n.d.

work between different processors, and is thus also sometimes referred to as the Multicomputer Era of the programmer.[2]

One very unusual feature of the early days of supercomputing was that developments took place entirely in the American private sector. It was only when the Europeans and the Japanese also started work on their own supercomputers that the US government began to take an active interest. Nevertheless, it was only in 1995 that the first formal US government policy – called the Accelerated Strategic Computing Initiative or ASCI – was announced. The European and Japanese initiatives, in contrast, were driven by their governments and universities.[3]

The chart below tabulates the progress of supercomputers through the two eras. On the X-axis is plotted the year of introduction of the captioned machines; and on the Y-axis the speed of each machine in Gigaflops, measured by the industry standard Linpack Benchmark. As the chart shows, speeds of supercomputers have been doubling every two years.



Source: *Supercomputers-The Amazing Race* by Gordon Bell (Microsoft Corporation, 2014), p. 7.

2 Bell, 2014

3 Ezell & Atkinson, April 2016

India's Supercomputing Efforts

The supercomputer effort in India began in the late 1980s, when the US stopped the export of a Cray supercomputer because of continuing technology embargoes. In response, the Indian government set up the Centre for Development of Advanced Computing (C-DAC) with the mission of building an indigenous supercomputer. In 1990, C-DAC unveiled the prototype of the PARAM 800, a multiprocessor machine, the first outcome of the new programme. PARAM was benchmarked at 5 Gflops, making it the second fastest supercomputer in the world at that time.

How China Achieved Dominance in Supercomputing

What, meanwhile, of China? Historical records show that China had developed an interest in HPC as early as the 1950s and 1960s. During the Mao era, even at the height of the excesses of the Cultural Revolution, and in spite of the removal of Soviet assistance after the Sino-Soviet split, the Chinese computer programme proceeded without let up. By the end of the 1960s, China was manufacturing its own integrated circuits and integrating them into indigenous third generation computers, making China in some respects even more advanced than the USSR.

In July 1972, barely four months after the epochal visit of US President Richard Nixon to China, a delegation of American computer scientists visited China at the invitation of the Chinese government, and spent three weeks with their Chinese counterparts. While they were suitably impressed by the strides made by the Chinese in mastering the technology, it was the perspective and objectives of the Chinese technology programme that really gave them pause.

The Chinese, it turned out, were not interested in the small and inexpensive "minicomputers" which were at that time taking the US and Europe by storm. What they were really interested in were the high speed machines such as the CDC Star, which were considered the state of the art in the early 1970s. It was evident to the American delegation that matching US capability in this area was a major objective of the Chinese. The delegation made this observation in the report they subsequently published in the journal *Science*.[4]

The Chinese interest in supercomputing thus seems to have been established very early and remained constant during the decades of political turmoil in the 1960s and 1970s. This interest was institutionalized very substantially in March 1986, when Deng Xiaoping initiated the famed '863' programme to acquire parity with the US, and with the rest of the world, across a range of high technology sectors. For supercomputing to develop, a host of other industries and sectors had to develop as well, such as semiconductor manufacture, design of integrated circuits, expertise in

4 The narrative in this and the previous two paragraphs is drawn from Tom Mullaney, "The Origins of Chinese Supercomputing And an American Delegation's Mao-Era Visit," Foreign Affairs, 4 August 2016, at https://www.foreignaffairs.com/articles/china/2016-08-04/origins-chinese-supercomputing

the mining and refining of rare earths, etc. All of these were integrated well into the 863 programme.[5]

It took two decades for these efforts to bear fruit. In 2006, Chinese supercomputers entered the Top 500 list for the first time. At that point, India had eight supercomputers on the list, which was otherwise dominated almost entirely by the Americans, albeit with strong competition from the Japanese at the top of the list. 10 years later, in 2016, China leads the Top 500 list with 169 machines, including the Sunway TaihuLight, the world's fastest at 93 petaflops as mentioned earlier. The US comes second, with 165 machines. Europe as a whole has about 110 machines, and Japan barely 40, although it is to the Japanese credit that the average speed of their supercomputers is the highest. India, unfortunately, has stayed nearly static with only nine systems in the Top 500 list.

Supercomputers are the second sector where China has established global leadership, the first being rare earths mining and refining, in which it holds a 95 per cent market share. But China's growing dominance in the supercomputer sector displays capabilities that go well beyond the specialized mining and refining technologies that characterize the rare earths sector.

Prerequisites for Making the Fastest Supercomputer

Developing the world's fastest supercomputer requires capabilities that start with pure science – specifically quantum physics and the electrodynamics of semiconductors. Allied with this is the requirement of a highly educated and competent cadre of computer scientists who understand the complexities of such abstract computer science concepts as the 'theory of computation' and are able to apply these concepts to developing efficient algorithms that can solve very complex real world problems. Building up a cadre of scientists with such specialized knowledge requires decades of effort, which the Chinese have systematically put in. This needs to be combined with the capacity to design Very Large Scale Integration (VLSI) integrated circuits, including complex microprocessors that are as good as, if not better than, American products.[6] A host of networking and connectivity technologies that enable large numbers of processors to operate efficiently in parallel – the Sunway has over 10 million parallel processors – need to be mastered for the design to even reach the prototype stage.

Many seemingly unconnected technologies are associated with supercomputers. For example, HPC machines consume enormous amounts of power – the Sunway alone consumes as much as 28 MW. It is to the credit of Chinese scientists that the home-grown processors used in the Sunway are actually three times as energy efficient as the nearest American equivalents. The physical design of the machine, including the cooling system, is itself a mechanical and metallurgical engineering challenge.

5 MOST, China, 2016

6 Orr & Thomas, 2014

Finally, for supercomputers to be effective, they need to be loaded with a large suite of specialized software packages, ranging from operating systems that cater for the multiprocessor environment to the application suites capable of executing algorithms that help solve the truly complex real world problems such as weather forecasting, very big data analysis, biomedical modelling, and of course security-related applications such as cryptography, advanced aerospace engineering and weapon systems design.

Future Trends in Supercomputing

This raises the questions: Do countries like India stand a chance in this race? And, what can they do? The answers may lie in a careful analysis of future trends.

The Chinese mastery of the wide range of technologies positions them well for winning the next race in supercomputers, which is breaking the "exascale barrier". In simple terms, this is the race to determine who first succeeds in constructing a supercomputer that is capable of a speed of one exaflop per second, or one thousand million Gigaflops, one Gigaflop itself being one thousand million floating point operations per second. There are four countries in the race – China, the US, France and Japan. China looks well set to win the race in the year 2018.[7] France and Japan have both indicated that they would achieve the objective by 2020, and the US has conservatively indicated 2023. But the US has also stated that it expects to regain long term leadership.[8]

The exascale barrier is a landmark for supercomputers for reasons that go beyond the mere desire to be the first. Supercomputers operating at such incredible speeds will encounter a variety of barriers that previous generations of designers did not have to contend with. For example, the network and interconnectivity hardware that allows millions of processors to operate in parallel will have to speed up by an order of magnitude to accommodate exascale performance. Similarly, the cooling system will become a central design constraint – a statement that supercomputer engineers are wont to make is that future HPC machines may need their own independent nuclear reactor for power supply and cooling!

What India Needs to Do

All this brings back into focus the need for innovation. One outstanding feature of the supercomputer sector is that innovation is always taking place across the entire cycle, from new theories of computation to the design of chips and to new forms of software. Unlike other sectors which stabilize based on commercial considerations sooner or later, the innovation pot is always boiling over in the case of supercomputers. This is both a daunting barrier and an exciting opportunity for

7 Trader, 2016

8 Thibodeau, 2016

countries like India. There are several imperatives if India is to regain some measure of competitiveness in this strategically vital sector.

First, India must move away from the perspective which it has allowed to dominate, namely, that the application of supercomputers is more important than supercomputer technologies themselves. In this perspective, it does not matter whether an HPC machine is indigenous or imported, as long as it is usefully applied. This perspective ignores the strategic importance of supercomputers and the abundant evidence that all major countries view these technologies as critical.

Second, India must understand that it is possible to start from the current state of the art itself. There is no need to entirely retrace the path already taken by China and the other countries. Using technological expertise that is available with the global network of Indian and Indian-origin scientists and engineers, it is possible to start from a baseline which is already advanced. In addition, the software skills and personnel base that India has built up in the public and private sectors can be effectively leveraged to propel innovation on the software components of supercomputer technology.

Third, India has to understand that supercomputer research always requires fundamental research into the next stages of computing. Thus, going beyond the exascale barrier might require new approaches that are right now only in the theoretical stage – quantum computing, for example, has been only spoken about in research forums, but may well turn out to be the basis of the next leap forward. The time frames required to operationalize and commercialize nascent technologies are shrinking, and this is something that needs to be factored into the Indian approach.

Fourth, India should set itself clear objectives of what it wants to achieve in this strategically significant sector. The Chinese perspective is telling – over 50 years ago, China set itself the clear objective of parity with the United States. While the setting up of the National Supercomputer Mission in 2015[9] is a laudable first step, it needs to be followed up by the identification of clear objectives and allocation of adequate resources. Within a Mission perspective, it should be possible to cut down bureaucratic red tape and allow scientists and engineers to take bold and radical steps without fear of reprisal.

Finally, it needs to be appreciated that supercomputers are strategic in the most important sense, namely, the creation of an ecosystem that extends well beyond the boundaries of science and technology and has the capacity to transform the country. A strong supercomputer sector leads to capability in a variety of other fields, from semiconductor manufacturing and precision engineering to optimal strategies for agricultural production, urban planning and the like. All this would be in addition to the national security related applications where India cannot afford to be dependent on foreign expertise. Building up capability in this sector requires active government leadership to catalyse the establishment of a vibrant academic infrastructure where research at the frontiers of physics and material sciences, computational mathematics and computer science are encouraged, to establish strong partnerships with industry for technology transfer and commercial

9 Cabinet Committee on Economic Affairs India, 2015

exploitation, and finally to create widespread awareness of the possibilities and potential of supercomputers. In the more advanced countries, using supercomputer resources has become routine for a large and increasing percentage of Fortune 500 companies. In China, the Sunway TaihuLight installation is intended to function as a public service, with access available to all. It may be simpler for India to catch up with these countries than is commonly imagined. What is required are bold decisions that aim at reaching comparative parity within the next decade.

References

- Bell, G. (2014). *The Amazing Race- History of Supercomputing 1960-2020.* San Francisco, California: Microsoft Corporation.
- Cabinet Committee on Economic Affairs India. (2015). *National Supercomputing Mission (NSM): Building Capacity and Capability.* New Delhi: Press Information Bureau, Government of India.
- Ezell, S. J., & Atkinson, R. D. (April 2016). *The Vital Importance of High Performance Computing to US Competitiveness.* Washington DC: Information Technology and Innovation Foundation.
- MOST, China. (2016). High Technology R&D Program (863 Program).
- Mullaney, T. (2016, August 4). The Origins of Chinese Supercomputing. *Foreign Affairs*.
- Orr, G., & Thomas, C. (2014). *Semiconductors in China: Brave new world or same old story?* McKinsey & Co.
- Thibodeau, P. (2016, June 21). US to have 200-peatflop supercomputer by early 2018. *ComputerWorld*.
- Top500. (n.d.). *Top500.* Retrieved from www.top500.org: www.top500.org
- Trader, T. (2016, May 2). *China Sets Ambitious Goal To Reach Exascale By 2020.* Retrieved from HPC Wire: http://www.hpcwire.com/2016/05/02/china-focuses-exascale-goals/

About the Authors

Prabha Rao is Senior Fellow at the Institute for Defence Studies and Analyses (IDSA).

The Institute for Defence Studies and Analyses (IDSA) is a non-partisan, autonomous body dedicated to objective research and policy relevant studies on all aspects of defence and security. Its mission is to promote national and international security through the generation and dissemination of knowledge on defence and security-related issues.

Disclaimer: Views expressed in IDSA's publications and on its website are those of the authors and do not necessarily reflect the views of the IDSA or the Government of India.

© Institute for Defence Studies and Analyses (IDSA), 2015

Institute for Defence Studies and Analyses
1, Development Enclave, Rao Tula Ram Marg
New Delhi 110 010 India
T +91-11-2671 7983 F +91-11-2615 4191
contact.idsa@nic.in www.idsa.in
Twitter @IDSAIndia
www.facebook.com/InstituteforDefenceStudiesAnalyses