

# Chapter 20

## What Supercomputing Will Be Like in the Coming Years

Mehmet Dalkilic

Indiana University, Bloomington, USA

### ABSTRACT

*This chapter is an abridged sort of “vision statement” on what supercomputing will be in the future. The main thrust of the argument is that most of the problem lies in the trafficking of data, not the computation. There needs to be a worldwide effort to put into place a means to move data efficiently and effectively. Further, there likely needs to be a fundamental shift in our model of computation where the computation is stationary and data moves to movement of computation to the data or even as the data is moving.*

### INTRODUCTION

Indiana University has had a history of possessing some of the most powerful supercomputers for of any University world-wide. So, it is fitting to do exercise some visioning into the future of supercomputing holds for faculty who routinely use a supercomputer. The newest supercomputer, named Big Red 200 (BR200), coinciding with Indiana University’s 200<sup>th</sup>, anniversary is among the first next-generation supercomputer based on HPE’s Cray Shasta architecture. BR200 operates at nearly 6 petaFlops which, at the time of this writing, makes it the 32nd most powerful supercomputer in operation world-wide and the most powerful university-owned and operated AI supercomputer in the U.S.

### A Simple Observation From Feynman

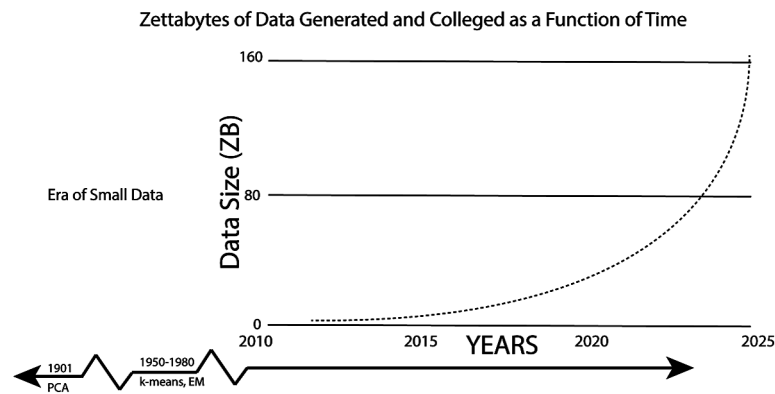
Richard Feynman (1918-1988) was a Nobel winning physicist who worked not only in physics, but in other areas as well, *i.e.*, computing (Feynman et al., 1998). In this work, he makes a prescient prediction which we will paraphrase where  $t$  is time:

DOI: 10.4018/978-1-7998-7156-9.ch020

$$\lim_{t \rightarrow \infty} \frac{\text{Computation}_t(\text{Data})}{\text{Move}_t(\text{Data})} = 0$$

This means, in words, as scientists move toward building increasingly more powerful computing machines, the time to move the data will be the limiting factor, not the time of the computation. This problem was seen several years back *e.g.*, (Coughlin, 2018). A projected timeline showing the disparity between the growth of data in zettabytes

Figure 1. Amount of data produced and predicted vs. when current machine learning and AI algorithms were developed. It is estimated that all of human speech ever spoken could be captured with ~42 zettabytes.



and when the current, most popular AI/ML algorithms were created is shown in Figure 1. The difference is startling when one realizes that the most popular data reduction technique, PCA (Principle Component Analysis) is 120 years old. Even with Moore's Law still reasonably true, for the next couple of years, the fact remains that data growth will continue. The traditional means of moving data to the computation rather than moving (multiple) computations to the data will have to change. The infrastructure that researchers currently have moves data at nearly  $2/3$  the speed of light for short distances, but, almost amusingly, using *physical* transportation of the data remains faster than sending it electronically. What does this mean for the supercomputer?

## Monolithic Supercomputer

This inundation of data might mean that the standard, monolithic supercomputer's days are number except for specialized types of computation, *e.g.*, weather. With the resurgence of AI/ML pervading almost every STEM discipline, and many non-STEM disciplines as well, more and more people demand access to computation for a myriad of purposes with one thing in common—big data. Outside of those who use supercomputers, there is a misconception that computation is the limiting factor. This fact is significant because it is the public who will decide what to fund—an easily understandable, yet more powerful, supercomputer or infrastructure to shuttle data more effectively, or even wholly new architectures where computation is done on the data as it is moved. What will it be? Though Quantum Computing is lurking

1 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/what-supercomputing-will-be-like-in-the-coming-years/273410](http://www.igi-global.com/chapter/what-supercomputing-will-be-like-in-the-coming-years/273410)

## Related Content

---

### Non-Contact Abnormal Physiological Status Detection During Sport and Training

Jinlin Yang (2021). *International Journal of Distributed Systems and Technologies* (pp. 16-26).

[www.irma-international.org/article/non-contact-abnormal-physiological-status-detection-during-sport-and-training/287860](http://www.irma-international.org/article/non-contact-abnormal-physiological-status-detection-during-sport-and-training/287860)

### Small World Architecture for Building Effective Virtual Organisations

Lu Liu and Nick Antonopoulos (2009). *Grid Technology for Maximizing Collaborative Decision Management and Support: Advancing Effective Virtual Organizations* (pp. 190-211).

[www.irma-international.org/chapter/small-world-architecture-building-effective/19345](http://www.irma-international.org/chapter/small-world-architecture-building-effective/19345)

### Soft-Checkpointing Based Hybrid Synchronous Checkpointing Protocol for Mobile Distributed Systems

Parveen Kumar and Rachit Garg (2013). *Development of Distributed Systems from Design to Application and Maintenance* (pp. 87-100).

[www.irma-international.org/chapter/soft-checkpointing-based-hybrid-synchronous/72248](http://www.irma-international.org/chapter/soft-checkpointing-based-hybrid-synchronous/72248)

### Extended Prefix Hash Trees for a Distributed Phone Book Application

Fabian Stäber, Gerald Kunzmann and Jörg P. Müller (2009). *International Journal of Grid and High Performance Computing* (pp. 57-69).

[www.irma-international.org/article/extended-prefix-hash-trees-distributed/37513](http://www.irma-international.org/article/extended-prefix-hash-trees-distributed/37513)

### Introduction to Serverless Computing

Radhika Kavuri and Satya Kiranmai Tadepalli (2024). *Serverless Computing Concepts, Technology and Architecture* (pp. 1-19).

[www.irma-international.org/chapter/introduction-to-serverless-computing/343717](http://www.irma-international.org/chapter/introduction-to-serverless-computing/343717)