



IN-MEMORY COMPUTING

Matthew N. O. Sadiku, Adebowale E. Shadare, and Sarhan M. Musa

Roy G. Perry College of Engineering

Prairie View A&M University

Prairie View, TX 77446

ABSTRACT

In-memory computing is the storage of information in the main random access memory (RAM) of servers rather than in databases operating on comparatively slow disk drives. It is computing that stores data in RAM across many connected devices and processes it in parallel to speed processing. This allows enterprises to analyze data in a matter of seconds instead of hours. This paper provides a brief introduction to in-memory computing.

Keywords: In-memory computing, Near-memory computing, Memory-centric architecture.

I. INTRODUCTION

The rise of the Internet of Things (IoT), cloud computing, and social media has created demand for applications that need to be extremely fast. Traditionally, data is placed in storage; when needed the data will be accessed and acted upon in the computer's memory. As volumes of data increase, so does the time needed to access. In-memory computing (IMC) (also known as near-memory computing or memory-centric architectures) does away with the traditional concept of storage. Instead, data is stored directly in the computer's memory. Thus, IMC involves storing data in the random access memory (RAM) (or flash memory) of servers instead of in databases running on relatively slow speed. RAM-stored data is available instantaneously, while data stored on disks is limited by network and disk speeds. IMC is compared with the traditional computing architecture in Figure 1 [1]. Recently, the cost of RAM has continued to fall and this helps deliver in-memory speed plus massive scalability and high availability.

Companies are upgrading their IT architecture to take advantage of the low latency processing offered by in-memory computing. With an increased availability of high-performance multi-core processors, the drop in memory prices and this has made in-memory computing economical among a wide variety of applications.

II. BACKGROUND

Thus, IMC comprises of three components: in-memory computing, in-memory databases, and in-memory data grid.

- *In-memory computing:* In-memory computing platforms have emerged, bringing together in-memory data grids, in-memory processing, in-memory database, and streaming analytics into a single, unified platform. In-memory processing is an emerging technology for processing of data stored in an in-memory database. Since data is placed in random-access memory (RAM), in-memory processing allows data to be analyzed in real time. In-memory processing is at a lower cost compared to traditional BI tools [2].
- *In-memory databases:* In-memory databases (IMDB) use random access memory (RAM) for data storage, which results in faster response and processing. The key feature of IMDB is its ability to load and execute all data in memory. IMDB requires a substantial amount of memory since all data must fit into memory. It has been touted as the cure for database performance problems [3].
- *In-memory data grid:* In-memory data grids (IMDG) store all of their data in memory. By utilizing system memory rather than spinning disks, IMDG are much more faster than traditional systems. They are designed with memory-first and disk-second approach. Integrating in-memory computing into in-memory data grids has brought many of the benefits of parallel super-computing to the enterprise.

III. APPLICATIONS

In many real-life applications, in-memory computing was an enabling technology.

Facebook uses a lot of in-memory infrastructure. New memory-driven applications include business, financial services, big data, Internet of things (IoT), mobile computing, power grid, cloud/SaaS, retail, healthcare, supply chain, manufacturing, online gaming, bioinformatic, and sciences.

- *Business:* The increasing use of technology in business is generating more and more data. Managing data to help achieve business objectives is the most important task facing any organization today [4]. IMC enables business customers such as banks, retailers, and utilities to analyze huge volumes of data on the fly. It enables extremely fast processing which helps businesses to increase performance. Marketing enterprises can use in-memory computing to deliver real-time personalization. Financial companies use IMC for market surveillance to improve fraud detection.
- *Big Data:* As data volumes increase and become more complex, organizations must find ways to process such a huge amount of information in a timely manner. While big data offers enormous potential for an organization, it can create significant latency when users must access data. Traditional big data analytics tools are limited by the speed limitations of storage systems which can be spinning physical disks which may be electronic, optical, or magnetic. Adopting in-memory in big data improves exponentially since IMC can provide ultra-fast access to big data. IMC tools are designed to help organizations analyze their vast quantities of business data. Any organization with big data (such as utilities or finance) needs to adopt in-memory computing at some point in order to continue to function [5].
- *Power Grid:* An in-memory computing can help power grid address the hard-disk and network I/O bottleneck problem, where data processing and computation response speed are crucial. IMC means using distributed in-memory data grid that allows one to store data in computer memory and process the data in parallel directly in the memory. The IMC-based power grid analysis approach is generic and can be applied to many power grid analysis scenarios [6].

IV. BENEFITS AND CHALLENGES

The most obvious benefit of in-memory processing is its speed. When data is stored in RAM, applications run 1,000x faster because data does not need to be retrieved from disk and moved into RAM for processing. Another major benefit of a memory-centric architecture is that it avoids the need to wait for all the data to be reloaded into RAM in the event of a reboot [7]. IMC can deliver data-crunching benefits across applications and do so in real time. It can transform business IT and enable companies to crunch and analyze large volumes of data in near real-time. In-memory computing handles well unstructured data such as social media news feeds. Several organizations have taken advantage of these benefits for purposes from maximizing sales to analyzing gene sequences.

Although keeping all data in-memory works well for many applications, it does not work for all. An apparent limitation of many IMC adoptions is that all data must fit in memory. A large amount of RAM is needed for an in-memory configuration for large numbers of users and this affects the hardware costs. Security is a major challenge since in-memory tools expose large amounts of data to end users. Other challenges include:

how to manage data movements in the new infrastructure; how to organize data in memory; how to achieve a desirable level of persistency for memory data; how to manage the hybrid memory All the challenges are essential for IMC to effectively apply to big data processing [8].

V. CONCLUSION

In-memory computing refers to the use of direct memory instead of disks for the purpose of storage and computation. It is a promising emerging framework to accelerate data-intensive processing by striving to keep the entire data in RAM. The shift from traditional, hard disk-enabled data warehouses to IMC-enabled data warehouses reduces the number of layers of data analysis. Many organizations have turned to IMC to solve their speed and scale challenges. Although the concept of IMC is well known, its deployment is still limited to innovators. IMC is clearly the future of computing. More information on IMC can be found in books in [3,9-10] and several books available at Amazon.com.

REFERENCES

- [1] P. P. Chougule, B. Sen, and T. D. Dongale, "Realization of processing in-memory computing architecture using quantum dot cellular automata," *Microprocessors and Microsystems*, vol. 53, 2017, pp. 49-58.
- [2] "In-memory processing," *Wikipedia*, the free encyclopedia
https://en.wikipedia.org/wiki/In-memory_processing
- [3] W. E. Yu, "In-memory computing—Evolution, opportunity and risk," *ISACA Journal*, vol.5. 2013.
- [4] "In-memory computing yields real-time insights from big data," *Strategic Perspectives & Insights for Executives*, vol. 13, no. 7, Feb. 2013.
- [5] A. Bowling, "Can in-memory computing answer the big questions about big data?"
<https://www.computerweekly.com/opinion/Can-in-memory-computing-answer-the-big-questions-about-Big-Data>
- [6] M. Zhou and D. Feng, "Application of in-memory computing to online power grid analysis," *IFAC PapersOnLine*, 2018, pp. 132–137.
- [7] A. Kleinfeld, "Memory-centric architectures: What's next for in-memory computing," June 2018.
<https://thenewstack.io/memory-centric-architectures-whats-next-for-in-memory-computing/>
- [8] X. Shen, R. Lovas, and X. Liao, "Editorial for the special issue on in-memory computing," *Journal of Parallel and Distributed Computing*, vol. 129, 2018, p. 322.
- [9] G. Blokdysk, *In-Memory Computing*. Emereo Publishing, 2nd ed., 2018.
- [10] G. Pulipaka, *Big Data Appliances for In-Memory Computing: A Real-World Research Guide for Corporations to Tame and Wrangle Their Data*. High Performance Computing Institute of Technology, 2015.

ABOUT THE AUTHORS

Matthew N.O. Sadiku is a professor at Prairie View A&M University, Texas. He is the author of several books and papers. He is an IEEE fellow. His research interests include computational electromagnetics and computer networks.

Adebowale Shadare is a doctoral student at Prairie View A&M University, Texas. He is the author of several papers.

Sarhan M. Musa is a professor in the Department of Engineering Technology at Prairie View A&M University, Texas. He has been the director of Prairie View Networking Academy, Texas, since 2004. He is an LTD Sprint and Boeing Welliver Fellow.

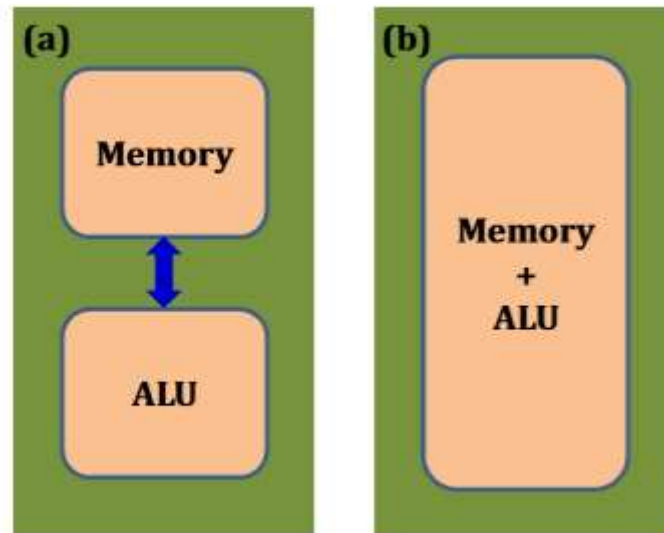


Figure 1 Computer architectures: (a) traditional: separate memory and processing unit. (b) In-memory computing: both memory and processing unit in single unit [1].