In-Memory Computing

ECAR-WG Technology Spotlight

This bulletin is one of a series of papers from ECAR working groups designed to help institutional leaders learn about and understand the implications of emerging technologies in higher education. These technologies have been identified as the “Top 10 Confusing Technologies” in the ECAR report Higher Education’s Top 10 Strategic Technologies for 2015. Other papers and related resources are available at the research hub for Higher Education’s Top 10 Strategic Technologies for 2015.

In-memory computing helps make big-data analytics possible by keeping data in a server’s RAM rather than removing it to databases on separate disk drives. Keeping it at the cached layer means not only that it can be retained in its original format but also that it can be searched, found, retrieved, and processed at much higher rates than data kept on disks, particularly when dealing with large amounts of data. In-memory computing has been made possible by a combination of factors: an increase in computer processing speeds, alongside a decrease in prices for RAM memory storage.

The speed of accessing and processing data this way makes real-time reporting possible, as well as on-the-fly analysis. In-memory computing eliminates the need—for performance reasons—to organize and index the data in ways tailored to the reporting we want to do. As a result, it also removes the need for processes such as ETL (extract, transform, load) from the process. Removing ETL means that data can be accessed and used faster (because effectively all latency has been removed from the system itself with an in-memory approach) and that adding new data to existing data stores is greatly simplified.

Importance to Higher Education

Higher education, like all business sectors, is seeing a drastic increase in the amount of data that are produced and stored and is seeing significant value in retaining those data for analytics purposes. One reason for this increase is the shift of teaching and learning technologies to online methods, both for in-person courses and for online courses and MOOCs. These platforms are generating tremendous amounts of data, and traditional databases and approaches are insufficient for tackling the volume of data generated. In addition, higher education research creates and needs to process increasingly large data sets, in hard sciences such as biology and chemistry and also in social sciences, humanities, and elsewhere. The Internet of Things (IoT) is also likely to have a huge impact on campus environments and their various systems (lighting, cooling, labs, etc.), along with a resultant potential for new sources of massive amounts of data.
data. Processing all that real-time information requires in-memory database solutions. In-memory solutions might also be able to help in the area of administrative systems, enabling ERP solutions to combine transactional (online transaction processing, or OLTP) and analytical (online analytical processing, or OLAP) databases, for example, thereby simplifying these systems and reenvisioning data warehousing.

In-memory solutions present tools that can help higher education tackle the problem of managing these vast amounts of data and improve outcomes as a result. Finding connections and patterns across large amounts of data becomes possible and fast, enabling real-time analytics and data-driven decision making. As was noted in Forbes, “Speed is absolutely crucial for predictive analytics to succeed. [With in-memory computing], the Internet of Things…will become manageable. You will know what’s going on in near real-time—rather than waiting around.”3 From the perspective of a research institution, in-memory computing will allow researchers to explore expanded solution spaces to problems (e.g., research modeling complex proteins) that just cannot be accomplished without this technology.

A significant question, however, that needs to be asked is, How important is it for real-time access to data and analytics? If it isn’t absolutely necessary, in-memory computing may not be as crucial for higher education as it is in other industries—at least not at this time. The role that consumer IT platforms play in shaping user expectations will certainly and eventually have an impact on how user expectations for IT solutions in general (including in higher education) and for real-time information as a must.

**Current Landscape**

Gartner selected in-memory computing as one of its Top 10 Strategic Technology Trends for 2013, saying that “The possibility of concurrently running transactional and analytical applications against the same dataset opens unexplored possibilities for business innovation. Numerous vendors will deliver in-memory-based solutions over the next two years[,] driving this approach into mainstream use.”4 Applications are being developed to take advantage of in-memory computing; for instance, SAP HANA is an in-memory computing platform that is currently available for higher education use.5 Other software solutions, such as the GridGain In-Memory Data Fabric, provide a unified API to connect to multiple data stores (e.g., SQL, NoSQL, and Hadoop).6 In the EDUCAUSE 2015 Top 10 Technologies Survey, when asked about in-memory computing, only 2% of respondents noted that they have it in place, with another 6% planning or implementing it now. An additional 12% are tracking it for potential future implementation (see figure 1).

![Image](image.jpg)

*Figure 1. Results of the 2015 Top 10 Strategic Technologies survey, when respondents were asked about in-memory computing*
When You Should Expect It

In-memory computing capabilities exist and are being used now, but primarily outside higher education. Specialized research may begin looking at in-memory computing to support individual projects, but the use of this technology on a wide scale and outside academic research projects is unlikely in the near future. As institutions aim to do more with big data and research, however, particularly for real-time analytics, it is a technology that should be tracked. Importantly, institutions should begin having conversations about how to develop the skill sets that will be needed to support and nurture the use of these technologies in their environments. Further, as architectures for in-memory databases grow and their scale of use increases, the network may become a potential bottleneck in scaling these systems, prompting discussions about increased network prominence and what higher education may need to do to improve network infrastructure.

Reviewers

ECAR would like to thank Jim Alves-Foss, Director, Center for Secure and Dependable Systems, University of Idaho, and Mark Silis, Associate Vice President, Information Systems and Technology, MIT, for their very helpful and significant feedback to this paper.

Related Resources

- TIBCO Software: In-Memory Computing
- IBM: What is in-memory computing?
- “In-Memory Analytics,” white paper, Yellowfin, April 2010.
Notes

1. Although *in-memory computing* is by far the most common term for this technology, it has also been called *in-memory processing*. In addition, the related term *in-memory analytics* refers to the use of this technology specifically for analytics, most commonly for business intelligence (see “In-Memory Analytics”).

2. For a look at price per Gb from 1980 to 2009, see “A History of Storage Cost.” See also Agam Shah, “Memory Prices to Fall This Year after Stabilizing in 2014,” *Computerworld*, January 14, 2015, which points to competition and excess inventory as drivers.


5. See SAP HANA In-Memory Platform.