
# *Research Report*

## The Arrival of Affordable In-Memory Database Management Systems

***Executive Summary***

The enterprise computing marketplace is about to enter a new era of computing: the era of *affordable* in-memory database management systems (IMDBMS). Three factors are driving this rapidly evolving market segment:

1) *Input/output* (I/O) – Accelerated I/O subsystems and fast memory channels are allowing gigabytes of data to be fed to fast "designed-for-data" processors where that data can be processed quickly in-memory;
2) *Database acceleration* – Database accelerator software is allowing large volumes of data to be compressed and read in columnar form – delivering results exponentially faster than using traditional row-based database techniques; and,
3) *Recognition of the importance of Big Data* – Enterprises now realize that analyzing "Big Data" can help analysts discover new insights, create competitive advantages, and better serve their customers.

*New I/O Subsystem Interfaces, Fast Memory Channels and Fast, Multithreaded Processors*

From a hardware perspective, three important advances have taken place this year:

1. *There have been major improvements in I/O speed* – enabling data to be delivered to the processor a lot faster than traditional I/O subsystems. IBM's Coherent Accelerator Processor Interface (CAPI) is an example of a major I/O speed improvement. This interface makes it possible to attach large volumes of Flash cache (up to 40 TB) to a POWER8 processor – where it appears to be "in memory" (as opposed to being Flash storage).
2. *The use of RDMA* – Several vendors are starting to make use of a communications protocol known as Remote Direct Memory Access (RDMA) to streamline communications (using a zero copy approach) between I/O devices and the processor;
3. *Memory channel speed has been improved* – again enabling more data to be delivered more quickly to processors, thus accelerating the processing of large volumes of data; and,
4. *Processors have gotten faster* (particularly IBM's recently-introduced POWER8 microprocessor that can now process 8 threads per cycle).

***This combination of faster speeds and feeds combined with faster processors and large shared memory/cache support has enabled the creation of a new generation of lower cost IMDBMS servers.***

*Database Acceleration in Software*

From a software perspective, database vendors such as IBM, Oracle and SAP have found new ways to accelerate the analysis of large Big Data databases. *Clabby Analytics* has

already analyzed the database accelerators provided by Oracle (Exadata), IBM (BLU Acceleration), and SAP (Hana) – see this report – but there are numerous other IMDBMS vendors including Altibase, Exasol, Kognito, McObject, ParStream, Quartet FS, VMware and VoltDB that have entered the IMDBMS marketplace.

> ***In short, vendors are finding ways to accelerate the speed at which a database can parse data – resulting in exponential decreases in the time it takes to perform analysis and achieve a result.***

*Changing Business Needs*

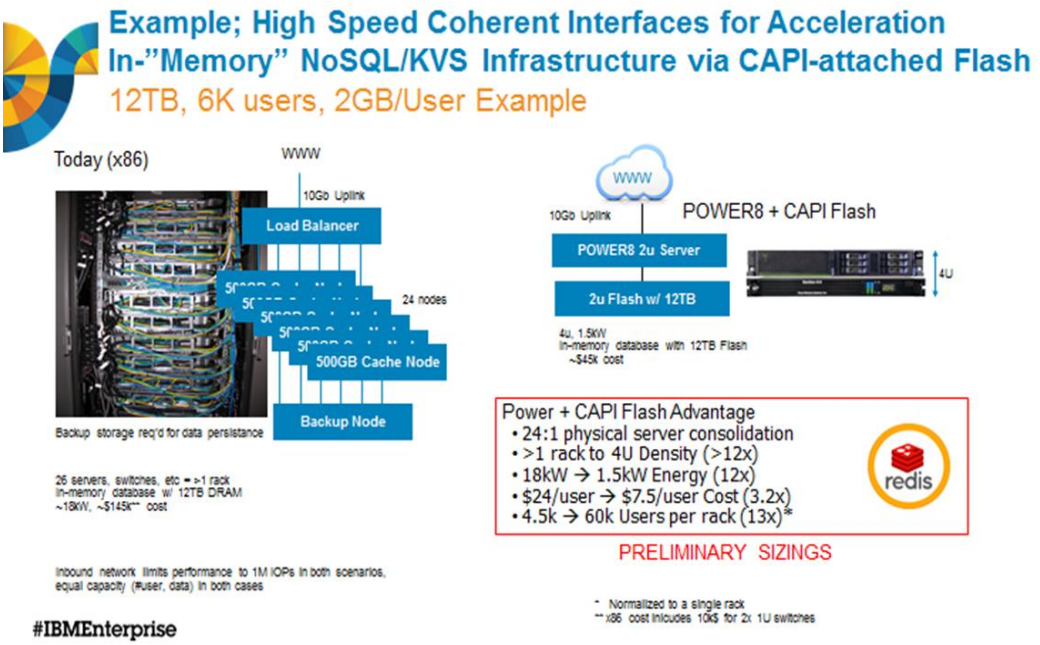
Enterprises now recognize that valuable insights can be derived from the analysis of the volumes of data that the enterprise has already captured. Enterprise executives want to use Big Data analysis to derive new insights; to create competitive advantages; to derive new insights into business operations; and to better service customers and/or to better understand customer behavior. With lower cost in-memory Big Data analytics servers now available, business executives finally have access to affordable in-memory/cache servers they need to analyze large volumes of Big Data in real-time.

In this *Research Report*, *Clabby Analytics* takes a close look at the major changes that have taken place over the last year in hardware and software that are helping to drive improved analytics processing performance while driving down the cost of IMDBMS servers.

***Affordable***

Figure 1 provides a rough cost comparison of an in-"memory" (meaning Flash is being used like memory) Power System NoSQL/KVS configuration as compared with a more traditional x86 in-memory environment.

***Figure 1 – A Traditional x86 In-Memory Environment Compared to Power Systems w/CAPI***




***Source: IBM Corporation – October, 2014***

Important preliminary sizing information can be found in the red box on the lower right of Figure 1. What this information shows is that a typical in-memory x86 environment with associated back-up storage for persistence, 12 TB of DRAM (memory), two 1U switches and load balancing software costs about $24/user. This compares with an IBM Power System with 12 TB of Flash used as memory cache that costs about $7.50 per user. What this means is that a Power Systems solution can cost 3.2x less than an x86-based in-memory solution. (Note: Flash solutions attached to the memory channel are also available in the market – so costs are coming down on the x86 side too).

***At Clabby Analytics we believe that the cost of in-memory systems that use Flash as memory cache will entice many IT buyers who have been unable to afford traditional, expensive in-memory configurations (with expensive DRAM) to buy these new, lower-cost Flash-based solutions. We think new price-point will ignite the market for in-memory database management systems.***

***The Changes in Hardware That Support IMDBMS***

Four changes in hardware have enabled larger, shared memory support and have also provided a means to drive memory/cache prices downward – resulting in a new generation of lower cost in-memory server architectures. As described in the *Executive Summary*, these three changes are: 1) an accelerated I/O subsystem; 2) faster memory channels; 3) the use of RDMA protocol and associated adapters; and, 4) faster processors.

*Accelerating I/O: The Impact of CAPI*

IBM's new CAPI model represent a more direct ways for I/O devices to communicate with processors. In short, what CAPI does is its allows external adapters to communicate with the same memory addresses that the processors use – allowing those adapters to move data in and out of a POWER8 processor at speeds that are much faster than traditional I/O channels.

*Memory Channel Acceleration*

The closer data is to a given processor, the faster it can be processed (which is why the amount of levels 1, 2, 3 and 4 cache located on the processor chip or off-chip is important). After processor cache, the next level of "closeness" to the processor has traditionally been main memory. In Intel systems, the amount of shared main memory that could be addressed in recent years topped-out at between 2TB and 4TB; whereas POWER7 and now POWER8 can reach 16TB of main memory.

The amount of memory that can be addressed is important – but the speed at which data can be sent to the processor is equally important because slow memory channel speed means that the processor will have to wait for data. In systems design, the goal is to balance the amount of data the processor can handle with the speed at which a memory channel can feed those processors:

- The memory bandwidth speed of a Xeon v2 E5 processor is about four times slower than the memory bandwidth speed of a POWER8 processor – but the Xeon can only process 2 threads per clock cycle as opposed to the 8 threads that a POWER8 can handle – so the Intel memory speed to processor thread ratio looks pretty good

(Xeon v2 can process a quarter as many threads as a POWER8 – which is why a four times slower memory channel speed is an good ratio for this environment).

- Ironically, the POWER8 processor can process 4 times as many threads as the Xeon – and its memory channel speed is about 4 times faster than a Xeon v2-based server. So again, the memory channel speed looks good for this environment.

*The Use of RDMA*

A well-balanced, well-tuned system does not waste clock cycles. So, to ensure that processors are kept busy, the I/O and memory subsystems need to feed the processors a steady stream of data to be processed. But traditional approaches to I/O communications are fraught with communications and management overhead that impedes the flow of data to its destination. For instance, using the tradition I/O flow model, device drivers are called, copy or pin source data is derived, memory mapped I/O needs to take place, acceleration needs to take place, polls and interrupts need to complete, copy or unpin results need to be correlated – and then data from the device driver can be delivered. All of this activity represents tens of thousands of commands and calls – and results in a lot of processor management/communications overhead just to manage device interaction with a processor.

***RDMA uses an approach to communications called "zero copy" which eliminates a lot of the abovementioned communications activities – thus significantly speeding-up communications between the processor and underlying devices.***

*Fast Processors*

2014 has been a big year in the server microprocessor marketplace with big "built-for-data" microprocessor announcements in February and April of this year:

- In February, 2014, Intel introduced a new version of its x86-based Xeon processor that can support up to 32 sockets, with 15 cores and 1.5 TB of memory per socket. This means is that Intel server designs can potentially support up to 48 TB of data in memory in a single scale-up server (but note, this is not shared memory). Also, from a processing power perspective, Intel's Xeon v2 processors can execute 2 threads per core (so a 15 core Xeon processor can execute 30 threads – for a total of 960 threads in a 32 socket server).

- In April, IBM introduced three new versions of scale-out POWER8 servers that support up to 1TB of main memory per socket (IBM's initial POWER8 servers are 2 and 4 socket configurations – but larger, scale-up multi-socket configurations will be released shortly). One particular POWER8 module is offered in a 12 core processor configuration that can process 8 threads per core (or 96 threads per clock cycle – the same number of threads as in the Intel 32-socket server described above. This POWER8 12 module offers an equivalent amount of processing power in a single node).

***What all of the specifications in the last three sections mean is that new generation Intel Xeon v2 and IBM POWER8 microprocessors are now capable of rapidly processing very large volumes of data held in main memory and in Flash cache. These vendors have created specialized interfaces to speed the flow of data from cache to the processors (using interfaces such as IBM's CAPI). And, as a result, new large cache and memory systems configurations are coming to market that can provide accelerated analytics processing at a lower cost (because Flash memory costs significantly less than traditional main memory).***

***Software Advances in In-Memory Databases***

Server/processor-side improvements are only part of the reason that the IMDBMS market is poised for strong growth. Over the past few years we've seen dramatic improvements in database design that have greatly accelerated the speed at which very large databases can be read and analyzed.

For instance, IBM's BLU Accelerator can read traditional row-based data – or can restructure that data to be read in columns – delivering results from 10X to hundreds of times faster than traditional databases. SAP, Oracle, Altibase, Exasol, Kognitio, McObject, ParStream, Quartet FS, Volt DB and VMware have also adapted their software products to support large in-memory configurations.

As we observed in this report, we see database vendors using three approaches when architecting new Big Data/IMDBMS solutions. They focus on:

1. Finding ways to organize and compress data such that large amounts of data take up less space – and find a way to read compressed data to speed query completion;
2. Using efficient algorithms to accelerate the speed of Big Data analysis; and,
3. Selecting a system environment that provides balanced resource utilization such that CPU power, memory, input/output (I/O), networks and storage all work together in a balanced fashion in order to generate query results as expeditiously as possible.

***The bottom line in comparing Big Data/IMDBMS solutions is that there are some commonalities in the ways that vendors approach Big Data analytics – but there are also some very distinct differences. These differences manifest themselves in database query processing speed; number of concurrent users that can be supported; the affect that query complexity can have on system performance; system efficiency/optimization; and manageability.***

***Summary Observations***

What is driving the adoption of IMDBMS systems? The first major driver is *increased database processing speed* due to accelerated access to data held in main memory and low cost solid state Flash cache. The second driver is the ability to *read much larger amounts of data (Big Data) more quickly*. The ability to read more data more quickly results in deeper insights and better analytical results (because more data has been culled to produce a given result). Enterprises now recognize the competitive importance of analyzing Big Data – and with new generation IMDBMS environments, suddenly the cost to analyze Big Data has decreased significantly.

**The Arrival of Affordable I-Memory Database Management Systems**

Growth in the previous generation of Intel x86- and POWER-based systems had been constrained by the cost of large volumes of random access main memory. But with the introduction of new I/O connectivity options such as IBM's CAPI – and with the introduction of much lower cost Flash memory (that can serve as high-speed data cache) – previously price prohibitive IMDBMS systems have become much more affordable for businesses large and small.

***To date, in memory databases have been used primarily in the energy and telecommunications industries where real-time analytics are needed to examine sensory data within pipes, to manage energy output or to monitor communications equipment. But, with the availability of more affordable in-memory systems designs, we expect strong growth in new markets including advertising and retail environments (where large volumes of geospatial and locational data can be analyzed to predict and predetermine consumer behavior), and in financial and telecommunications environments where real-time data will be analyzed in real time in order to create competitive advantage. And these new split-second analysis environments represent just the tip of the iceberg. As the sales volume for real-time in-memory systems increases – and as managed service providers bring new Database-as-a-Service environments to market – expect prices to decrease even further, making in-memory databases affordable across a wide spectrum of small, medium and large businesses.***

***Clabby Analytics***
***http://www.clabbyanalytics.com***
*Telephone: 001 (207) 846-6662*

*© 2014 Clabby Analytics*
*All rights reserved*
*November, 2014*

***Clabby Analytics is an independent technology research and analysis organization. Unlike many other research firms, we advocate certain positions — and encourage our readers to find counter opinions — then balance both points-of-view in order to decide on a course of action. Other research and analysis conducted by Clabby Analytics can be found at: www.ClabbyAnalytics.com.***