

HRG Assessment: SAP HANA IBM eX5 Appliance and DB2

Customers of real-time event driven always on businesses expect availability (fulfillment) of products and services with little wait time. Real time responsiveness gives the customer the impression that they have more control. As systems react to user input in real time customers feel empowered, appreciated, and better served. This type of responsiveness gives the business user greater decision making autonomy in responding to evolving customer requirements and competitive pressures while driving effective decision making in the enterprise down to the point of customer contact. However, this can have significant implications for the enterprise and supply chain. To successfully satisfy customer requirements and meet new demands placed on your organization and suppliers will require a comprehensive solution.

The SAP HANA IBM eX5 Appliance in combination with IBM DB2 is a comprehensive solution that delivers the real time event driven business results required when implementing market responsive event driven supply chain, ERP, and CRM functionality. This gives the business more control throughout the entire sales cycle from development and promotion to fulfillment as systems respond in real-time. Customers feel empowered and better appreciated. How much control you give customers depends on your business model, tactics, and strategy and is no longer limited by available techniques and technologies. Allowing clients to customize what they buy has significant and wide ranging implications for the entire enterprise and its end-to-end supply chain.

SAP HANA provides an evolutionary response to change by allowing organizations to rapidly analyze operational, customer behavior, and market data for competitive advantage and unique market insights. SAP HANA functions as a data repository or warehouse for SAP BusinessObjects™ and other solutions that require analytic speed and accuracy. During a recent SAP HANA on IBM hardware benchmark HANA processed 10,000 queries per hour running against 1.3 TB of data.

With DB2, SAP and SAP HANA customers gain real-time business intelligence and accelerated performance that translates directly to competitive advantage with substantial cost savings and administrative efficiencies. DB2 drives reduced storage costs by leveraging database compression in conjunction with fast backup and recovery capabilities. The combined IBM SAP HANA solution lets customers harmonize business and IT goals, strategies, tactics, and resources as change occurs.

The SAP HANA IBM eX5 Appliance in combination with IBM DB2 meet today's real time data requirements while addressing on-going budget and legacy investment constraints. In the highly competitive global business arena continually increasing data volume, variety, and velocity are the norm. Low transactional latency once the unique concern of the financial services industry is emerging as a "top of the list" issue for most on-line enabled businesses, and other highly competitive market segments requiring real time analysis and intelligence.

IBM eX5

IBM with its eX5 Servers brings to market a major solution building block element to meet continually increasing workload requirements. These eX5 compute infrastructure solutions in combination with Intel's hyper-threading, multi-core, extreme memory, E 7 processors satisfy the large memory, SMP, and parallelism requirements that occur simultaneously in many real time high volume (Big Data) application environments.

IBM eX5 server technology significantly advances the state of information technology and delivers significant benefits from the available memory capacity of IBM eX5 in combination with the improvements in transaction throughput enabled by the new Intel® E7 series processors. IBM customers will realize significant competitive advantage in faster ROI, reduced TCO, and improved time to result.

IBM silicon allows processors on eX5 systems to access extended memory very quickly. The IBM Enterprise X-Architecture chip is in its fifth generation with eX5 and leverages decades of IBM experience in integrating microelectronics to create first-of-a-kind silicon solutions.

IBM's eXFlash technology replaces older hard disk drive storage subsystems. eXFlash reduces storage costs by up to 97% and delivers up to 30x more local database performance for DB2. With the addition of IBM's Systems Director capabilities customers can pre-configure servers, remotely re-purpose systems and set up automatic updates and recoveries.

IBM eX5 SAP HANA Appliance

IBM's eX5 HANA appliance provides the computational throughput and very large memory capacity needed to make Big Data useful. The current business environment is characterized by on-going growth in the volume of raw and processed data as well as an insatiable demand for compute power. Growing competitive pressures and customer requirements drive the need for faster more comprehensive Business Intelligence and Analytical solutions. The quest for increased productivity, lower TCO, and faster ROI further characterize today's global business environment.

For an in-memory appliance, such as SAP HANA, a system's main memory, its capacity, and its performance play an important role. The Intel Xeon processor E7 family has a memory architecture that is well suited to the requirements of such an appliance.

IBM bases their Appliance offering for SAP HANA on the x3690 X5 and the x3950 X5 servers.

The IBM System x3850 X5 a 4U rack-optimized server forms the basis of the x3950 X5. These systems are purpose built for compute and memory intensive solutions like SAP HANA. The x3850 X5 and x3950 X5 both scale to four processors and 2 Terabytes (TB) of RAM. With the MAX5 attached, the system can scale to four processors and 3 TB of RAM. Two x3850 X5 servers can be connected together for a single system image with eight processors and 4 TB of RAM.

The IBM System x3690 X5 a 2 socket 2U server brings the benefits of eX5 features and performance to the IBM eX5 SAP HANA appliance. The x3690 X5 uses eXFlash internal SSD storage and all x3690 X5 configurations for SAP HANA use eXFlash for high IOPS log storage and data storage.

IBM General Parallel File System

IBM General Parallel File System (GPFS), a high-performance shared-disk file management solution, can prefetch data into its buffer pool, issuing I/O requests in parallel to as many disks as necessary to achieve the peak bandwidth of the underlying storage-hardware infrastructure. GPFS can be configured to use low-latency disks for index operations and high-capacity disks for data operations of a relational database and remain available in the case of a disk or server failure. In addition GPFS can be set to recover from node, storage, and other infrastructure failures without operator intervention.

IBM GPFS delivers the following benefits for the IBM Systems solution for SAP HANA:

- ✓ Provides a stable cluster-capable file system for SAP HANA.
- ✓ Enhances storage device performance through data striping across devices
- ✓ Supports the scale-out a single HANA system without adding external storage and presents this cluster as a single HANA image.
- ✓ Delivers automated (once configured) high-availability and disaster recovery capabilities.

The IBM Systems Solution for SAP HANA can be scaled up to increase the capabilities of a single system by adding more components, or scaled out using IBM GPFS to use multiple systems in a cluster as a single logical HANA system. This scale out solution is enabled by two redundant 10 Gb Ethernet networks with one allocated for cluster-internal communication of the SAP HANA software and the second network being allocated for cluster-internal communication of GPFS, including replication

SAP HANA Appliance

The IBM eX5 based SAP HANA appliance supports the analysis of large volumes of data in memory without the need to materialize aggregations. This appliance is a combination of hardware and software that is provided as an optimized platform for SAP HANA. The appliance is comprised of the SAP HANA database, repository files for the SAP HANA studio, and the SAP HANA client libraries for the connection of applications to the SAP HANA database. Typically the SAP HANA studio and client libraries are installed on a separate system. The Software Update Manager (SUM) framework enables the automatic download and installation of SAP HANA updates from the SAP Marketplace and other sources using a host agent and facilitates distribution of the Studio repository to users. In addition the Lifecycle Management (LM) Structure for SAP HANA which is a description of the current HANA installation is used by SUM during the automatic updates process as a point of reference.

SAP HANA Database

The SAP HANA database combines row-based, column-based, and object-based data formats and exploits the parallel processing capabilities IBM eX5 and underlying Intel E7 series processor capabilities and functions. SAP HANA is capable of both OLAP and OLTP workloads. At the core of the SAP HANA database is its relational database engine which supports both row-based and column-based data structures as dictated by data compression and other customer requirements. As interfaces the SAP HANA database supports SQL (JDBC/ODBC), MDX (ODBO), and BICS (SQL DBC). Additionally, the calculation engine supports in memory in the database Analytics, Business Intelligence, and a range of

calculations. Application functionality is loaded in memory rather than moving the data to the application layer. A business functions library is included that can be called by applications to perform business calculations. The SAP HANA-specific SQL Script language allows data-intensive application logic to be moved into the SAP HANA database for more efficient execution.

When SAP HANA is used as a data warehouse for new applications the consumer of the data replicated to SAP HANA is the source system. The source system can then use the SAP HANA database for analytics. Two examples of this are SAP Smart Meter Analytics, and SAP COPA.

SAP HANA can be used either as a side-car data warehouse or as the database for SAP NetWeaver Business Warehouse (BW) and Business-Suite (CRM, ERP, SCM) replacing the traditional database server of a SAP NetWeaver BW and Business-Suite (CRM, ERP, SCM) installation. In order to use SAP HANA as the database for SAP NetWeaver BW it is necessary have to installed v 7.30 of NetWeaver BW as well as version 7.30 SP5 SP03 of SAP HANA. Installation, migration of data, cost of the appliance, and SAP licenses need to be planned for. While SAP HANA is much faster than traditional database technologies that speed comes at a cost. When considering SAP HANA HRG recommends that an evaluation and assessment be done in order to ensure optimal results in terms cost, ROI, and data center impacts. While it is true that SAP HANA will ultimately help control or reduce head count it is equally true that making changes of any significance to a Data Center architecture requires more resources while the project is active. This requires planning and a formalized impact assessment.

When SAP HANA is used as a data mart (side-car) rather than as a database server it is necessary to populate the SAP HANA database with operational and or transactional data. This data needs to be fed into SAP HANA through replication from the from the source system. In this case the traditional database server, such as DB2, acts as the primary server. By using DB2 as the primary database server customers benefit from the performance enhancements provided by HANA without the cost of a major software upgrade. This method of deployment buys some time for those customers with previous generation SAP products and third party applications as they decide how to best integrate SAP HANA as part of their strategic IT direction. It should be noted here that DB2 supports most SAP software. For those customers that only want to use SAP HANA for part of their solution using DB2 in this manner offers the best of both worlds in terms of avoiding migration and software upgrade costs. The combination of SAP HANA and DB2 as a solution set leverages DB2's adaptive compression, multi-temperature storage, pureScale® capabilities, and Oracle compatibility features to provide excellent benefit in terms of controlling cap ex, reducing TCO, and increasing ROI.

Replication

The replication methods that can be used to feed data to SAP HANA are: 1) log-based replication (Sybase Replication Server), 2) Trigger-based replication using the SAP LT system, and 3) ETL (Extract, Transform, and Load). Using replication, data from third party (non-SAP) applications can be fed into SAP HANA.

In order to realize the true benefits of SAP HANA data needs to be replicated from disparate operational and transactional sources into SAP HANA quickly and efficiently. DB2 works directly with SAP HANA to provide automatic, real-time replication. SAP HANA establishes a real-time link with IBM DB2 and using Sybase Replication Server technology DB2 data is automatically replicated into SAP HANA and is immediately available for analysis.

Sizing

Using T-shirt sizes for SAP HANA simplifies sizing and limits the number of supported hardware configurations in order to reduce complexity and control quality. Through its tightly controlled SAP HANA appliance platform certification and software licensing process SAP ensures a more consistent customer experience.

IBM System x workload-optimized models for SAP HANA, 2012 models

SAP T-shirt size	XS	S	S+	M and M+	L Option
In-memory compressed data	64 GB	128 GB	128 GB	256 GB	512 GB
Server memory	128 GB DDR3	256 GB DDR3	256 GB DDR3	512 GB DDR3	1024 GB DDR3
Log Storage	10x 200 GB 1.8" MLC SSD (combined log and data)	10x 200 GB 1.8" MLC SSD (combined log and data)	1.2 TB High IOPS adapter	1.2 TB High IOPS adapter	1.2 TB High IOPS adapter
Data Storage			8x 900 GB 10 K SAS	8x 900 GB 10 K SAS	8x 900 GB 10 K SAS
Number of CPUs	2x Intel Xeon E7-	2x Intel Xeon E7-	2x Intel Xeon E7-	4x Intel Xeon E7-	8x Intel Xeon E7-
eX5 models	(x3690 X5) 7147-HAx or 7147-H1x	(x3690 X5) 7147-HBx or 7147-H3x or 7147-H2x	(x3950 X5) 7143-HAx or 7143-H1x	(x3950 X5) 7143-HBx or 7143-H2x	(x3950 X5) Combine 7143-HBx or 7143-H2x with 7143-HCx or 7143-H3x)

Licensing

SAP HANA for IBM eX5 systems is provided as an appliance from IBM. However, the HANA software license needs to be acquired from SAP. SAP HANA licenses are based on the amount of main memory for SAP HANA incremented in 64 GB chunks. This software is available in three editions: 1) The SAP HANA platform edition (basic), 2) SAP HANA enterprise edition, and 3) the SAP HANA extended enterprise edition.

The IBM SAP HANA appliance comes with the complete software stack, including the operating system, GPFS, and the SAP HANA software. Due to the nature of the software stack, and dependencies on how the IBM Systems solution for SAP HANA is used at the customer location, the software stack cannot be preloaded completely at manufacturing. Therefore installation services are required.

DB2 10 for Linux, UNIX, and Windows

IBM DB2 is an excellent database choice for SAP solutions that are deeply integrated with DB2 throughout the entire software lifecycle. Customers benefit from this integration with faster ROI, lower TCO and faster certification and implementation of new versions of DB2.

IBM and SAP have a long history of cooperative development built around IBM DB2 and SAP software integration which allows DB2 to seamlessly work with SAP application solution software including the new SAP HANA software platform. The net result of this integration is that users can implement HANA for those workloads where real-time analytics and BI provide real benefit to the business and use DB2 for the majority of business operation related workloads where real-time is not required. In this way customers can benefit from both offerings in a manner that constrains capital expenditures but does not constrain business benefits.

DB2 allows SAP customers flexibility in terms of the operating environments and hardware platforms they can use. This translates into legacy investment protection such that a customer can utilize the SAP HANA software platform solution on certified x86 Intel based systems running SUSE Linux to meet their real-time in memory analytics and BI requirements while protecting their previous technology investments through the tight functional integration between IBM DB2 and their installed base of SAP application solution software. Additionally, utilizing SAP software migration tools in an optimized manner means that, in most cases, SAP applications and underlying data can be quickly and easily migrated from Oracle to DB2 with few if any modifications. This in turn allows SAP customers to be able to take advantage of DB2's lower license costs and tight integration with SAP software. Ultimately this type of migration will facilitate the migration of appropriate data and applications to the SAP HANA software platform as more of the SAP software portfolio is enabled on HANA. In this manner customers can take advantage of SAP HANA real-time in-memory speed where most appropriate while leaving more traditional business operation workloads and data in place thereby delivering extreme performance while controlling costs.

IBM DB2 10 helps lower administrative, application development, and storage costs, while maintaining a highly reliable and highly available database environment. DB2 functionality helps to minimize development time and reduce development costs, because application developers no longer have to build this functionality into their applications. This functionality is described below:

Adaptive compression:

Adaptive compression is an enhanced compression technology that adapts to changing data patterns, yielding high compression ratios. DB2 10 Adaptive compression features page-level compression dictionaries and leverages previously available row compression technology found in earlier releases of DB2 for Linux, UNIX, and Windows. Page-level dictionaries automatically adjust to data changes on a page while data remains available as page-level compression dictionaries automatically adapt to changing data values over time. Customers are able to perform online and offline-table reorganizations as needed to ensure optimal compression ratios. However this is required infrequently compared to previous versions of DB2.

Multi-temperature storage:

Identify and manage data based on its "temperature" or access requirements in order to optimize storage costs. DB2's Multi-temperature storage functionality lets customers control their storage costs by placing dormant data in archive and their hottest data on Flash or SSD dynamically. Now with DB2 10 customers can match their data to the most appropriate storage tier based on user and application data access patterns and requirements.

IBM DB2 Near-Line Storage (NLS) solution for SAP NetWeaver Business Warehouse

Customer experience informs us through case studies that DB2 NLS reduces the size of the online DB2 database by up to 90% significantly improving SAP BW performance. NLS works at SAP BW application level to control the size of the online database by moving “aged data” to Near-Line-Storage. DB2 NLS effectively constrains and minimizes SAP HANA license and related HW costs because less data needs to be held in RAM.

Using the DB2 Near-Line Storage (NLS) feature “aged” data is stored in a database rather than moving it to an off-line archive like tape. Near-line storage of aged data allows a reduction of database size resulting in improved speed of performance for the online SAP BW system. Requiring less online database volume, the SAP BW performance gets faster and the required RAM capacity for SAP HANA is constrained

DB2 NLS works at the SAP BW application level. Once NLS has been set up to check the age of data based on business specific parameters it is transparent to the DB2 database and incurs no overhead at the database level. DB2 NLS for SAP BW helps control and effectively limit software license and related hardware costs.

Availability enhancements:

IBM DB2 pureScale on System x® leverages the IBM xArchitecture® and provides a cost-effective, open, reliable and secure IT environment that can easily scale as workloads demand. DB2 pureScale provides continuous availability in the event of a sudden hardware or software failure. Availability enhancements include; high availability disaster recovery (HADR) where multiple-standby databases provide availability functionality that is transparent to applications. DB2 10 offers support for up to three HADR standby databases. Establish a principal standby database tightly synchronized with the primary database for high availability, and allocate two auxiliary standby databases for disaster recovery. DB2 pureScale® provides continuous availability in case of unplanned maintenance or a catastrophic failure.

IBM® PureData System for Transactions in combination with DB2 pureScale and tightly integrated with SAP software solutions provide a highly reliable and scalable database appliance to reduce datacenter complexity, improve time to value, reduce the cost of data management, and constrain IT capital expenditures.

Oracle compatibility:

DB2 for Linux, UNIX, and Windows provides high levels of code compatibility for Oracle applications. In most cases Oracle applications can run on DB2 with no source code changes. DB2 SQL compatibility support for the Oracle database introduced with DB2 9.7 includes new scalar functions and expanded support for triggers and user-defined data types to facilitate migrating Oracle applications to DB2.

Conclusion

IBM's eX5 running the latest Intel® Xeon® E 7 series processors is purpose built to maintain high through put, provide large memory capacity, and access to high quality data for business intelligence and analytics. High end IBM eX5 systems with expanded memory and high speed solid state disk drives are ideal solutions for meeting emerging high-throughput large data analytics and BI customer requirements.

The IBM eX5 SAP HANA appliance is a viable alternative to the Oracle Exalytics In-Memory Machine while DB2 presents a sound alternative to the Oracle database. With the combined IBM eX5, SAP HANA, and DB2 solution SAP and IBM are well positioned to harvest Oracle's installed base.

Cloud Computing will increasingly be utilized for the analysis of data to facilitate access to advanced BI and analytics capabilities without requiring significant investments in application development or compute technology. The IBM eX5 SAP HANA appliance is an excellent choice as a delivery platform for business Intelligence and Analytical Cloud Computing Service delivery.

HANA and DB2 each have their own unique complimentary advantages. HANA brings real time speed and predictable performance based on a rigorously certified solution comprised of Intel E 7 series processors, IBM eX5 server technology and SUSE Linux. DB2 is less expensive than HANA and delivers excellent TCO and ROI benefit. DB2 is available on a wide variety of hardware platforms and operating environments. Using DB2 with HANA allows customers to benefit from HANA's real-time in memory analytics while controlling costs by using DB2 for the majority of SAP applications where real-time is not a requirement.

Harvard Research Group is an information technology market research and consulting company. The company provides highly focused market research and consulting services to vendors and users of computer hardware, software, and services. For more information please contact Harvard Research Group:

Harvard Research Group™

PO Box 297
Harvard, MA 01451 USA

Tel. (978) 456-3939

Tel. (978) 925-5187

e-mail: hrg@hrgresearch.com

[*http://www.hrgresearch.com*](http://www.hrgresearch.com)