

HEALTH INFORMATION SYSTEMS MOBILIZED BY NOSQL SOLUTIONS

MARKLOGIC WHITE PAPER • JANUARY 2016

Moving the system of record from legacy mainframe to Enterprise NoSQL
on Intel®-based solutions



THE OPPORTUNITY FOR COST EFFECTIVE TRANSFORMATION

Big Data increasingly impacts decision-making throughout the healthcare industry. Enterprise NoSQL database solutions enable patient- and consumer-centric care by integrating all relevant information in a single, searchable database. Healthcare providers who increasingly rely on Electronic Medical Records (EMRs) can now have real-time access to everything from lab tests to physician notes to academic journals.

Though mainframe architecture served the data, warehousing, transaction, and system of record needs of health services organizations for an impressively long time, it cannot support this enhanced level of access and usability.

As service providers, clinicians, and payers look to the healthcare information systems of the future, continued utilization of the legacy mainframe presents a costly and time consuming hurdle. Although maintaining the current installed system may seem to be the most affordable decision for the immediate financial cycle, analysis of the insidious increases in maintenance costs along with lost opportunities build a strong case for migrating sooner rather than later.

DATA IS THE NEXT COMPETITIVE EDGE

Big data is the topic of conversation across all industries, and for good reason. As far back as 2011, a McKinsey & Company report found that the leaders in big data usage consistently outperformed the rest of their industry in both revenue and earnings¹. In a 2015 survey of executives from large global companies, 90 percent had medium to high investment in big data. Two-thirds reported that big data and analytics have had significant, measurable impact on revenues².

¹ "Big Data, The Next Frontier for Innovation, Competition, and Productivity," McKinsey Global Institute, 2011.

² Press, Gil, "6 Observations from a New Survey on the State of Big Data Analytics," Forbes.com, September 4, 2015. Found at <http://www.forbes.com/sites/gilpress/2015/09/04/6-observations-from-a-new-survey-on-the-state-of-big-data-analytics/>.

This trend toward Big Data solutions has led to more and more requests from physicians, medical researchers, and regulators to conceive of information in a new way. This new way of looking at care delivery can stretch the resources of the mainframe. In many cases, the task is entirely beyond the capabilities of the existing system of record.

Modern healthcare delivery systems require additional patient and consumer data capabilities. The new standard expects flexible, inclusive, and distributed data.

THE GROWING COST OF DOING NOTHING

Future opportunity can be hard to quantify, but current costs are a serious and measurable issue. While there is a short-term cost to adopting a new technology capable of flexible and creative data implementation, the status quo will be increasingly more expensive. The maintenance of a legacy mainframe is becoming costlier by the year.



As an organization's legacy system ages and widespread use of mainframe technology fades further into history, it will cost more and more to maintain the same level of service.

As the skillsets to maintain the equipment and codebase sunset, programmers and administrators are more difficult to find and increasingly expensive. Following a major data breach, the United States Office of Personnel Management released its 2014 Strategic Information Technology Plan. Along with concerns about security, the report prioritizes migrating away from the legacy mainframe systems because cost of support is expected to rise 10–15 percent

annually³. And this doesn't even take into consideration opportunity costs of not being able to use mainframe data in new big data solutions.

THE ALTERNATIVE: DISTRIBUTED DATA CENTERS

While the administration and operation costs of legacy mainframes, including maintenance and new applications, are getting higher, the capabilities and return on investment (ROI) of Intel-based alternatives increase. The Intel-based data center solution ecosystem is being revolutionized by innovation in management, energy efficiency, and performance. Modern Intel-based data centers are able to run more services per staff hour, megawatt, and clock cycle than ever before.

INTEGRATED DATACENTER TECHNOLOGY

Intel storage, network, and compute technologies are increasing the productivity and performance of data centers.

- **Scalable Processors Designed for Data Centers**
For example, the Xeon E5-2600 v3 product family, which is designed for the workloads of a data center, includes numerous enhancements that provide performance increases up to three times that of the previous generation⁴, world-class energy efficiency, and enhanced security.
- **Intelligent and Efficient Storage** – The Intel SSD DC P3608 Series solid state drives (SSD) deliver high-performance and low latency. Non-Volatile Memory Express (NVMe) and eight lanes of PCIe 3.0 eliminate bottlenecks in high performance computing workflows, accelerate databases, and provide business insights through real time analytics. The unique dual controller architecture scales efficiently with multiple core Intel Xeon processors.
- **Integrated Network Systems** – With 1 billion Intel Ethernet ports shipped and more than 30 years of

innovation, Intel offers a comprehensive portfolio that includes controllers, adapters, IP, and switch silicon. Since the inception of Ethernet, Intel has continued to help drive Ethernet technology by providing technological advances and active leadership in standards committees and industry associations.

NOSQL: SKIPPING ETL AND SAYING YES

For decades, the art of data has been designing database schemas and the craft of extract, transform, and load (ETL) processes. Standardizing available patient-generated information, physician inputs, and other fields is traditionally a key part of any new data integration or consolidation project on a legacy mainframe healthcare system. Mainframe data often needs to be copied into a different application-centric database or data warehouse, requiring an ETL process.

ETL projects need to be vetted for accuracy. The content and process need to be kept up-to-date. ETL can be resource intensive, both for staff and compute. For the agency, ETL is cost, delay, and constraints.

The transform portion in particular is resource intensive and complex. In the transforming of data, some aspects of the information can be lost or compromised. For instance, if a building height is calculated using different parameters in different databases and those need to be stored together, the variances might be lost.

When an organization adopts a NoSQL approach to data management, the database imports content in its native state. The ETL process is replaced by the significantly less involved “extract and load.” Sometimes termed “schema-agnostic,” Enterprise NoSQL document databases do not require extensive design, planning, or architecture before introducing a new data source.

PATIENT-CENTRIC CARE RELIES ON EMERGING BIG DATA APPLICATIONS

With modern processing and data tools, healthcare providers are discovering new value in their existing datasets. They create more patient-centric care by combining information from disparate data sources, often from partners and resources outside of the organization. Enterprise NoSQL is a compelling enabler of future data applications. An opportunity that requires

³ Strategic Information Technology Plan, U.S. Office of Personnel Management, 2014.

⁴ As of September 8, 2014. New configuration: Hewlett-Packard Company HP ProLiant ML350 Gen9 platform with two Intel Xeon Processor E5-2699 v3, Oracle Java Standard Edition 8 update 11, 190,674 SPECjbb2013-MultiJVM max-jOPS, 47,139 SPECjbb2013-MultiJVM critical-jOPS. Baseline: Cisco Systems Cisco UCS C240 M3 platform with two Intel Xeon Processor E5-2697 v2, Oracle Java Standard Edition 7 update 45, 63,079 SPECjbb2013-MultiJVM max-jOPS, 23,797 SPECjbb2013-MultiJVM critical-jOPS <http://www.spec.org/jbb2013/results/res2014q1/jbb2013-20140121-00050.html>.

MarkLogic / Enterprise NoSQL Database Platform

POWERFUL	AGILE	TRUSTED
<ul style="list-style-type: none"> Native JSON Store Native XML Store Native RDF Triple Store Geospatial Support Full-text Search Flexible Indexes Bitemporal Real-time Alerting Semantic Inference Tiered Storage Server-side JavaScript Fully Transactional 	<ul style="list-style-type: none"> Scalable and Elastic Cloud Ready (AWS) Hadoop and HDFS REST API SQL Support Multi-OS Support Schema Agnostic Samplestack MarkLogic Content Pump XA Transactions Ad-hoc Queries Index Across Data Types 	<ul style="list-style-type: none"> Performance at scale LDAP and Kerberos Security Security Certifications Monitoring and Management Configuration Management 24/7 Engineering Support ACID Transactions Flexible Replication Customizable Backup Customizable Failover Point-in-time Recovery Atomic Forests

Enterprise NoSQL is a trusted platform that cost-effectively responds to modern expectations.

the consumption of JSON, XML, unstructured files, text, RDF triples, geospatial, or any binary can be implemented rapidly without the pre-work of ETL and schema design.

A million heterogeneous documents can be corralled to answer a question like “How effective was this year’s flu vaccine in the metro area?” No matter the format all participants in the health services industry use, data and documents from a variety of sources can be combined to maximize service availability.

ENTERPRISE NOSQL SERVING PATIENT 360°

Designed for durability and scalability, Enterprise NoSQL technology from MarkLogic supports massively scaled applications around the world. MarkLogic® Enterprise NoSQL is the platform for organizations in need of a system of record capable of high volume and high flexibility. Many companies are already realizing the benefits of MarkLogic Enterprise NoSQL.

TOP 5 BANK

32 million live deals	1,600 requests per second
-----------------------	---------------------------

BBC STREAMING 2012 OLYMPIC COVERAGE

55 million global browsers	2.8 petabytes
----------------------------	---------------

HEALTHCARE.GOV

5,000 highly complex transactions a second	11+ million Americans with health coverage
--	--

CONDÉ NAST

7+ million rich media assets	<1 second search results
------------------------------	--------------------------

“ With MarkLogic, it is taking one hour to stage content, a few minutes to push live, and less than a second to run queries and analysis. This is a tremendous coup for our business as well as our technology.”

- Beverly Jamison, Senior Director, IT Architecture and Publishing Solutions, American Psychological Association

CASE STUDY: AMERICAN PSYCHOLOGICAL ASSOCIATION MOBILIZES ITS DATA

With over 150,000 members to serve, the American Psychological Association (APA) website gets more than a million visits per day. Its legacy mainframe could not adequately correlate the 160,000 articles and 3 million abstracts for researchers to do their jobs properly.

The APA migrated to a MarkLogic NoSQL database solution—speeding response times, improving reliability of results, and reducing the number of servers from ten to three.

With MarkLogic, it’s easier for the APA development team to add features and create better workflows.

DISTRIBUTED AND FLEXIBLE DATA CENTER ARCHITECTURE

When dealing with multiple service providers, new regulations, and changing patient needs, healthcare organizations need to be able to affordably grow and transform IT infrastructure. With a mainframe, one can add processing and storage, but adding secondary locations is not so easy. Architecting for bursts and peaks means essentially purchasing compute for the maximum need.

Organizations adopting Intel® Xeon® processor-based data centers can manage growth and opportunity using modern flexible computing techniques. Virtualization of servers and infrastructure allows for an unprecedented ability to maximize resource use.

Bursts in demand and new initiatives mean that an IT organization needs to be able to cost-effectively increase compute capacity, perhaps for only a brief

period of time. Applications developed on Intel-based solutions have options. Along with reallocating resources within the data center, hybrid cloud solutions can extend the application to a platform as a service (PaaS) or an infrastructure as a service (IaaS) vendor. MarkLogic is architected to handle elasticity and new distribution deployments without downtime.

ENTERPRISE NoSQL + INTEL-BASED SOLUTIONS

- **Are ready for new integrations**
- **Are designed for distributed workloads**
- **Scale up and down elastically**
- **Lower management costs**
- **Lower development costs**
- **Lower TCO costs**
- **Are architected for high availability**

MIGRATING THE SYSTEM OF RECORD FROM A LEGACY MAINFRAME TO NOSQL

Achieving the promise of a new technology usually means a migration process. Migrating a system of record, the heartbeat of an organization, is no small task and not one to be taken lightly. As the market moves away from mainframe solutions, actionable guidance and help is available to the organization that embarks on a migration path.

Intel offers best practices for the migration planning and execution in its paper “Migration from UNIX/RISC and Mainframe to Intel-Based Solutions.”

HSBC RE-HOSTS MODERN MAINFRAME APPLICATIONS

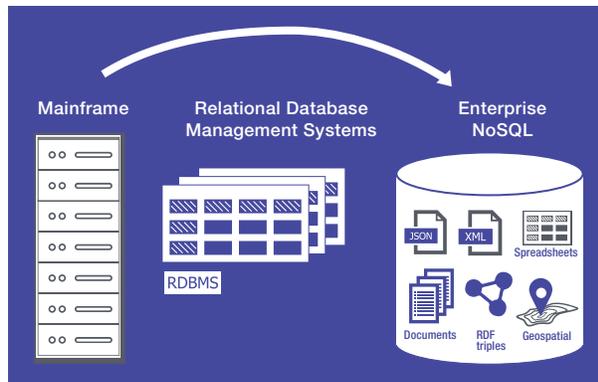
HSBC encountered performance and reliability issues for a key mainframe after deploying a new suite of loan applications. Migrating the applications to Intel Xeon processor-based servers delivered:

- A better experience for customers and branch office personnel due to improved performance and higher uptime
- Lower costs and greater headroom, with 70 percent lower monthly service charges and a 2,000 MIPS reduction in mainframe workloads
- Improved business flexibility with an infrastructure that is easier to scale and adapt

Read the full case study for more information about the HSBC methodology and migration experience at <http://www.intel.com/./data-center-xeon-hsbc-whitepaper.pdf>.

For more guidance on migration to an Intel architecture data center, visit the Intel [IT Center for Server Migration](#)

LEAPFROG RDBMS: ENTERPRISE NoSQL IS A MORE COMPELLING DESTINATION



Migrating to Enterprise NoSQL provides greater capabilities and a less disruptive migration path than RDBMS.

The legacy mainframe applications that are still running today have most likely survived despite previous

challengers. Throughout the years, IT departments have evaluated migrations to relational database management systems (RDBMS). Many times, they chose to leave the mainframe in place.

Since the 1990s, relational databases have become the expected database solution for most data projects. Still, there is a huge base of applications still running on thousands of mainframe systems around the world. They were not migrated for very good reasons: migration is costly, disruptive, and risky.

Enterprise NoSQL, along with offering extended capabilities, has a safer, less intensive migration path and offers greater data agility to ensure the data is future-proofed for the next 40 or more years. With a more flexible information storage, less transformation means less time and less risk.

IT CAN BE DONE; IT HAS BEEN DONE: COPYBOOKS TO XML

While data is stored and managed differently in a modern NoSQL database, the XML-friendly structure allows for some logical migration paths. Ken Krupa, MarkLogic CTO, utilized the COBOL from his early career to [document a process](#) that will reliably ingest COBOL copybooks into the MarkLogic Enterprise NoSQL database.

MEASURING SUCCESS

Ultimately, success of all health care IT is the more efficient delivery of better patient care.

Enterprise NoSQL now enables the patient- and consumer-centric healthcare of the future. Providers access more usable, reliable patient information to improve outcomes. Organizations leverage a unified platform to reduce the size and complexity of infrastructure. Payers access integrated, heterogeneous data to reduce fraud and better identify best practices. For the future, an organization can realize the value of an integrated data repository that provides access to all aspects of the organization's data profile. An Enterprise NoSQL solution is designed to quickly integrate new data sources as they appear, no matter the format or source.

For the IT department and executives, the financial bottom line is an obvious gauge of success. Reducing the cost of support and of developing new applications can be a path to realizing return on investment for the migration efforts.

ABOUT MARKLOGIC

For more than a decade, MarkLogic has delivered a powerful, agile, and trusted Enterprise NoSQL database platform that enables organizations to turn all data into valuable and actionable information. Organizations around the world rely on MarkLogic's enterprise-grade technology to power the new generation of information applications. MarkLogic is headquartered in Silicon Valley and has offices throughout the U.S., Europe, Asia, and Australia. For more information, please visit www.marklogic.com.

ABOUT INTEL

Intel (NASDAQ: INTC) is a world leader in computing innovation. The company designs and builds the essential technologies that serve as the foundation for the world's computing devices. As a leader in corporate responsibility and sustainability, Intel also manufactures the world's first commercially available "conflict-free" microprocessors. Additional information about Intel is available at newsroom.intel.com and blogs.intel.com, and information about Intel's conflict-free efforts is at conflictfree.intel.com.