



HRG Assessment

Vblock™ Unified Data Solution for HealthCare

Next generation and leading HealthCare organizations and institutions need to facilitate data sharing among professionals to expedite diagnosis and effective treatment. HealthCare providers looking for solutions to address their current requirements should remain mindful of what the future may hold in store. The increased use of remote monitoring devices to collect and transmit patient data (regardless of location) to healthcare professionals contributes significantly to the growth in HealthCare related Big Data. This telemetric healthcare data enables care givers to monitor and track individuals and enables early intervention for conditions as they worsen.

Key drivers for the growing use of remote, mobile, and embedded telemetric healthcare devices are:

- Aging populations
- The predicted increase in patients with chronic conditions
- Continually increasing demand for medical services
- The ongoing shortage of physicians and qualified clinicians
- The evolution of, and easy access to, on-line technologies in health care.

The use of remote, mobile, and embedded telemetric healthcare devices improves access to care, for elderly and rural patients. The effective use of these devices helps to reduce the cost of care by reducing the number of hospital readmissions and associated transport costs while improving the access to, and responsiveness of care givers. It is now the norm for patients to communicate with doctors, nurses, or caregivers using on-line chat, a web portal, or email. This in turn continues to drive increases in the volume of EHR and other patient centric data.

VCE's VCAP (Vblock™ Converged Analytics Platform) HealthCare solution can be used to improve the interoperability of dissimilar local, state, federal and international EHR systems as a result of EDMT's unique ingest and metadata index capabilities. Ultimately the VCAP HealthCare Solution could be used to match a patient's health records across the complete gamut of settings in which care and services are provided. This allows clinicians, members of a care team, or even patients to view the complete picture of an individual patient's care requirements. HealthCare management and delivery mechanisms will continue to evolve driven largely by individual use of mobile smart phone and remote monitoring technologies to manage, and report real-time health information and data.

Unified Data Strategy

What is Unified Data? Unified Data is the combination of Big Data and Enterprise data (sometimes called "Little Data") that has been ingested, indexed, and stored in a searchable and query-able enterprise-class database. Unified Data can then be cross-searched and cross-analyzed using currently available data analysis tools and techniques, as well as, SQL, "R" and other in-database analytics tools and methods.

Things to consider when developing a Unified Big Data Strategy are:

- Data Variety – structured, unstructured, and real time
- Data Volume – trillions of transactional records, “Internet of Things”, and files
- Data Velocity – the ability to ingest and store the volumes of data at the rate required
- Choice of database – data should be stored in an ACID-compliant repository
- Running analytics and searches on (and across) structured and unstructured data from multiple sources

The effective use of Unified Big Data and VCE’s VCAP HealthCare Solution holds the potential for improving care delivery, individual health and wellness while helping improve policies and remove barriers related to the delivery of health care spanning local, state, national, and international boundaries.

HealthCare Data Sources

Based on recent discussions HRG has learned firsthand from HealthCare professionals and their IT counterparts that they are becoming increasingly aware of the benefits of having access to Unified Big Data specific to HealthCare as they strive to improve care, improve outcomes, decrease cost, increase efficiency, and deliver a reasonable return on investment (profit). With the introduction of solutions like VCE’s VCAP Solution for HealthCare no data, regardless of type, location, or format is beyond integration. As HealthCare providers continue to collect increasing volumes of data from their operations, patients, websites, and other sources keeping track of that data and analyzing it to provide meaningful and actionable information will provide significant benefits for patients, providers, and payers. Some of the sources of Big Data are:

- The continually increasing volumes of Health and Wellness related data and transactions originating from the healthcare “Internet of Things” comprised of smart phone health and wellness apps, embedded devices with monitoring capabilities, wearable activity and vital sign monitoring devices, and a growing number of other devices and appliances.
- Social media sites used for sharing ideas, opinions, images, documents, related to providers, clinicians, treatments, insurers, and such.
- Emerging technology developments facilitate organizational change, drive productivity, increase creativity, speed new healthcare product and services development, and create data
- Patient – provider, web portal, e-mail, SMS based interactions, as well as, device-to-device transactions all create data

VCE’s VCAP Unified Data Solution for HealthCare can help facilitate the effective diagnosis and treatment of illnesses when symptoms are first presented in a clinical or remote setting. This solution can be used to help identify global syndromic health threats and developing pandemics (like the recent Ebola outbreak) as they emerge. Additionally, this solution can be used to help inform first response teams for most effective intervention, containment, and treatment thus avoiding untold suffering and unnecessary deaths.

Leading HealthCare providers need to become actively involved and encouraged to engage and participate in pilot project programs designed to enable this type of next generation use of Unified Big Data for HealthCare toward the goal of delivering lower cost, more effective, and more responsive health services.

VCAP now and in the future:

The current challenge for HealthCare Provider Information Technology executives is to ingest, integrate, analyze, and administer all available health data (clinical, historical, genomic, and more) from existing and ongoing health records from past 36 months or longer as appropriate regardless of form, format, or location of data. Once the requisite data has been ingested and cleaned the data could be made available via on-line cloud based applications and application interfaces in such a way as to ensure near real time assessment of individuals presenting symptoms to practicing clinicians and physicians at the time of observation. Using Unified Big Data and data analytics tools the practitioner will

be able to compare an individual's symptoms against all known and similar cases in the integrated data. This will allow individual practitioners to more rapidly and effectively identify underlying illnesses and to prescribe the most effective therapies.

Looking to the future of HealthCare ideally a Unified Big Data HealthCare Solution like VCE's VCAP Solution for HealthCare will be used to capture and make available to practitioners all known and available traditional and non-traditional therapies that have proven to be effective in the treatment and cure of specific illnesses. More specifically such a solution will:

- Identify all appropriate and relevant data sources and determine the best way to capture that data.
- Require an information technology architecture that will scale to fit future workload requirements as demand increases.
- Identify and make available, either directly or through APIs, all best in class and most appropriate data analytics and clinical assessment tools
- Incorporate the collection of available medical data from Smart Phone apps and medical telemetric devices like EKG, heart rate, blood pressure, blood sugar, and other such embedded and wearable monitors.
- Ensure total security and no risk of personal identifier data being accessed and misused – read as completely hacker proofed
- Ensure that the data and affiliated applications and tools are available 24x7.
- Capture and analyze current medical processes and combine with patient care data for total individual and population based process and outcome analysis, optimization, and improvement.

Advances in the state of Life Sciences research applications such as Bioinformatics, Genomic Research, and Translational Medicine are driving corresponding workloads such as genomic sequencing, genome segment assembly, segment alignment, and secondary data analysis. Each of these workloads and many more gain significant benefits from the scalability and flexibility of VCE's VCAP Solution. VCAP brings extreme data ingest speed and capacity for mixed workload multiple source and format BigData Analytics. HealthCare organizations will realize significant benefit through faster ROI, reduced TCO, and improved time to result.

While today's institutions and organizations conducting the full genome sequence experiments do provide primary analysis, they may not do the secondary analysis. It is this focused secondary analysis that looks for the very small percent of DNA variants which is the basis for research discoveries and insights. This is a precursor for the development and delivery of clinical applications to identify, diagnose, prevent and cure disease. This area of secondary data analysis is emerging as the area where right sized compute and data intensive solutions like VCE's VCAP Solution can significantly address the computational, data handling, and analytical requirements needed to make genomic data useful to Life Sciences companies, physicians, and patients.

Genome Sequencing – A Key Data Source

Whole or full genome sequencing of the DNA bases or segments that constitute each human genome produces very large volumes of data which then has to be stored for secondary processing and analysis such as the identification of genetic markers which serve as indicators for a specific disease. This type of secondary analysis requires significant high performance compute power, memory, and storage capacity.

High throughput full genome sequencing is becoming a reality with the development and ongoing refinement of technologies such as the nanopore "DNA transistor". The promise of nanopore based genomic sequencing is to sequence whole strands of DNA, dramatically increase sequencing throughput and accuracy and move the cost point of sequencing a single human genome to below \$100 per genome. The result will be a significant increase in the volume of available genome sequence data. The population of a human genome data repository will dramatically improve the identification of the genetic and proteomic basis for disease before a disease presents itself. This in turn will facilitate proactive prevention and treatment through life style counseling, personalized medicine, and even unique custom personalized drug therapies. The patent for Nanopore technology is held by Harvard University and Oxford Nanopore Technologies.

Personalized Medicine – The Future of HealthCare

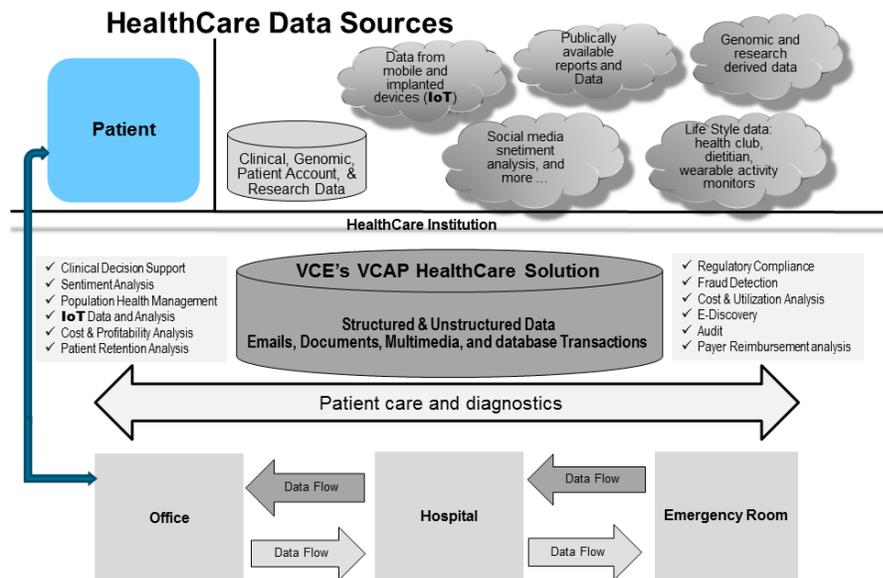
In HRG’s opinion genomic sequencing is a prerequisite to the development of true personalized medicine which will be based on the combination of data derived from patient clinical history, personal history, family history, and genomic sequencing. Reducing the cost of sequencing a complete human genome is a necessity for the field of personalized medicine to open up. As the repository of genome sequencing data is more completely populated researchers will be able to analyze anomalies or differences in one genome as compared to the base level mapping of the entire human genome. Through this type of analysis researchers will be able to identify and decipher disease specific genetic markers resulting in individual / personalized preventative and therapeutic medical applications. The field of personalized medicine will not realize its full potential until these genomic and clinical history data repositories have been established and populated. Security, privacy, disaster recovery, availability, and other emerging concerns will have to be addressed in order to pave the way for fulfilling the promise that personalized medicine holds.

Personalized medicine will be based on the combination of personal, historic, clinical, and genomic data in a single source. This could take the form of a machine readable card, subcutaneous RFID chip, or such device that an individual could carry with them. This device will have the individuals’ unique genome encoded along with other relevant personal information. When this individual goes to a pharmacy, for example, a pharmacist will be able to cross check the encoded genome, clinical historical data, and existing prescriptions in order to ensure that any new prescription does not conflict with existing conditions or treatments. This type of information when available will be a game changer in terms of enhancing the physician's clinical therapeutic efficacy.

Unified Data Solution for HealthCare

VCE with their Vblock™ Systems provide a solution that meets all of the big data challenges head on. The Converged Analytics Platform solution consists of the Vblock™ System 300 and 700 families combined with BMMsoft EDMT and SAP IQ. The Unified Big Data challenge is to ingest, index, archive, manage, analyze, and unify high volume, high velocity, structured, unstructured, static, and streaming data from a continually increasing variety of sources and locations.

Workloads and applications that benefit as a result of implementing a VCAP Unified Big Data HealthCare solution (see diagram – HealthCare Data Sources) are: Clinical Decision Support, Sentiment Analysis, Population Health Management, IoT Data and Analysis, Cost & Profitability Analysis, Patient Retention Analysis, Regulatory Compliance, Fraud Detection, Cost & Utilization Analysis, Risk Management, e-Discovery, and Audit (both financial and regulatory compliance).



Vblock™ Converged Analytics Platform for Unified Data

The Vblock™ Converged Analytics Platform Solution is an integrated platform of compute, storage and connectivity infrastructure for ingesting, indexing, and storing Unified Data. With the Vblock™ Solution customers can run searches and queries against the indexed unified columnar data store.

Vblock™ Systems deliver the efficiency and business agility of virtualization and cloud computing, integrating industry leading compute, network, and storage technologies.

Vblock™ Systems are designed, engineered, manufactured and certified to leverage storage technologies from EMC, networking and compute from Cisco, and, optionally, VMware virtualization software. Cisco UCS servers and Nexus switches support innovations like Unified Fabric, embedded management, and policy based computing.

The Vblock™ Converged Analytics Platform solution, a services based solution, uses EMC’s VNX storage for the Vblock™ 300 series platforms, and EMC’s VMAX product for the Vblock™ 700 series platforms. EMC’s VNX provides Fully Automated Storage Tiering (FAST) for the efficient and intelligent use of various data storage sub-systems.

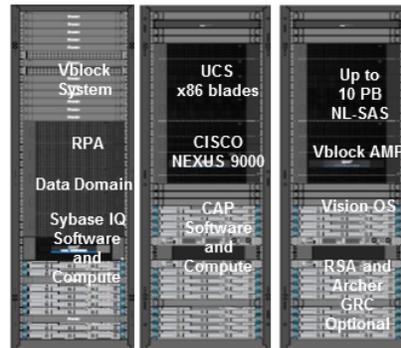
The Vblock™ System with the Converged Analytics Platform solution is designed to handle mixed enterprise class application workloads and benefits significantly from Vblock’s integrated operational, management, and administrative simplicity.

VCE Vision™ intelligent operations provides the systems management for Vblock™ Systems, optimizes services for a converged infrastructure, integrates directly with VMware technologies, and provides an extensible API for leading management tools. VCE Vision’s features include virtualization optimization, converged operation and an open management API.

VCE Vblock™ Systems are delivered fully configured 45 days from the time of a customer order and they are fully operational within 48 hours of delivery. All Vblock™ physical and logical builds are completed and tested to ISO quality standards at the factory prior to shipment to the customer.

VCE Single Point of Contact Support includes a built in problem escalation process that seamlessly assigns any unresolved service tickets to the appropriate Cisco, EMC or VMware engineering support personnel. VCE told HRG that their first line support staff typically resolves 96% of calls without having to engage personnel from Cisco, EMC or VMware.

VBLOCK™ CONVERGED ANALYTICS PLATFORM (CAP)



VBLOCK™ Enterprise Class Solution

Delivering Maximum Scalability & Flexibility	Vblock™ 340	Vblock™ 720
Compute	3840 cores	11520 cores
Compressed Data Capacity	21 PB	35 PB
Queries/Day	20+ M	60+ M
Ingestion/Day	12 PB	36 PB



BMMsoft EDMT® Unified Data Solution

BMMsoft's EDMT® Solution ingests and indexes structured and unstructured data using up to 3 million independent data processing channels that ingest this data into a single analytic SAP IQ data repository. EDMT® automatically creates and stores metadata for the ingested content. The EDMT Universe connector for BusinessObjects, lets BusinessObjects users access, view, and analyze all of the structured and unstructured data in an EDMT archive.

EDMT stands for Emails, Documents, Multimedia, and database Transactions. EDMT® in combination with SAP IQ 16 is at the heart of the Vblock™ Converged Analytics Platform (VCAP). This converged solution delivers the performance, reliability and scalability needed for real-time ingesting, indexing, accessing, and cross analysis of extreme volumes of heterogeneous Big Data. In recent customer test, VCAP loaded 23 PB (23,000 TB) of mixed data using only 42 ETL channels at speed of 3 PB per day. During the same test, over 100 billion documents, files and emails were loaded – in addition to over 350 Trillion database and sensor records. EDMT was configured to use external storage to store actual files/images of those 100 billion files (whose metadata were loaded in EDMT/IQ) in external file systems i.e. EMC Isilon or WORM – or over 100 PB of external objects fully managed, versioned and searchable by VCAP.

EDMT supports enterprise-wide searches across disparate mixed data types. Virtually any type of electronic data regardless of format or location can be ingested, indexed, and queried in close to real time. When ingesting streaming data EDMT does not perform the extract and transform operations which means that even higher rates of data ingest can be realized making ingested data available for search, query, and analysis in close to real-time with ingest latency as low as 20 milliseconds. Using EDMT normal business level queries typically return results in the sub second range while highly complex queries can return a result in from 1 to 300 seconds.

Both EDMT and SAP IQ use column data stores providing significant benefits in terms of data compression, improved query/analytics speed, and better data security. EDMT's use of Multiplexing is an excellent fit with the SAP IQ technology as both benefit from the full SAN connectivity between EDMT, SAP IQ and storage.

EDMT and SAP IQ comply with the ACID rules for database architectures providing high-quality data that is critical for enterprise applications and reliable verifiable analytic results based on clean and complete data. EDMT uses unadulterated or “accent-free” SQL that ensures precise data definition, ACID compliant data relationships, and precise query results.

Using EDMT businesses can run unified real-time text analysis, SQL analysis, and cross-analysis of mixed data types. With the EDMT Solution real-time monitoring, instant cross-analysis of new and historical data and real-time response to market changes, product problems, customer dissatisfaction, litigation, audit, fraud threats, and competitive threats can all be handled from a single system with no impact on production systems. EDMT can enforce data retention policies to meet EIS and regulatory compliance requirements while leaving the source data in its original state. Furthermore, EDMT is unique in that it is compatible with all other enterprise class SQL based applications.

EDMT's high availability features support the implementation of multi-site active-active disaster recovery sites. Remote replication to multiple sites ensures that problems with one or more replication targets or channels will not impact replication. According to BMMsoft the latency between the primary site and the secondary site(s) can be configured to be as low as 2 seconds.

SAP IQ 16

SAP IQ, a column oriented database, provides significant storage compression, query speed and performance advantages when compared to traditional row oriented databases. Column oriented DBMS outperform traditional row-based database management systems running analytic workloads on average by a factor of 100 times.

With SAP IQ 16, SAP introduced a native MapReduce API, Hadoop integration, Predictive Model Markup Language (PMML) support, and an expanded library of statistical and data mining algorithms that leverage the power of distributed query processing across a Massively Parallel Processing (MPP) grid based on SAP IQ's Multiplex technology.

New SAP IQ APIs enable the implementation of proprietary algorithms that run in-database. SAP claims that running proprietary algorithms in database delivers greater than 10 times the performance acceleration as compared to existing approaches. Additional improvements have been made for text data compression and bulk data loading interfaces.

With the release of SAP IQ 16 SAP introduced In-Memory Row-Level Versioning Store extended to IQ Multiplex for scale out IQ deployments. Point-in-Time Recovery restores an IQ backup plus all committed database operations to a specified point in time. An OData (Open Data Protocol) Server that lets web clients communicate with an SAP IQ database server using the OData interface. SAP IQ Cockpit, a new graphical administration tool for on-board management and monitoring of SAP IQ, is the first phase in the evolution of SAP Control Center towards enterprise scale administration and monitoring of all SAP database technology.

The SAP IQ and BMMsoft EDMT Multi-node designs are very tightly coupled. The SAP IQ shared disk / shared database is made possible because SAP IQ is not a partitioned database. With SAP IQ you can add storage without being required to add additional servers just as you can add servers without having to add storage. *(It is worth reiterating that with SAP IQ 16 storage can be added independently from servers and vice versa which is NOT possible with Hadoop.)* This architecture allows the addition of servers of any size to the SAP IQ grid.

EDMT and SAP IQ both use MPP "shared disk" architecture. Adding or removing servers ("nodes") with SAP IQ is straight forward not requiring data reorganization or re-partitioning. A single query in either EDMT or SAP IQ grid can be spread across all nodes, or a predefined subset of all nodes, or use just one server. This functionality provides substantial flexibility in terms of workload management and user isolation. All available nodes can be regrouped within 1 second (according to BMMsoft) to provide support for a big query.

The nonstop functionality of both SAP and EDMT is closely tied to their multiplex capability. In the event of a node failure or replacement there is no requirement for the admin to take any action to reestablish access to "lost data" because all of the nodes in the SAP IQ MPP environment "see" and access all of the shared data directly. For disaster recovery EDMT's Multi-site Replication feature is used to replicate all shared EDMT data to the disaster recovery site.

Conclusion

Today hospital emergency room staffs are overburdened resulting in long treatment wait times. A decision support system that can help streamline patient intake and diagnosis based on VCAP's ability to unify and analyze data from disparate sources could help alleviate this situation.

Personalized medicine, based on the combination of data derived from patient clinical history, personal history, family history, and genomic sequencing, will be a game changer in terms of enhancing the physician's clinical therapeutic efficacy and providing for more effective and complete health and wellness care.

Unified Big Data and data analytics tools will enable individual practitioners to more rapidly and effectively identify underlying illnesses and to prescribe the most effective therapies. With VCAP's data ingest speed and capacity for multi-source and multi-format Big Data Analytics HealthCare organizations will realize significant benefit through faster ROI, reduced TCO, and improved time to result.

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

e-mail: hrg@hrgresearch.com

<http://www.hrgresearch.com>