HRG Assessment

Vblock™ Unified Data Solution

Today the challenge is to find opportunity in change, take advantage of change to prosper, and learn to deal with broader based competition, global demographics, and changing regulations. The current business climate in the financial services sector is highly competitive and increasingly fast paced. Decision makers have little time for reflection and even less time for making effective decisions on best available information.

Financial services firms are challenged by having to support accelerated product and services innovation while simultaneously and aggressively reducing the amount of time, money and skilled resources required to deploy and administer core infrastructure systems. The Vblock™ Converged Analytics Platform Solution was conceived and designed specifically to address this challenge.

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 can 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.

Data integration and analytics is a required part of any viable business strategy designed to overcome the challenges of managing in an environment where increasing rates of change and innovation are normal. Today for a strategy to be effective it must recognize the importance of data and the need to ingest, index, and integrate Big Data (defined as structured and unstructured, streaming and static data from a variety of sources) for analysis.

Things to consider when developing a Unified Data Strategy:

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

Good information based on the best available data is central to analytics, planning, and decision-making. The key to developing and implementing a successful Unified Data strategy is to understand the business, know what data is available, define what information is needed, and what results are desired. The ability to rapidly respond to competitive and market place change as it occurs is a key benefit of an effective Unified Data strategy and is rapidly becoming a requirement throughout the financial services sector.

Copyright © 2014 Harvard Research Group, Inc.

Data Sources

Executives, directors, and managers in financial services firms understand the benefit of having access to the complete range of corporate and market data without having to be exposed to its underlying complexity. Today, no data, regardless of type, location, or format is beyond integration. It is the nature of businesses to grow and to accumulate data, and those businesses that interpret historical business and market data accurately and apply that knowledge to the analysis of current business and market data in the shortest elapsed time will tend to be market leaders. As companies collect increasing volumes of data from their business operations, websites, marketing initiatives, sales campaigns –"Little Data"- and other sources (Big Data) keeping track of that information and making it useful to the business through data analytics can provide significant distinctive competitive advantage. Some examples of the sources and uses of Big Data follow:

- The continually increasing volumes of data and transactions originating from personal mobile computing devices including smart phones, tablets, and other untethered mobile computing devices
- Social media sites used for sharing ideas, opinions, images, documents, and promoting products and services
- Emerging technology developments facilitate organizational and business change, drive productivity, increase creativity, speed new product development, and create data
- Person-to-person Web, e-mail, and SMS based interactions, and device-to-device transactions create data
- Data from the "Internet of Things", POS systems, smart cards, video feeds, SCADA systems, process control devices, and a growing number of other devices and appliances
- Analysis of consumers' use of credit cards, leading to KYC (know your customer), to *not* only determine what marketing approaches credit card issuers may pursue, but to also look for signs of fraudulent activity or possible misuse
- Analysis of trading patterns to analyze counter-party risk and assess exposure to market conditions
- Analysis of regulatory compliance and solvency, as an outcome of post-trading settlement activity
- Analysis of portfolio make-up, existing investment categories to determine areas for leverage or rebalancing risk
- Analysis of international fund transfers between financial institutions to determine compliance exposure and potential money laundering or suspicious terrorist activity

Benefits

When Business Intelligence (BI) tools and analytics run on Unified Data many previously unachievable results and benefits will realized by firms in the Financial Services sector.

- Identification of new business opportunities
- Better product innovation, shorter development cycles, and reduced time to market
- Better focused / more effective sales and marketing initiatives
- Better targeting of prospective customers
- Improved customer satisfaction and retention
- More accurate and timely risk assessment and risk avoidance
- New market / new opportunity identification
- Reduced risk and better regulatory compliance
- Enhanced profitability
- Better targeted and more effective competitive responses

Copyright © 2014 Harvard Research Group, Inc.

and more …

Three Financial Services Solution Areas for Unified Data

What follows is a brief overview and discussion of three key financial services application solution areas. Each of these areas describes a function or division that most financial services firms have today. Each of the following data centric solution areas will benefit significantly from the implementation of an effective Unified Big Data Strategy and the ability to perform real-time text analysis, SQL analysis, and cross-analysis of mixed data types.

Risk Management

In the financial services and insurance industry sectors risk management is about consciously taking the risks we want to take, for a fair price, without taking on too much risk or being blindsided by the unforeseen. Risk management helps a firm maximize its return and general prosperity. Poor risk management practices can quickly lead to ill-conceived financial commitments, undetected fraud, insolvency and collapse. Vital to business success and mandated by regulatory law, risk management is a firm wide, board-level concern.

Effective risk management means the difference between financial success and business failure. Volumes of data must be analyzed including market prices, historical experience, interest rates, and numerous other risk factors (Big Data). The data must be reliable, instantly accessible, continually updated, and selectively replicated. Effective Risk Management solutions must protect investments in proprietary applications, while enabling continual evolution to new technologies and architectures.

Enterprise risk management or corporate risk management is focused on investigation, fraud examination, and identification of security threats, intellectual property loss, environmental risks, product deficiencies and many more potential areas of enterprise or corporate liability exposure.

Audit and Compliance

An audit is the review and examination of a company's financial statements, assets and liabilities and may include a physical audit. The auditor must report his findings based upon the evidence examined expressing an opinion on and providing credibility to the financial statements. In order to facilitate the audit process thereby limiting related costs one approach is to establish a document retention policy and establish an electronic stored information archive. By anticipating an auditor’s information requirements and then archiving that data the time and labor required to perform a financial audit may be reduced and controlled. This same line of reasoning applies with regard to industry and government regulations and providing requisite proofs of compliance. This level of preparedness will result in the reduction of time and costs required to support either a financial audit or a regulatory compliance audit. The type of data (Big Data) captured and archived will be similar to that required for eDiscovery.

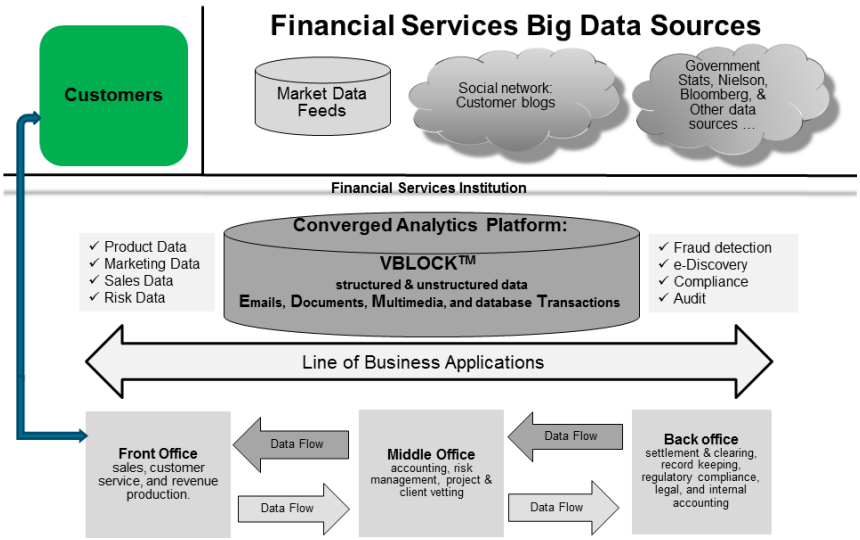
eDiscovery

Electronic Discovery refers to preservation, collection, review and production of electronic records in connection with litigation or regulatory requirements. Specific rules regarding the production of Electronically Stored Information ("ESI") have been established. The federal rules (and their state counterparts) require potential litigants to have a clear and comprehensive understanding of their ESI so that they can search and produce such ESI if necessary as part of the litigation process. ESI can include large volumes of data such as image data, audio files, video, e-mails, documents, database transactions, SMS, and text (Big Data). If a litigant does not know where its ESI is stored, and has no consistent document retention policies this adds unnecessary expense, lengthens the time necessary to comply with discovery requirements that typically demand fast production times. The lack of ESI and consistent document retention policies has in some cases resulted in sanctions and the award of substantial punitive damages, and unnecessary publicity leading to loss of confidence among the consumers and trading partners.

Copyright © 2014 Harvard Research Group, Inc.

Unified Data Solution for Financial Services

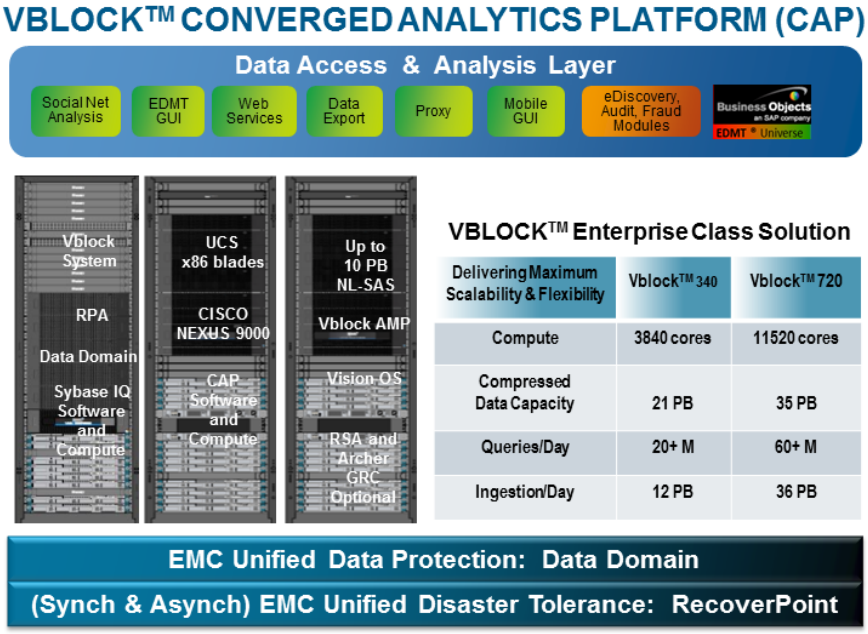
VCE with their Vblock™ Systems provide a solution that meets all of the big data challenges head on. The Converged Analytics Platform solution consists of the Vblock™ System 300 and 700 families combined with BMMsoft EDMT and SAP IQ. This is a strong partner based solution that meets the Unified Big Data challenge head on. The challenge is ingesting, indexing, archiving, managing, analyzing, and unifying high volume, high velocity, structured, unstructured, static, and streaming data from a continually increasing variety of sources and locations.



Vblock™ Converged Analytics Platform for Unified Data

The Vblock™ Converged Analytics Platform Solution is unique in its capability to deliver 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™ 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

Vblock™ Systems are designed, engineered, manufactured and validated 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

Copyright © 2014 Harvard Research Group, Inc.

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 intelligent systems management of Vblock™ Systems, optimizing services for a converged infrastructure and integrating directly into 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 time of 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's first line support staff typically resolves 96% of calls without having to engage personnel from Cisco, EMC or VMware.

BMMsoft EDMT® Unified Data Solution

BMMsoft's EDMT® Solution ingests, indexes, and transforms structured and unstructured data using up to 3 million independent data processing channels that can then be stored and managed by a single analytic data repository such as SAP IQ. 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 the EDMT archive.

BMMsoft's EDMT Solution effectively addresses the following four Big Data IT challenges:

- **Volume -** A single Jupiter system stored 12 PB of input data into a single database and EDMT's federated option allows 10s and 100s of such systems to operate as single logical entity.
- **Velocity** – The near real-time data loading speed and query speed of EDMT addresses the data velocity challenge with typical data loading speeds of 3 PB/day, data loading latency that is under 100 milliseconds and query speeds that in many cases are under 1 second enabling near real-time risk assessments, response to competitive market moves and more …
- **Variety -** Data from POS systems, smart cards, video feeds, SCADA systems, process control devices, sensor/RFID, enterprise SQL records, emails, files, images and a growing number of other devices and appliances – loosely defined by catch all phrase "the internet of things - IoT " has been loaded, indexed, and searched.
- **Value** – the TCO of the solution is orders of magnitude lower than most other enterprise solutions and equal to or lower than Hadoop.

The name 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. This converged solution delivers the performance, reliability and scalability required for real-time ingesting, indexing, accessing, and cross analysis of extreme volumes of heterogeneous Big Data.

Copyright © 2014 Harvard Research Group, Inc.

EDMT's use of Multiplexing is an excellent fit with the SAP IQ technology and both solutions benefit from the full mesh connectivity between EDMT and SAP IQ.

When EDMT is coupled with SAP IQ ingesting real time streaming market tick data storage and then running in database analytics the performance and throughput delivered is a testament to this close partnership. In addition, 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.

EDMT supports enterprise wide searches across disparate mixed data types with its high speed Extract, Transform, and Load (ETL) and indexing capability. 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, such as financial market 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.

EDMT and SAP IQ use column data stores providing significant benefits like excellent data compression, improved query/analytics speed, and better data security.

Both 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.

EDMT® is a standalone application specifically designed for ingesting, indexing, archiving, and searching of large volumes of structured and unstructured data. Using EDMT businesses can run real-time text analysis, SQL analysis, and cross-analysis of mixed data types. With the EDMT Solution near 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 customers can capture, index and store data, and 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 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. Replication to sites that have "fallen behind" will automatically resume from the failure point and eventually catch up with the working sites. 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 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

Copyright © 2014 Harvard Research Group, Inc.

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 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 and Recommendation

Financial Services firms and managers today are faced with increased regulatory and competitive pressure which directly translates into the requirement to capture, store and analyze continually increasing volumes of data, both structured and unstructured. The Vblock™ Converged Analytics Platform has been purposefully architected by VCE to help businesses most efficiently and effectively meet the increased and more detailed reporting requirements of regulatory bodies. Additionally, this solution is very effective in capturing real-time, historical, structured and unstructured data for the analysis of, and response to, new and emerging competitive challenges.

Right now industry leaders are exploring ways to respond to opportunities and challenges as they happen in real-time. This means enabling applications and automated systems to take full advantage of real-time data analysis in order to propose a tactical response based on that real-time analysis.

The VCE Vblock™ Converged Analytics Platform is a solution that was devised and architected for this evolving future where unified data analysis tools and applications software will be able to leverage real-time and historical data that reside on the same system. This new offering from VCE is able to ingest, index and store vast volumes of compressed data (100s of petabytes) with very low levels of latency. Additionally, this new VCE solution is highly scalable and able to accommodate the most demanding data ingest and storage requirements. This is a solution that deserves attention. HRG highly recommends that financial services firms and managers give this solution a close examination. We believe that you will be impressed with this new platform based solution and the full capability that this partner based team comprised of VCE, BMMsoft, and SAP brings to market.

Copyright © 2014 Harvard Research Group, Inc.

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

Copyright © 2014 Harvard Research Group, Inc.