

DATA MANAGEMENT USE CASES IN FINANCIAL REGULATION

MARKLOGIC WHITE PAPER

Following the financial crisis of 2008, a number of regulatory laws and standards were introduced in an effort to reduce the “systemic risk” a bank poses to the financial system. This whitepaper details a series of use cases and lessons learned around key imperatives related to the European Union’s BCBS 239 Principles for Effective Risk Data Aggregation and Risk Reporting.



REDUCING SYSTEMIC RISK

Following the financial crisis of 2008, laws such as the EU's European Market Infrastructure Regulation and the US's Dodd-Frank Wall Street Reform and Consumer Protection Act were introduced in an effort to reduce the "systemic risk" a bank poses to the financial system. Never again can a bank be allowed to become "too big to fail." Global financial institutions must now measure how exposed they are to possible losses, and either take action against those risks, or increase their cash and/or collateral kept in reserve should something go wrong.

In order to measure a bank's exposure, common calculations must be applied to all trading activities – but reconciling information across the data silos within the organization is a daunting challenge.

DATA AGGREGATION & REPORTING PRINCIPLES

The Basel Committee on Banking Supervision (BCBS) has developed regulatory standards – essentially, guidelines and principles – to help financial institutions comply with the new laws. In particular, regulation number 239, entitled *Principles for Effective Risk Data Aggregation and Risk Reporting*, has four core principles:

- **Risk-data aggregation capabilities.** Banks should develop and maintain strong risk-data aggregation capabilities to ensure that risk-management reports reflect the risks in a reliable way.
- **Overarching governance and infrastructure.** A bank should have in place a strong governance framework, risk-data architecture and IT infrastructure.
- **Risk-reporting practices.** The content of risk reports should be accurate, clear, and complete, and reports should be presented to the appropriate decision makers in time to allow for an appropriate response.
- **Supervisory review, tools, and cooperation.** Supervisors should determine whether the principles achieve the desired objectives.

IMPERATIVES & SOLUTIONS

Several data- and technology-related imperatives arise from the BCBS 239 principles, and this whitepaper describes six key lessons learned from our work with global financial institutions that are deploying solutions to meet those imperatives.

#1. RELATIONAL DATABASES ARE NOT SUITABLE FOR COMPLEX MODELS

“A bank should design, build and maintain data architecture and IT infrastructure which fully supports its risk data aggregation capabilities and risk reporting practices not only in normal times but also during times of stress or crisis, while still meeting the other Principles”

BCBS 239 – Section I, Principle 2

USE CASE IMPERATIVE: REGULATION & RESTRUCTURING

Like all major Banks after the financial crisis, the derivatives business of a leading global financial services firm faced problems adapting to the complexity of new regulations and real-time supervisory requirements, which resulted in poor performance and high cost associated when continuing to run their operational trade store on a common relational database model—at the heart of their then-\$70 trillion derivatives business.

Each derivative processing function – i.e. settlement, validation, clearing, etc. – was stored in a separate database, resulting in more than 20 Sybase* relational databases. Derivative trade data came in heterogeneous XML documents, which were difficult to decompose into relational tables. As a result, the bank relied upon significant ETL processes to process and query derivative information, generating opaque CLOBs and limiting the middle and back office’s ability to keep up with the rate of contractual innovation. Analysts looking for a comprehensive view of a counterparty would have to navigate more than a dozen different applications.

SOLUTION

The bank turned to MarkLogic in 2009 to deploy a new operational data store based on the MarkLogic® Enterprise NoSQL database platform. MarkLogic’s native XML support made it possible for the bank to quickly integrate and manage its XML derivative trade data without ETL delays, streamlining workflow processes and simplifying much of the existing architecture.

The MarkLogic platform makes it possible for hundreds of users to perform both structured and ad hoc queries over trade data, and the system integrates with multiple front office applications. With MarkLogic’s flexible data model, it is much easier for the back office to accommodate changes to the structure of trade documents without downtime or schema changes. The operational data store provides full ACID compliance and government-grade security to ensure data reliability, integrity, and consistency, and integrates with a number of other applications to ensure a streamlined trade settlement process.

The end result is a successful and growing deployment that powers more than \$80 trillion in trades, persists more than a billion documents, and supports more than 10,000 reads and updates per second.

#2. BANKS NEED TRUSTED ENTERPRISE FEATURES TO RUN THEIR BUSINESS

“A bank should be able to generate accurate and reliable risk data to meet normal and stress/crisis reporting accuracy requirements. Data should be aggregated on a largely automated basis so as to minimize the probability of errors.”

BCBS 239 – Section II, Principle 3

USE CASE IMPERATIVE: ACCURACY & INTEGRITY

A tier-one global investment bank was struggling to provide its risk and compliance personnel with a quick way to access regulated trade data. Trade data – trade events, transaction data, historical trade reports, etc. – was stored across more than 30 trading systems, including many built on Oracle* DB. The trade data came in multiple XML formats including FPML, a standard XML format used in the derivatives industry. An individual risk analyst interested in retrieving all the data related to a single trade would have to navigate multiple websites, each secured with a separate log in. Alternatively, the company relied on significant ETL processes to provide compliance analysts with a comprehensive report of trade information.

The bank was interested in integrating and managing all of its trade data in one platform, making it easier for risk and compliance analysts to quickly assess risk and respond to external regulatory requirements. The company spent significant effort to bring multiple data types and sources on an Oracle-backed project, but it was unsuccessful.

SOLUTION

MarkLogic’s multi-model, enterprise-ready database made it possible for the bank to easily handle structured and unstructured data in one platform.

Specifically, MarkLogic’s native XML support allowed the company to ingest the company’s heterogeneous XML data without any ETL delays. As data is ingested, it is automatically indexed, making this information available to end-users quickly. Using MarkLogic’s built-in full text indexes, risk and compliance analysts are able to pull up comprehensive trade information at sub-second speeds. Range indexes are used to keep track of trade data versions, and to ensure data consistency across data stored in MarkLogic and other source systems. Trade data can be surfaced through a dashboard, providing compliance analysts with a visual and intuitive view of trade data.

Additionally, trade data is distributed across multiple downstream applications in different formats, providing multiple business lines with access to the same, previously siloed information. With the operational and transactional MarkLogic platform, the bank can also support performing trade decorations in place, providing a single source of truth for the bank to perform critical functions such as uncleared margin calculations.

The flexibility that MarkLogic provides makes it possible for the company to accommodate new data and queries over time. The first production deployment went live within six months, and now over 50 terabytes of trade data – and billions of documents – from more than 30 trading systems are currently stored in MarkLogic.

#3. COMPLEX RELATIONSHIPS REQUIRE SEMANTICS

“A bank should establish integrated data taxonomies and architecture across the banking group, which includes information on the characteristics of the data (metadata), as well as use of single identifiers and/or unified naming conventions for data including legal entities, counterparties, customers, and accounts.

Roles and responsibilities should be established as they relate to the ownership and quality of risk data and information for both the business and IT functions.”

BCBS 239 – Section I, Principle 2, 33

USE CASE IMPERATIVE: TAXONOMIES & LINKAGES

A leading provider of business process outsourcing solutions wanted to deliver a settlement and clearance management solution that could quickly and accurately:

- Bring “trade” details into a single representation of a trade
- Link parties, instruments, transactions
- Provide current views of Positions and Trades Statuses in real-time
- Deliver data to client systems and applications in desired formats

SOLUTIONS

The vendor turned to MarkLogic to help with the required delivery of real-time positions and status, and linking parties to instruments to transactions to provide complete and accurate data. First, MarkLogic allows them to quickly ingest data as-is and eliminate the dependency on traditional ETL processes. Second, MarkLogic Semantics provide a method to bridge the relational and non-relational worlds in the three key areas of document linking, schema mapping, and master data management.

#4. REGULATION & OPERATIONS ARE INTERTWINED

“Risk Management reports should cover all material risk areas within the organization. The depth and scope of these reports should be consistent with the size and complexity of the bank’s operations and risk profile, as well as the requirements of the recipients.”

BCBS 239 – Section III, Principle 8

USE CASE IMPERATIVE: COMPREHENSIVENESS

A tier-one global investment bank wanted to provide its employees with an integrated view of operational data, providing a global perspective of the bank’s operations in one platform in order to improve how the bank executes tactical, daily operations.

The challenge was that the bank’s information was stored in various formats – ranging from COBOL index files, to Excel spreadsheets – across thousands of databases and tens of thousands of applications. Significant ETL processes were required to generate a unified view of trade event data for a single trade, for example. As an alternative, employees had to navigate through multiple sources and data systems to pull all the information related to a transaction or event – like assessing the liquidity available to an institutional customer. Even then, identifying the context around any one transaction or event would be incredibly difficult as duplicate information would be stored in different formats without clear detail for reconciliation.

SOLUTION

The bank leveraged MarkLogic’s flexible schema to integrate and manage varied data from multiple sources, without significant ETL delays. With MarkLogic, the bank has built a single canonical data model, allowing the company to deliver information from thousands of sources to employees and applications across the globe. Multiple bank teams – risk management group, compliance group, etc. – have reaped the benefits of a single database and search platform, saving time once spent on information exploration, as well as providing a stronger risk profile and more consistent exposure and liquidity numbers.

The bank is now able to meet compliance and other external regulatory requirements with stronger, contextualized reporting.

REQUIREMENTS FOR A COMPREHENSIVE VIEW

Flexible Model

The power to express everything from Excel spreadsheets to Oracle tables to COBOL indexes

Transactional Consistency

The operational data store for questions about client liquidity

Tunable Security

Limit who can see reference data based on existing permissions

#5. IT ASSETS ARE INTEGRAL TO ASSESSING RISK

“A bank should be able to capture and aggregate all material risk data across the banking group...”

BCBS 239 – Section II, Principle 4

USE CASE IMPERATIVE: COMPLETENESS

It's not just trading activities that can put a banking group at risk. The IT infrastructure components that support the operations across an organization can also be a source of risk, and banks must be able to quickly assess and address any issues. A top 5 global financial services firm found it impossible to get an overarching view of all of their IT assets, as data was scattered across many silos – including 100+ data sources and 200+ object classes (e.g. Servers, Racks, Databases, Switches, etc.).

SOLUTION

The firm developed a technical asset inventory system that serves as a centralized inventory for the firm's IT assets, including all of their relationships to specific infrastructure such as data centers, racks, and servers.

Using MarkLogic, they were able to integrate all of the data silos and provide a real-time search and query interface to better analyze status, cost, impact analysis, gaps and redundancies, and much more. MarkLogic also adds additional value by enriching the data with semantics to help analyze the relationships by creating a knowledge graph of the assets, allowing users to ask questions like, “Show me the applications in the investment banking domain that are hosted on either physical or virtual hosts running on Dell servers that are more than two years old.”

Some of the key capabilities of the system include:

- Identify affected software systems and business units when issues arise
- Cost reporting at a software system level
- Impact analysis and reporting for data center and infrastructure changes
- Google-like access to software system and infrastructure information
- Efficient data entry with reduced inconsistencies across various compliance initiatives
- Identification of gaps and redundancies across technology assets

The firm was able to deploy the solution into production within six months.

#6. GETTING A VIEW OF THE WHOLE BANK REQUIRES EXTREME SCALE

USE CASE IMPERATIVE: INTEGRATED CUSTOMER DATA & TRANSACTIONS, AT SCALE

Although BCBS 239 focuses on investment banks, the largest global systematically important banks (G-SIBs) also have consumer banking arms with similar data aggregation and transparency requirements. The consumer and community banking division of a top 5 global bank needed a better way to get complete, accurate information to their employees – and in turn, to regulators and customers. They wanted a system that would:

- Allow users to simultaneously search across multiple systems containing separate but related information
- Eliminate the need for manual aggregation of search results
- Associate documents to business processes

The repository was going to start “small” – with about 2 billion documents – but needed to be able to grow to at least 80 billion documents and cover most of the retail side of the bank including mortgages, credit cards, home equity loans, and auto loans. And document ingest and search needed to be fast – 40 million documents per day, and query results returned in less than 500 milliseconds.

SOLUTION

The bank – a G-SIB who had previous experience working with MarkLogic on a large project in their investment banking arm – has turned to MarkLogic to implement a division-wide document metadata repository and provide a Centralized Search Index engine to enable cross-system search results with result management to include faceted filtering and guided navigation. MarkLogic will deliver numerous benefits to the bank:

- **Timely** – First production deployment expected within 6 months
- **Flexible** – Integrates 35 primary systems for mortgage information alone, with ability to harmonize data inside the database as needed
- **Cost efficient** – Elasticity with commodity hardware, and reduced TCO from a platform that combines content and indexing
- **Enterprise ready** – Security, Consistency, HA/DR
- **360° view of retail banking customers** – delivering fast, complete, and accurate results across billions of records

ENTERPRISE DATA INTEGRATION REQUIRES ADVANCED FEATURES

Today, organizations want a unified, actionable, 360 view of data. Unfortunately, data is spread across disconnected databases and data integration lags the speed of business. The underlying problem is that the traditional approach to data integration using relational databases with ETL sacrifices agility, timeliness, and cost.

MarkLogic is the world’s best database for integrating data from silos. Organizations around the world rely on MarkLogic— an operational and transactional Enterprise NoSQL database platform—to integrate their most critical data and build innovative applications on a 360-degree view. MarkLogic makes it easy to get data in and easy to get data out, and it is 100 percent trusted to run critical business operations.

WHY MARKLOGIC?

- Fast Time to Results
- Ask Anything Universal Index
- Trusted Data & Transactions
- Enterprise-Grade Security
- Scale-Out Commodity Hardware
- Lightning Fast & Real-Time

With MarkLogic, a typical data integration project goes up to four times faster. And, since MarkLogic is transactional, you can build real-time, operational applications on your integrated data.

Large and complex data integration initiatives – like those powering the solutions described in this whitepaper – require a combination of features and performance that only MarkLogic can provide, including:

- Scale-out commodity hardware. Scalable and elastic without over-provisioning or over-spending
- Tiered storage. Minimize the cost of storage as the database expands
- Semantics. Enhance search and define data relationships
- Bitemporal. Time-based history and audit trail of data
- Real-time alerting. Unlimited alerts on new data at time of ingestion
- 100% trusted. ACID transactions, automated HA/DR, Common Criteria certified security

MORE INFORMATION

New and changing regulatory and competitive pressures require financial services institutions to make better use of the massive amounts of disparate and intricate data they use and produce each day. But, you can leverage all of your data to reach your enterprise goals with less time and expense than you once thought. Learn how MarkLogic is helping organizations around the world to integrate their data silos and deliver a unified, actionable, 360-view of their data.

PLAN FOR SUCCESS WITH HIGH-STAKES DATA PROJECTS

Your high-stakes data projects don't have to end – as analysts predict – in failure. Learn how to leverage all of your data to reach your objectives with less time and expense than you might imagine.

<http://www.marklogic.com/resources/plan-success-high-stakes-data-projects/>

BEYOND RELATIONAL

Learn why relational databases are ill-suited to handle today's data challenges.

<http://www.marklogic.com/resources/beyond-relational/>

SOLVE COMPLEX DATA CHALLENGES WITH MARKLOGIC

A new generation technology designed for today's modern data types, MarkLogic is the only Enterprise NoSQL database. Built with a flexible data model to store, manage, and search all of your data, without sacrificing any of the data resiliency and consistency features of RDBMS.

Visit us on the web to learn more:

<http://www.marklogic.com/solutions/financial-services/>

© 2016 MARKLOGIC CORPORATION. ALL RIGHTS RESERVED. This technology is protected by U.S. Patent No. 7,127,469B2, U.S. Patent No. 7,171,404B2, U.S. Patent No. 7,756,858 B2, and U.S. Patent No 7,962,474 B2. MarkLogic is a trademark or registered trademark of MarkLogic Corporation in the United States and/or other countries. All other trademarks mentioned are the property of their respective owners.

MARKLOGIC CORPORATION

999 Skyway Road, Suite 200 San Carlos, CA 94070

+1 650 655 2300 | +1 877 992 8885 | www.marklogic.com | sales@marklogic.com



999 Skyway Road, Suite 200 San Carlos, CA 94070

+1 650 655 2300 | +1 877 992 8885

www.marklogic.com | sales@marklogic.com