Cloud Solutions for Centralized Reference Data Management
Srinivasa Kumar Yerchuru
Srinivasa heads the Capital Markets practice at TCS and has over 20 years of experience working with banking and financial services customers. He spearheads large strategic programs across capital markets and other market initiatives. His areas of interest include investment banking, private wealth management and asset management. He holds a master’s degree from the Indian Institute of Science, Bangalore.

Ganesh Padmanabhan
Ganesh has seven years of experience in IT solutions and consulting in the Capital Markets domain. He has worked with leading market infrastructure firms on initiatives related to corporate actions and settlements. Ganesh holds a master’s degree in business administration from the S. P. Jain Center of Management, Singapore.

Ramakoteswara Rao T
Ramakoteswara Rao has over 15 years of experience working with manufacturing, retail, banking and capital markets customers across the US, the UK and Europe. His areas of interest include large program management, wealth management and related industry initiatives. Rao holds a master’s degree in mechanical engineering from IIT, Madras.
Due to the repercussions of the global financial crisis, most banking and financial institutions are seeking innovative ways to improve efficiencies and contain costs. Financial institutions are examining their business processes and identifying new functions that can be managed effectively by their outsourcing partner. Reference Data Management (RDM) currently revolves around data quality, integration between data elements and flexibility to meet regulatory demands. Over the years, centralization and standardization of data management and reducing risks and costs have been the prime motives for transformation programs when the technology has been available at a reasonable cost. Such transformations have resulted in increased automation and higher straight through processing (STP) with a special focus on centralized solutions instead of siloed data processing methods.

Today, managing reference and master data offers no differentiation and there is no competitive cost advantage in maintaining systems separately and liaising with multiple entities such as data vendors, technology vendors, internal IT teams and client teams. Despite use of improper reference data resulting in fixes on trade data, many capital market firms consider the cost of maintaining reference data within the firm as an unnecessary overhead. As a result, capital market firms are adopting a platform based approach for RDM that has resulted in increased process efficiencies though no measures have been initiated to improve the data quality. Thus, it is necessary for the firms using data from various sources to develop their own expertise to improve the data quality through strong governance mechanisms and information management tools.
Despite having failed to address data quality issues, many data vendors and technology solution providers are offering a flexible approach to RDM, as the demand for a centralized utility for data management grows. A centralized utility would help to address the issues related to system deployment and efficiencies. The emergence of a centralized utility model for reference data management is currently evolving and industrialization of such a model requires consensus and collaboration between multiple market participants, data vendors, technology solution providers and regulators. In the interim, firms need to explore options to adopt next generation technology services such as RDM as a service (managed service model) and cloud or data virtualization techniques. One option is to adopt a cloud-based centralized enterprise data management solution that ensures improved risk management, regulatory compliance and maximizes operational efficiency across the enterprise. To counter increased pressures, financial institutions are evaluating the option of cloud-based hosting, maintenance, monitoring, and support of reference data systems in an attempt to understand if it can result in improved cost and operational efficiencies. The objective is to explore the options of promoting on-demand data sharing between trading, sales, reference data, market data, and financial control systems.
Contents

1. Introduction 6
   The Current Scenario 6
   Challenges in Reference Data Management 6
   Reference Data Management Solutions Available Today 7
   Comparing Data Management Models 8

2. Setting up a Flexible and Scalable RDM solution 9
   Centralized Data Management – A Strategic Approach 10
   Benefits of Cloud Solutions for Reference Data Management 11
   Data Management through a Centralized Reference Data Utility 13

3. Conclusion 14

4. Further Reading 15
Introduction

Over the years, most research, products and solutions have focused on distributed reference data platforms which can be scaled at the enterprise level. Most financial institutions are engaged in integrating the enterprise data management program with reference data including legal entity data, instrument data, pricing, corporate actions, standing settlement instructions, and client data. The concept of a single or centralized reference data platform was triggered by the need to reduce manual intervention, speed up the on-boarding process, simplify risk analysis, expedite regulatory reporting and improve automation, thereby achieving higher operational efficiencies.

The Current Scenario

Reference data is consumed across the securities processing value chain. Despite technological advances, the problems related to faulty reference data, misinterpreted corporate actions and higher reconciliation efforts continue to exist leading to increased operational risks and costs. Organizations incur lower costs by fixing faulty transactions at the initial stages of the trade process as compared to later stages. Losses due to inaccurate processing of corporate action, bad data or misinterpretation of data account for five to eight percent of firms’ operational costs.

Today, the reference data space is teeming with products including Eagle PACETM, and GoldenSource among others. Firms can now choose from multiple reference data products for each business line, and are thus spending more on sourcing data from various data vendors, leading to high maintenance costs and complex work flows, requiring a large number of resources to maintain the processes and systems.

Although expenditure on reference data can be lowered through outsourcing, the implementation cost continues to rise as the solution will need to integrate with legacy applications. Additionally, these products cannot address some challenges related to data acquisition, maintenance, and managing and distributing data to the core applications. Firms often duplicate processes to perform the same tasks for an identical outcome and these processes lack the ability to use straight through processing (STP) for timely trade settlement.

Challenges in Reference Data Management

Some of the specific challenges that surround RDM across the securities value chain from an operational perspective include:

- Time consuming customer on-boarding process and lack of a single view of client and product details resulting in limited up-sell and cross-sell opportunities for the financial institutions
- Inaccurate statements due to inconsistent client and product data leading to poor client reporting and management
- Pricing errors and ineffective pre-trade analysis due to erroneous data resulting in inaccurate quotes
- Poor risk management and assessments due to inaccurate counterparty and trade book data
- Increased settlement errors, reconciliation issues, trade breaks and incorrect payment instructions and higher manual interventions due to delay in capturing market data
Poor real time risk reporting systems due to lesser P&L control, market risk exposure, duplicate trades and trade failures due to non-availability of data

Poor client services as a result of inaccurate calculation of brokerage fees, commissions and stamp duties

Manual processing of corporate actions, entitlements and client notifications resulting in increased reconciliation of corporate action processes.

Constantly evolving regulatory environment with emerging regulations like FATCA and LEI significantly impacting reference data

Existence of multiple data sources and processing of data in silos leading to inefficiencies

To reduce the data quality issues, firms need to integrate:

- Different types of reference data such as financial instrument data, indexes, Exchange Traded Funds (ETF)s, legal entity identifiers and corporate actions data with enterprise data
- Corporate actions and counterparty data with the firm’s data management system

Reference Data Management Solutions Available Today

Large firms adopt siloed data systems and source data from different service providers based on consumption by different business units leading to duplication of reference data. Following a merger or an acquisition, firms expand and grow geographically and adopt siloed solutions resulting in different RDM systems and databases that can lead to integration difficulties. In addition to the above complexities, firms invest in complex platforms to store varied market data including tick size, price and curve rates among others.

While the use of reference data products from leading vendors like Asset Control, Eagle PACETM, Informatica - Siperian, Oracle Siebel and GoldenSource is increasing, many firms still continue to maintain reference data in legacy systems and integrate with web-based solutions through Service Oriented Architecture (SOA) or message bus. The reference data products customize the meta data model and adopt either a centralized or federated approach based on business needs. Table 3 shows some of the models through which firms manage reference data.

<table>
<thead>
<tr>
<th>Models</th>
<th>Solution</th>
<th>Approach</th>
<th>Descriptions</th>
<th>Maturity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared Service</td>
<td>In premise</td>
<td>Federated</td>
<td>A centralized expert team comprising the software vendor, operations and application maintenance teams support operations and maintain the RDM platform and delivery systems</td>
<td>High</td>
</tr>
<tr>
<td>Managed Services</td>
<td>In premise</td>
<td>Federated and Centralized</td>
<td>A dedicated ADM vendor offers related end to end service in-premise with platform implementation and support</td>
<td>Medium</td>
</tr>
<tr>
<td>Hosted Solutions</td>
<td>Out of premise</td>
<td>Centralized</td>
<td>RDM assets are taken out of the client premise and the entire platform is hosted in the vendor’s data centre. ADM and infrastructure services are provided from vendor premise</td>
<td>Low</td>
</tr>
<tr>
<td>Utility</td>
<td>Out of premise</td>
<td>Centralized</td>
<td>Build and develop industry renowned platform managed by a single vendor</td>
<td>Extremely low</td>
</tr>
</tbody>
</table>

Table 3: Reference Data Management Models
With regard to outsourcing, many firms are moving to a shared services/managed services model for reference data system management. Adopting industry products and a shared services model for maintenance offers no competitive advantage; hence firms look for a more commoditized model for reference data. Using a combination of tools for data management, data resource firms and third party data service vendors are providing normalized data through a managed services model offering Reference **Data as a Service**. The third party data vendor would manage commoditized reference data by providing a single standardized golden copy of the reference data by normalizing, cleansing, validating and aggregating data from multiple sources and streaming it to customers through a managed service.

**Comparing Data Management Models**

In general, firms adopt two types of data management models, federated or centralized. Though the centralized model is considered to be traditional and easy to implement, changes in business requirements make it complex leading to a longer development cycle. On the other hand, federated architecture offers quick turnaround in implementation but leads to silo-based processing and integration challenges. Table 4 compares the federated and centralized data management models.

![Table 4: Comparison of Federated and Centralized Data Management Models](image)

Both the models have pros and cons and firms need to adopt the right model, framework and architecture to attain a flexible and cost-efficient data management solution. These models primarily help in addressing deployment and efficiency issues but do not address the data quality challenges currently faced by the firms. To improve data quality, firms need to develop expertise by creating centers of excellence and using the best information management tools. By adopting a centralized model, firms can generate a customized golden copy of reference data to meet demand and eliminate process redundancies and fragmented data sources.
Setting up a Flexible and Scalable RDM solution

To meet the current business needs under volatile conditions, securities firms need to adopt the right approach, framework, data architecture and model. The right approach should have the ability to offer cost efficiencies along with flexibility to scale up infrastructure capabilities. Table 5 depicts the various phases of maturity on a timescale in managing reference data by describing the salient features of the recommended model, its benefits and limitations.

<table>
<thead>
<tr>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centralization</td>
<td>Cloud based delivery system</td>
<td>Utility based model</td>
</tr>
<tr>
<td>Short term</td>
<td>Medium Term</td>
<td>Long Term</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>■ Centralize data management through the implementation of a platform technology (like GoldenSource, Eagle PACETM) and build a centralized team for managing the data capture and distribution. Alternatively, create a data governance model and data management framework</td>
<td>■ Transfer the RDM assets to a vendor to host the infrastructure and applications using a private or hybrid cloud solution</td>
<td>■ Expand income streams by supporting a single data platform management and mutualize operation support in a cloud environment</td>
</tr>
<tr>
<td>■ Continue with a data hub, maintain cross reference and centrally manage the sourcing and distribution of services and mutualize the operations and support for RDM for the existing federated architecture</td>
<td>■ Adopt a vendor-specific on-demand or cloud based data management solution while retaining data delivery on-premise</td>
<td>■ Increase efficiencies through process and technology automation with lean data sourcing methods and expand the platform/utility to service multiple clients.</td>
</tr>
<tr>
<td>■ Move the common reference data to a central repository and retain the specialized reference data managed by the respective LOBs</td>
<td>■ Develop a scalable model by creating centers of excellence and build expertise by using information management tools to improve data quality</td>
<td>■ Requires a standardized meta data model and governance mechanism</td>
</tr>
</tbody>
</table>

**Benefits**

- Improves process efficiencies through centralization
- Reduces costs by eliminating redundancies in sourcing and replication of data
- Provides more control and improves data integrity
- Reduces TCO by hosting the RDM asset in a cloud environment
- Utilizes resources effectively
- Reduces the time for implementation for future business demands and needs
- Achieves economies of scale
- Provides a more flexible, configurable and scalable solution
- Results in faster implementation, information updates and time to market.

**Limitations**

- Requires exception based processing
- Requires longer lead time for implementation
- Cannot achieve economies of scale
- Creates integration challenges for the downstream applications between cloud and non-cloud applications and between various data providers for sourcing data
- Presents concerns over data privacy and cloud security
- Requires a standardized meta data model and governance mechanism
- Requires more effort for providing customized solutions and complexity in segregation of data

Table 5: Recommended approach, benefits and limitations for managing reference
Centralized Data Management – A Strategic Approach

Many investment banks have implemented multi-year enterprise data management programs aimed at integration of reference data including legal entity data, instrument data, pricing, corporate actions and standing settlement instructions and client data.

A scalable, flexible platform based solution capable of providing cost and operational efficiencies that reduces manual intervention, speeds up the on-boarding process, simplifies risk analysis and regulatory reporting and increases automation is necessary to meet firms’ business needs.

A strategic approach would be to automate reference data management across the value chain by defining a centralized model as shown in Figure 2.

![Figure 2: Centralized Reference Data usage across securities settlement value chain](image-url)
Figure 3 is a schematic representation of a centralized RDM solution.

Benefits of Cloud Solutions for Reference Data Management

Market data is widely consumed by investment banks, hedge funds, wealth management firms and other proprietary trading firms, and is a critical input for their trading strategies spanning across global markets. Hence, many exchanges such as NASDAQ, NYSE, and CME have started offering public market data as a cloud computing service to market participants.

The main challenges in maintaining an internal market data infrastructure include latency issues, constant changes and modifications to feed handlers and caching systems, and high cost and effort to adapt to minor changes. Besides, data management for corporate actions, which requires additional effort to capture, route, scrub, reconcile, store and redistribute data within the organization, is also a challenge. Also, there are challenges within the organization to centralize reference data management and build the necessary network and services infrastructure to distribute it globally to consuming applications.

Though cloud adoption has gained momentum, challenges around data sensitivity and security continue to exist and need to be addressed. A cloud-based technology solution can deliver cost savings in the following areas – data center, development efforts, testing cycles, operational efforts in managing data - and also help improve scalability.
Multiple platform providers offer cloud-based or on-demand data scrubbing solutions and deliver the clean data in the required format or load it to the enterprise data warehouse. Cloud-based solutions can be game changers only if they are scaled to a centralized utility on a multi-tenant platform. For instance, if security master data can be managed in the cloud, data can be delivered directly to the downstream systems thereby eliminating storage requirements, on-premise maintenance of the securities master data and related infrastructure. An alternative means of accessing data is through a web browser. Cloud-based MDM or RDM solutions will completely reshape the data service vendor, technology partner and outsourcing service provider landscapes with serious business implications.

Using a cloud-based model, firms can address some of the key concerns and issues in the reference data management domain:

- Weak/legacy infrastructure and custom-built RDM system
- Continuously changing regulatory landscape
- Constant changes to downstream systems
- Modifications to feed handlers and caching systems

Firms can shift to a cloud-based solution in two ways:

**Option 1** – Move from a standalone MDM or RDM software to a vendor-specific on-demand or cloud-based solution where data management is performed on the cloud

**Option 2** - Host the entire RDM infrastructure on a cloud data center, perform data management activities and transfer the normalized data to an operational data store, which resides on client premises, and then transfer the data to downstream applications

A cloud-based RDM solution should be able to deliver data via sophisticated web Application Programming Interfaces (APIs) that can easily integrate into on-premise operational data stores. This mechanism also helps address the data privacy and security concerns.

The solution should have the capability to integrate many data sources such as data vendor and exchange, and central securities depositories. The data management solution should ensure data integrity by adopting logical and rule-based data cleansing processes. The solution should be able to aggregate data from different sources, enrich, scrub and perform reconciliation at a record level, and then deliver normalized data in different formats – files or feeds— for consumption or to an operational data store residing in the client premise. This data can be further distributed for consumption to various applications. The data synchronization and replications to the operation data stores should be tuned with defined frequencies so that data is refreshed in real time or intraday depending on the applications’ consuming mode and integration techniques.

Organizations can derive multiple business benefits from cloud-based RDM solutions including improved operational efficiency owing to a higher degree of automation and reduced total cost of ownership (TCO). Cloud-based RDM solutions also help lower business risk by reducing trade exceptions and breaks and eliminate errors associated with manual processes. Additionally, by using flexible web APIs, real-time and referential market data is integrated into a spreadsheet enabling advisors to consider cutting the cord on terminals altogether. The inherent flexibility of cloud-based RDM solutions help improve time-to-market by allowing modifications to the application to quickly respond to changes in compliance requirements.
Data Management through a Centralized Reference Data Utility

The concept of a centralized utility for reference data management is evolving as most firms recognize that managing reference data in-house or through an outsourcing vendor is neither competitive nor a core activity. Reference data such as instrument data, corporate actions data and counter party data are the same for all firms and can be applied universally in the market place. In the current scenario, many firms incur costs and expend effort when they receive faulty data or files from the existing vendor. This effort is duplicated across the market place and resolves common issues facing all firms. By adopting a centralized utility, firms can manage their resources and capital effectively and focus on core activities. Furthermore, many regulators are aligning their regulations and policies to encourage the establishment of centralized data repositories to minimize systemic risks. A few of these are listed below.

- Securities and Exchange Commission’s (SEC) EDGAR corporate filing, a repository of corporate financial information
- European Central Bank (ECB) is proposing a reference data utility
- The Committee of European Securities Regulators’ (CESR) development of its own central reference data repository

The effort of building data repositories is overlapping across geographies and adding complexities by introducing additional standards and silo based processing of data, transactions and position information. On the other hand, the industry is aiming and looking at a single global solution, standard and mechanism for reference data management.

In lieu of a single RDM solution, some market infrastructure firms are collaborating with technology providers to provide normalized and clean data through a centralized utility based on the managed services model. These utilities validate, correct, enrich and maintain business entity reference data from various sources and provide a single feed or file to the investment firms which in turn take care of distributing the data to downstream applications. Firms such as Deutsche Borse’s/DTCC Avox subsidiary are supporting a collaboration of over 25 financial institutions, including investment banks and asset management companies, which share their independently sourced business entity data to ensure accuracy and compatibility.

Some of the utilities that could emerge in the industry are:
- Securities master utility to manage equity and fixed income instruments, issuers, benchmarks and indices
- Market data and prices utility to manage prices, yields, ratings, durations, interest rates, foreign exchange rates, historical prices, closing prices, time series, volatilities, correlations, and factors
- Counter party data management utility to manage counter parties, funds, trading entities, agents, guarantors, transfer agents, registrars, paying agents and custodian identities
- Corporate actions data utility to manage corporate events like stock splits, proxy notifications, mergers, tenders, dividend and interest declarations, and capital distributions
- Client delivery/settlement instruction data bases to manage the standing settlement instructions
To build and provide data management solutions and services through a centralized utility, the following aspects should be considered:

- Collaborative approach and cooperation of industry participants to develop a multi-tenanted model with minimal switching costs.
- Consensus on further standardization and frameworks
- A universally acceptable meta data model and architectural standards
- Identification of critical data sources, rationalization as per industry demand and addressing jurisdictional factors
- A mechanism to integrate client systems and data stores
- Data privacy/security requirements in managing the data repository
- A solution to quickly respond to changing markets, demands and regulations and spend less time expanding into other geographies and asset classes
- Issues related to ownership, liability and sharing of legal risks

To realize success, a utility model requires critical mass and economies of scale. To achieve this, market infrastructure firms would need to take the driver’s seat to address industry problems in partnership with technology providers. Such an approach would be easier and quicker, and result in a viable solution to successfully manage and meet service levels. As market infrastructure firms such as Central Securities Depositories (CSDs) in Europe are losing revenue to market initiatives like Target 2-Securities (T2S), building a centralized RDM utility on a managed services model and offering data as a service to market participants can potentially become a new revenue stream for CSDs.

Conclusion

With the increase in the diversity of data and volume of transactions, firms are struggling to identify and resolve the underlying cause of reference data mismatches that result in costs to fix them. Most investment firms are in favor of outsourcing both the technology and operations in the reference data management domain. There has been increase in the maturation of reference data management with increased focus on improving efficiencies.

Cost savings coupled with the need for a reliable and scalable distribution mechanism across applications have been important drivers for data consumption. We understand that the cost of fixing broken trades or transactions is increasing exponentially and the main reason for such an increase is inefficiencies in the data management process. Outsourcing software implementation and mutualization of operations has resulted in poor economies of scale and inferior data quality. Hence there is scope for commoditization of centralized reference data management using a cloud computing solution.
Data service providers such as Interactive Data Corporation and XSP have moved to the cloud and adopted a SaaS application to provide financial institutions with access to Global Corporate Actions Data. With the market shaping up to embrace centralized reference data management, cloud-based solutions will enable firms to reduce operating costs and complexity. Such a solution would focus on servicing internal and external customers and help them become result driven, agile and extensible in distributing the reference data to downstream applications with minimal impact on applications that are affected due to reference data change.

In general, RDM is transforming from complex, federated, on-premise systems to a combination of complex, federated, on-premise and cloud-based systems. With applications consuming data from reference or master data management databases moving to cloud-based and ‘as a Service’ platforms, it is imperative that RDM follows suit. As a result, the data management function will soon be commoditized and evolve into the Data as a Service (DaaS) model.

Further reading

2. EuroClear - Jo Van de Velde, Philippe Chambadal: A New Central Data Utility To Manage Data Challenges
3. Financialintergroup - Infrastructure issues in the securities industry: The case for a central counterparty for data management
4. EDM Council, Data Quality — Comparison and Analysis of Securities Reference Data, 6th January, 2009
About TCS’ Banking and Financial Services Business Unit

With over four decades of experience working with the world’s leading banks and financial institutions, TCS offers a comprehensive portfolio of domain-focused processes, frameworks, and solutions that empower organizations to respond to market changes quickly, manage customer relationships profitably, and stay ahead of competition.

Our offerings combine customizable solution accelerators with expertise gained from engaging with global banks, regulatory and development institutions, and diversified and specialty financial institutions.

TCS ranked #2 in the 2013 FinTech 100 rankings of top global technology providers to the financial services industry and counts 12 of the top 20 global financial institutions among its clientele. From retail and corporate banking, capital markets, market infrastructure, and cards, to risk management and treasury, TCS helps organizations achieve key operational and strategic objectives.

Contact
For more information about TCS’ Banking and Financial Services Unit, visit:
bfs.marketing@tcs.com
or visit http://www.tcs.com/industries/banking/Pages/default.aspx

Subscribe to TCS White Papers
Feedburner: http://feeds2.feedburner.com/tcswhitepapers

About Tata Consultancy Services (TCS)
Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match.
TCS offers a consulting-led, integrated portfolio of IT and IT-enabled infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model™, recognized as the benchmark of excellence in software development. A part of the Tata Group, India’s largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at www.tcs.com

IT Services
Business Solutions
Consulting