Data Integrity for the Investment Management Industry
Abstract

The quality of data is a critical component of any enterprise relying on the data to make critical business decisions. All Business Intelligence and Analytics is only as good as the underlying data. It is therefore paramount to ensure that the integrity of data is maintained across all systems. The purpose of this white paper is to assert the importance data integrity as it relates to the Investment Management industry. It also highlights Kuberre Systems’ Data Integrity Program (DIP), a product of extensive field experiences dealing with data quality problems and ways to address them. Data Integrity is a non-trivial task and requires an extremely disciplined approach and tools to execute well. Data Integrity Program incorporates all those along with a flexible model that customizes the program to the unique needs of each enterprise. The program leverages a global delivery and support model to continuously monitor and address any issue that is identified in a proactive way – hence maintaining the end-users’ confidence in the data quality.

Introduction

Managing critical information assets is challenging within most Investment Management firms, especially when the firm is experiencing data volume growth at an exponential rate. In addition to collecting intra-day transaction data and use of historical data, there are nightly feeds from various disparate sources that need to be maintained. Data integrity issues exacerbate the problem. Even though a larger number of these organizations realize the magnitude and criticality of these issues, addressing them in time remains a challenge. Suspicious data quality can cause severe downstream repercussions, resulting in investment managers making imperfect decisions or incorrect forecasts leading to increase in costs, loss of customers, failure to meet the regulatory compliance and ultimately jeopardizing the firm’s reputation.

Poor data quality can manifest itself in many ways. The increasing number of corporate actions – mergers, acquisition, splits, dividends etc and the volume of data in nightly feeds – require a close observation and immediate intervention in case of any anomaly. The systems supporting Investment Management are usually very tightly integrated and if bad data is received and processed by one, other systems also get affected. It is imperative that proper checks are put in place to avoid any such proliferation of bad data in the systems as the researchers and managers are relying on the quality of the data to make critical financial decisions for their firms and clients.
These issues vary in complexity, frequency of occurrence and their impact on the business as illustrated in Figure 1. If not addressed in time, these can impact one or more of the following - Market Caps, Industry Assignments, Cross Vendor Links, Total Returns, Duplicate securities, Security Histories, Active Exposure, Attribution, Cash available for investment etc.

In a large number of firms, the lack of data integrity has made Portfolio Managers shy away from using the business intelligence systems that have been put in place and revert back to using their spreadsheets to make decisions. IT and other organizations are struggling to get the data epidemic in control and identify these issues in time.

Once identified, these issues invariably result in high degree of intervention from Operations and/or IT organizations, who, in turn have to rely on developers’ expertise to analyze and fix the problems. This is a distraction for valuable resources that otherwise are on deadlines for other core development related tasks. It is often too late by the time developers get to solve the problem.

The key to overcome this huge problem is to have a well structured data integrity program which ensures that appropriate processes are in place and data quality is being monitored proactively across the organization by skilled personnel. Unlike other support groups, data integrity is not isolated and is entangled with various other groups such as production monitoring, applications development and application support. An effective data integrity program unifies all those roles and is able to address the data quality issues - in time.
The ‘DIP’

Kuberre Systems offers a comprehensive Data Integrity Program (DIP) specifically for the Investment Management firms. The program leverages a disciplined methodology, knowledge based framework, technology tools and a global team of business analysts and developers to deliver Data Quality Support with Certainty.

The Global Support team ensures that the systems are being monitored all the time and any issue that is uncovered is addressed by them in time. It is only the rare case that an issue is escalated. The team is also available for a 24x7 Level 1/Level 2 support via telephone, fax or email. The support team is constantly enriching the data integrity process by adding more monitoring scripts or validation rules.

A sample set of data elements and processes that can be managed and monitored using the DIP include –

- **Security Master**
  - Corporate Actions
    - Consistency of corporate actions such as id changes across systems viz. Security Master, Accounting and Trading as well as vendors such as Compustat, Worldscope, Barra and IBES.
    - Validation of corporate actions on fundamental items such as prices, shares and total returns
  - Industry classifications over time
  - Security type classifications
Duplicate security instances cleanup and the corresponding effects on historical holdings/transactions

Post process characteristics calculations such as percent owned, accounts with the maximum exposure to a given security risk characteristics

Handling of dually listed securities and the distinct setup of these across vendors and their impact on internal systems

- **Time series links to external vendor data**
  - Differentiating vendor treatments to corporate actions such as Mergers and Acquisitions – Survivor preference, continuous history versus different histories for back-from-bankruptcy cases
  - Ability to expire vendor links

- **Holdings/Transactions**
  - Ensuring all of the positions are loaded correctly taking corporate actions affect into consideration
  - Making sure the market values deviations are within an expected range
  - For a time series analysis, security links to the security master are consistent with the corporate actions
  - Post process calculations such as percent of portfolio, percent of equity that impact portfolio level characteristics.

- **Data Reconciliation** between Custodian, Broker and Internal Systems with impacts to portfolio management and Compliance.
  - Trade, Holdings and Cash reconciliation

- **Performance Measurement**
  - Impact of corporate actions on trades/holdings
  - Issues with Cancel/Rebooks and the corresponding impact on Holdings
  - Issues with back-dated transactions

- **Post Data Load Analytics**
  - Portfolio level characteristics - performance, risk analysis, sector/country/industry exposures
  - Performance Attribution
  - Bottom up returns versus official returns – investigations, support, maintenance
Methodology

The DIP methodology is based on Six Sigma’s DMAIC (Define, Measure, Analyze, Improve and Control) principles and is supplemented by 24x7 Global Support - See Figure 3. One of the keys to solving these data-related issues is to ensure that the enterprise has a good understanding of what those issues are, and how they affect their computing systems.

The objective of the "Define" stage is to collaborate with key client stakeholders, further validate the business needs and define the data quality objectives, current data processes, requirements, scope and the tangible deliverables for the future state. During the "Measure" phase, the key metrics and progress measurement criteria are defined. These measurements help track the program success are then presented on a weekly/monthly basis to all stakeholders.

The Analyze phase has a workshop involving personnel from both client and Kuberre. The workshop conducts a current state analysis and identifies all known issues. A baseline for data quality is established. Every operational process, nightly job and escalation procedures is documented. An ‘Improved’ blueprint is laid out describing the future state of the data quality and systems. Any re-engineering or scripts needs are identified in this phase. Right data governance structures are then put in place to establish "Control". At all stages an open communication with the client facilitates the execution of the program. Finally, data stewards need to be identified and assigned responsibility for keeping data elements (within their scope) clean on an ongoing basis.

Figure 3
DQAD and 24x7 Monitoring

Once the systems and process are in place, Kuberre’s global team provides 24x7 monitoring allowing them to proactively identify and address any data issue. In most cases these issues are uncovered and resolved even before the users get in to work in the mornings.

A critical component of the Data Integrity Program is DQAD (Data Quality Assurance Dashboard), a web based light weight information portal with very rich analytical functionality. In addition to serving as an objective reporting system to measure the quality of enterprise data on an on-going basis, it also provides ‘Certainty’ to answer complex questions such as “If I were to run an attribution over a period – should I be aware of any underlying data issues?” DQAD ties data loading, analytics computations and data integrity procedures to provide deep insight into data quality and answers to complex queries.

DQAD’s mantra is “Know What You Need” – and it provides a very visually appealing interface to relevant information and eliminates the need for mass broadcast emails which either get ignored by users or dilute the criticality of the problem.

Following is a sample set of elements that are available through DQAD –

- Current status
  - Data loads, Application(s), Data Reconciliation
  - By business functionality – holdings, alphas, attribution, batch reports etc.
  - Open issues

- Metrics measurement
  - Insight into current as well as historical issues, resource allocation and well documented resolution path – to give a clear understanding of what happens next on a particular issue

- Data Integrity insight
  - By business functionality – answers to some common queries such as –
    - “If I run attribution for YTD, are there any open issues in the underlying data?”
    - “I am accessing portfolio level characteristics for the last month, are there any open issues in the underlying data that I need to be aware of?”
    - “Security’s industry classification doesn’t look right – could someone confirm?”

- All pending issues
- All resolved issues
Benefits Summary

Kuberre Systems' Data Integration Program provides the most comprehensive support that addresses your Data Integrity with Certainty. It consolidates various support groups under one umbrella for a more efficient, timely support. This program addresses –

- Trading and Compliance Errors
- The inability of the production monitoring folks to take any corrective action and provide production support with intelligence - comprising of developers and business analysts
- Eliminate nightly calls to developers on call and replace it with a more efficient working environment
- Eliminate distractions to developers in addressing data integrity problems which dilutes their focus from their core development efforts
- Proactive notifications to the end users and senior management on potential data integrity, application related issues
- Issues escalated to internal support staff only in cases of extreme problem scenarios that haven’t been identified
- Insight into current status of systems and data quality with a simple to access and real time dashboard – envision the most popular web portal framework (Yahoo, CNN) containing all you need to know about system and the quality of data
- Increase the confidence level of the user base on internal applications by assuring data quality and integrity of the underlying data of the applications.

- A Global Delivery Model thus enabling
  
  - 24x7 support
  - Maintenance, including development
  - Proactive data integrity & data reconciliation - before markets open

To discuss how Kuberre Systems can help your organization with Data Integrity with certainty, please
Call 1.866.KUBERRE x204 or
Email at di-program@kuberresystems.com