



GemFire Enterprise Data Fabric: A Paradigm Shift Reference Data Management Solution

WHAT IS REFERENCE DATA?

Reference data is the fundamental data underlying and defining the customers, securities and transactions that flow through the world's financial systems. Reference data is used in the processing of transactions, in compliance measurement, analytics, risk management, and client reporting. Without this common data, trading, clearing and settling securities transactions would not be possible.

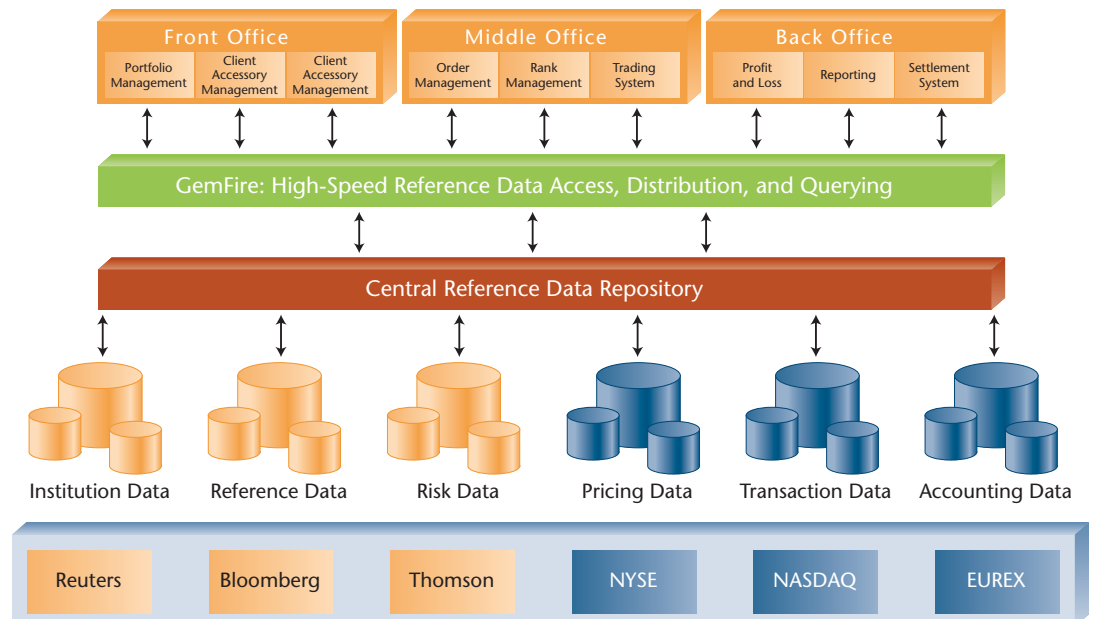
Inconsistent, inaccurate, or incomplete reference data is the primary cause for trade failures in the Financial Services industry. Exceptions in the trade lifecycle caused by unreliable reference data lead to increased operational risk, lost revenue opportunities, financial liability and the need for expensive manual trade duplication and reconciliation processes. Since artificial system delays are often introduced to make allowances for poor reference data systems, these systems are notorious for limiting the forward progress of automation and STP.

CENTRAL REFERENCE DATA REPOSITORY

Faced with both the financial imperative of providing high quality reference data and the new industry regulations complicating the management of legal-entity information, firms need to improve reference data management across the enterprise. A common

approach to addressing the challenges of reference data involves a central reference data repository. Twenty-four percent of firms have consolidated data across the enterprise and 55% of firms have efforts under way to consolidate data, though not across the entire enterprise (TowerGroup).

Enterprise Reference Data Architecture



REFERENCE DATA MANAGEMENT CHALLENGES AND SOLUTIONS

What makes reference data management so challenging? In a nutshell, having many legacy data silos complicates the process of data retrieval, normalization, and aggregation. Updates to reference data must be processed in real-time. Dependent applications need fine-grained, customizable caching, replication, and synchronization semantics. User-facing applications at the enterprise edge demand fast consistent access to golden reference data while global networks impose firewalls and add network latency bottlenecks. To complicate matters, reference data management must work across a wide range of applications and hardware platforms while providing 24x7 uptime.

While a central reference data repository is an ideal solution, it is not easy to achieve. For many enterprises, a single reference data repository is unachievable due to many factors. These include complexity, effort involved, time, cost and resources. To define, develop and manage a single central reference data repository is a huge undertaking.

Many firms find that after having undergone a major effort to create a central reference data repository, their data management challenges have not gone away. The central reference data repository brings forward some additional data challenges, including data that is overly normalized and huge performance issues for data access and distribution.

Enterprises address these new reference data challenges by creating separate optimized databases for each type of reference data. Although some of the original data challenges are reduced, they are not addressed completely. Unfortunately, performance issues continue to be a critical problem, as this results in slower response in applications, such as order processing, trade settlement, clearing, pricing and reconciliation systems, that are consumers of this data. Ultimately, this results in delays in critical decisions that investment managers, brokers, traders, stock exchanges, bankers, clearing houses need to make. These delays result in missed business opportunities.

THE NEXT GENERATION REFERENCE DATA MANAGEMENT SOLUTIONS

So what is missing from the reference data management solutions described above? There is a need to find solutions that solve the many challenges with a new perspective.

- High performance reference data management services
- Deliver consistent reference data in multiple format as required by applications
- Eliminate reference data repository performance issues
- Manage reference data sharing and data synchronization
- Support querying, filtering and rapid error detection in reference data
- Lower costs caused by reference data trade exceptions
- Cost effective reference data management

NEXT GENERATION REFERENCE DATA MANAGEMENT SOLUTIONS CAPABILITIES

What is the best solution for handling reference data? An optimal solution starts by exploiting the underlying nature of reference data. As relatively static, read-mostly information, reference data is largely stable information that can therefore be shared by multiple consumers without the need for worrying about frequent, fine-grained transactional updates. As a result, reference data is ripe content for high performance enterprise data caching techniques that maximize concurrent access to data, allowing dependant applications to share equivalent copies of the same information throughout all corners of the enterprise.

As an enhancement to existing IT infrastructure components, a solution designed explicitly to solve high performance data problems through data distribution, and streamlined caching services is a powerful tool for managing reference data. In addition to high performance, an optimal solution for refer-

INDUSTRY TRENDS

- Financial services firms are adopting automated trade processing capabilities to move toward STP
- Growing number of trades
- Increasing competitive threats
- Reduction in settlement times
- Volatility and unpredictability of volumes of stock trading
- Contracting spreads

ence data management also requires built-in capabilities for multiple data source integration, immediate and efficient system-wide updates and reliable, enterprise-wide information diffusion.

GEMFIRE ENTERPRISE DATA FABRIC (EDF): A PARADIGM SHIFT

REFERENCE DATA MANAGEMENT SOLUTION

GemFire Enterprise Data Fabric (EDF) is a flexible, high performance, cross-platform, and cost-effective solution for data distribution, caching, querying, analysis and virtualization across the enterprise in real-time. GemFire EDF is an optimal solution for the challenges of reference data because of the following:

- Provides high performance data virtualization capabilities.
- Delivers consistent reference data in multiple formats as required by applications.
- Eliminates reference data repository performance issues.
- Supports standards-based querying
- Supports real-time analysis of reference data for fault detection
- Manages reference data sharing and data synchronization across applications.
- Improves data consistency across applications and lowers costs caused by reference data trade exceptions.
- Provides a cost effective reference data management solution.
- GemFire-based solution can be an intermediate step for enterprises that do not have a central reference data repository yet.

GEMFIRE FEATURES AND BENEFITS

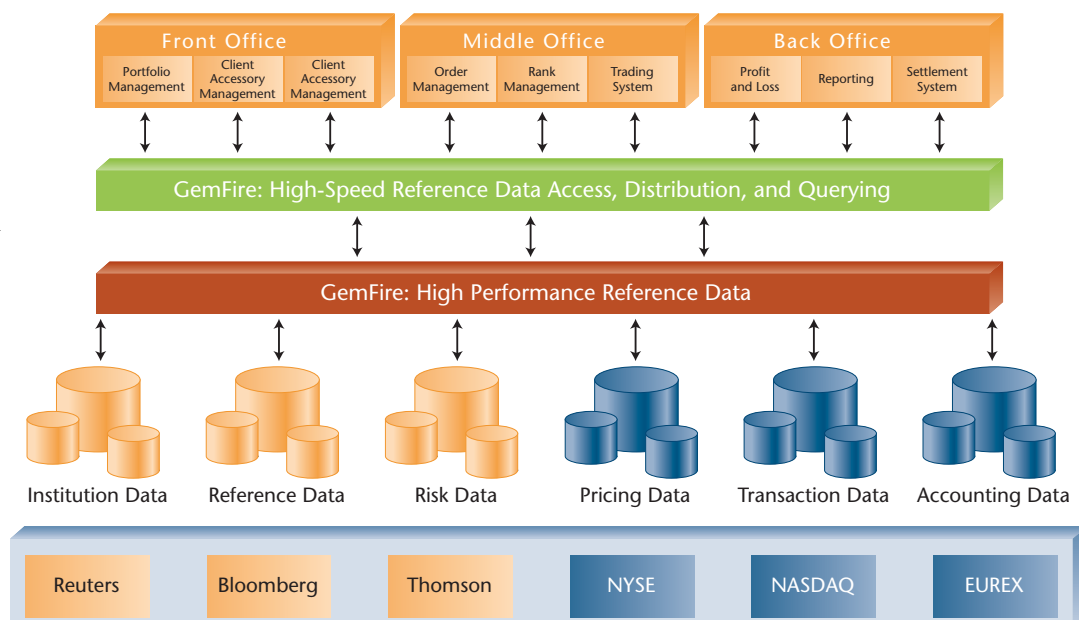
• Data Virtualization

Data feeds from multiple providers are integrated in a logical repository. GemFire works equally well with a centralized persistent data store or with a distributed, heterogeneous set of data stores. This provides user-facing applications with transparent push/pull access to all reference data through convenient APIs including XML, SOAP, and Java. GemFire decreases IT complexity, improves manageability and improves system reliability by abstracting applications away from their brittle connections to distributed data sources.

TOWERGROUP SURVEY RESULTS

- 45% of trade exceptions in automated processing systems are due to faulty reference data.
- 48% firm's instrument and master file reference data is contained in 10 or more systems.
- 60% of firm's projects to improve the reference data management were considered high priority.
- 58 full time equivalent employees are required to maintain reference data, on average.
- 83% of firm's are considering new technology investments to address the reference data management.

GemFire: A High Performance Reference Data Solution



- **Data Distribution to Enterprise Applications**

Applications reach selected reference data quickly by accessing data from local or nearby caches, creating flexible, configurable, adaptive data distribution to the enterprise edge - from Blades, Grids and Clusters to Global Networks - from server applications to edge-caching.

- **Enterprise Data Caching and Synchronization**

Reference data is kept cached near dependent applications and is synchronized automatically according to user-defined policy attributes. GemFire improves application accuracy and performance and saves money by reducing the need for manual processing of data updates or broken trades. Consistent information is automatically propagated to the edge, creating high quality reference data that ensures trade success and IT automation.

- **Querying**

The GemFire solution supports standards-based querying (OQL) on large data volumes. The ability to perform complex joins as part of a query enables information consumers to access relevant parts of reference data with ease and efficiency.

- **Distributed Data Management**

Semantic transformation logic is centralized for efficiency and flexibility; reference data is scrubbed, normalized, and enriched just once on behalf of consumer applications improving enterprise-wide consistency of reference data for all lines-of-business. Insulates applications from expensive, repetitive transformation logic.

- **Reference Data Cleansing**

GemFire can help in identifying data inconsistencies and faults as and when reference data is published into the system. This can be accomplished by monitoring incoming data and checking for relationships and expected patterns. Such approaches to cleanse the data prior to storage in reference databases can prevent incorrect data from becoming

pervasive in the system and avoid any errant decisions to be made based on this data.

- **High Data Availability**

Fault-tolerant mirroring ensures business continuity so financial service providers can offer the most competitive service level agreements with customers.

- **Scalability**

Cache eviction and disk-paging offer virtually unlimited scalability enabling aggregation of new reference data sources as business volume grows. Automated paging also acts like a 'safety valve', ensuring that applications will not fail when faced with an unprecedented or unplanned spike in data volumes.

- **Open Industry-Standards**

GemFire is based on open industry standards like Java, XML and web services, so customers can avoid vendor lock-in. GemFire Enterprise, a component of the GemFire EDF, is based on the JCache (JSR-107) standard for distributed caching.

KEY INDUSTRIAL REFERENCE DATA MANAGEMENT CHALLENGES

- Excessive cost of data management
- Reference data is spread across enterprise databases, making it inconsistent
- Multiple data formats and silos
- Querying and cleansing of reference data
- Central reference data repository has significant performance issues
- Reference data sharing and data distribution is complex
- Significant costs of fixing trade exceptions caused by reference data
- Increased operational risk

-
- TowerGroup Survey: Is the securities industry making progress on reference data management? September 2002, TowerGroup.
 - Reference Data: The key to quality STP and T+1. October 2001, TowerGroup, Reuters, Capco.



Corporate Headquarters:

1260 NW Waterhouse Ave., Suite 200 Beaverton, OR 97006 | Phone: 503.533.3000 | Fax: 503.629.8556 | info@gemstone.com | www.gemstone.com

Regional Sales Offices:

New York | 90 Park Avenue 17th Floor New York, NY 10016 | Phone: 212.786.7328

Washington D.C. | 3 Bethesda Metro Center Suite 778 Bethesda, MD 20814 | Phone: 301.664.8494

Santa Clara | 2880 Lakeside Drive Suite 331 Santa Clara, CA 95054 | Phone: 408.496.0242

Copyright© 2005 by GemStone Systems, Inc. All rights reserved. GemStone®, GemFire™, and the GemStone logo are trademarks or registered trademarks of GemStone Systems, Inc. Information in this document is subject to change without notice.