

# RED HAT JBOSS DATA VIRTUALIZATION

Your data, your way, at your business speed

DATASHEET

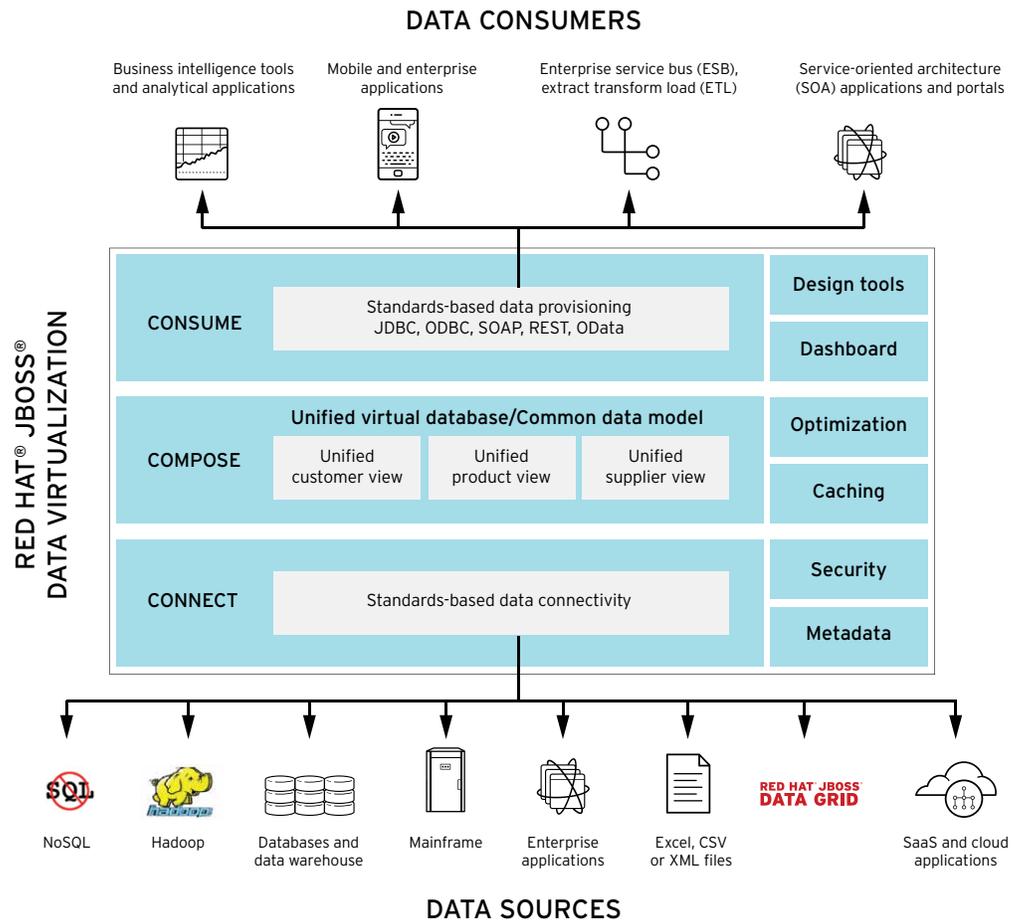
## BENEFITS

- More efficient and broader utilization of all enterprise data
- Better control and security of information
- Greater agility, faster time to solution
- Improved organizational productivity
- Improved IT efficiency
- Reduced development, maintenance, and hardware costs
- Increased return on data assets (ROA)

## FEATURES

- Provides standards-based read/write access to heterogeneous data stores in real time
- Speeds application development and integration by simplifying access to distributed data
- Transforms data structure and semantics through data virtualization
- Consolidates data into a "single view" without the need to copy any data
- Provides centralized access control, and auditing through robust security infrastructure

Data is the raw material of business – an economic input almost on par with capital and labor. However, data in most organizations is scattered across multiple operational and analytical systems. Big data with new sources, such as social media, cloud applications, and syndicated data services are on the rise, and many organizations are realizing that physical consolidation or replication of data is not practical for all data integration and business agility needs. Data needs to be made easily consumable by people who need it to advance the business.



JB0041



facebook.com/redhatinc  
@redhatnews  
linkedin.com/company/red-hat

## WHAT'S NEW

### BIG DATA INTEGRATION

- Hadoop integration with Hive and Impala support
- NoSQL database connectivity with full read/write access for Mongo DB, JBoss Data Grid, and Cassandra
- Apache Solr search integration

### CLOUD READY

- Available as part of OpenShift Online with simplified web development UI (developer preview)
- Cloud deployment on Amazon EC2 and Google Compute Engine
- Salesforce.com bulk API support for very large datasets

### PRODUCTIVITY AND DEPLOYMENT OPTIMIZATIONS

- Security audit log dashboard for complete view of data access activities
- Rapid development for custom data source connectivity
- Azul Zing JVM support for high-performance computing
- Query performance enhancement with dependent joins pushdown
- EAP 6.3 support for enhanced patching

## UNLOCK THE FULL POTENTIAL OF ALL YOUR DATA WITH LEAN DATA INTEGRATION

Red Hat® JBoss® Data Virtualization is a lean data integration solution that provides easy, real-time, and unified data access across disparate sources to multiple applications and users. JBoss Data Virtualization makes data spread across physically distinct systems – such as multiple databases, XML files, and even Hadoop systems – appear as a set of tables in a local database. This comprehensive platform enables agile data provisioning by providing the following functions:

- **Connect:** Access data from multiple heterogeneous data sources with different access methods and storage models.
- **Compose:** Easily create reusable, business-friendly data models and virtual unified views by combining and transforming data from multiple sources.
- **Consume:** Make integrated data available on demand for consumption by external apps through open standards interfaces.

JBoss Data Virtualization software implements all three steps internally, hiding all the complexities of the true location of the data or the mechanisms required to access or merge it, thus making it very easy for developers and users to work with data.

## DELIVER KEY I.T. INITIATIVES WITH AGILITY

The simplicity offered by JBoss Data Virtualization enables users to acquire actionable, unified information when they want, in the way they want, at the speed their businesses need. This leads to enlightened business execution and easy adaptation to changing business demands. Combined with ease of development, JBoss Data Virtualization supports a range of IT projects and initiatives.

**Self-service business intelligence (BI):** The virtual, reusable data model provides a business-friendly representation of data, allowing the user to interact with data without having to know the complexities of the database or where the data is stored, and allowing multiple BI tools to acquire data from a centralized data layer.

**Unified 360° view:** The virtual data model delivers a complete view of master and transactional data in real time. The virtual data layer serves as a unified, enterprise-wide view of business information that improves a user's ability to understand and use enterprise data.

**Agile service-oriented architecture (SOA) data services:** A data virtualization layer delivers the data services layer to SOA applications. JBoss Data Virtualization speeds both the creation of virtual data stores – without the need to touch underlying sources – and the creation of data services that encapsulate the data access logic. Data virtualization also allows multiple business services to acquire data from a centralized data layer and provides loose coupling between business services and physical data sources.

**Improved information control:** Data virtualization layers provide data firewall functionality. JBoss Data Virtualization improves data quality via centralized access control, robust security infrastructure, and reduction in physical copies of data, thus reducing risk. The metadata repository catalogs enterprise data stores and the relationships between the data in various data stores, enabling transparency and visibility.

## SUPPORTED DATA SOURCES

### ENTERPRISE RDBMS

- Oracle
- IBM DB2
- Microsoft SQL Server
- Sybase ASE
- MySQL
- PostgreSQL
- Ingres
- Maria DB

### ENTERPRISE DATA WAREHOUSE

- Teradata
- Netezza
- Greenplum

### BIG DATA

- Apache Hadoop
- HortonWorks
- Cloudera

### NOSQL

- Red Hat JBoss Data Grid (Infinispan)
- MongoDB
- Apache Cassandra (Tech preview)

### SEARCH

- Apache Solr

### ENTERPRISE AND CLOUD APPLICATIONS

- Salesforce.com
- SAP

Big data and cloud data integration: Data virtualization provides a rapid virtual integration approach that doesn't require replication of already "big" data sources. Many organizations are adopting cloud computing where each new cloud source must be integrated with the existing IT environment. Data virtualization solves this problem, allowing enterprises to maintain a complete view of internal and external information while taking advantage of attractive cloud economics.

## TRANSFORM INTO ENLIGHTENED BUSINESS

Red Hat JBoss Data Virtualization is a must-have for organizations seeking:

**Profitability growth and risk reduction.** JBoss Data Virtualization delivers unified and timely information to help your organization increase revenue, reduce costs, manage business risk, and reduce compliance penalties.

**Agility and productivity boosts.** Robust design and development environments let you respond faster to change and improve your staff's effectiveness and efficiency and realize faster time to value.

**Optimization of existing IT investments.** Data virtualization improves utilization of existing server and storage investments while reducing unnecessary data replication and the costs of duplication and associated infrastructure management.

## KEY CAPABILITIES AND FEATURES

### MODEL-DRIVEN DEVELOPMENT

JBoss Data Virtualization includes Teiid Designer, an Eclipse-based graphical tool, which models, analyzes, integrates, and tests multiple data sources to produce relational, XML, and web service views that show business data without programming. You can map from data sources to target formats using a visual tool, as well as resolve semantic differences, create virtual data structures at a physical or logical level, and use declarative interfaces that are compatible with and optimized for your applications.

### VIRTUAL DATABASE AS OPEN SEMANTIC MODEL

The virtual database is the semantic layer that resides between an organization's physical data sources and the user; it's a unified business representation of distributed data sources. It allows the user to interact with data with limited knowledge about it. The virtual database is created using familiar terminology to describe the business environment and allows the user to retrieve exactly the necessary data. This virtual database is accessible through standards-based provisioning mechanisms like SQL or web services, to any business intelligence tool or application.

### CENTRALIZED DATA SECURITY

JBoss Data Virtualization gives you the power to manage and monitor data services in a single unified environment and enforce and manage policies and roles across federated data for all data services. You can configure policies for data security, privacy, column-level data masking, and data sanitization of sensitive fields based on user roles. In addition to using the security capabilities of Red Hat JBoss Enterprise Application Platform, the software supports web services security and fine-grained access control for SQL data service and virtual table. In addition to column-level security, you get row-level security at the virtualization layer, which is independent of data sources. Transport and password encryption are available through SSL support.

**TECHNOLOGY CONNECTORS**

- Flat files, XML files, XML over HTTP
- SOAP web services
- REST web services
- OData services

**OFFICE PRODUCTIVITY**

- Microsoft Excel
- Microsoft Access
- Google Spreadsheets
- Apache POI for Excel

**SPECIALTY DATA SOURCES**

- ModeShape Repository
- Mondrian
- MetaMatrix
- LDAP

**SERVICE ENABLED DATA FOR SOA**

A data service approach hides the complexity of diverse data sources from the applications, enables the isolation of data-oriented logic in a data layer that is independent of applications, and renders the costly work of data transformation and integration much more valuable to the organization by making it more visible, reusable, and maintainable. Your IT team can build reusable logical data objects based on business entities (e.g., CUSTOMER, PRODUCT, and ORDER) and insulate applications from changes in data sources with a model-driven, standards-based data abstraction layer.

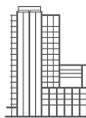
**PERFORMANCE OPTIMIZATION**

**Advance caching:** Support of several caching modes includes materialized views, result set caching, and code table caching that provide dramatic performance improvement. Configurable time-to-live, memory preferences, and updatability options are available for data caching.

**Query optimization:** Intelligent and automated techniques (e.g., cost- and rule-based query optimizer using information from source introspection, query capabilities, and constraints) include pushdown queries, dependent joins, projection minimization, partitioned aware unions, and copy criteria to optimize data sources join processing, and support for high-performance sub-queries. Comprehensive query trace support with manual plan override of automatic strategy selections for optimized query.

**METADATA REPOSITORY AND GOVERNANCE**

The OASIS SOA Repository Artifact Model and Protocol (S-RAMP) repository defines a common model for repositories and an interaction protocol to facilitate data sharing and use of common tools. The metadata repository catalogs enterprise data sources and the relationships between the data in various data stores, providing complete transparency and visibility. The repository supports Atom/REST-based client interface and XPath 2-based query language and integration with Maven.



**ABOUT RED HAT**

Red Hat is the world's leading provider of open source solutions, using a community-powered approach to provide reliable and high-performing cloud, virtualization, storage, Linux, and middleware technologies. Red Hat also offers award-winning support, training, and consulting services. Red Hat is an S&P company with more than 80 offices spanning the globe, empowering its customers' businesses.



facebook.com/redhatinc  
@redhatnews  
linkedin.com/company/red-hat

**NORTH AMERICA**  
1 888 REDHAT1

**EUROPE, MIDDLE EAST,  
AND AFRICA**  
00800 7334 2835  
europe@redhat.com

**ASIA PACIFIC**  
+65 6490 4200  
apac@redhat.com

**LATIN AMERICA**  
+54 11 4329 7300  
info-latam@redhat.com