

TIBCO® Data Virtualization

Reduce complexity and costs while gaining faster access to your data Integration connectors

Benefits

Economical

Integrate data reliably at a fraction of physical warehousing and ETL time, cost, and rigidity. Evolve rapidly when requirements change.

Immediate

Deliver up-to-the-minute data as needed, using advanced performance optimization algorithms and fine-grained security.

Business-friendly

Transform native IT structures and syntax into easy-to-understand, IT-curated data sets shareable via a self-service business directory.

Wide Range

Access data from distributed data sources including traditional enterprise, big data, cloud, and IoT. Use it across myriad analytics, self-service, business intelligence, and transactional applications.

Enterprise Grade

Support multiple lines of business, hundreds of projects, and thousands of users.

With data being the new competitive battleground, businesses that take advantage of their data will be the leaders, and those that do not will fall behind.

But gaining this advantage is a more difficult technical challenge than ever as your business requirements are ever-changing and your data is more-widely distributed across on-premises, big data, theInternet of Things (IoT), and the cloud. Traditional physical data integration, via data warehousing and ETL, is often too time consuming, too rigid, and too costly to support your dynamic business needs. There has to be a better way. There is and it's called data virtualization.

TIBCO® Data Virtualization software lets you integrate data with breakthrough speed and cost effectiveness. With it, you can build and manage virtualized views and data services that access, transform, and deliver the data your business requires to accelerate revenue, reduce costs and risk, improve compliance, and more.

TIBCO Data Virtualization is Java enterprise-grade middleware with a modular structure that supports all phases of data virtualization development, run-time, and management.

TIBCO Data Virtualization Modules

Studio is the agile modeling, development, and resource management tool that data-oriented developers use to model data, design, and view data services, build transformations, optimize queries, manage resources, and more. Easy to learn and use, Studio's graphical modeling environment provides a flexible workspace where queries are created and tested, as well as a data services repository where completed objects are published. Studio

also offers a rich set of transformations in addition to an easy-to-use transformation editor. Five languages complement Studio's graphical modeling capabilities: SQL, SQL Script, Java, XQuery, and XSLT.

Web UI provides a browser-based interface for the business users to discover and utilize self-service data within TIBCO Data Virtualization, using Data Catalog and Data Workbench functionality.

Adapters provide a wide range of data source connectivity for databases, files, big data, cloud sources, packaged applications, and more. Beyond schema-to-schema only connectivity, TIBCO Data Virtualization adapters integrate with data source optimizers to ensure more accurate queries and higher performance. The Data Source Tool Kit allows you to build additional custom adapters.

Cost-based and Rules-based Optimizers used by the Federation Engine work directly with data source optimizers to maximize query performance. The Objects Repository lets you manage your data services throughout their lifecycle. Myriad fine-grained Security capabilities, including authentication, authorization, and encryption, safeguard the delivery of sanctioned data only. Flexible Caching options enable higher performance and greater uptime. Quality helps ensure you deliver the best data possible. And built-in Governance features provide complete visibility, traceability, and control.

Manager is the administrative console used to set up user IDs, passwords, and security profiles, as well as view logs, check server activity, and more.

Deployment Manager lets you quickly and easily migrate entire projects in a single step, including their resources, cache settings, server configurations, security profiles, and more across instances to simplify and automate your development lifecycle.

Monitor provides a comprehensive, real-time view of your TIBCO Data Virtualization cluster. Monitor displays all the pertinent system health indicators required to help your IT operations staff guide corrective actions.

Active Cluster works in conjunction with load balancers to provide high availability and greater scale to meet your challenging service level agreements. Active Cluster simplifies complex operations management by automatically sharing resources, adjusting capacity on demand, and more.

Business Directory is a self-service data directory you can use to easily search, categorize, and consume IT-curated data sets developed using TIBCO Data Virtualization. Easy data set sharing and reuse helps accelerate good business outcomes while reducing IT workloads.

Discovery lets you go beyond simple data profiling to examine data, locate important entities, and reveal hidden relationships across distinct data sources. You can quickly build and display comprehensive entity-relationship (E-R) diagrams and data models so you can meet new business requirements faster and more easily.

Complete Data Virtualization Solution

Beyond TIBCO Data Virtualization software, TIBCO also provides a broad array of complementary advanced services, training, technical support, customer advisory program, knowledgebase, and partner offerings, so you get the complete solution needed to ensure your data virtualization success.

Deployment Options

TIBCO provides multiple options for deploying TIBCO Data Virtualization. You can install and run it on-premises, in your private cloud environment, or at a public cloud provider such as Amazon AWS, Google Cloud Platform, and Microsoft Azure. For AWS, TIBCO also provides an Amazon Machine Image (AMI) on the AWS Marketplace to simplify and accelerate deployment. TIBCO Data Virtualization can also be deployed in Kubernetes or Docker as containers.

Key Features

Development Environment

Tables 1 through 4 summarize key TIBCO Data Virtualization development capabilities.

Table 01 — Business Directory

Access and use data sets via an intuitive, self-service browser

FEATURE	DESCRIPTION
SEARCH OPTIONS	Find data sets using simple Google-like searches as well as advanced parameterized searches.
BROWSE	Find data sets using attributes such as data type, format, category, and more.
DATA PREVIEW	Preview specific data included within data sets.
SECURITY	See only the data sets you're supposed to see based on your TIBCO Data Virtualization enforced security profile.
CONSUMPTION	Use your favorite analytic/BI tools to consume data sets using Business Directory access information.
CATEGORIZATION	Organize large and/or diverse data sets into user-defined or system categories.
COLLABORATION	Allow business users and IT to interact via comments to improve data quality and utilization.
BUSINESS METADATA EXTENSIONS	Add custom definitions, properties, links, and status codes to enrich IT metadata with business metadata that your business users understand.
BUSINESS METADATA DISPLAY OPTIONS	Present business metadata in a shared area, an existing properties tab, or a new custom tab.
PERSONALIZATION	Receive email updates of changes and comments for data you care about.
PROVISIONING	Register additional TIBCO Data Virtualization instances to expand your data sets.
REST API	Access Business Directory programmatically to expose data sets to additional consuming applications.

Table 02—Modeling and Transformation

Design reusable views and data services within an agile, high-productivity development environment

FEATURE	DESCRIPTION
GRAPHICAL DEVELOPMENT ENVIRONMENT	Graphically model data, design view and data services, build transformations, optimize queries, manage resources, and more.
AGILE DEVELOPMENT METHODOLOGY	Develop and refine views and data services iteratively.
INTROSPECTION	Automatically probe physical data sources and select desired resources. Optionally inspect data sources interactively.
DATA DISCOVERY	Reveal data relationships across disparate entities using formal keys and fuzzy matching.
BOTTOM-UP MODELING	Design views and data services by combining data from disparate systems without worrying about underlying source access and format complexity.
FLEXIBLE MODELING AND TRANSFORMATION METHODS	Create views and data services graphically or via scripting languages as appropriate. SQL, SQL script, XQuery, XSLT, and Java functions.
SQL	Use familiar SQL standards.
SQL SCRIPT	Implement stored procedures using a familiar scripting language.
XQUERY	Create complex XML structures using a graphical XQuery editor.
XPATH TRANSFORMATION	Establish arbitrary complex mapping of XML schema elements to XML output.
JSON QUERYING AND TRANSFORMATION	Query and transform JSON data from Web services into a relational format.
ANALYTIC FUNCTIONS	Use a full set of analytic functions such as CORR, COUNT, NTILE, STDDEV, and VARIANCE.
CONTRACT FIRST DESIGN	Build data services using preexisting WSDLs and schemas.
CONTRACT LAST DESIGN	Define the Java wrapper first and then develop WSDL.
QUERY PLAN	View query execution plan steps and details.
VIEWS DEPENDENCY GRAPH	Graphically display dependencies between data sources and views and data services.
PHYSICAL TABLE CREATION	Create and drop physical tables within a designated data source.
LOCALIZED UI	Chinese and Japanese language support.

Table 03—Metadata Repository: Store and manage relevant metadata

FEATURE	DESCRIPTION
COMPLETE REPOSITORY	Manage resources such as data sources, views and data services and procedures throughout their life cycles.
PUBLIC METADATA API	Deploy a web services metadata API for easy access and sharing.
SCHEMA CHANGE NOTIFICATION	Receive notice when data source schemas change.
SOURCE METADATA	Access the metadata of the physical data source.
OPEN API	Open access to TIBCO Data Virtualization system libraries to enable custom scripting and orchestration.

Table 04—Version Control: Control the view and data service development lifecycle

FEATURE	DESCRIPTION
MIGRATION GUI	Add project folders directly to version control systems. Check-in and check-out folders or individual resources to track changes. Rollback revisions of folders or resources to prior versions.
RESOURCE LOCKING	Protect against inadvertent modifications and overwrites.
CHANGE HISTORY	Track changes made by users with annotations.
SOURCE CONTROL AND TRANSPORT	Manage artifacts and transports from development through production.
VERSION CONTROL SYSTEMS	Integrate directly with Apache® Subversion® and Git.

Run-time Environment

Table 05—Federated Query Engine: Run optimized queries across data sources

FEATURE	DESCRIPTION
FEDERATION ENGINE	Join and aggregate data that is vertically and horizontally partitioned.
DATA SOURCE	Leverage data source optimizers to ensure query accuracy and maximize query performance.
COST-BASED OPTIMIZER	Use statistics to create an optimal query plan that reduces unnecessary data flow across the network.
RULE-BASED OPTIMIZER	Allow users to specify exactly how they want to run a particular query.
SCHEDULING	Run queries based on set times.
ALERTTRIGGERS	Implement resource, event, and user-defined triggers. Use a published API to handle custom Java alerts.

Table 06—Performance Optimization Algorithms and Techniques: Optimize query performance

FEATURE	DESCRIPTION
COMPLETE SET OF JOIN ALGORITHMS	Automatically rewrite the query to use the most efficient join strategy (for example, hash join, sort-merge join, distributed semi-join, data-ship join, union-join flip, nested-loop join, and others).
SINGLE-SOURCE JOIN GROUPING	Run data-reducing joins at the data source rather than bringing the data across the network.
PREDICATE PUSH-DOWN	Push WHERE clause predicates down into the underlying data source to reduce data at the source.
FULL AND PARTIAL AGGREGATE PUSH DOWN	Push aggregate functions down to source when applicable.
SERIALIZATION OR PARALLELIZATION OF JOIN OPERATORS	Determine the proper join order and join algorithms based on estimated cardinality and join results derived from data distribution histograms.
PROJECTION PRUNING	Eliminate all unnecessary columns from fetch nodes in a query tree.

CONSTRAINT PROPAGATION	Distribute filters to multiple branches of the query plan, allowing data reduction by a single filter to potentially occur in multiple data sources.
SCAN MULTIPLEXING	Reuse data sets that appear in multiple places in a single query plan.
EMPTY SCAN DETECTION	Detect logical conditions that would produce empty data sets, and then eliminate those parts of the query plan prior to processing.
REDUNDANT OPERATOR CROPPING	Eliminate redundant or extraneous operators within a complex multiple-operator query.
BLOCKING OPERATOR PREFETCHING	Proactively run parts of the query plan that must finish before other parts of the query plan can continue, thereby increasing the overall responsiveness of the query.
CONNECTION POOL SHARING	Share access to data sources to avoid bottlenecks.
RESULTS STREAMING	Stream data to consuming applications as results are processed at the underlying sources.
HYBRID MEMORY AND DISK USE	Balance memory and disk use for optimal performance.
NATIVE XML SUPPORT	Support XML internally for fast parsing and joins.
API	Expose the query execution plan via JDBC/ODBC.

Table 07—Caching: Move data to predesignated caches to boost availability and performance

FEATURE	DESCRIPTION
EVENT-BASED REFRESH	Refresh caches based on defined business rules.
SCHEDULED REFRESH	Refresh caches at set times.
MANUAL REFRESH	Refresh caches on demand as needed.
INCREMENTAL REFRESH	Refresh part of a cache.

FULL REFRESH	Refresh the entire cache.
NATIVE BULK EXTRACT	Use native bulk EXTRACT functions in the source to extract data more efficiently than using SELECTs.
NATIVE BULK LOAD	Use native LOAD functions in destination-to-load and refresh caches more efficiently than using INSERTs.
PARALLEL LOAD	Use multiple threads to load caches in parallel.
CENTRALIZED CACHE TRACKING	Centrally track caching tables distributed across multiple data sources.
MULTI-TABLE CACHING	Avoid contention on cache refreshes, accelerate refresh speeds, and maintain non-stop availability using multiple tables per cache view.

Table 08—Data Access: Connect and use data from distributed sources

FEATURE	DESCRIPTION
DATABASES	Access popular databases via Open Database Connectivity (ODBC) and Java Database Connectivity (JDBC).
BIG DATA	Access Apache Hadoop® through Apache Hive™, Apache Impala®, and Apache HBase®.
MULTIDIMENSIONAL DATA	Access multidimensional data in SAP BW.
NOSQL AND CLOUD DATABASES	Access NoSQL and cloud database sources such as Amazon DynamoDB, Amazon Redshift, Cassandra, and MongoDB.
WEB SERVICES	Access SOAP over HTTP, XML over HTTP, and Java Message Service (JMS) services. A message pipeline allows interjection of custom logic during the web service request and response.
PACKAGED APPLICATIONS	Access SAP, Oracle E-Business Suite, Salesforce.com, and other applications using standard objects such as invoices, shipments, orders, customers, opportunities, and more.
SAAS APPLICATIONS	Access SaaS applications including Google Analytics, Microsoft Dynamics, and NetSuite.
COLLABORATION	Access collaboration apps such as email, Google Sheets, and Microsoft SharePoint.
SOCIAL MEDIA	Access social media sources such as Facebook, LinkedIn, RSS, and X.
JAVA API	Access non-relational sources using custom procedures.

DATA SOURCE TOOL KIT

Access a set of libraries of services that can be imported into your preferred Integrated development kit to facilitate and accelerate data adapter creation. Services include database mapping, data type mapping, syntax mapping, and function mapping. Together these services minimize custom code development.

Table 09—Data Delivery: Deliver requested data to myriad consuming applications

FEATURE	DESCRIPTION
DATABASE OBJECTS	Publish views for consumption through ODBC, JDBC, and ADO. NET.
WEB SERVICES	Publish data services in the form of WSDL for consumption using SOAP or SOAP over JMS. A message pipeline allows interjection of custom logic during the web service request and response.
REPRESENTATIONAL STATE TRANSFER (REST)	Publish data services in the REST format. REST CREATE, READ, UPDATE, and DELETE functions are supported.
OPEN DATA (ODATA) PROTOCOL	Publish data services in the OData format.
JSON	Publish JSON, including more formatting capabilities for XML- to-JSON translation, for both procedures and table outputs.
BI TOOL INTEGRATION	Create temporary tables in BI tools to store filters for visualizations or reports.
DDL SUPPORT FOR CLIENT APPLICATIONS	Support client applications that create tables using DDL statements via JDBC/ODBC/ADO. CREATE TABLE, CREATE TABLE AS SELECT, and CREATE TEMP TABLE syntax. Multiple physical data source mappings are supported for each published database to avoid conflicts between users when creating tables.

Table 10—Security: Protect sensitive data with fine-grained, multi-level security

FEATURE	DESCRIPTION
POLICY-BASED SECURITY	Apply authentication, authorization, and encryption rules via policies.
SINGLE SIGN-ON	Sign on once to access all integrated data sources and consuming applications.
ROW-LEVEL AUTHENTICATION	Control access to specific rows via granular permissions.
COLUMN-LEVEL AUTHENTICATION	Control access to specific columns via granular permissions.

AUTHENTICATION TYPES	Support basic, pass-through, Kerberos, SAML, and NTLM.	
COLUMN MASKING	Implement column masking rules to hide, replace, or obfuscate portions of a column's value depending on a user's level of access.	
SSL OVER HTTP WITH SUPPORT FOR MUTUAL AUTHENTICATION	Mutually authenticate published services, web services data sources, and Oracle databases. Certificate-based authentication and Web Services Security (WSS) authentication are supported.	
PASS-THROUGH	Use an existing user ID and password and pass through to TIBCO Data Virtualization for authentication.	
LIGHTWEIGHT DIRECTORY ACCESS PROTOCOL (LDAP)	Use security profiles from LDAP to authenticate user access to protected data sources.	
PLUGGABLE AUTHENTICATION MODULE	Use third-party systems for authentication.	
ENCRYPTION	Encrypt passwords and data in motion via TLS.	
ACCESS MANAGEMENT	Use TIBCO Data Virtualization software as the system of record for security roles and profiles.	

Table 11—Governance: Provide complete visibility, traceability and control

FEATURE	DESCRIPTION
DATA LINEAGE	Trace lineage from multiple data sources to a single data consumer.
COLUMN LINEAGE API	Access column lineage via an API.
WHERE-USED	Trace where-used from a single data source to multiple data consumers.
LOGGING	Track system and/or user activity.
STANDARDS ENFORCEMENT	Implement internal and industry data standards.
OPEN API	Open access to TIBCO Data Virtualization system libraries to enable custom scripting and orchestration.

Table 12—Data Quality: Help ensure correct and complete data

FEATURE	DESCRIPTION	
STANDARDIZATION AND CONFORMATION	Create views and data services that conform to agreed standards.	
ENRICHMENT AND AUGMENTATION	Extend views and data services with additional data.	
VALIDATION	Validate data sets with users prior to publishing views and data services.	
OBJECT REUSE	Share views and data services to ensure consistent data definitions.	

Management Environment

Tables 13 through 15 summarize key TIBCO Data Virtualization management capabilities.

Table 13—Management: Administer and manage for efficient operations

FEATURE	DESCRIPTION	
MANAGEMENT CONSOLE OPTIONS	Access the management console through TIBCO Data Virtualization Studio or a web browser.	
USER SETUP	Set up user and group profiles.	
SECURITY	Enable multiple forms of security to increase data protection.	
SCHEDULING	Schedule TIBCO Data Virtualization activities flexibly.	
DEPLOYMENTS	Manage tasks related to development, management, configuration, and versioning.	
REAL-TIME SYSTEM INDICATORS	Monitor critical system metrics and tune for optimal performance including memory usage, query plans for currently running and past requests, data sources, and caches.	
USAGE METRICS	Deliver usage activity detail to your reporting tool of choice via an open API.	
SIMPLE NETWORK MANAGEMENT PROTOCOL (SNMP)	Allow monitoring by third-party systems via SNMP API.	

Table 14—Active Cluster: Cluster servers to meet reliability, availability and scalability SLAs

FEATURE	DESCRIPTION
FLEXIBLE CLUSTER DEPLOYMENT	Expand capacity and improve availability simply by adding new nodes to an existing cluster or adding new clusters.
ACTIVE-ACTIVE CLUSTERING	Distribute TIBCO Data Virtualization workloads across cluster nodes in conjunction with load balancers.
SHARED CLUSTER CACHE	Improve overall cluster performance by coalescing redundant data source hits and reducing data latency.
REPLICATED METADATA REPOSITORY	Replicate metadata across clusters to simplify management.
RESTORE CLUSTER NODES	Restore cluster nodes using Web Manager, cluster_util script, or API.

Table 15—Deployment Manager: Automate migration or promotion of artifacts, configurations, and settings

FEATURE	DESCRIPTION	
RESOURCE MIGRATION	Migrate or promote (create/update/delete) artifacts from one TIBCO Data Virtualization instance to another.	
CACHE SETTING MIGRATION	Migrate or promote cache table names, caching methods, refresh method, and cache policies and schedules.	
SERVER CONFIGURATION MIGRATION	Replicate server configurations, for example, enabling and disabling triggers.	
USER/GROUP MIGRATION	Migrate or promote user and group IDs, security profiles, and other user and group information.	

Table 16—Web UI: Browser-based interface to search and self-service data

FEATURE	DESCRIPTION
DATA CATALOG	Search and discover data within TIBCO Data Virtualization. Support search by metadata such as column, annotations, description, etc.
DATA WORKBENCH	Self-service data by creating data flows in Data Workbench. Flows can be created by joining, filtering or aggregating existing data sources, and then publishing the flows for external consumption.

Technical Specifications

Tables 17 through 18 summarize TIBCO Data Virtualization technical specifications and standards support.

Table 17. Supported Platforms: Leverage standard operating systems and high performance servers

TIBCO DATA VIRTUALIZATION STUDIO CLIENTS	SERVER	JVM
· Microsoft Windows	 CentOS Cisco Unified Computing System Oracle Linux and Red Hat compatibility mode Red Hat Enterprise Linux AS SUSE Enterprise Linux Microsoft Windows Solaris 	· 64-bit

Table 18—Data Source Adapters: Simplify data access while ensuring query accuracy and performance

STANDARD DATA SOURCE ADAPTERS

GENERAL PURPOSE ODBC AND JDBC DATA SOURCE ADAPTERS

CRM AND ERP ADAPTERS

- · Apache Drill
- · Apache Impala
- Custom Java procedure
- Cloudera CDH
- Data direct mainframe
- Files (cache, delimited, and XML)
- Greenplum
- Hadoop/Hive
- Hbase
- · Hortonworks HDP
- · HSQLDB
- · IBM DB2
- · IBM DB2 z/OS
- · Informix
- · LDAP
- · Microsoft Access
- · Microsoft Excel
- · Microsoft SQL Server
- Mock File Delimited
- MySQL
- Netezza NPS
- · Oracle
- · Odata
- · Parstream
- PostgreSQL
- · SAP HANA
- · Sybase
- Sybase IQ
- · Teradata
- TIBCO Data
 Virtualization
- · Vertica
- · WSDL
- XML (flat files over HTTP)

- ADO.NET
- · ODBC for Linux, AIX, HP-UX and Solaris
- · JDBC
- · Microsoft Windows
- Teradata
- Vertica
- Data Source Tool Kit
- Software Development Kit for Data Source Adapter Development

COLLABORATION ADAPTERS

- Email
- · Google Apps
- · Google Sheets
- · Microsoft Active Directory
- Microsoft SharePoint (onpremises and online)
- Microsoft SharePoint Excel services

NOSQL AND CLOUD DATABASE ADAPTERS

- · Amazon DynamoDB
- Amazon RedShift
- Cassandra
- Couchbase
- Google BigQuery
- · HBase
- MongoDB
- Google Cloud Storage

- CRM and ERP Adapters
- Microsoft Dynamics CRM (on-premises and online)
- Microsoft Dynamics
 GP
- Microsoft Dynamics NAV
- NetSuite CRM
- NetSuite ERP
- Oracle EBS
- Salesforce.com
- Siebel

MARKETING AUTOMATION ADAPTERS

- Google Adwords
- Google Analytics
- HubSpot
- Marketo
- Oracle Eloqua

SAP ADAPTERS

- SAP Netweaver
- mySAP Business
 Suite
- · SAP BW
- SAP Business Explorer (BEx)

SOCIAL MEDIA ADAPTERS

- Facebook
- LinkedIn
- RSS
- . X

Table 19—Additional Technical Specifications: Take advantage of industry standards

DELIVERY INTERFACES	SECURITY PROTOCOLS	CACHE REPOSITORIES
- ADO.NET - ODBC - Hadoop - JDBC - REST - SOAP - SOAP and JMS TIBCO Enterprise Message Service™ - SOAP and JMS Sonic MQ ENTERPRISE SERVICE BUSES - OpenMQ - Sonic MQ - Enterprise Message Service™ WEB SERVICES PROTOCOLSNET - OData - REST and JSON - SOAP - WSDL - WSI - XPath - XQuery - XSLT - XML (flat files or over HTTP)	 Base64 Kerberos NTLM SAML SSL WS-Security DIRECTORY SERVICES Active Directory Oracle Directory Server Enterprise Edition Novell eDirectory TIBCO Data Virtualization SOFTWARE DEVELOPMENT STANDARDS SQL 92 and 99 Unicode JDK J2EE JNDI 	 File Greenplum HSQLDB IBM DB2 Microsoft SQL Server MySQL Netezza Oracle PostgreSQL SAP HANA Sybase ASE Sybase IQ Teradata Vertica DATA SHIP SOURCES AND TARGETS IBM DB2 Microsoft SQL Server Netezza Oracle PostgreSQL Sybase IQ Teradata

