Cisco Information Server 7.0

"At Pfizer, we have all the data integration tools that you can find on the market. But when senior execs come to me daily with key project/resource questions whose answers will determine the courses of action we'll take in running our business, my team uses the rapid deployment methods that are built within CIS. This reduces each project from 4-6 weeks to 2-3 days."

—Dr. Michael Linhares, Director of Business Operations, Pfizer

The Cisco® Data Virtualization Suite is data integration software that makes your enterprise a more agile business. It lets you quickly build logical business views of data scattered throughout your enterprise and delivers greater insight into your organization. Cisco Data Virtualization increases the value of your network and other IT assets, without the long delays of data replication and physical consolidation traditionally required to achieve your goal of a unified view of the business. The Cisco Data Virtualization Suite empowers your organization to achieve profitable growth, risk reduction, productivity, and effectiveness.

The Cisco Information Server (CIS) is the foundation of the Cisco Data Virtualization Suite. It is a Java-based server that accesses existing data noninvasively, federates disparate data, abstracts and simplifies complex data, and delivers the results as data services or relational views (that is, logical business views) to consuming applications such as business intelligence, analytic tools, and other information dashboards. With advanced query optimization technology, Cisco Information Server delivers extremely high performance.

The Cisco Data Virtualization Suite also consists of complementary components and options that make the Cisco Information Server more powerful: for example, enhancing its ease of use and scalability.

CIS Business Directory is an intuitive web-based interface for business users to easily access and use business data. Business Directory provides a self-service directory of virtualized business data contained in CIS to empower users to search, browse, and collaborate on all available data, and categorize large, diverse data sets. Users can use their preferred analytic or BI tools to obtain the desired data in Business Directory to drive their business decisions and actions without the need to have a higher level of technical expertise to view the data using CIS Studio.

Cisco Information Server Studio is the primary modeling, view, service development, and resource management environment used by data-oriented developers familiar with data management concepts such as entity relationship (E-R) diagrams, SQL, and JavaScript. The modeling environment presents a familiar resource-tree view of available physical data sources, a workspace area in which queries are created and tested, and an area in which views can be published for use by consuming applications.

Cisco Information Server Manager is the administrative console for data virtualzation. Manager lets administrators set up user IDs, password, security profiles, view logs, check status of underlying resources, and more. Manager could be accessed via Studio or through a Web browser.

Cisco Information Server Discovery enables you to go beyond profiling to examine data, locate important entities, and reveal hidden relationships across data sources. You can use that knowledge to quickly build and display comprehensive E-R diagrams, data models, and live data, to validate business requirements with end users faster and more easily.

Cisco Information Server Monitor provides a comprehensive, real-time view of your Cisco Data Virtualization Suite environment. Whether the environment is a single server or a cluster of servers, Cisco Information Server Monitor displays all the pertinent system health indicators necessary to assess current conditions. If processes slow down or operations fail, your IT operations staff can use these insights to guide the actions required to meet service-level agreements (SLAs).

Cisco Information Server Adapters provide connectivity to critical data sources in categories such as Collaboration (Sharepoint, Google Docs), Social Media (Facebook and Twitter), Marketing Automation (Marketo, Eloqua), CRM/ERP, SAP and many more. The adapters enable standard, SQL-based query or web-services access by any front-end or reporting tool into your packaged applications. CIS adapters include the most common application objects, so you can more easily deliver the data required to meet your business needs. Cisco also provides a feature-rich toolkit for developing CIS adapters to non-standard data sources. The Data Source SDK (software development kit) enables more sources to be virtualized to enrich analytics, BI, and othe information applications. Our vendor neutral approach opens up CIS to a larger community of consultants and developers to add greater value to the platform.

The Cisco Information Server Active Cluster feature allows substantial scaling of your data virtualization deployments and maintains continuous availability of your data services. The Cisco Information Server Active Cluster feature enables you to fulfill SLAs by easily increasing capacity on demand, simplifying scaling, and improving the manageability of your data services environment.

Deployment Manager simplifies and automates CIS deployment management by allowing administrators and users to quickly and easily migrate resources, cache settings, server configurations, security profiles and other information from one CIS instance to another (for example, promoting resources and settings from the development environment to staging to production). For large-scale deployments, Deployment Manager is a critical tool to help minimize deployment risks and promote enterprise software implementation best practices as development implementations are moved into production.

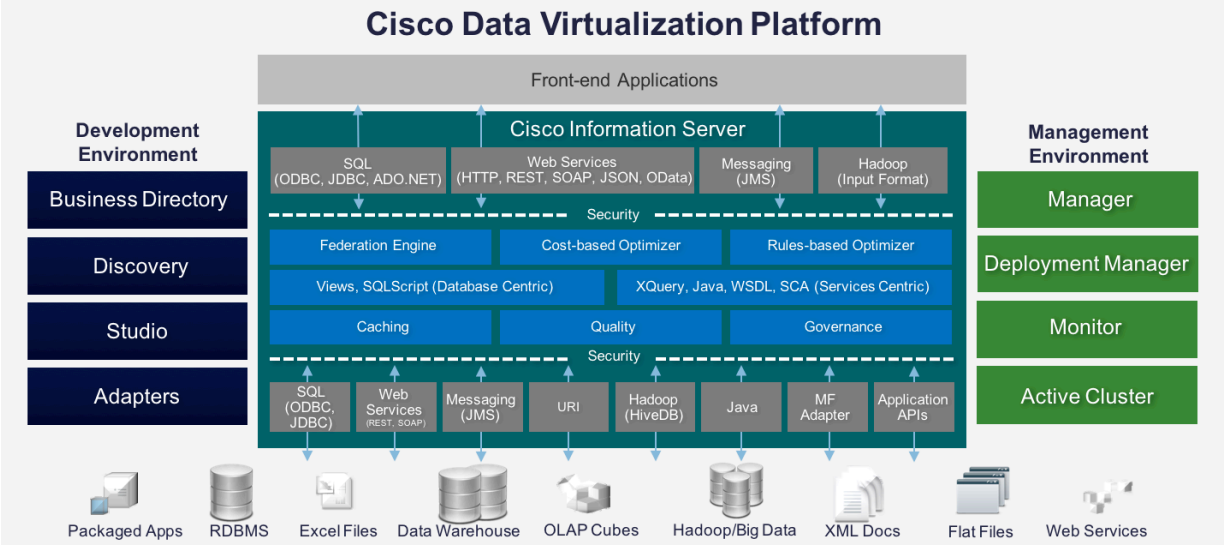

Figure 1. Cisco Information Server Architecture

Main Features

Development Environment

Tables 1 through 7 summarize the main features for the development environment.

Table 1. Business Directory: Access and Use Business Data Through an Intuitive Web-Based Interface

Feature	Description
Search	Simple Google-like searches to advanced parameterized searches make it easy to find the data you need.
Browse	Browse by data type, format, category, etc. to learn all about the available data.
Use	Use your favorite analytic/BI tools to query the desired data set, using the access information provided by Business Directory.
Security	Users only see the data you're supposed to see based on your existing security profile.
Categorize	Group data into system or user-defined categories to organize large, diverse data sets.
Document	Add custom definitions, properties, links and status codes to enrich content about data.
Collaborate	Business users and IT can interact via comments to improve quality and usefulness of the data.
Personalize	Users can watch the data they need to receive email updates of changes and comments.
Provision	Registering more Cisco Information Server instances makes even more IT-curated data available to users.
API	REST API exposed for external consumption, allowing programmatic access to the Business Directory server.
Data preview	Supports a data preview for tables; uses the existing security setup on CIS.
Custom properties	Present business metadata in a shared area at the top of the page, on an existing Properties tab, or on a new custom tab.

Table 2. Modeling: Design Views and Services Using a High-Productivity Development Environment

Feature	Description
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 by combining data from disparate systems without worrying about underlying source access and format complexity.
Custom logic	Create views by using scripting languages for added convenience and flexibility.

Contract first	Define Web Services Description Language (WSDL) first and then develop the Java wrapper.
Contract last	Define the Java wrapper first and then develop WSDL.
Views dependency graph	Graphically display the data sources on which views depend, and the views on which data sources depend.
Query plan	View all the steps and details related to the processing of a view or procedure.
SQL script	Implement stored procedures using a familiar scripting language.
SQL 99	Support of the latest SQL standards.
Physical table creation	Create and drop physical tables in a single data source.
Analytic functions	Support a full set of analytic features (CORR, COUNT, NTILE, STDDEV, VARIANCE, etc.).
Localization	Supports Chinese and Japanese.

Table 3. Transformation: Transform Complex Data Structures with an Easy-to-Use Editor

Feature	Description
XPath transformation	Establish arbitrary complex mapping of XML schema elements to XML output.
XML shaping	Transform data from tabular to hierarchical, and from hierarchical to tabular. Access XML data with SQL.
XQuery	Create complex XML structures using a graphical XQuery editor.
Contract first development	Build services using preexisting WSDL and schemas.
JSON querying and transformation	Query and transform JSON data from Web services into a relational format.

Table 4. Metadata Repository: Store and Manage Metadata

Feature	Description
Complete repository	Manage resources such as data sources, views, and procedures throughout their lifecycles.
Public metadata API	Deploy an API containing a collection of web services and built-in procedures.
Schema change notification	Receive notice when data source schemas change.
Metadata exposure	Access the metadata of the physical data source.
Open API	Deploy a web services–based metadata API for easy access and sharing.
Temp tables	BI tools and consuming apps can store filters for visualizations or just create working space.

Table 5. Version Control: Manage Development and Changes in a Controlled Manner

Feature	Description
Version-based import and export	Save and restore resources from directories and files.
Resource locking	Protect against inadvertent modifications and overwrites.
Change history	Track changes made by users with annotations.
Source control and transport	Manage data artifact sources and transport from development environments to production environments.
Multi-container support for DDL	Avoid conflicts when multiple users access the same tables.

Table 6. Governance: Provide Complete Visibility, Traceability, and Control

Feature	Description
Data lineage	Trace data from the data source to the data consumer. An API is available for third-party systems to access log data.
Logging	Implement full activity tracking, or track only selected components of the server.
Standards enforcement	Apply industry data standards in views and services development.
Policy-based security	Apply authentication, authorization, and encryption.

Table 7. Data Quality: Help Ensure Correct and Complete Data

Feature	Description
Standardization and conformation	Create views that meet agreed standards.
Enrichment and augmentation	Extend views with data from other sources.
Validation	Verify data values and user requirements.
Masking	Present data within contextual formatting.

Run-Time Server

Tables 8 through 13 summarize the main features for the run-time server.

Table 8. Query Engine: Run Optimized Queries on a Single Data Source or Across Multiple Disparate Data Sources

Feature	Description
Federation engine	Join and aggregate data that is vertically and horizontally partitioned.
Cost-base optimizer	Use statistics to create an optimal query plan that reduces unnecessary data flow across the network.
Rule-base optimizer	Allow users to specify exactly how they want to run a particular query.
Hybrid memory and disk use	Balance memory and disk use for optimal performance.
Transformation	Shape data using XQuery, XSLT, Java, and SQL functions.
Alerts	Implement resource, event, and user-defined triggers. Use a published API to handle custom Java alerts.
Scheduling	Run queries based on set times.

Table 9. Performance Optimization Algorithms and Techniques: Optimize Query Performance for Large and Complex Data Sets

Feature	Description
Complete set of join algorithms	Select and employ the most efficient join strategy for a given situation (for example, hash join, sort-merge join, distributed semi-join, data-ship join, and nested-loop join) to help ensure the most efficient data processing.
Single-source join grouping	Run data-reducing joins in the data source rather than bringing the data across the network.
Predicate push-down	Push WHERE clause predicates all the way down into the underlying data source to reduce data in the source.
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.
Results streaming	Stream data to consuming applications as results are obtained and processed from the underlying sources.
API	Expose execution plan via JDBC/ODBC.

Table 10. Caching: Move Data to a Designated Storage Location to Boost Availability and Performance

Feature	Description
Event-based refresh	Update the cache based on defined business rules.
Scheduled refresh	Update the cache based on set times.
Incremental refresh	Update a part of the cache based on triggered changes.
Manual refresh	Update the cache on demand as needed.
Native data source load	Use target repository native load functions to load and refresh the cache.
Parallel load	Use multiple threads to load the cache in parallel.

Table 11. Data Access: Connect and Expose Data from Diverse Sources

Feature	Description
Connection pool sharing	Share access to data source to avoid bottlenecks.
Big data	Access Hadoop through Hive and MPP-based analytic appliances such as IBM Netezza and HP Vertica.
Collaboration	Access collaboration apps such as Email, Google Spreadsheets, and Microsoft SharePoint.
Databases	Connect to standard databases using Open Database Connectivity (ODBC) and Java Database Connectivity (JDBC).
Marketing Automation	Access marketing automation services such as Google Analytics, Marketo, and Oracle Eloqua.
Multidimensional data sources	Access multidimensional data sources such as SAP NetWeaver BW using Cisco Information Server adapters.
Native XML support	Support XML internally for fast parsing and joins.
NoSQL and Cloud DBs	Access sources such as Amazon DynamoDB, Amazon RedShift, and MongoDB.
Packaged applications	Connect to SAP, Oracle E-Business Suite, Salesforce.com, and other applications through their approved APIs using Cisco Information Server adapters.
Social Media	Access Social Media sources such as Facebook, LinkedIn, RSS, and Twitter.
Web services	Consume Simple Object Access Protocol (SOAP) over HTTP and Java Message Service (JMS). XML over HTTP is supported. A message pipeline allows interjection of custom logic during the web service request and response.
Java API	Access non-relational sources using custom procedures.
Data Source SDK(software development kit)	Access a set of libraries of services that can be imported into your preferred Integrated development kit to facilitate and accelerate CIS data adapter creation. Services include database mapping, data type mapping, syntax mapping, and function mapping minimize custom code development.

Table 12. Data Delivery: Deliver Data to Consuming Applications

Feature	Description
Database objects	Publish data models in the form of views and procedures 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.

Table 13. Security: Support Multiple Forms of Security to Increase Data Protection

Feature	Description
Single sign-on	Sign on once to access all integrated applications and data sources.
Row-level authentication	Implement data access authentication to the row level.
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 Cisco Information Server 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.
Access management	Use Cisco Information Server as the system of record for security roles and profiles.

Management

Tables 14 through 16 summarize the main management features.

Table 14. Management: Administer, Manage, and Optimize for Efficient Operations

Feature	Description
Multiple-access management console	Access the management console through Cisco Information Server Studio or a web browser.
Real-time system indicators	Monitor critical system metrics and tune for optimal performance. • Monitor the memory use of the Cisco Information Server. • View the query plan for currently running and past requests. • Check the status of all underlying data sources and cached resources.
Scheduling	Schedule loads of individual cached resources and groups of resources using policies.
Security	Set up user profiles and groups that support multiple forms of security to increase data protection
Deployment	Manage tasks related to application development, management, configuration, and versioning.
Simple Network Management Protocol (SNMP) support	Allow monitoring by third-party systems.
View usage metrics	View usage activities data are available for analytics.

Table 15. Active Cluster: Substantially Scale CIS Deployments

Feature	Description
Active/active clustering	Provides maximum scalability of the enterprise platform and allows companies expand capacity on-demand by simply adding new servers to the cluster.
Shared cluster cache	Improves overall cluster performance by coalescing redundant data source hits and reducing data latency.
Replicated metadata repository	Simplifies deployment of large clusters and improves manageability of the overall solution
Restore cluster nodes	Using Web Manager, cluster_util script, or API

Table 16. Deployment Manager: Automate Migration or Promotion Artifacts, Configurations, and Settings

Feature	Description
Resource migration	Transfer or promote (create/update/delete) artifacts from one CIS instance to another.
Cache setting migration	Transfer or promote cache table names, caching methods, refresh method, cache policies and cache schedules.
Server configuration migration	Replicate server configurations (for example, enabling and disabling triggers).
User/group migration	Transfer or promote user/group IDs, security profile, and other information.

Specifications

Table 17 Summarizes the Cisco Information Server specifications.

Table 17. Specifications

ODBC and JDBC

- ADO.NET
- iODBC 3.521 for Linux, AIX, HP-UX, and Solaris
- JDBC 3.0 and 4.0
- Microsoft Windows
- Teradata 14.10.00.17 JDBC
- Vertica 6.01.200 JDBC

Standard Data Sources

- Apache Drill
- Cisco Information Server
- Custom Java procedure
- Cloudera CDH4, CHD5.3
- Cloudera Impala 1.0, 2.0
- Data direct mainframe
- Files (cache, delimited, and XML)
- Greenplum 3.3 and 4.1
- Hbase 0.98
- Hadoop/Hive 0.10, 0.12, 0.13, 0.14
- Hortonworks HDP 2.1, 2.2
- HP Neoview 2.3 and 2.4
- HSQLDB 2.2.9
- IBM DB2 9 (Type 2 and 4), 10.5 (Type 4)
- IBM DB2 z/OS 9 and 10
- Informix 9.x
- LDAP 3
- Microsoft Access
- Microsoft Excel
- Microsoft SQL Server 2008, 2012, 2014
- Mock-File-Delimited
- MySQL 5.1 and 5.5
- Netezza NPS 6.0, 7.0, 7.2
- Oracle 11g, RAC, and 12c
- PostgreSQL 9.0, 9.1
- SAP HANA SPS 09
- Sybase 12, 12.5 ASE, 15, 15.5 ASE
- Sybase IQ 15 and 15.2
- Teradata 13, 13.10, 14, 14.10, 15
- WSDL 1.1
- Vertica 6.1
- XML (flat files over HTTP)

Data Ship Sources and Targets

- IBM DB2 LUW v9.5 (target only)
- MS SQL Server 2008, 2012, 2014
- Netezza 6.0 and 7.0
- Oracle 11g and 12c
- PostgreSQL 9.0, 9.1
- Sybase IQ 15
- Teradata 13, 13.10, and 14
- Vertica 5.0 and 6.1

Delivery Interfaces

- ADO.NET
- ODBC 3.521
- Hadoop
- JDBC 2 SE 1.5.0, 1.6.0, 2.0, and 3.0
- SOAP 1.1 and 1.2
- SOAP and JMS: TIBCO EMS and Sonic MQ
- REST

Enterprise Service Buses

- Sonic 7.5
- TIBCO EMS 4.4
- OpenMQ 4.4

Web Services Protocols

- .NET 2.0, 3.0, 4.0, and 4.5 (client side)
- OData
- REST and JSON
- SOAP 1.1 and 1.2
- WSDL 1.1
- WSI 1.0 and 1.1
- XPath 1.0 and 2.0
- XQuery 1.0
- XSLT 1.1 and 2.0
- XML (flat files or over HTTP)

Directory Services

- Active Directory 2008, 2012
- Oracle Directory Server Enterprise Edition 11.1
- Cisco Information Server
- Novell eDirectory 8.8

Other Standards

- SQL 92 and 99
- Unicode support
- JDK 1.6, J2EE 1.3, and JNDI

Cache Repositories

- File
- Greenplum 4.1
- HSQLDB 2.2.9
- IBM DB2 LUW 9.5 and 10.5
- Microsoft SQL Server 2008, 2012, 2014
- MySQL 5.1 and 5.5
- Netezza 6.0 and 7.0
- Oracle 11g, 11g R2, and 12c
- PostgreSQL 9.1 (default)
- SAP HANA SPS 09
- Sybase ASE 12.5 and 15.5
- Sybase IQ 15.2
- Teradata 13, 13.10, and 14
- Vertica 5.0 and 6.1

Security

- Base64
- Kerberos
- NTLM
- SSL
- WS-Security

Platforms

Clients for CIS Studio

- Microsoft Windows 7 and 8

Server

- CentOS 5.10, 6.5
- Cisco Unified Computing System™ (Cisco UCS®)
- IBM AIX 6.1.0, and 7.1.0
- Oracle Linux 5.10 and 6.5 Red Hat compatibility mode
- Red Hat Enterprise Linux AS 5.10, and 6.5, 7
- SUSE Enterprise Linux 11.3, 12
- Microsoft Windows 2003 Datacenter and Web Editions, 2008 Enterprise R2, and 2012 Standard and R2
- Solaris 5.10, 10 (SPARC)

JVM

- 64-bit

Table 18. Additional Adapters

Collaboration • Email • Google Apps • Google Sheets • Microsoft Active Directory • Microsoft SharePoint (On-premise and online) • Microsoft SharePoint Excel Services **NoSQL and Cloud DBs** • Amazon DynamoDB ○ Amazon RedShift • Cassandra • Couchbase • Google BigQuery • HBase • MongoDB	**CRM and ERP** • Microsoft Dynamics CRM (On-premise & Online) • Microsoft Dynamics GP • Microsoft Dynamics NAV • NetSuite CRM • NetSuite ERP ○ Oracle EBS ○ Salesforce.com ○ Siebel **Marketing Automation** • Google Adwords • Google Analytics • HubSpot • Marketo • Oracle Eloqua	**SAP** • SAP Netweaver ○ mySAP ○ SAP BW ○ SAP Business Explorer (BEx) **Social Media** • Facebook • LinkedIn • RSS • Twitter

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA CIS 7.0.4 7/16