

Vertica Overview

A deploy-anywhere SQL database designed for advanced analytics, speed, and elasticity.

With the continued growth of data volumes and citizen data scientists' broader use of analytics, many companies are re-examining their systems to meet modern demands. Analytics is rapidly evolving. New data from IoT sources, social media, Weblogs and data streams, gas and electrical grids, and mobile networks is being collected in massive data sets. This gives organizations a new opportunity to become truly data-driven, provided they can manage the new data growth and discover the patterns and trends that can lead to both business opportunities and repeat business from their customers.

Product Overview

OpenText™ can help you and your team modernize your data warehouse, deploy analytics in a hybrid cloud environment, and democratize data and analytics to provide further access. In addition, OpenText can unify how your company powers its analytics by offering an open, scalable and elastic database with many powerful features.

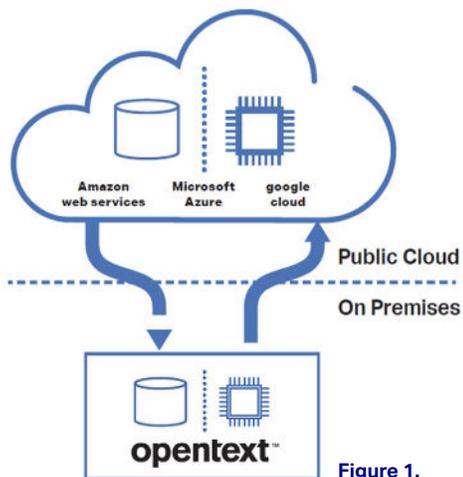


Figure 1.

Product Highlights

Standards-Based Analytics

OpenText offers complete SQL-based analytical functions that make the best use of all the memory and all the resources of the OpenText cluster. SQL is our first and most utilized language for OpenText, making it friendly to analysts and data scientists alike. While OpenText offers some important differences behind the scenes, such as column data storage and massively parallel processing, users won't have to write special queries or consider data structure at all. Instead, you run your SQL, and it runs fast.

Advanced Analytics

With OpenText, you can finally get machine learning into production. OpenText supports cluster-optimized ML algorithms, R, and Python. Data scientists and analysts can build their models using their preferred tools and languages, then leverage OpenText to power them on bigger data sets. In-database machine learning addresses every step in the ML process.

More experienced teams can import models built and trained in other platforms and languages—like TensorFlow, Spark, Python, and SPSS—via the PMML (Predictive Model Markup Language) format. With PMML model export, you can export models created in OpenText for scoring in other systems, such as edge nodes for IoT use cases. Data science and data engineering teams can finally operationalize machine learning while using their tools to leverage their data at scale fully.

Performance and Speed

Exceptional analytical performance requires efficient data loading, analytical query speed, and high concurrency. OpenText is built from the

Key Benefits

Standards-based Analytics:

Speaks SQL, R, and Python so everyone in your organization has access to analytics

Advanced Analytics:

Includes twice the number of analytical functions, including time-series, geospatial and embedded ML

Performance and Speed:

Massively parallel and proven to handle the world's biggest workloads and most number of queries

Efficient:

Through proprietary design, it uses fewer nodes and less storage than any other database

Hybrid Cloud:

SaaS, Public Clouds, Private Clouds, Kubernetes, and bare metal deployments offers analytics when and where you need it

Ecosystem-friendly:

Connects to ETL and BI tools and hundreds of middleware applications

Data Lake/Data Warehouse:

Store data in our optimized repository, or leverage your own Parquet files for analytics

Secure:

Encrypts data in the database and data in-motion, as well as a host of other features

ground up to blast you through all three. It offers superior load throughput, blazingly fast analytical speed, and high concurrency for multi-user querying. You get speedy query results, reducing query times from hours to minutes or minutes to seconds—something outdated row-store technologies cannot achieve. Its modern architecture gives you the flexibility and scalability to ramp up as workloads increase. OpenText can even isolate workloads to serve multiple departments without duplicating data. OpenText's sub-clustering capability allows workload isolation, eliminating resource competition between BI and data science projects. End-to-end machine learning management—from data prep to deployment. No need to download and install separate packages.

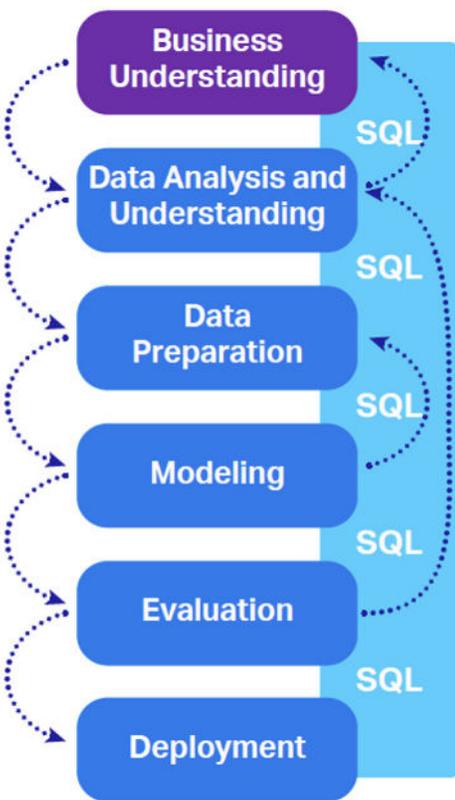


Figure 2. Vertica built-in machine learning process flow.

Efficiency

OpenText offers better compression and impressive optimizations. As a result, you will require less hardware and storage than comparable data analytics solutions. In addition, our optimized architecture means that queries run 10-50X faster than on most platforms, while you can store between 10-30 times more data per server.

Efficiency comes in the form of tools, too. OpenText™ Vertica™ Database Designer improves operations and reduces resource consumption while improving your team's projection designs for consistently faster queries. It analyzes your logical schema, sample data, and your sample queries. Vertica Database Designer creates a physical schema design (a set of projections) that you can deploy automatically or manually. It can be used by anyone, even users without specialized database knowledge.

Hybrid Cloud

Vertica helps you avoid getting locked into a single cloud vendor—you can use the tools of your choice and take full advantage of the underlying infrastructure you already have in place. It offers portability across multi-cloud, on-premises environments, and data lakes. OpenText runs on popular public clouds, including Google Cloud Platform (GCP), Azure, AWS, Alibaba, and VMware clouds. However, it can run in a similar configuration on-premises while leveraging a range of local object stores, such as Apache Hadoop HDFS for communal storage, MinIO, Vast, Dell ECS, Scality, H3C, and Pure Storage. It creates a tough, flexible platform for running a company's analytical and computing workloads, because applications can run simultaneously on multiple environments in a hybrid cloud infrastructure. OpenText can use public clouds and private data centers, and it offers the flexibility to switch at an instant.

Ecosystem-Friendly

Every release of OpenText™ Vertica™ is certified and tested with visualization and ETL tools. It supports popular SQL and Java

Database Connectivity (JDBC)/Open Database Connectivity (ODBC). You can preserve years of investment and training in these technologies because all popular SQL programming tools and languages work seamlessly. Leading BI and visualization tools are tightly integrated, such as Tableau and MicroStrategy, as well as popular ETL tools like Informatica, Talend, Pentaho, and more.

Data Lake/Data Warehouse

OpenText can analyze data wherever it resides—HDFS or Cloud Object Storage—and in all popular formats—ORC, Parquet, JSON, or ROS (native Vertica). You do not need to port data simply for analytics from one location to another. Instead, you save time and realize data insights sooner with the data in place. OpenText also does data analysis on complex data types like Maps and Arrays, Structs in Parquet on S3, HDFS, open SQL-based analytics, and new use cases.

Security

OpenText provides end-to-end security with support for industry-standard protocols and partner solutions such as LDAP, Kerberos, TLS, FIPS 140-2, AWS IAM, and Apache Sentry. OpenText uses a layered security model and provides multiple security authentication authorization mechanisms. Authentication and access is controlled by passwords stored with SHA, LDAP, Kerberos, and SSL certificates. It allows fine grained access control to row and column data, including column masking. It keeps an audit trail, natively exported to other security domains for analysis and persistence.

Full and Complete Analytics

OpenText offers more built-in analytics features than any other data analytics platform:

- Many powerful SQL commands not found on lesser databases, including time-series and geospatial
- Simple SQL execution of predictive analytics (ML/AI) gives analysts the key to machine learning

- OpenText supports Python and lets you use the OpenText cluster to speed through your Python code
- Prepare data with functions for normalization, outlier detection, sampling, imbalanced data processing, missing value imputation, and more
- Create, train, and test advanced machine learning models on massive data sets
- Evaluate model-level statistics including ROC tables and confusion matrices
- Support for familiar programming languages lets you develop user-defined extensions (UDx) with C++, Java, zPython, or R.
- Built-in machine learning algorithms support classification, clustering, and predictive applications. These include linear regression, logistics regression, k-means, naïve Bayes, support vector machines and random forest.

OpenText's integration with Voltage form at preserving encryption technology ensures end-to-end data protection, securing data in use, in motion, and at rest. Voltage SecureData by OpenText™ adds a data-centric layer to OpenText's layered security model, facilitating compliance with new and emerging data privacy regulations such as CCPA and GDPR. Voltage encryption reduces the risk of breach by securing sensitive data while preserving the data's length and character set. Encryption protects structured data such as tax ID, name, address, GPS location, IP address, date of birth, and salary—all typical personal and protected health information types. The decryption of protected data is controlled by policy, including requiring LDAP authentication and authorization if needed.

Figure 3. OpenText's integration with the Voltage product for additional enhanced security.

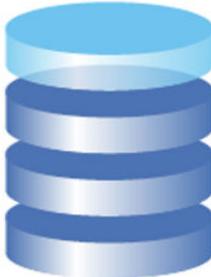
Protection of data everywhere it goes





Voltage SecureData Cloud provides security in the cloud across Hybrid IT systems. NiFi integration enables IoT protection at the edge

Protection that scales with big data





Voltage SecureData delivers protection that scales with the growth of nodes, data volumes and data types

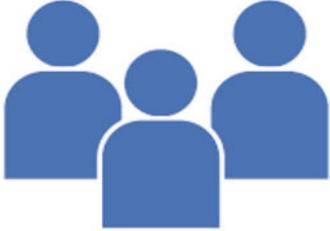
Protection enabling data usability for analytics





Data protected by Voltage Hyper FPE preserves usability for analytic insights and supporting business processes

Protection for sharing data with 3rd parties





Voltage SecureData granular policy controls allow many users to access protected data and only a few to expose sensitive data, if required

Choose OpenText-as-a-Service, or Customer-Managed Vertica

OpenText™ Vertica™ Accelerator is our data analytics SaaS offering, which provides all the functionality of OpenText™ Vertica™'s Unified Analytics Platform, available as software as a service on AWS. It runs in your own Amazon cloud account, and OpenText experts will help you with onboarding. All of your data analysts and data scientists can tap the power of OpenText advanced analytics and machine learning without worrying about scaling, upgrades, monitoring, or support.

The Vertica Unified Analytics Platform is our customer-managed offering—the revolutionary column-store, relational database designed from the ground up for analyzing massive data sets, whether on-premises, on Hadoop, in the clouds, in any combination of those. OpenText's multi-parallel processing (MPP) and shared-nothing architecture remains unrivaled for speed and concurrency, and offers in-database machine learning capability for teams seeking advanced AI solutions without compromising workloads for critical business intelligence SLAs.

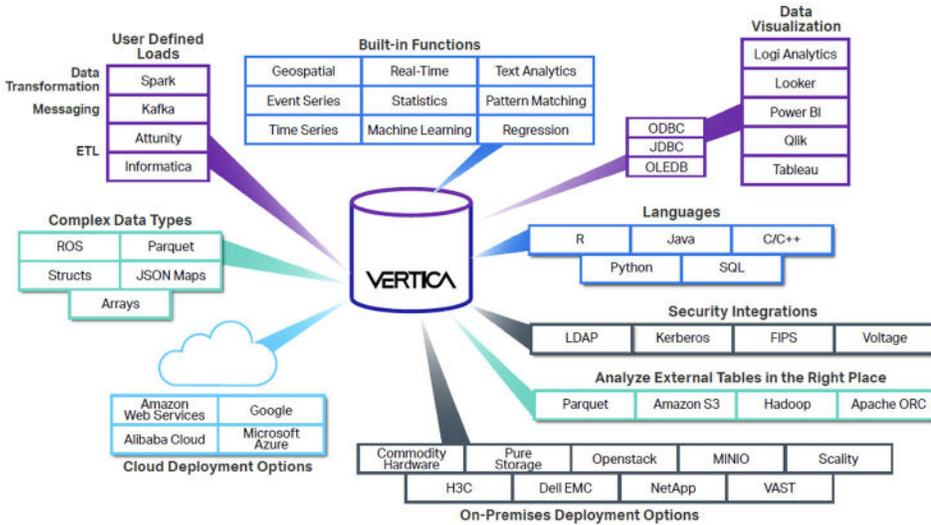


Figure 4. Vertica's open architecture and rich ecosystem.

Try Vertica Accelerator OR the Vertica Community Edition, for Free

Your data analytics needs are unique, and always changing. Your analytics data warehouse should be flexible enough to meet your demands today, while helping you see into your business's future.

Try Vertica, and make your concept a reality.

Try Vertica Accelerator (and watch the demo) at: www.vertica.com/accelerator

Try the customer-managed Vertica at: www.vertica.com/try

Learn more on our website: www.vertica.com

Find free Vertica training and certification courses on Vertica Academy: academy.vertica.com

Connect with Us

[OpenText CEO Mark Barrenechea's blog](#)

