Vertica Advanced Analytics Platform

The Vertica Advanced Analytics Platform is consciously designed with speed, scalability, simplicity, and openness at its core and architected to handle analytical workloads via a distributed compressed columnar architecture.

Product Overview

Vertica Advanced Analytics Platform provides blazingly fast speed (queries run 10–50X faster), exabyte scale (store 10–30X more data per server), openness, and simplicity (use any business intelligence [BI]/ETL tools, Hadoop, etc.)—at a much lower cost than traditional data warehouse solutions and a better time to market than unproven open source solutions.

Handling Today's Massive Data Volumes

In modern data infrastructures, data comes from everywhere: business systems like CRM and ERP, IoT sensors, tweets and other social media data, Web logs and data streams, gas and electrical grids, and mobile networks to name a few. With all this data generated from so many places, companies are turning to disparate, lower-cost storage locations to store and manage these volumes, adding complexity and creating silos as they seek to perform advanced analytics on that data at scale.

To manage this volume of data, newer, more innovative Big Data analytics platforms have emerged to keep up with the sheer size and complexity but some struggle to reach all of the data in a unified fashion and others offer only very limited infrastructure options. Only an advanced analytics platform that can reach all of the data without any underlying infrastructure restrictions will meet the needs of the most data-driven enterprises. From data monetization to customer retention to compliance to

traffic optimization, enterprises that embrace Big Data analytics platforms are changing the dynamics of industries from retail to health care to telecommunications to energy and beyond.

Key Benefits

What Are the Key Technology Requirements of a Big Data Analytics Platform?

So, just what should you look for in a data analytics solution to address today and tomorrow's data challenges? Consider the following:

- **Analyze huge data volumes in a unified manner:** You are likely looking to analyze data at unlimited scale combined with the need to store your data in the right place at the right time. Today, the scale may be terabytes or petabytes. Tomorrow, you may be thinking about exabytes.

- **Meet business expectations:** Users don't want to wait for results. Your solution must provide the scalability to meet service-level agreements (SLAs) and business needs with the best TCO and fast ROI.

- **Embrace popular tools:** If your Big Data analytics relies on extract, transform, load (ETL) tools or SQL-based visualizations, your analytics platform must provide robust and powerful SQL and also be certified to work with all of your tools—not just those from your primary vendor or limited to a single infrastructure platform.

Key Features

At the core of the Vertica Advanced Analytics Platform is a column-oriented, relational database built specifically to handle today's analytic workloads. Unlike commercial and open-source row stores, which were designed long ago to support small data, the Vertica Advanced Analytics Platform provides customers with:

- Complete and advanced SQL-based analytical functions to provide powerful SQL analytics

- A clustered approach to storing Big Data, offering superior query and analytic performance

- Better compression, requiring less hardware and storage than comparable data analytics solutions

- Flexibility and scalability to easily ramp up when workloads increase

- Better load throughput and concurrency with querying

- In-database machine learning algorithms and R, Python extensibility

- Less intervention with a database administrator (DBA) for overhead and tuning

The Technology That Makes Vertica So Powerful

Vertica is built from the ground up to handle the challenges of Big Data analytics. With its massively parallel processing system, it can handle exabyte scale, and has done so in some of the most demanding use cases in the industry. Because it's a columnar store and offers aggressive compression of data, it delivers very fast Big Data analytics, taking query times from hours to minutes or minutes to seconds as opposed to outdated row-store technologies built for an earlier era. Finally, Vertica provides very advanced SQL-based analytics from graph analysis to triangle counting to Monte Carlo simulations to geospatial and more. It is a full-featured analytics database system. And all of this can be applied to your "hot" data loaded directly into Vertica for the most demanding use cases as well as to data in external locations like HDFS and Amazon S3, ensuring a unified analytics view without the need for data movement.

Vertica's in-database machine learning capabilities allow users to take advantage of Big Data while simplifying and speeding up their predictive analytics processes to make better-informed decisions, compete more effectively, and accelerate time-to-insight. From data prep to deployment, Vertica supports the entire machine learning process, and allows models to be deployed across Vertica clusters, a key requirement for solutions that embed Vertica as their analytics engine.

Every release of Vertica is certified and tested with visualization and ETL tools. It supports popular SQL, and Java Database Connectivity (JDBC)/Open Database Connectivity (ODBC). This enables users to 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, MicroStrategy, and others and so are all popular ETL tools like Informatica, Talend, Pentaho, and more.

Vertica offers maximum scalability for large-scale Big Data analytics. It is uniquely designed using a memory-and-disk balanced distributed compressed columnar paradigm, which makes it exponentially faster than older techniques for modern data analytics workloads.

Securing Critical Vertica Data

Vertica 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. Vertica uses a layered security model and provides multiple security, authentication authorization mechanisms. Vertica uses TLS to establish a secure connection between the client machine and the server. Authentication and access can be controlled by passwords stored with SHA, LDAP, Kerberos, and SSL certificates. Further, Vertica's security model is built on ANSI standard Role based access control, where privileges are assigned to roles to which users can be assigned. Vertica also has an access policy mechanism that allows fine-grained access control to row and column data, including column masking. Finally, Vertica's logging mechanisms ensure there is an audit trail of operations performed and this trail can be natively exported to another security domain for analysis and persistence.

Vertica also includes high-performance integration with Voltage SecureData, an industry leading security solution to provide end-to-end protection, securing data at rest, in-motion and in use. With Voltage format-preserving encryption and tokenization technologies, protection is enabled at the data field and subfield level. This preserves characteristics of the original data, including numbers, symbols, letters, and numeric relationships such as date and salary ranges. It also maintains referential integrity across distributed data sets so joined data tables continue to operate properly. With Voltage SecureData, the majority of analytics can be performed on the de-identified data in its protected form. Data scientists need not have access to live payment card, personal, or protected health information in order to deliver business insights.

The Broadest Deployment and Consumption Models

Available on-premise, on Hadoop, and in the clouds, Vertica is a unified, single analytical engine that offers proven Big Data analytics at unmatched speed and scale, forever free from underlying infrastructure.

- Vertica can be deployed on premise in your own data center, in a private cloud or in public clouds including Amazon, Azure, Google, or VMware clouds. It's simple— just take your Vertica Enterprise license and deploy the same unified engine on the clouds. If you need extra capacity and have no time to stand up onpremise hardware, this is an attractive option.

- The same unified Vertica engine can be deployed natively on Apache Hadoop. Vertica SQL on Apache Hadoop accelerates data exploration and SQL analytics while running natively on an organization's preferred Hadoop distribution.

In the Clouds: Vertica software is optimized and preconfigured to run on Amazon, Azure, Google, and VMware clouds. Vertica provides users, the agility and extensibility to quickly deploy, self-provision, and integrate with a wide variety of BI and ETL software tools. With the flexibility to start small and grow as your business grows, Vertica enables you to transition your data warehouse to the cloud, to on premises, and back seamlessly. With this level of agility, there's no need to compromise.

On premise: Vertica Advanced Analytics Platform is a "shared-nothing," distributed database designed to work on clusters of cost-effective, off-the-shelf servers, and its performance is scaled simply by adding new servers to the cluster. The grid architecture of Vertica reduces hardware and scaling costs substantially (by 70 to 90 percent) when compared to traditional
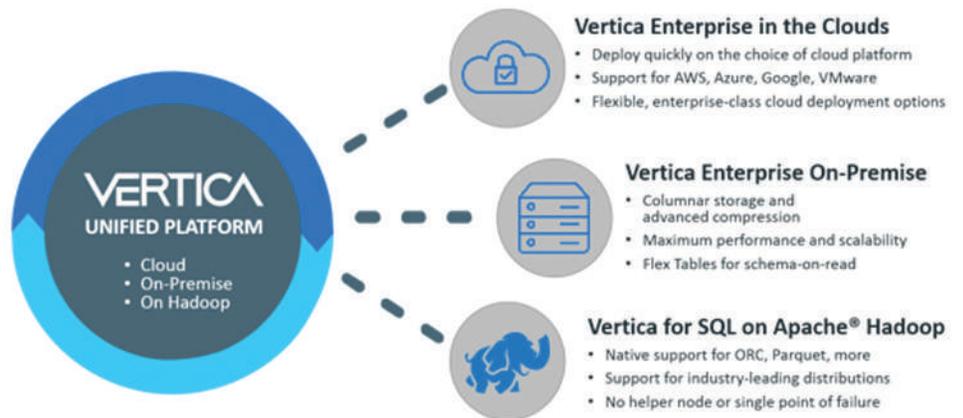
databases that require "big iron" servers with many CPUs and SANs. Clustering also speeds up performance by parallelizing querying and loading across the nodes in the cluster for higher throughput.

On Hadoop: When used together with Hadoop, Vertica for SQL on Apache Hadoop installs directly in your Hadoop cluster and empowers your organization to use a powerful set of data analytics capabilities and do far more than either platform could do on its own. It offers no single point of failure because it's not reliant on a helper node to query. It even reads native Hadoop file formats like ORC, Parquet, Avro, and others, and writes to Parquet. By installing the Vertica SQL engine in the Hadoop cluster, you can tap into advanced and comprehensive SQL on Hadoop capabilities, complete 100 percent of the TPC-DS queries without modification, and run on any Hadoop distribution.

Monetizing and Making Big Data Matter to Us All

The Vertica Advanced Analytics Platform enables organizations to skip the hype and extract value from Big Data. Here are some examples of organizations that have capitalized on their most strategic asset—their data—with Vertica:

- **Intuit**—Processes billions of transactions to deliver highly personalized and rapid returns for millions of TurboTax tax preparation users.

Vertica Enterprise in the Clouds
- Deploy quickly on the choice of cloud platform
- Support for AWS, Azure, Google, VMware
- Flexible, enterprise-class cloud deployment options

Vertica Enterprise On-Premise
- Columnar storage and advanced compression
- Maximum performance and scalability
- Flex Tables for schema-on-read

Vertica for SQL on Apache® Hadoop
- Native support for ORC, Parquet, more
- Support for industry-leading distributions
- No helper node or single point of failure

VERTICA
UNIFIED PLATFORM
- Cloud
- On-Premise
- On Hadoop

- **Conservation International**—Helps scientists assess the impacts of climate, people, and land use by comparatively analyzing sites and species on 86 million records in near real time.

- **Cerner**—Improves patient care by analyzing clinician's efficiency in the Emergency Room (EMR) and saved 500 lives to date based on sepsis alert modeling.

- **Anritsu**—Replaced their legacy EDW with Vertica as their new embedded analytics database to implement predictive analytics solutions that had been theoretical until now and achieved 351% ROI with a payback of just 4 months.

- **Guess**—Delivers essential daily store reports via mobile devices for accurate sales tracking, improvements in merchandise allocation and distribution, and insightful customer purchasing behavior.

Try it and make your concept a reality. The Vertica next-generation high-performance SQL analytics engine is available as three integrated offerings to meet your varying needs—on premise, in the cloud, or on Hadoop. Your needs are unique, so your analytics database should be too. Evaluate Vertica today at: **www.vertica.com/try/**

Learn More At
www.vertica.com/

VERTICA | **MICRO FOCUS®**