GIGAOM

Vertica

Data Warehouse

WILLIAM MCKNIGHT

GIGAOM

Vertica
Data Warehouse

1. Summary

Vertica in Eon Mode is a fully American National Standards Institute (ANSI) SQL-compliant relational analytical database management system that separates compute from storage. Version 9.3.1 is the latest release, including support for Pure Storage to deliver dynamic workload capabilities to on-premises data centers as well as key features like improvements to subcluster management, increased control and tuning of the Depot, and enterprise workload management enhancements to isolate workloads for resourcing and provisioning nodes for ephemeral workloads.

2. Technical Considerations

Vertica is built on Massively Parallel Processing (MPP) and columnar-based architecture. It has its own flavor of SQL called Vertica SQL or VSQL, which has many built-in predictive analytics, machine learning, and advanced analytics functions like geospatial, time series, pattern matching and text analytics functions. Vertica supports complex data types like arrays, maps, and STRUCT, all of which can be nested as well. It can read ORC and Parquet complex types into Vertica and export Vertica data to Parquet.

Queries access optimized local cache in the compute nodes, which is called the Depot, where you can, if you wish, indicate which data you want in. The Depot is actually a form of data storage that is tightly coupled within the compute node for hot data. You can pin tables and partitions in the Depot until an eviction needs to occur.

Largely using the Depot well, Vertica employs a host of performance tricks, leaving the customer to only just to write the query. Though not especially difficult to administer, a Vertica customer does have to administer the "projections" which are database objects that are optimized collections of table columns and live aggregated data.

As might be expected for a database that leads with its column orientation, Vertica's compression rates are very high. Vertica can fully encrypt data, which may become a useful function in the era of the General Data Protection Regulation (GDPR). Article 34 states that fully encrypted data doesn't need to be reported. Integration with Micro Focus' Voltage SecureData product extends this function to Format Preserving Encryption (FPE), allowing analytics without any risk to sensitive data.

3. Deployment Considerations

Vertica in Eon Mode is available on Amazon Web Services (AWS) and also works on-premises – on Pure Storage FlashBlade – which also separates compute from storage, an industry first for on-premises separation. Vertica also offers Enterprise Mode,which has tightly coupled compute and storage resources, for on-premises and cloud deployments that have stable workload requirements. There is also a free community edition, with no time limit and storage of 1 TB of data across three nodes.

Vertica in Eon Mode uses a dedicated Amazon S3 bucket for storage, with a varying number of compute nodes spun up as necessary to meet the demands of the workloads. Vertica keeps cluster status during shut down/startup activity and can restore or replace failed nodes quickly. There is also a status for extended sleep known as hibernation.

All deployments are based on a single database software engine in three editions – Vertica SQL-on-Hadoop (VSQLoH), Express, and Premium: functions are throttled for VSQLoH and Express. There are three licensing options:

- Usage-based, hourly pricing directly from AWS Marketplace

- A perpetual or subscription software license to deploy on-premises and in any major public cloud

- Vertica perpetual or subscription licenses can be purchased on a per-node or per-terabyte basis

If you want to analyze data stored in Hadoop or Amazon S3 data lakes, there is a Vertica External Data add-on charge. Development, test environments, backups, and replication are free, which is unique.
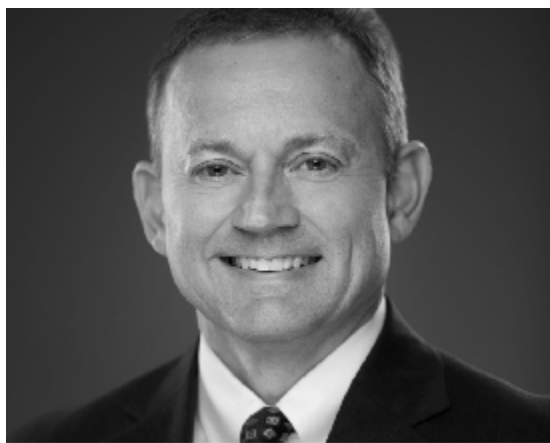
4. Conclusion

Vertica is a longstanding steady database management system (DBMS) player that keeps innovating. It is an ANSI SQL-compatible relational database with market-leading exploitation of the columnar orientation, separation of compute and storage, and hybrid on-premises/multi- cloud development.

In consequence, it offers a realization of great performance for high-volume workloads. In particular, it has found resonance as a relational alternative for storing the massive amounts of information generated by Internet of Things (IoT) applications.

Vertica is so confident of its performance of Vertica in Eon Mode that it has published multiple benchmark reports, comparing its analytical query performance and concurrency to rivals with greater market share.

5. About William McKnight

An Ernst & Young Entrepreneur of the Year Finalist and frequent best practices judge, William is a former Fortune 50 technology executive and database engineer. He provides Enterprise clients with action plans, architectures, strategies, and technology tool selection to manage information.

William McKnight is an Analyst for GigaOm Research who takes corporate information and turns it into a bottom-line producing asset. He's worked with companies like Dong Energy, France Telecom, Pfizer, Samba Bank, ScotiaBank, Teva Pharmaceuticals and Verizon — Many of the Global 2000 — and many others. William focuses on delivering business value and solving business problems utilizing proven, streamlined approaches in information management.

He is a frequent international keynote speaker and trainer. William has taught at Santa Clara University, UC-Berkeley and UC-Santa Cruz.

6. About GigaOm

GigaOm provides technical, operational, and business advice for IT's strategic digital enterprise and business initiatives. Enterprise business leaders, CIOs, and technology organizations partner with GigaOm for practical, actionable, strategic, and visionary advice for modernizing and transforming their business. GigaOm's advice empowers enterprises to successfully compete in an increasingly complicated business atmosphere that requires a solid understanding of constantly changing customer demands.

GigaOm works directly with enterprises both inside and outside of the IT organization to apply proven research and methodologies designed to avoid pitfalls and roadblocks while balancing risk and innovation. Research methodologies include but are not limited to adoption and benchmarking surveys, use cases, interviews, ROI/TCO, market landscapes, strategic trends, and technical benchmarks. Our analysts possess 20+ years of experience advising a spectrum of clients from early adopters to mainstream enterprises.

GigaOm's perspective is that of the unbiased enterprise practitioner. Through this perspective, GigaOm connects with engaged and loyal subscribers on a deep and meaningful level.

7. Copyright