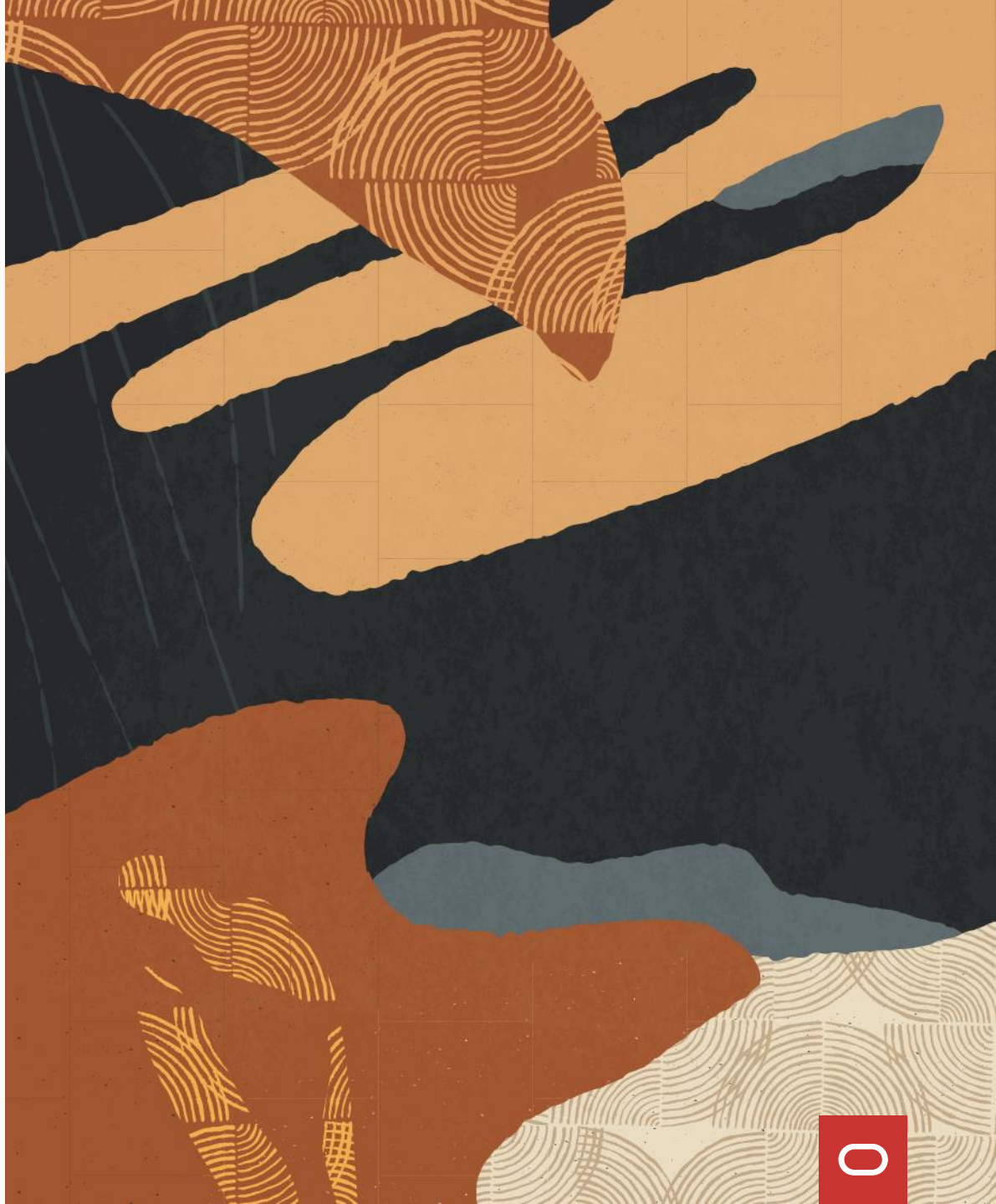
Vertica vs Oracle DB

Competition overview

Presenter Name

Presenter Title



Overview of Vertica

Vertica technical overview

- Analytical, column-oriented database
- founded in 2005 by the database researcher Michael Stonebraker, Andrew Palmer, Ralph Breslauer and Christopher P. Lynch
- Two on-premise editions: **community** (free) and **enterprise (commercial)**.
- Massive Parallel Processing architecture
- Usually considered as extremely easy to install
- Internally based on PostgreSQL, but with its own, significant extensions

Vertica technical overview

Vertica on clouds

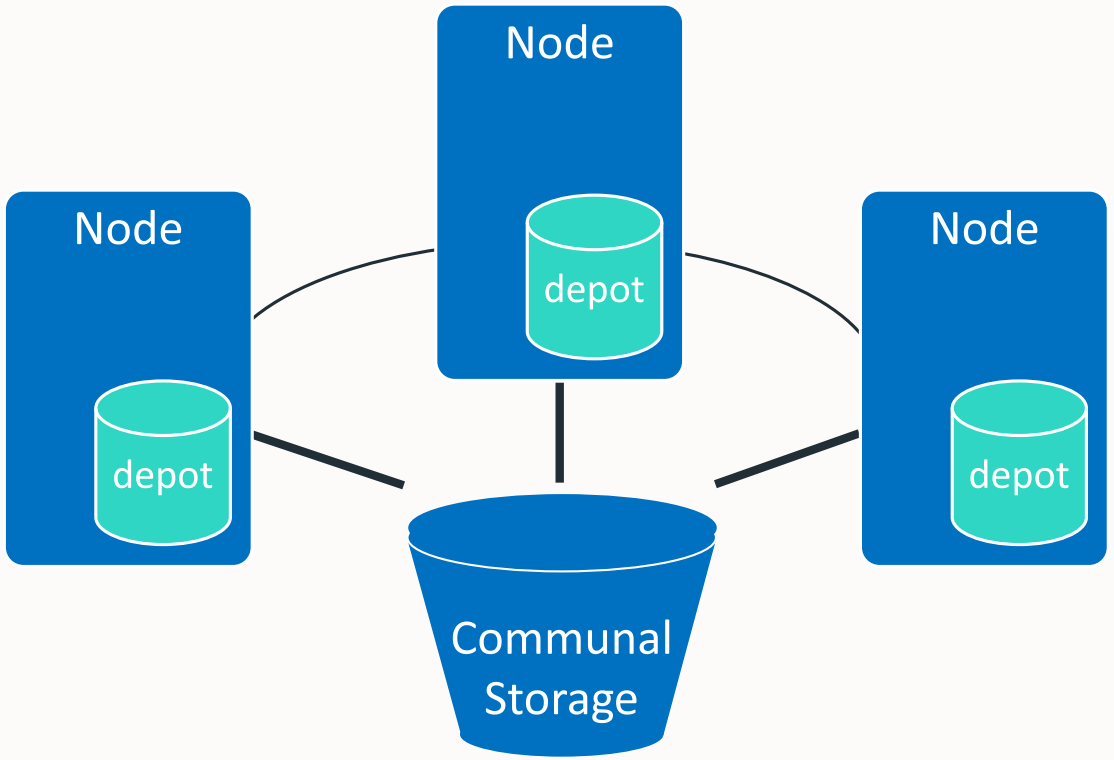
- Available in AWS, Azure and Google Cloud marketplaces
- Architecture based on massive parallelization, exactly like in case of an on-premise installation
- Compute and storage provisioned independently
- The same functionality as in case of on-premise Enterprise Edition

Vertica technical overview

Architecture Overview

EON mode

- Uses shared object storage
- Cloud editions use appropriate solution offered by a cloud provider (example: in AWS – S3)
- Separates the persistent data storage from the compute resources

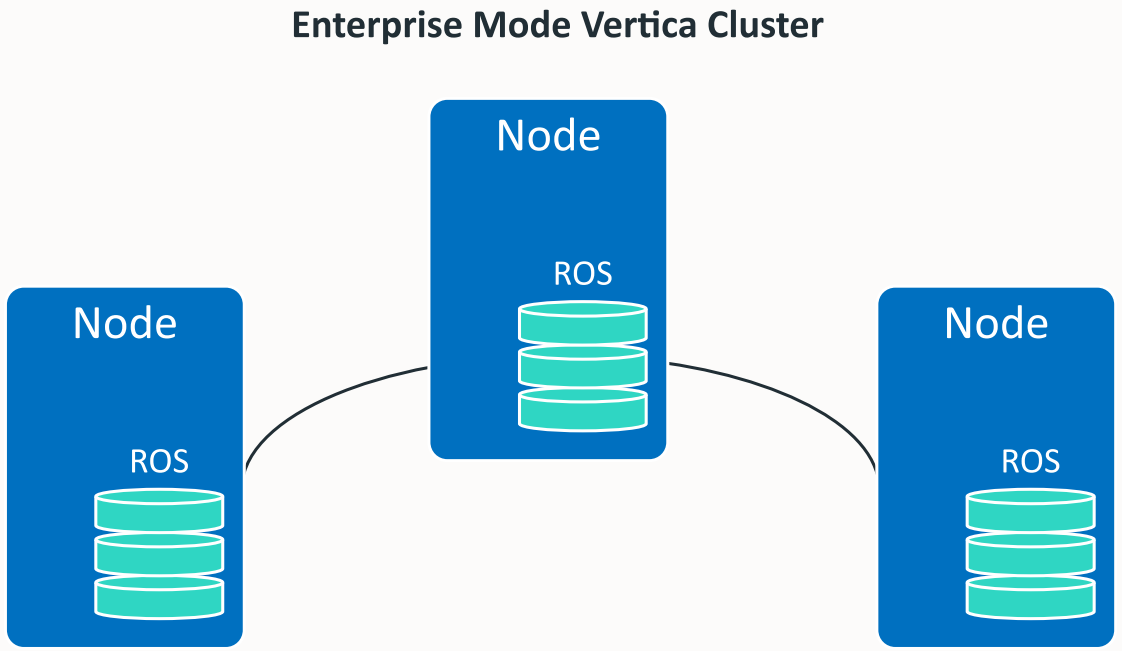


Vertica technical overview

Architecture Overview

Enterprise Mode

- Does NOT use shared storage
- Every node is responsible for storing and processing a portion of the data
- When a node is added or restored after being offline it automatically queries other nodes to update its local data



Vertica technical overview

EON vs Enterprise modes

Chief advantages of...	Where database mode is...	
	Eon	Enterprise
Cloud	•Easily scaled up or down to meet changing workloads and reduce costs. •Workloads can be isolated to a subcluster of nodes. •Virtually no limits on database size. Most cloud providers offer essentially unlimited data storage (for a price).	Works in most cloud platforms. Eon Mode works in specific cloud providers.
On-premises	•Workloads can be isolated to a subset of nodes called a subcluster. •Can increase storage without adding nodes (and, if the object store supports hot plugging, without downtime).	No additional hardware needed beyond the servers that make up the database cluster.

Vertica technical overview

Columnar Database

- Logically uses tables and ***projections*** for storing the data
- A ***projection*** is a set of a table columns
- Projections offer
  - Data compression
  - Distribution of the data across the cluster
- ***Superprojection*** stores all data for a table, while a non-super projection can store only subset of columns.
- ***A query-specific projection*** stores only subset of columns used by a query
- ***An aggregate projection*** stores aggregation of the data
- A segmented projection distributes the data across the cluster nodes using hash algorithm
- A non-segmented projection can be used to replicate the data across the cluster nodes (so called **K-safe projection**)

Vertica CE vs EE

Comparison

	Vertica CE	Vertica EE
Number of nodes	3	No limits (*)
Volume of the data	1TB	No limits (*)
Hadoop integration	N/A	Vertica for SQL on Apache Hadoop
Availability	Docker image or Virtual Machine image	Package for self-installation (rpm, etc)

(*) with exception of techincal limitations, discussed late in this deck

Note:

As part of the Vertica CE license, you agree to the collection of some anonymous, non-identifying usage data. This data lets Vertica understand how customers use the product, and helps guide the development of new features. None of your personal data is collected. For details on what is collected, see the Community Edition End User License Agreement.

Source: Vertica official documentation

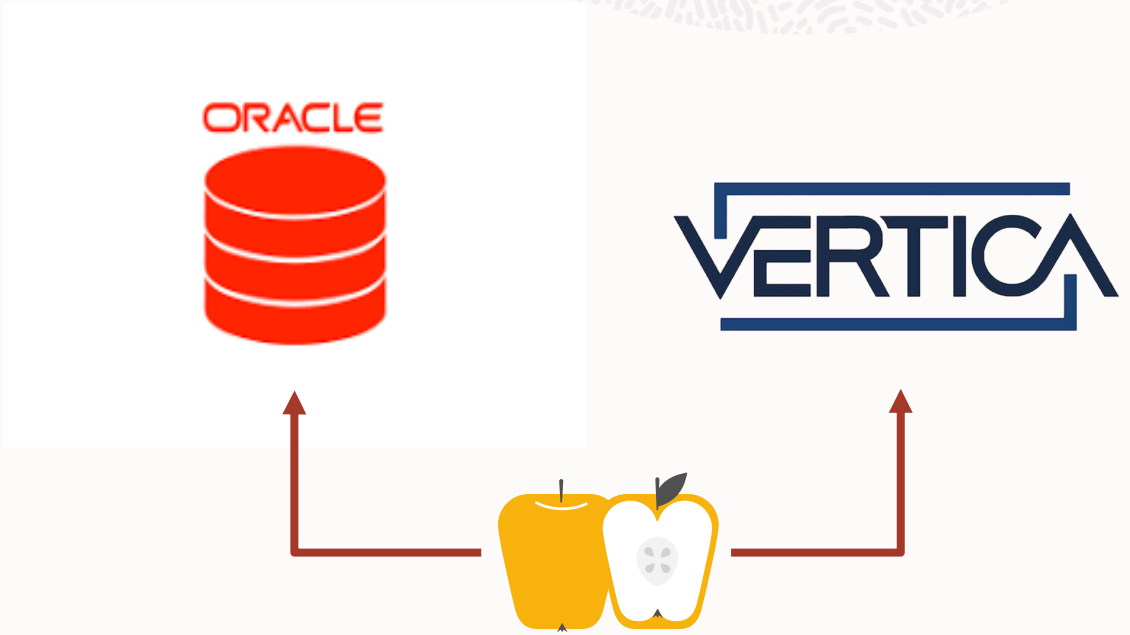
Comparison of Oracle Database and Vertica

Oracle Database vs MongoDB

Let's compare **Oracle Database** with **MongoDB**
But …

Let's compare **apples to apples**

- **OracleDE 23c** vs **Vertica CE**
- **OracleEE** vs **Vertica EE**



Oracle Database Developer Edition 23c vs Vertica Community Edition

Feature	Oracle DE	Vertica CE
Security	✓	✓
Data Encryption	✓ (TDE)	Integration with Voltage Secure Data (third party, commercial solution)
Enterprise User Management	✓ (Centrally Managed Users allows for integration with Active Directory)	✓
Auditing	✓ (standard auditing, FGA, Privilege Analysis, Transparent Sensitive Data Protection)	Auditing limited to access restrictions
Scalability & Performance	✓	✓
Partitioning	✓	Range and hash partitioning, no list partitioning, no composite partitioning
In-Memory	✓ (InMemory Column Store, InMemory Aggregation)	✓
Data Compression	✓ (Advanced Compression, Advanced Index Compression, Prefix Compression)	✓
Clustering	X	Up to 3 MP nodes in CE working in Active-Active mode, extremely easy configuration
Backup and Recovery, HA and DR solutions	✓ (RMAN, trial recovery, fast-start fault recovery, flashback table & database, cross-platform backup&recovery)	✓ similar functionality to rman standard backup and restoration, K-safe projections provide replication of the data across the cluster nodes
Low-Code IDE	✓ (APEX)	x
Management tools	✓ (Enterprise Manager, SQL Developer)	✓ Management Console (OEM equivalent), Administration Tools (Text UI), vsql (CLI)
Other	**No customer data collected**	**As part of the Vertica CE license, you agree to the collection of some anonymous, non-identifying usage data. This data lets Vertica understand how customers use the product, and helps guide the development of new features. None of your personal data is collected. For details on what is collected, see the Community Edition End User License Agreement.**

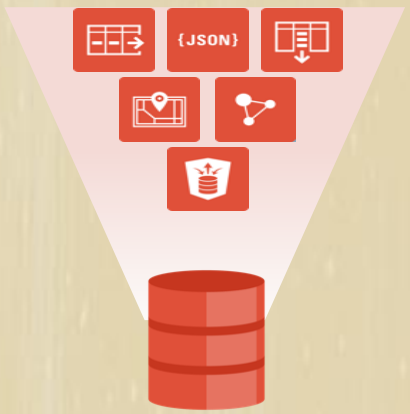
Oracle Database Enterprise Edition vs Vertica Enterprise Edition

Feature	Oracle EE	Vertica EE
Security		
Data Encryption	✓ (TDE)	Integration with Voltage Secure Data (third party, commercial solution)
Enterprise User Management	✓ (Centrally Managed Users allows for integration with Active Directory, LDAP&Kerberos authentication and autorization)	LDAP based authentication
Auditing	✓ (standard auditing, FGA, Privilege Analysis, Transparent Sensitive Data Protection)	Auditing limited to access restrictions
Data Redaction/Data Masking & Subsetting/VPD/Label Security	✓	X
Scalability & Performance		
Clustering	✓ (RAC for share-everything&AA configuration, ADG for share-nothing, DML redirection)	Up to 128 MP nodes in CE working in Active-Active mode, extremely easy configuration
Sharding	✓	x
Partitioning	✓	Range and hash partitioning, no list partitioning, no composite partitioning
In-Memory	✓	✓
Data Compression	✓ includes storage, network and in-memory compression	✓
Parallel operations	✓ Parallel Query, DML and DDL. Adaptive DOP, resource manager allows for controlling DOP for different consumer groups	Parallel queries and load operations supported.
Backup and Recovery, HA and DR solutions	✓ details here	✓ similar functionality to rman standard backup and restoration, K-safe projections provide replication of the data across the cluster nodes
Low-Code IDE	✓ (APEX)	x
Management Software	✓ (Enterprise Manager + packs, SQL Developer)	✓ Management Console (OEM equivalent), Administration Tools (Text UI), vsql (CLI)

Oracle Database vs Vertica limitations

Limitation	Oracle Database (all editions)	Vertica (all editions)
Purpose	Multi-purpose DBMS, can act as an OLTP, DSS or mixed-workload system	Data Warehouse only
Supported data models	relational, object-oriented, multidimensional. spatial, document-oriented, graph, support for API specific to a particular data model, like SODA, PGQL, etc.	Relational only. Unstructured data, like JSON documents can be loaded into flexible-schema tables. No support for non-SQL processing.
Max number of cluster nodes	100	128/3 (EE/CE)
Max table size	unlimited	The smaller of: 2^64 rows per node 2^63 bytes per column
Max row size	unlimited	32,768,000 bytes
Tables/projections per database	unlimited	Limited by physical RAM, as the catalog must fit in memory
Concurrent connections per node	Unlimited	Default of 50, limited by physical RAM (or threads per process), typically 1024
Server-side functions	Functions, Procedures, Triggers, packages, Java stored programs, full support for server-side Java Script	Functions, Procedures, Triggers, np packages (PLvSQL based on PlpgSQL)
Transaction processing	Always ACID	ACID, model similar to OracleDB, no autonomous transactions

Summary



Modern applications need to generate value from data in new ways

- They are built using new development methodologies and technologies
- But it complicates Database Architecture by using multiple single-purpose databases
- Development focused on Integration

Oracle Database allows for easy start developing document-oriented applications

- Provides multiple APIs
- Fully supports document-oriented data model based on JSON specification
- Provides multiple enterprise-scale solutions, which increase its availability, security and performance
- Supports multiple workload types (OLTP, DSS, mixed workload)

Vertica database puts pressure on columnar-based data processing, data caching and massive parallelization, offering limited support for modern data-centric application development

- No support for non-relational data models
- Data-Warehouse–only database
- SQL-based APIs only
- No low-code development environment