VERTICA
by opentext™

Vertica

Architecture Overview

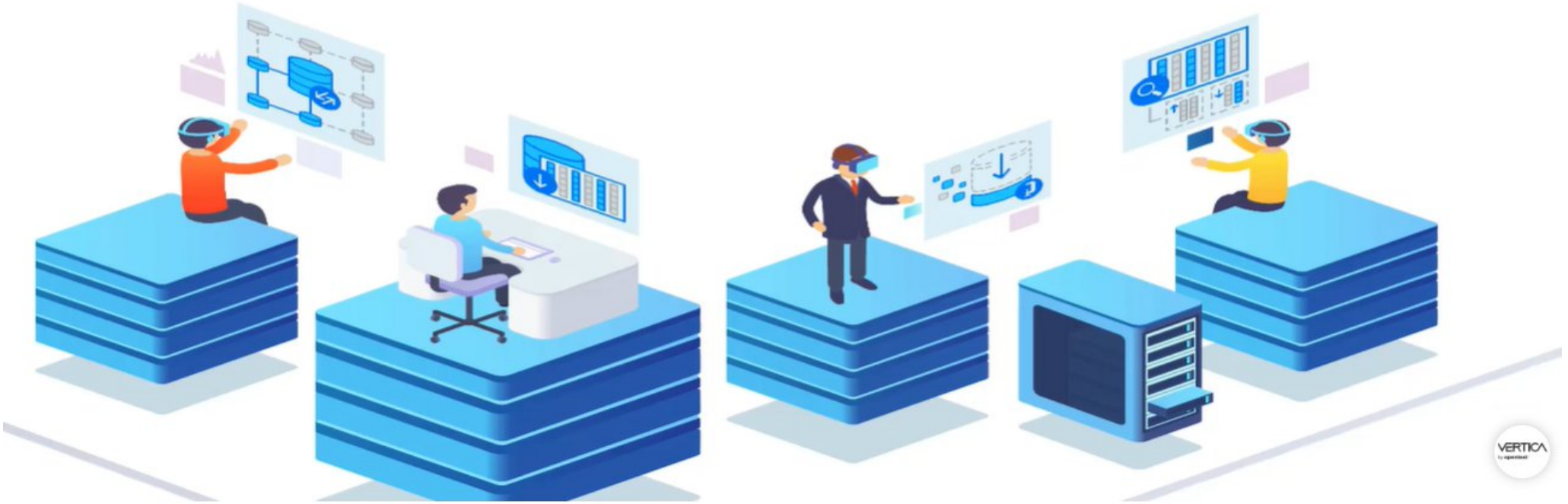
The Analytics Platform

Massively Parallel Processing
Run queries in parallel across infrastructure and scale-out linearly for faster performance or more users

Columnar Storage
Increase query speed over traditional row-based storage systems by reading only the necessary data

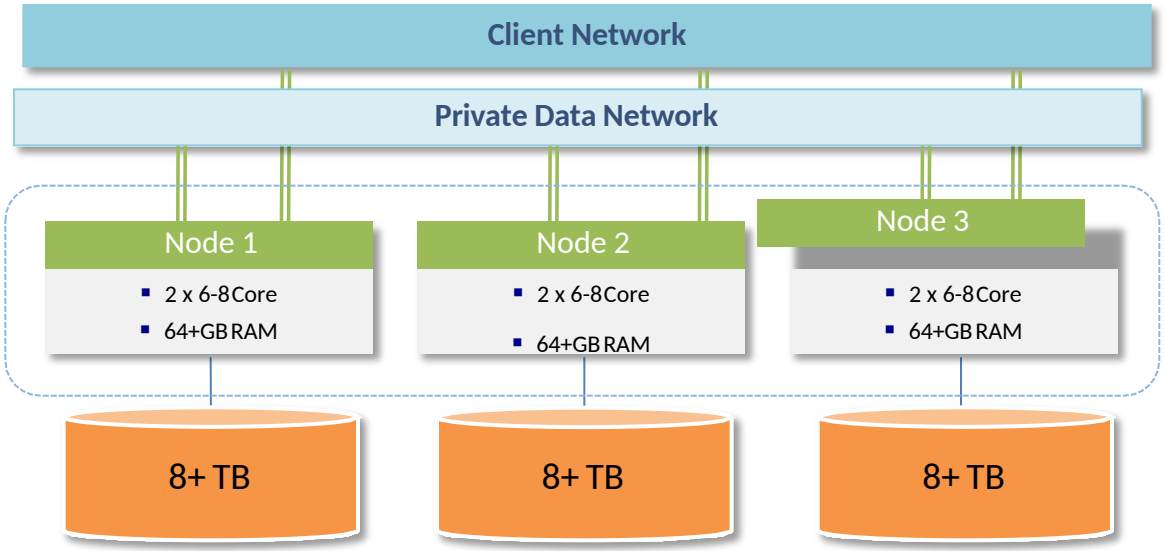
Advanced Compression
Advanced compression algorithms reduce disk space by up to 90% and improve performance by reducing costly I/O

Optimized Projections
Store data in format and structure optimized to improve performance on mostly frequently-run analytics



Massively Parallel Processing (MPP)

- Parallel design leverages data projections to enable distributed storage and workload
  - Active redundancy
  - Automatic replication, failover and recovery
- Shared-nothing, grid-based database architecture provides high scalability on clusters of commodity hardware



Nodes are Peers

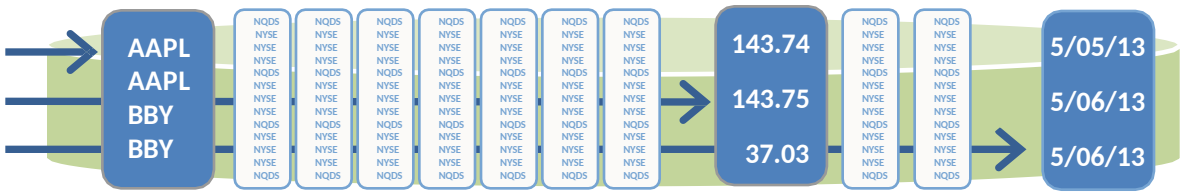
- No specialized nodes
- All nodes are peers
- Query/Load to any node
- Continuous real-time load and query

Column Storage

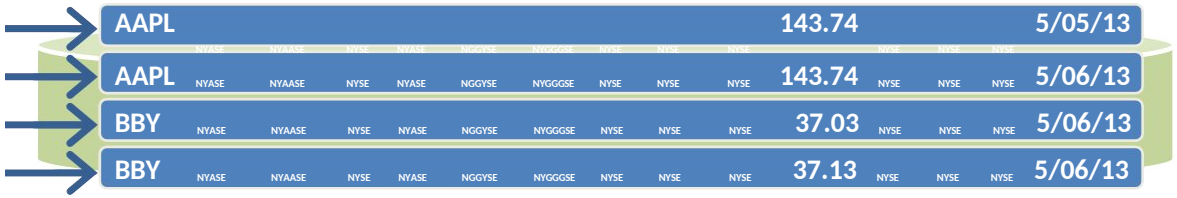
- Vertica organizes data for each column
  - Each column is stored separately on disk
  - Only reads the columns needed to answer the query
- Significant reduction of disk I/O

```
SELECT
avg(price)
FROM tickstore
WHERE symbol
= 'AAPL'
or
date = '5/06/13'
```

Column Store - Reads 3 columns

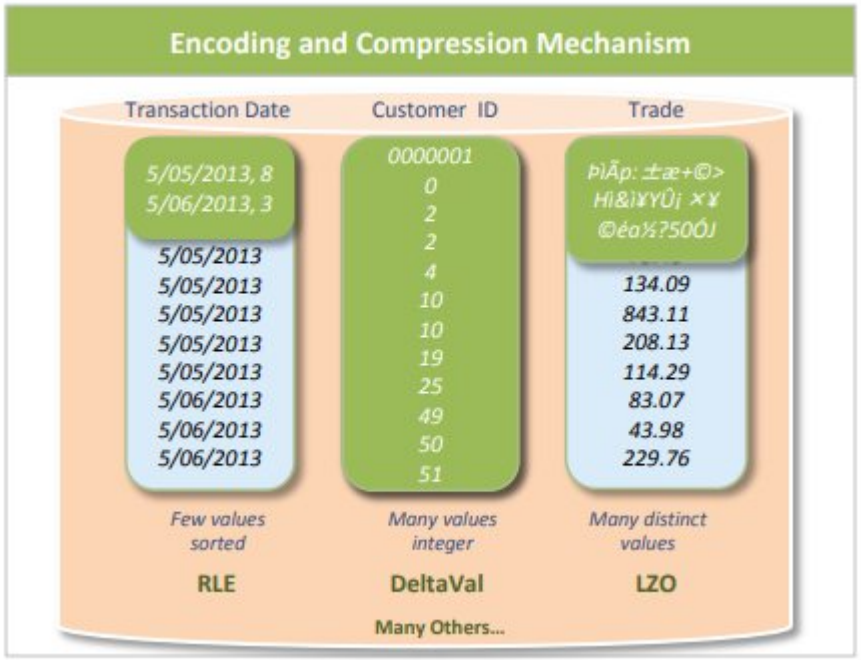


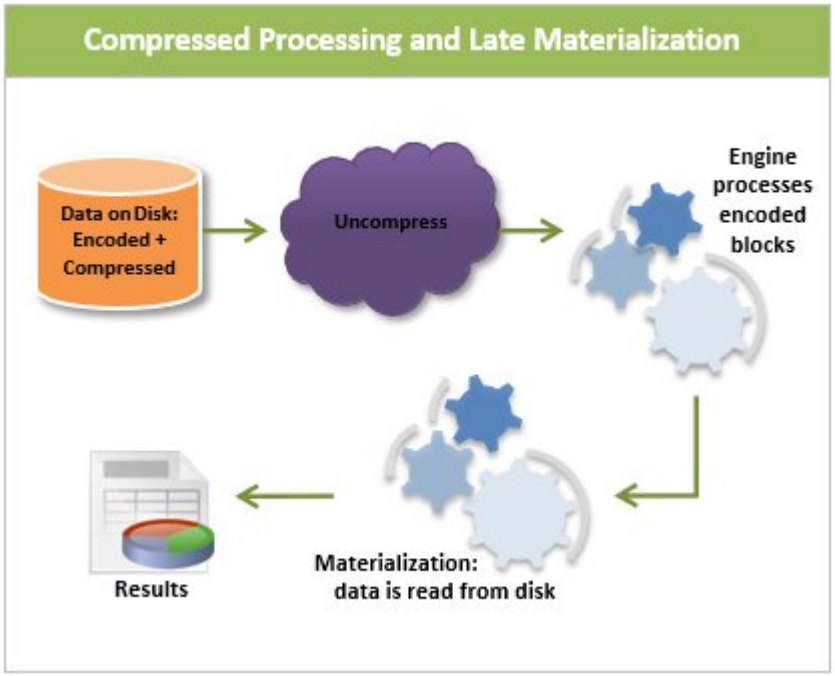
Row Store - Reads all columns



Advanced Compression and Encoding

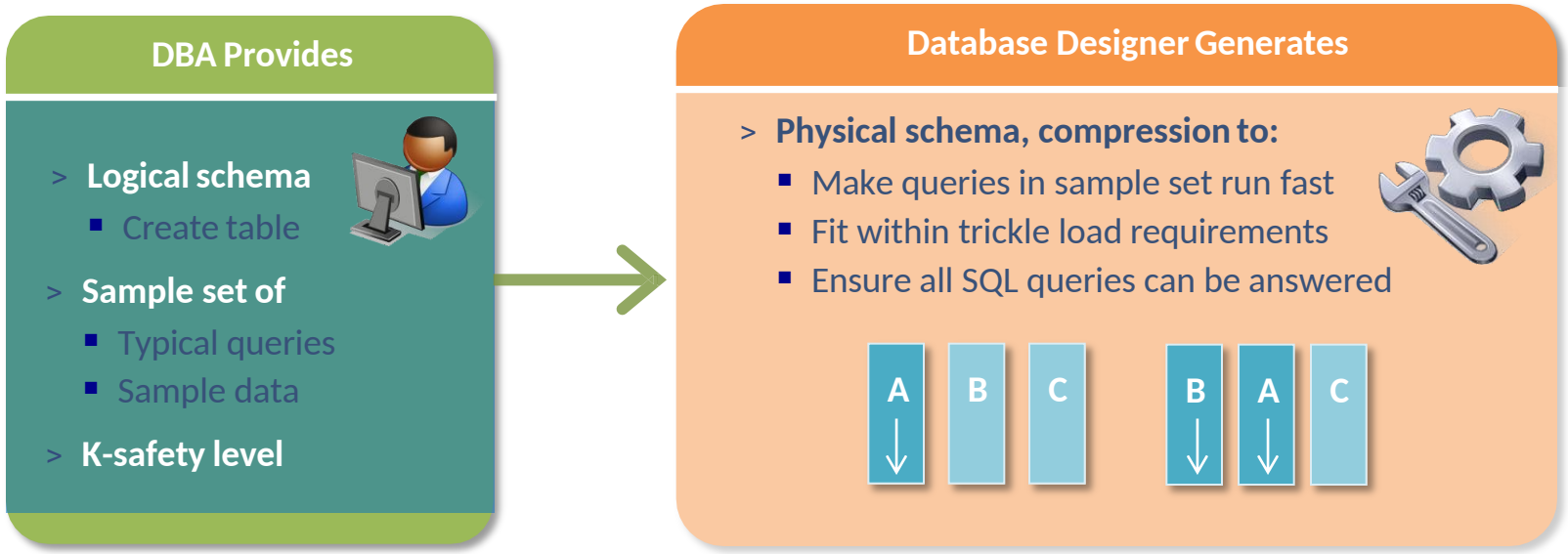
- Slower disk I/O is replaced with fast CPU cycles and aggressive encoding and compression
- Sorting and cardinality help determine encoding





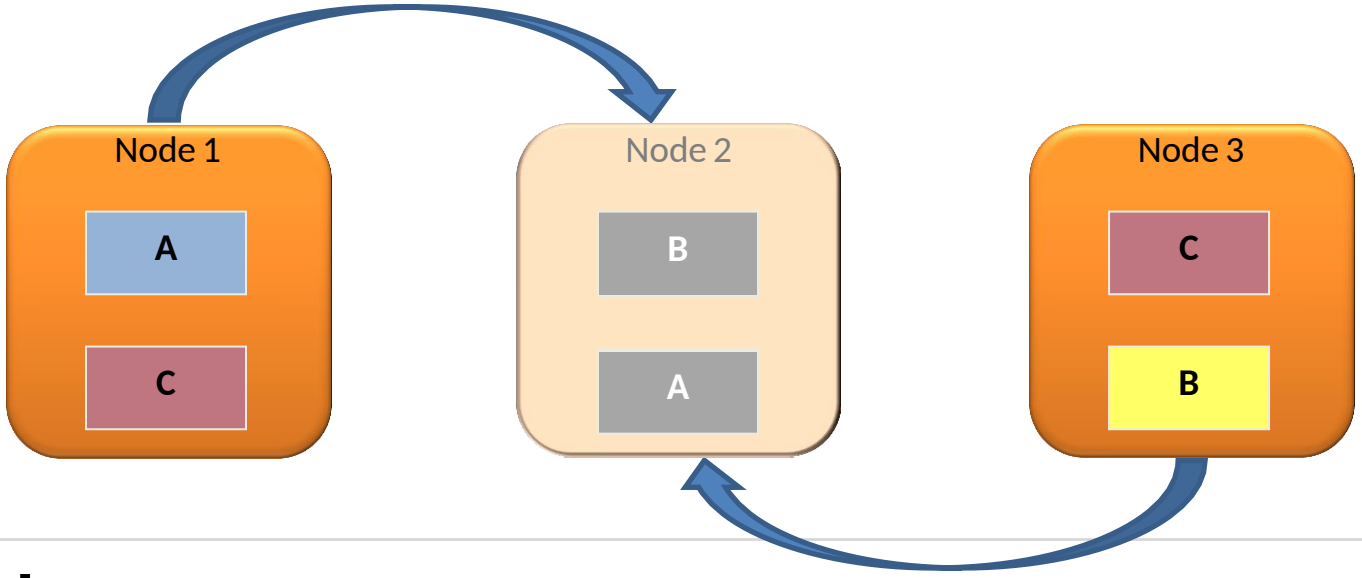
Automatic Database Design

- Database Designer (DBD) recommends a physical database design that provides the best performance for the user's workload
  - Analyzes your logical schema, sample data, and sample queries
  - Minimizes DBA tuning
- Run anytime for additional optimization, without stopping the database



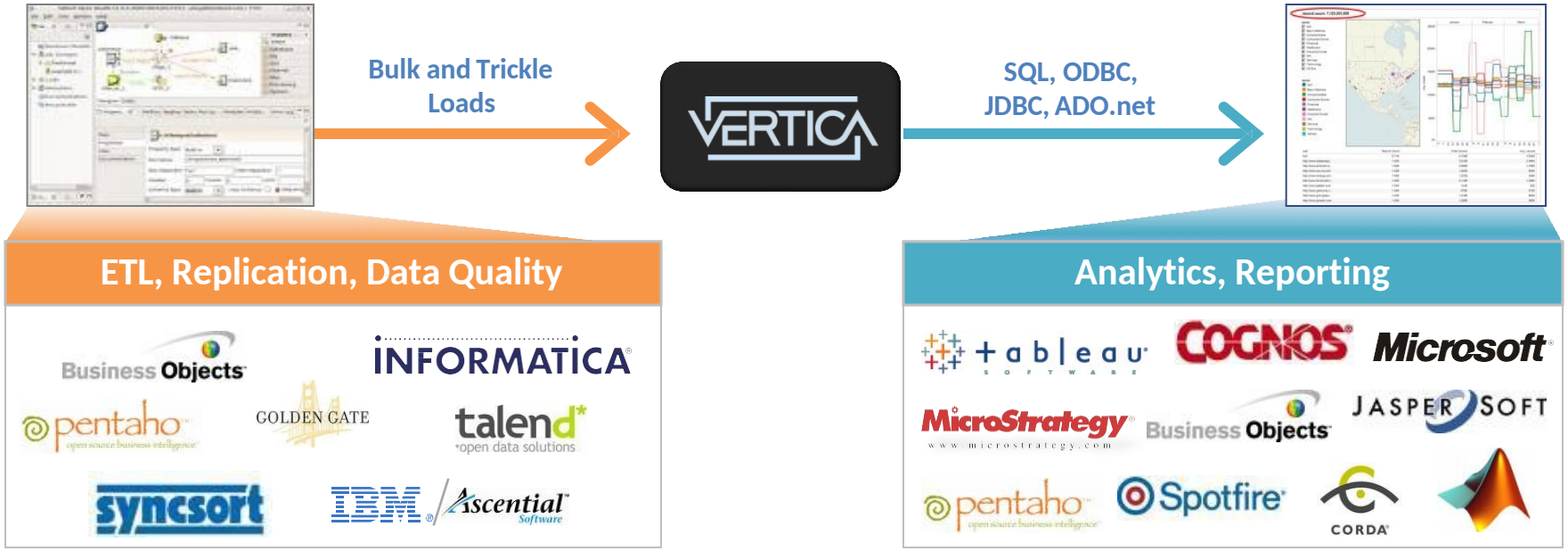
High Availability

- RAID-like functionality within database
  - If a node fails, a copy is available on one of the surviving nodes
  - No need for manual log-based recovery
- Always-on Queries and Loads
  - System continues to load and query when nodes are down
  - Automatically recovers missing data by querying other nodes



Native SQL and Application Integration

- Standard SQL Interface
- Simple integration with Hadoop and existing BI and ETL tools
  - Supports SQL, ODBC, JDBC and majority ETL and BI reporting products
- Leverages existing investments to lower Total Cost of Ownership (TOC)



opentext™

Thank you