

THE PRIVATE DATA CLOUD

by Yellowbrick 

Private Data Clouds provide all the strengths of data clouds we know and love while addressing the shortcomings that make such clouds inappropriate for use cases with security, data residency or cost concerns. At the core, all data clouds have a cloud-optimized storage, a rich SQL data warehouse with separate storage and compute, on-demand elasticity, cross-cloud replication and data sharing.

Yellowbrick invented a newer, fully cloud native data stack on Kubernetes, avoiding the pitfalls of multi-tenant platforms by running in customers' own cloud accounts (BYOC model) and on-premises private clouds. Yellowbrick customers can meet all data residency, localization and sovereignty requirements while enjoying substantial performance improvements and cost savings by running on their own cloud infrastructure.

Optimized Storage

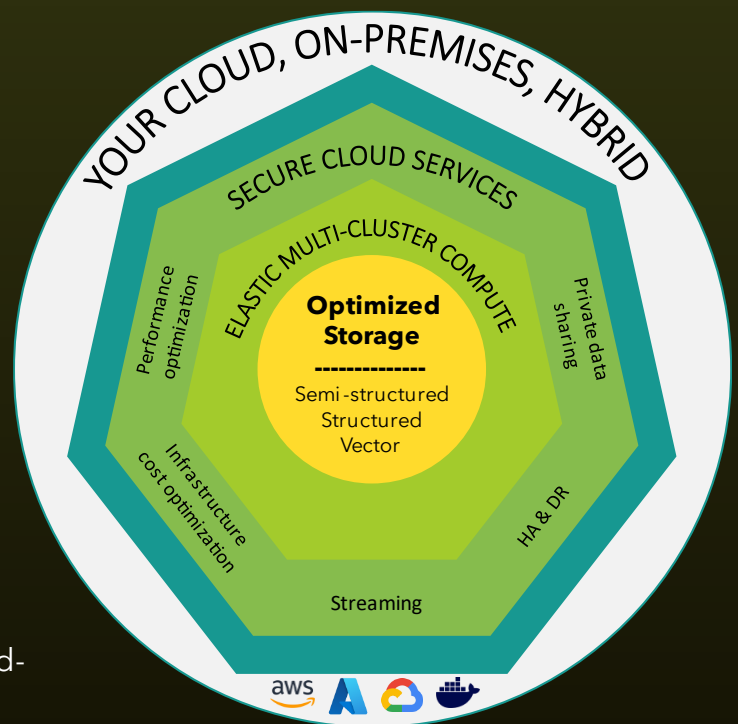
- Cloud object and on-premises storage of structured, semi-structured and vector data
- Hybrid row and compressed column store
- Automatic sorting, clustering, and partitioning
- Tiered scan-resistant caching for fast retrieval
- Support for storing, querying and flattening JSON
- Database sizes from gigabytes through 10PB

Elastic Compute Clusters

- Create, suspend, resume and dynamically resize separate compute clusters using SQL
- Automatic load balancing for scalability
- Partition and isolate workloads across clusters e.g. an ETL cluster, data science cluster, BI or ad-hoc cluster
- Usage metering using SQL

Performance Optimization

- In-memory analytics performance
- LLVM vectorized query execution
- Hundreds of concurrent queries (up to 1000 QPS) for thousands of active users
- Rich workload management
- Database-optimized OS with kernel bypass and DPDK network stack



Private Data Sharing

- Read-only snapshots of databases can be shared to on-prem or cloud using SQL
- Incrementally republish and share updates

Streaming

- Insert rows into the database in real-time without micro-batching
- Rich ecosystem integrations: JDBC/OBJC, Kafka, Spark, Oracle GoldenGate, Fivetran, Striim, Informatica