



Translytical Databases: A Framework for Evaluation and Use Case Analysis

Presented by: William McKnight

“#1 Global Influencer in Big Data” Thinkers360

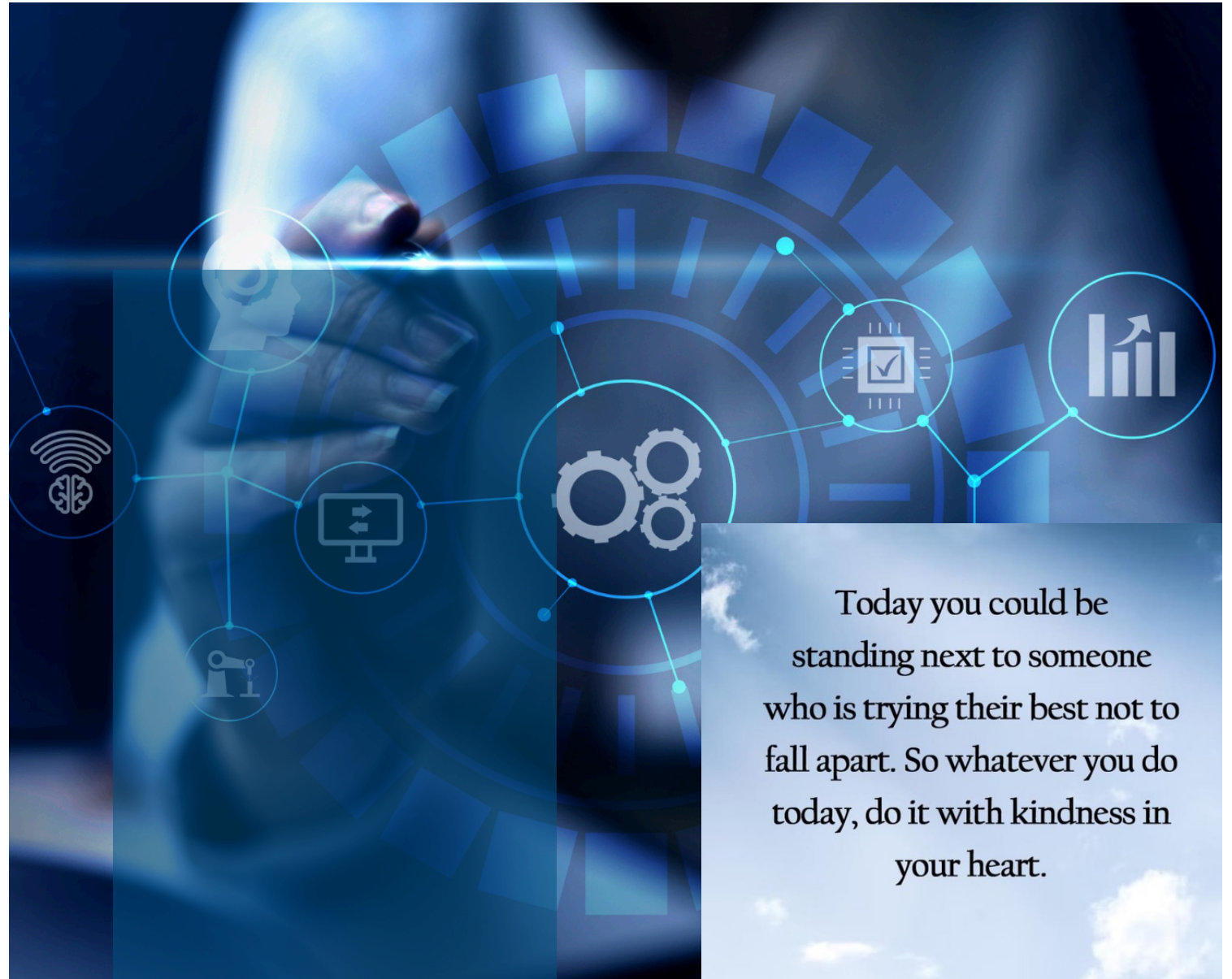
President, McKnight Consulting Group

3 X **Inc 5000**



 /in/wmcknight

www.mcknightcg.com
(214) 514-1444



Today you could be standing next to someone who is trying their best not to fall apart. So whatever you do today, do it with kindness in your heart.

Dataversity Analytics Architecture with William McKnight 2025

1. 2025 Trends in Analytic Architecture
2. Standardizing Data Collaboration: The Role of Open Table Formats in Data Architecture
3. From Models to Data: How Gen AI is Changing the Game for Data Scientists and Data Teams
4. The Rise of Automated MDM: How AI and Machine Learning are Revolutionizing Master Data Management
5. The Data Observability Advantage: Unlocking the Secrets to Reliable, High-Quality Big Data
6. Promising AI Use Cases for the Enterprise in 2025
7. Edge Computing Evolved: Introducing the Zero-DBA, Zero-ETL Embedded Database
8. The ROI of AI is (usually) there – Let's Run the Numbers
9. Translytical Databases: A Framework for Evaluation and Use Case Analysis
10. From Pre-Trained to Fine-Tuned: How to Get the Most Out of Vector, RAG, and Small Language Models
11. Designing a Data Platform for the Future: Principles, Patterns, and Best Practices for Data Fabrics and Data Meshes
12. AIOps: Applying DevOps to Competitive Advantage

McKnight Consulting Group Partial Technology Implementation Expertise

Big/Analytic/Vector/Mixed Data Management



Data Movement and APIs



Data Management



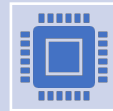
Operational/Transactional Data Management



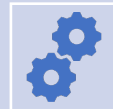
Decisions, Decisions, Decisions



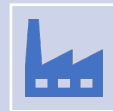
Unprecedented variety of data store choices to meet the needs of the varied workloads in an enterprise today



Enterprises have many needs for databases, including cache, operational, data warehouse, master data, ERP, analytical, graph data, data lake, and time series data



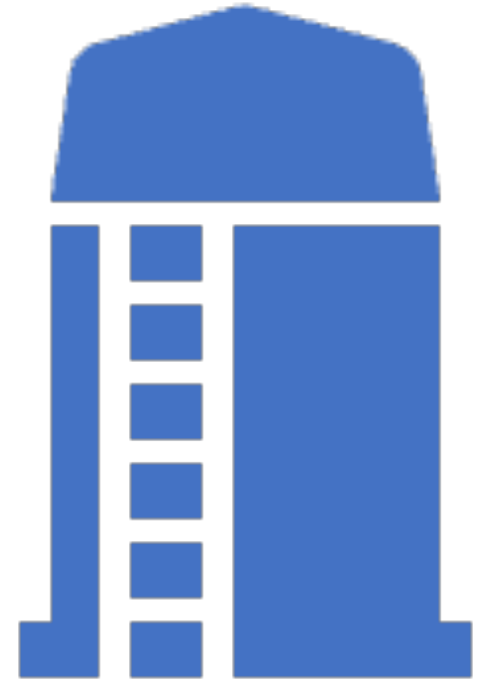
While vendor offerings have exploded in recent years, in due time frameworks will integrate components into what amounts to a single offering for multiple workloads, perhaps even for the enterprise



But what if price-performant offerings for adjacent workloads in an enterprise *have* materialized?

Growing complexity for managing these many data workloads enterprises need

- **Silos and Fragmentation:** Different data types and profiles often reside in separate systems, creating information silos. This makes it difficult to get a holistic view of operations and uncover hidden connections.
- **Limited Skills:** Analyzing unstructured and semi-structured data requires new approaches like machine learning and natural language processing.
- **Scalability Concerns:** The sheer volume of data, especially unstructured, can overwhelm existing infrastructure. Enterprises need scalable solutions to handle the ever-growing data deluge.



OLTP vs OLAP

OLTP

- Process business interactions as they occur
- Support limited query
- Focus on IUD/individual transactions
- Low latency and high throughput needed
- ACID compliance
- Normalized data model

OLAP

- Analytics/complex analysis
- Offload of processing from OLTP
- Dimensional data model
- Lite data modification from source
- Complex queries, frequently long-running
- Large data accumulation

Analytics Defined



- Analytics is the process of utilizing data to enhance business processes.
- Analytics is deeper than simple knowledge; they have depth.
- There's Analytic Projects and...
- There's Analytics Added to projects

Capability Requirements

- Analytics on live data, recent data and historical data
- Real-time analytics calculated from across data domains
- Pre-calculated data
- Live analytics usable operationally
- A seamless platform
- Meet operational SLAs



Benefits of Real-Time Analytics



- Speed to Insight
- Customer Experience
- Operational Excellence
- Deeper Understanding



Translytical Use Cases

Translytical Use Cases

- Portfolio Management
- Wealth Management
- Fraud Analytics
- Risk Management
- Algorithmic Trading
- Crypto Exchange
- SC/IoT Analytics
- Real-Time Customer Experience
- Network Telemetry
- Geolocation Analysis
- Field Support Optimization
- Ad Optimization & Ad Serving
- Streaming Media Quality Analytics
- Real-Time Recommendations
- Video Games
- Telemetry Processing
- IoT & Smart Meter Analytics
- Predictive Maintenance
- Geospatial Tracking

Next Best Offer/Touch

- Need to incorporate not only analytics through last night, but also today, all morning, last hour and last second into screen render
- Need to incorporate not just the user's data but all users' data
 - Need to correlate user to other users instantly
- Only AI can operate at the needed scale



Financial Markets

- Billions of API Requests Daily
- Need 5-10ms Average Query Response
- Data to include:
 - Real-time and historical stock price
 - Cryptocurrencies, Forex, Commodities, Currencies, Premium Data
- Front-Office Traders Need Real-Time Analysis

Healthcare

- Genomic medicine
- Virtual visits
- Tele-health and AI Triage
- AI Diagnostics
- Robotics Automating Lab Work



Retailer

- Better & personalized product recommendations for consumers based on session data, historical order data, and trending products.
- Continuous and automatic retraining the recommendation (ML) engine.
- Near real-time data integration from the retail application to the analytical platform.
- Identify potential compliance issues with customer data, classify and tag sensitive data with labels, and track how sensitive data is being used from the data source to the reports.
- Integration with other systems such as their ERP, email and instant messaging platforms with the analytical solution to get a full 360 view over their business operations and to improve customer satisfaction.

Manufacturing

- Predictive Maintenance
- Quality Control
- Supply Chain Optimization
- Production Planning and Scheduling
- Anomaly Detection
- Yield Optimization
- Inventory Management

Asset Management



- Predictive Maintenance
- Risk Assessment
- Portfolio Optimization
- Asset Performance Monitoring
- Anomaly Detection
- Compliance Monitoring
- Return on Investment (ROI) Analysis

Security Surveillance

- Anomaly Detection
- Facial Recognition
- Object Detection
- Predictive Analytics
- Intrusion Detection
- Access Control
- Incident Response





Translytical Architecture

Data Architecture Needs for Translytical Workloads

- Fast Streaming Ingest (millions of events/second)
- Low Latency
- High Concurrency (thousands of concurrent users)
- Unlimited Storage
- Pipelines
- Transactional Consistency



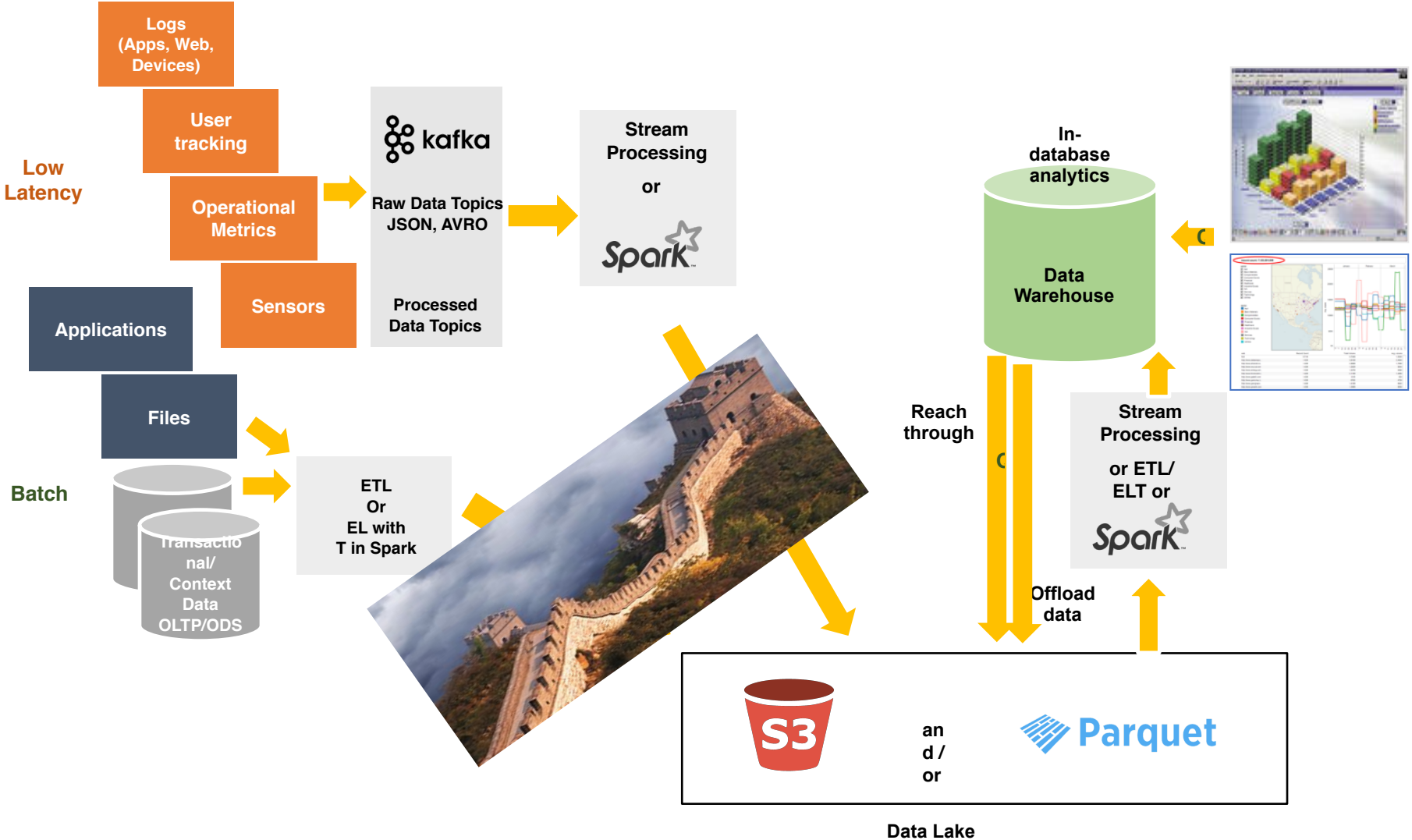
David Shiffman, Ph.D. 🇺🇸
@whysharksmatter.bsky.social

Someone asked me why sharks stick their dorsal fins out of the water. The answer is simple.

Sharks' dorsal fins stick out of the water sometimes because their fins are on top of and therefore higher than the rest of their body, and sometimes their body is near the surface of the water.



Data Architecture Ill-fit for Translytical



Event-Driven Architectures



- Kafka (like) Connector
- Realtime Pub/Sub Messaging Platform
- Edge Computing
- Selective Feed to a Data Lake

Single Product Architectures

- Single Table Storage for Transactions and Analytics
 - Fast IUD and Query
 - Simplified Data Architecture
 - Reduced Data Movement
- Rowstore + Columnstore



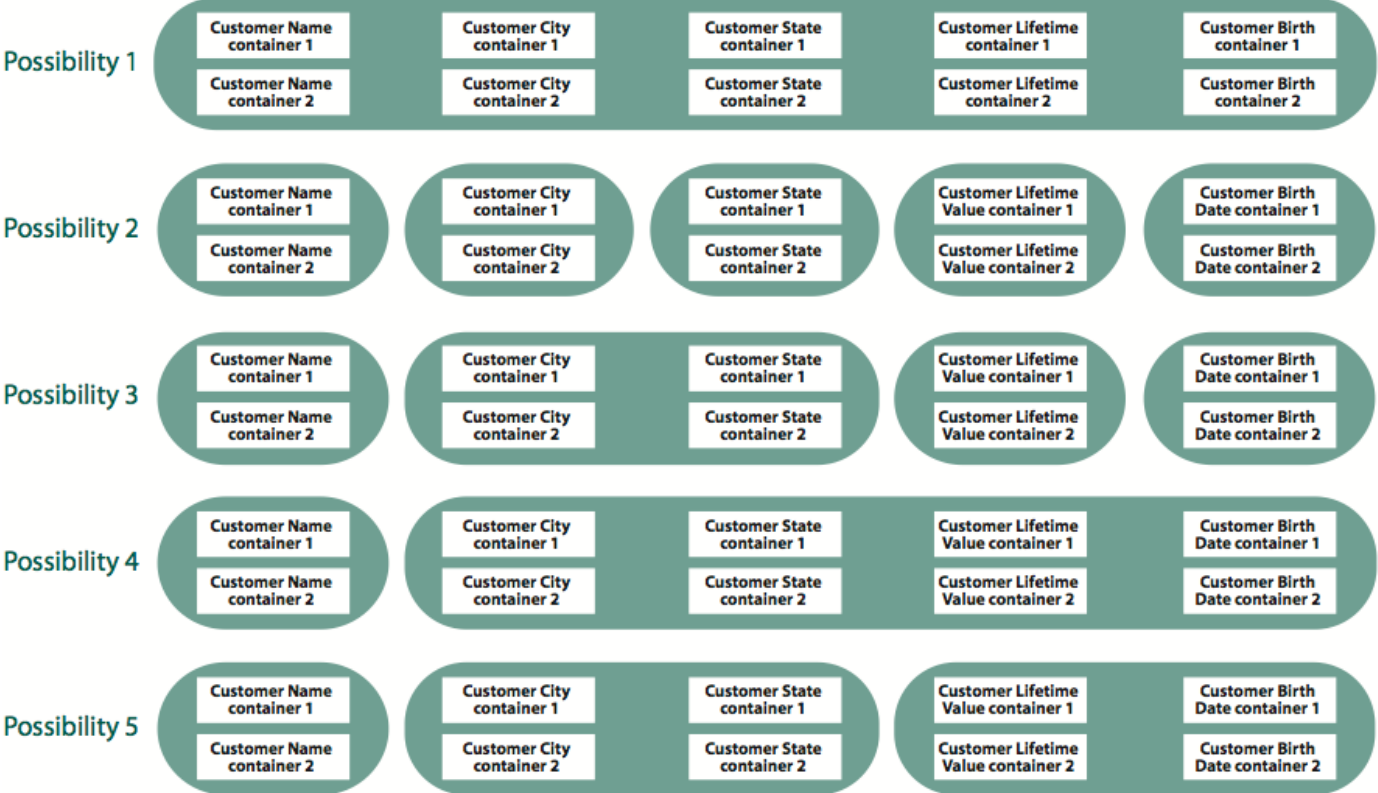
Column Stores



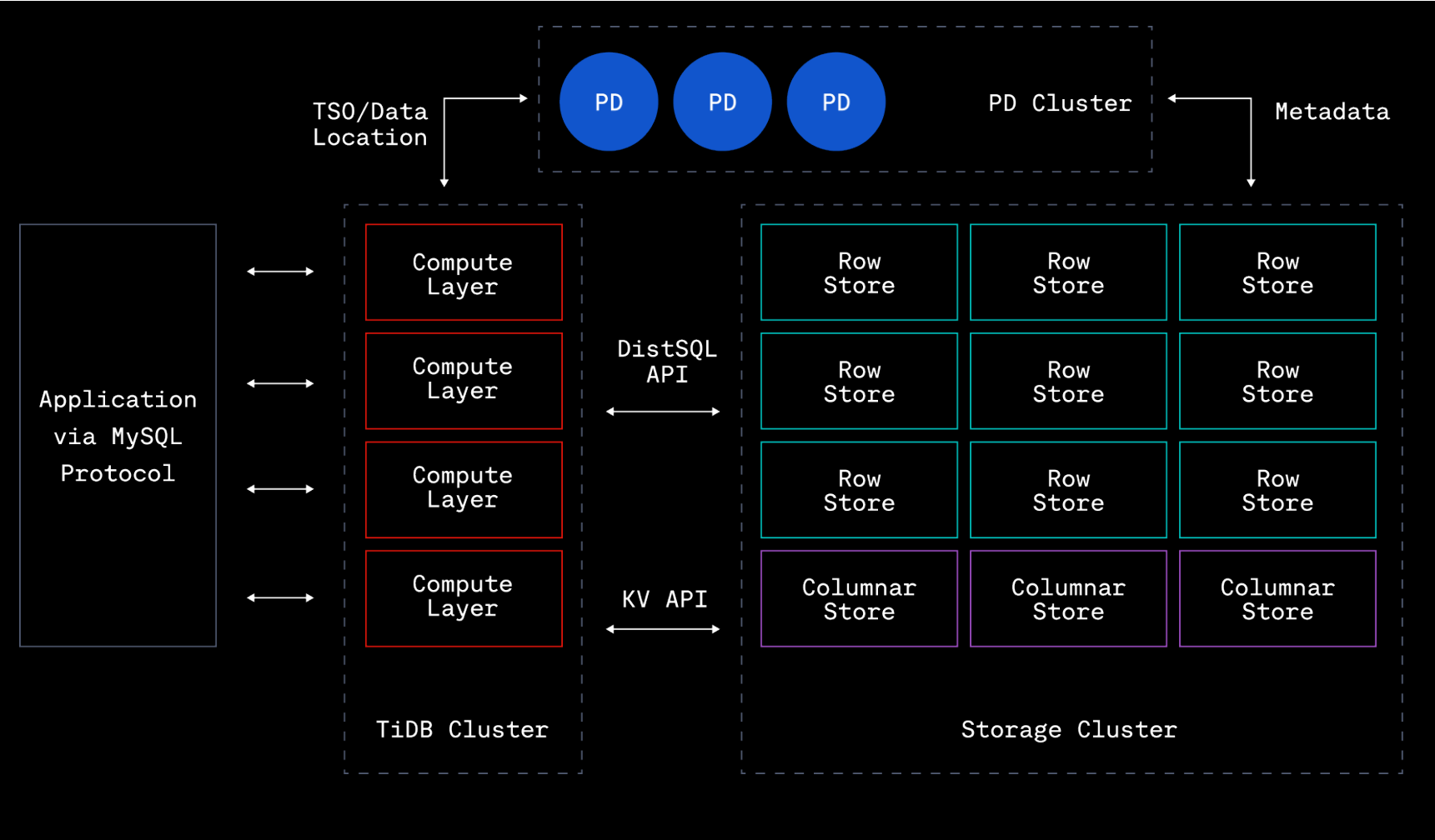
- Data model:
 - A big table, with column families
 - Map-reduce for querying/processing
- Schema-lite
- No single point of failure
- Operational simplicity
- Closest NoSQL implementation to RDBMS

Columnstore

- Most HTAP use two storage types internally: an in-memory rowstore and a disk-based columnstore
- Columnstore:



HTAP Architecture



Multi-Vendor Architectures

- **MySQL or PostgreSQL (OLTP) + Apache Cassandra (Column Store for OLAP):** Use a traditional relational database for transactional workloads and Apache Cassandra's column-store capabilities for analytical workloads.
- **MySQL or PostgreSQL (OLTP) + Apache Pinot (OLAP):** Combine a relational database for transactional workloads with Apache Pinot's real-time distributed OLAP datastore for analytical workloads.
- **MongoDB (OLTP) + Apache Spark (OLAP):** Combine MongoDB's NoSQL database for transactional workloads with Apache Spark's distributed computing engine for analytical workloads.

Benchmark

1. A MySQL, Analytic db, ETL, Redis, and Kafka Stack

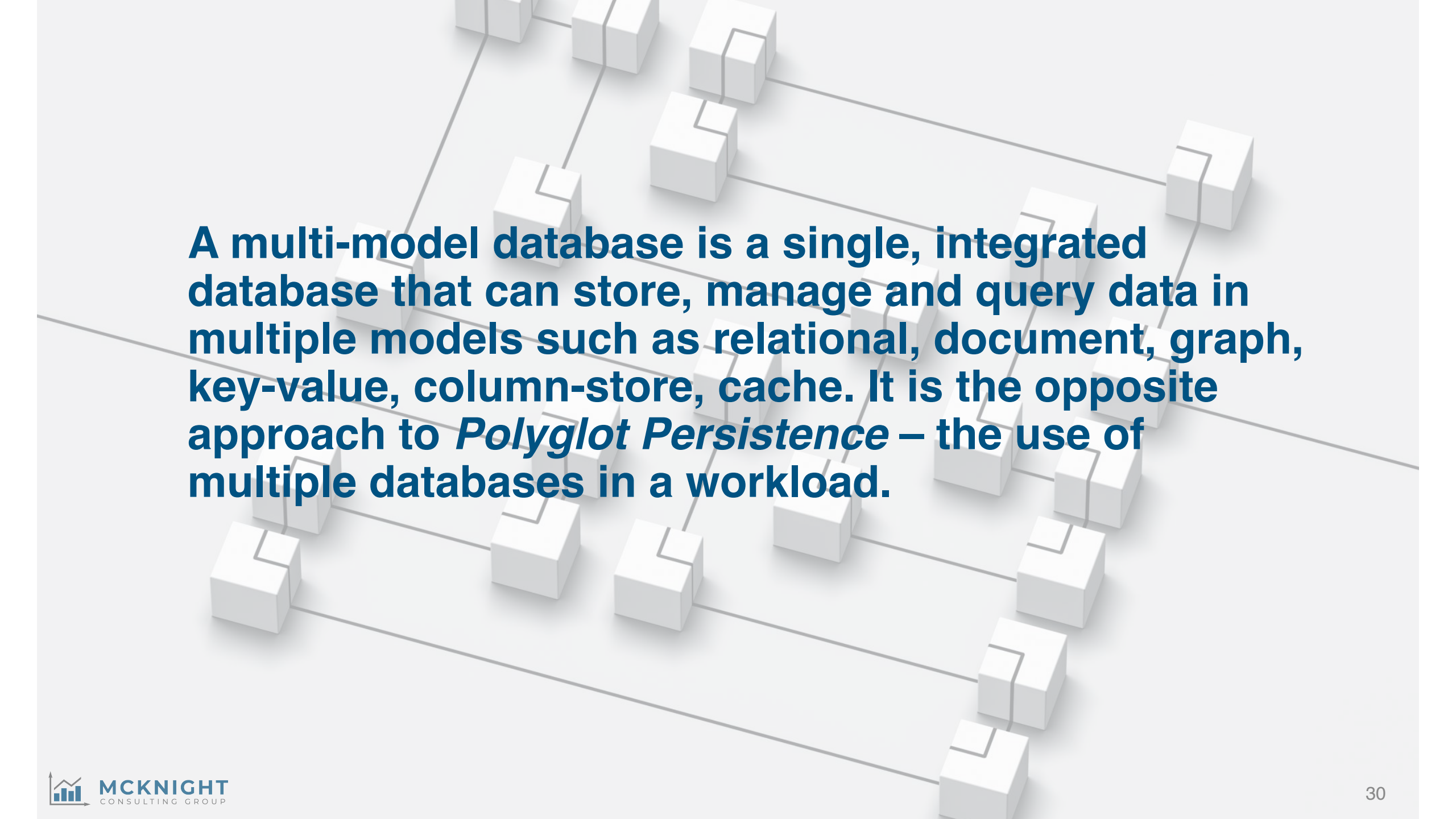
2. A PostgreSQL, Analytic db, ETL, Redis, and Kafka Stack

3. A HTAP database with Kafka Stack

- In the TPC DS-like workload, an analytic db was superior, both with maintenance and without maintenance
- Platform costs favor the 1 db by 1.9x over 1 analytic db and 2.5x over the other in Year 1.
- Development costs are 2.5x - 3x and Production Costs are 2.1x - 2.5x for the analytic db solutions.
- We calculated the annual costs of the platform stacks and the Time-Effort Costs (People Costs, Development Costs and Production Costs) and concluded that the 1 db is 2 times cheaper than 1 analytic stack and 2.5 times cheaper than the other over 3 years running enterprise-equivalent workloads.
- We found a single database competitiveness in operating effectively, actually putting it in a winning position for both transactional and operational workloads.
 - The use of a single database facilitates operational analytics and offers an efficient approach for any organization.



Multi-Modal Databases



A multi-model database is a single, integrated database that can store, manage and query data in multiple models such as relational, document, graph, key-value, column-store, cache. It is the opposite approach to *Polyglot Persistence* – the use of multiple databases in a workload.

Popularity of Multi-Model Databases

- 7 of the top 8 are “Multi-model”
- 9 of the top 10 DBMS on db-engines are “Multi-model”.
- 16% of the total are “Multi-model”.

Top 16 systems in ranking, September 2025

Rank			DBMS	Database Model	Score		
Sep 2025	Aug 2025	Sep 2024			Sep 2025	Aug 2025	Sep 2024
1.	1.	1.	Oracle	Relational, Multi-model	1170.62	-50.08	-115.97
2.	2.	2.	MySQL	Relational, Multi-model	891.77	-23.69	-137.72
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	717.32	-36.84	-90.45
4.	4.	4.	PostgreSQL	Relational, Multi-model	657.17	-14.08	+12.81
5.	5.	5.	MongoDB	Document, Multi-model	380.50	-15.08	-29.74
6.	6.	7.	Snowflake	Relational	190.19	+11.29	+56.47
7.	7.	6.	Redis	Key-value, Multi-model	145.17	-2.02	-4.25
8.	8.	9.	IBM Db2	Relational, Multi-model	124.19	-3.12	+1.14
9.	9.	14.	Databricks	Multi-model	124.06	+8.25	+39.82
10.	10.	8.	Elasticsearch	Multi-model	118.26	+3.99	-10.53
11.	11.	10.	SQLite	Relational	107.88	-4.72	+4.53
12.	12.	11.	Apache Cassandra	Wide column, Multi-model	106.98	-1.53	+8.04
13.	13.	15.	MariaDB	Relational, Multi-model	91.46	-2.13	+8.02
14.	14.	12.	Microsoft Access	Relational	83.61	-4.15	-10.15
15.	15.	17.	Amazon DynamoDB	Multi-model	80.28	-3.20	+10.22
16.	16.	16.	Microsoft Azure SQL Database	Relational, Multi-model	79.18	+3.34	+6.23

Example Use of Databases

- Oracle – Relational
- MongoDB - Document
- Neo4J – Graph
- Pinecone – Vector
- InfluxDB – Time Series
- MarkLogic – RDF
- Oracle Spatial - Spatial





What do these Trends mean to an Enterprise

Simpler Data Management

- **Unified Platform:** Instead of managing separate databases for different data workloads (OLTP, OLAP, etc.) and data types, enterprises can use a single database, simplifying administration and reducing complexity.
- **Reduced Silos:** Data can be stored together even if it has different structures, fostering better data integration and breaking down information silos within the organization.
- **Flexibility:** A HTAP multi-model database allows storing various data workloads within the same system, catering to diverse processing needs across different departments.



Improved Data Analysis

- **Unified Queries:** Querying across different data processing within a single platform. This simplifies complex data analysis tasks that might require joining data from disparate sources.
- **Deeper Insights:** By bringing various data processing and types together, enterprises can gain a more holistic view of their operations and uncover hidden patterns or trends that might be missed with isolated data stores.



Other Benefits

- **Reduced Costs:** Consolidating into a single platform can potentially lead to cost savings compared to maintaining multiple specialized databases.
- **Faster Development:** The ability to store and query different data types within a single system can streamline application development, allowing enterprises to bring products and services to market faster.



Summary

Need

- Growing complexity for managing these many data workloads enterprises need.

Workloads

- Separation of OLTP and OLAP not as true anymore.

Use Cases

- Use cases include Fraud Analytics, Real-Time Customer Experience, Predictive Maintenance, Risk Management, Algorithmic Trading.

Database Requirements

- Translytical databases require specific fast streaming ingest of millions of events per second, low latency, high concurrency to support thousands of users, unlimited storage, robust pipeline management, and transactional consistency to ensure data accuracy and reliability.

Architecture

- Typically HTAP stores data as both row and columnar. The HTAP architecture determines where to retrieve the data.

Multi-Model

- A multi-model database is a single, integrated database that can store, manage and query data in multiple models such as relational, document, graph, key-value, column-store, cache

Benefits of Simplification

- Benefits of these simplifications include simpler data management, improved data analysis, reduced costs and faster development.



Translytical Databases: A Framework for Evaluation and Use Case Analysis

Questions on data & AI strategy,
architecture, evaluating products
or formulating projects?

Email:
wmcknight@mcknightcg.com

Call/Text: 214-514-1444

www.mcknightcg.com

