MemSQL:
The World's fastest Operational DB

Lejo Jacob, Solutions Architect, San Francisco

August 15, 2019

MEMSQL

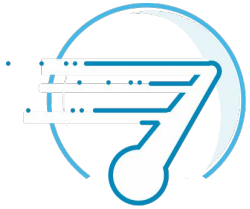
About Me:

Lejo Jacob

MemSQL Solution Architect West Lead

2 Years @ MemSQL
3 years at another Analytics company
Over 10 years at various Financial companies

The Journey with MemSQL



ACCELERATE
Your Enterprise

SIMPLIFY
Your Architecture

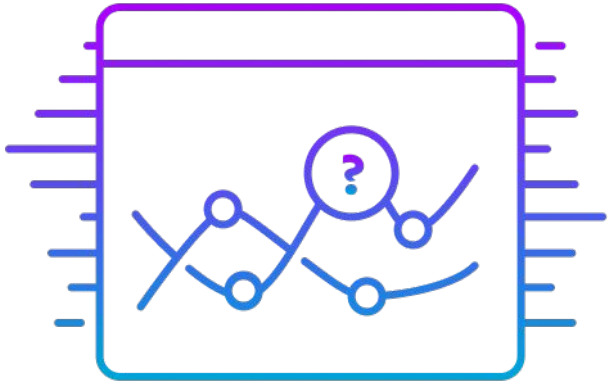
WIN
With MemSQL

MEMSQL

What is MemSQL?

Scalable relational database for transactions and analytics

What is Operational Analytics?



Operational Analytics is associated with the data that Companies need to improve their existing operations.

Companies are starting to view it as a strategic priority. Operational Analytics plays an important role in driving profits or creating competitive advantage.

MemSQL is in 5 of the top 10 banks in US

20X ETL Processing

72TB Per Day

20 Million Upserts per Second

MEMSQL

2 of the Top 3 Telcos

20X ETL Processing

72TB Per Day

20 Million Upserts per Second

MEMSQL

12 of the Fortune 50

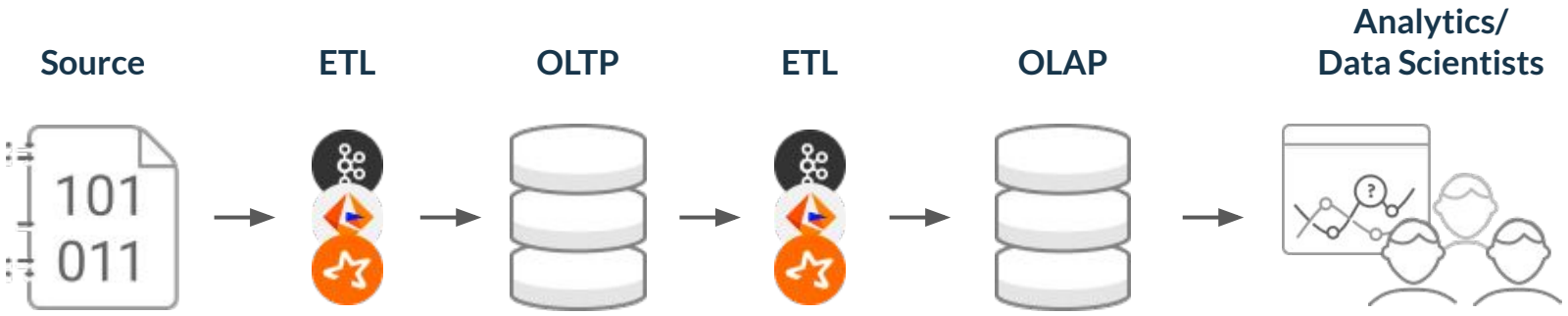
20X ETL Processing

72 TB Per Day

20 Million Upserts per Second

MEMSQL

Latency: Typical Enterprise Environments

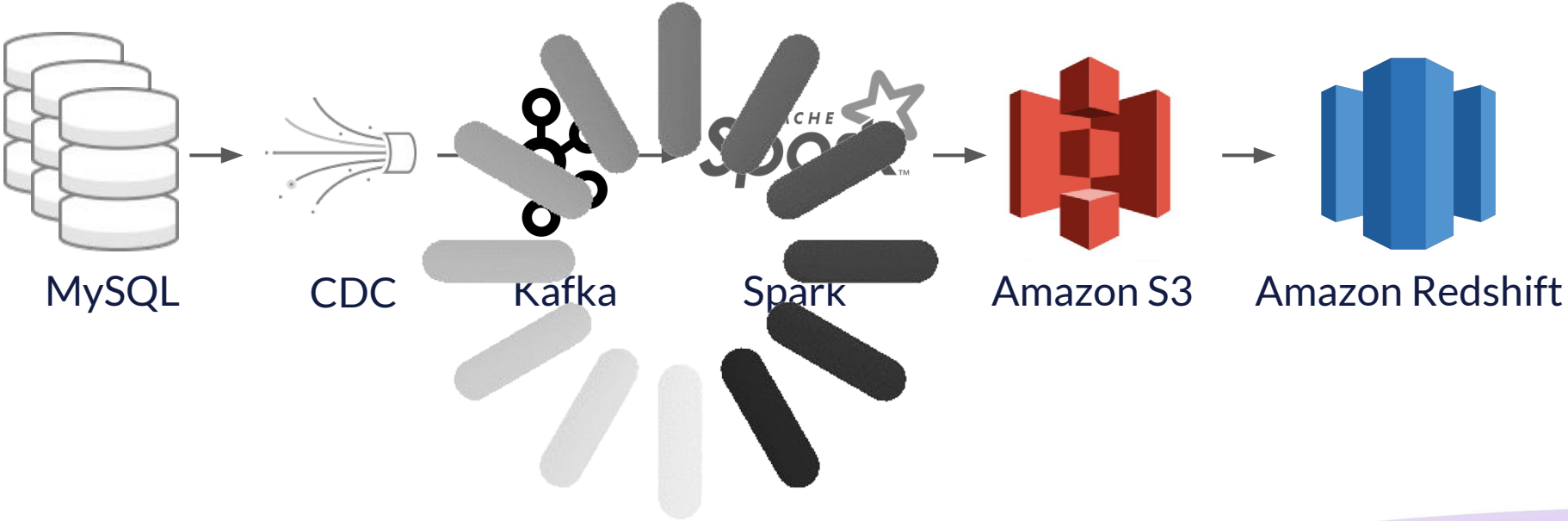


MEMSQL

Sounds Familiar?

MEMSQL

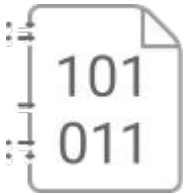
Welcome to the Spin-Zone !



MEMSQL

MemSQL: The No Spin Zone

Source



Tick Data, Sensor, IoT, Pos

Real-Time Platform

Streaming Ingest, Rowstore & Columnstore

Analytics/ Data Scientists

MEMSQL

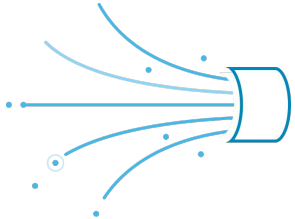
ACCELERATE

Your Enterprise

MemSQL Powers Real-Time Enterprises

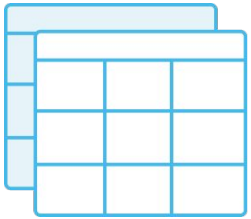
A single platform for all real-time data needs

Scalable SQL



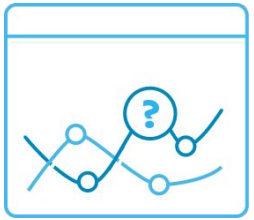
Integrated streaming with Pipelines

Transactional Database Workloads



High volume transactions for structured and semi-structured data

Data Warehouse Workloads



Fast, scalable SQL for immediate analytics

MEMSQL

MemSQL in One Slide

Scalable SQL

SIMPLE

- Full ACID features
- Lock free: OLTP, OLAP, HTAP
- ANSI SQL
- Document/JSON
- Geospatial

In-Memory & Solid-State

REAL-TIME

- In-Memory Rowstore
- Solid-State Columnstore
- Stream directly to rowstore or columnstore

Distributed

LOW COST

- Shared nothing, massively parallel
- Scale-out with commodity hardware

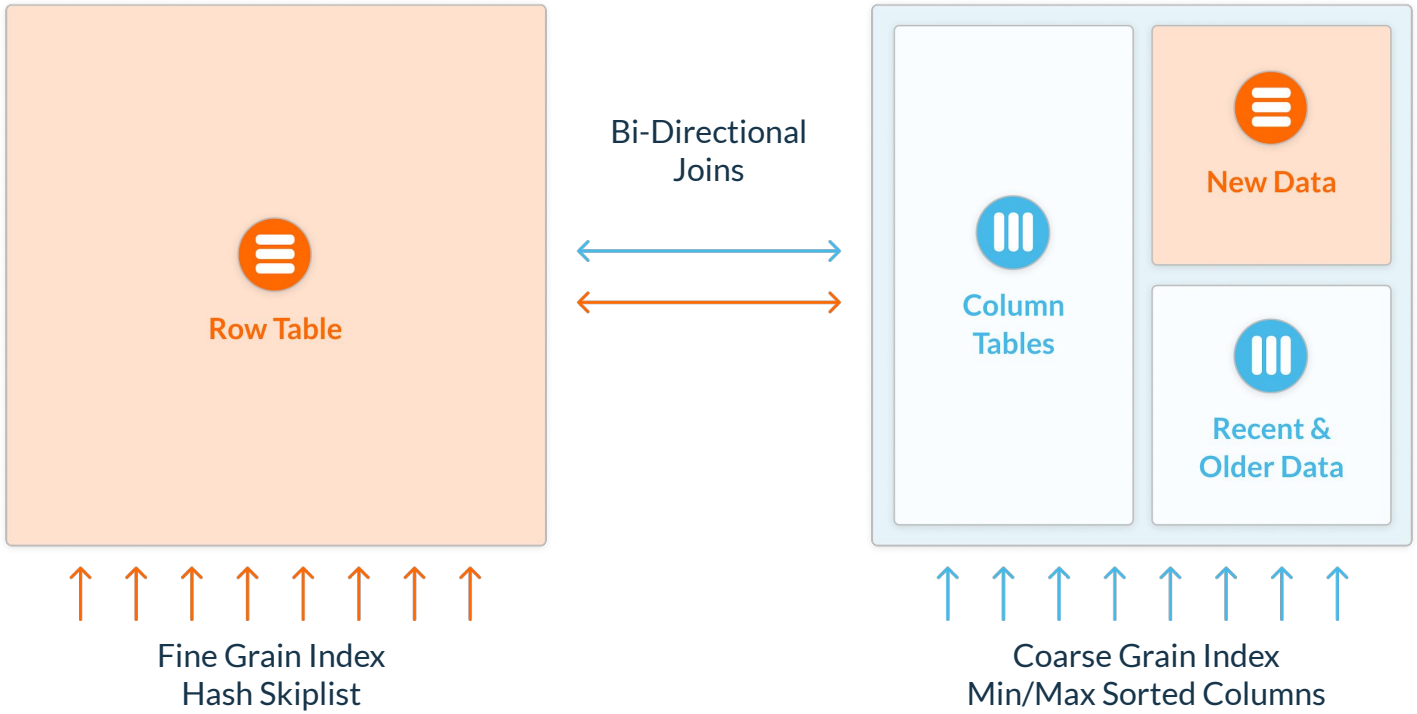
Datacenter or Cloud

FLEXIBLE

- Deploy on-premises
- Cloud agnostic
  - Amazon, Microsoft, Google, Digital Ocean

MEMSQL

Hybrid Transactional & Analytical Workloads



MEMSQL

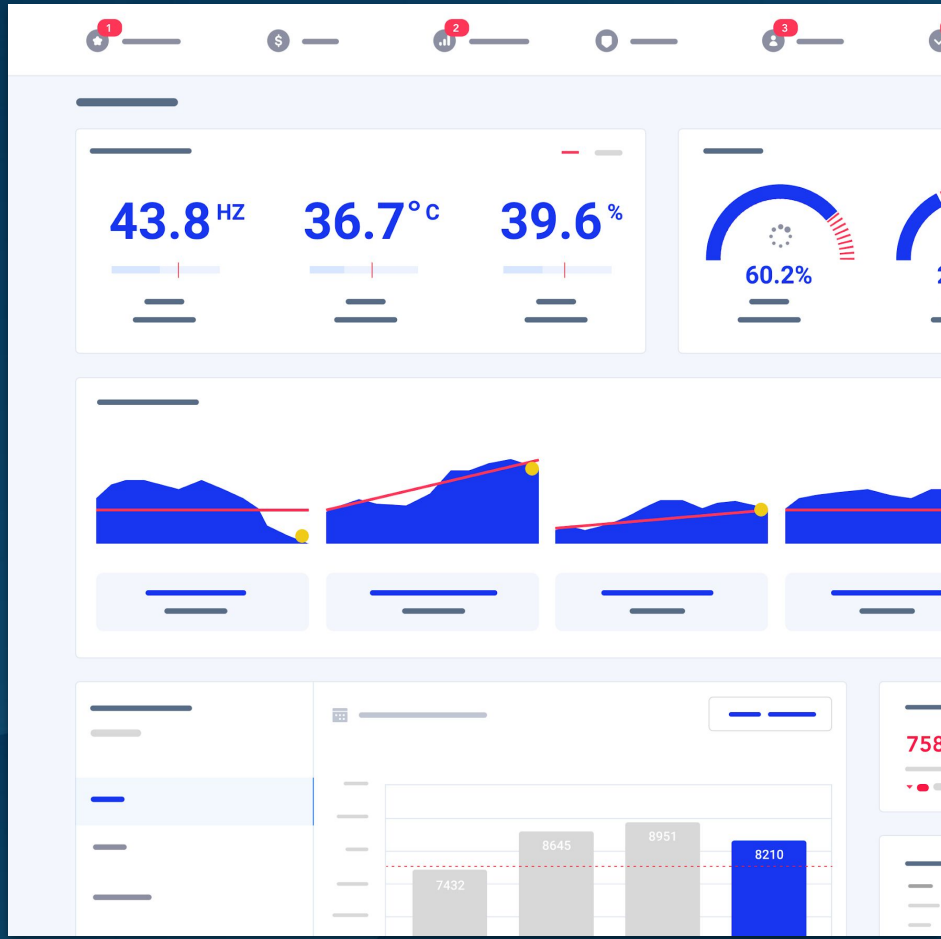
MemSQL Operational Analytics Use-cases

MEMSQL

MEMSQL + Top US Oil Producer

IOT Analytics
Predictive Maintenance, Supply Chain Analytics, and Utility Grid Management

Using MemSQL a top energy provider ingests drill bit data and reports back key indicators within minutes allowing operators to respond to live conditions preventing costly bit failures and outages

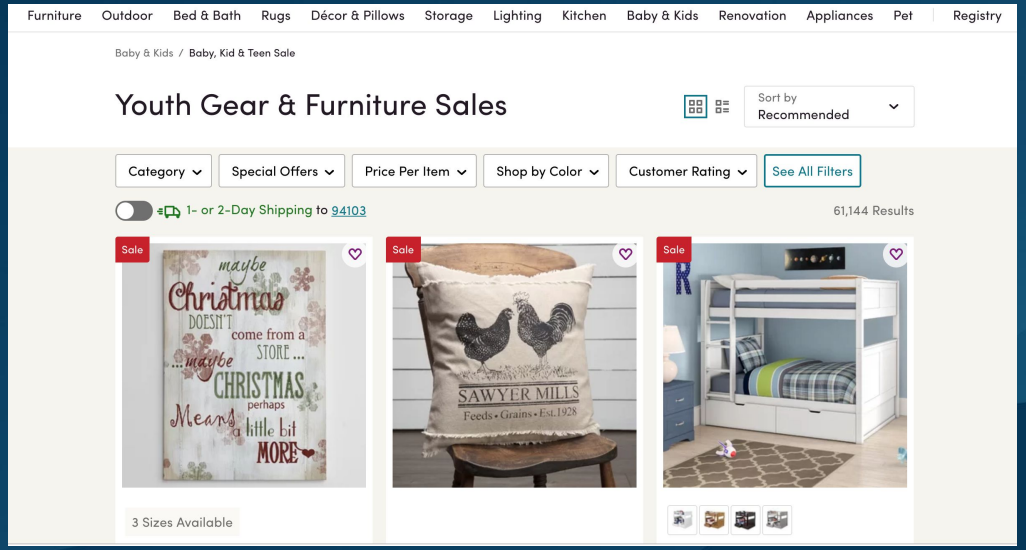


MEMSQL

MEMSQL + Top E-Commerce Company

Real-time A/B Testing

Using MemSQL a top retail company is able to A/B test their retail Website and ensure that the A/B experiments does not adversely impact revenue.
If the experiment is impacting revenue they can stop the experiment immediately.
MemSQL directly helps Customer's bottomline



MEMSQL

MEMSQL + Top US Financial Company

Operational Analytics on 100s of OLTP DBs

Customer has 100s of SQL Server database to support their end Customers.
ETL from these database is slow.
Using MemSQL they are able to CDC data from all the SQL Server DBs into one Cluster of MemSQL and run Operational Analytics on top of them

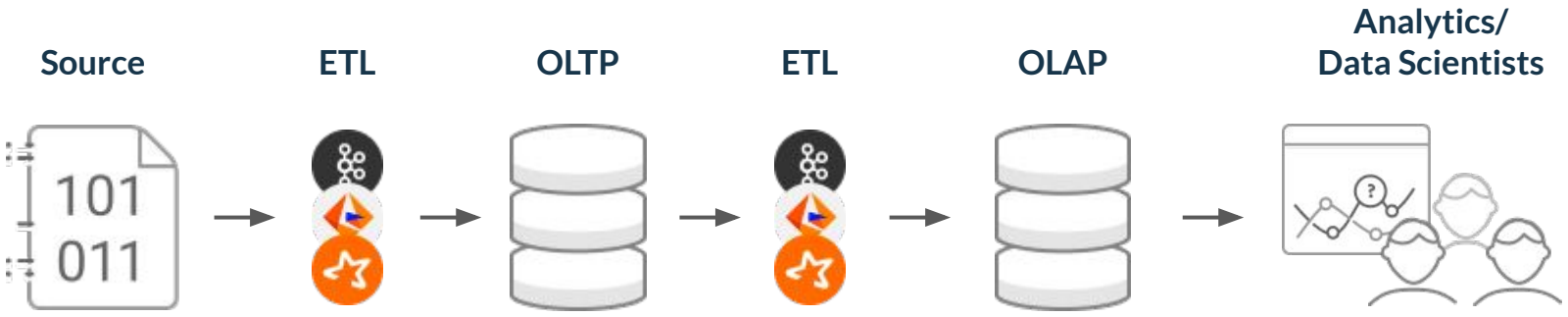
Database Comparison

	Postgres	Postgres Arora	Redshift with Spectrum	Snowflake	MemSQL	Comments
API Latency	10-20 minutes (50%+ timeouts)	1-2 secs	2 secs	< 1 seconds	100-200 ms	For V Large Account
RBAC	every 2-5 min	every 15 mins	every 15 mins	every 15 mins	Realtime lookup	
Concurrency	60-80	40-80	10-20	8-16	1000+ tested	
Simplicity	350+ Tables (Lots of agg tables)	100+ Tables (Lots of agg tables)	70 tables (Some agg tables)	50 tables (Some age tables)	25 tables (No agg tables)	On the fly Agg No micro ETL's
Time to report on new data	1-2 mins	5 mins	2-3 mins	2-3 mins	1-2 sec (Kafka integration)	Some tech needs micro batching in S3
Scalibility	++	+++	+++	+++++	+++++	
Operational Factors	+++	+++	++	+++++	+++	Cost + Maintainance
Ease of Integration	+++	+++	++++	+++++	+++++	with current components S3, Kafka etc + Data structures
Perf tuning options	+++++	+++++	+++	++	+++++	

SIMPLIFY

Your Architecture

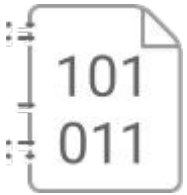
Latency: Typical Enterprise Environments



MEMSQL

MemSQL: Real-Time Enterprise

Source



Tick Data, Sensor, IoT, Pos

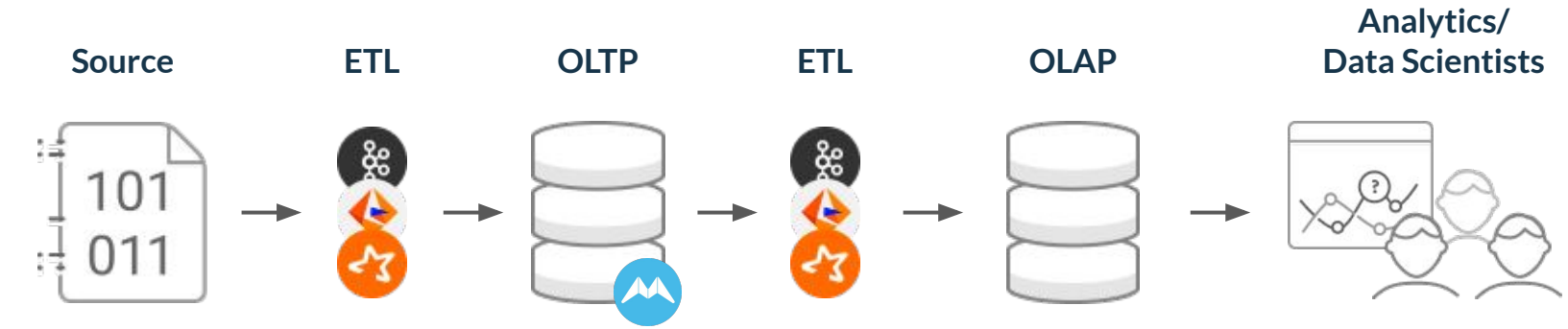
Real-Time Platform

Streaming Ingest, Rowstore & Columnstore

Analytics/ Data Scientists

MEMSQL

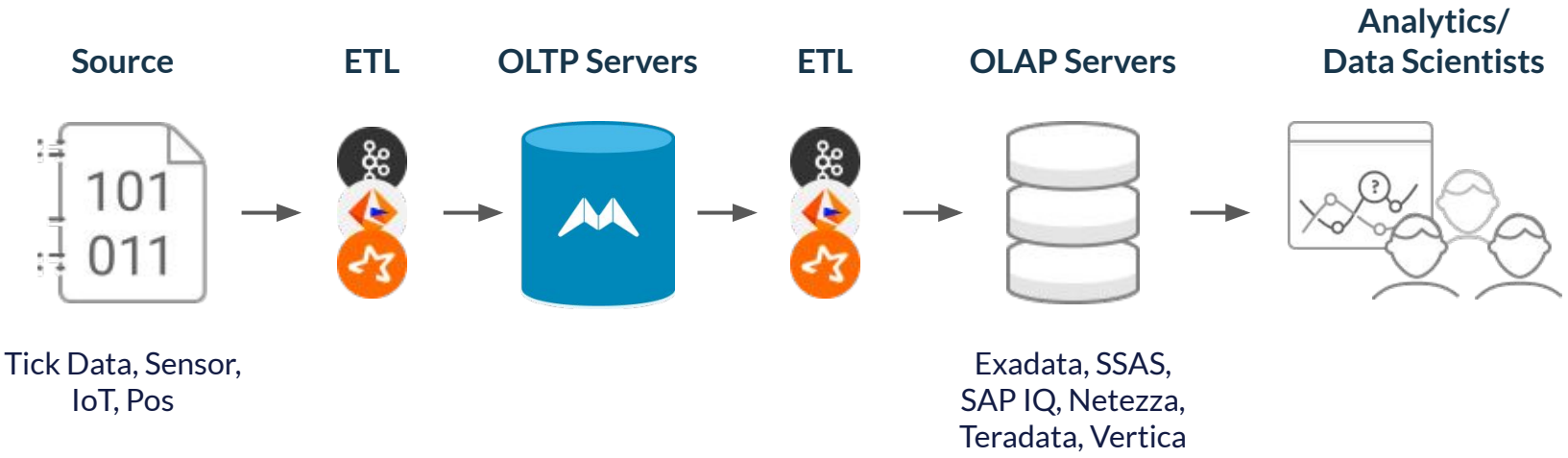
MemSQL: OLTP Augmentation



Augment your legacy OLTP DB with MemSQL

MEMSQL

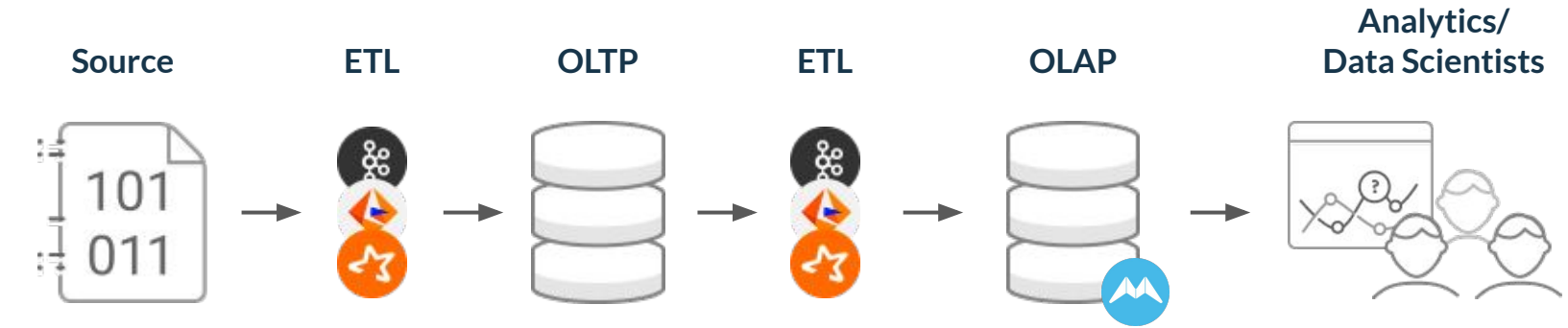
MemSQL: OLTP Replacement



Replace your legacy OLTP DB with MemSQL

MEMSQL

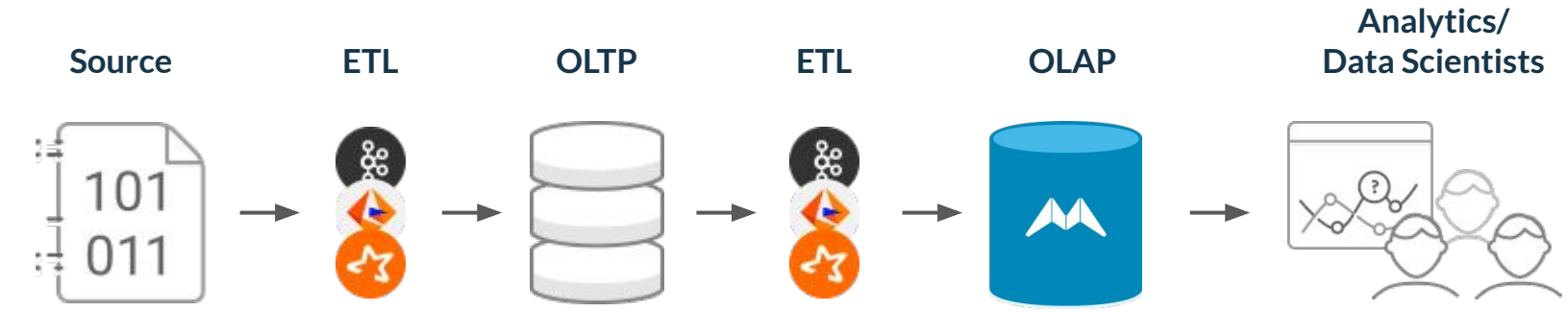
MemSQL: Legacy EDW Augmentation



Augment your legacy OLAP DW with MemSQL

MEMSQL

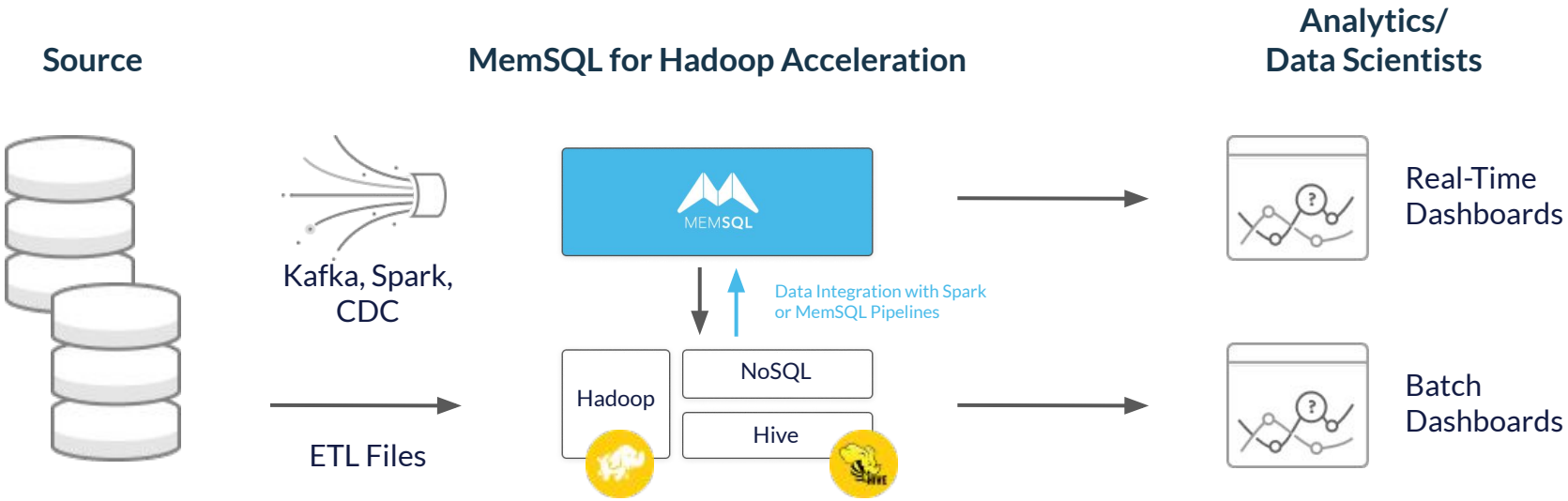
MemSQL: EDW Acceleration



Replace your legacy OLAP DW with MemSQL

MEMSQL

MemSQL: Hadoop/Datalake Augmentation



Augment your Datalake with MemSQL

MEMSQL

WIN

with MemSQL

MemSQL Resources

http://docs.memsql.com

http://blogs.memsql.com

http://www.youtube.com/memsql

http://www.slideshare.net/memsql

Find me on LinkedIn!

https://www.linkedin.com/in/lejojacob

MEMSQL

MemSQL Deep Dive

Hybrid Transactional & Analytical Workloads

Row Oriented Tables

Transactional

- Fine grain Aggregation
- Seek, Update, Delete Millions of Rows with GBs/TBs of data
- No compression

Column Oriented Tables

Analytical

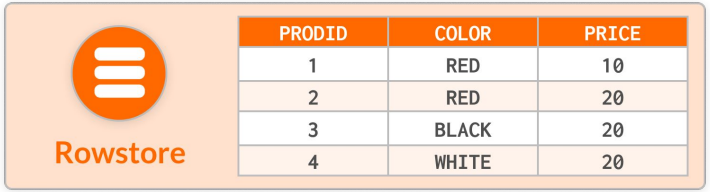
- Bulk Aggregation
- Seek, Update, Delete Billions of Rows with TBs/PBs of data
- 5x to 10x compression

Massive concurrent streaming ingestion with transactional row level MVCC

Hybrid Transactional & Analytical Workloads

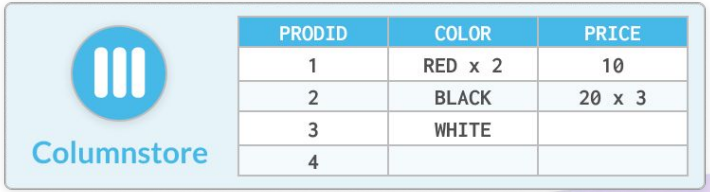
Row Oriented Tables

- Best for OLTP
- Millions to Billions of rows
- Seek specific rows
- Hash and B-tree index
- Few milliseconds response
- OLTP & Batch Insert, Update, Delete
- Row lock on update and delete



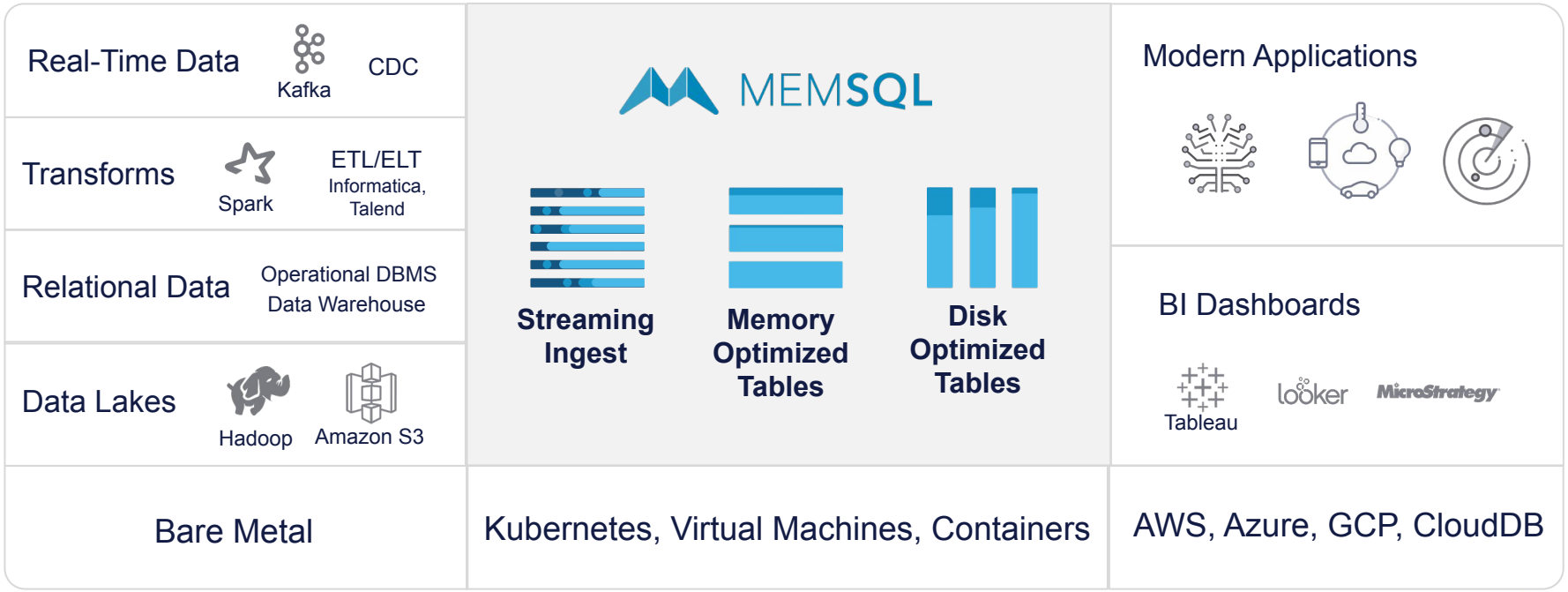
Column Oriented Tables

- Best for OLAP
- Billions to Trillion rows
- Scan many rows
- Sort and Min/Max index
- 10s milliseconds response
- Batch Insert, Update, Delete
- Segment lock on update and delete



MEMSQL

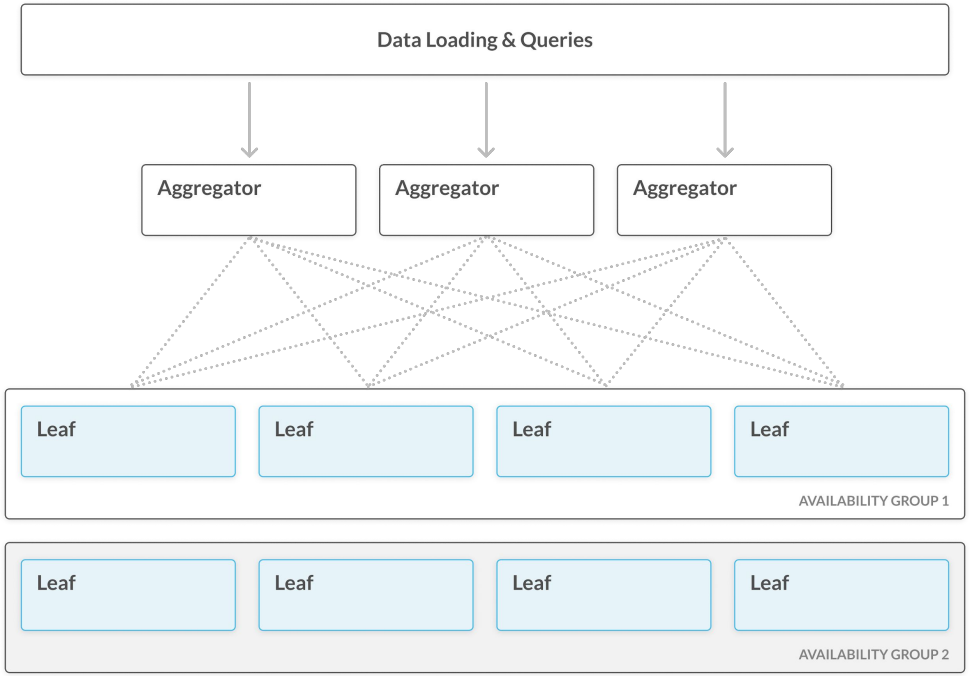
Ecosystem Overview



MEMSQL

MemSQL Distributed Architecture

Shared Nothing Architecture | Distributed Query Optimizer | Highly Available, Fault Tolerant | Commodity Hardware



Aggregator Linear Scalability

- High concurrent OLTP and OLAP
- High throughput data ingestion

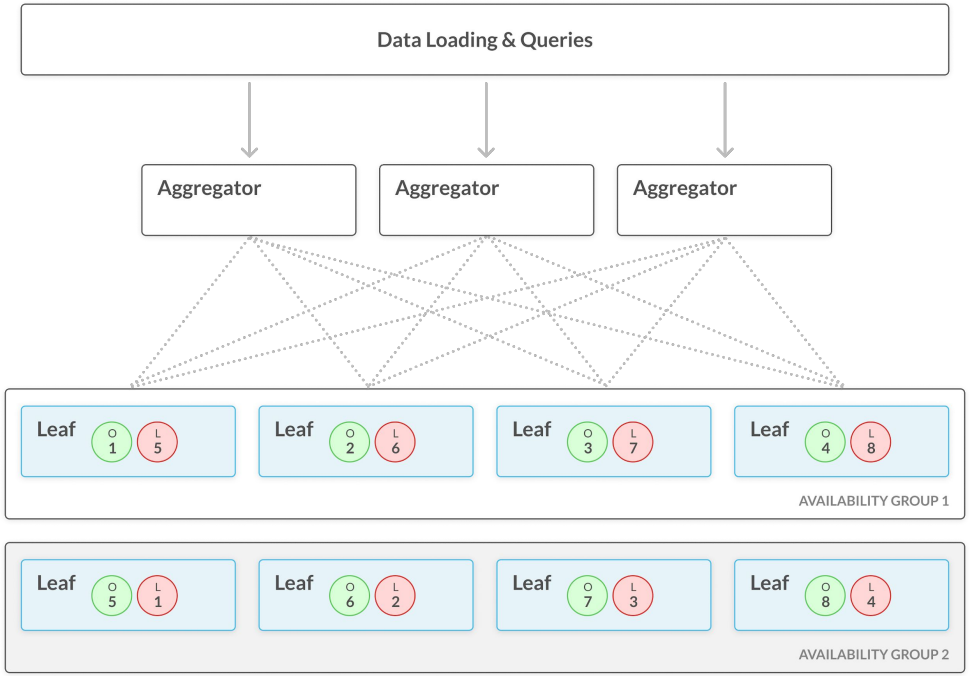
Leaf Linear Scalability

- High data volume
- High throughput ELT

MEMSQL

MemSQL Distributed Architecture

Shared Nothing Architecture | Distributed Query Optimizer | Highly Available, Fault Tolerant | Commodity Hardware

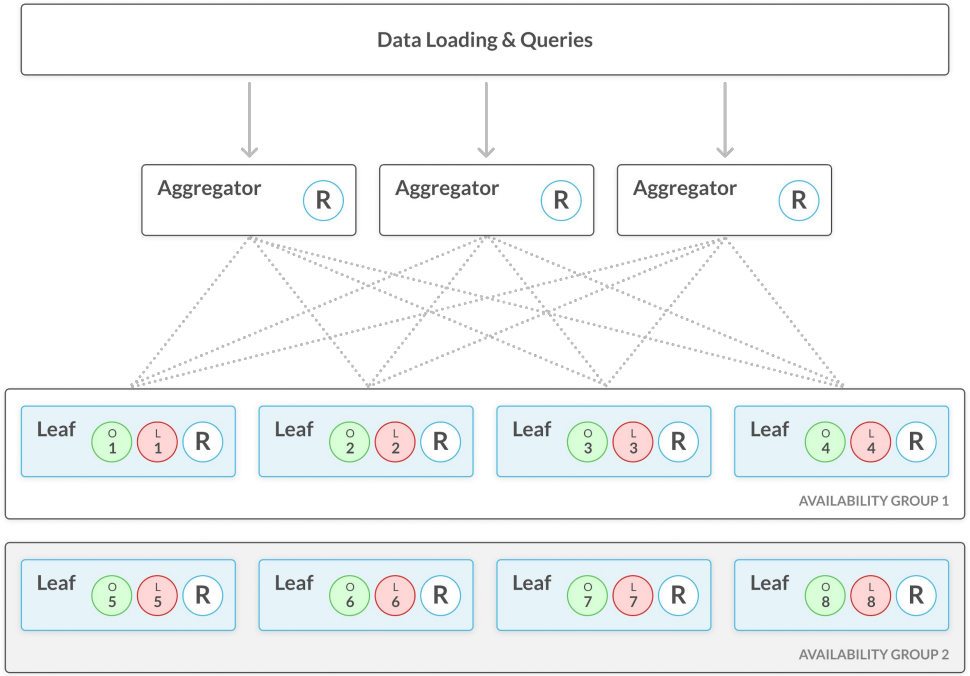


Data is replicated between Leaf Nodes for HA

MEMSQL

MemSQL Distributed Architecture

Shared Nothing Architecture | Distributed Query Optimizer | Highly Available, Fault Tolerant | Commodity Hardware

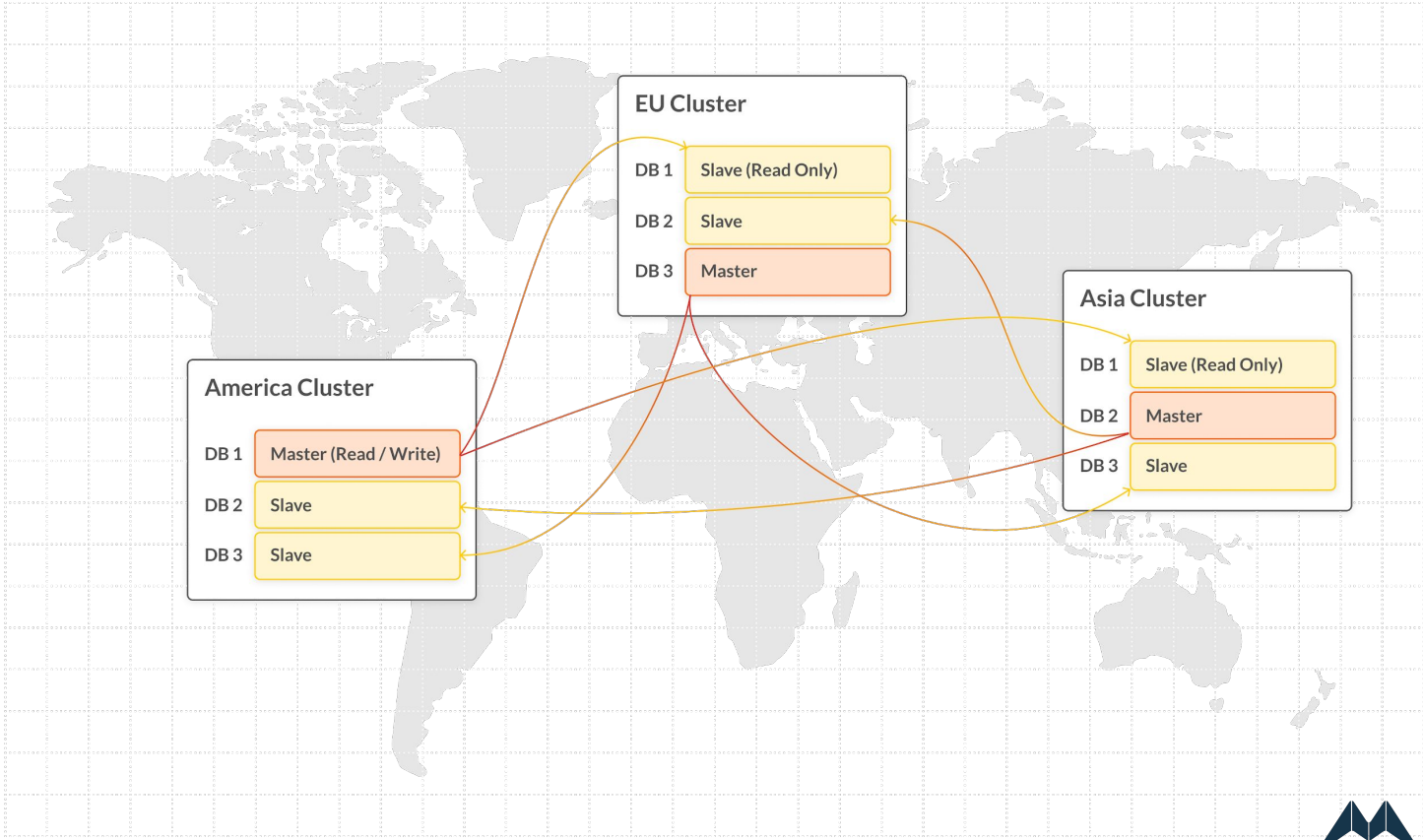


Reference Tables are duplicated on all nodes

NOTE: Dimension Tables are good candidates for Reference Tables

MEMSQL

Fine Grain Database (Schema) Level Replication



MEMSQL

Thank You!

MEMSQL