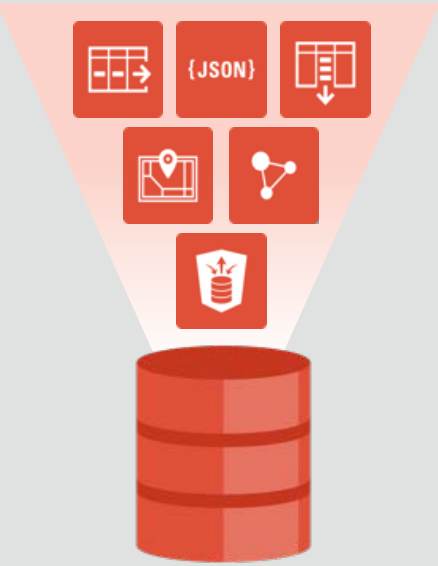
ORACLE

Data Architecture Dilemma

Many Single-Purpose Databases?

Or

A Converged Autonomous Database?



Tirthankar Lahiri

Senior Vice President, Data and In-Memory Technologies

A Brief History of Enterprise Data Management

HISTORY

Mainframe Era

Mid 60s – Early 80s

Client-Server Era

Early 80s – Mid 90s

Internet Era

Mid 90s – Mid 2000s

Cloud Era

Mid 2000s – Present

These are the Voyages of the Modern Enterprise

On its continuing mission …

To explore strange new data models and unique new workloads …

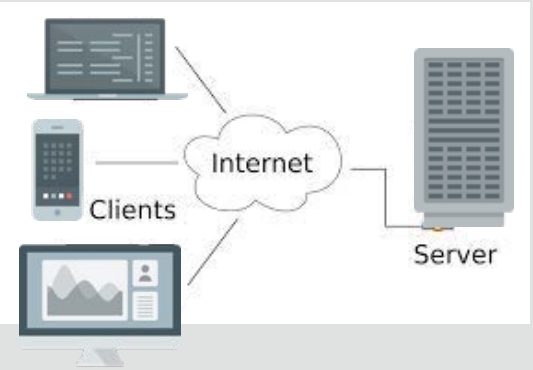
To Boldly Go Where No Database has Gone Before

Mainframe Era: The War of The Paradigms

- Hierarchical (IMS) or Network Models (Codasyl)
  - Data organized as trees or graphs with links
  - Data navigation was imperative:
  - Programmers described how to access data

- The RDBMS (e.g. System R) was an upstart competitor
  - Data is organized into logical sets with relationships
  - Data navigation is declarative
  - SQL describes the question, not how to get the answer

Client-Server Era: The RDBMS *Wins*



- **The RDBMS became the defacto data management standard**

- Key reason: **Developer Productivity**
  - Declarative language and data model: Simple development
  - Standard language and APIs: Reuse developer skillsets
  - Consistent Transactions: Database guaranteed atomicity
  - Integrity of data automatically enforced by constraints

- Developers can focus on **Application Logic (i.e. Development)**
  - Data access, cleanliness, consistency handled by the database

Internet Era: New Extreme Requirements

- **"The internet changes everything"** – Larry Ellison
- **"The network is the computer"** – Scott McNealy

New Requirements	
Volume	Exponential growth of data
Scalability	Massive online user populations, much larger than enterprise
Availability	Requirement to be always on, 24x7 access to data
Variety	Many different data models (text, XML, Graph, Spatial) Object, XML databases appeared but vanished once mainstream DBs implemented them
Security	Much larger attack surface area than enterprise

Cloud Era: Developers and Microservices

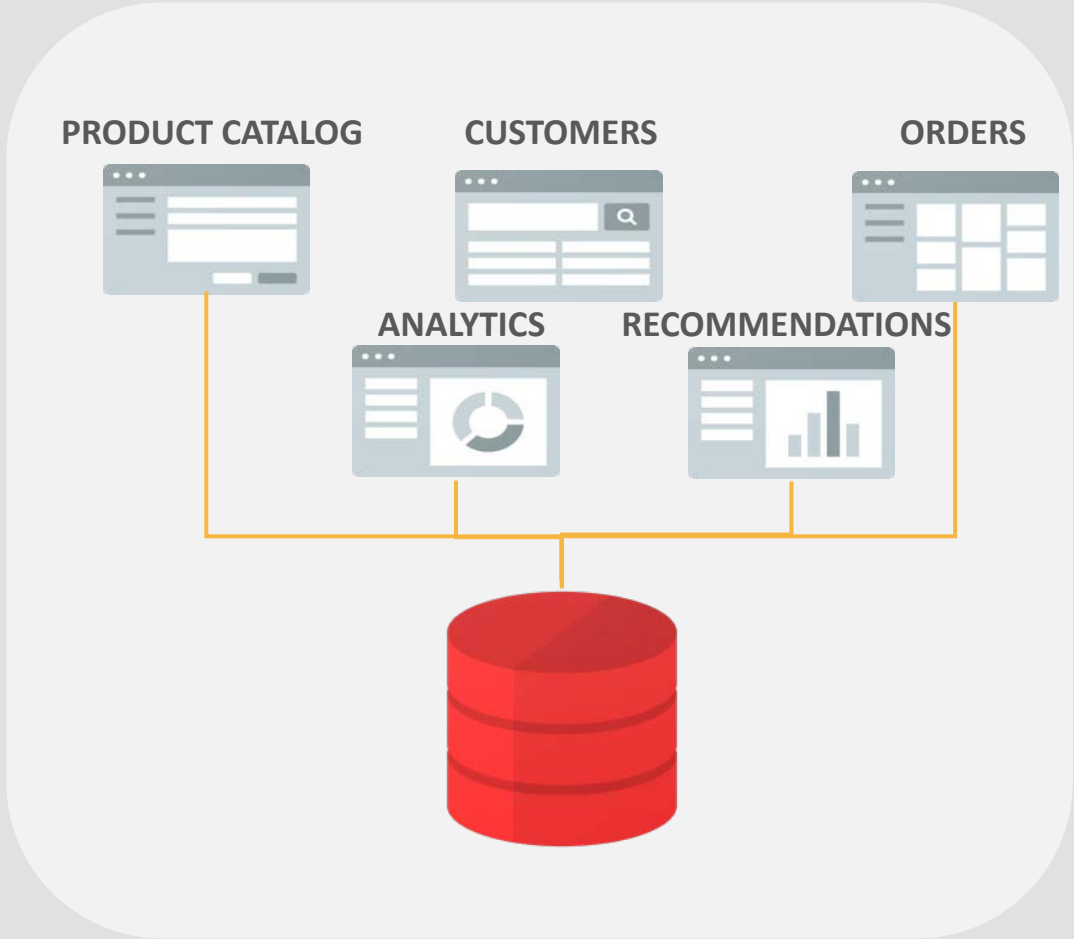
“The Cloud Changes Everything”

- Deployment, Administration, specially **Development**
  - Empowers and Democratizes Development
  - Developers are also users (of database services)
  - Developer communities grow to planet scale

- Key transformation: Lightweight **Microservices**
  - Microservices encapsulate functionality of specific modules
  - Modular, loosely coupled and communicate via events
  - Agile: Can be developed and upgraded independently
  - Highly available: Microservices can fail independently

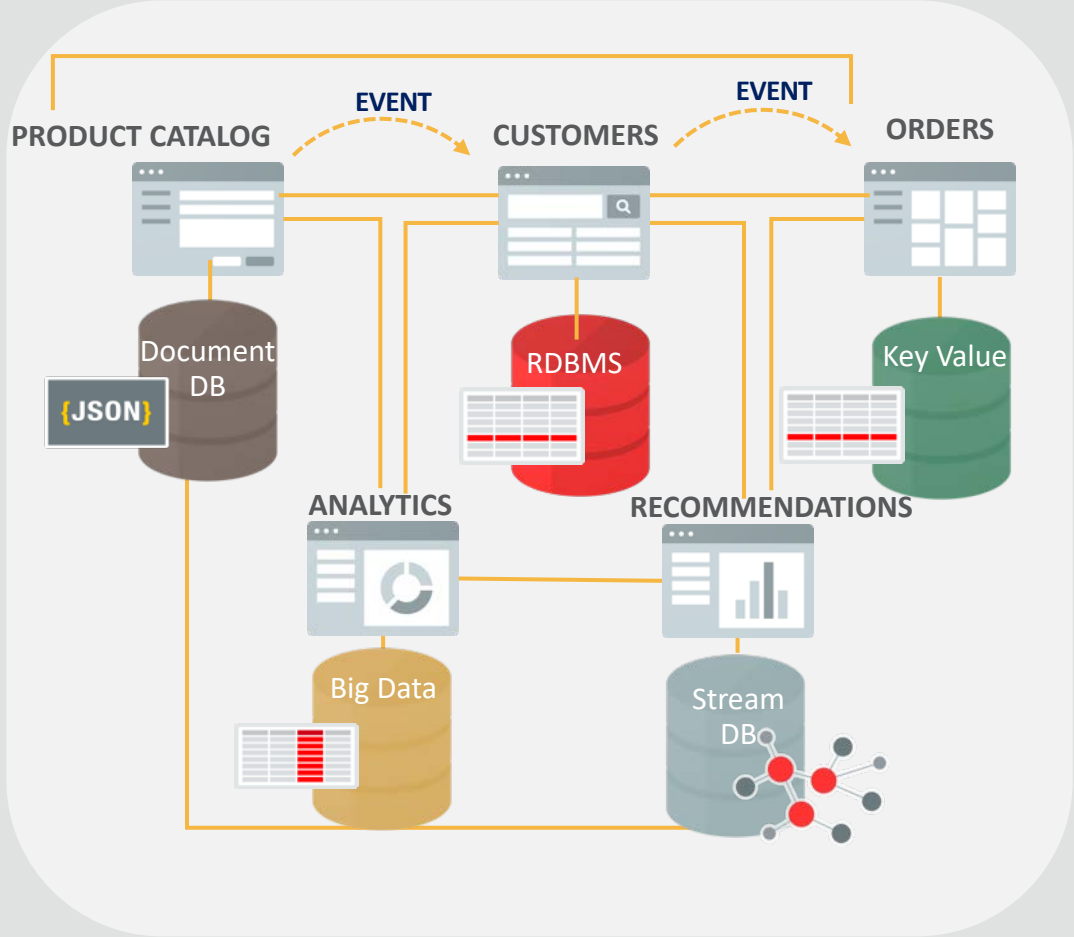
Cloud Era: Issues with "Monolithic" Databases

- Single Model database architecture sometimes slowed development:
  - Developers needed consensus on Data Model and Schema
- Single Database architecture sometimes could not meet cloud scale requirements



Cloud Era: Single Purpose Makes a Comeback

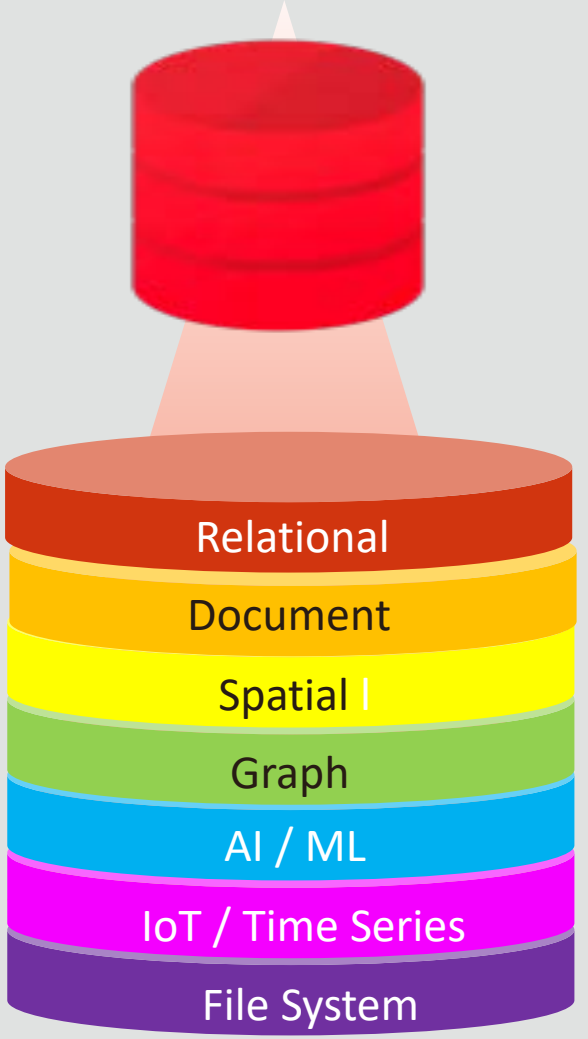
- Trend started by web giants:
  - Build Custom DBs for specific uses
  - Google Big Table, Facebook Cassandra, etc.
- NoSQL products began to proliferate
  - MongoDB, Couchbase, Redis, Neo4J …
  - Different data models, languages, APIs
- Provided scalability & performance
  - With tradeoffs in functionality



Cloud Era: Oracle Becomes a Converged Database

- Oracle Database continued to evolve far beyond a traditional Relational DBMS

- **Converges many** data models and workloads:
  - Class Leading Document Database with JSON, Text, XML
  - Class Leading Database for AI and Machine Learning
  - Class Leading Spatial and Graph Database
  - Class Leading support for IoT, Times Series, Binary Data, Database Resident Filesystem, Objects, etc.

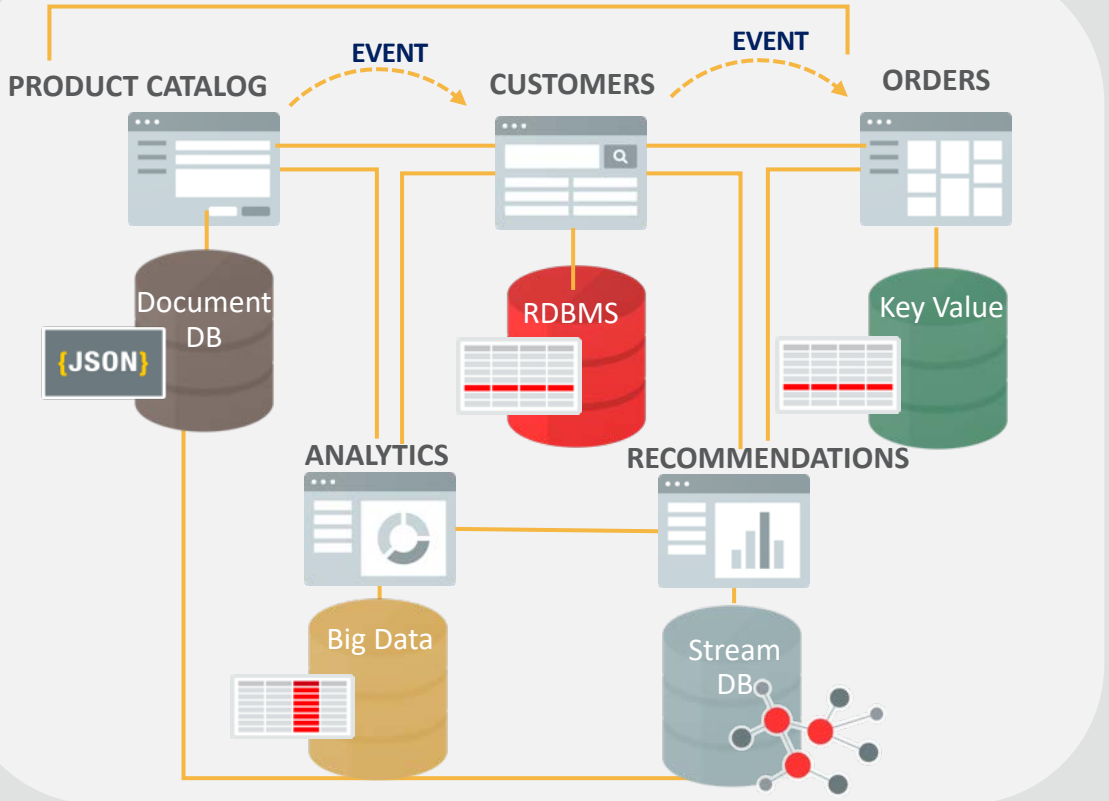


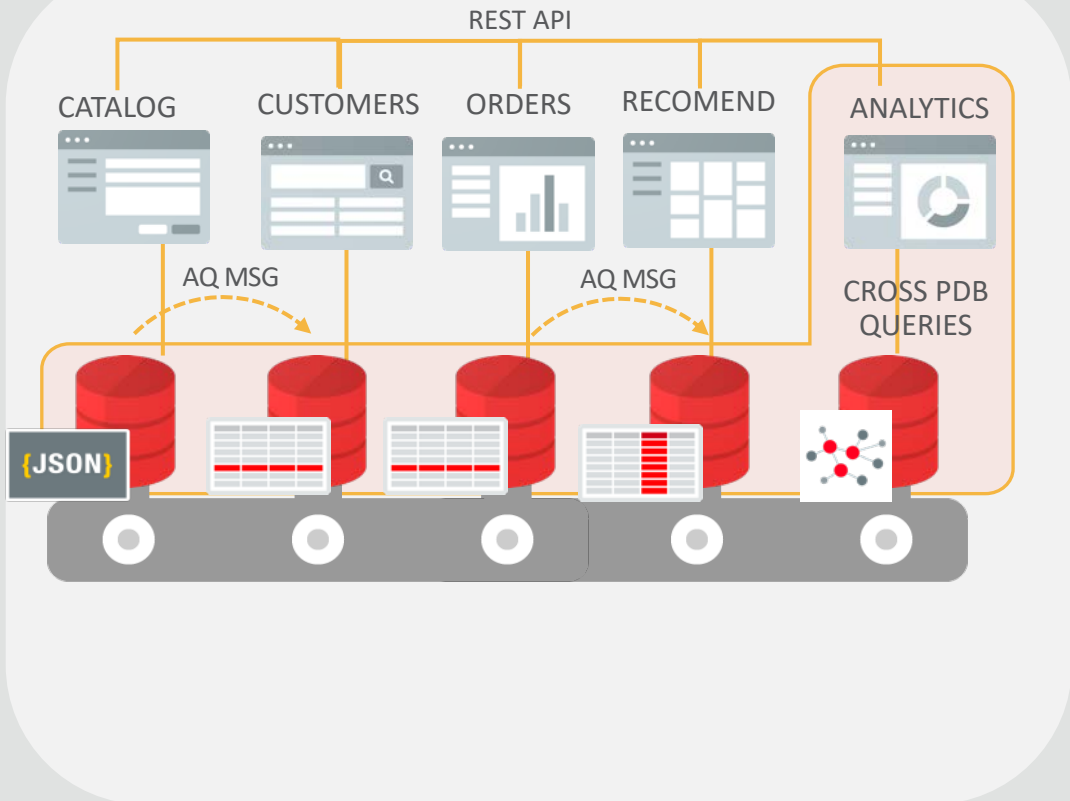
Meet the Present Day

Enterprise Data Architecture Dilemma

Today's Enterprise Data Architecture Dilemma

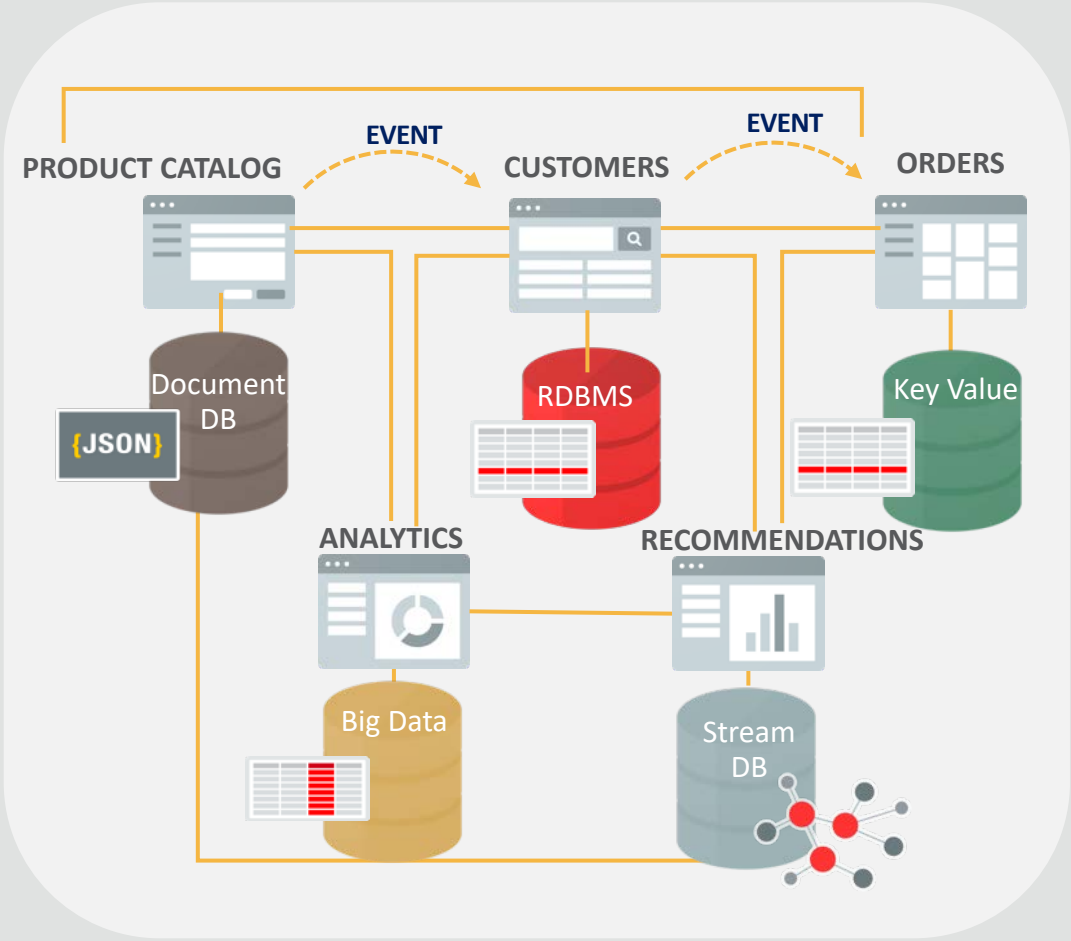
MANY SINGLE-PURPOSE DATABASES OR A CONVERGED DATABASE ?





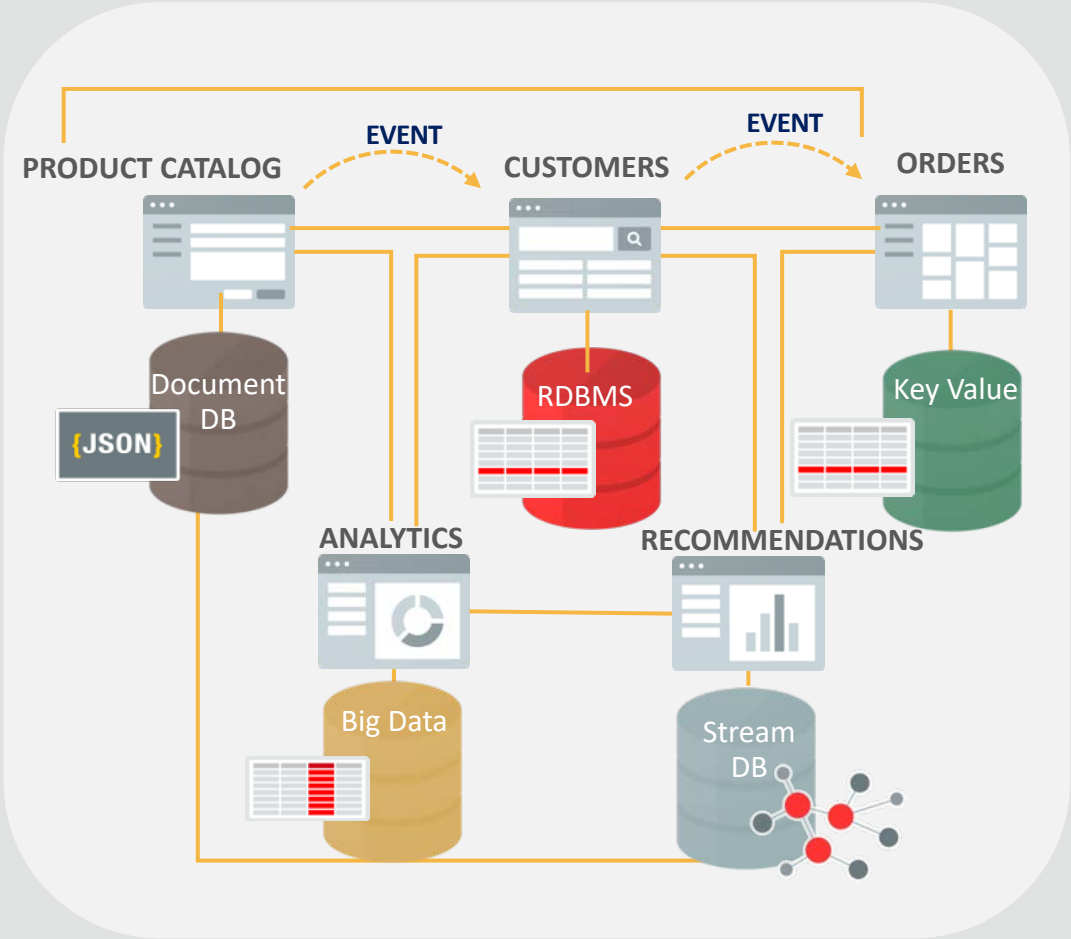
Single Purpose Databases: Microservice Friendly?

- Single-purpose engines may seem better for micro-services
- Each function has a separate database optimized for that use
- Functions are independent, and independent databases are also failure independent



Single Purpose Databases: Separating Myth from Reality

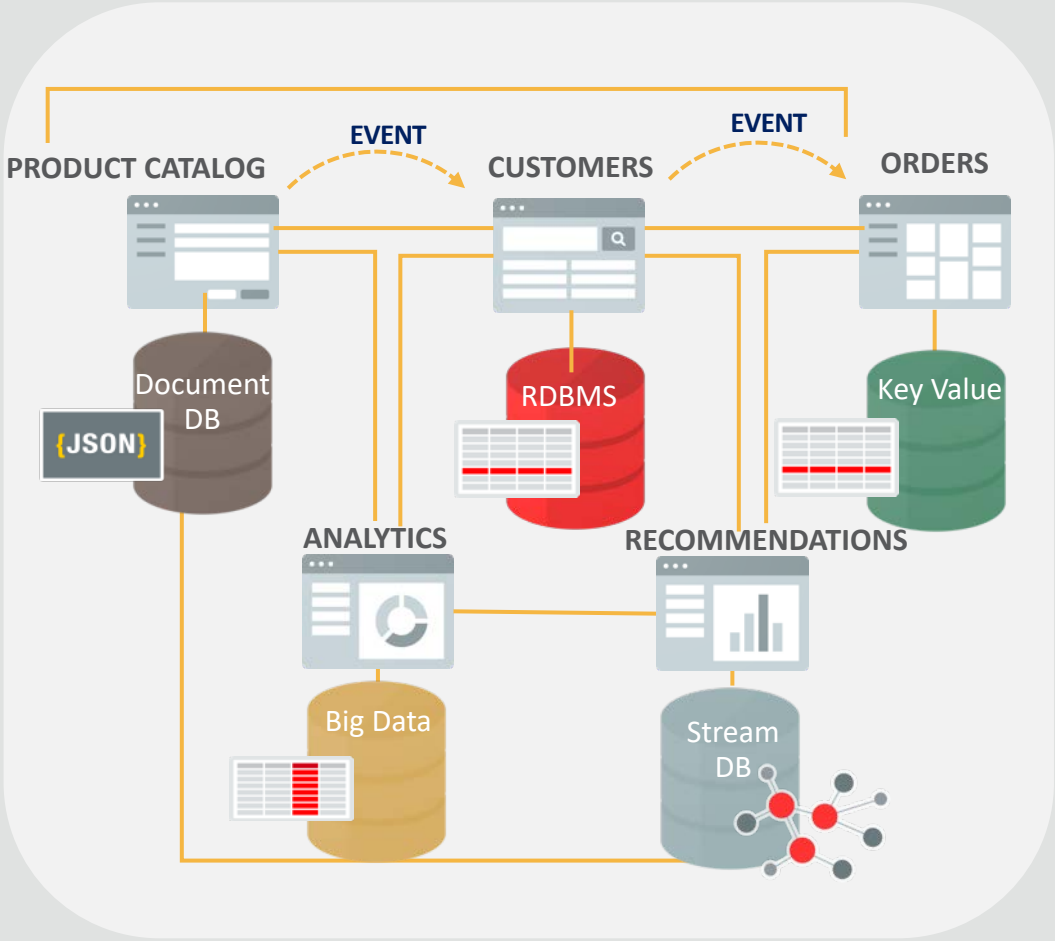
- Myth #1: Microservices **need** separate specialized databases
- **Fact:** They may need separate algorithms and data models
- Possible with single-purpose databases and a converged database



Single Purpose Databases: Missing Functionality

- Many of them lack:
  - Native data Integrity
  - Sophisticated Indexing
  - Full featured Transactions
  - Enterprise Security
  - Standard Declarative Languages

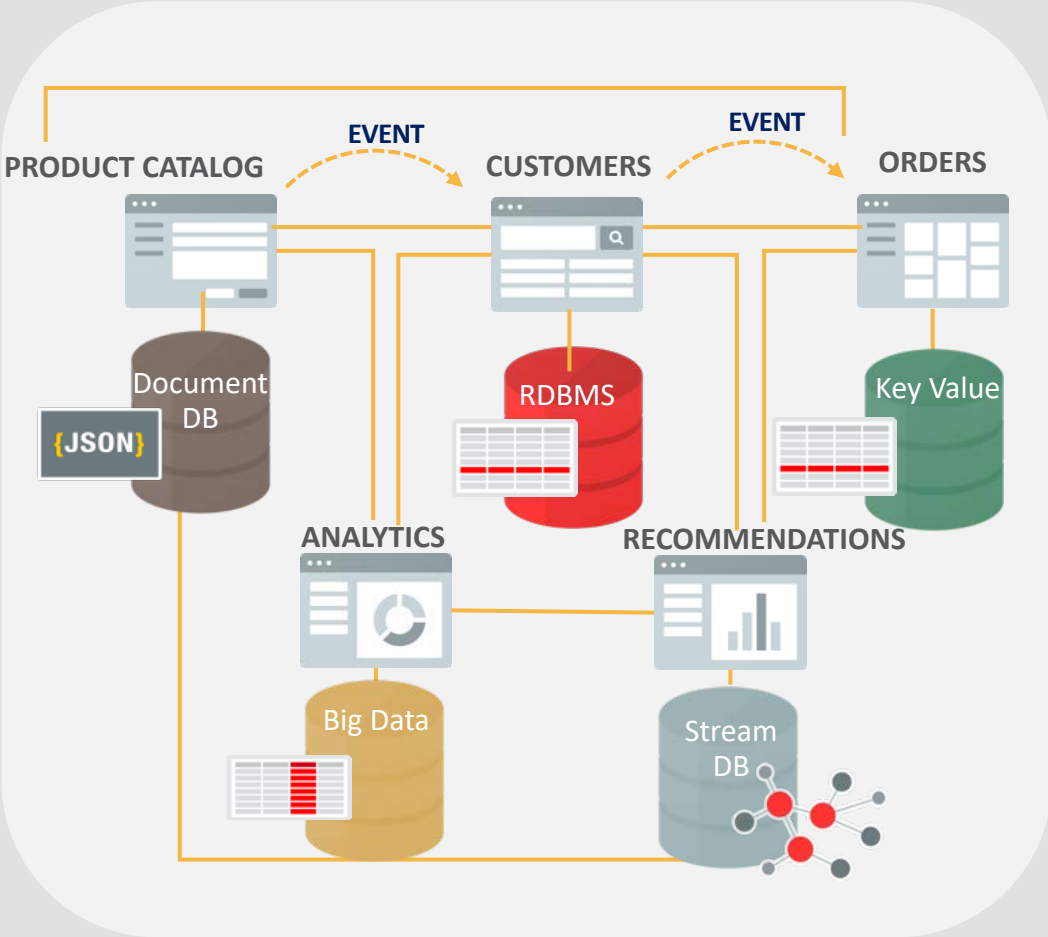
- Each is unique, need special skills



Single Purpose Databases: Infrastructural Downsides

Administration, Integration challenges:

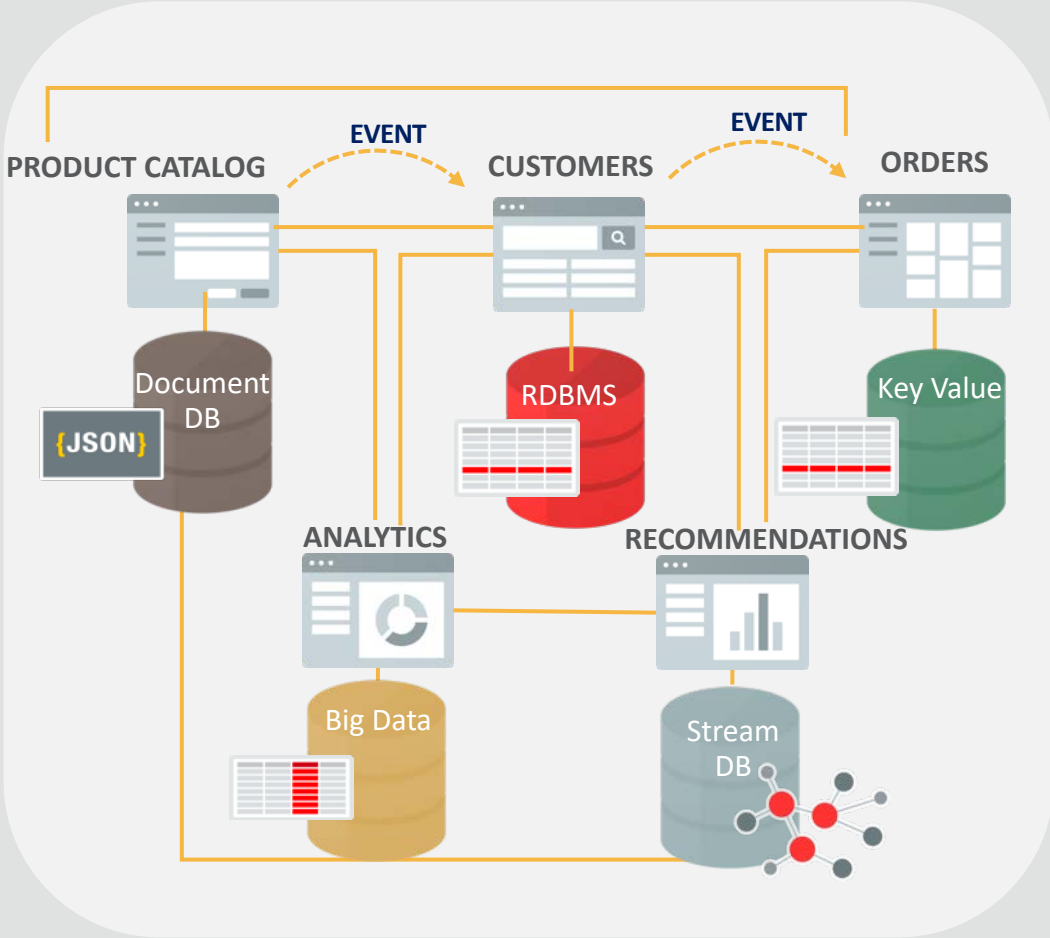
- Consistency of data
- Cross product compatibility
- Separate security patches
- Separate upgrade cycles
- Complex integration for Analytics
- Consistent backup of data is hard



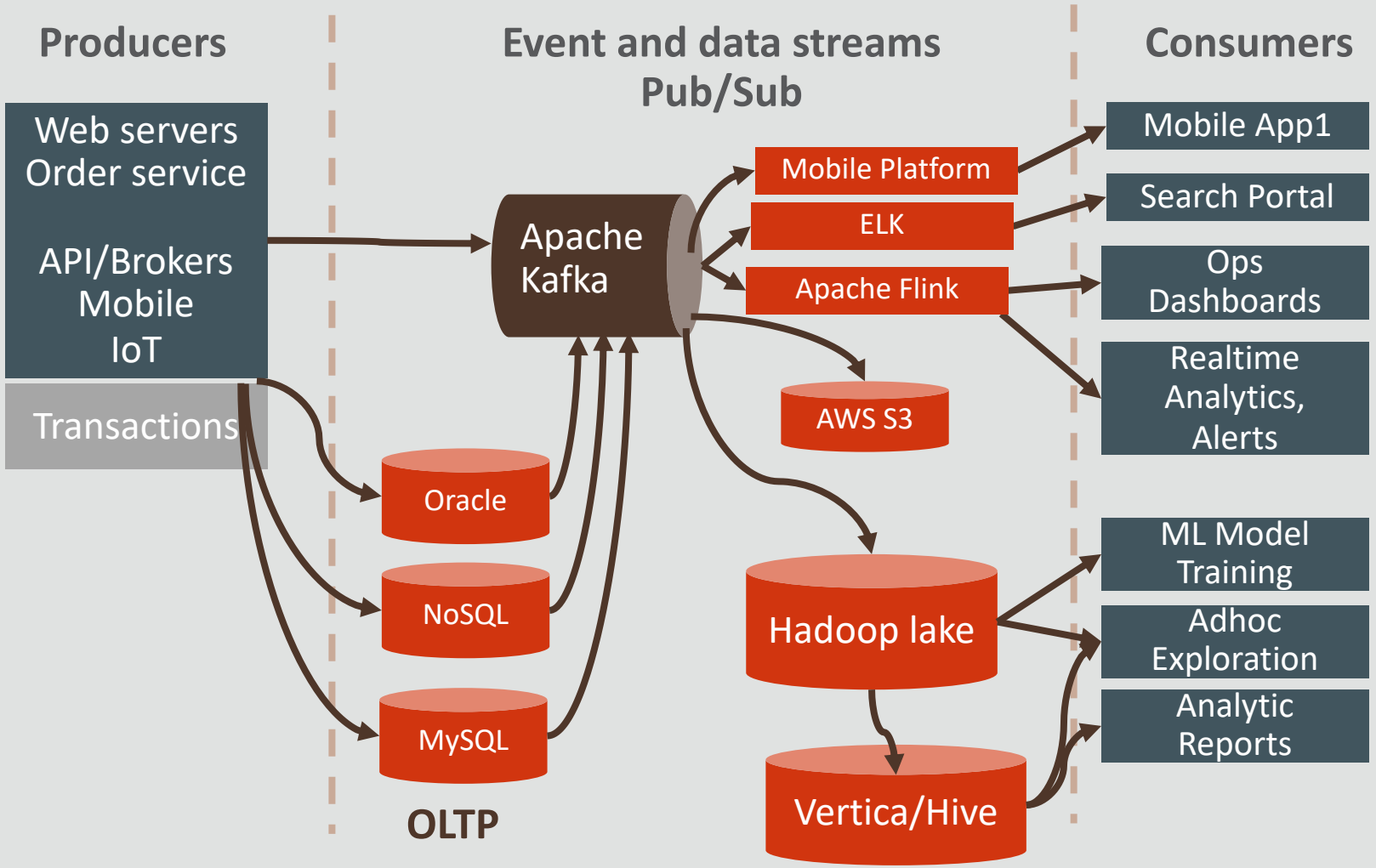
Single Purpose Databases: Macro-Complex Microservices

- Microservices must manage enormous amounts of application state

- Application is responsible for:
  - Atomicity of workflows
  - Failure handling
  - Consistent security policy



Real-World Macro-Complexity



Macro-Complexity

- Multiple technologies
- Multiple data stores
- Data copied multiple times to do analytics
- Compromises security
- Compromises data consistency
- Complex to maintain
- Need highly skilled developers to build & keep running

Can a Converged Database Ever Scale Like a Single-Purpose Database?

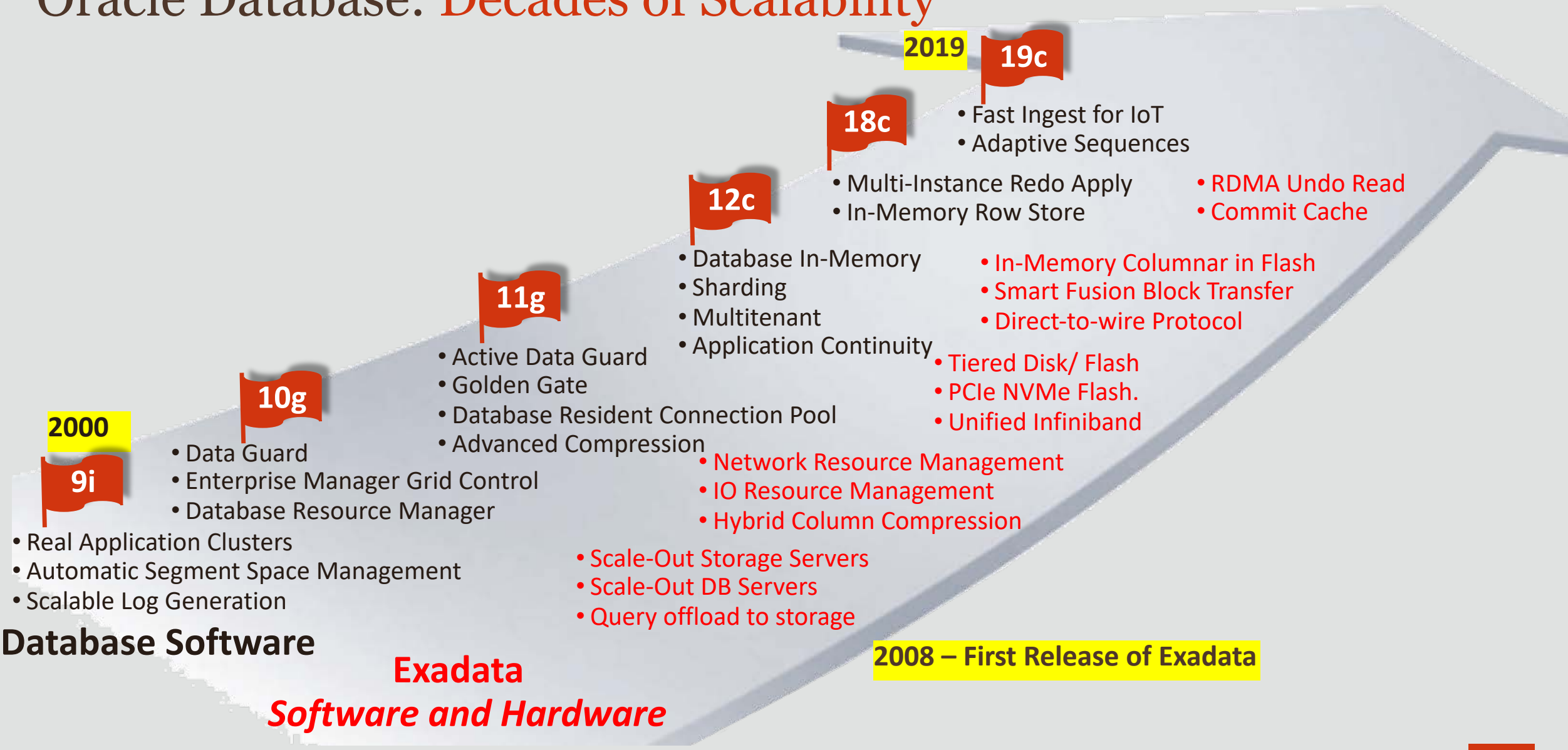
YES

Scalability of Oracle Database

Myth #2: An RDBMS is too complex to scale effectively

- Oracle Database has made major scalability advances over decades
- Every Oracle Database release has tackled difficult scaling problems
- Major architectural features for Scale-Up, Scale-Out, Fault tolerance

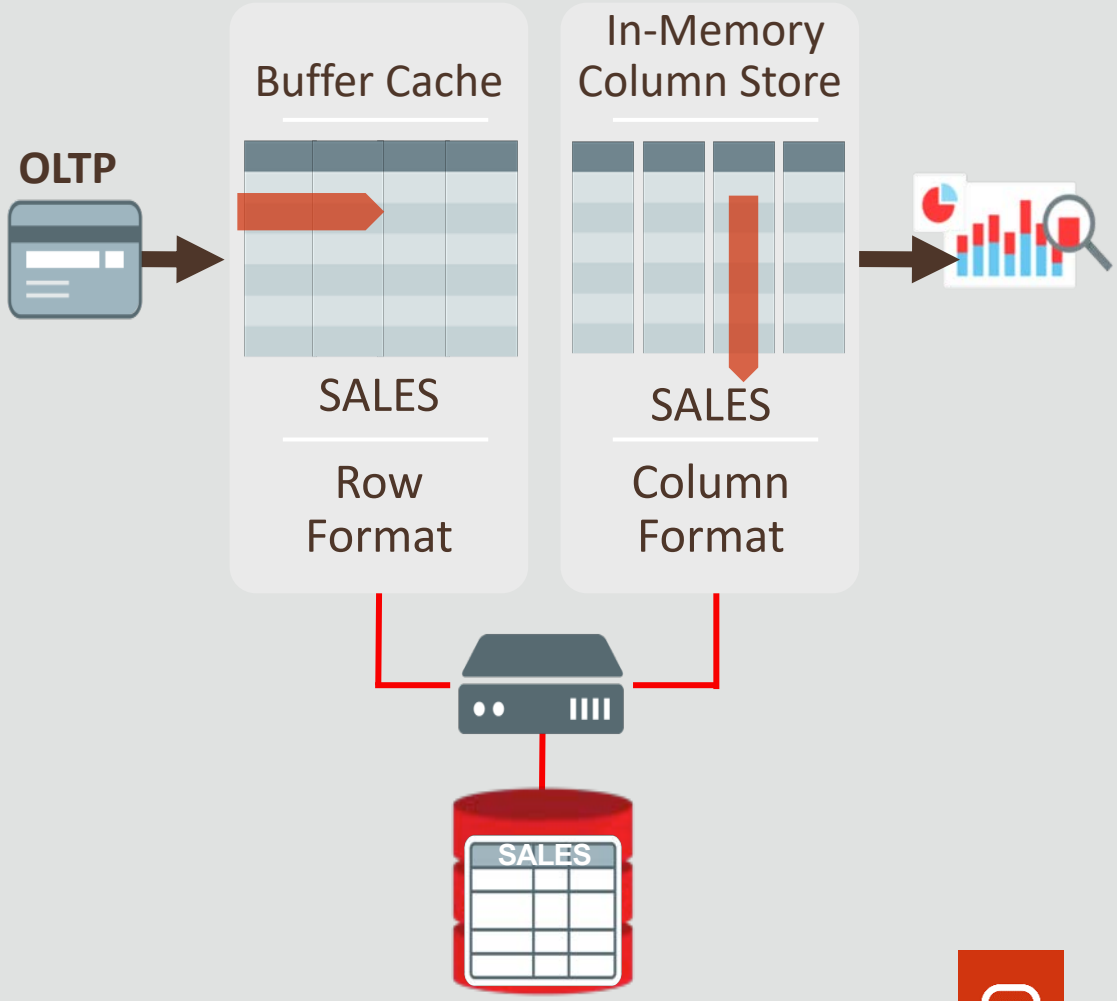
Oracle Database: Decades of Scalability



Database In-Memory: Converged Analytics and OLTP

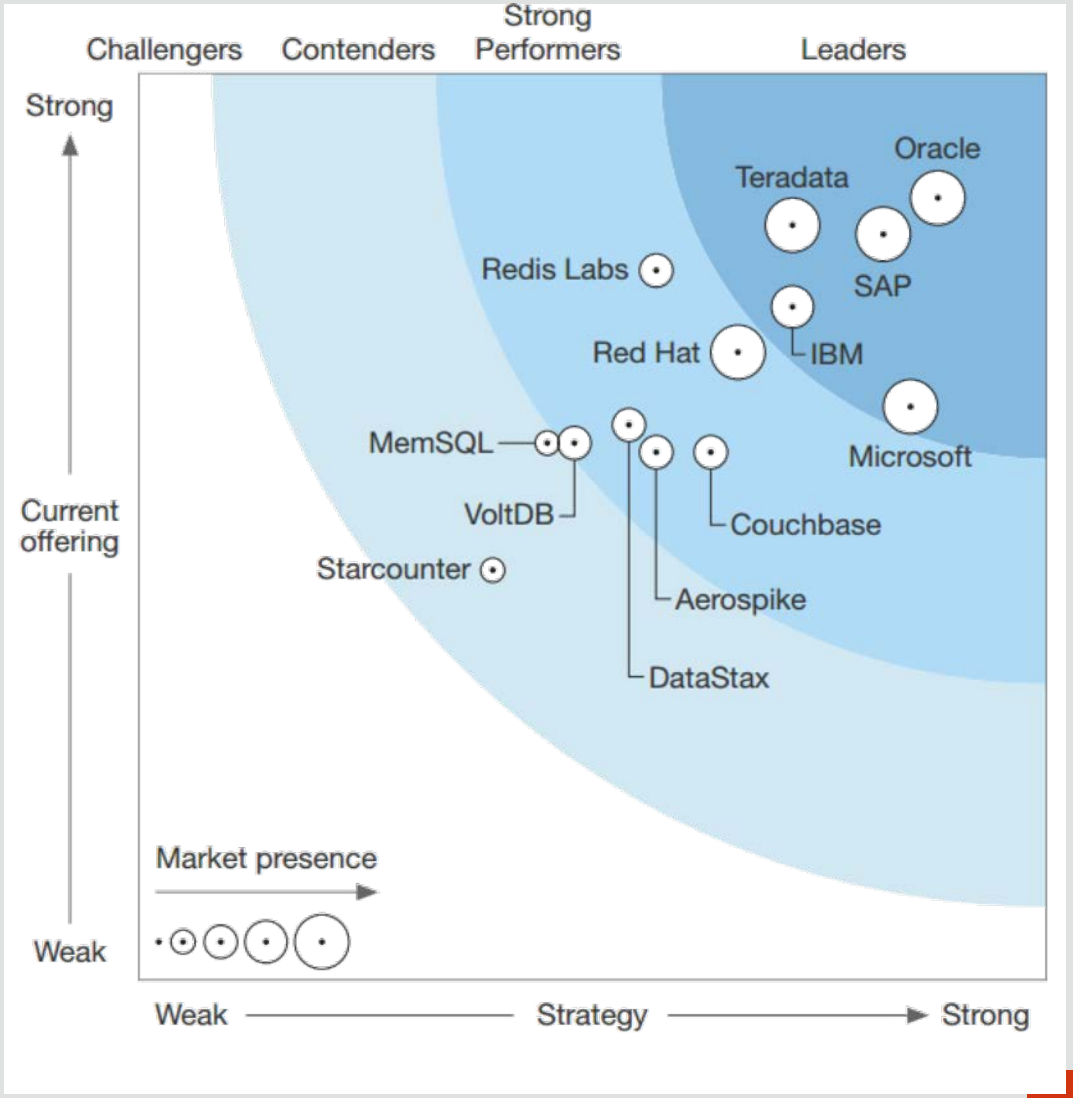
- Fast OLTP and Real-Time Analytics in the same converged database

- Industry-first dual format architecture:
  - Row format ideal for OLTP, column format is ideal for Analytics
  - Same table in row and column format simultaneously
  - No data movement needed
  - Analytics against against real-time data



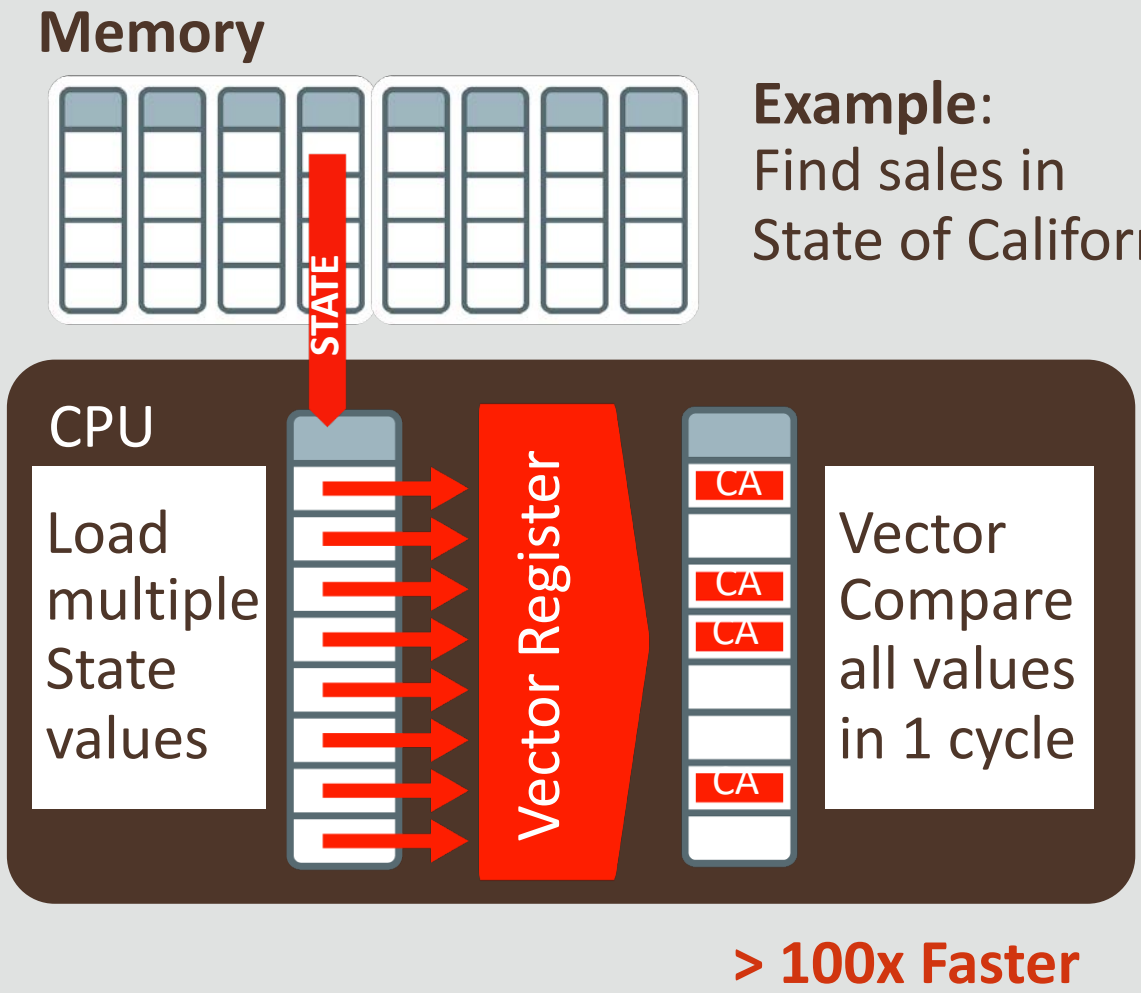
The Forrester Wave™: In-Memory Databases, Q1 2017

Oracle In-Memory Databases Scored Highest by Forrester on both Current Offering and Strategy



http://www.oracle.com/us/corporate/analystreports/forrester-imdb-wave-2017-3616348.pdf

In-Memory Enables SIMD Vector Processing

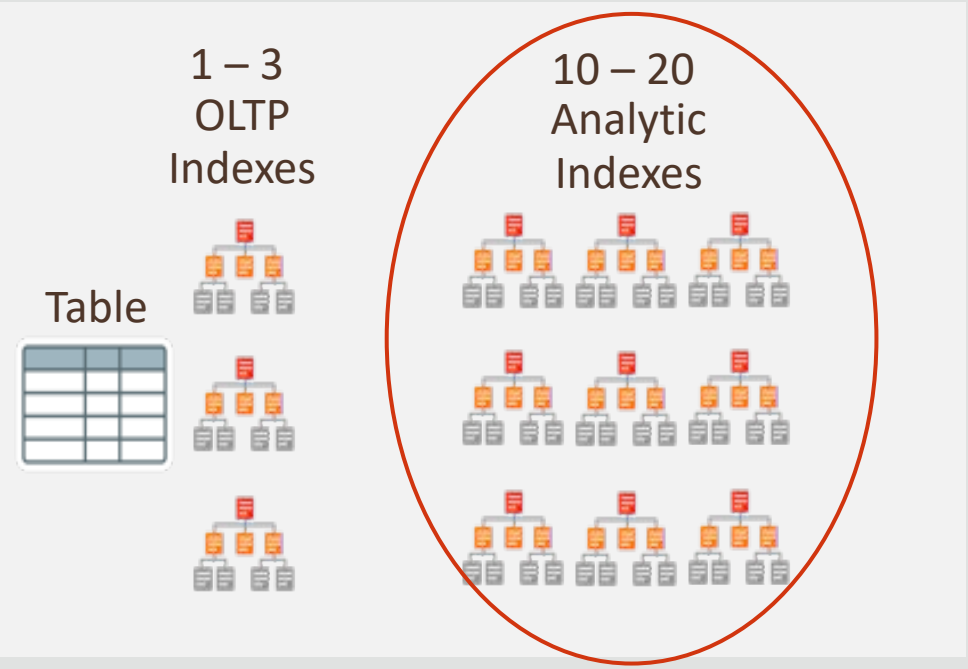


Example:
Find sales in
State of California

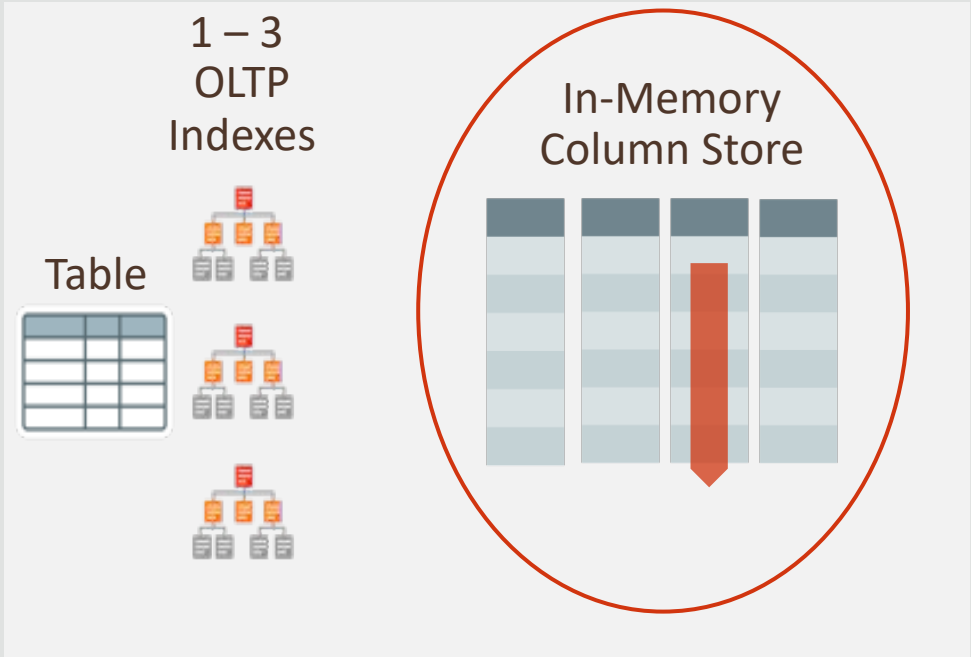
> 100x Faster

- Column format benefit: Need to access only needed columns
- Process multiple values with a single SIMD Vector Instruction
- Billions of rows/sec scan rate per CPU core
  - Row format is millions/sec

Accelerates Converged Workloads



REPLACE

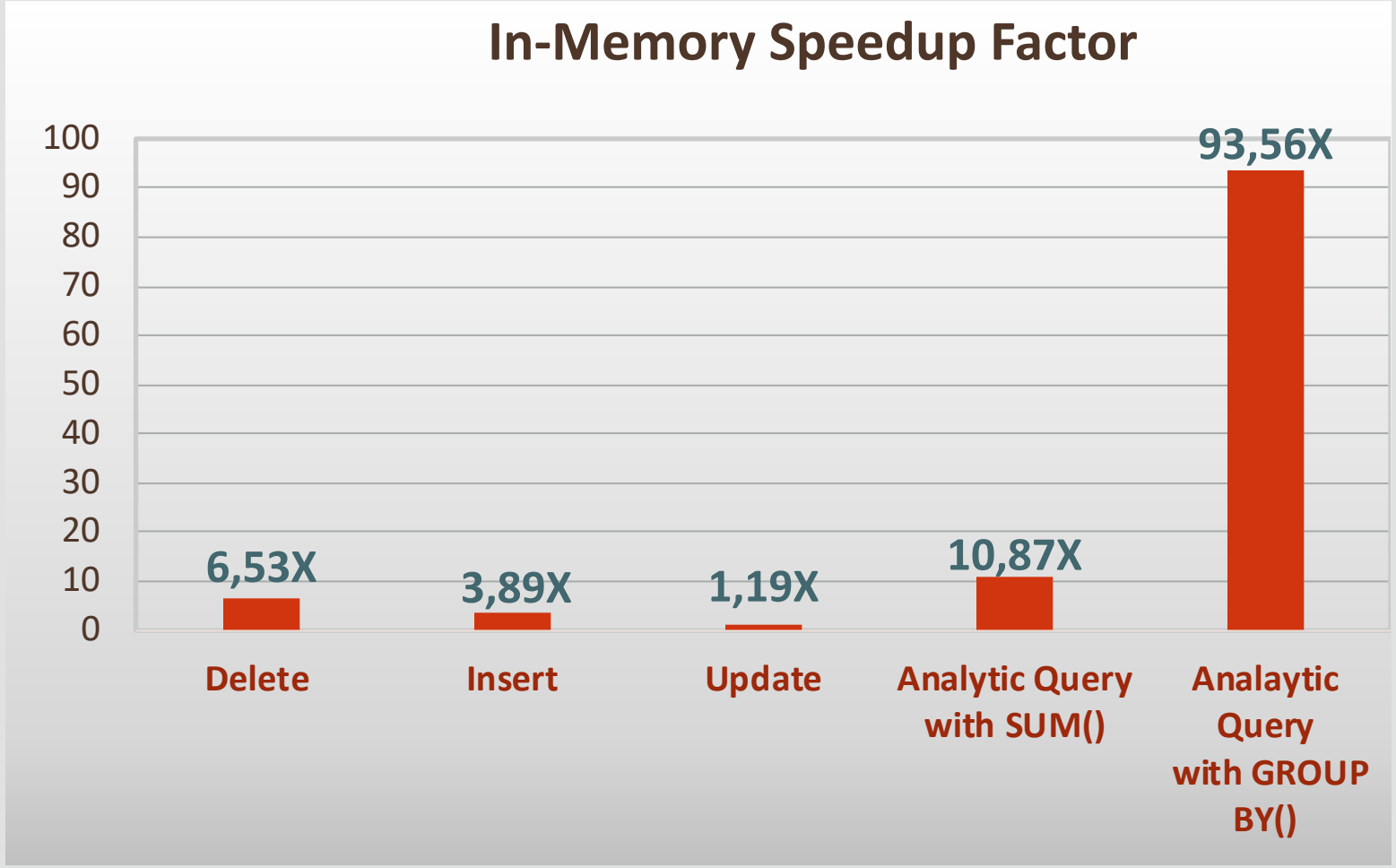


- Inserting one row into a table requires updating 10-20 analytic indexes: Slow!
- Fast analytics only on indexed columns
- Analytic indexes increase database size

- Column Store not persistent so updates are: Fast!
- Fast analytics on any columns
- No analytic indexes: Reduces database size

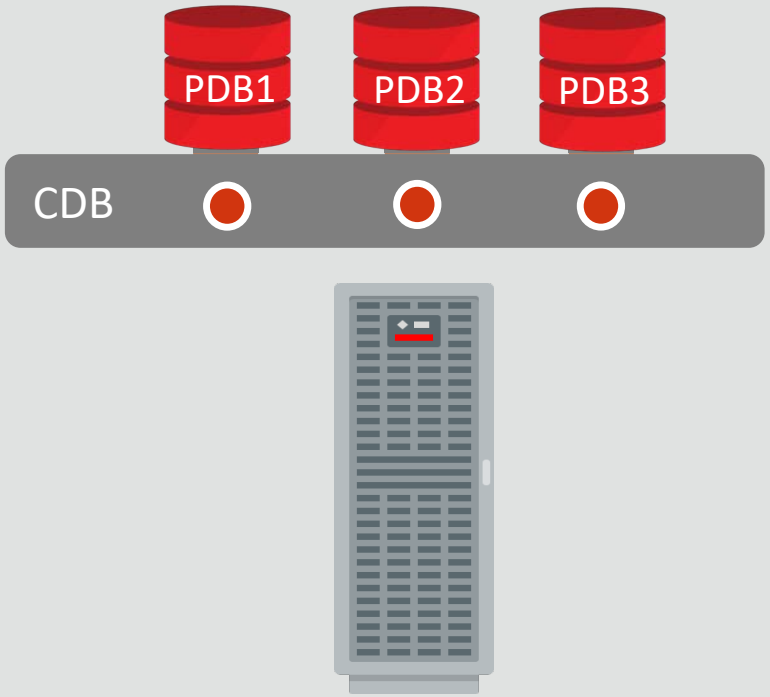
Example Speedup from Database In-Memory

- Synthetic converged workload consisting of DML and Analytic Queries
- Comparison between conventional indexes and Database In-Memory
  - Faster DMLs due to reduced index maintenance
  - Inmemory column store also speeds up Analytics

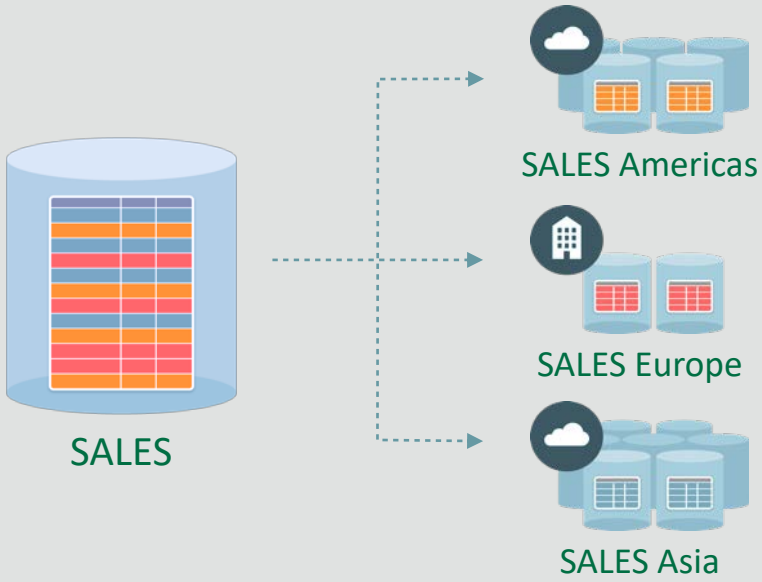


Multitenant: Efficient Cloud Scale Deployments

- Many tenant *Pluggable Databases* (PDBs) share a single Container Database (CDB)
- Completely application transparent
- Efficient resource sharing across tenants
- Manage many tenants as one



Sharding: A Hyperscale Database Architecture



One giant database divided into several smaller databases (shards)

- Divide massive databases into independent units
- Native SQL for up to 1000 Shards
  - Send SQL to correct shard based on shard-key
  - Cross shard queries
- Online reorganization of shards
- Sharding makes Oracle Database **Hyperscale**
  - Linear scalability of capacity, throughput, user population
  - Improves availability since shards are fault isolated
  - Manage of 100s of shards as a single logical database

Real-World Hyperscale Workloads

- Korea's number one mobile operator
- 65 billion transactions per day
- 18TB of data per day
- All data processing occurs on Oracle Database running on Exadata

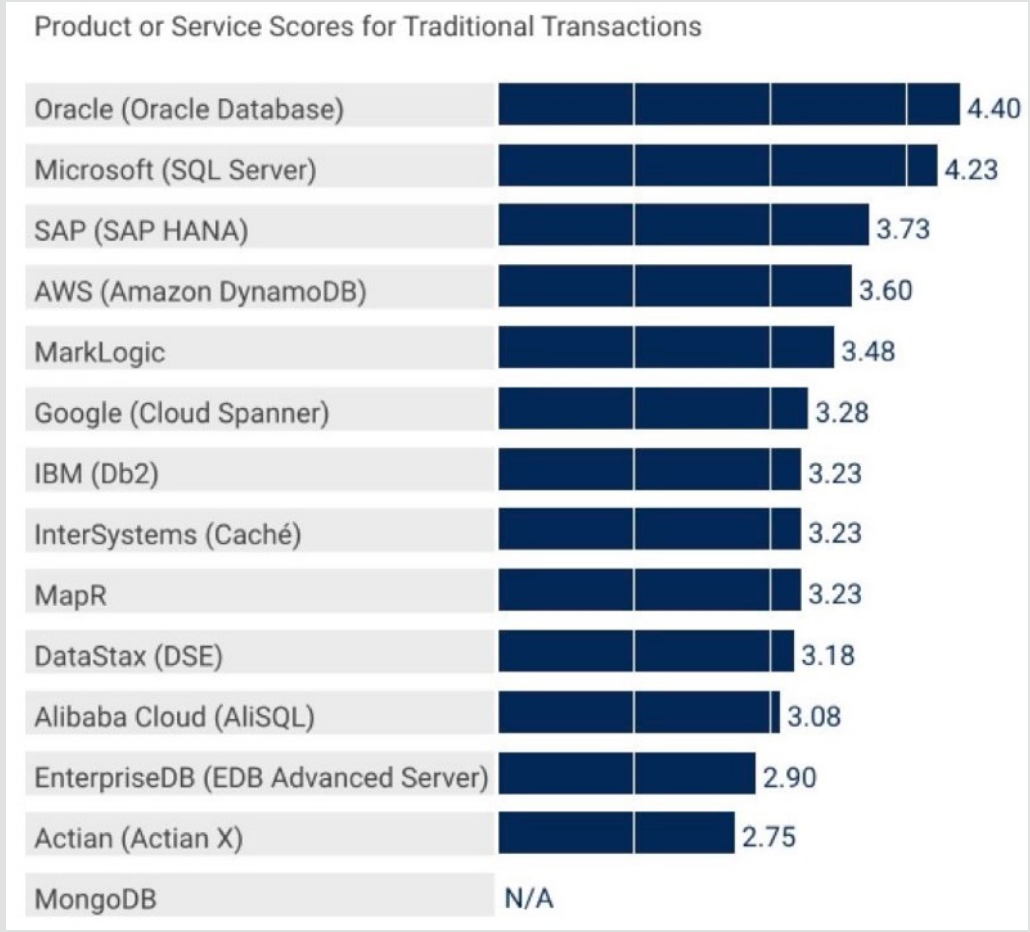
- One of world's largest law enforcement orgs
- ~3 billion transactions per day
- ~32 billion queries per day
- Database is over 1PB
- Deployed on Oracle Database on Exadata

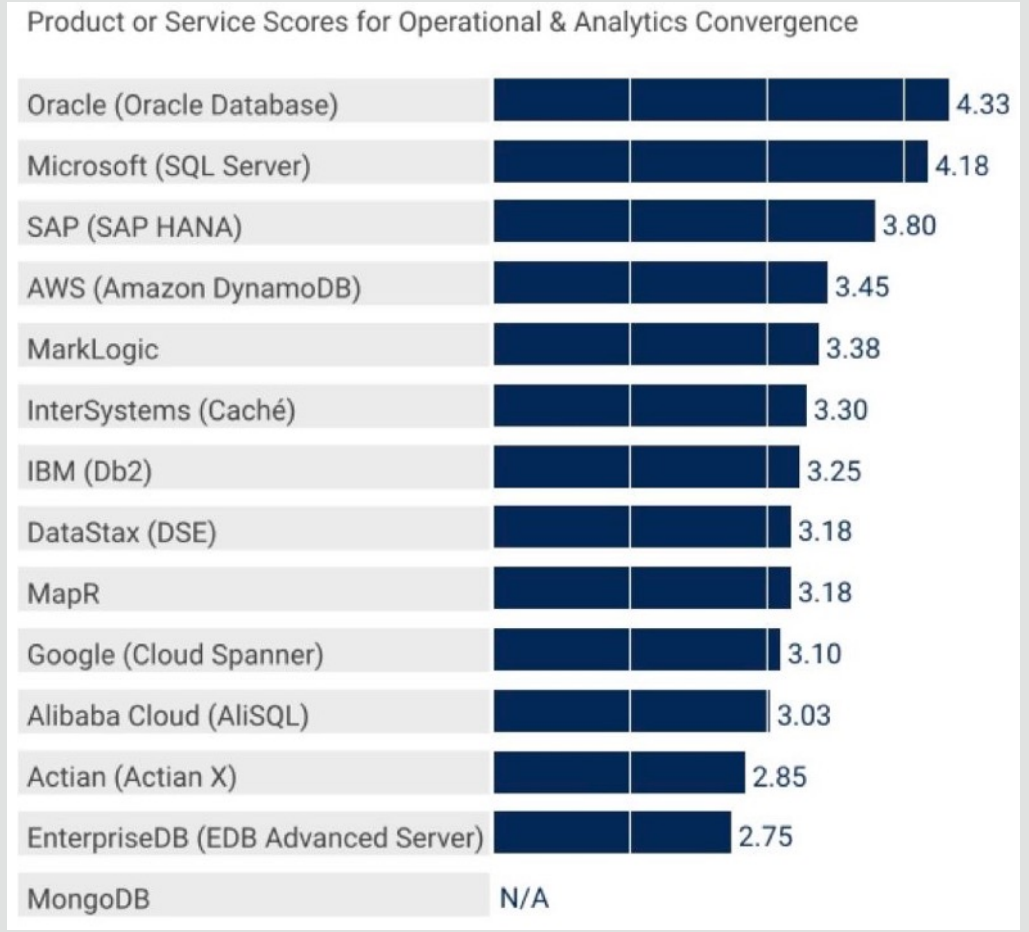
- World's largest stock exchange
- ~1 billion database transactions per day
  - 180,000 messages/sec
- ~ 15 TB of data per day
- All data captured and processed in an Oracle Database on Exadata

Can a Converged Database Deliver the Functionality of Many Specialty Databases?

YES

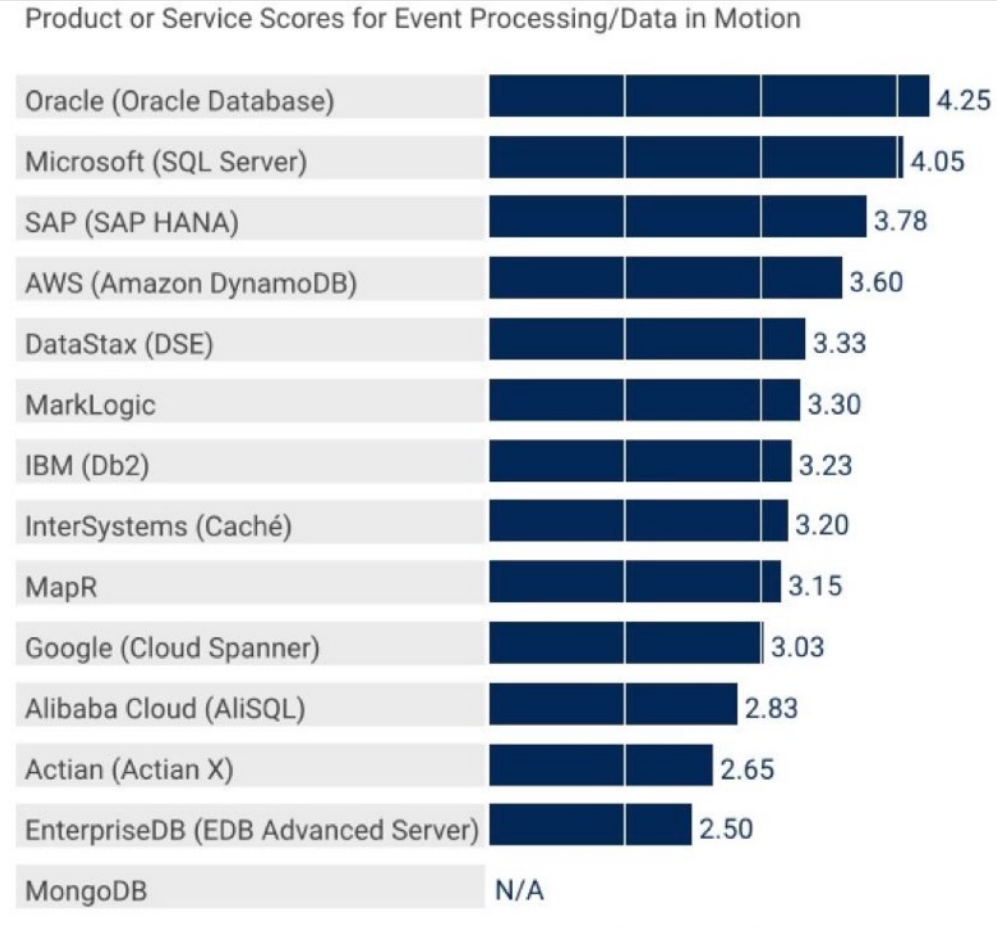
Oracle Database Ranks First for all OPDBMS Use Cases

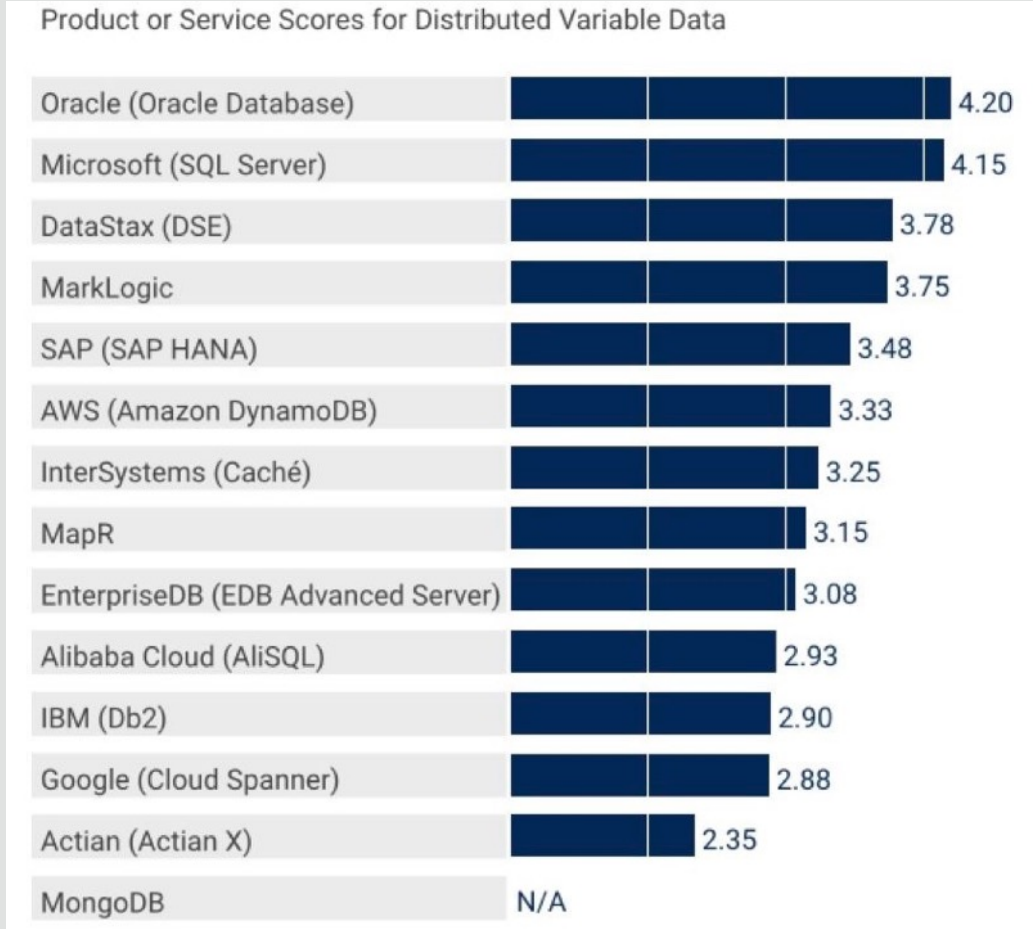




Gartner: *Critical Capabilities for Operational Database Management Systems* Donald Feinberg, Merv Adrian and Nick Heudecker, October 23, 2018
These graphics were published by Gartner Inc. as part of a larger research document and should be evaluated in the context of the entire document. The Gartner document is available upon request from Oracle. Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only thosee vendors with the highest ratings of other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

Oracle Database Ranks First for all OPDBMS Use Cases





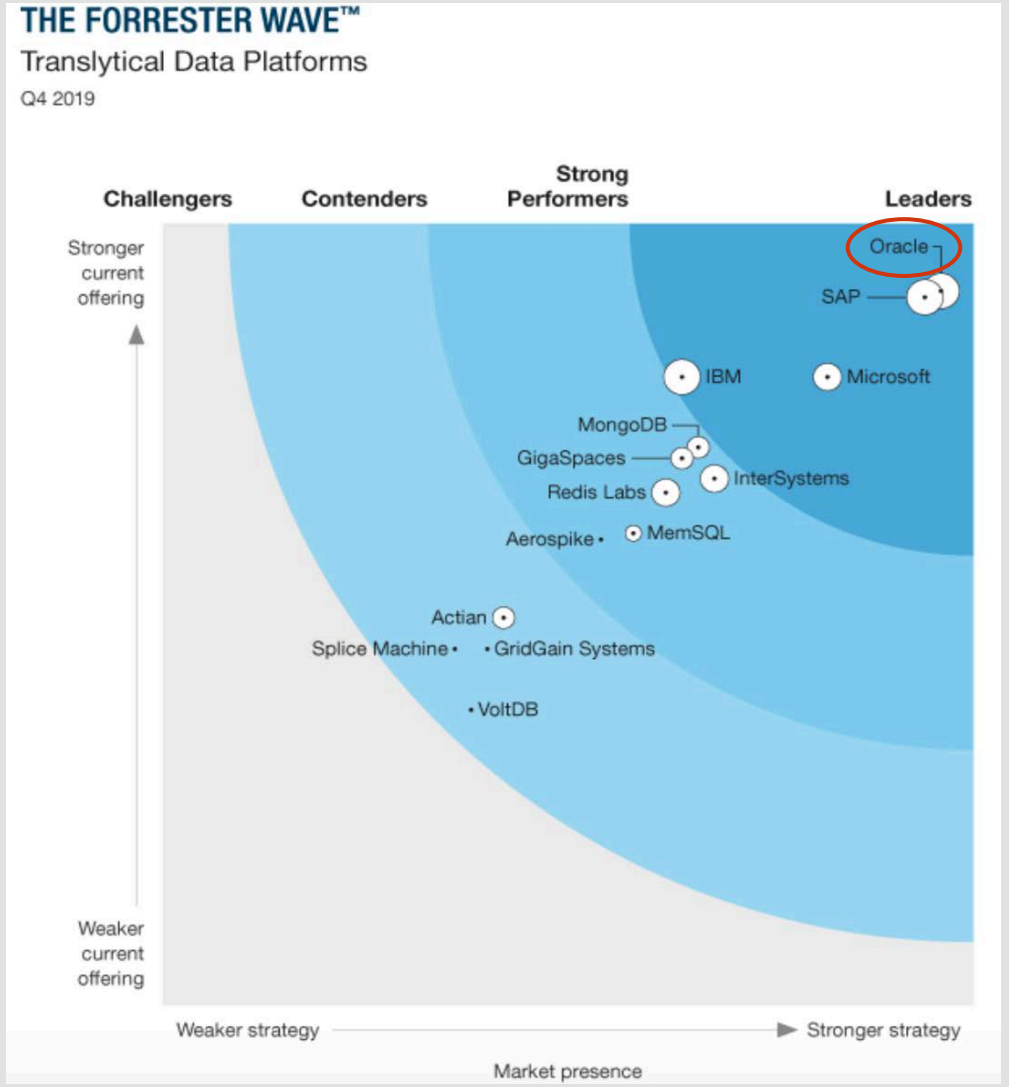
Gartner: *Critical Capabilities for Operational Database Management Systems* Donald Feinberg, Merv Adrian and Nick Heudecker, October 23, 2018
These graphics were published by Gartner Inc. as part of a larger research document and should be evaluated in the context of the entire document. The Gartner document is available upon request from Oracle. Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only thosee vendors with the highest ratings of other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

Forrester Wave
Translytical Data Platforms, Q4 2019

Oracle Position: Leader

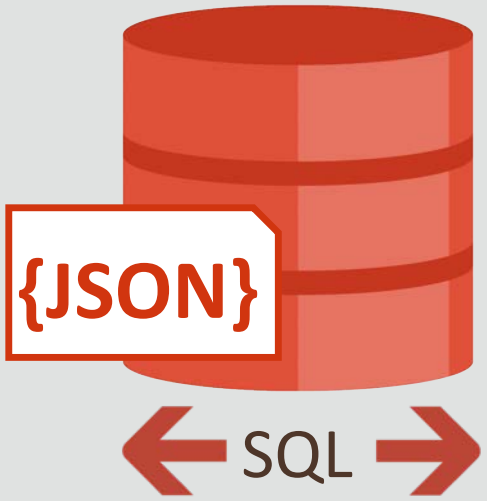
Oracle ranked highest on both Axis

- "Unlike other vendors, Oracle uses a dual-format database (row and columns for the same table) to deliver optimal translytical performance."
- "Customers like Oracle's capability to support many workloads including OLTP, IoT, microservices, multimodel, data science, AI/ML, spatial, graph, and analytics"
- "Existing Oracle applications do not require any changes to the application in order to leverage Oracle Database In-Memory"
- "Customers like the platform's ease of use, ease of expanding existing Oracle applications to take advantage of translytics, general data security capabilities, and technical support, as well as the Cloud At Customer offering"



Published: October23, 2019
Analysts: Noel Yuhanna, Mike Gualtieri

JSON in Oracle Surpasses Document Specialists

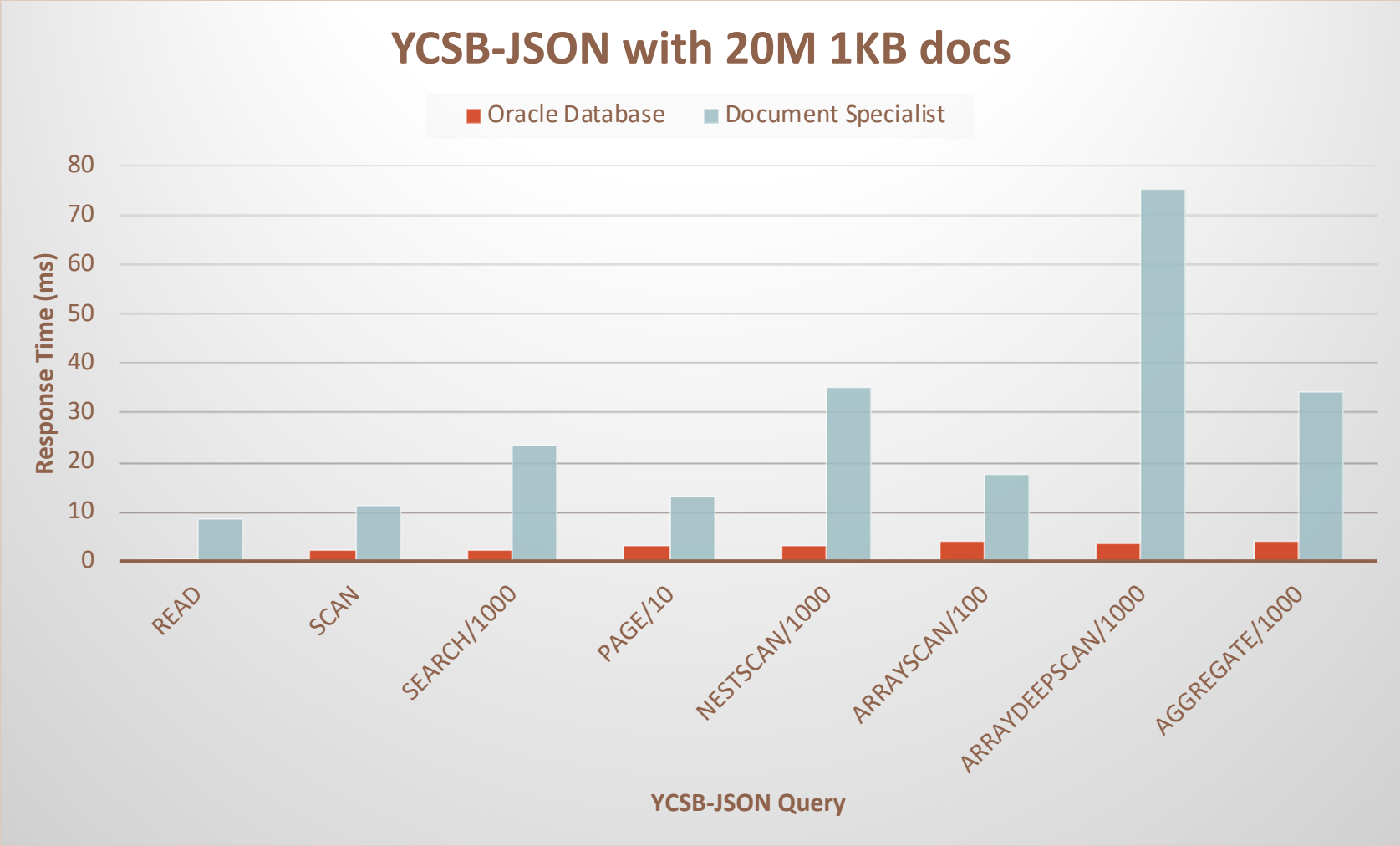


- Developer friendly schema-less data storage
- Simple SQL Syntax with JSON paths

```
select c.json_doc.name.last, c.json_doc.address.city
  from customers c;
```

- All existing Oracle features work with JSON Data
- Index any JSON column using Functional Indexes
- In-Memory Columnar Processing also allows 30-60x faster analytics on JSON data

JSON in Oracle Surpasses Document Specialists:



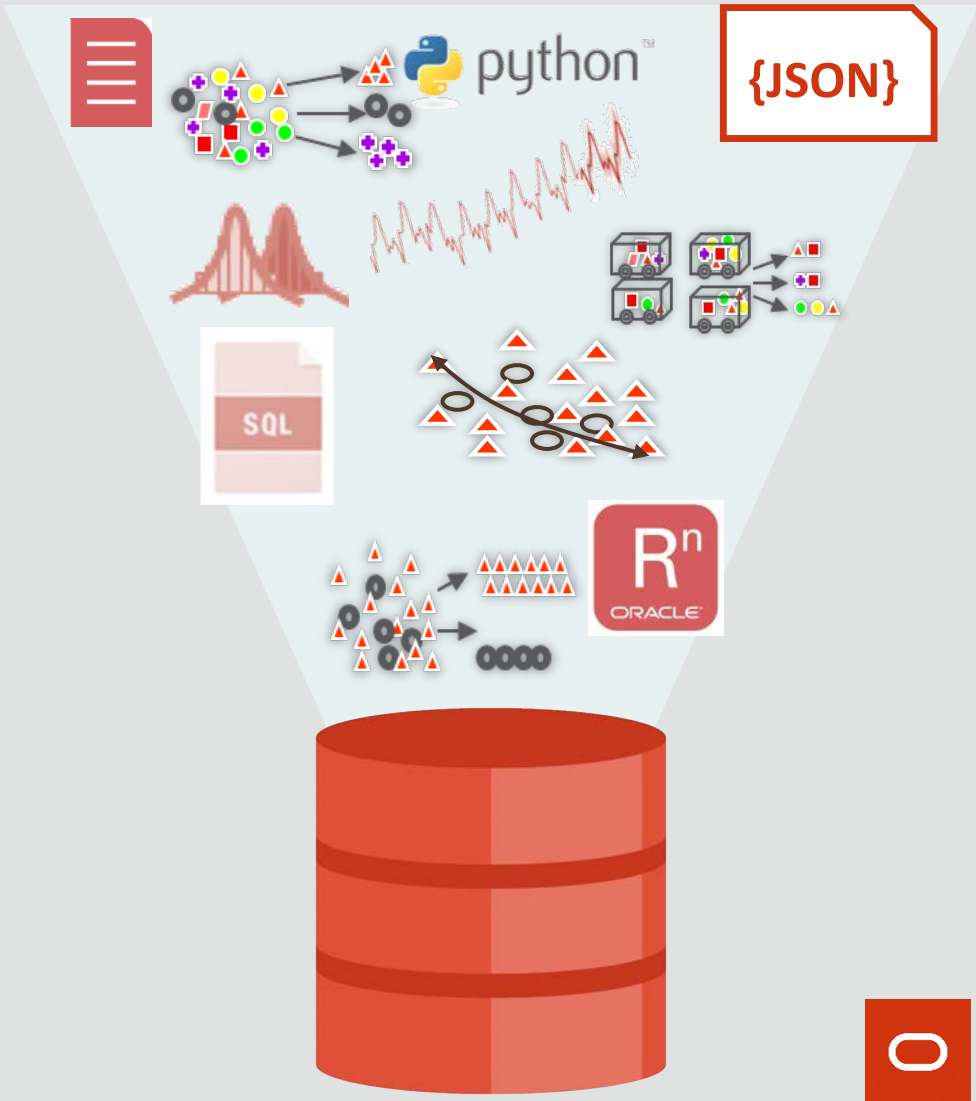
Methodology: Single client executing queries repeatedly with 1s sleep time. (20M x 1KB documents)

Source of YCSB-JSON queries:
https://blog.couchbase.com/ycsb-json-benchmarking-json-databases-by-extending-ycsb

Test Hardware	
Machine	½ X7 cell (each cell is 2x Xeon Silver 4114 CPU @ 2.20GHz, 10 cores, 2 th/core =40 cpus, 1024k L2, 13M L3, 187Gb RAM)
CPUs	**16 cores**
Memory	**115GB**

Machine Learning in Oracle Surpasses AI Specialists

- Many algorithms and models with Oracle Advanced Analytics option
- Machine learning models and algorithms run inside Oracle Database
  - Data stays in-place
  - Massively parallel execution
- Flexible model building
  - SQL, R or Python
  - Oracle Data Miner
  - Oracle AutoML



Oracle Big Data Spatial & Graph Surpass Specialists

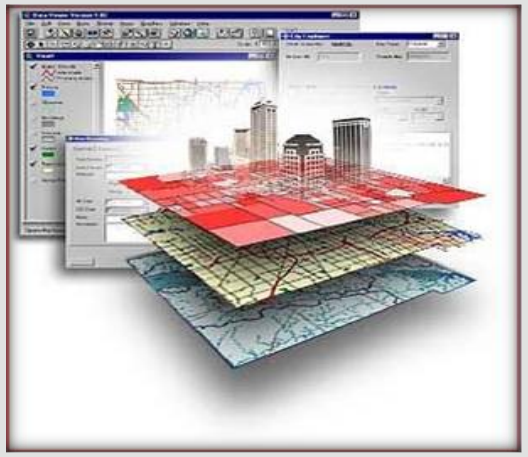
Spatial Analysis:

Location Data Enrichment

Proximity and containment analysis, Clustering

Spatial data preparation (Vector, Raster)

Interactive visualization



Property Graph Analysis:

Graph Database

In-memory Analysis Engine

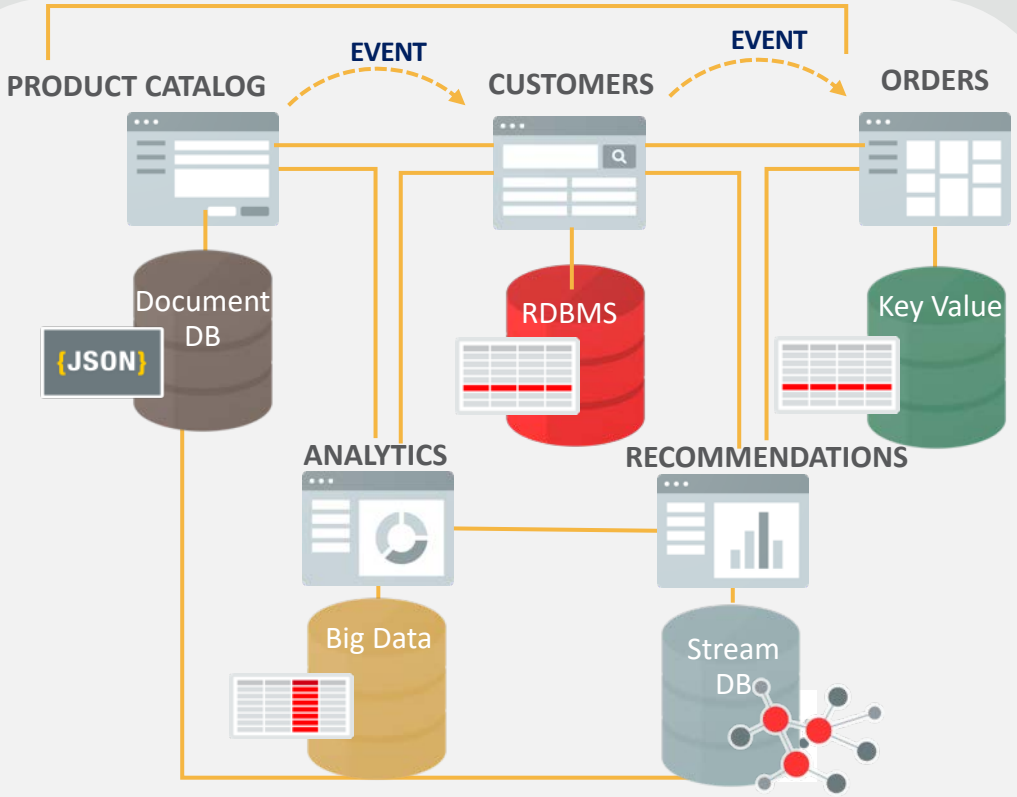
- Scalable Network Analysis Algorithms
- Declarative PGQL
- Developer APIs

Unique Advantage for a Converged Database: Synergy

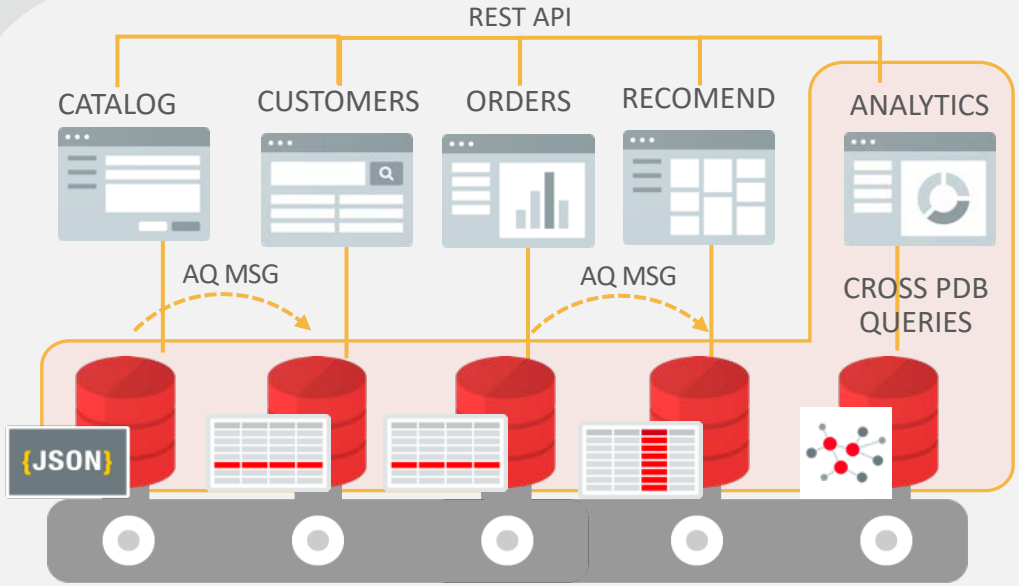
- Combined benefit exceeds the sum of individual benefits
- Convergence allows synergy across applications
- Data can be processed by many different algorithms, e.g.:
  - Relational analytics on document data
  - Fraud detection using Machine Learning inside a database transaction
  - Spatial search inside a relational application

Converged Database => Convergent Microservices

Single Purpose Databases:
DIVERGENT MICROSERVICES



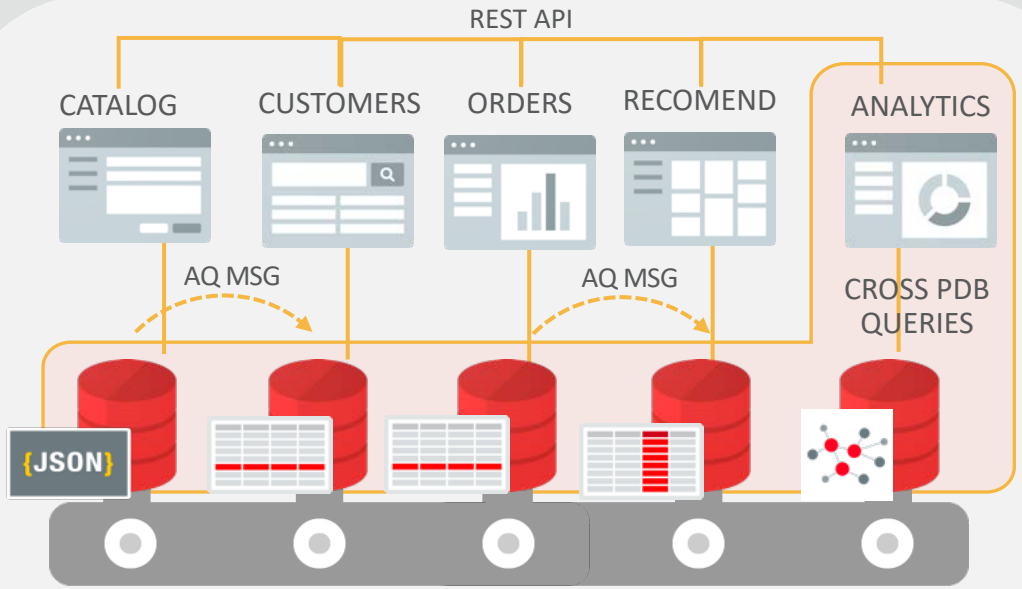
Converged Database:
CONVERGENT MICROSERVICES



Oracle Database for Convergent Microservices

- Microservices can be stateless
  - Push application state to database
  - Push application events into database
  - Database manages atomicity
  - Database manages recovery
- Consistent security policy is seamless and easy to implement

CONVERGENT MICROSERVICES



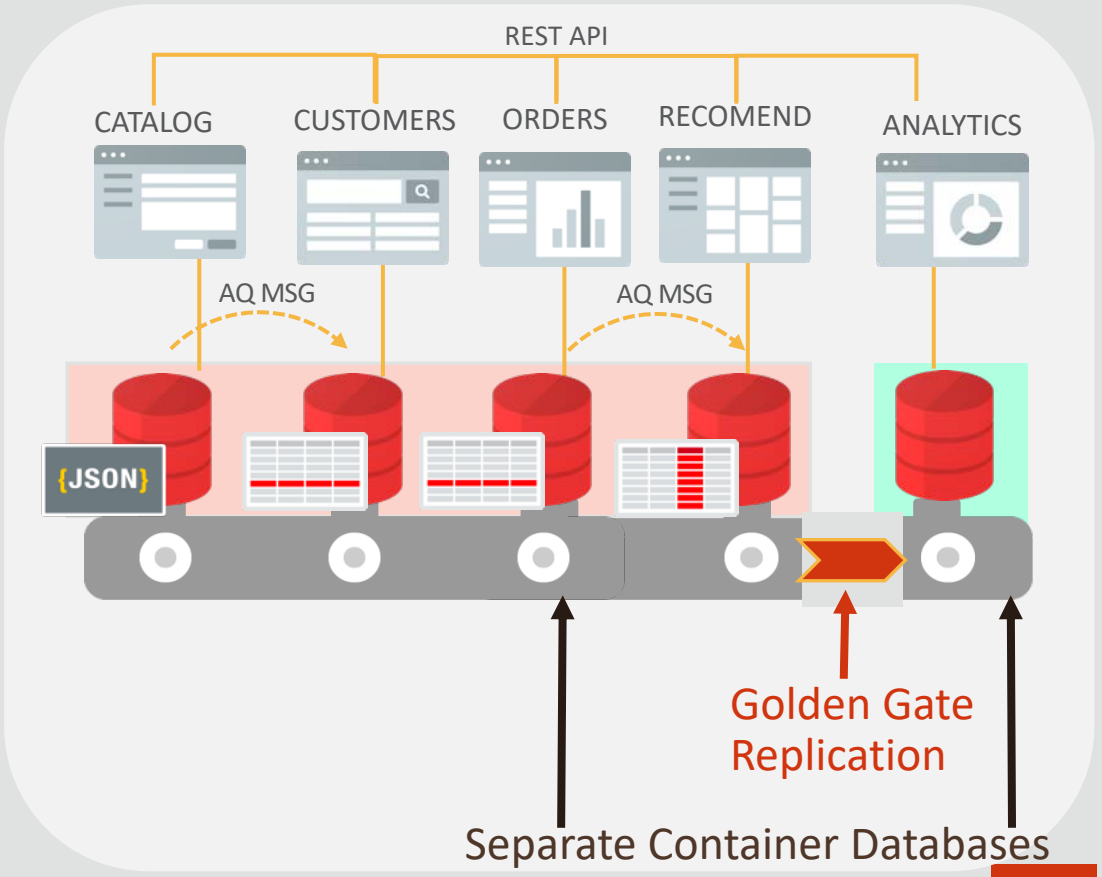
Oracle Database for Convergent Microservices

Myth #3: Converged database architecture requires all data to be in a single database

Can still use separate databases for isolation (e.g. OLTP and Analytics)

Multitenant allows **maximum flexibility** for microservices

- Can share a tenant database
- Or use separate tenant databases
- Or use separate container databases

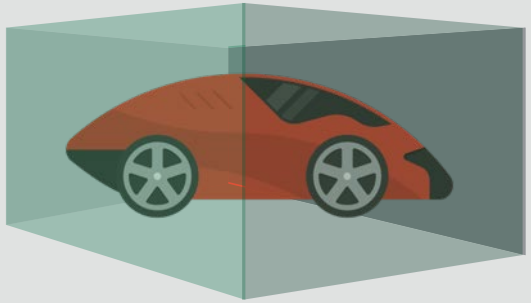


What about a Hyperscale Converged Database Augmented by Machine Learning?

Autonomous Database: Ultimate Converged Platform

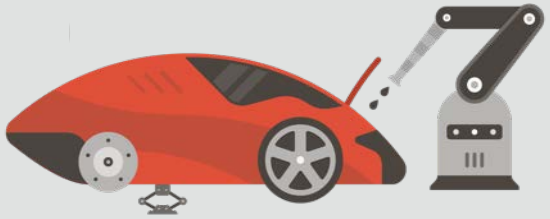
Self-Driving

- Scale-out database with fault-tolerance and DR
- Runs on enterprise-proven Exadata platform
- Full compatibility with existing enterprise databases



Self-Securing

- Automatically applies security updates online
- Secure configuration with full database encryption
- Sensitive data hidden from Oracle or customer admins



Self-Repairing

- Recovers automatically from any failure
- 99.995% uptime including maintenance
- Elastically scales compute or storage as needed

Autonomous Database Machine Learning

Diagnostics, recovery and optimizations for each layer of the deployment stack

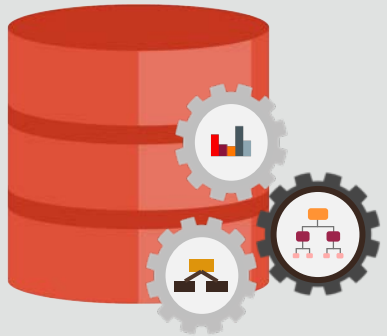
Database Infrastructure

Detection and recovery of failed/sick server, storage or switch/link

Database Operations

Hang Management
Anomaly Detection
Bug Identification and Prioritization

Workload Optimizations



Query Optimizer
Real-time statistics
Automatic Indexing

Autonomous Database

Thousands of New Customers

ATP

DATA INTENSITY

10X faster *provisioning, 5X faster reporting with elastic scalability*

MESTEC
MANUFACTURING PERFORMANCE. REDEFINED.

600% *performance improvement on ½ CPUs with Azure Interconnect*

Henry Ford Health System

Provisioning time reduced from ***2 days to 3 minutes***

veritrans

Guaranteed performance, availability, and security of PoS system for 1,000 clients

ADW

SKY

90% reduction in time*-to-market and 90% less maintenance*

agea

50% cost reduction *versus Teradata on-prem, eliminated maintenance*

droptank

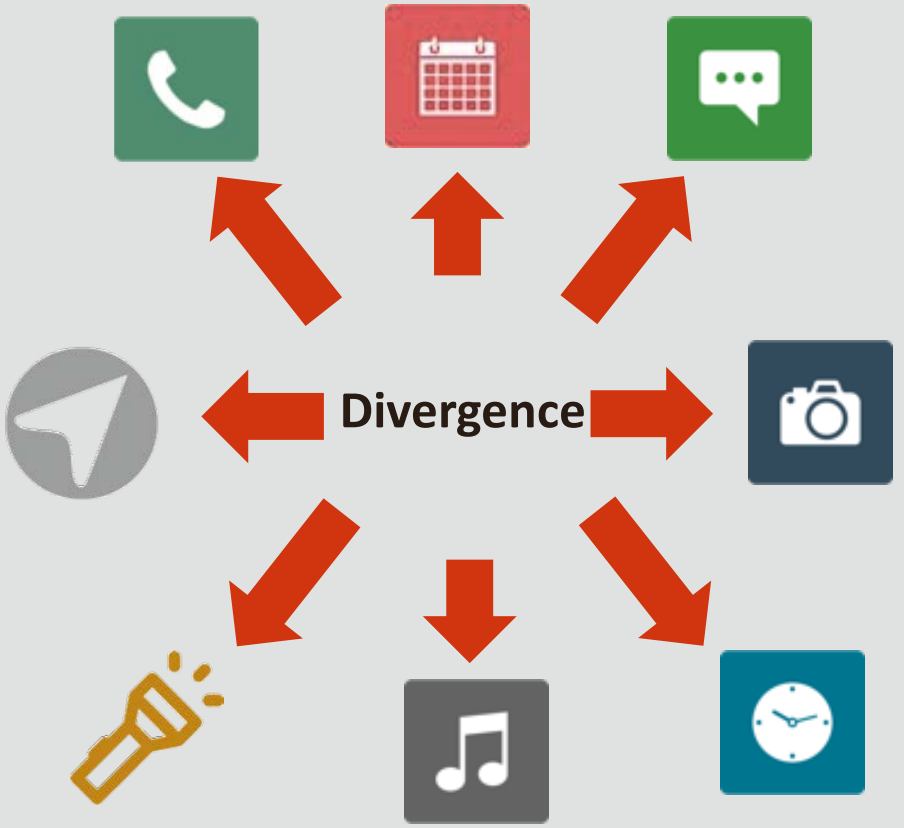
Scale 3,500 to 10,000 stores without hiring DBAs

SEUR
dpdgroup

75% cost reduction *in integration and flexibility to scale up 50%*

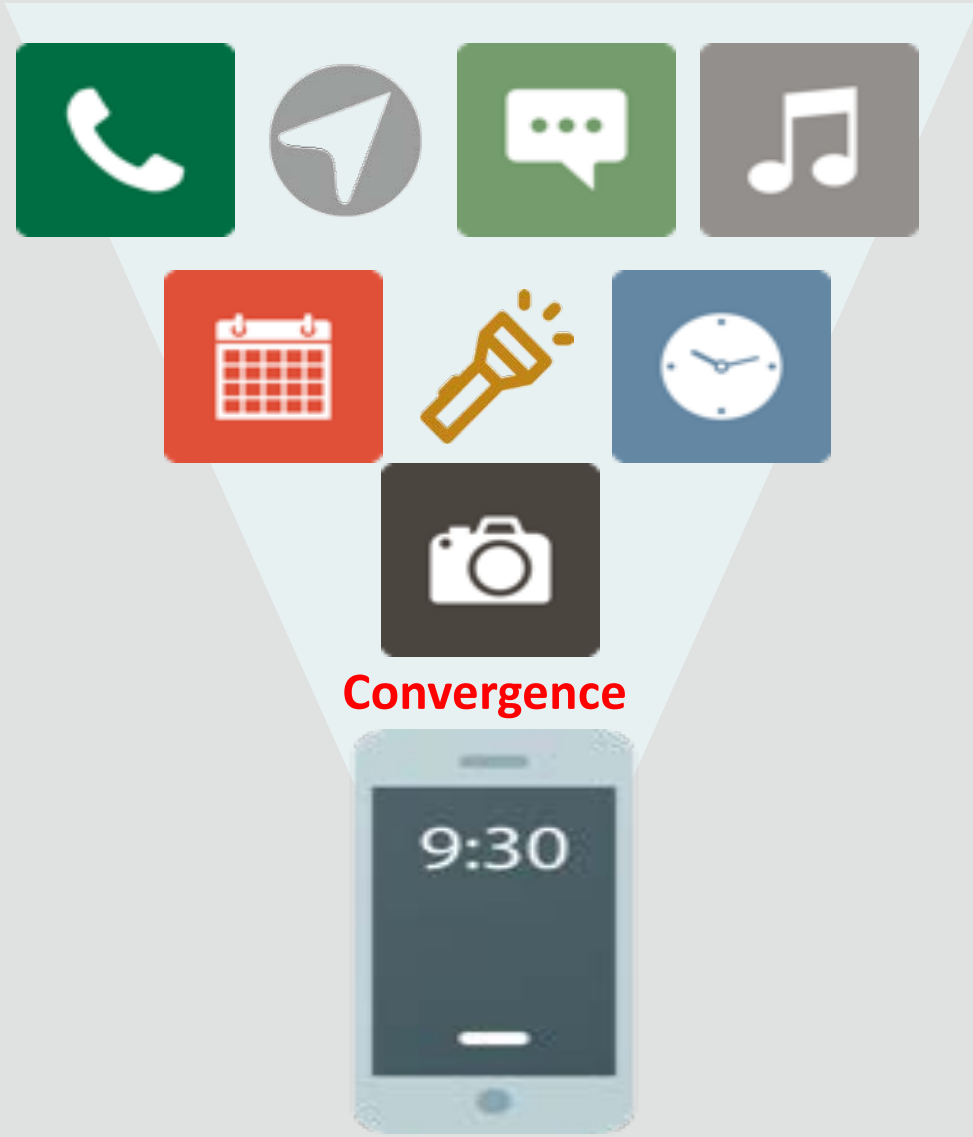
Summary: An Analogy

- In the 90s and early 2000s:
  - Many specialty single purpose devices
  - Calculators, phones, GPS, etc.



Summary: An Analogy

- In the 90s and early 2000s:
  - Many specialty single purpose devices
  - Calculators, phones, GPS, etc.
- The smartphone replaced most specialty devices and provided simplicity and synergy in a converged device
  - E.g. Calendar auto syncs from email
- Similarly: **A converged database vastly simplifies enterprise architecture**



ORACLE®