GOOGLE ALLOYDB VS. AMAZON AURORA VS. AZURE COSMOSDB

Taras Kloba, Kyrylo Prykhno, PGConf 2022

softserve


amazon
web services
Aurora
Microsoft
Azure
CosmosDB
Google
Cloud
AlloyDB

1 INTRO

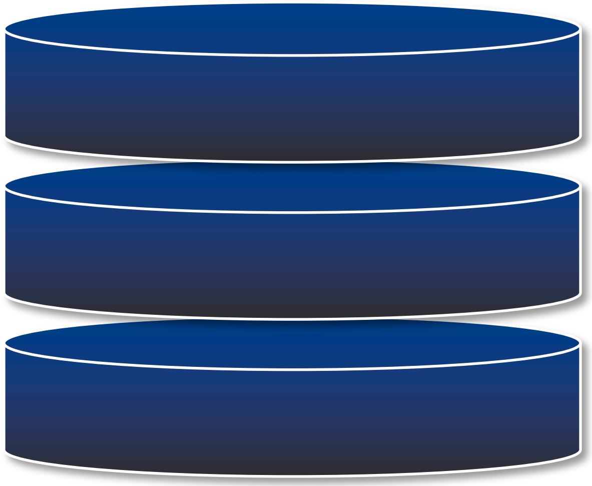
softserve

Our customers want to modernize from their legacy proprietary databases …

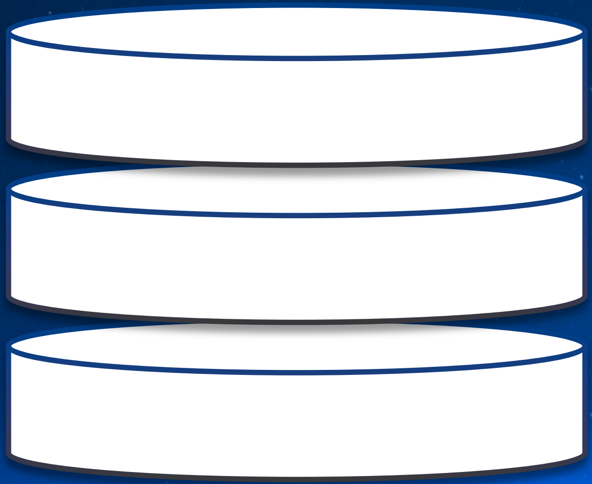
Oracle

SQL Server

DB2



to standardize on open source



softserve

Cloud offers organizations **agility**, **cost savings**, and **differentiated capabilities**.

That's why **75% of all databases** are expected to be in the cloud this year.

softserve

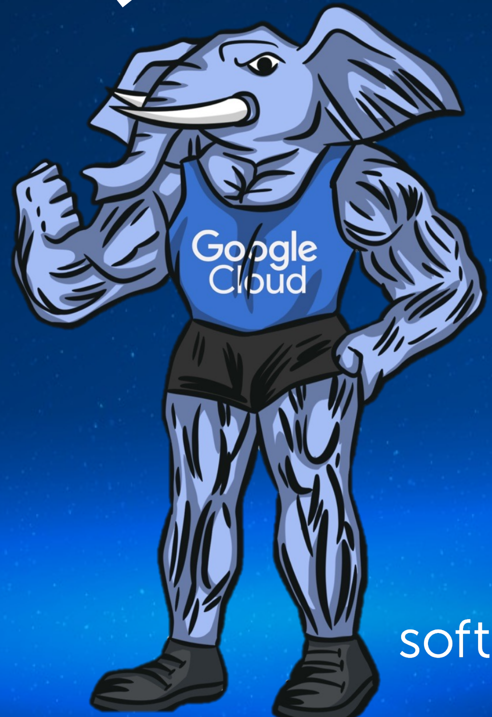
Source: Gartner: The Future of the DBMS Market

Hybrid
Transactional
Analytical
Proccesing

2

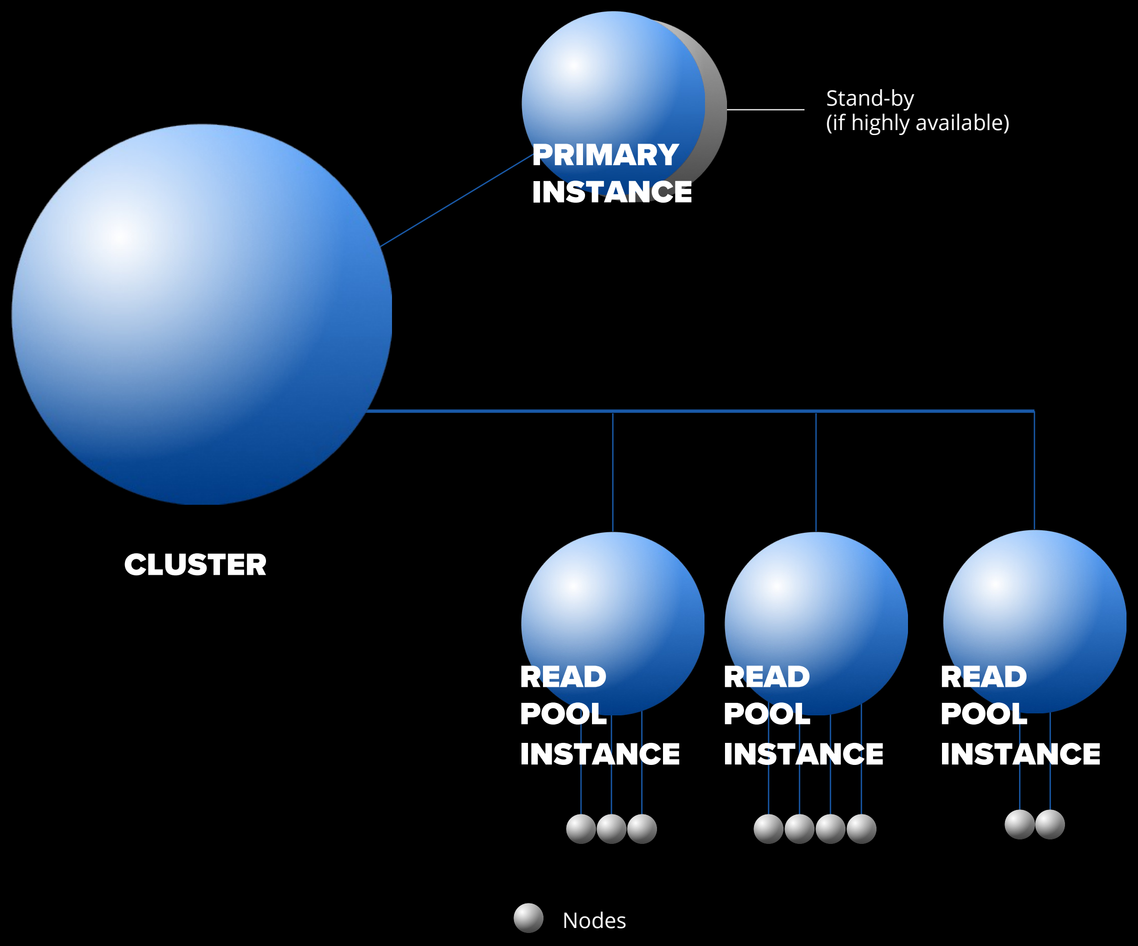
ALLOYDB FOR POSTGRESQL

ALLOYDB FOR POSTGRESQL



softserve

ALLOYDB HIERARCHICAL STRUCTURE



- A **cluster** contains all the resources for a PostgreSQL deployment.
- A **primary instance** provides the read/write connection point for the databases in a cluster. Every cluster has one primary instance.
- A **read pool** instance provides a read connection point for database data in a cluster.

Introducing AlloyDB, a PostgreSQL-compatible cloud database service

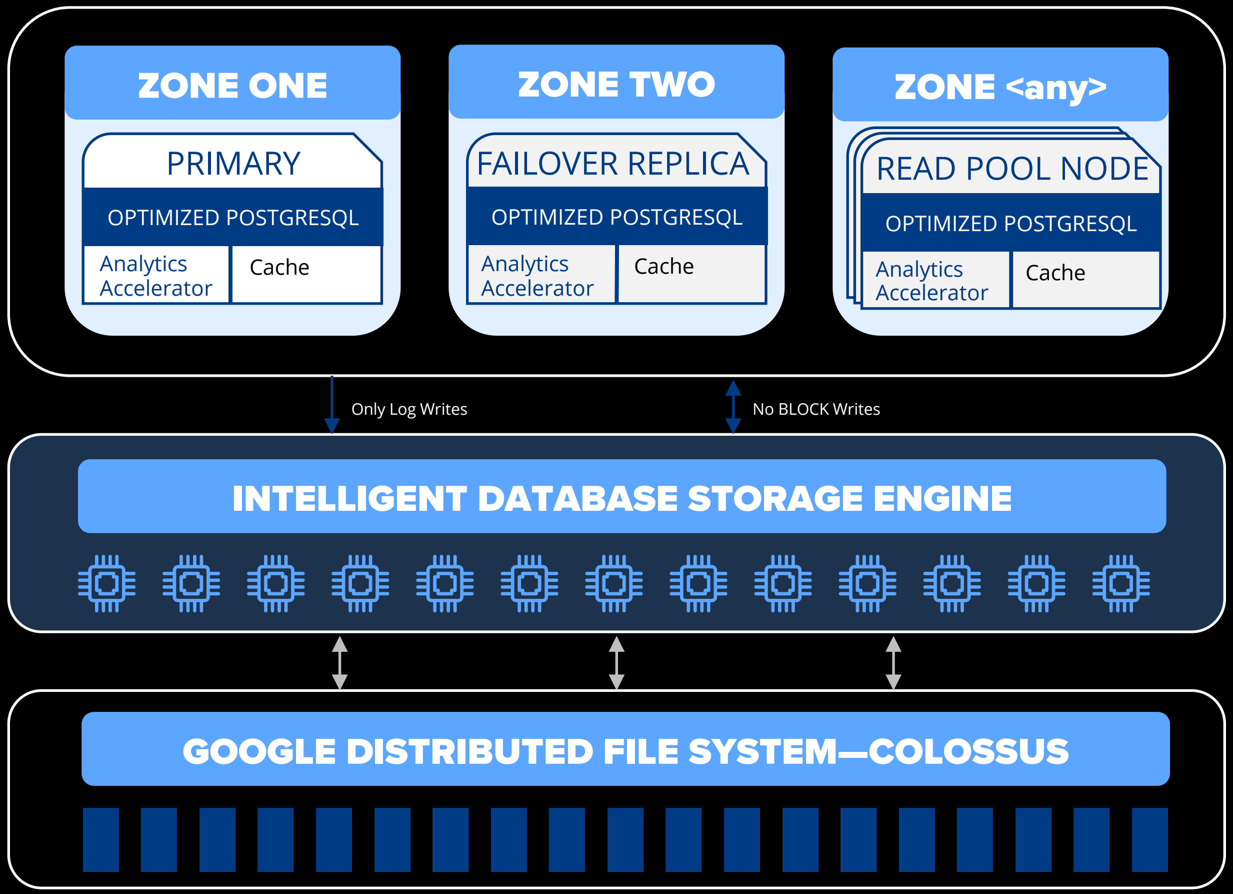
softserve

ALLOYDB ARCHITECTURE

INTELLIGENT DATABASE STORAGE DESIGNED AND OPTIMIZED FOR POSTGRESQL

Powers fast, predictable performance by eliminating I/O bottlenecks and offloading to storage service.

Regional storage improves cluster availability with fast, bounded failover and enables low-lag-to-read replicas.

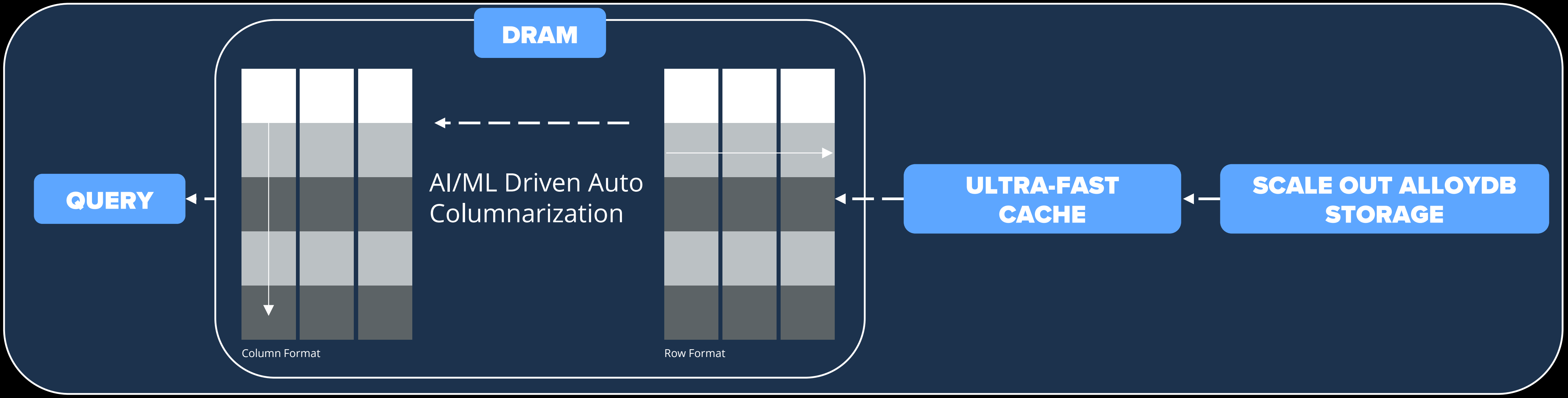


softserve

ALLOYDB ARCHITECTURE

FAST AND PREDICTABLE PERFORMANCE

Intelligent, workload-aware dynamic data organization leverages both row-based and column-based formats. Format layers of cache ensure excellent price-performance.



ADDITIONAL FEATURES

EASY TO MANAGE

- Automatic vacuum management
- Automatic memory management
- Automatic storage tiering
- Automatic data columnarization and query rewrite

FULLY POSTGRESQL COMPATIBLE

- Fully compatible with PostgreSQL 14
- Over 175 flags supported
- Over 50 extensions supported
- Move your existing PostgreSQL application as-is, with no code changes

PREDICTABLE AND TRANSPARENT PRICING

- No licensing or opaque I/0 charges
- Great price-performance
- Right-size instance when needed
- Pay-for-what-you-use storage

3

AWS AURORA

softserve

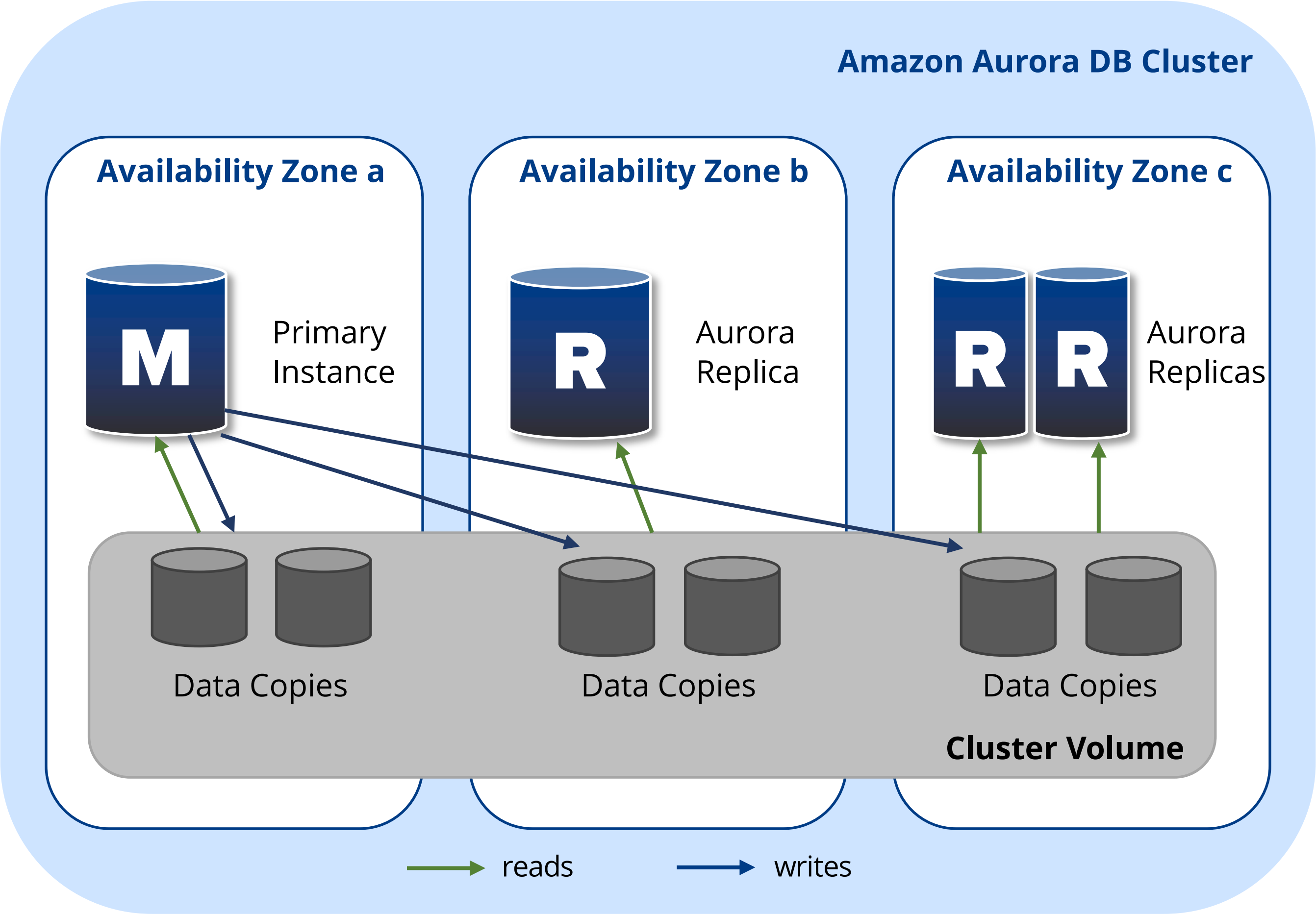
AMAZON AURORA
BASIC ARCHITECTURE

Amazon Aurora is a cloud-based relational database engine that combines the speed and reliability of high-end commercial databases with the simplicity and cost-effectiveness of open-source databases.

Basically, they have taken PostgreSQL and MySQL and replaced the storage layer with a proprietary layer that allows it to be distributed.

AWS re:Invent 2021 - Deep dive on Amazon Aurora

AMAZON AURORA DB CLUSTER



PRIMARY DB INSTANCE

Supports read and write operations and performs the data modifications to the cluster volume. Each Aurora DB cluster has one primary DB instance.

AURORA REPLICA

Connects to the same storage volume as the primary DB instance and supports only read operations. Each Aurora DB cluster can have up to 15 Aurora Replicas in addition to the primary DB instance.

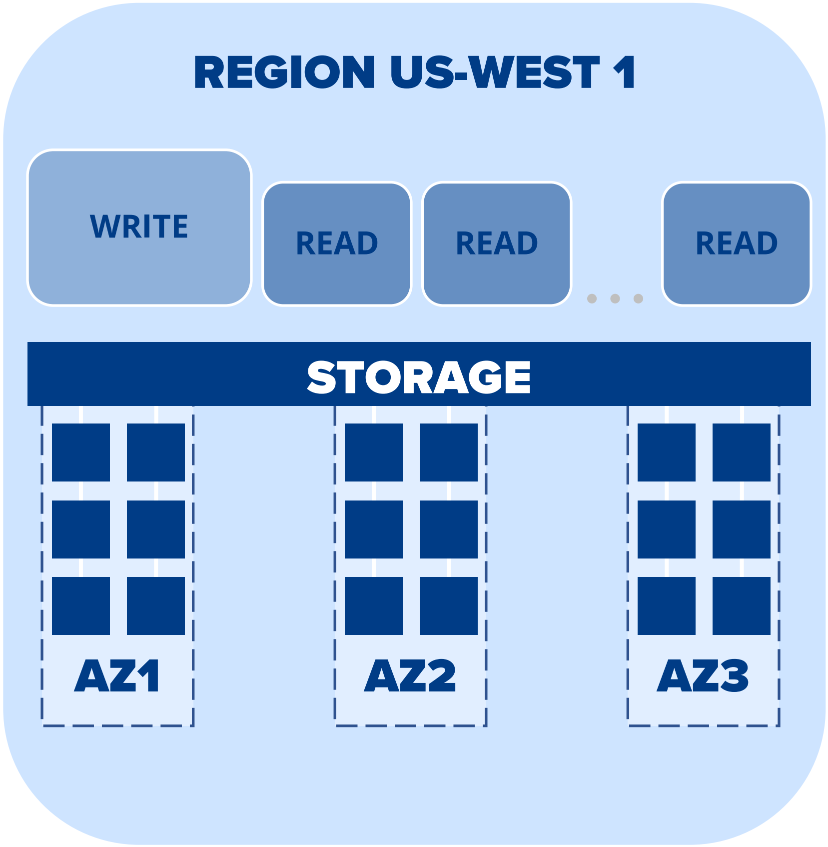
BASIC ARCHITECTURE

WRITE NODE
A single node or endpoint that makes all write requests for the database.

READ NODES
Multiple read-only endpoints to meet your read throughput requirements, typically deployed across multiple AZs.

STORAGE LAYER
A collection of machines with SSD spread across multiple AZs. Data is written here six times.

The trick is … everything is decoupled instead of writing locally to attached persistent storage, it writes to this custom, distributed storage layer.



BASIC ARCHITECTURE

AURORA: 4 OF 6 QUORUM WRITES

The storage layer will commit a transaction when four of six copies are written.

They do this so that the database can survive the loss of an availability zone. If two copies were in an AZ, the data can still have four copies.

Reads are guaranteed when three writes are written.

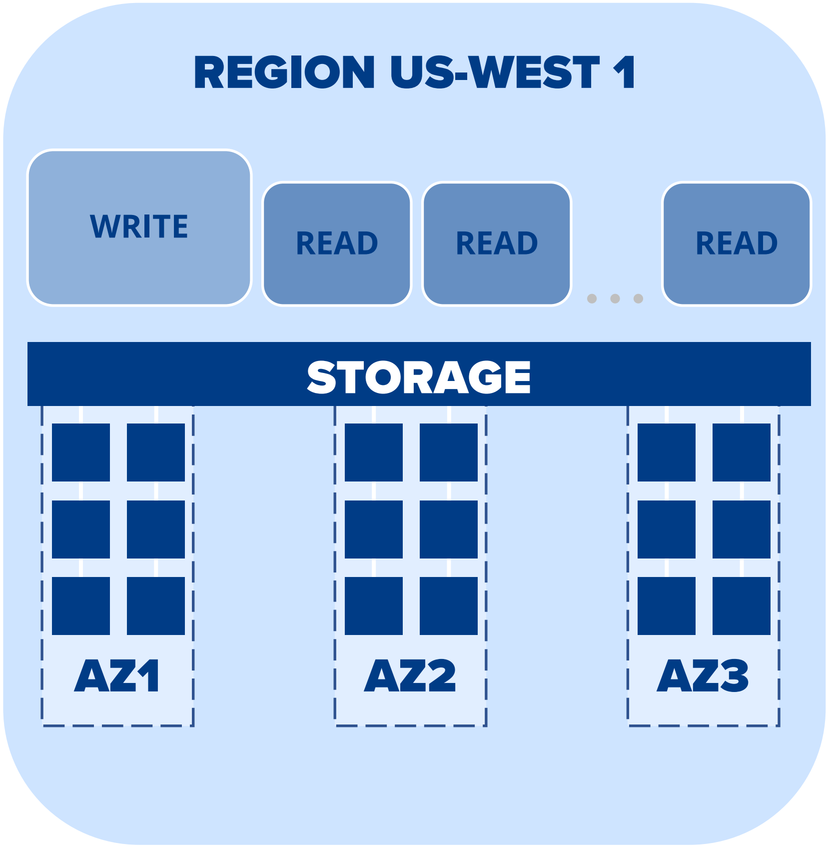
BASIC ARCHITECTURE

SCALE (FOR READS IN A SINGLE REGION)

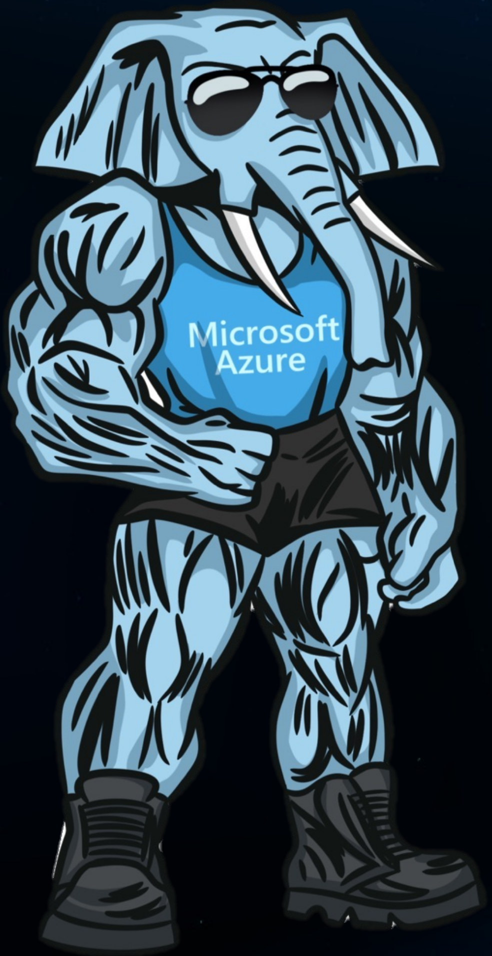
In order to scale Aurora, you simply add more instances on top of the shared storage, and they all have immediate access to all data written to the disk.

RESILIENCE (FOR READS IN A SINGLE REGION)

If a read node fails, it can just be recycled, and all queries can just be directed to other instances while it recovers.

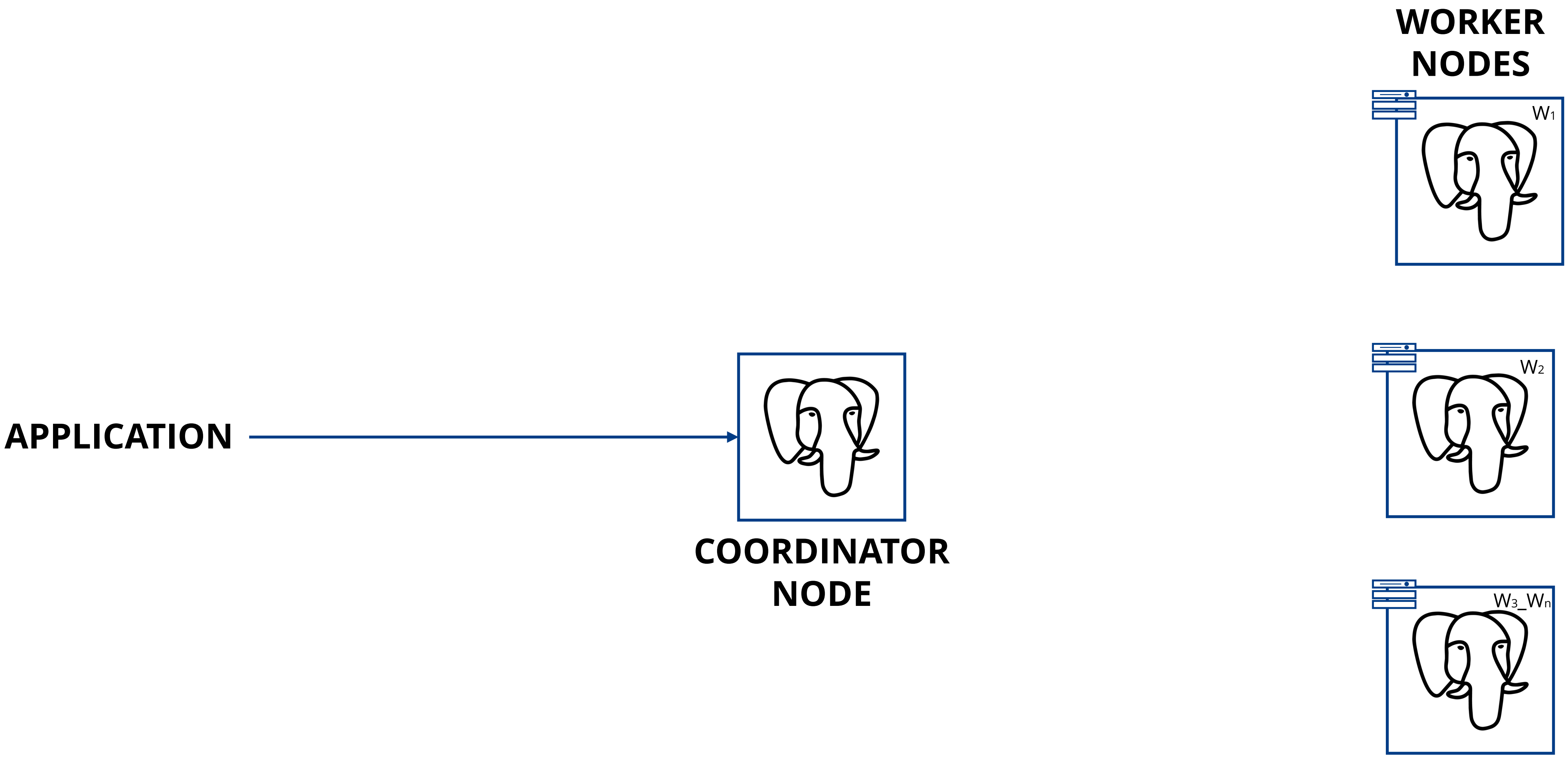


softserve



AZURE DATABASE FOR POSTGRESQL—HYPERSCALE (CITUS) IS NOW AZURE COSMOS DB FOR POSTGRESQL

Oct 12, 2022



A Citus cluster consists of multiple PostgreSQL servers with the Citus extension.

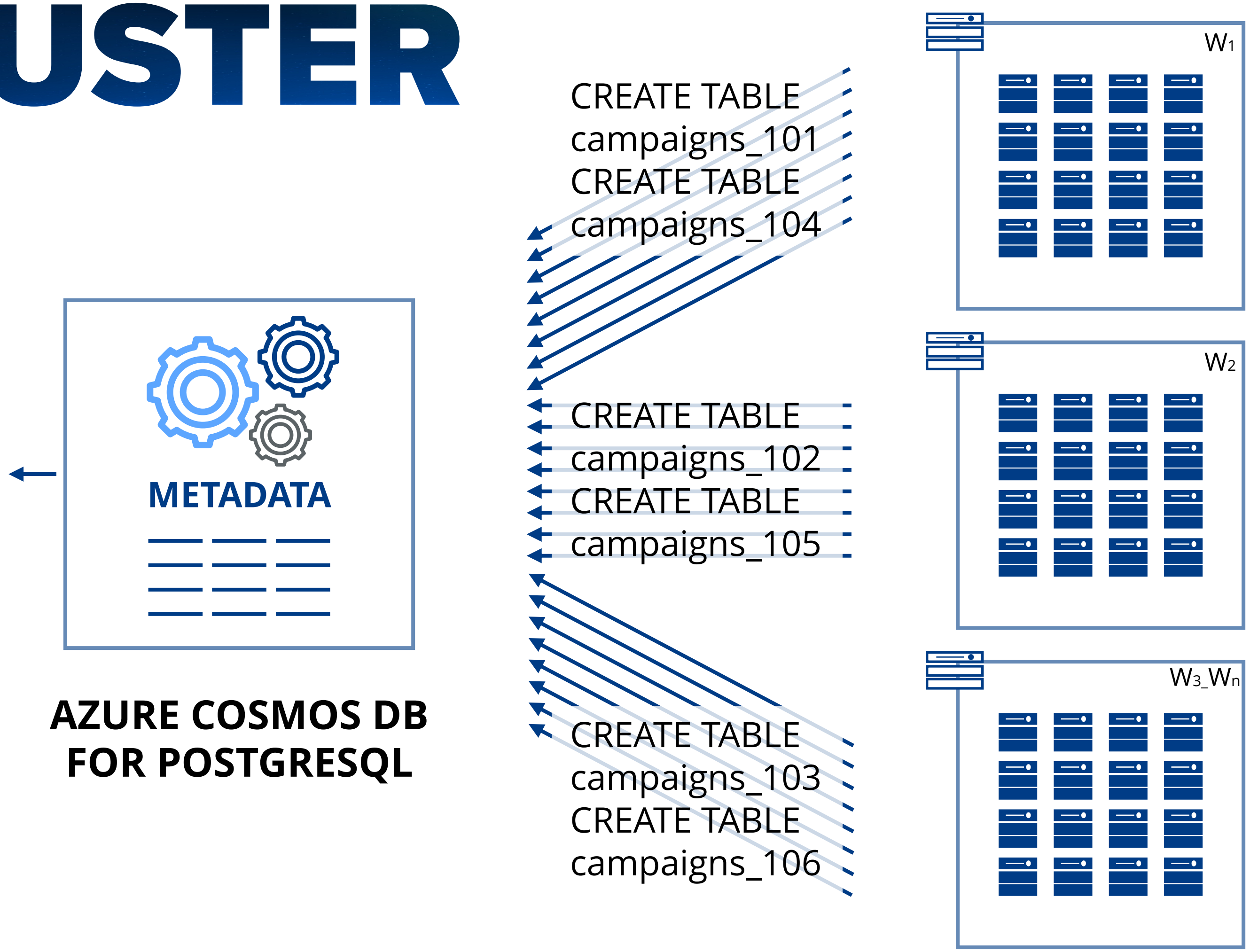
softserve

[Distributed Postgres: How to build a multi-tenant SaaS app with Citus]

HOW CITUS DISTRIBUTES QUERIES ACROSS THE DATABASE CLUSTER

APPLICATION

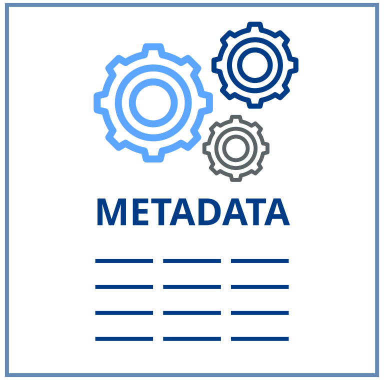
```
CREATE TABLE campaigns (...) ;
        SELECT create_distributed_table(
                'campaigns' , 'company_id' ) ;
```



HOW CITUS DISTRIBUTES TRANSCATIONS IN A MULTI-NODE CLUSTER

APPLICATION

```
    BEGIN ;
   UPDATE  campaigns
      SET  started = true
    WHERE  campaign_id = 2;
   UPDATE  ads
      SET  finished = true
    WHERE  campaign_id = 1;
COMMIT ;
```

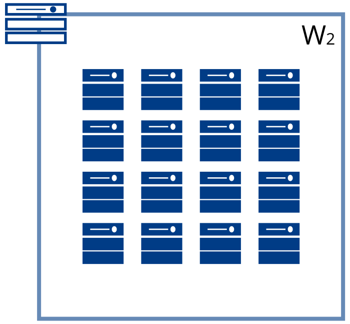


COORDINATOR NODE

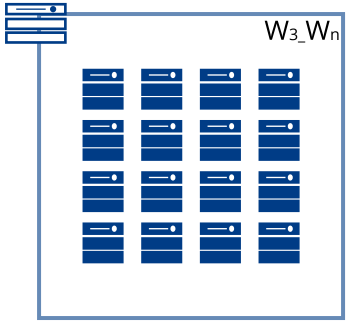
callbacks:

- pre-commit
- post-commit
- abort

BEGIN …
assign_distributed_
transaction_id …
UPDATE campaigns_102 …
PREPARE TRANSACTION…
COMMIT PREPARED…



BEGIN …
assign_distributed_
transaction_id …
UPDATE campaigns_203 …
PREPARE TRANSACTION…
COMMIT PREPARED…



softserve

ALL THE FUNCTIONS OF POSTGRES AVAILABLE TO CITUS CLUSTER

- JSONB
- Joins
- Functions
- Constrains
- Indexes:
  - B-tree
  - Gin
  - Brin
  - Gist
- Partial indexes
- Other extensions
- Rich datatypes
- Foreign data wrappers
- Window functions
- CTEs
- Full text search
- pg_stat_statements

softserve

4 BENCHMARKING

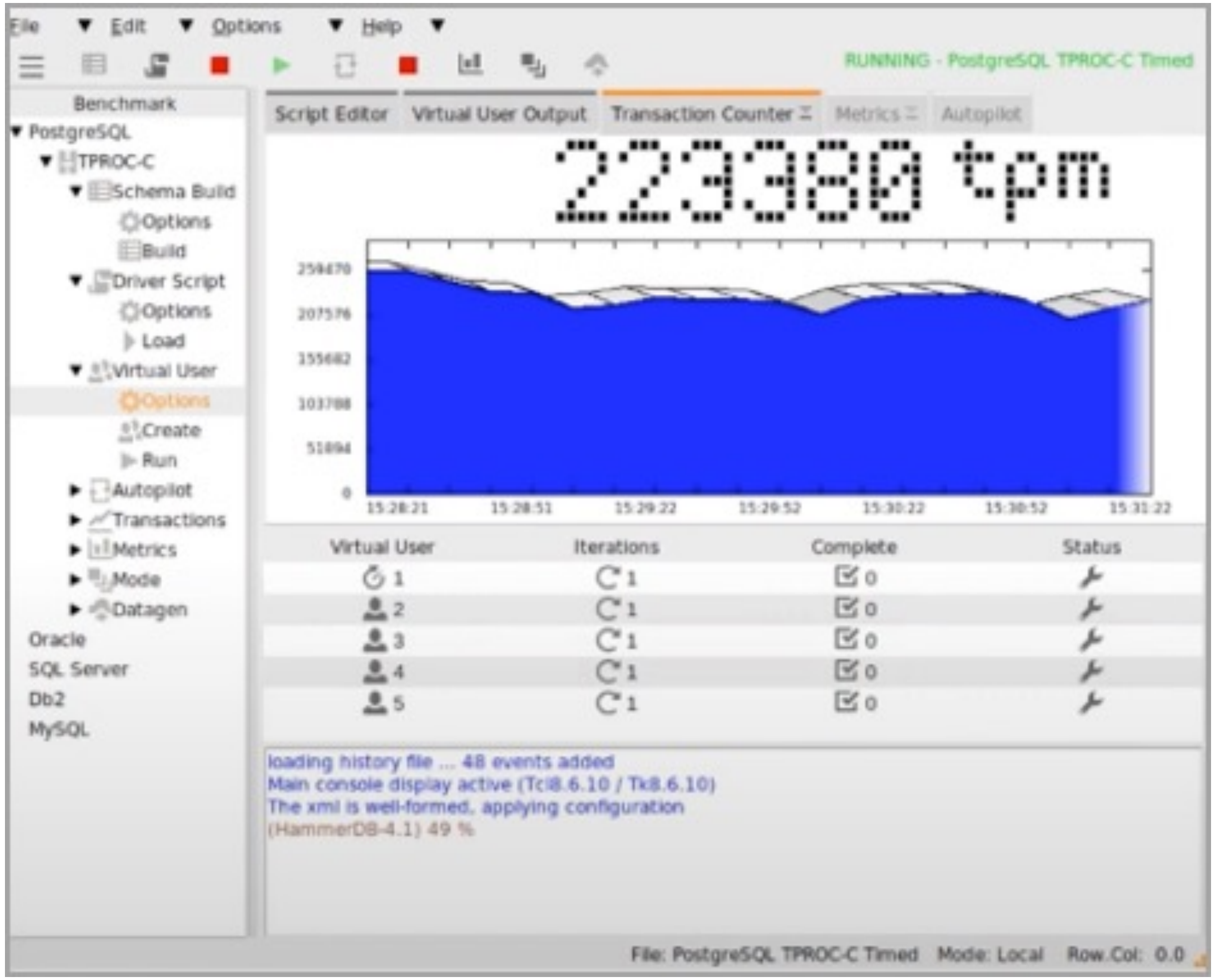
softserve

MY MAIN ADVICE WHEN RUNNING PERFORMANCE BENCHMARKS FOR POSTGRES IS: “AUTOMATE IT!”

Jelte Fennema

WHAT IS HAMMERDB?

- Not a database!
- Leading open-source tool for benchmarking relational databases
- Interfaces:
  - Graphical
  - Command Line
  - Web REST interfaces
- Industry standard benchmarks
- High performance and scalability

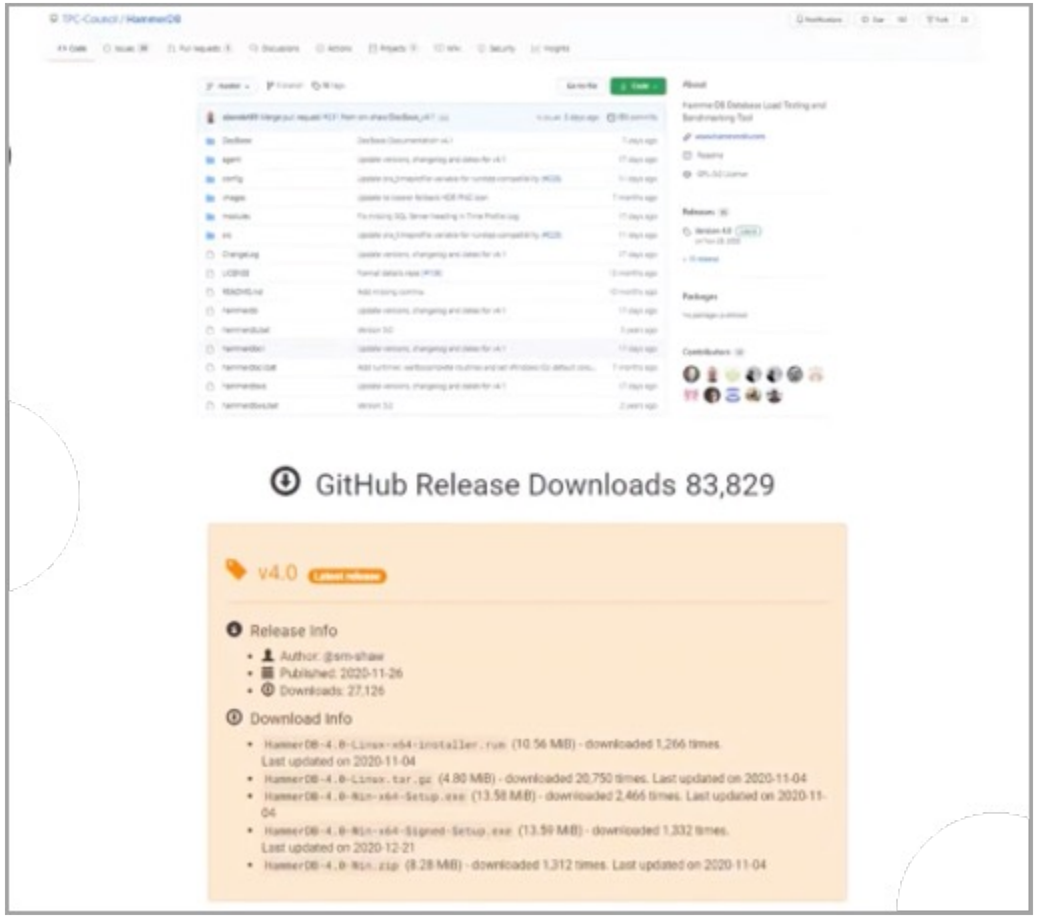


https://hammerdb.com

softserve

TPC OPEN SOURCE

- Hosted by TPC Council since 2019
  - Industry standard body for database benchmarks
- TPC-OSS subcommittee
  - Oversees and approves changes
- V4.1 Released on April 22, 2021
- Source code on GitHub
- Binaries @ GitHub Releases
  - https://www.hammerdb.com/download.html
- Client natively supports Linus and Windows on x64
  - GUI & CLI on both Linux and Windows
- GitHub Release Downloads
  - https://www.hammerdb.com/stats.html
- Test databases on any platform



SUPPORTED WORKLOADS

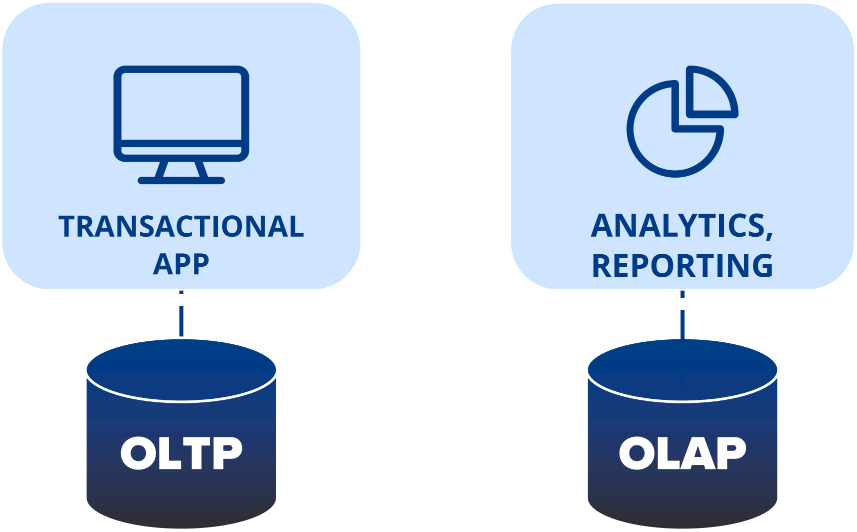
TPROC-C = OLTP

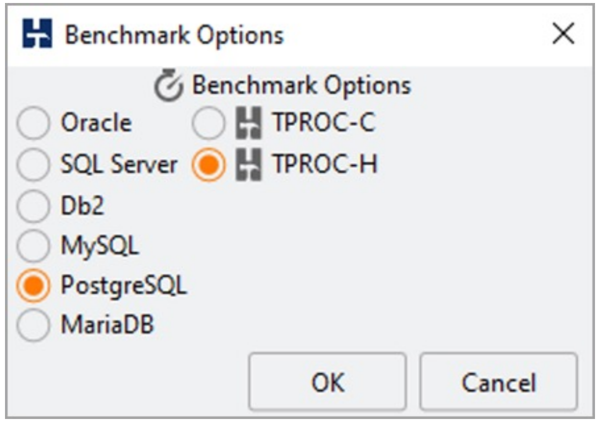
- Transactional workloads. Row oriented, high read and write throughput
- Derived from TPC-C

TPROC-H = OLAP

- Analytic, Decision Support
- Focus on ETL
- High bandwith reads and minimal writes
- Derived from TPC-H

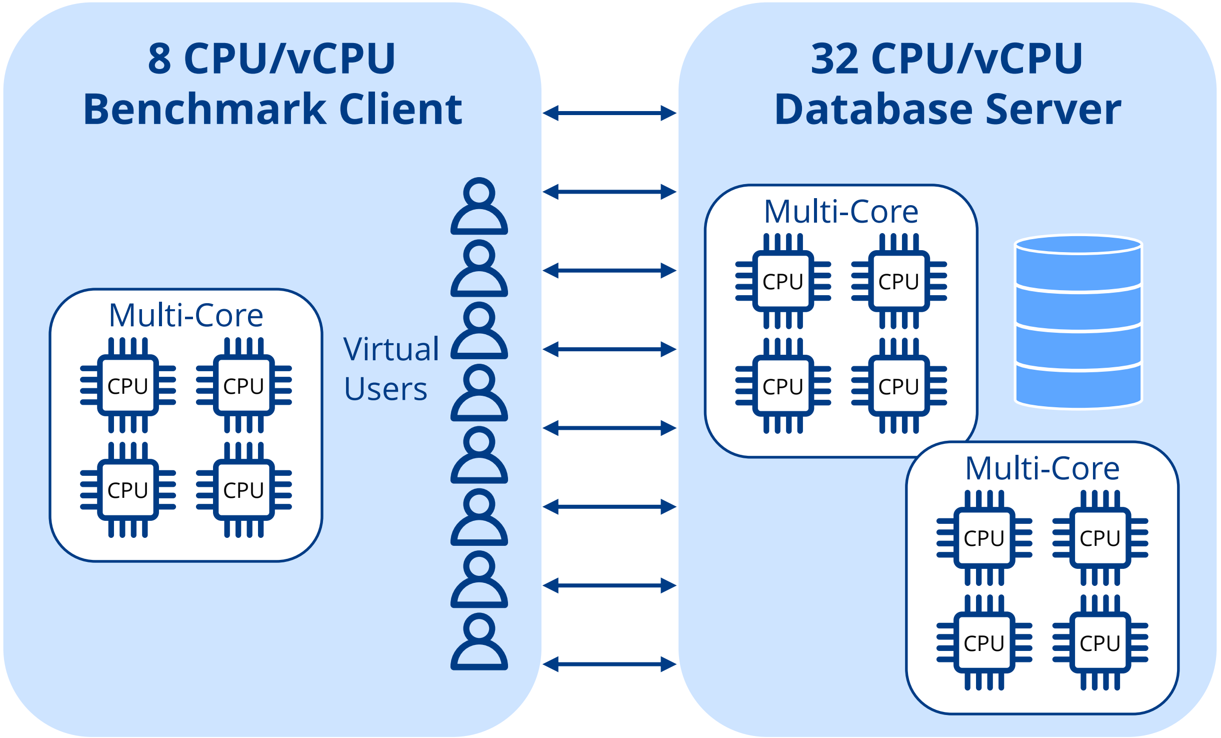
Using TPCC/TPC-C, TPCH/TPC-H for derives workloads not permitted (trademark violation)





KEY DATABASE BENCHMARKING CONCEPTS

- Parallel benchmarking software
  - Concurrency control must be in database, not in client
- Complex workloads designed to scale and test RDBMS concepts
  - Locking and latching
- Cross reference workloads across multiple database engines
  - Validate concepts
- HammerDB up to 6-7 NOPM on commercial database engines on two socket servers
  - High confidence levels that bottlenecks are in database software not HammerDB



SCHEMA BUILD CHOICES

SCHEMA BUILD

- Creates tables
- Creates and loads data
- Creates Indexes
- Creates functions/stored procedures
- Gathers statistics

NUMBER OF WAREHOUSES

- Define according to system scale
- Entire schema scaled based on warehouse count

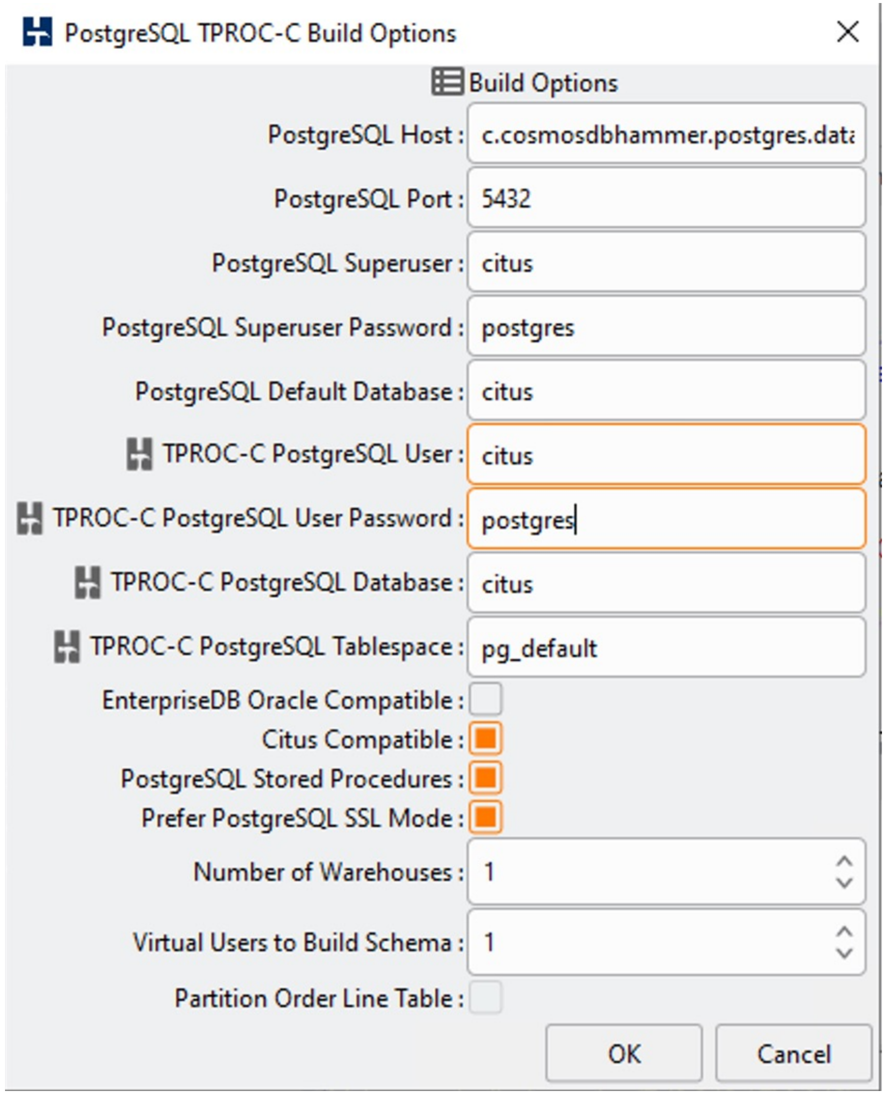
STORE PROCEDURES

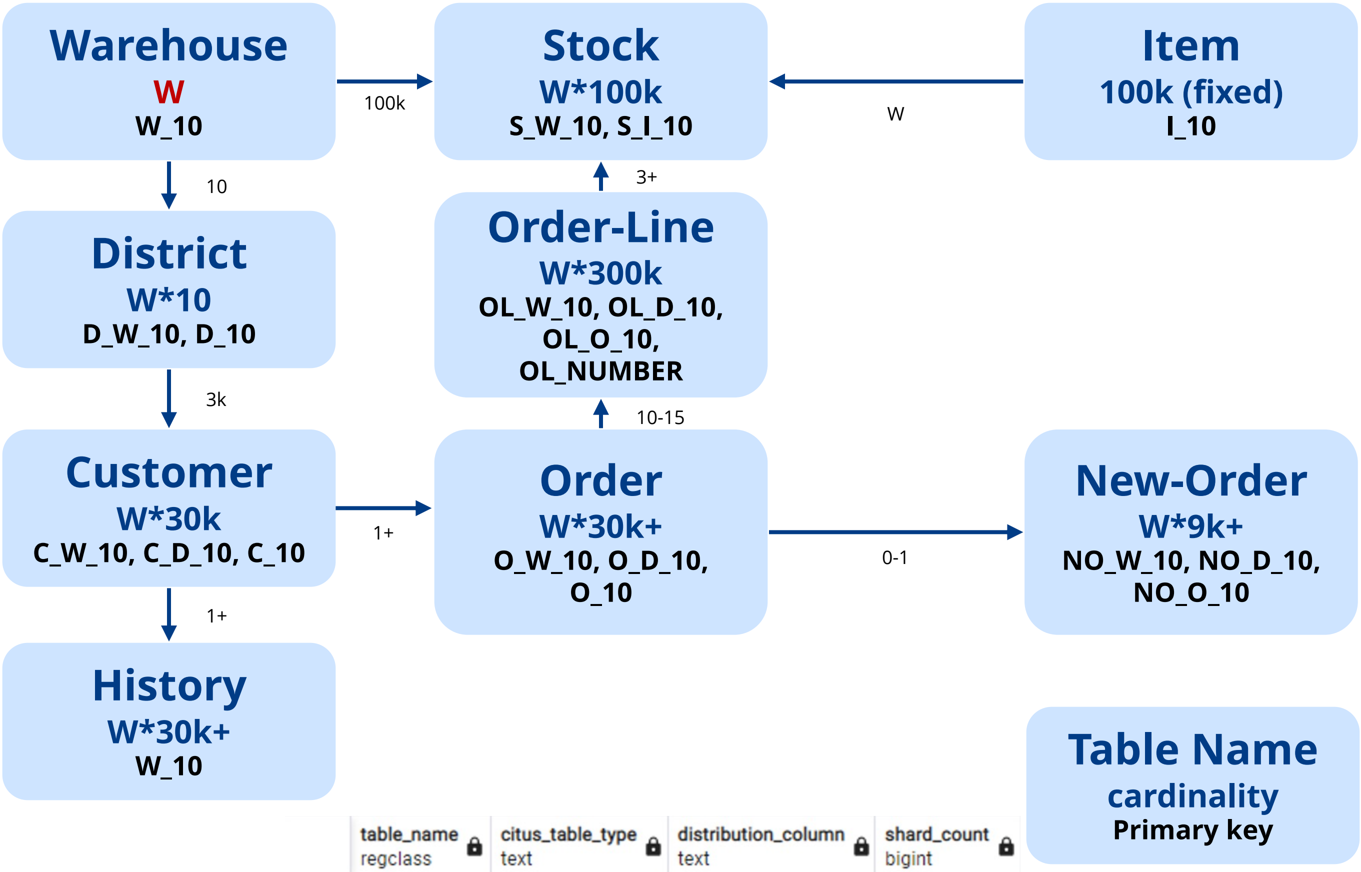
- New Order
- Payment
- Delivery
- Stock Level
- Order Status

VIRTUAL USERS TO BUILD SCHEMA

- Schema creates and loads data in parallels
- Use number of CPU cores/threads on HammerDB client

SCHEMA BUILD CHOICES





	table_name regclass	citus_table_type text	distribution_column text	shard_count bigint
1	customer	distributed	c_w_id	32
2	district	distributed	d_w_id	32
3	history	distributed	h_w_id	32
4	item	reference	<none>	1
5	new_order	distributed	no_w_id	32
6	order_line	distributed	ol_w_id	32
7	orders	distributed	o_w_id	32
8	stock	distributed	s_w_id	32
9	warehouse	distributed	w_id	32

softserve

UNDERSTANDING RESULTS: NOPM vs TPM

```
Vuser 1:  Test complete, Taking end Transactional Count.
Vuser 1: 140 Active Virtual Users configurated
Vuser 1: TEST RESULT : System achieves 1722391 NOPM from 5216947 PostreSQL TPM
```

NOPM

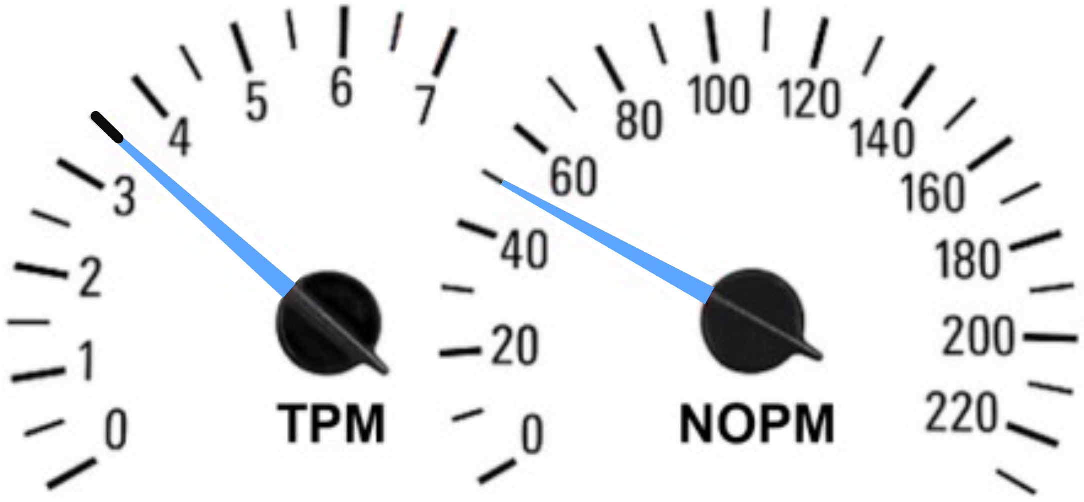
- How fast you are going
- Close relation to official tpmC

TPM

- How hard your engine is working

COMPARING PERFORMANCE

- NOPM can be compared between engines
- TPM can only be compared across the same engine
- TPM useful engineering metric to compare statistics



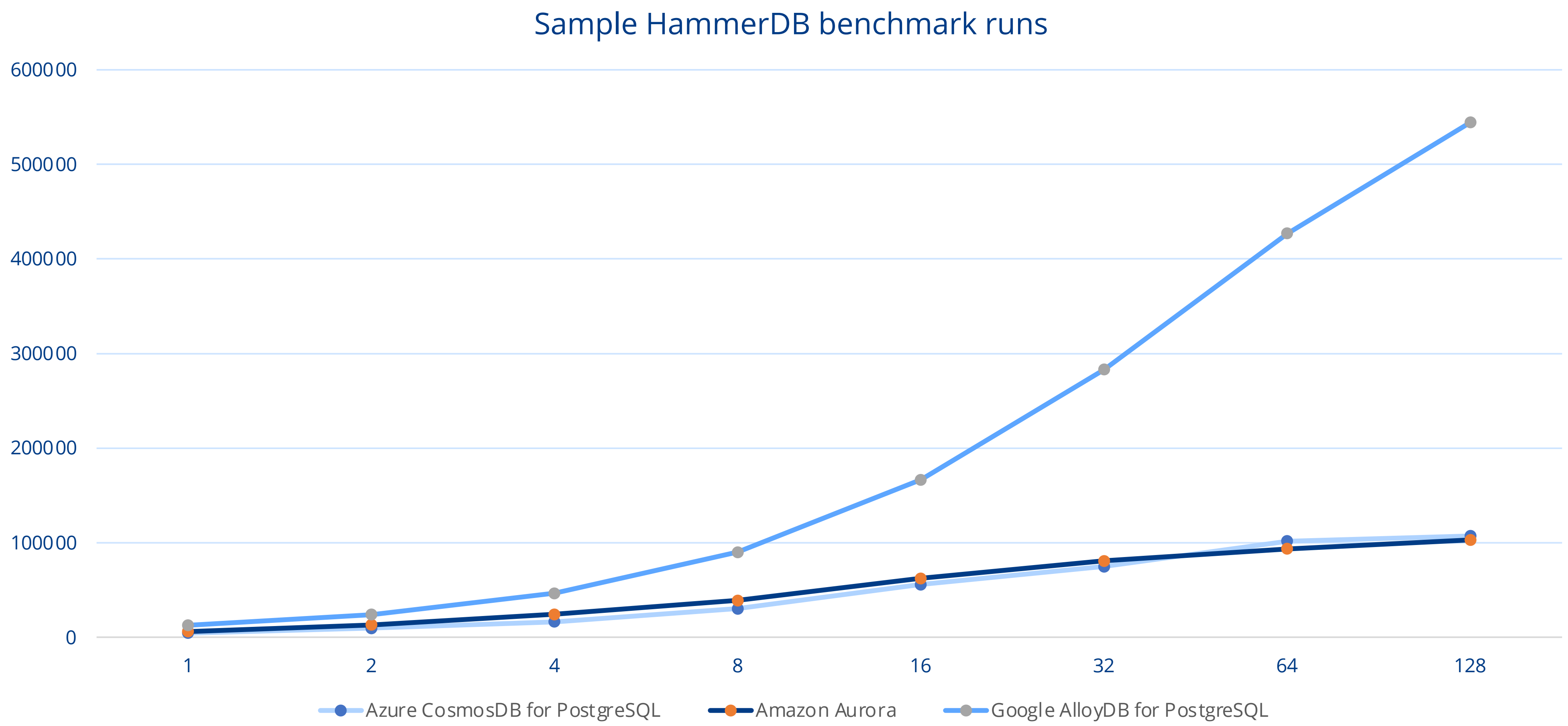
softserve

TABLE WITH COMPARISON

LIMITS	**GCP ALLOYDB**	**AWS AURORA**	**AZURE COSMOSDB FOR POSTGRESQL**
Max database storage per cluster	5 TiB	128 TiB	2 TiB per worker nodes (max 40 Tib)*
Max read pool nodes/workers per cluster	20 (15 if all nodes are of the 64 vCPU machine type)	Up to 15 Aurora Replicas in addition to the primary DB instance	Up to 20 workers
Maximum concurrent connections	Up to 240’000	Up to 16’000	Up to 2’000*
PostgreSQL compatibility	14	14	15
Cost per month	USD 8’619,26	USD 3’671.24 + IOPS	8’992.47
Hardware Configuration	vCPU: 32 – RAM: 256 GB	vCPU: 32 - RAM: 256 GB	Coordinator: 4 vCPU, 16 GB RAM Worker node: 2 nodes x 16 vCPU, 128 GB RAM

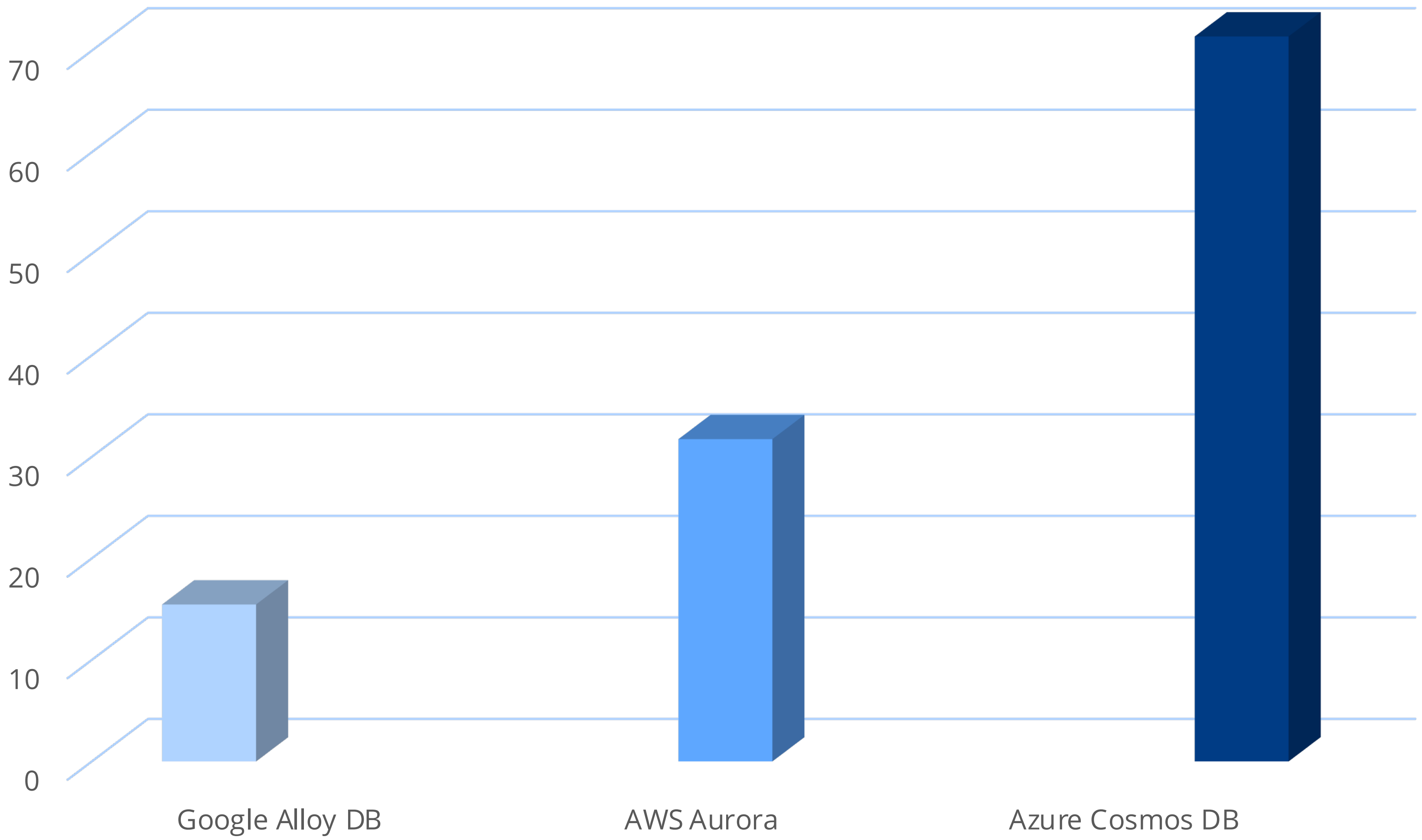
softserve

OLTP TESTING
(TPROC-C DERIVED FROM TPC-C)



softserve

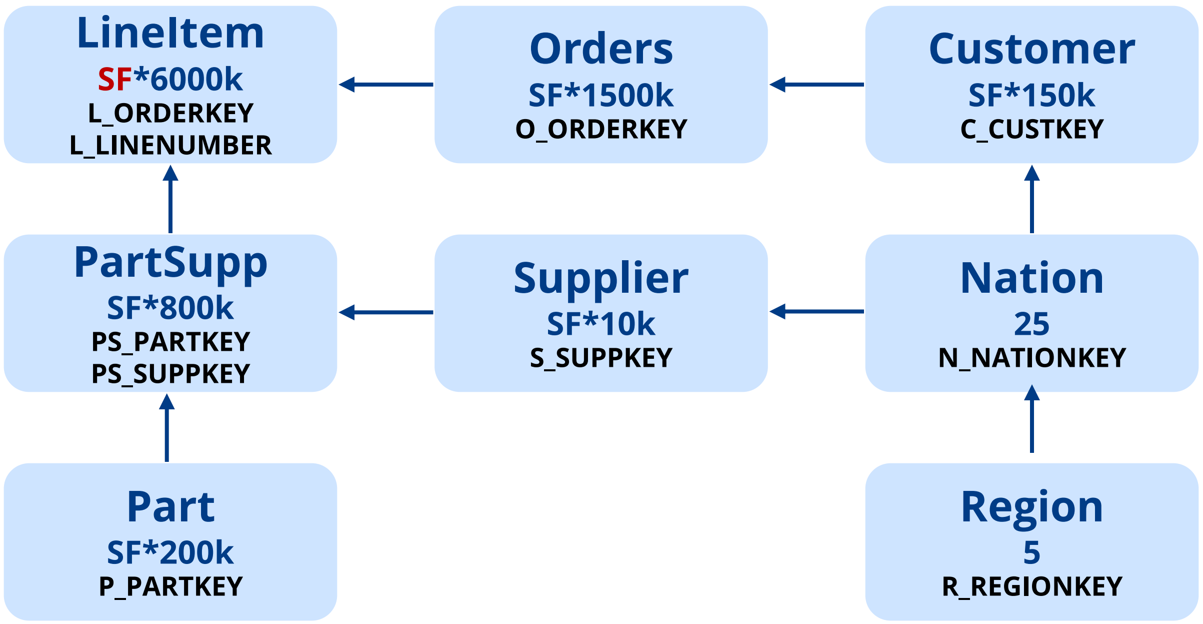
PRICE/PERFORMANCE RESULTS
(PRICE/KNOPM)

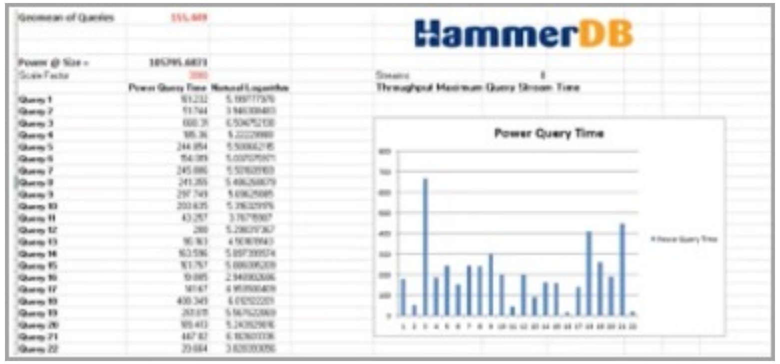


softserve

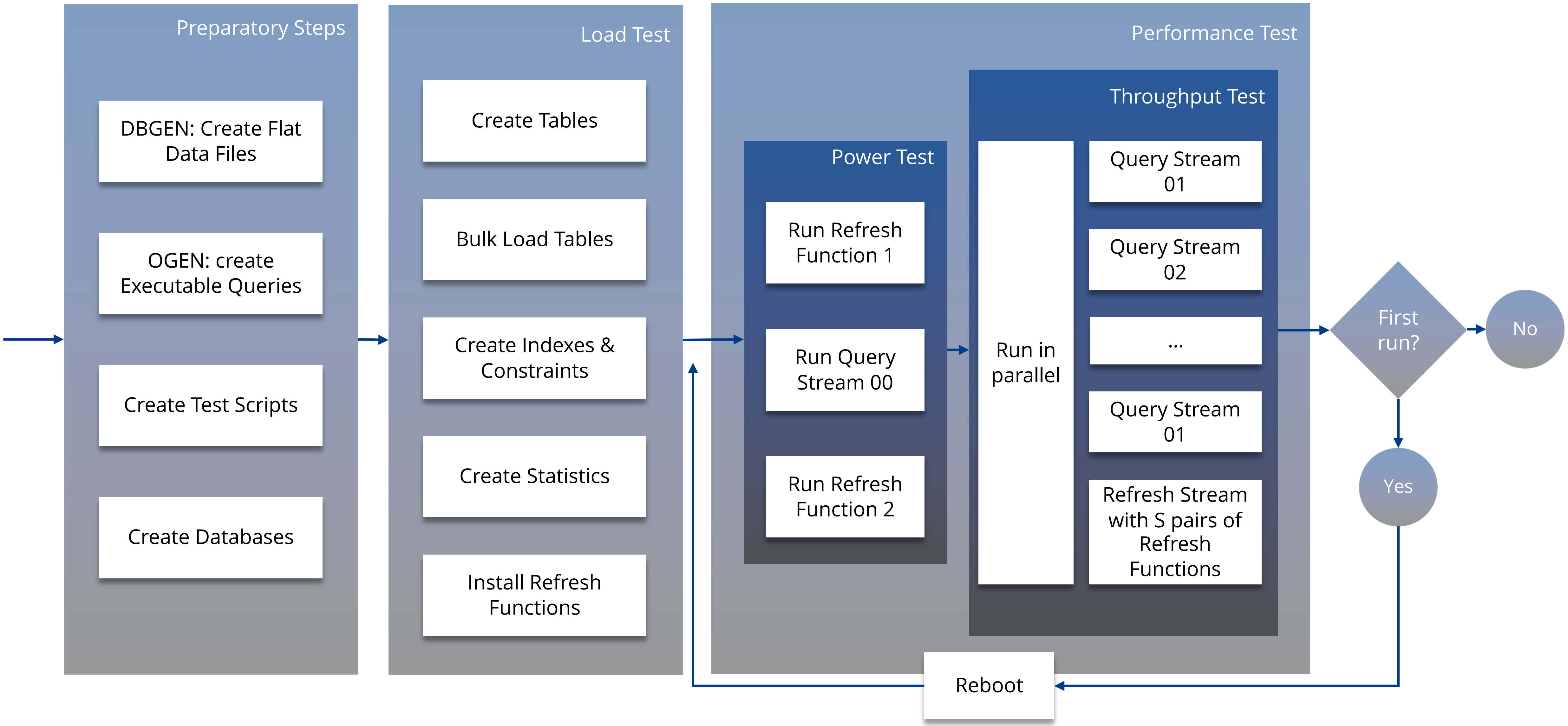
ANALYTICAL TESTING

- TPROC-H for Analytics
- Cloud Queries
- Stream of 22 Complex queries
- PostgreSQL parallel query
- Columnstores
- More complex skill set required



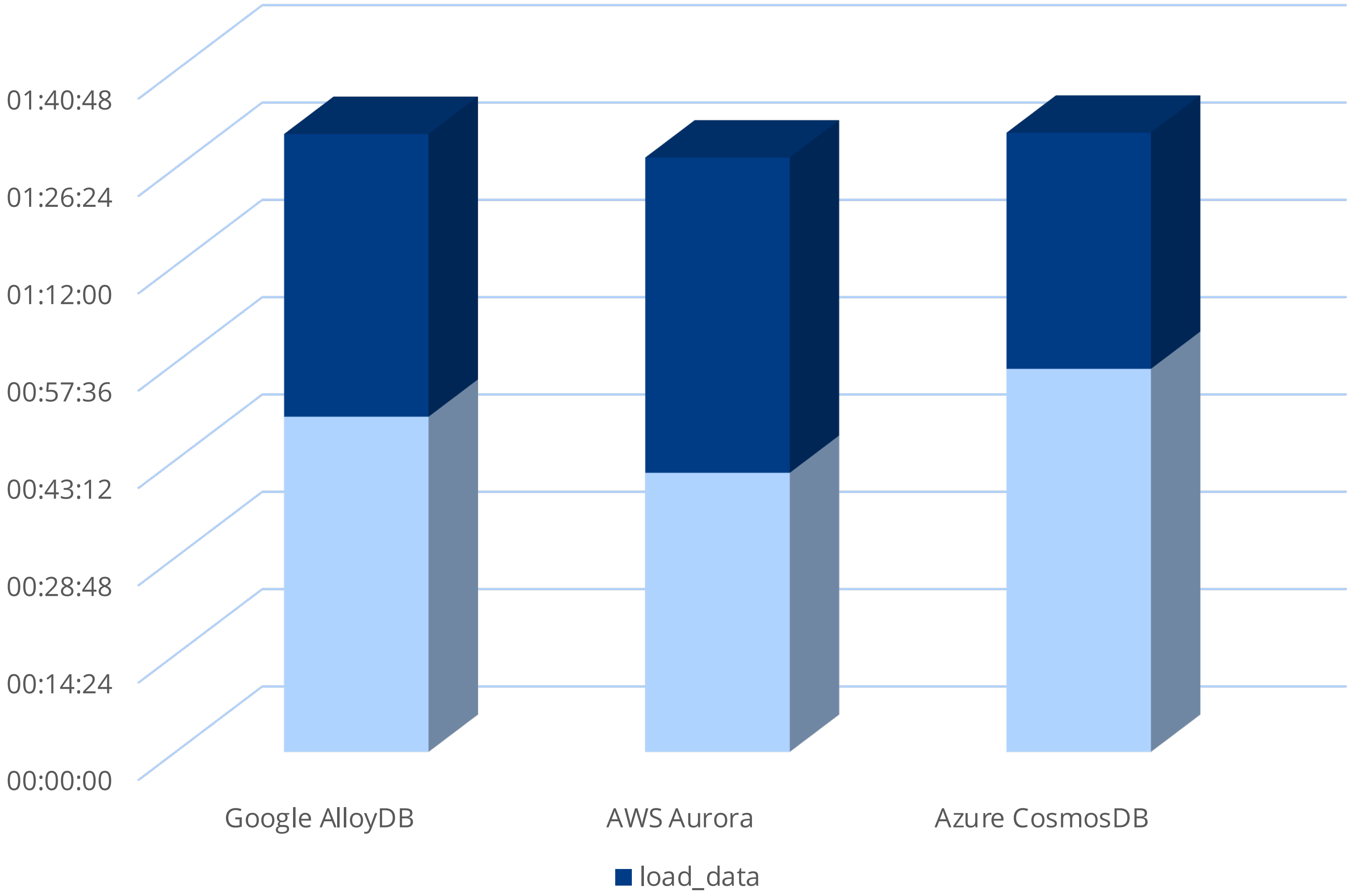


TPC-H PROCESS



softserve

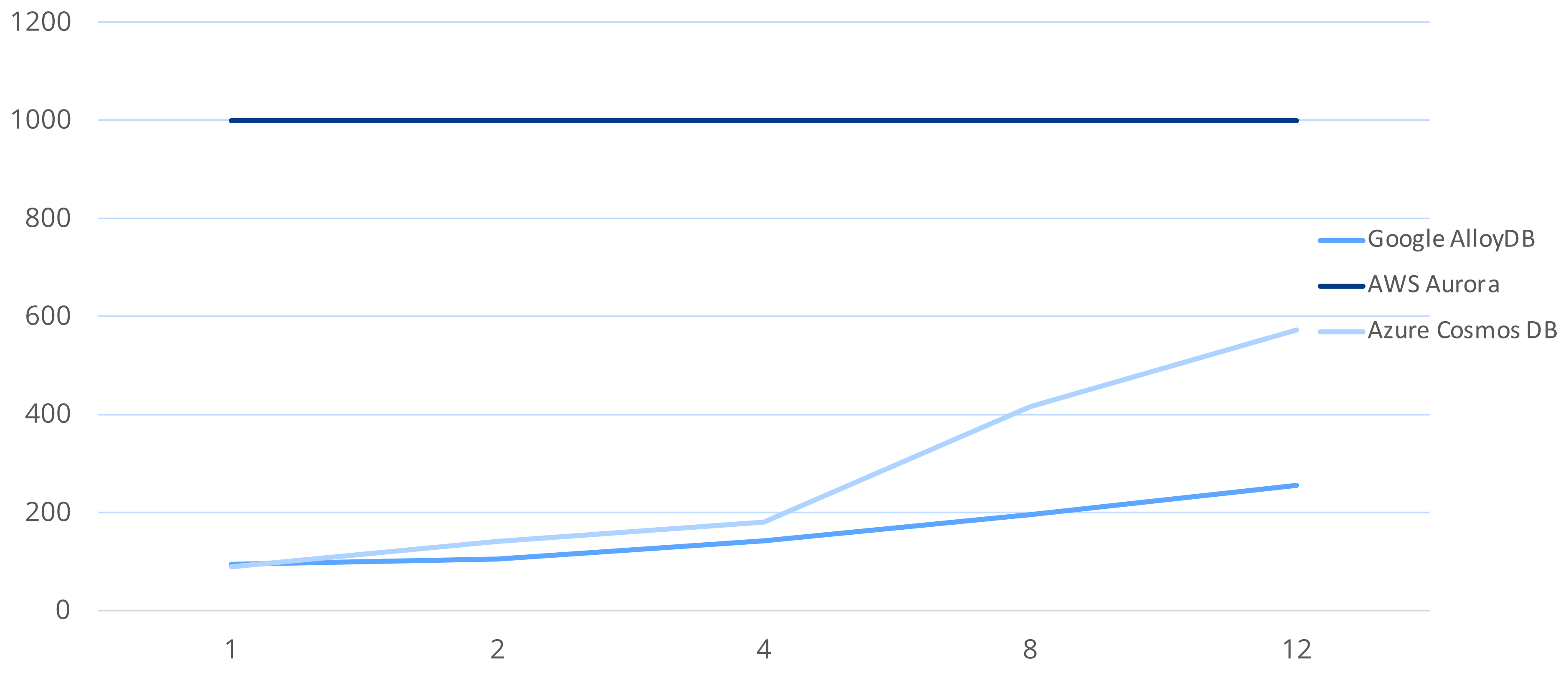
LOAD & INDEX DATA TEST



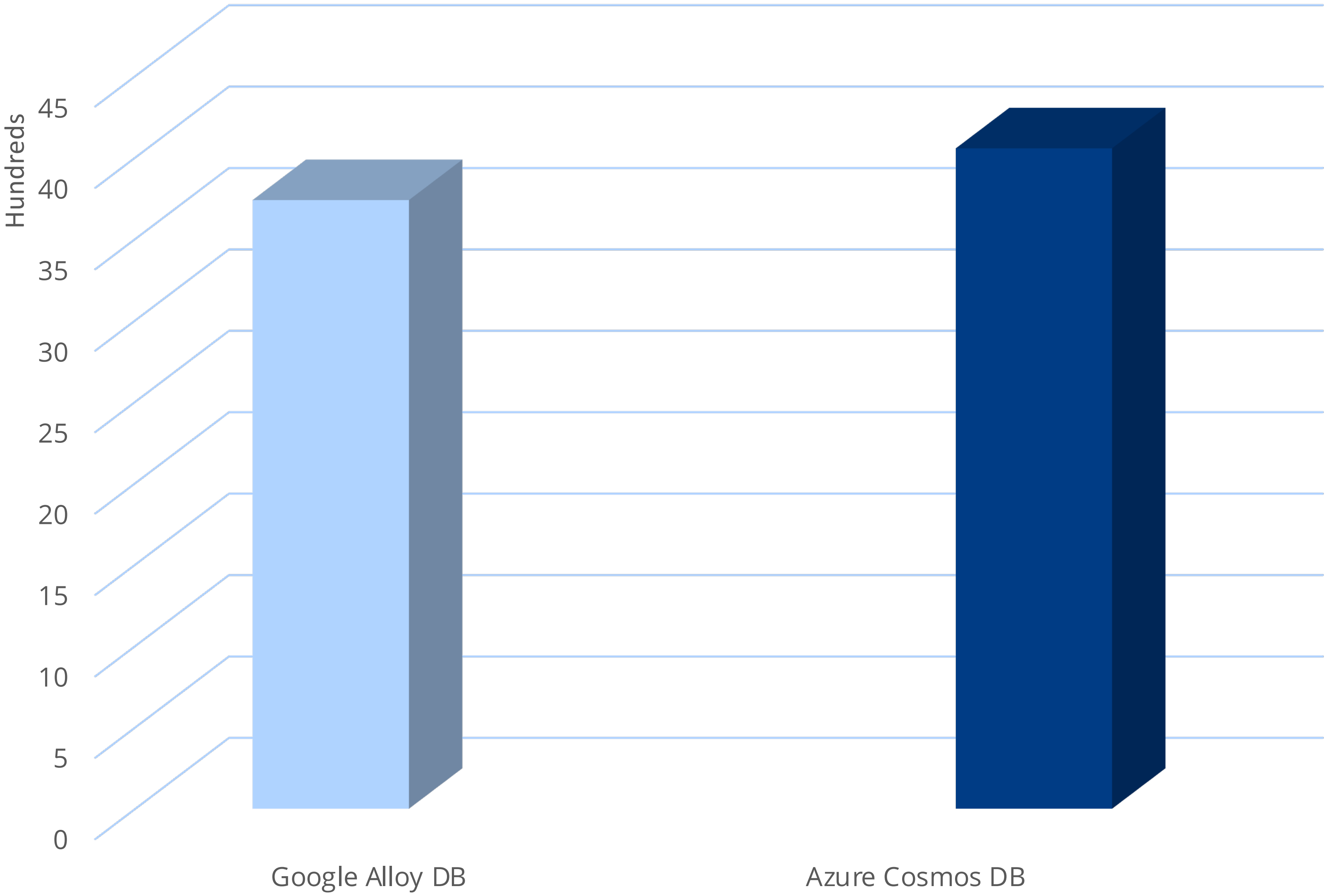
softserve

OLAP TESTING
(TPROC-H DERIVED FROM TPC-H)

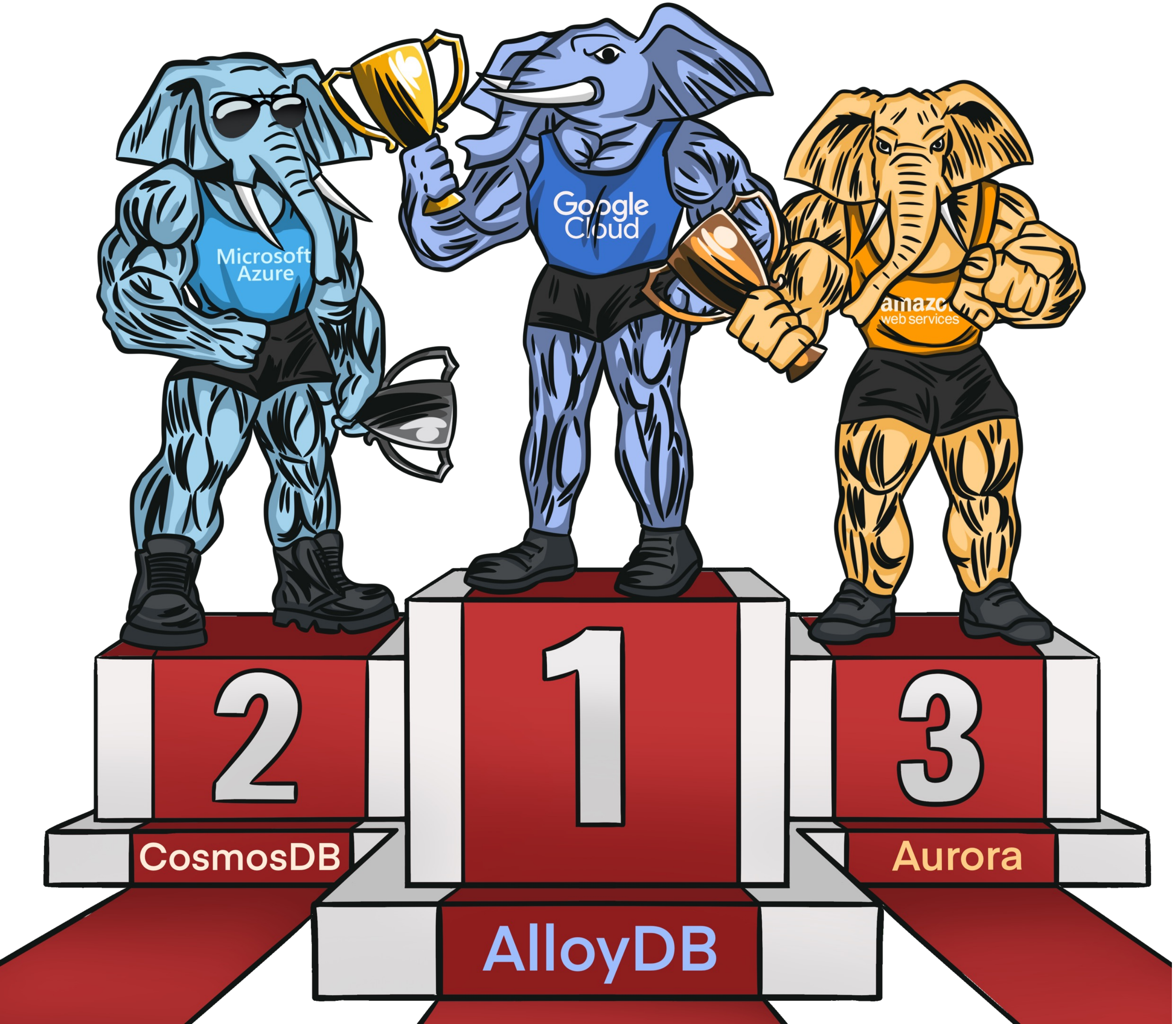
The lower—the better.



PRICE/PERFORMANCE RESULTS
(PRICE/QPHH)



softserve



SUMMARY

- Run PoC(s) to get practical experience and build confidence
- Do full-scale architectural exercises, with “How do I do X?” questions instead of “Can I do X?”
- Try to approach cloud vendors for the best pricing offer

softserve

CONTACT ME DIRECTLY

TARAS KLOBA
Big Data Engineering Manager
at SoftServe Inc.