

OSA CON 26

Paying Rent on Your Own Data

Patrick McFadin
Founder and Principal Advisor
McFadin Data and AI Advisory

Nov 2, 2026 | AWS Builder Loft & Online (osacon.io)

You built it. You filled it.
And every month you pay rent on it.

Your data. Someone else's engine, someone else's catalog, someone else's bill.

Every layer below the query **went the same way.**

LAYER	NEUTRAL INTERFACE	BECAME COMMODITY	VALUE MOVED TO
Operating systems	POSIX, the syscall ABI	Linux	Distros, then clouds
Web serving	HTTP	Apache httpd, nginx	The application
Orchestration	The Kubernetes API	Kubernetes everywhere	Platforms built on top
Transactional DBs	SQL, the Postgres wire protocol	PostgreSQL	Managed services, extensions
Analytics	**Open table formats, Arrow**	**Happening now**	**That's this talk**

Open was always possible. **Easy wasn't.**

THEN: OPEN OR EASY, PICK ONE

EASY

Rent a warehouse

One vendor. One bill.

OPEN

Run Hadoop or Spark

A team of specialists.

NOW: BOTH

Engine	DuckDB · DataFusion
Catalog	**Polaris · Unity**
Format	Iceberg · Delta
Storage	Your bucket

PART 1

The trade **changed.**

Storage went commodity. The markup moved to compute.

Storage went commodity. **Compute didn't.**

PLATFORM	STORAGE, PER TB-MONTH	COMPUTE METER (LIST)
Snowflake	~$23–40	**$2–4 per credit**
Databricks SQL	Your S3, ~$23	**$0.70 per DBU (serverless)**
Google BigQuery	~$23.50	**$6.25 per TiB scanned**
Redshift Serverless	~$24	**$0.375 per RPU-hour**
Microsoft Fabric	~$23 (OneLake)	**$0.18 per CU-hour**
Raw S3 Standard	**~$23**	**Bring your own**

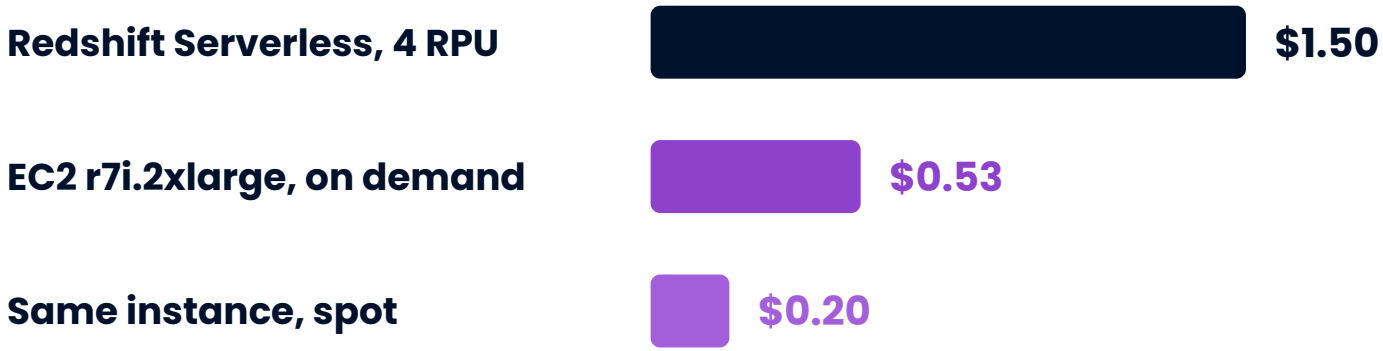
> **Storage prices converged on S3. The margin is in the meter.**

US list prices, Oct 2026, before discounts. Snowflake, Databricks and Fabric via third-party pricing guides.

The easy button **has a markup.**

Serverless fixed idle. Every hour you do use still carries a premium.

SAME 64 GB OF MEMORY, PER HOUR (US-EAST-1)



About 2.8x the bare machine for the same memory.

AND THE MINIMUMS

Snowflake: 60 seconds billed every time a warehouse resumes.

Redshift Serverless: 60-second minimum charge.

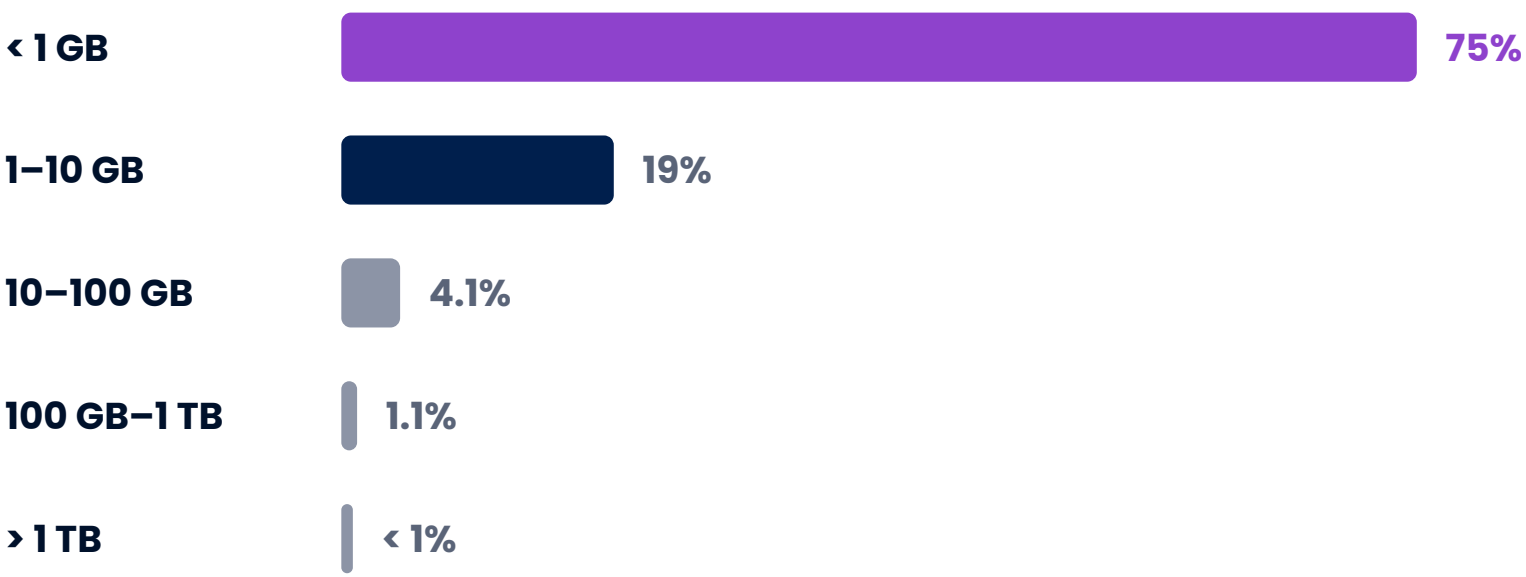
BigQuery: $6.25 per TiB scanned, however efficient the query.

> **Serverless competes on price only if you can point it at your own storage.**

Sources: AWS Redshift docs (1 RPU = 16 GB, 60 s minimum); EC2 pricing via Holori (Nov 1, 2026); Snowflake docs.

Most queries are **laptop-sized.**

Bytes scanned per query, from Amazon's own sample of the Redshift fleet.



93%

of users never ran a query that scanned more than 10 GB.

Only 0.03% of queries scanned more than 10 TB.

Source: Jordan Tigani (MotherDuck), "Redshift Files" (Aug 2024), analyzing Amazon's Redset sample (VLDB 2024). A biased sample, and the author sells DuckDB.

PART 2

Shop for **compute.**

Open source split storage from compute. Now you can buy them separately.

Any engine can now **read your bucket.**

FORMAT

Iceberg is the contract

v3 GA on Snowflake and Databricks (May 2026).

Delta tables readable as Iceberg.

ENGINE

The engine is a library

Arrow, DataFusion, DuckDB.

Trino and Spark for the big jobs.

HARDWARE

One box is enough

Up to 32 TiB of RAM in a single cloud instance.

Distributed is now a choice.

STORAGE

The bucket coordinates

S3 compare-and-swap (2024).

Diskless Kafka accepted (KIP-1150, 2026).

Your data stays put. The engine has to earn every query.

Sources: Snowflake release notes (May 2026); ASF; AWS; Aiven (Mar 2026). KIP-1150 is accepted, not released.

Route each job to **the cheapest engine.**

NEW RELIC · APRIL 2026

Snowflake to Iceberg on S3

1,000+ datasets moved to Iceberg, AWS Glue catalog, Spark on Kubernetes, Kafka writing straight to tables.

Ran both in parallel and checked row-level parity before cutover.

35–52% lower annual data platform spend.

OKTA

DuckDB in front of Snowflake

DuckDB inside AWS Lambda pre-processes security data before it reaches the warehouse.

About 1M invocations a day, peaks of 250 GB a minute.

Motive: a Snowflake bill that kept climbing.

Not rip-and-replace. Offload: keep the warehouse for the 1% that needs it.

Sources: New Relic engineering blog (Apr 29, 2026); Okta, Data Council 2024 talk. Okta published no dollar figures.

PART 3

The new **math.**

AI is cutting the labor side of the equation.

Someone built **the exit ramps.**

The vendors built them to win customers. Independents built them so you can move.

SNOWFLAKE · JUN 2026

AIM migration agent

AI-assisted SQL conversion, testing and validation.

DATABRICKS

Lakebridge

Free migration tooling from 10+ warehouses.

Claims up to 80% of tasks automated.

INDEPENDENT

Eden.dev

Sits in the traffic path: copies data, mirrors writes, cuts over.

Zero app changes. Postgres, MySQL, MongoDB, Redis.

When moving is cheap, lock-in stops being a moat.

Sources: Snowflake (Jun 5, 2026); Databricks (Jun 2025); Redis blog on Eden (Feb 19, 2026). Automation figures are vendor claims.

The specialist team is now **a tool call.**

THE LOOP THAT USED TO NEED A HADOOP TEAM

Provision › **Deploy** › **Benchmark** › **Profile** › **Tune**

EXAMPLE

easy-db-lab

Lab clusters, benchmarks and flame graphs on AWS. An agent drives the whole loop from the CLI.

> **The labor tax was this loop. Agents are good at loops.**

github.com/rustyrazorblade/easy-db-lab

AI makes building cheap. **Running is still on you.**

THE LAB

Let agents loose

Build it, break it, benchmark it.

PRODUCTION

Humans at the gate

Promote only what the lab proved.

13.8% of SRE tasks solved by the best agents in IBM's ITBench.

> **Agents in the lab. Humans at the gate.**

Source: IBM ITBench (ICML 2025).

The new math: **own the storage, shop the compute.**

RENTING

The walled garden

+ **Storage** at S3 prices or above
+ **Compute** at a markup, about 2.8x
+ **Minimums** on every small query
+ **Renewal** priced on your lock-in

OWNING

Open, in your account

+ **Storage** at S3 prices, your bucket
+ **Compute** at market price, per job
+ **Labor** falling, as agents do the lab work
+ **Exit** a cost you control

> **Labor was the line that kept renting ahead. That line is falling.**

Where the rent **gets collected now.**

01	**The catalog**	Open format, closed front door.
02	**Data gravity**	The format is portable. Petabytes and egress are not.
03	**The control plane**	The ops you never learned to run.

> **Watch these when you sign. They are the new markup.**

Own the slow-moving layers. **Rent the fast ones.**

And make sure the open layer you own stays open, or you've just found a new landlord.

OWN

Expensive to move

- Your bucket
- The format
- The catalog
- Policies as code

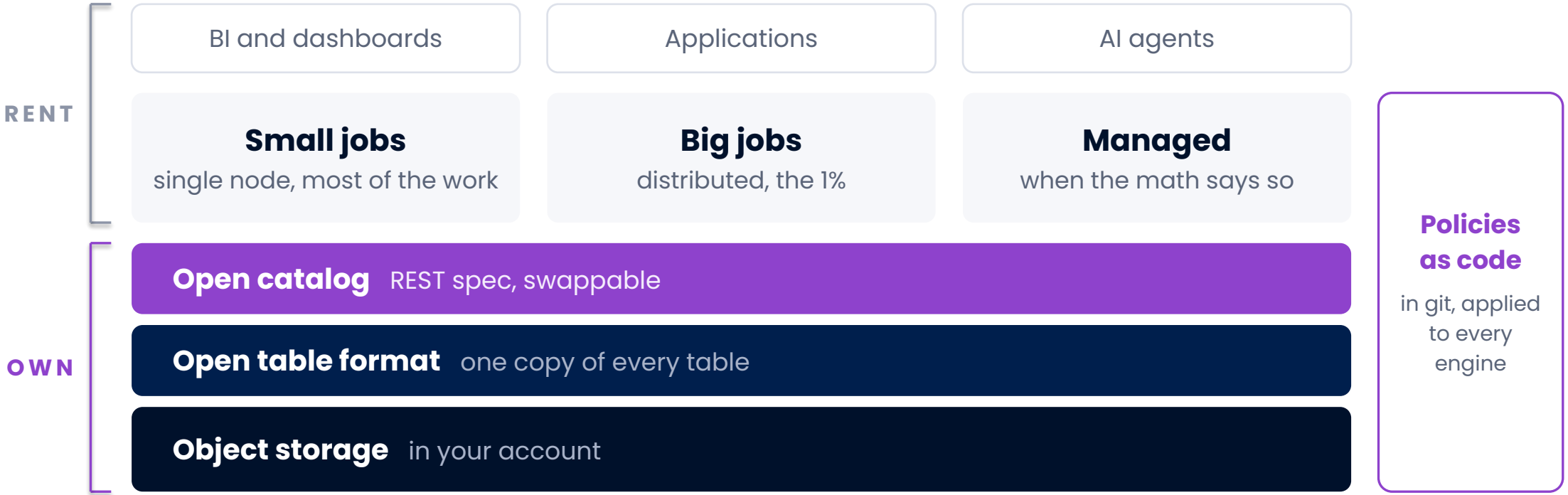
RENT

Cheap to swap

- Engines
- Warehouses
- Brokers
- AI tools

> **Renting is fine. Renting the thing you can't move is the trap.**

The open **reference architecture.**



Every engine reads the same tables. Switching is a choice, not a migration.

Own the land. **Rent the tools.**

01
The trade changed.
Storage is commodity. The markup is on compute.

02
Shop for compute.
Own the bucket. Pick the cheapest engine per job.

03
AI cuts the labor bill.
Agents in the lab. Humans at the gate.

McFadin
DATA AND AI ADVISORY

Patrick McFadin patrick@mcfad.in

Questions?

MCFAD.IN