

数据技术嘉年华

// Data Technology Carnival

开源 · 融合 · 数智化 — 引领数据技术发展 释放数据要素价值

DM openGauss PolarDB PostgreSQL MongoDB DB2 SQLite
OceanBase GreenPlumCassandra MariaDB Hive HBase Teradata

Oracle MySQL SQL Server Redis
OSCAR Claims X-DB iBASE Haisql-memcache
skyTSDb Kingwow TrendDB Cedar DragonBase
PDW HotDB Server OushuDB Gridsum ZETA
TafDB GeminiDB TDengine ArgoDB
MogDB Shentong Megawise TeleDB Sinodb
GreatDB KingDB LongDB ChronusDB RadonDB
UXDB CloudTable TSDB HUABASE HighGoDB
EsqynDB AnalyticDB SequoiADB ArkDB
GoldendB AlisQL CynosDB OpenBASE QuantumDB
MySQL SQL Server RedisTDSQL H2 LevelDB Percona
Oracle RedisDynamoDB GBase Redshift CouchDB
Aurora

LevelDB Percona TBase Kingbase
Sinodb DynamoDB GBase Redshift CouchDB
GreenPlum DM openGauss PolarDB
TIDB Neo4j Informix OceanBase
Aurora TDSQL H2 Memcached Sybase HANA
Cassandra MariaDB Hive HBase Teradata
PostgreSQL MongoDB DB2 SQLite

CynosDB TenBASE QuantumDB Ando
Timesten K-DB GoldendB AlisQL
UABASE HighGoDB Huayisoft HashDB
HybridDB Kunda GreatDB KingDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB

LongDB ChronusDB
RadonDB CloudTable TSDB CTSD
EsqynDB AnalyticDB SequoiADB
ArkDB



中国DBA联盟
All China DBA Union



墨天轮

Cost-Intelligent Data Analytics in the Cloud

Huanchen Zhang



清华大学
Tsinghua University



交叉信息研究院
Institute for Interdisciplinary
Information Sciences

Why Are Cloud Databases Different?

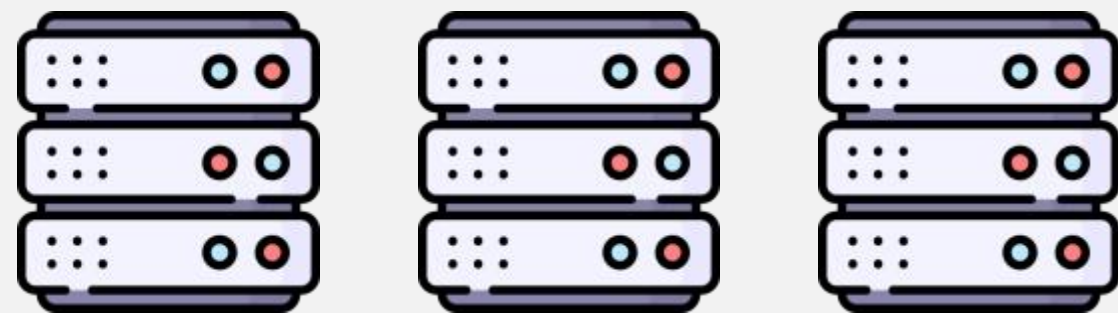
Traditional

0110
1001
1010

Why Are Cloud Databases Different?

Traditional

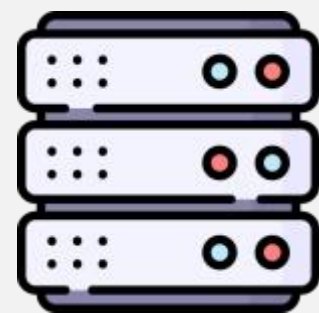
0110
1001
1010



Why Are Cloud Databases Different?

Traditional

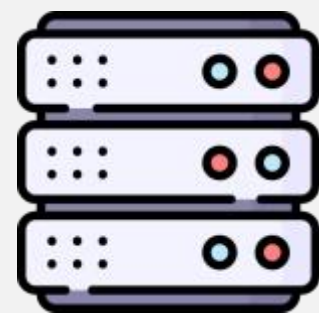
0110
1001
1010



Why Are Cloud Databases Different?

Traditional

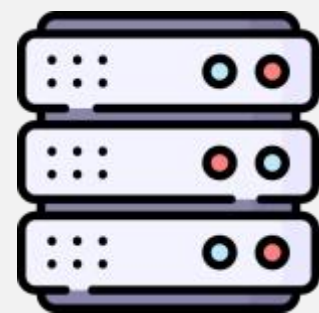
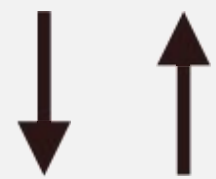
0110
1001
1010



Why Are Cloud Databases Different?

Traditional

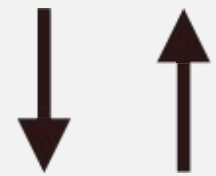
0110
1001
1010



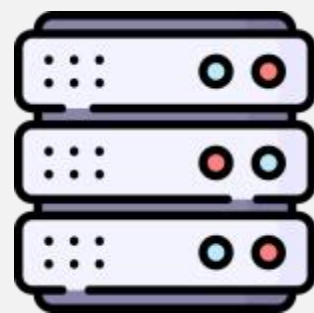
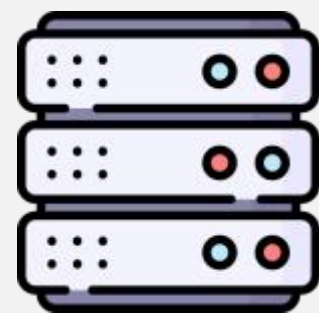
Why Are Cloud Databases Different?

Traditional

0110
1001
1010



\$\$\$

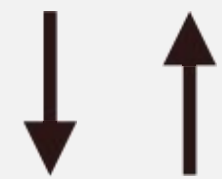


\$\$\$

Why Are Cloud Databases Different?

Traditional

0110
1001
1010



\$\$\$



\$\$\$

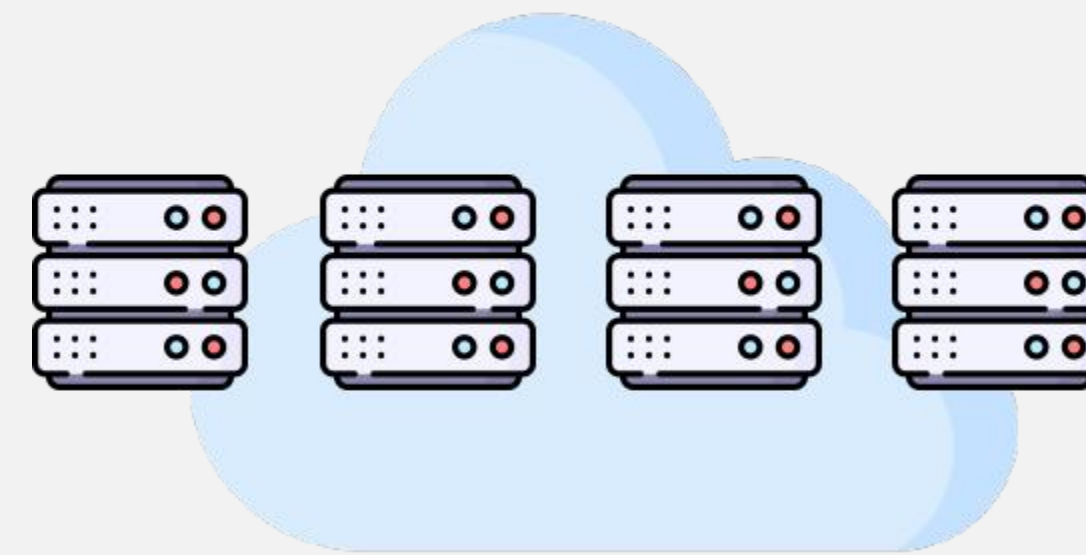
Cloud-Native

0110
1001
1010



\$

pay-as-you-go



...

Elasticity

Database Optimization in the Cloud Era

User Profit

□

Database Optimization in the Cloud Era

User Profit

Utility

\$ Cost

$$\Pi = U(p) - C$$

Database Optimization in the Cloud Era

$$\begin{array}{l} \text{User Profit} \\ \Pi \end{array} = \begin{array}{l} \text{Utility} \\ U(p) \end{array} - \begin{array}{l} \$ \text{ Cost} \\ C \end{array}$$

Traditional $C_{\text{sun}k} + C_e$

Fix, Large

Database Optimization in the Cloud Era

Traditional

$$\text{User Profit } \Pi = \text{Utility } U(p) - \text{\$ Cost } C$$

p t $C_{\text{sun}k} + C_e$

Fix, Large

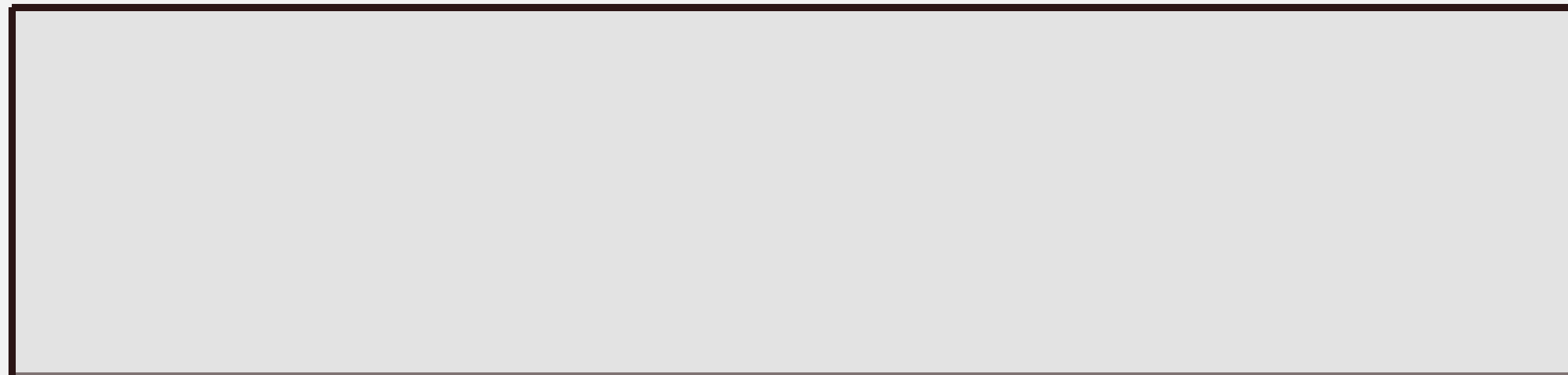
Database Optimization in the Cloud Era

	User Profit	Utility	\$ Cost
	Π	$= U(p)$	$- C$
Traditional		$p \cdot t$	$C_{sunk} + C_e$ <div> Fix, Large </div>
Cloud-Native		p	ΔC <div> Enabled by Elasticity </div>

Database Optimization in the Cloud Era

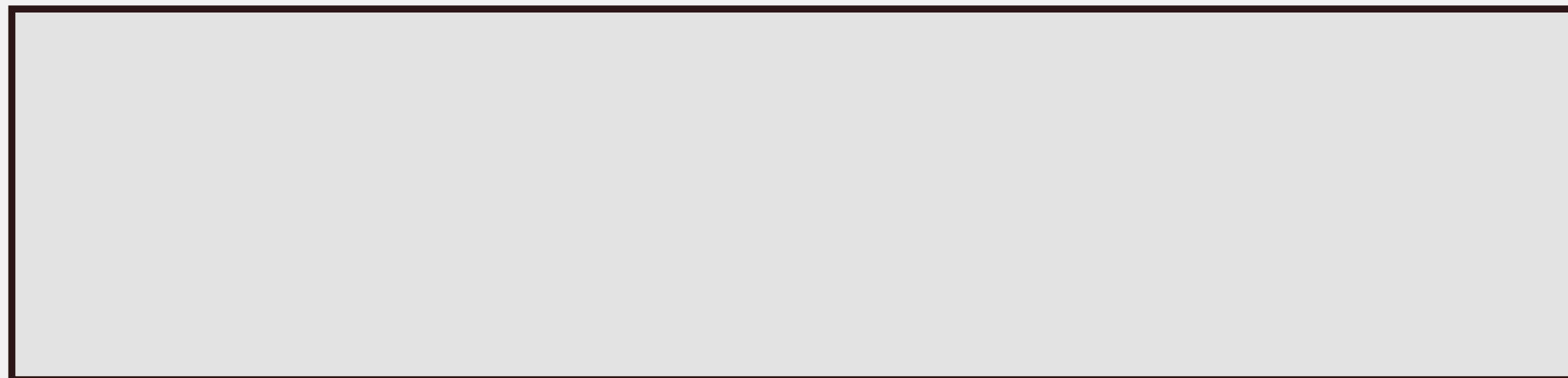
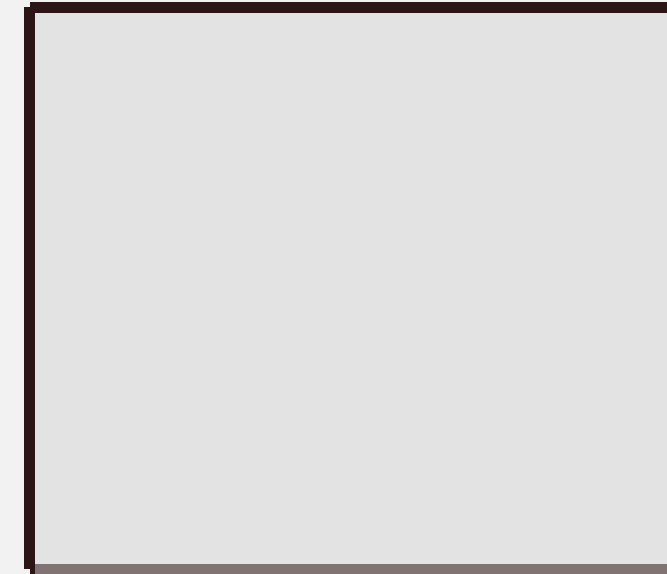
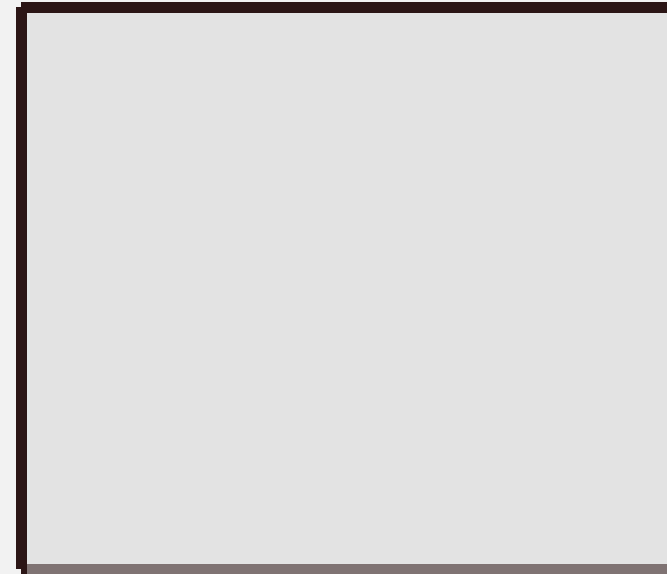
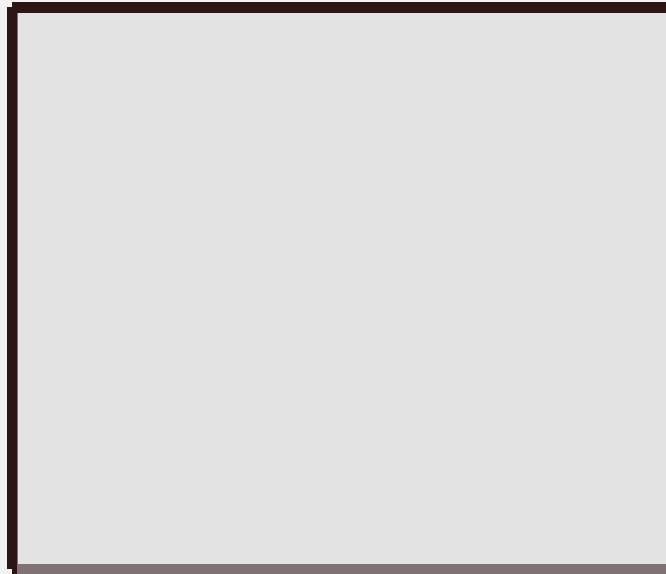
	User Profit	Utility	\$ Cost
	Π	$=$	$U(p) - C$
Traditional		$p \quad t$	$C_{\text{fixed}} + C_e$ <div> Fix, Large </div>
Cloud-Native		p	ΔC <div> Enabled by Elasticity </div>
		Bi-Objective Optimization	

Cost From Different Perspectives



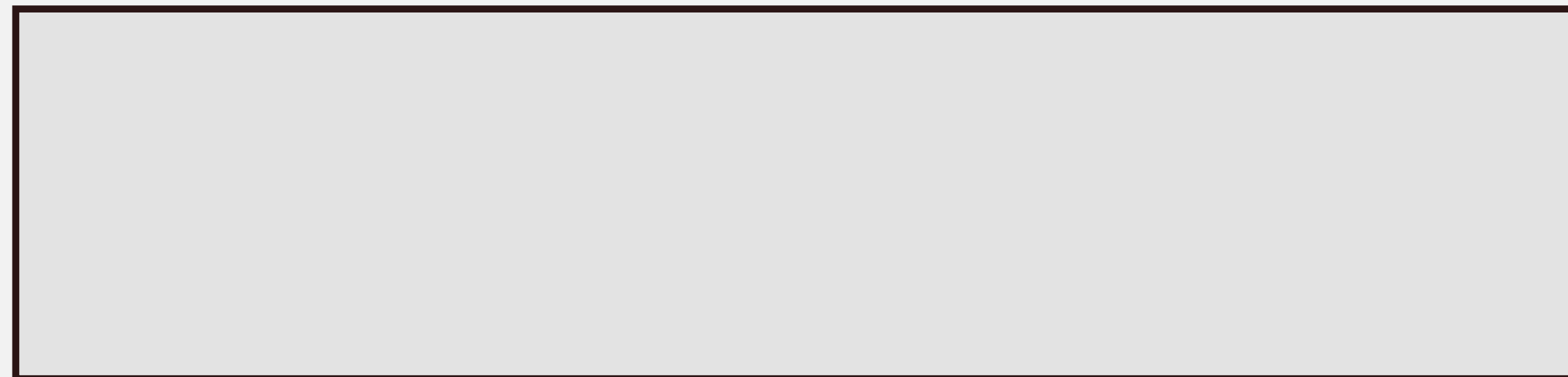
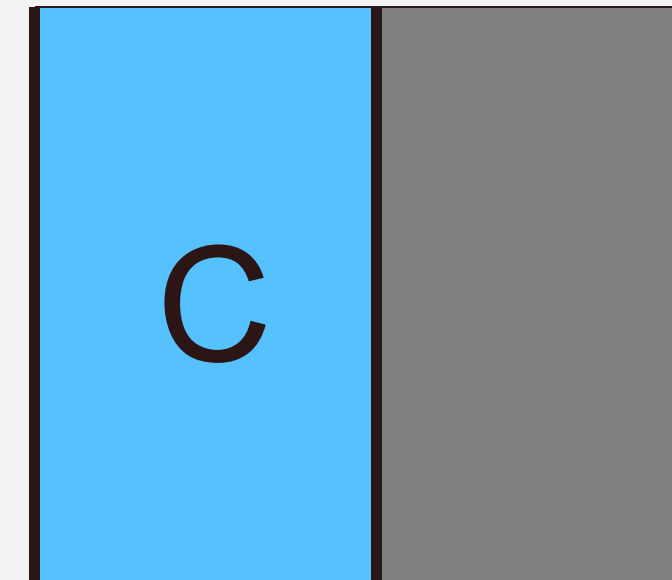
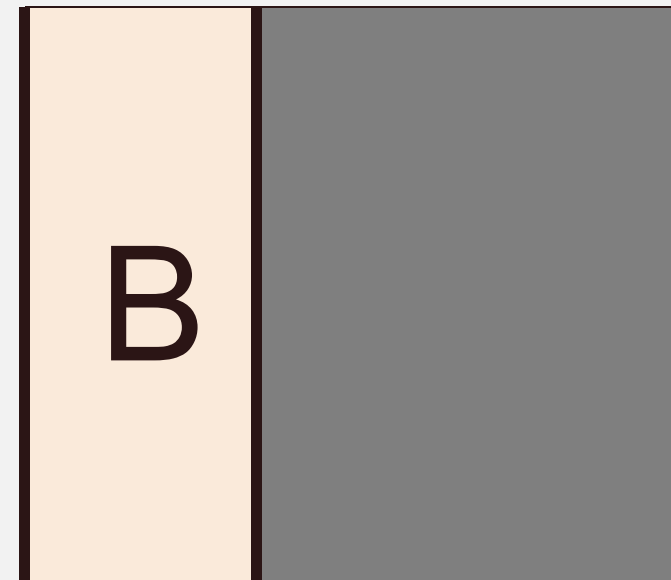
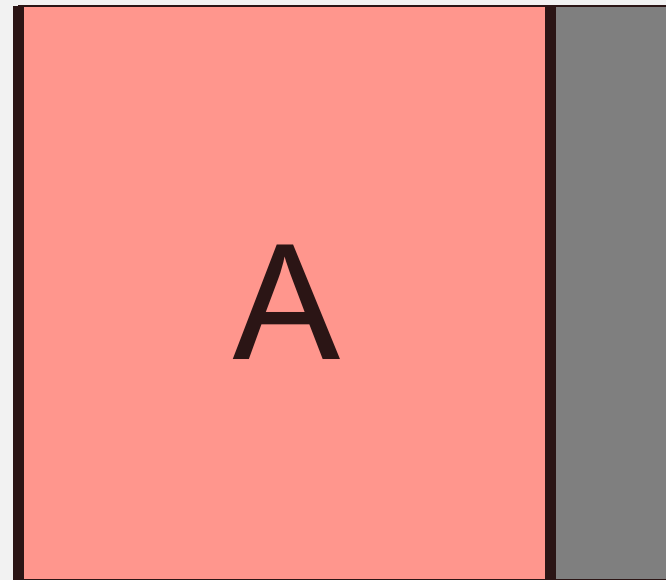
Resource Pool

Cost From Different Perspectives



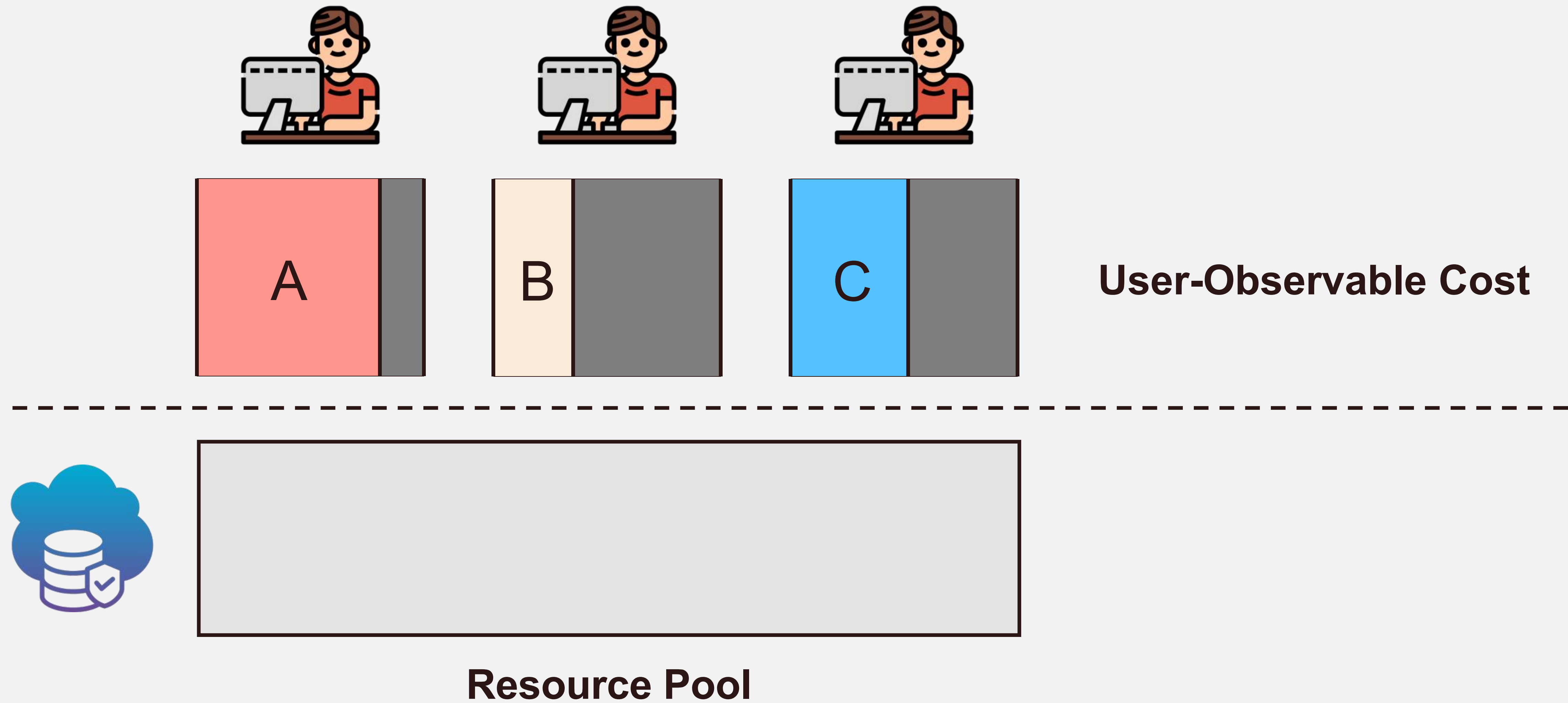
Resource Pool

Cost From Different Perspectives

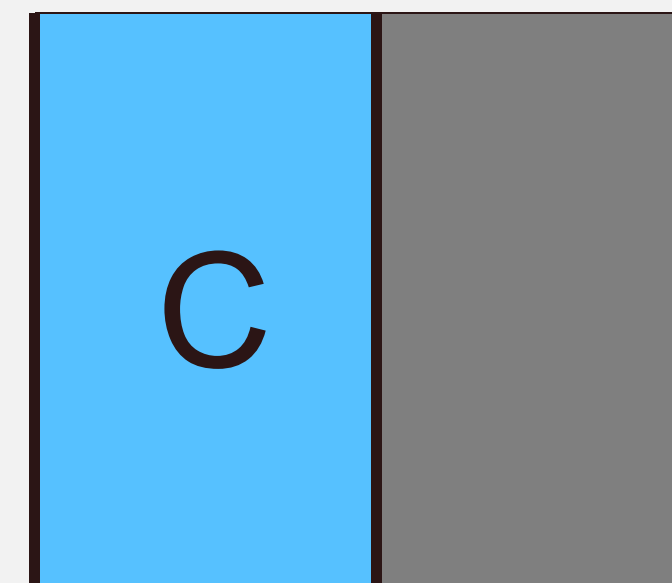
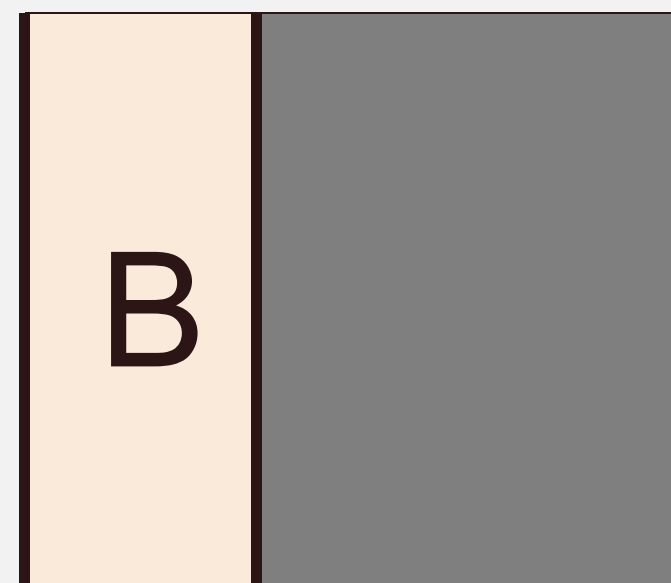
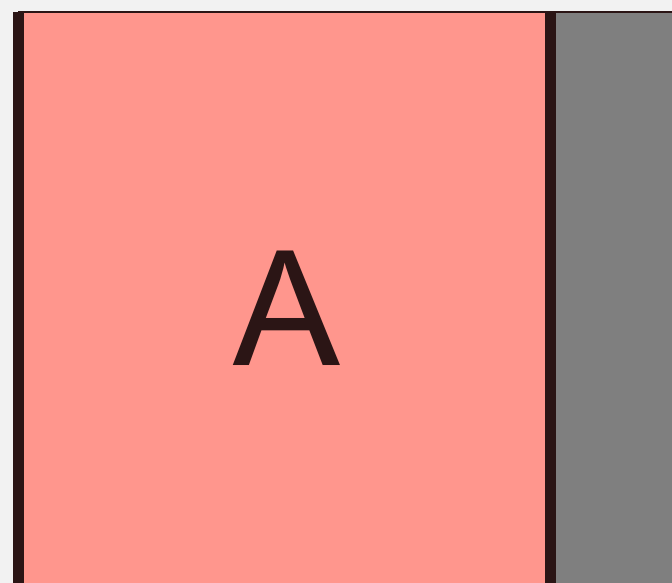


Resource Pool

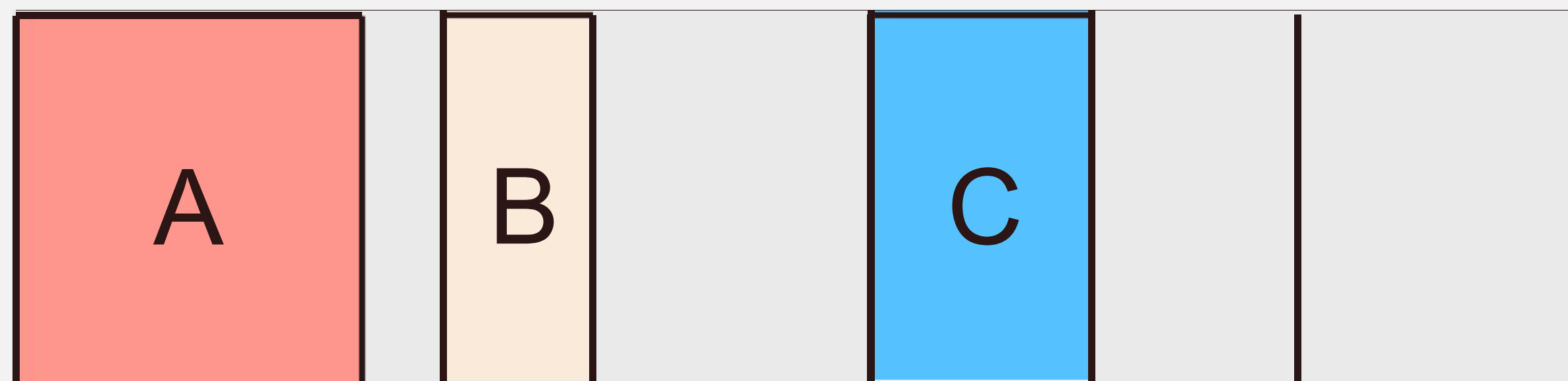
Cost From Different Perspectives



Cost From Different Perspectives

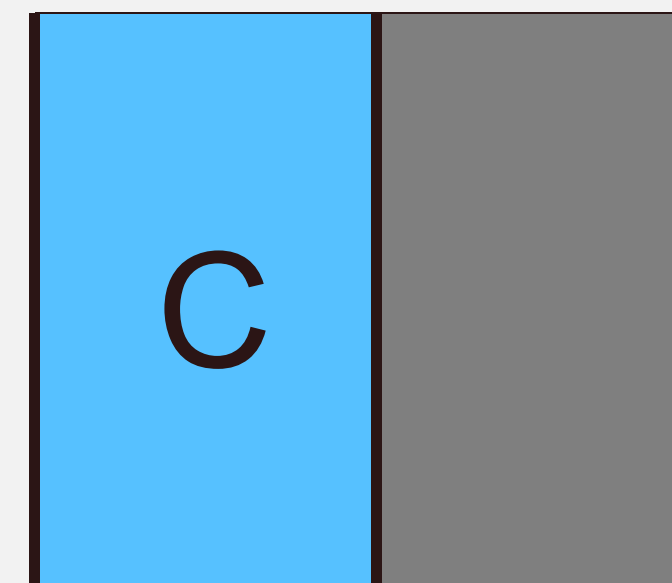
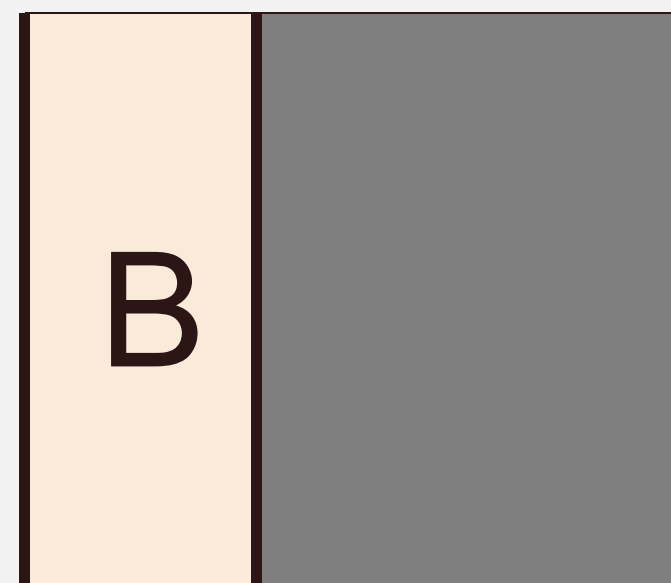
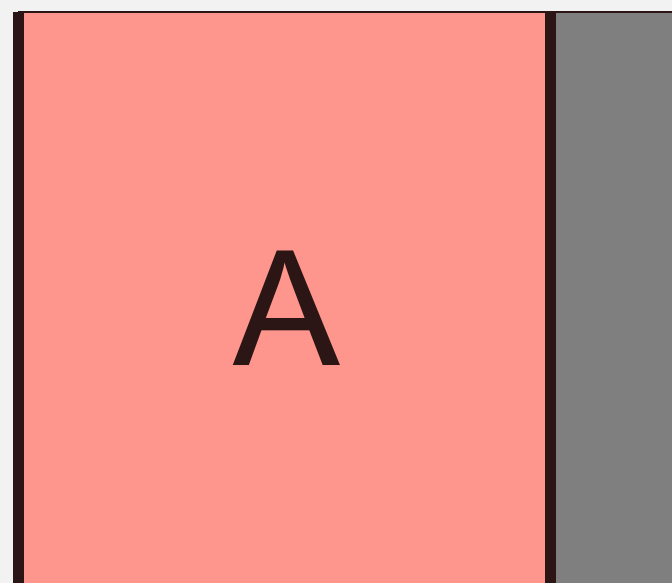


User-Observable Cost

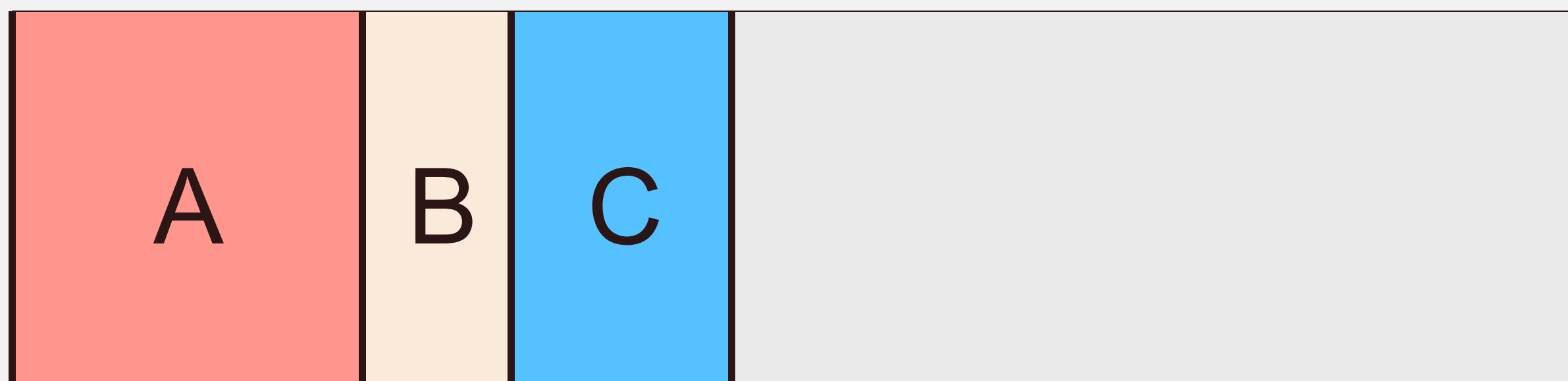


Resource Pool

Cost From Different Perspectives

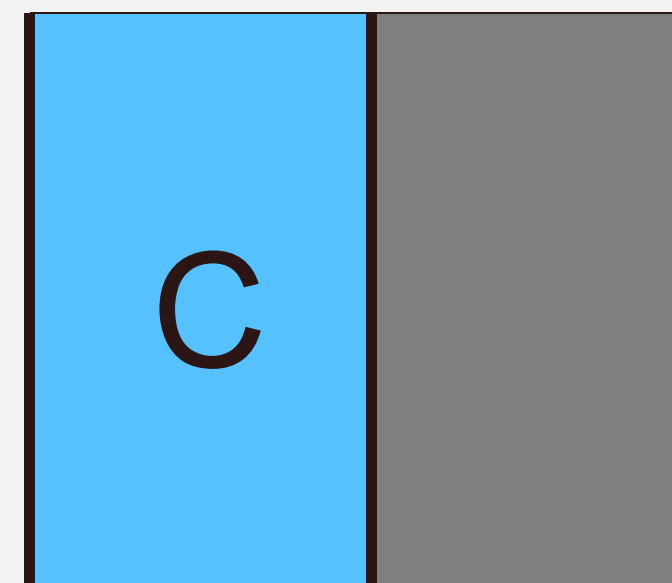
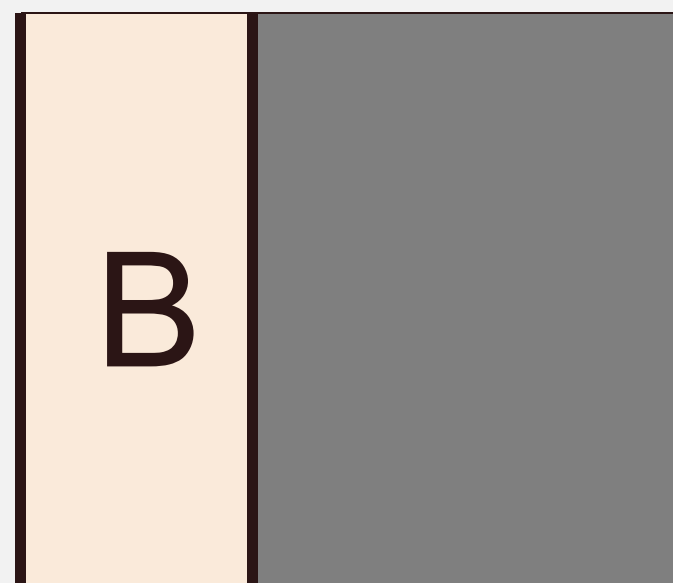
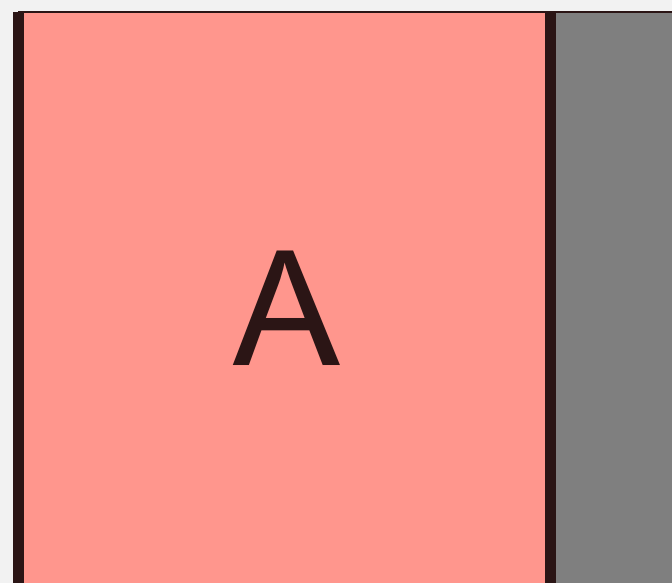


User-Observable Cost

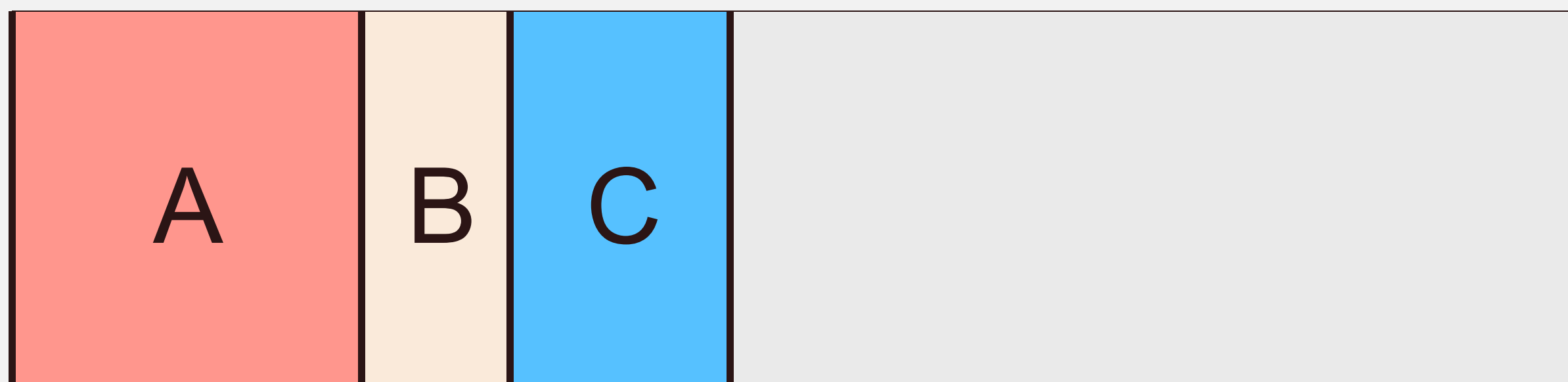


Resource Pool

Cost From Different Perspectives



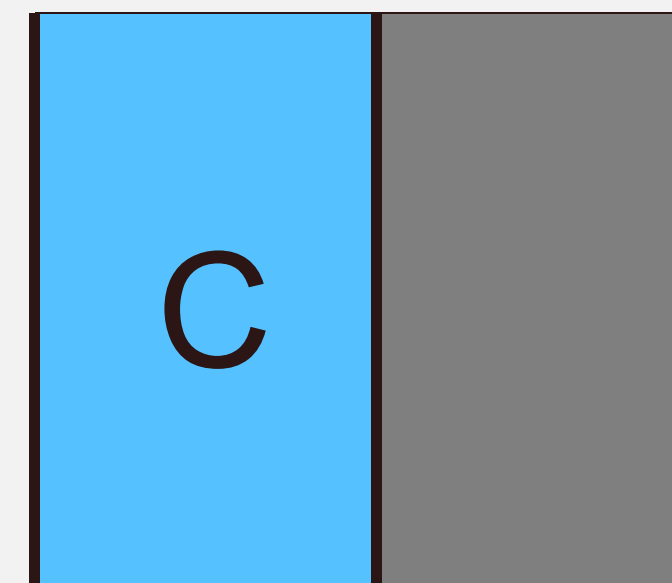
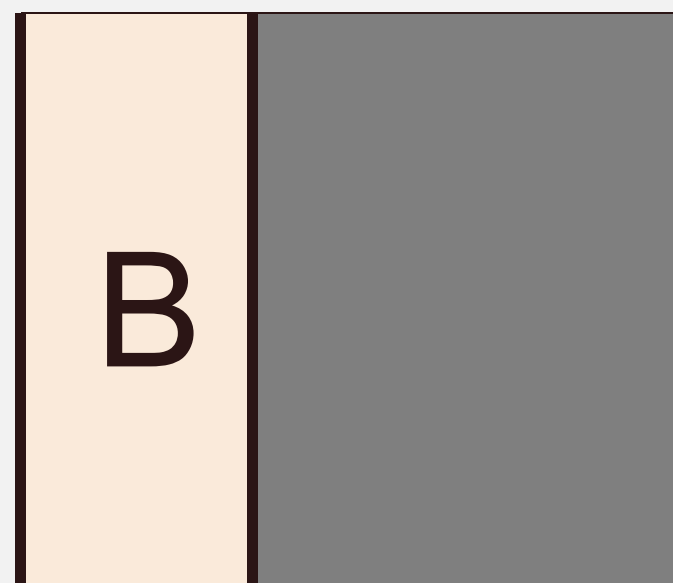
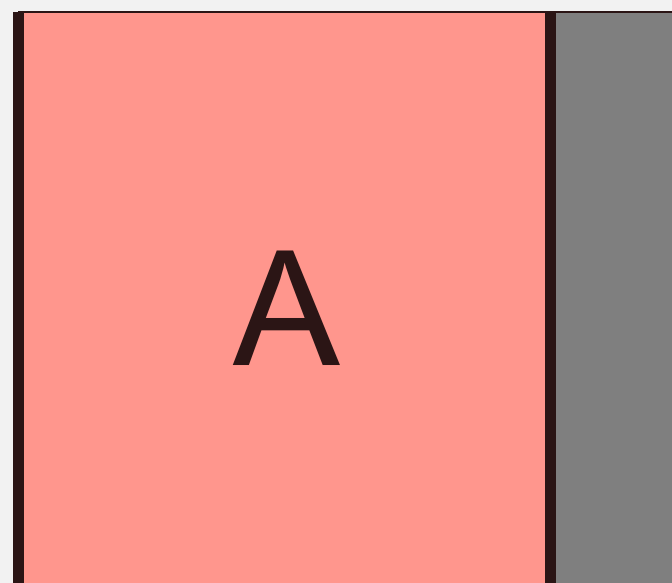
User-Observable Cost



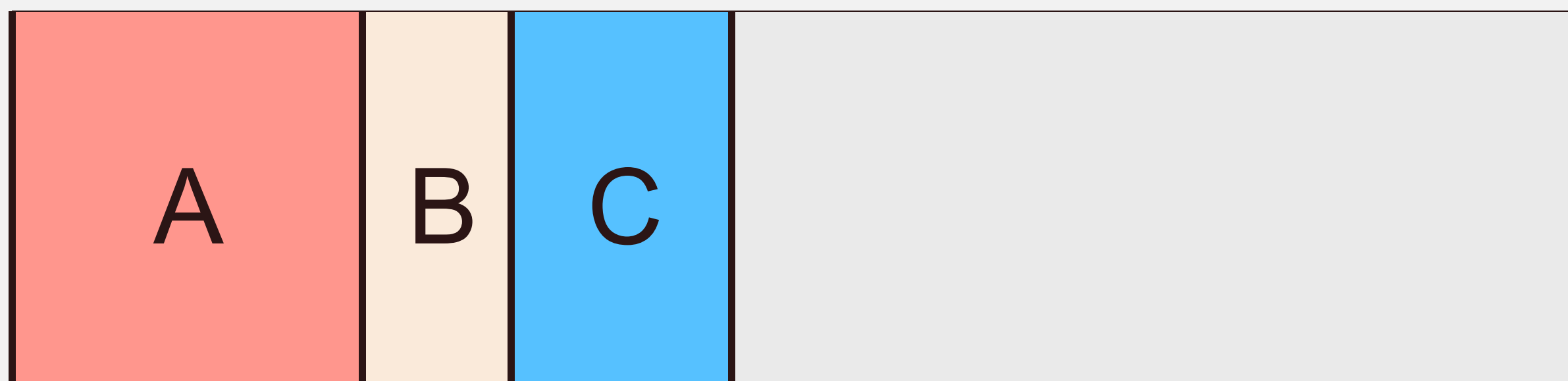
Provider-Observable Cost

Resource Pool

Cost From Different Perspectives



User-Observable Cost

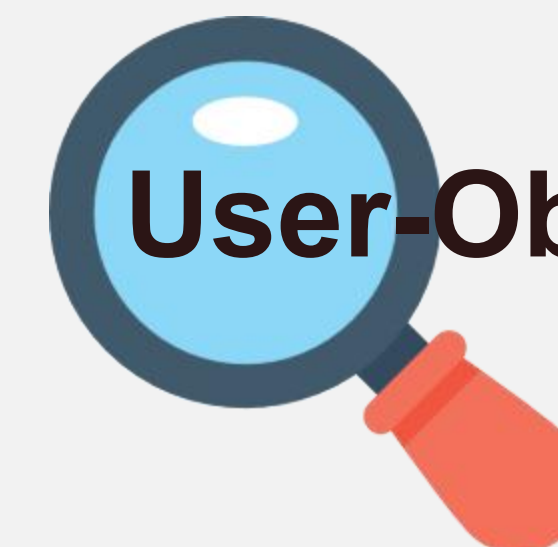
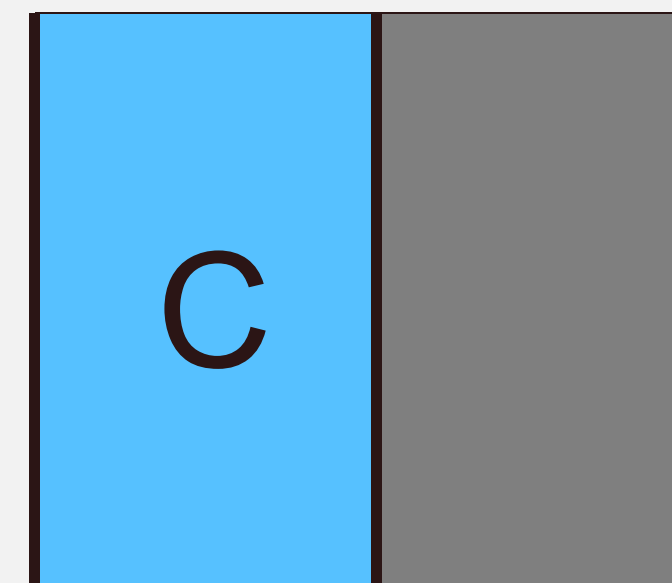
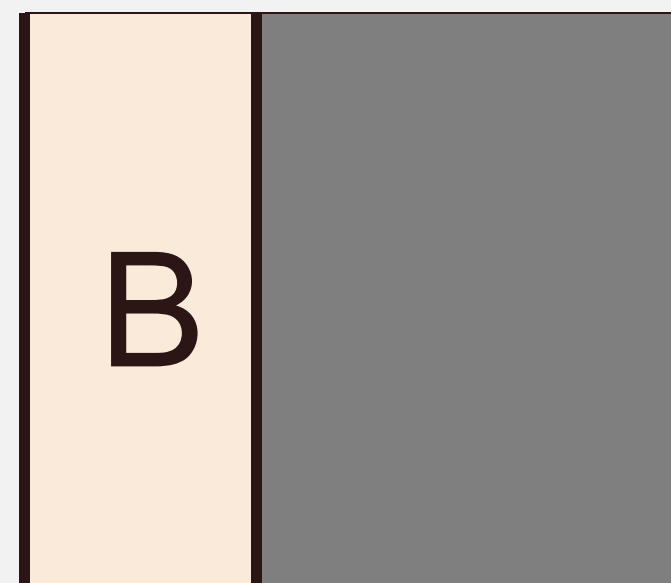
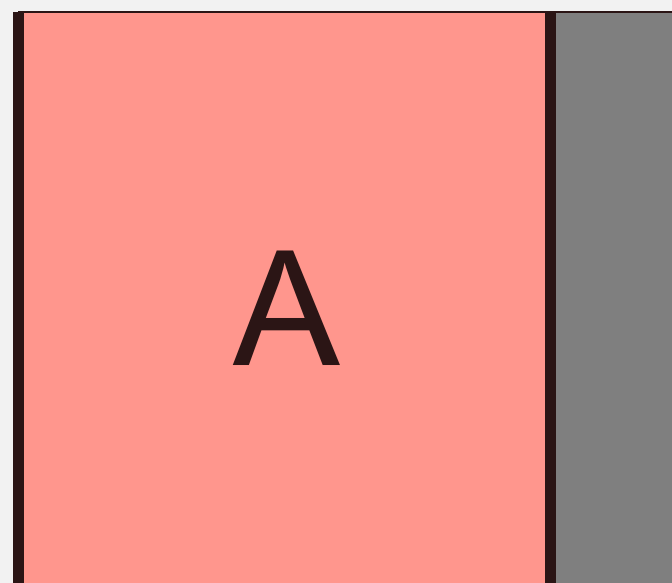


Provider-Observable Cost

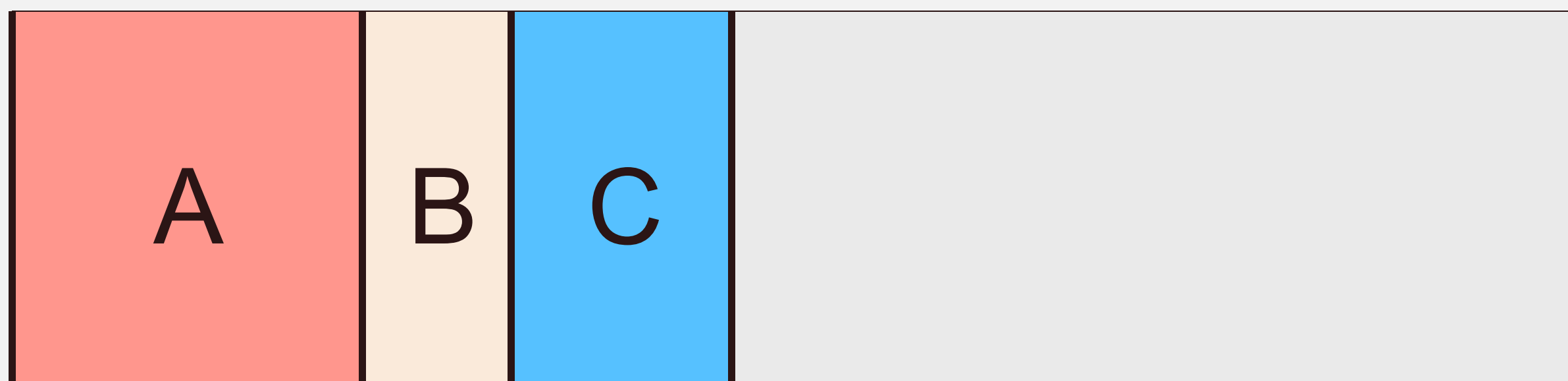
[Multi-Tenancy Techniques]

Resource Pool

Cost From Different Perspectives



User-Observable Cost



Provider-Observable Cost

[Multi-Tenancy Techniques]

Resource Pool

Cost Control Is Still Difficult

Workload



Cost Control Is Still Difficult

Workload



Choose your cluster size



1 server A



2 server A



4 server A

8 server A

1 server B

2 server B

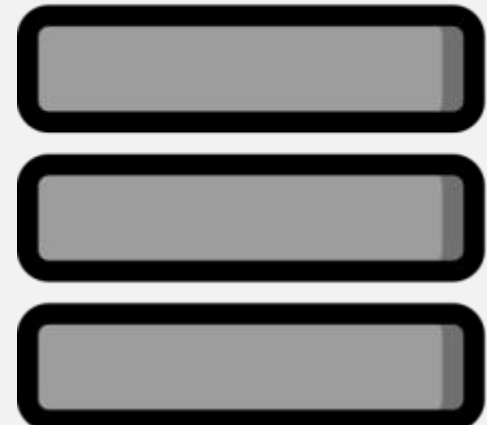
⋮

Cost Control Is Still Difficult

Workload



Choose your cluster size



- 1 server A
- 2 server A
- 4 server A
- 8 server A
- 1 server B
- 2 server B
- ⋮

→ Fixed cluster size the entire workload

Cost Control Is Still Difficult

Workload



Choose your cluster size



1 server A

2 server A

4 server A

8 server A

1 server B

2 server B

⋮

→ Fixed cluster size the entire workload

→ Users tend to over-provision

Cost Control Is Still Difficult

Workload



Choose your cluster size



1 server A



2 server A



4 server A

8 server A

1 server B

2 server B

⋮

→ Fixed cluster size the entire workload

→ Users tend to over-provision

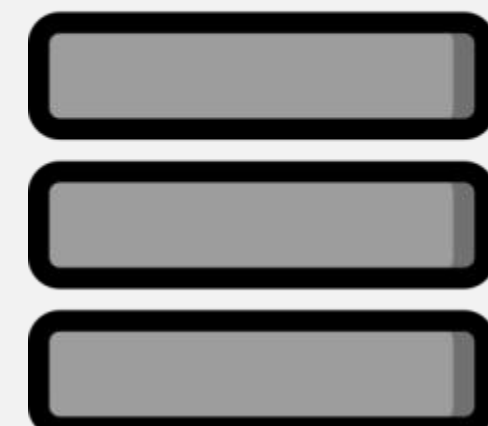
Resource Waste!

Cost Control Is Still Difficult

Workload



Choose your cluster size








- 1 server A
- 2 server A
- 4 server A
- 8 server A
- 1 server B
- 2 server B

⋮

DBA



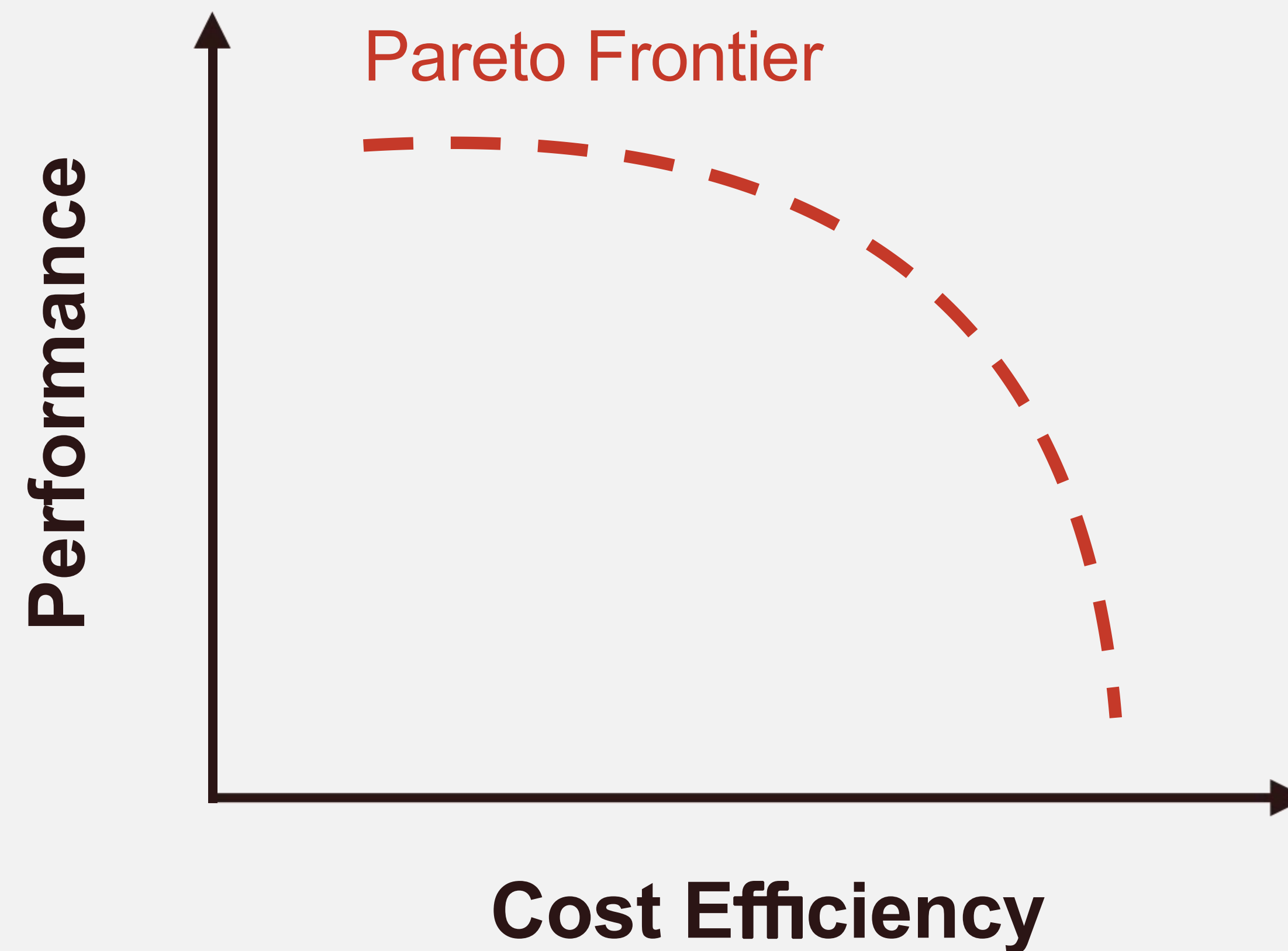
\$\$\$

-  Build Indexes
-  Build Materialized Views
-  Re-partition Data
-  Change Data Format
-  Re-train a Learned Module

Cost Intelligence



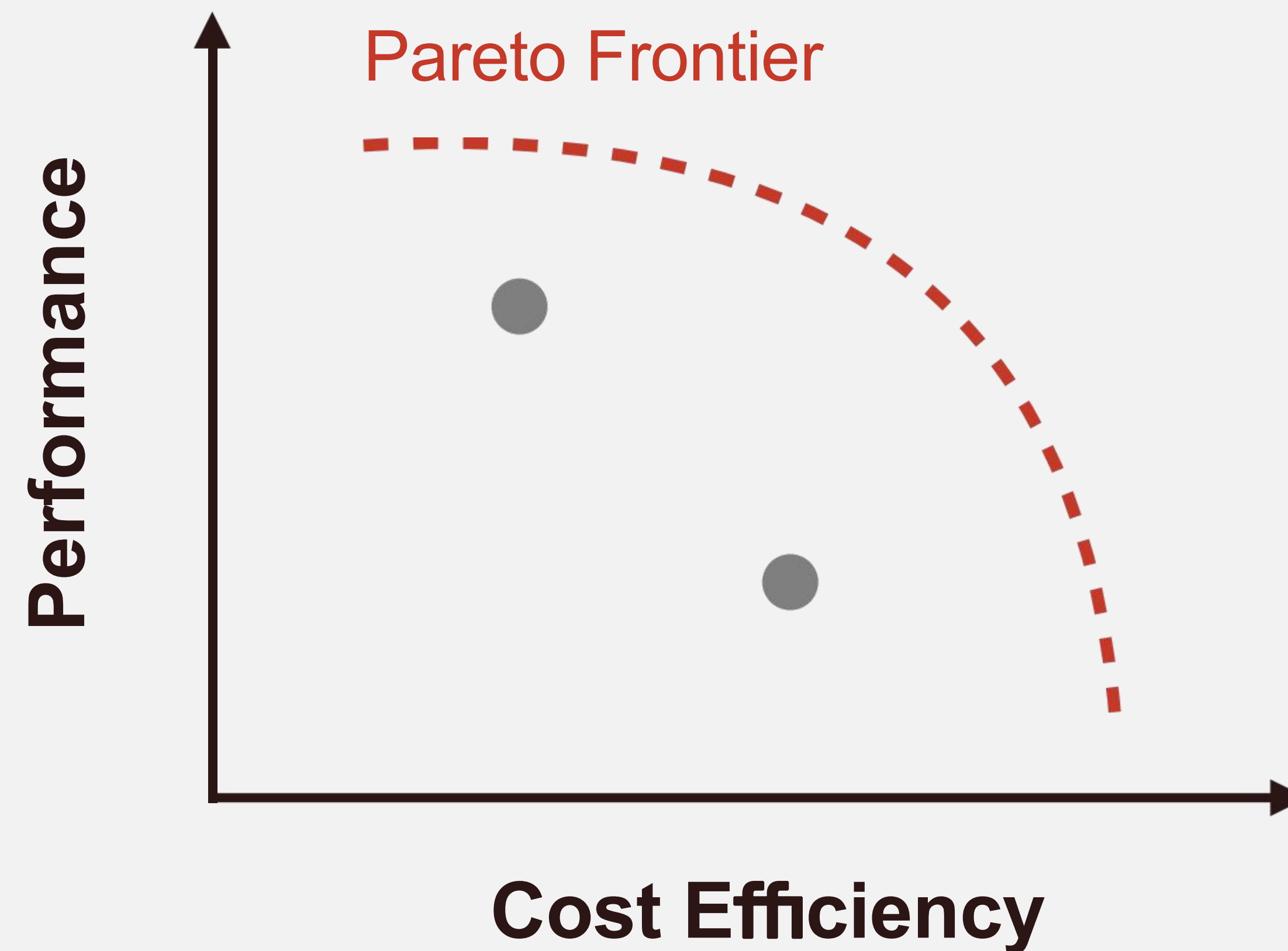
The system's ability to **self-adapt** to stay at **Pareto-optimal** in the performance-cost trade-of under diferent workloads and user constraints.



Cost Intelligence



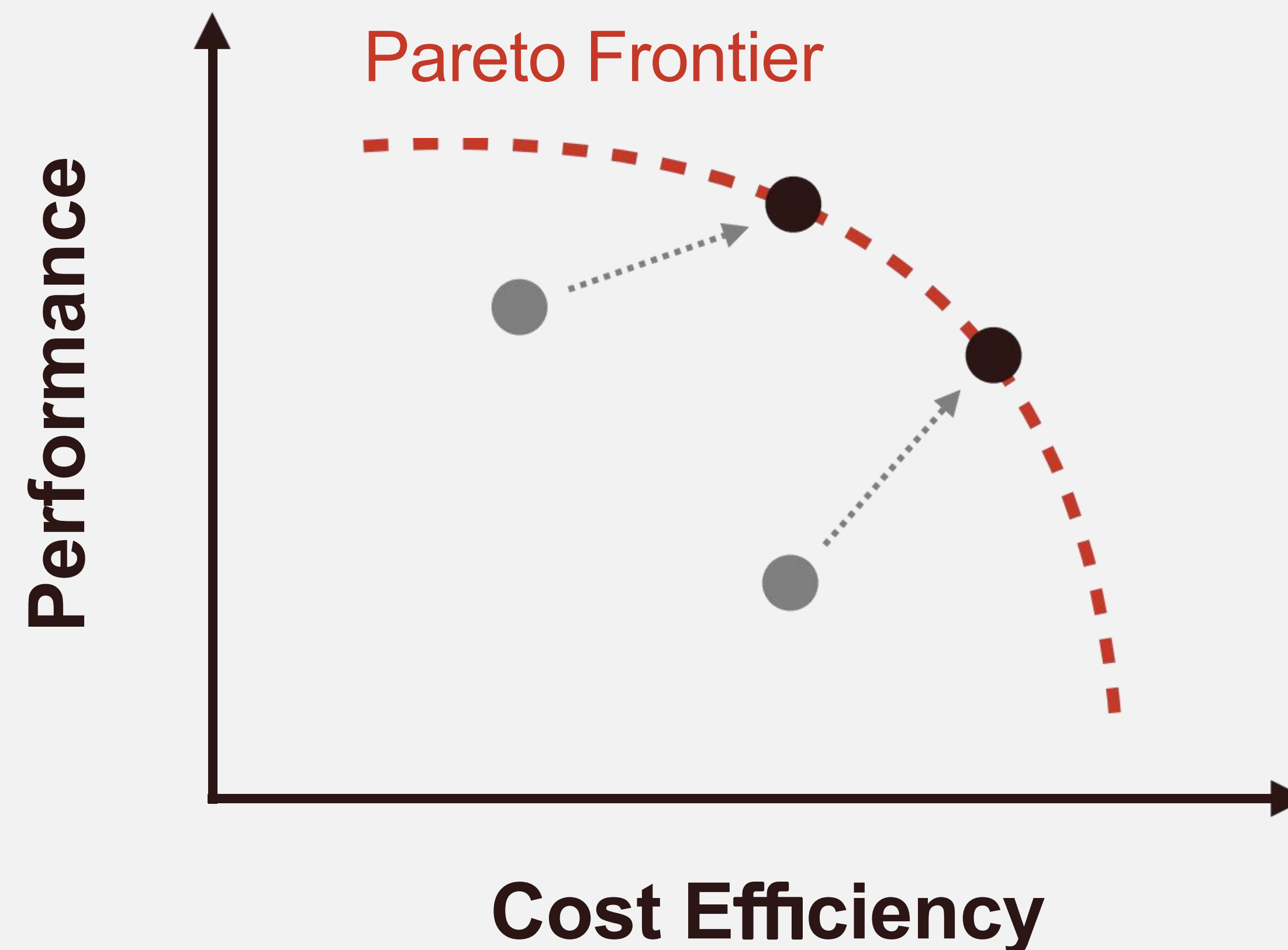
The system's ability to **self-adapt** to stay at **Pareto-optimal** in the performance-cost trade-off under different workloads and user constraints.



Cost Intelligence



The system's ability to **self-adapt** to stay at **Pareto-optimal** in the performance-cost trade-of under diferent workloads and user constraints.



Interface of a Cost-Intelligent System

Workload



Choose your cluster size



1 server A



2 server A



4 server A

8 server A

1 server B

2 server B

⋮

DBA



\$\$\$



Build Indexes



Build Materialized Views



Re-partition Data

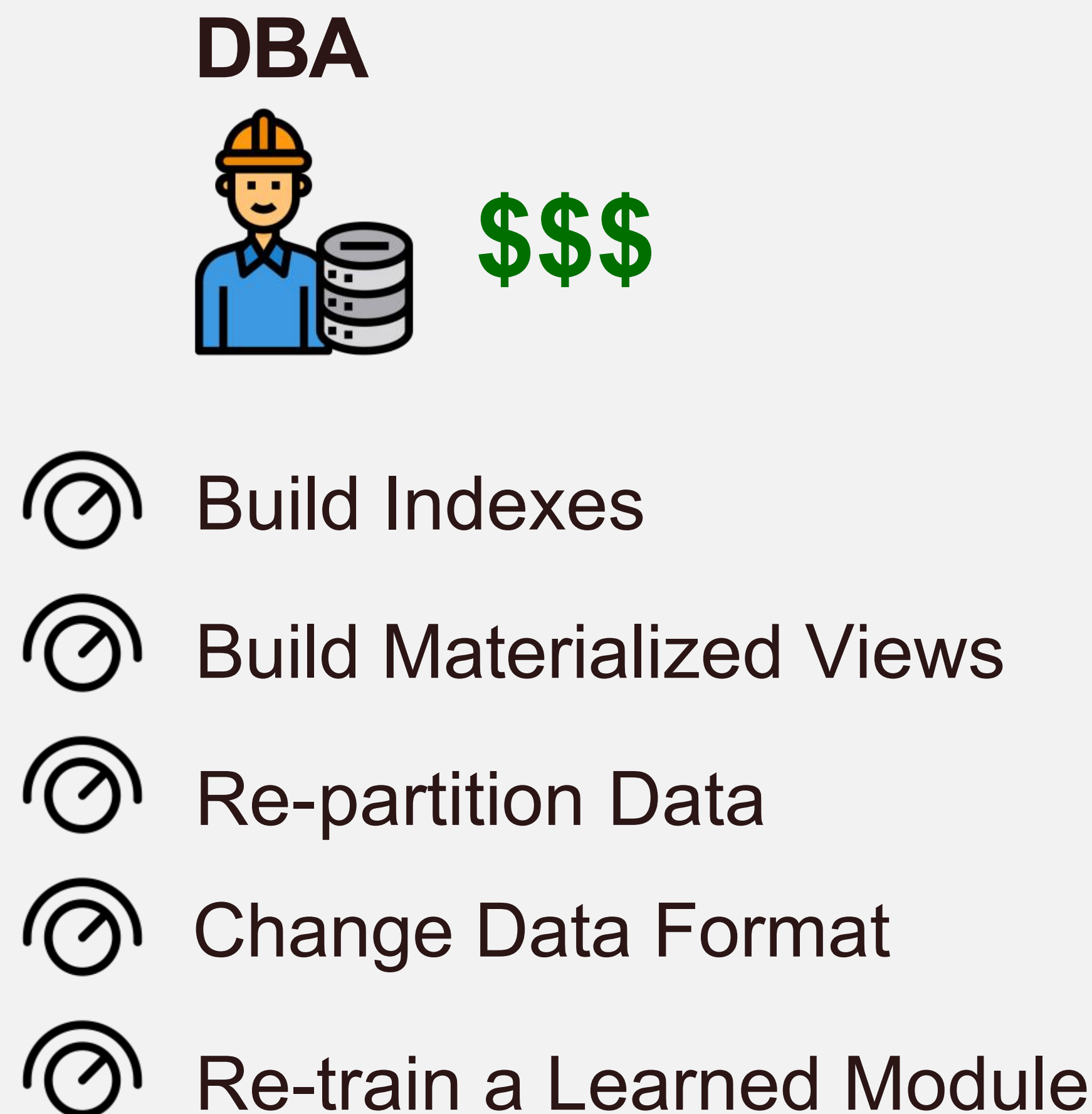
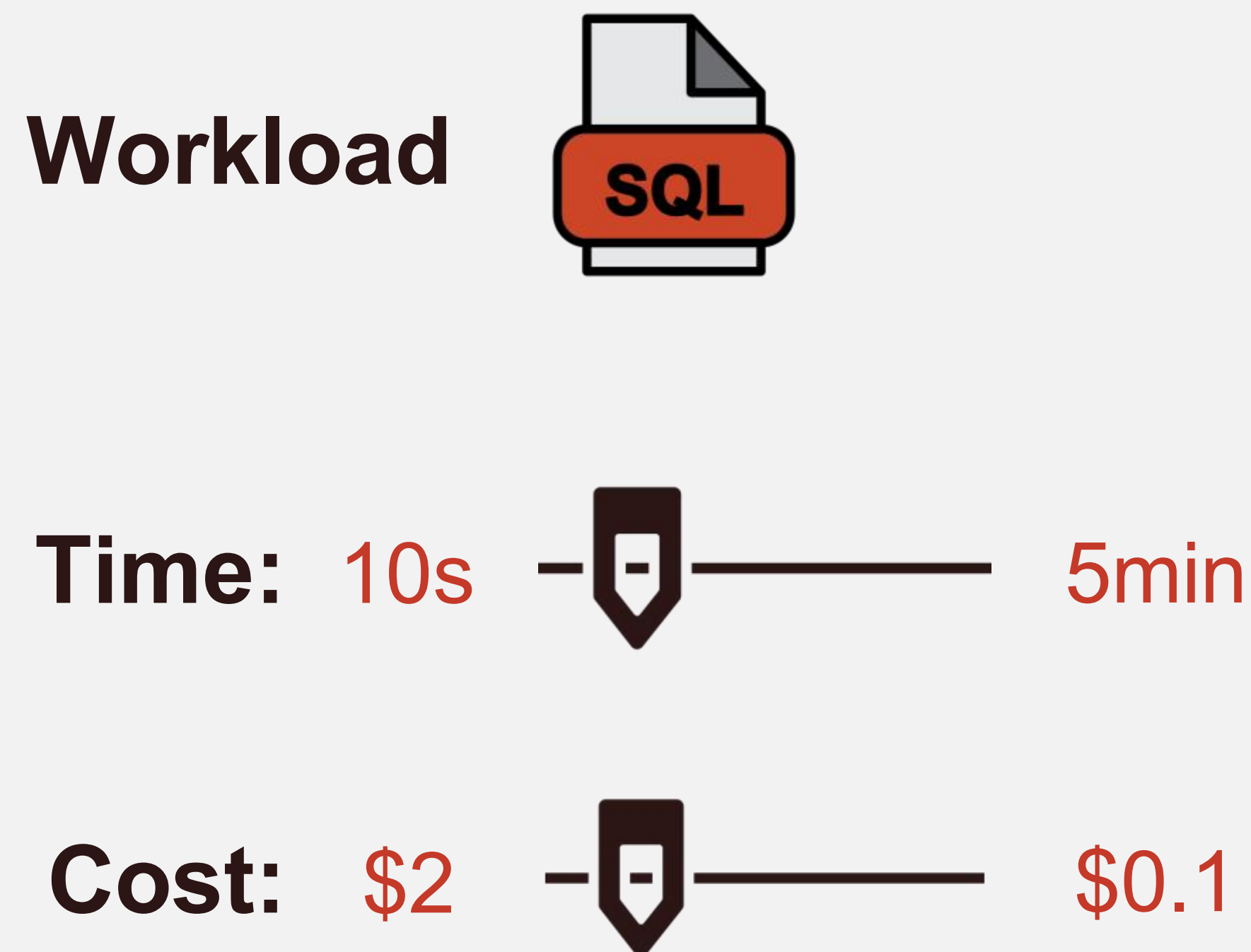


Change Data Format

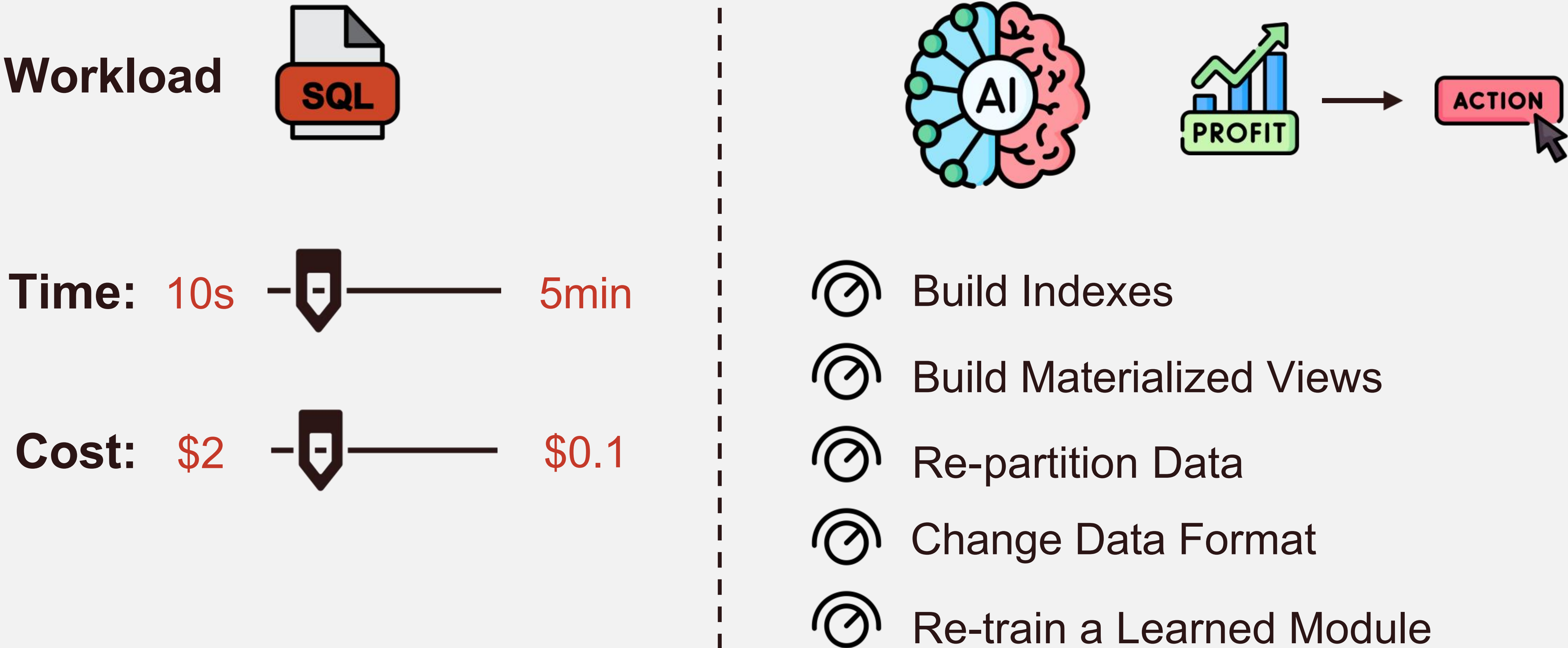


Re-train a Learned Module

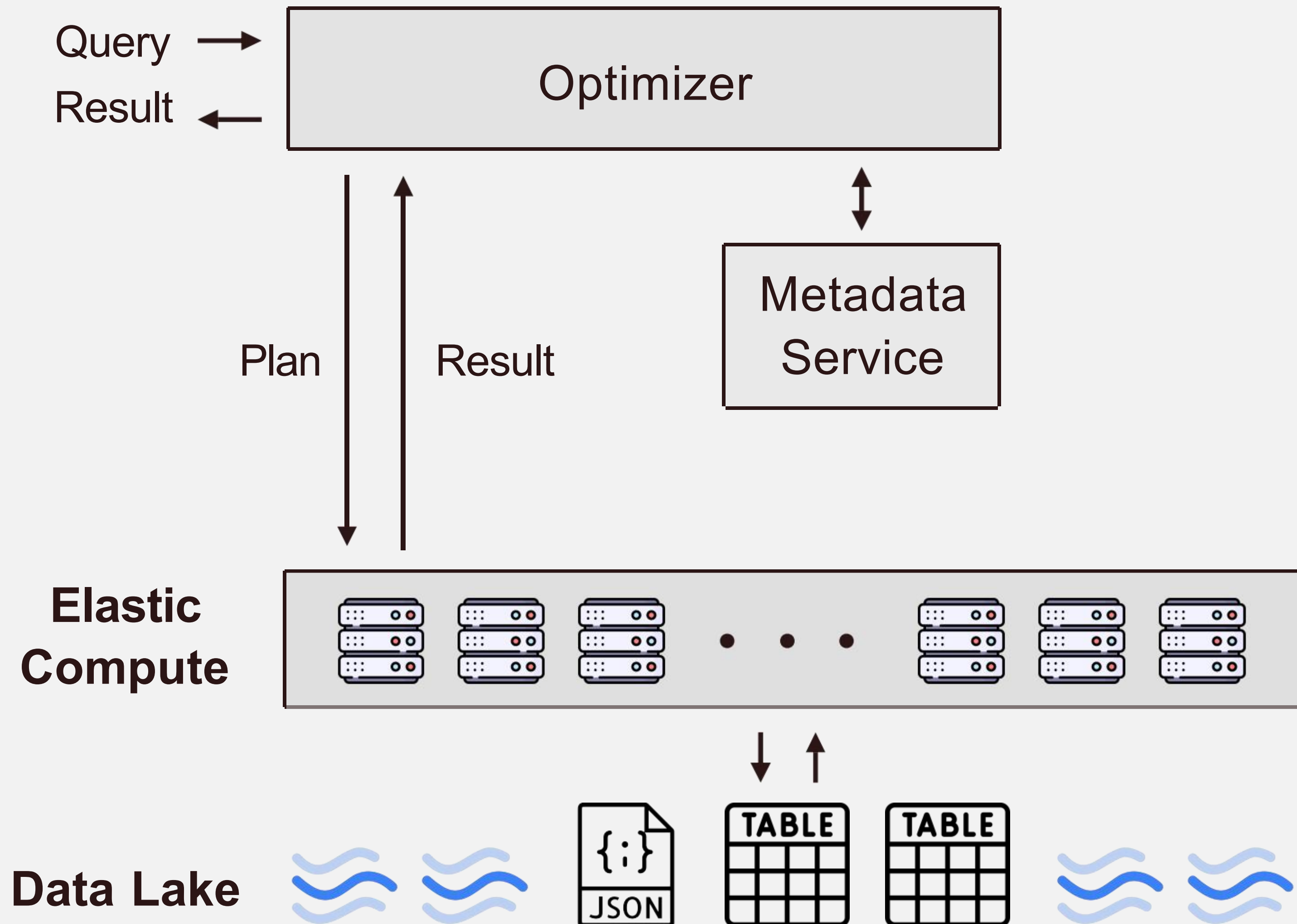
Interface of a Cost-Intelligent System



Interface of a Cost-Intelligent System



Base System Architecture



Automatic Resource Deployment

Workload



Automatic Resource Deployment

Workload



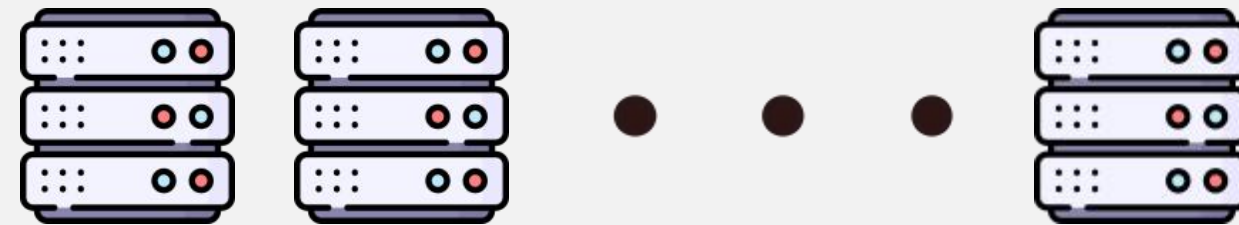
Config 1



×

100 min

Config 2



×

1 min

100 servers

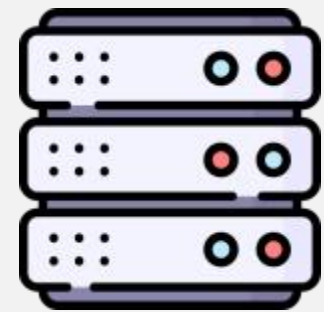
Same \$ Cost

Automatic Resource Deployment

Workload

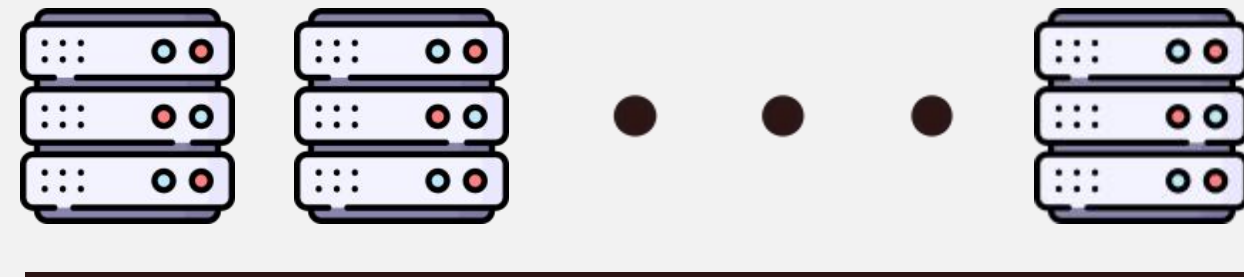


Config 1



× 100 min

Config 2



× 1 min

100 servers

Same \$ Cost

100x performance boost!



Automatic Resource Deployment

Workload

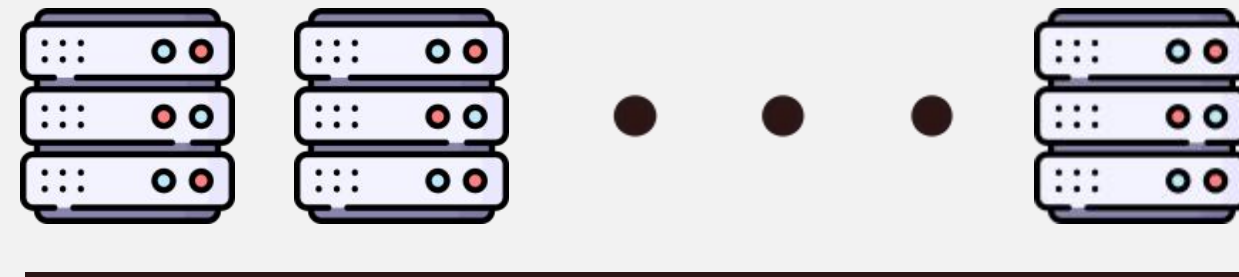


Config 1



100 min

Config 2



100 servers



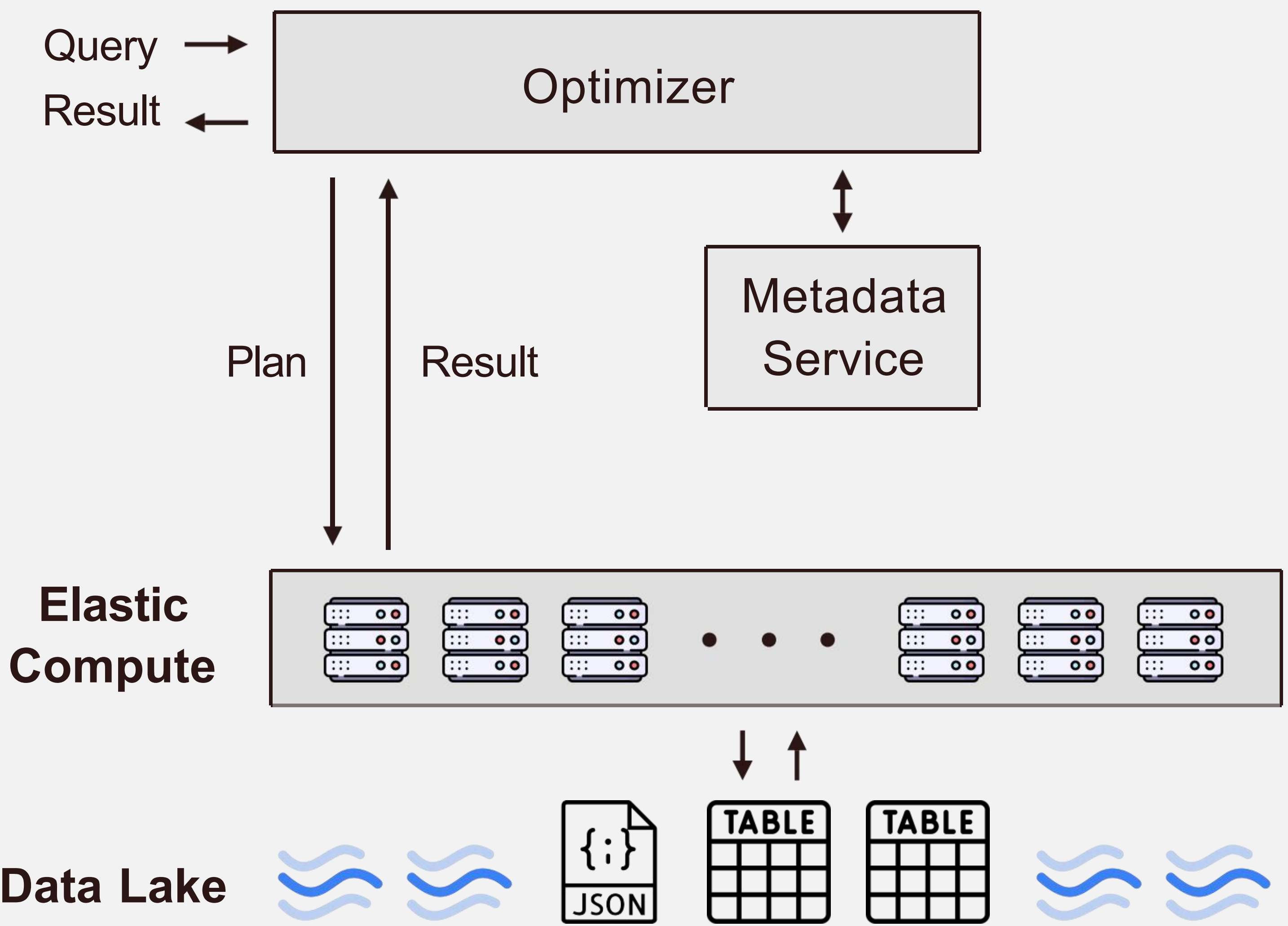
100 min



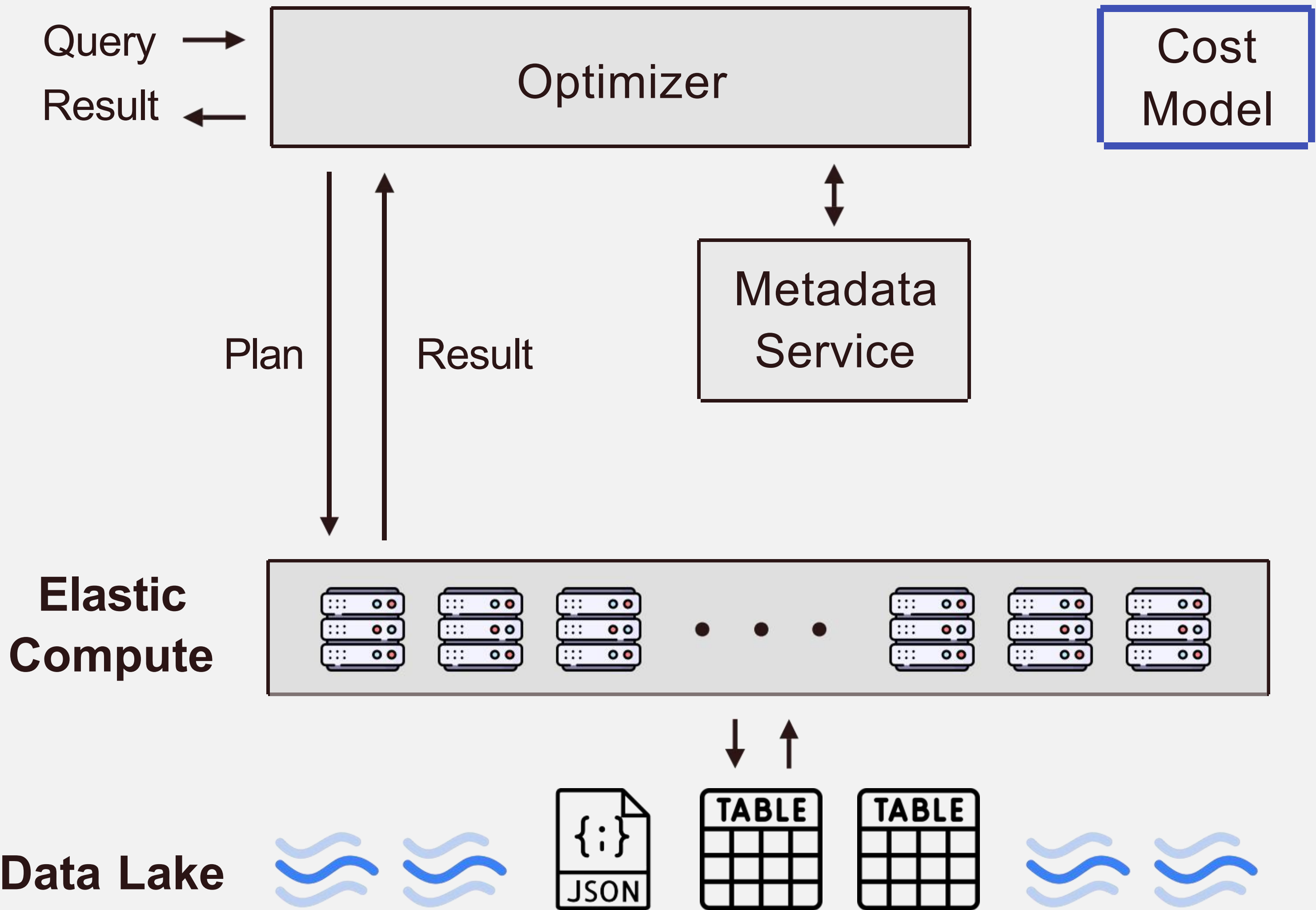
100x \$ Cost

Same performance

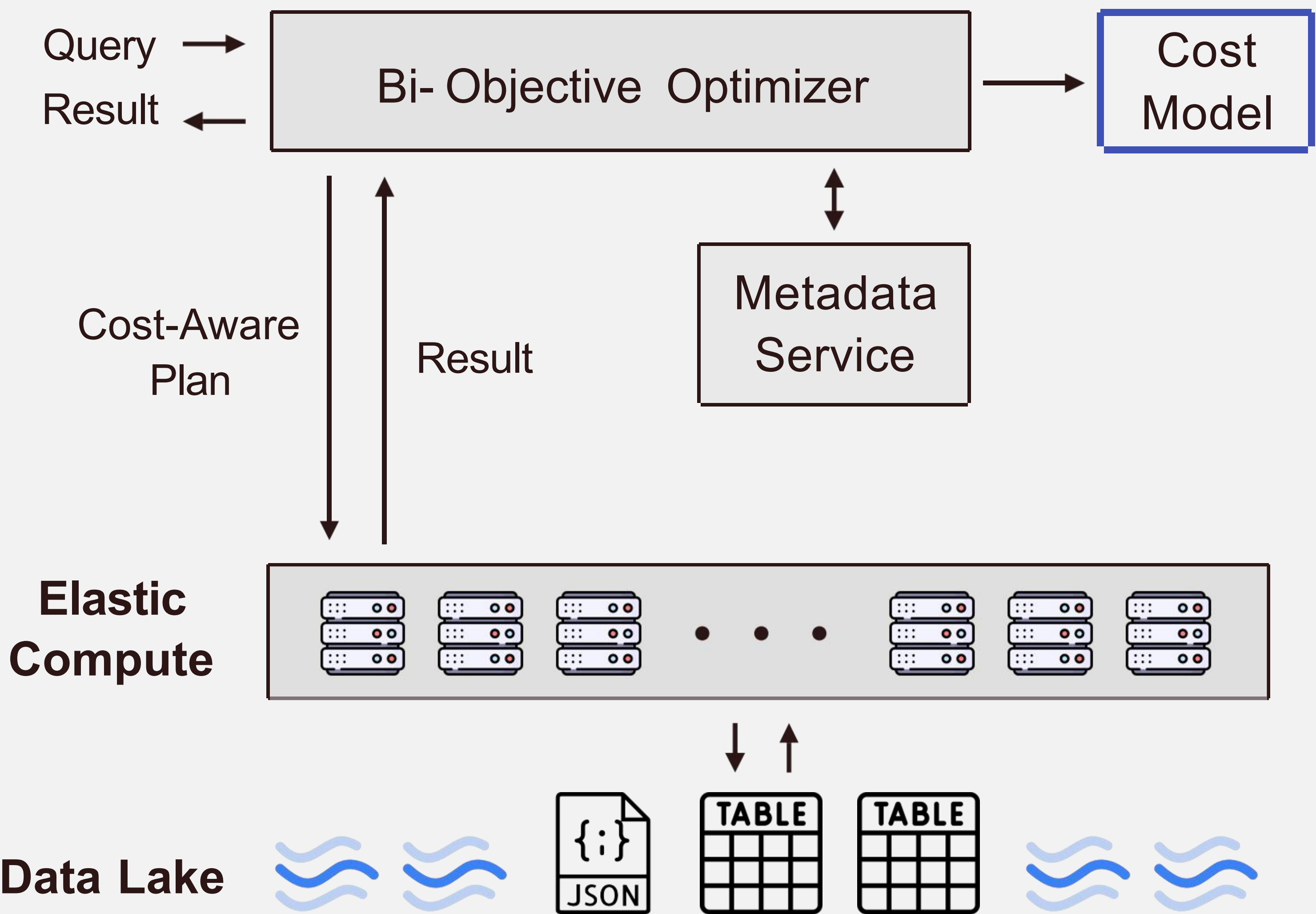
System Architecture



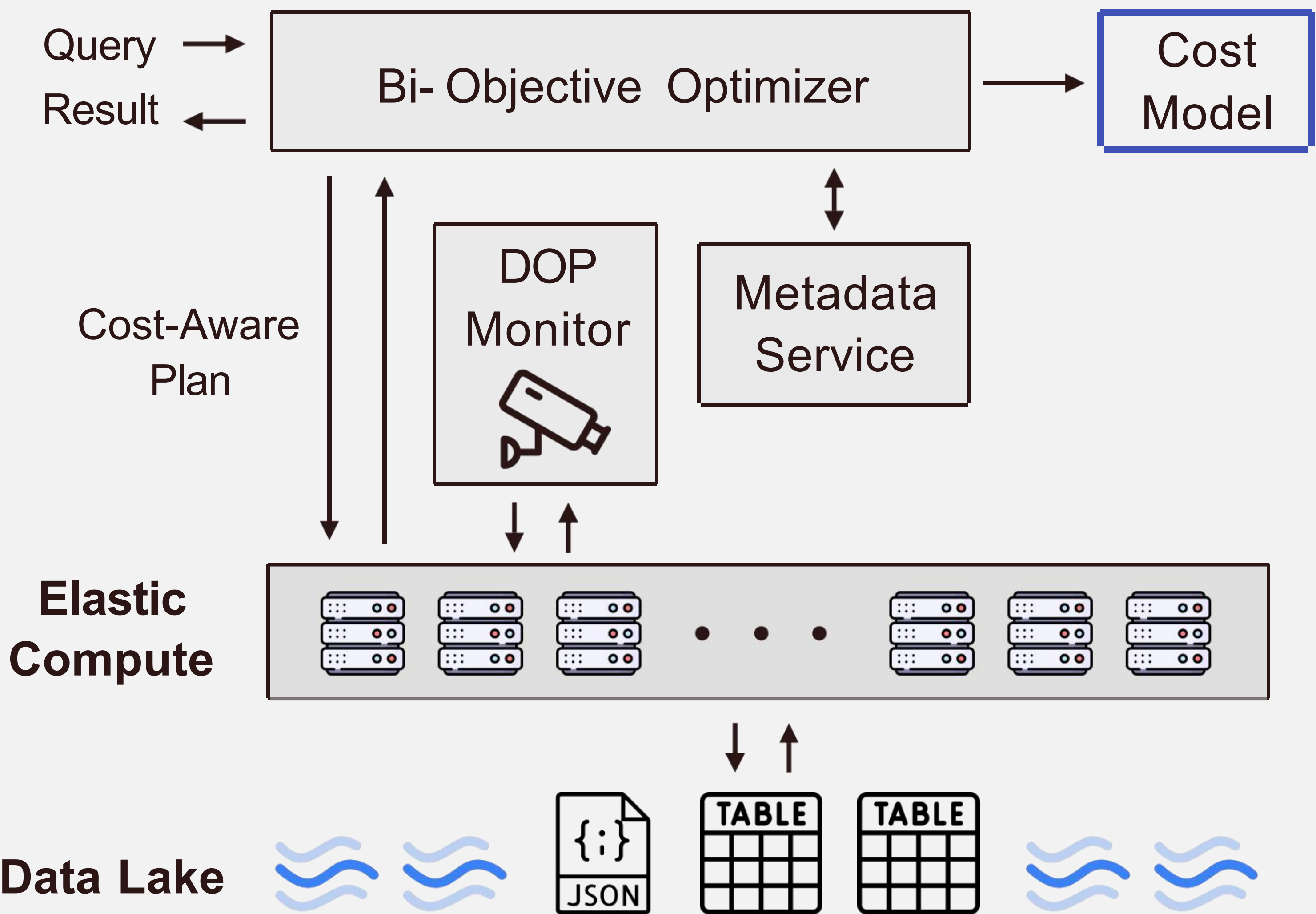
System Architecture



System Architecture



System Architecture



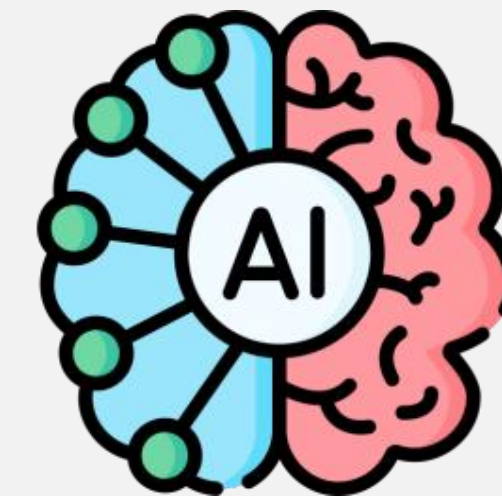
Cost-Oriented Database Auto-Tuning

- 🕒 Build Indexes
- 🕒 Build Materialized Views
- 🕒 Re-partition Data
- 🕒 Change Data Format
- 🕒 Re-train a Learned Module

DBA

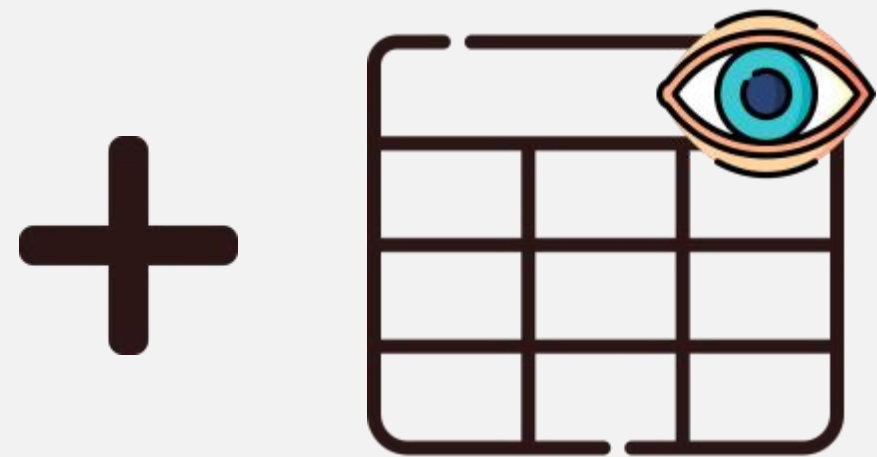


?



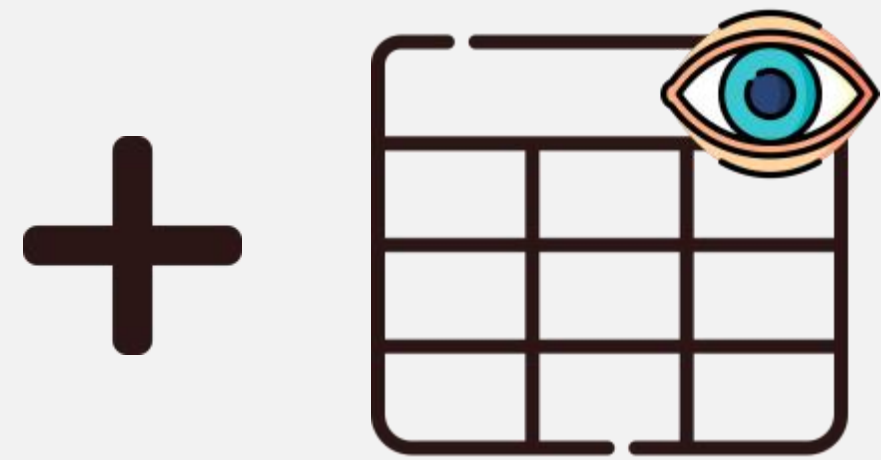
Cost-Oriented Database Auto-Tuning

Under Fix Resources



Cost-Oriented Database Auto-Tuning

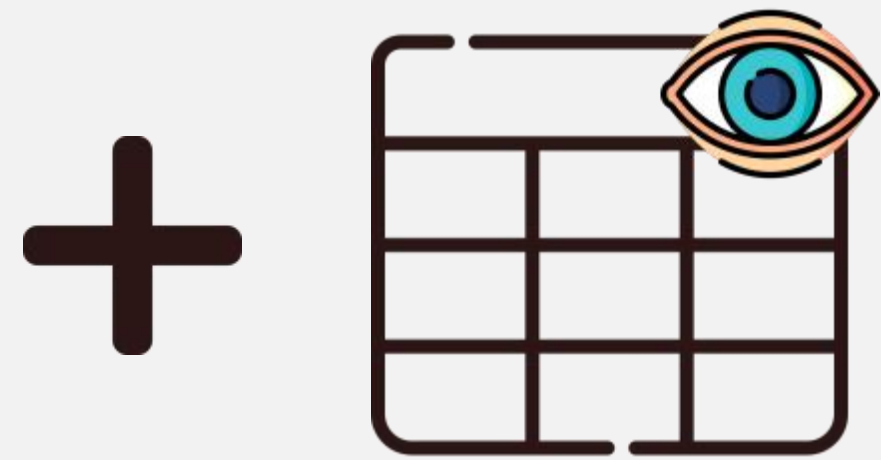
Under Fix Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Cost-Oriented Database Auto-Tuning

Under Fix Resources



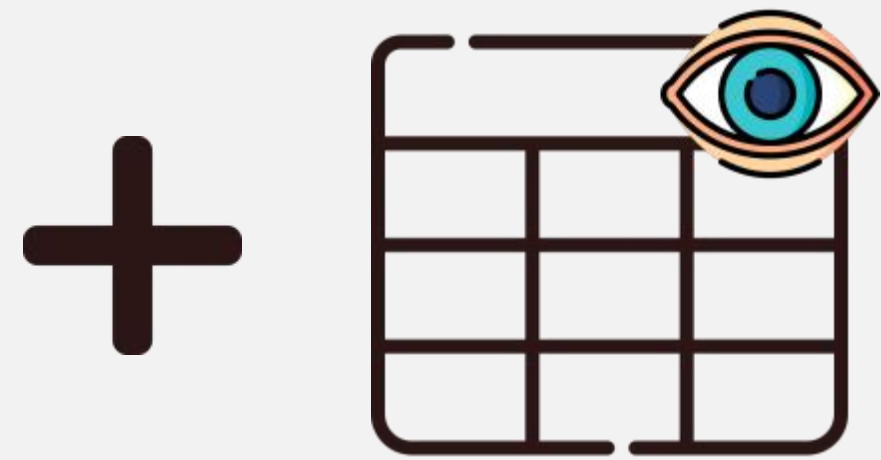
Read Perf:

Write Perf:

- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Cost-Oriented Database Auto-Tuning

Under Fix Resources



Read Perf: +

Write Perf:

- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Cost-Oriented Database Auto-Tuning

Under Fix Resources

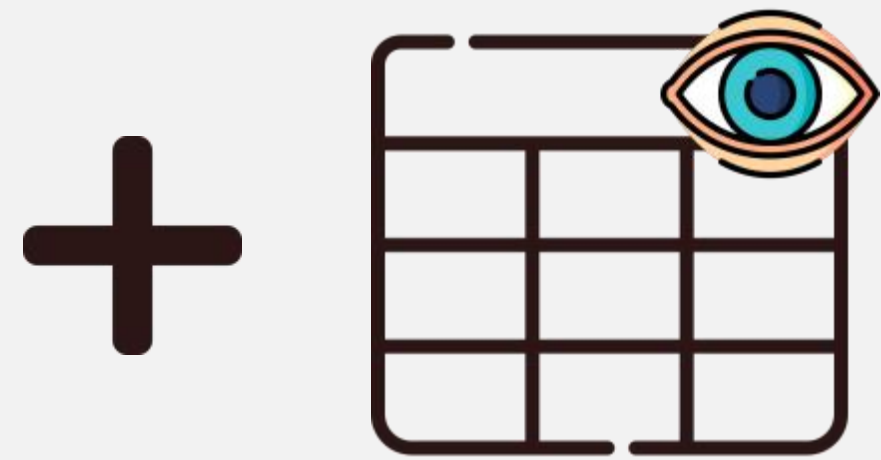


Read Perf: **+** **-**
Write Perf: **-**

- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Cost-Oriented Database Auto-Tuning

Under Fix Resources

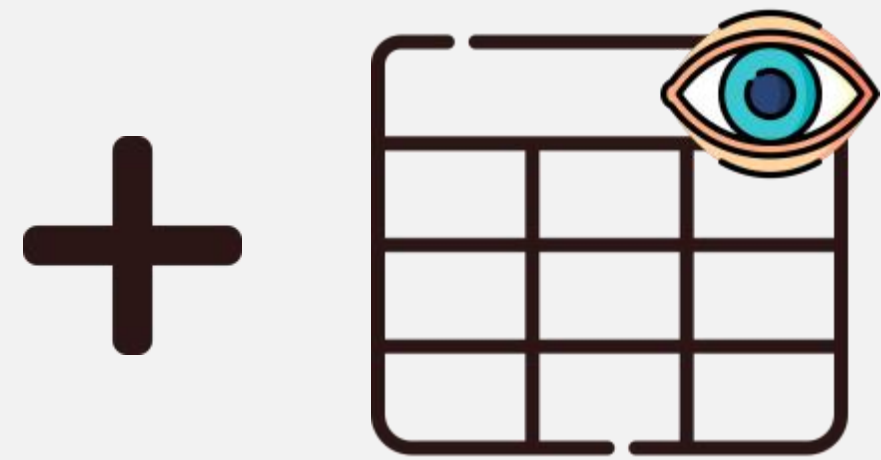


Read Perf: + - -
Write Perf: - - -

- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Cost-Oriented Database Auto-Tuning

Under Fix Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

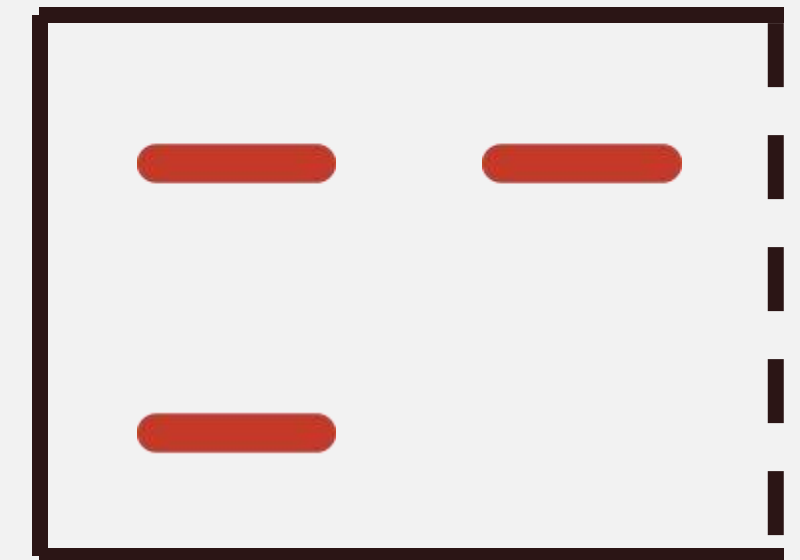
Resource Contention

Read Perf:

+

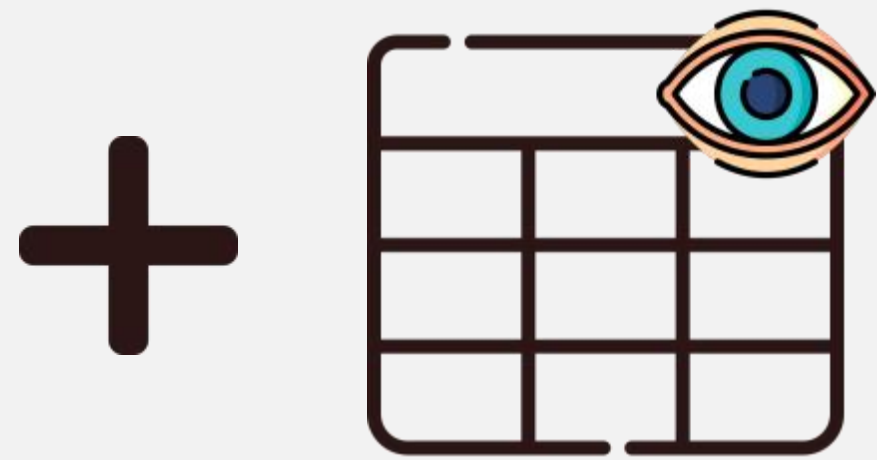
Write Perf:

-



Cost-Oriented Database Auto-Tuning

With Elastic Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Read Perf:

Write Perf:

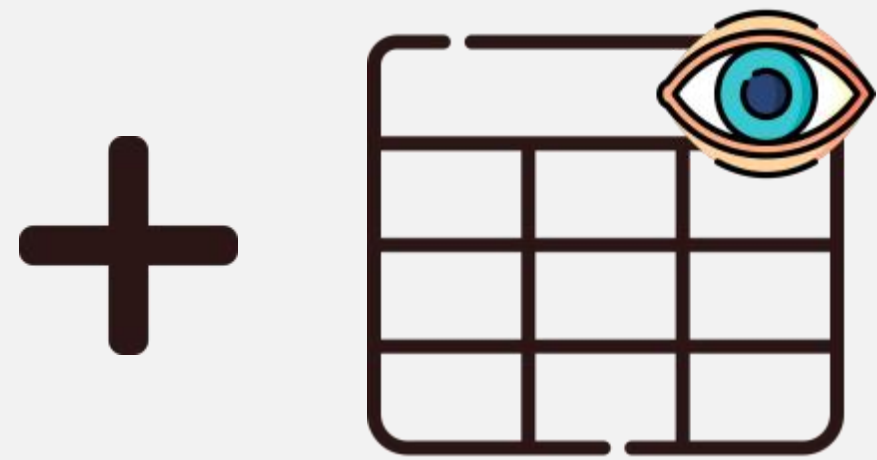
Read Cost:

Write Cost:

Storage Cost:

Cost-Oriented Database Auto-Tuning

With Elastic Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Read Perf: +

Write Perf:

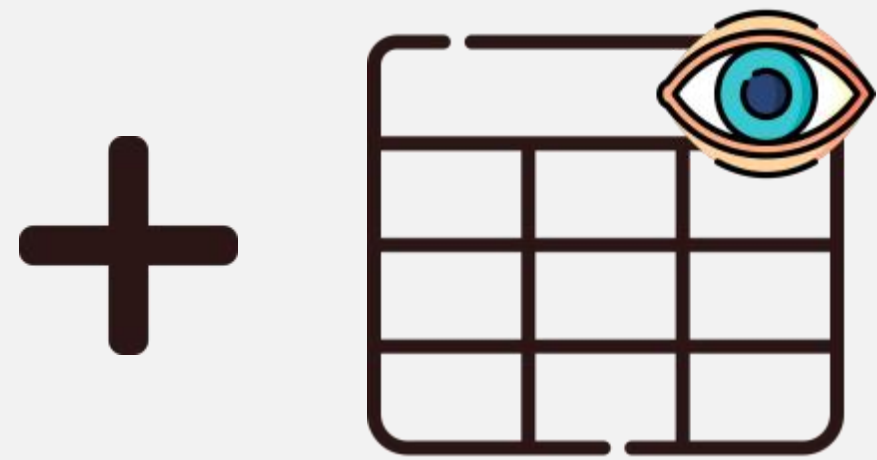
Read Cost: ↓ X

Write Cost:

Storage Cost:

Cost-Oriented Database Auto-Tuning

With Elastic Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Read Perf: +

Write Perf: Same

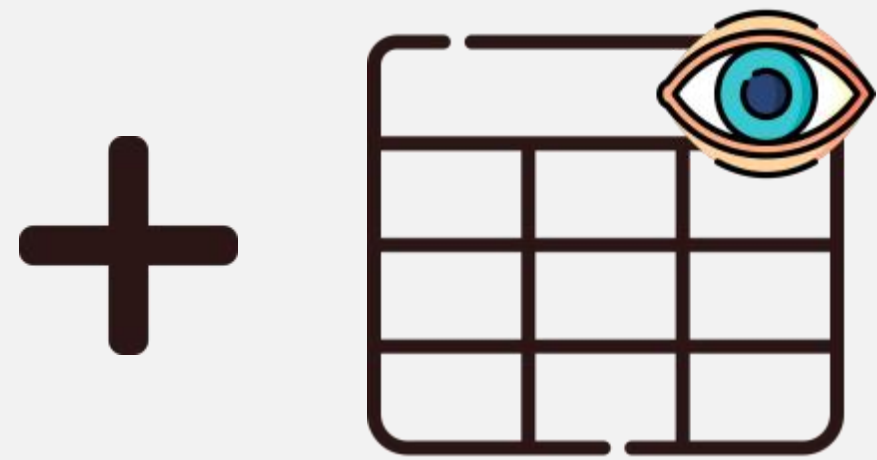
Read Cost: ↓ x

Write Cost: y ↑

Storage Cost:

Cost-Oriented Database Auto-Tuning

With Elastic Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Read Perf: +

Write Perf: Same

Read Cost: ↓ x

Write Cost: ↑ y

Storage Cost: ↑ z

Cost-Oriented Database Auto-Tuning

With Elastic Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Read Perf: +

Write Perf: Same

Read Cost: ↓ X

Write Cost: y ↑

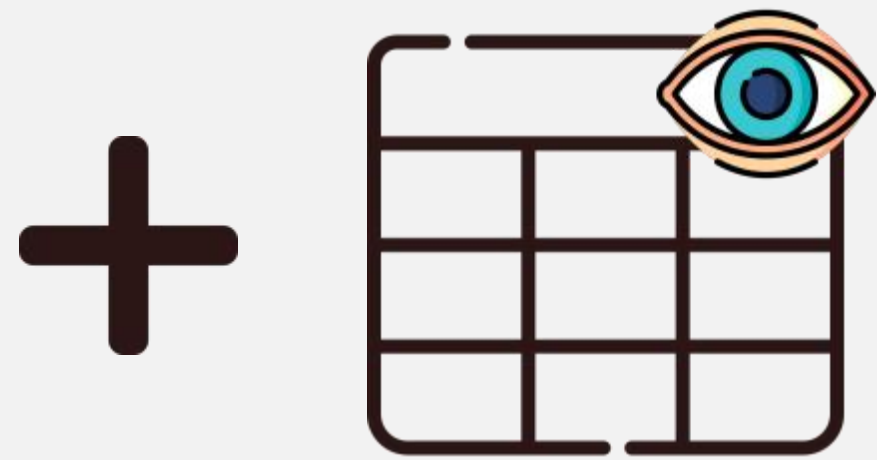
Storage Cost: ↑ z

$$x - y - z > 0$$



Cost-Oriented Database Auto-Tuning

With Elastic Resources



- Speeds up a subset of queries
- MV update slows down writes
- MV takes extra space

Read Perf: +

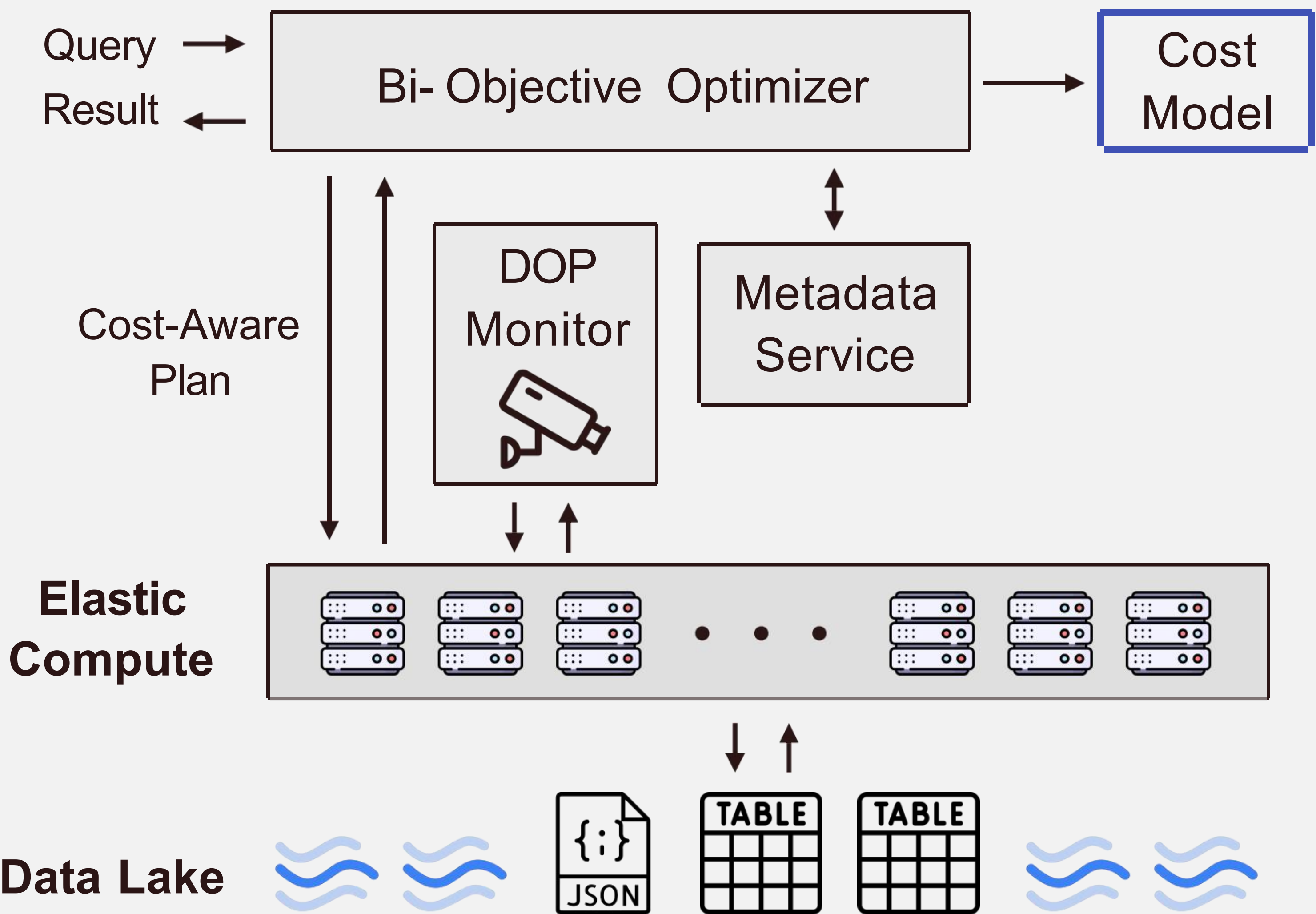
Write Perf: Same

Key Challenges:

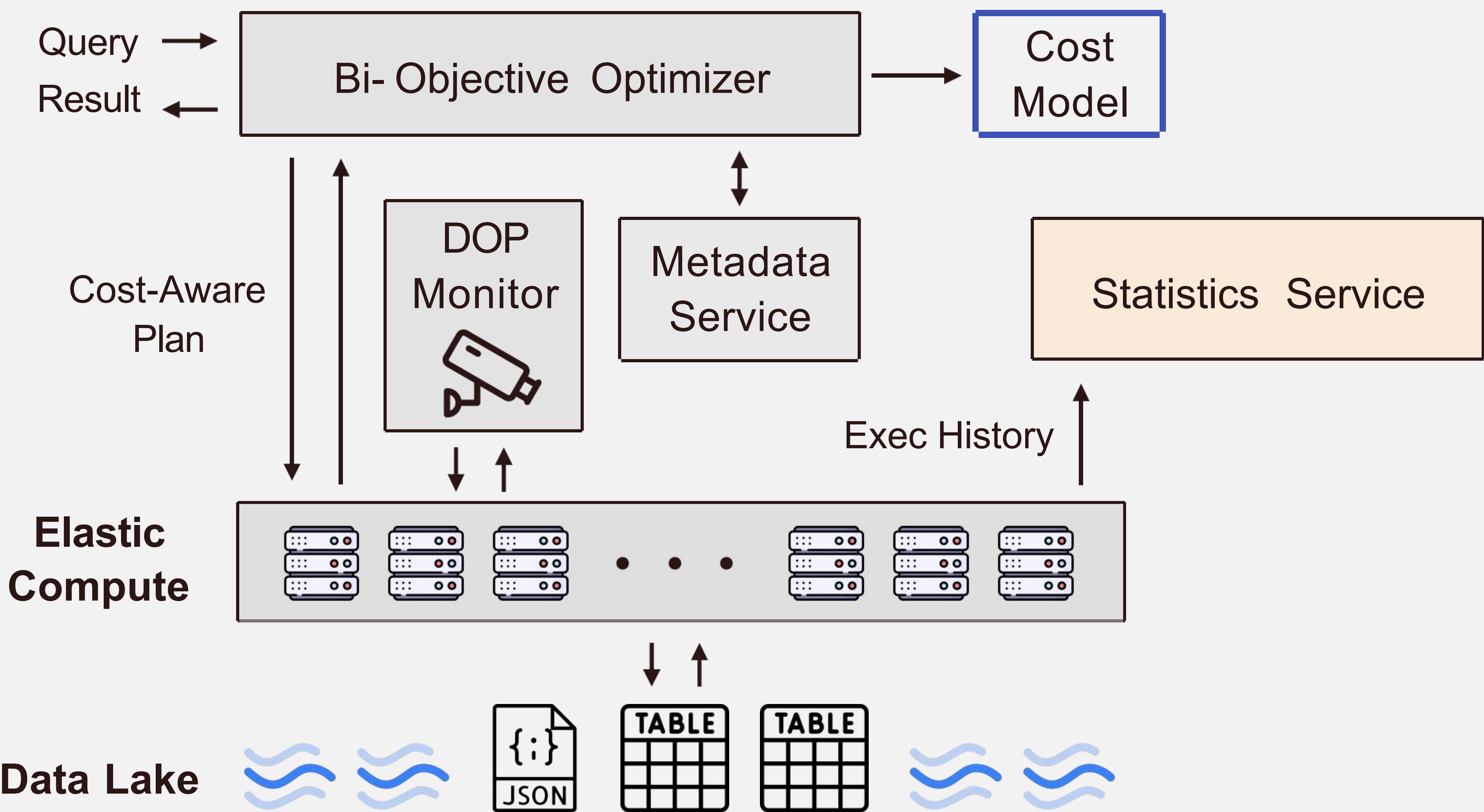
Accurate Cost Estimation

Accurate Workload Estimation

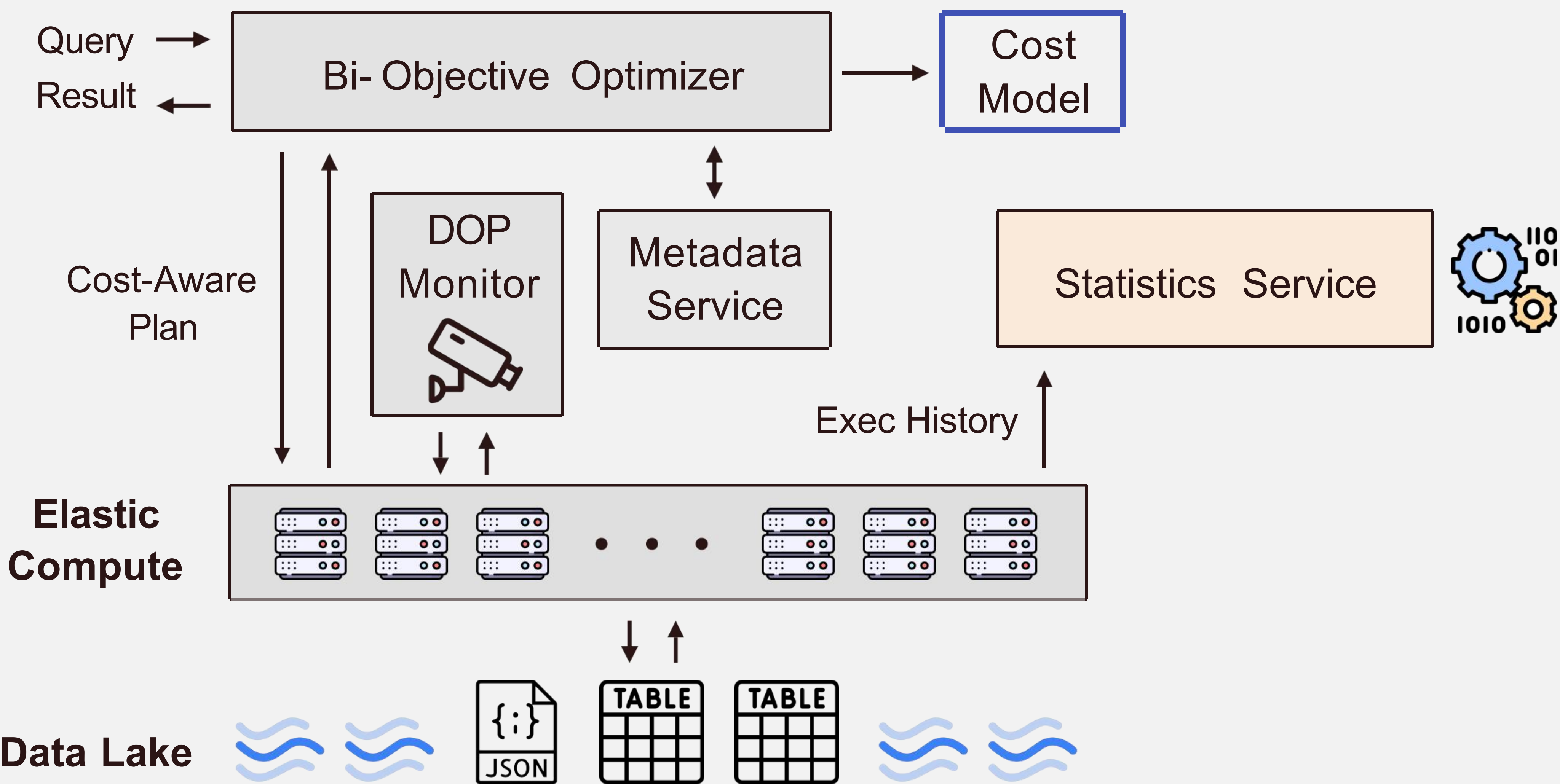
System Architecture



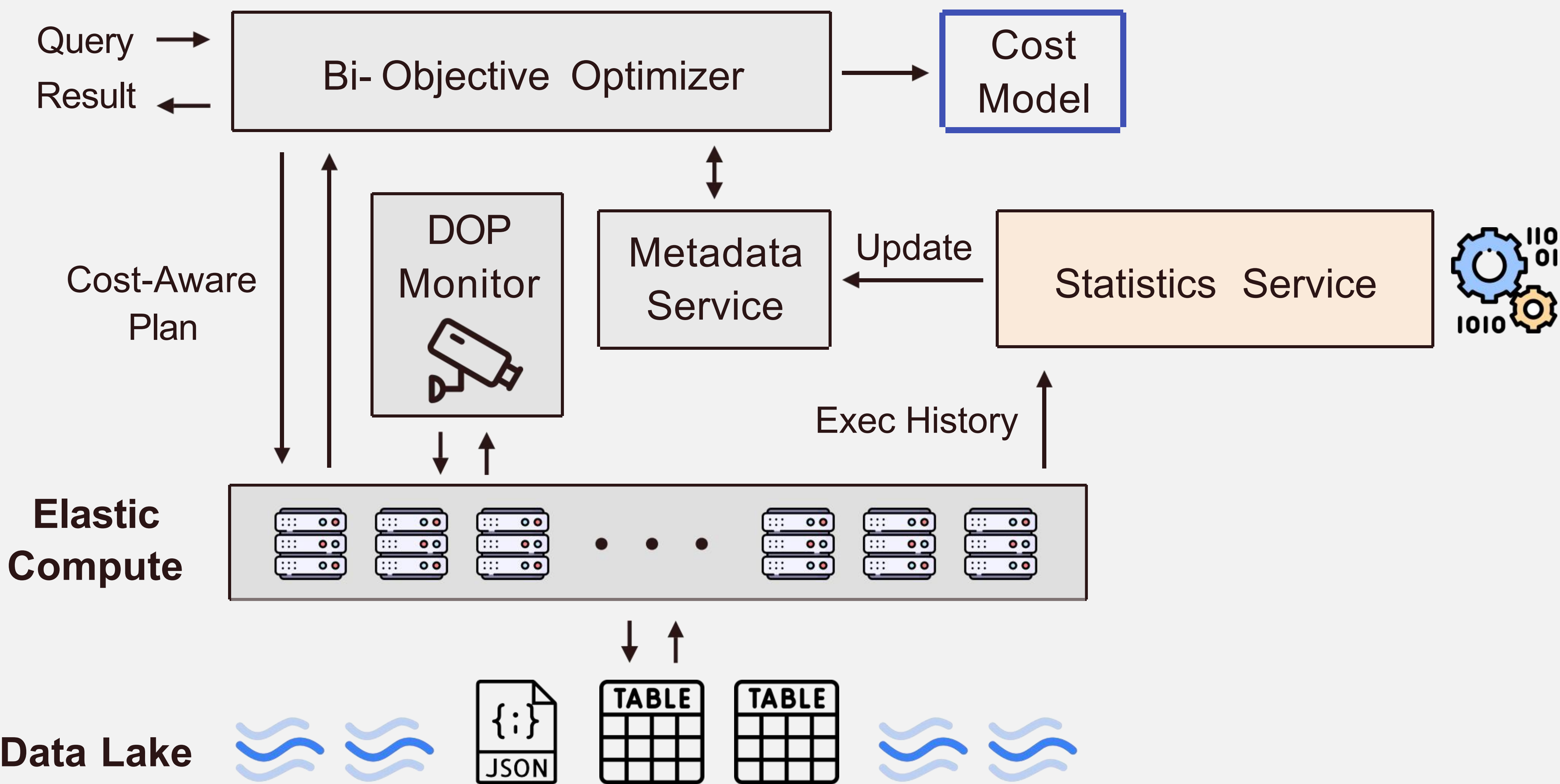
System Architecture



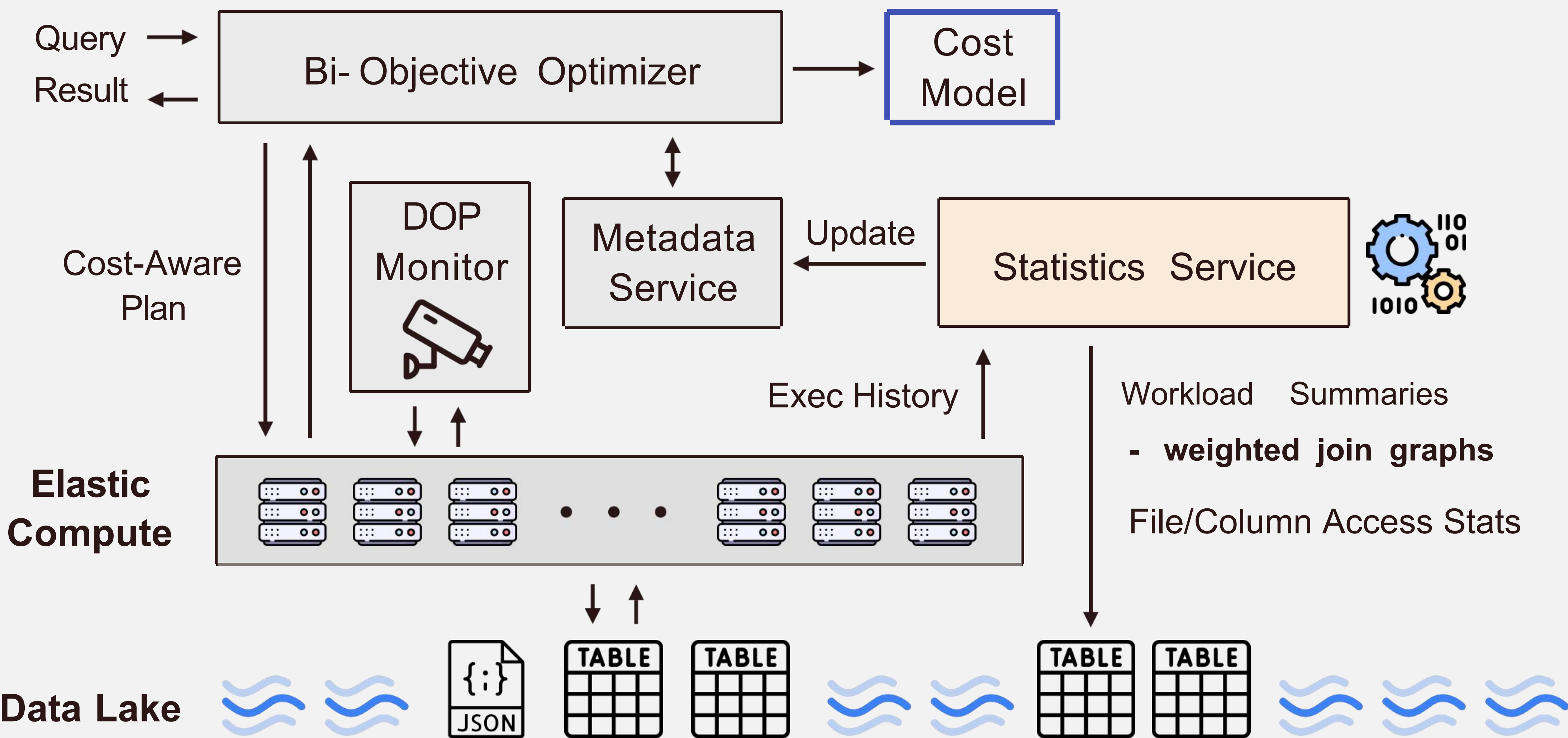
System Architecture



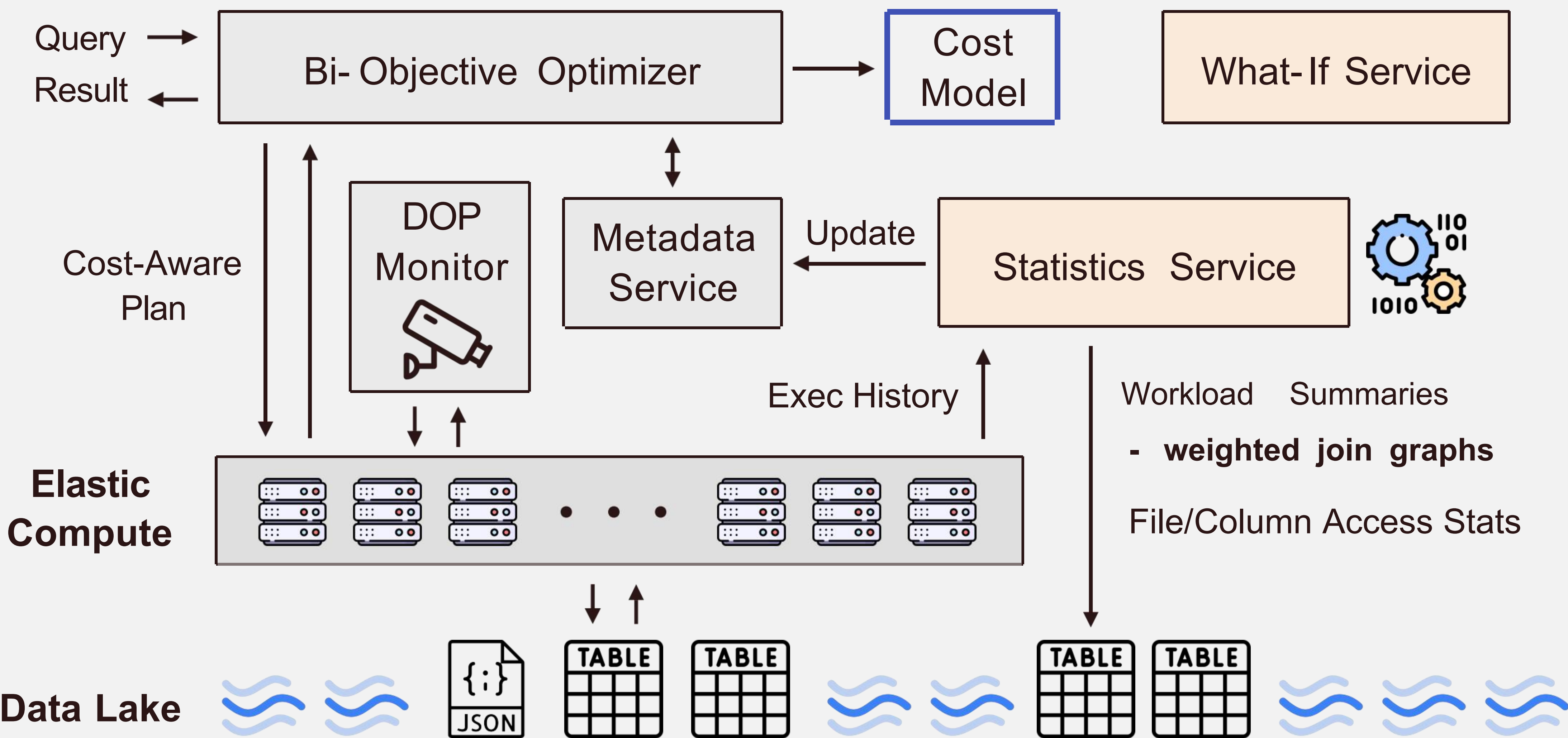
System Architecture



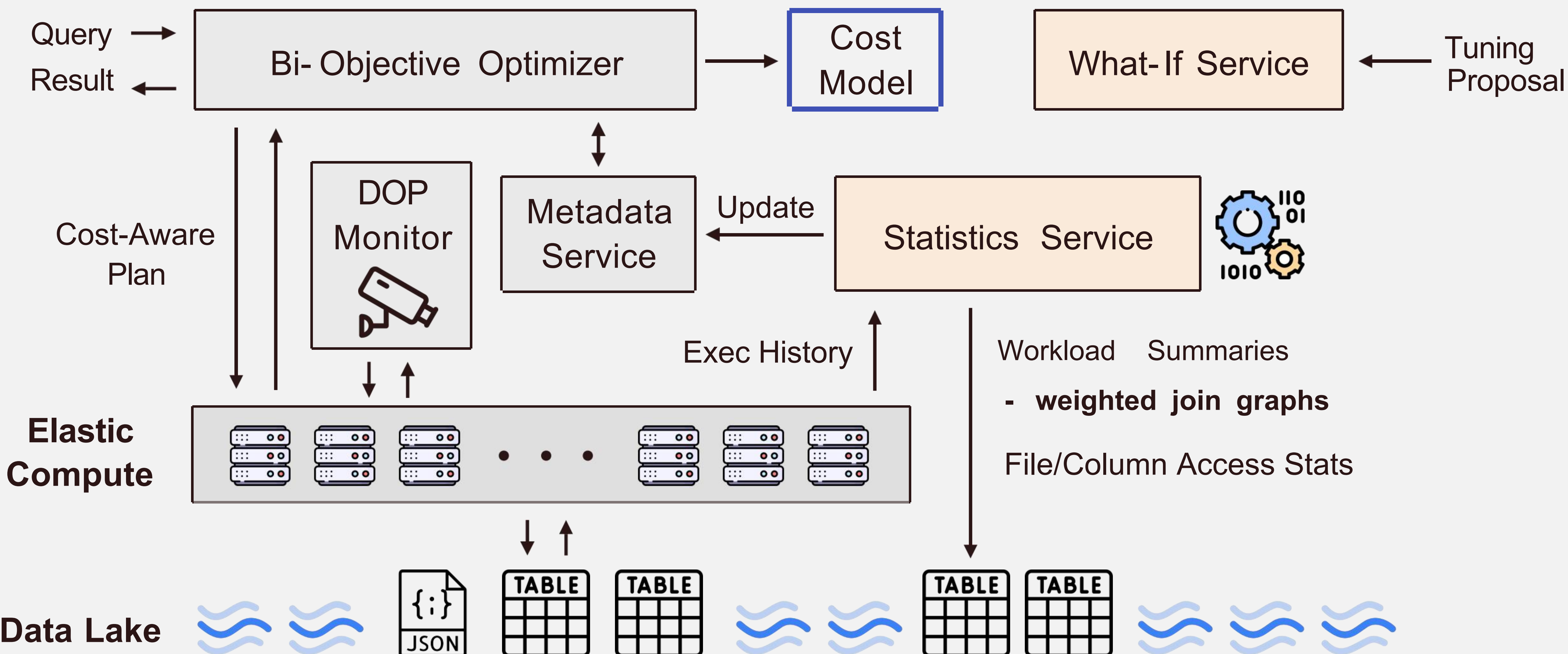
System Architecture



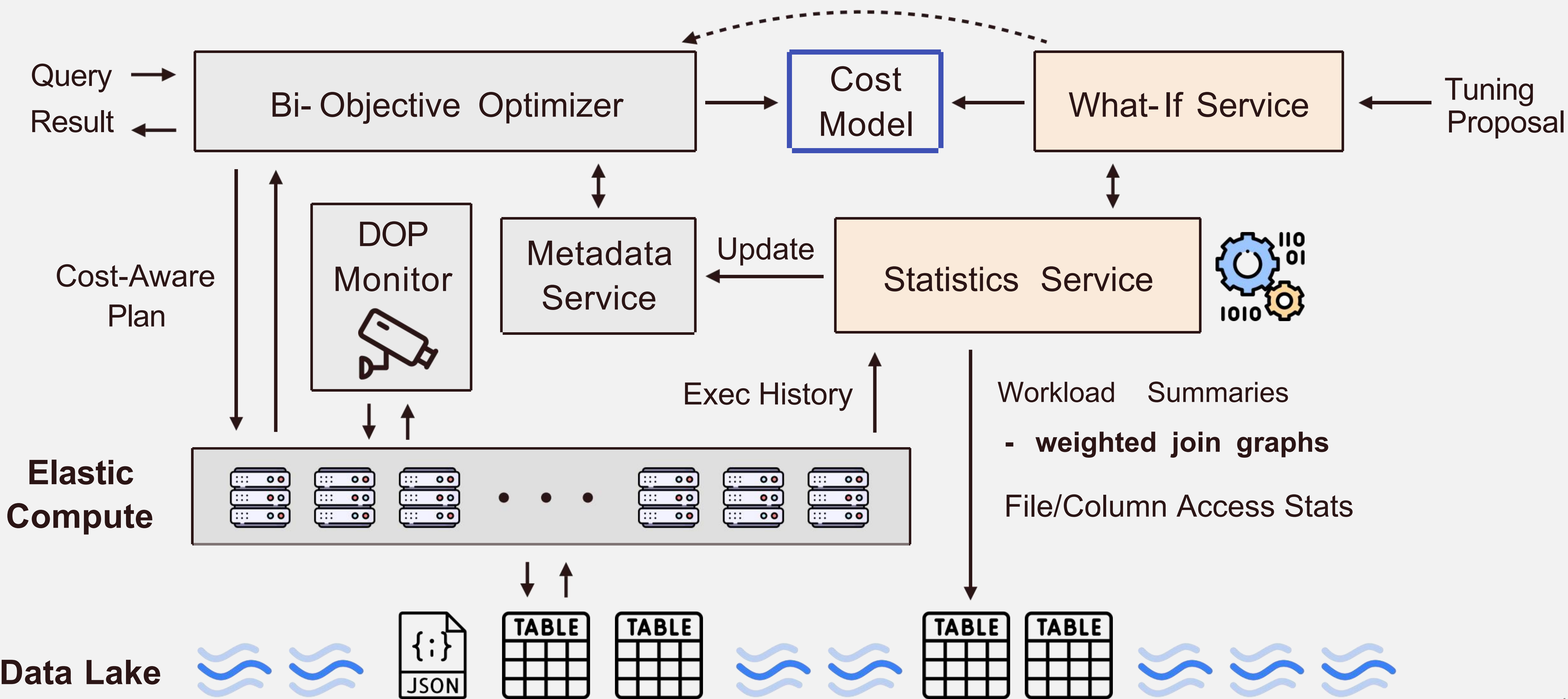
System Architecture



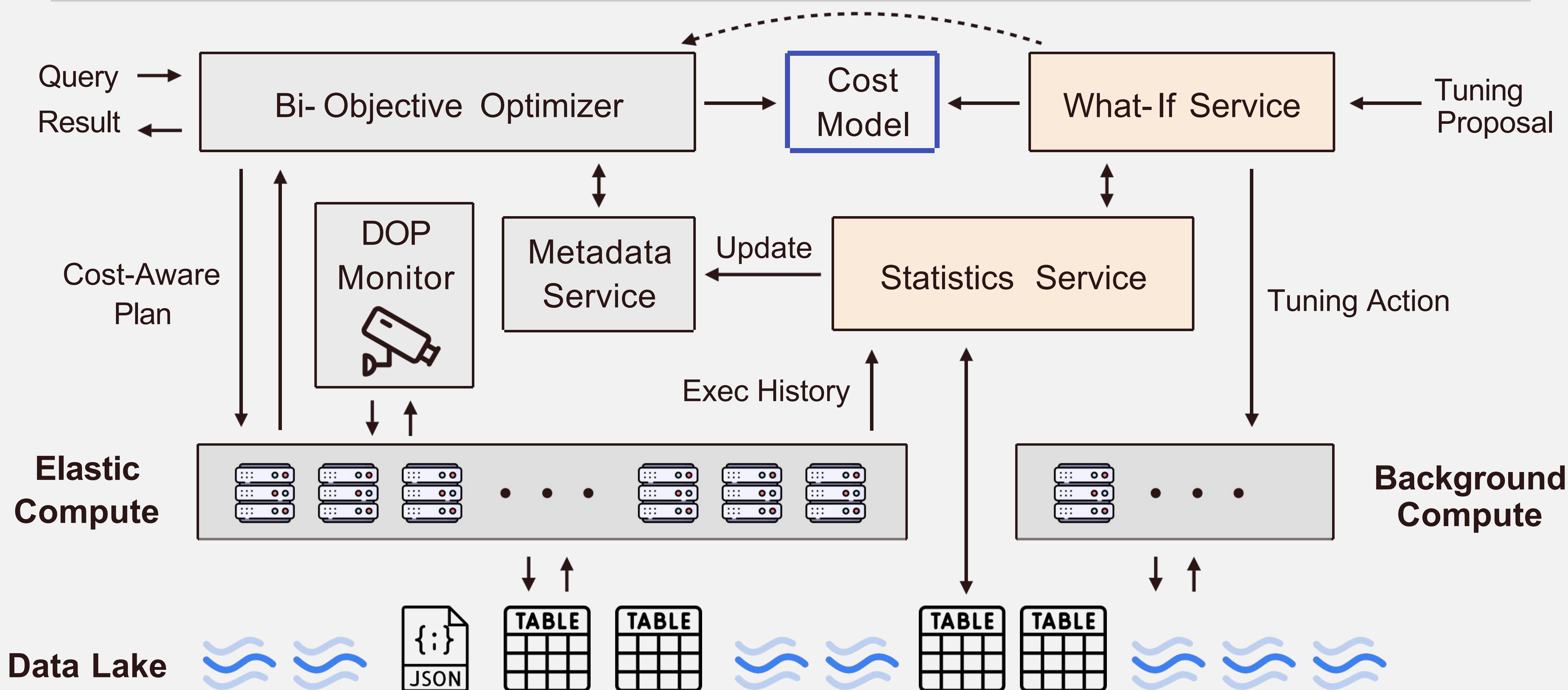
System Architecture



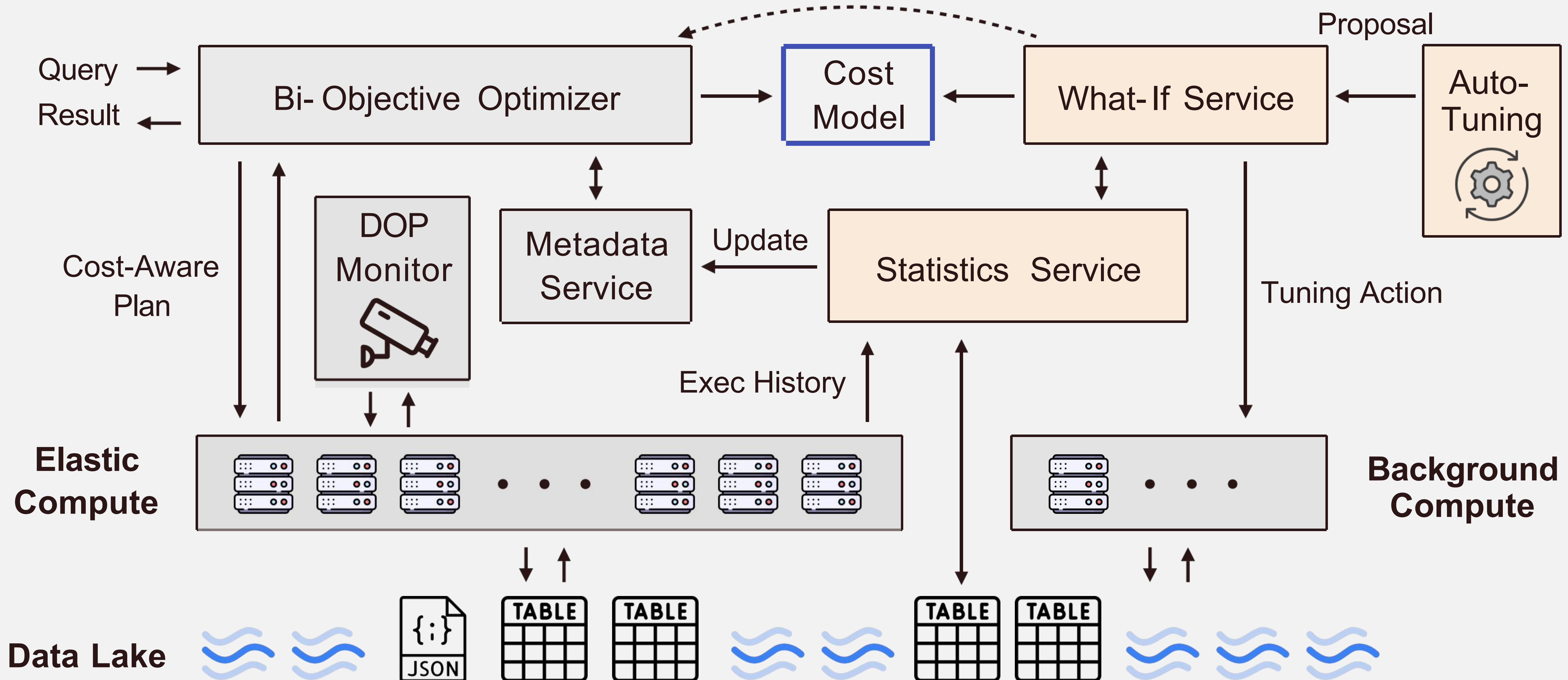
System Architecture



System Architecture



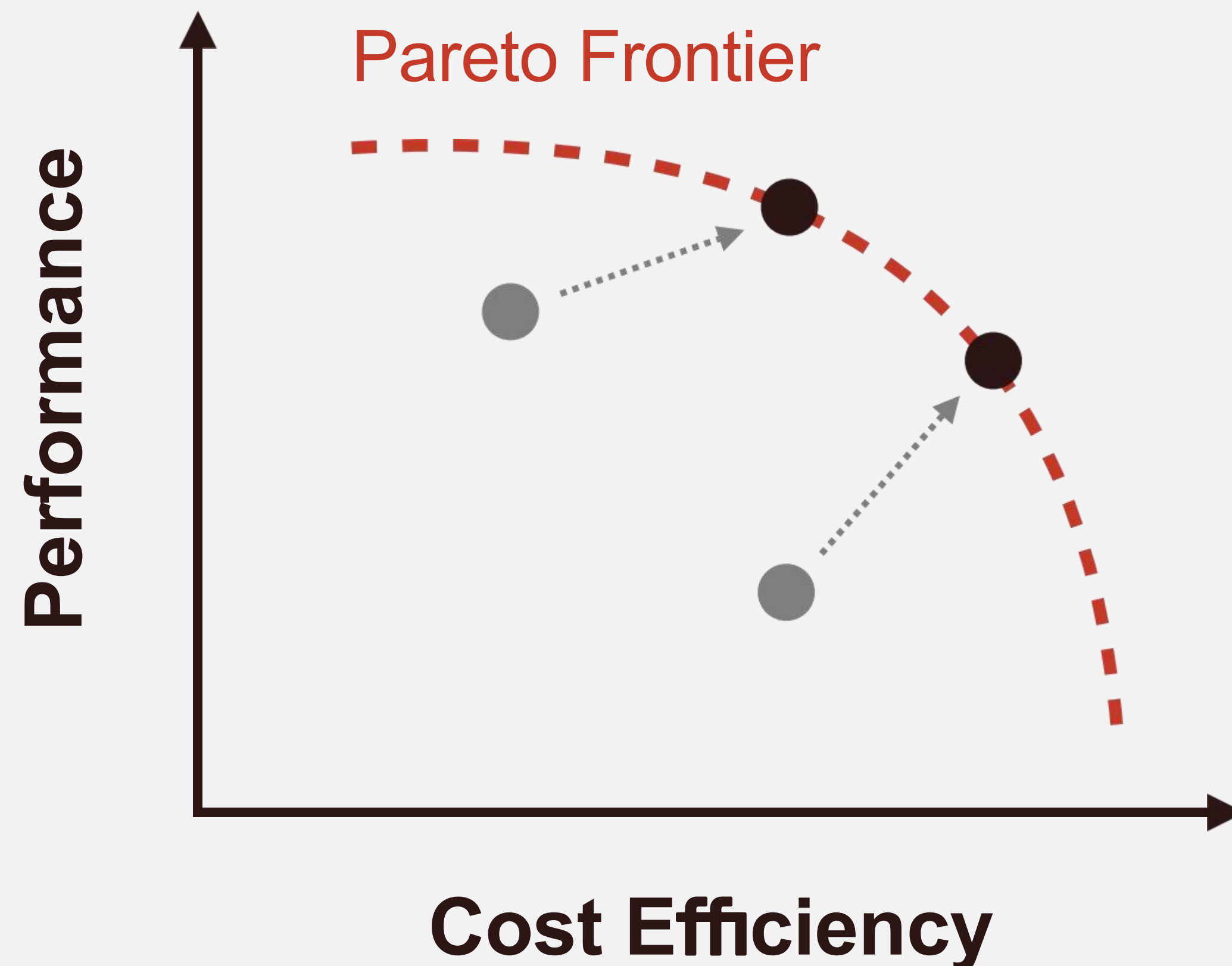
System Architecture



Towards Cost Intelligence



Cost is as important as performance in cloud-native databases



Time: 10s —  — 5min

Cost: \$2 —  — \$0.1

