



When free is not really free

Towards Benchmark as a Service?

Avrilia Floratou (University of Wisconsin – Madison)

Jignesh M. Patel (University of Wisconsin – Madison)

Willis Lang (University of Wisconsin – Madison)

Alan Halverson (Microsoft Jim Gray Systems Lab)

Motivation

On-site
Relational Database
Systems
(RDBMSs)

Complexity of:

- Choosing
- Configuring
- Maintaining
the RDBMS and server

Users must also
consider the long-term:

- License costs
- Maintenance costs
- Administration costs

Motivation

Cloud-based Database-as-a-Service (DaaS)

DaaS provider responsible for:

- Operating
 - Administering
- the RDBMs and server.

User pays “on-demand”
based only on the computing
resources that are actually
consumed.

DaaS offerings with multiple
pricing options (e.g based on
different types of compute
resources) can address a
variety of users’ needs.

Motivation

Cloud-based
Database-as-a-Service
(DaaS)

Do the current DaaS offerings actually simplify the process of running a database workload?

Users must make an upfront decision of choosing a DaaS offering while the long-term performance and cost consequences are harder to figure out.

An example pricing model

Monthly
Storage fee
per GB

\$0.1

Type of compute resources (DBInstance class)	Hourly Machine Usage fee	Hourly License Cost Fee	Commercial DBMS	Hourly License Cost Fee	Open-Source DBMS
Small	\$0.11	\$0.04	\$0.16	\$0	\$0.11
Large	\$0.44	\$0.20	\$0.64	\$0	\$0.44
HM XL	\$0.65	\$0.20	\$0.85	\$0	\$0.65
HM Double XL	\$1.30	\$0.40	\$1.70	\$0	\$1.30
HM Quadruple XL	\$2.60	\$0.80	\$3.40	\$0	\$2.60

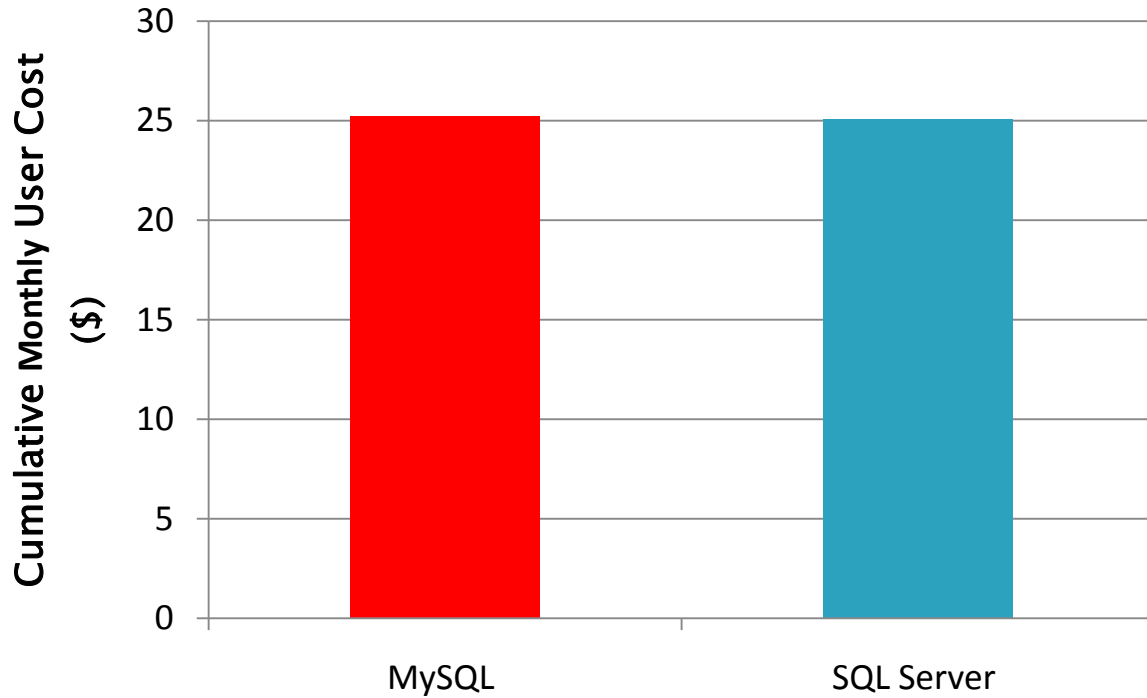
Example

Wisconsin Benchmark Query 21

```
Insert into TMP  
Select min (unique3) from  
TABLE1  
Group by onePercent
```

	SQL Server	MySQL
Monthly Storage fee	\$25	\$25
Hourly Machine Usage fee	\$1.30	\$1.30
Hourly License Cost fee	\$0.65	\$0
Execution Time	185 sec	621 sec

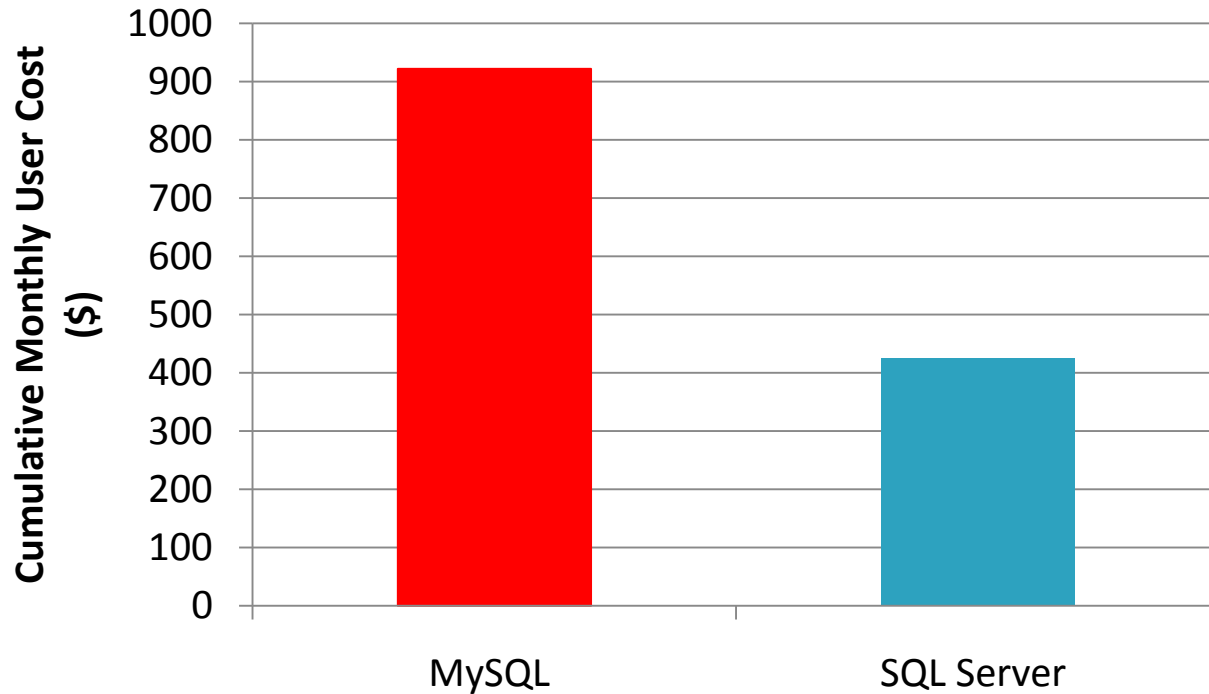
Example



#Repetitions/Month = 1

The total cost is dominated by the storage fee.

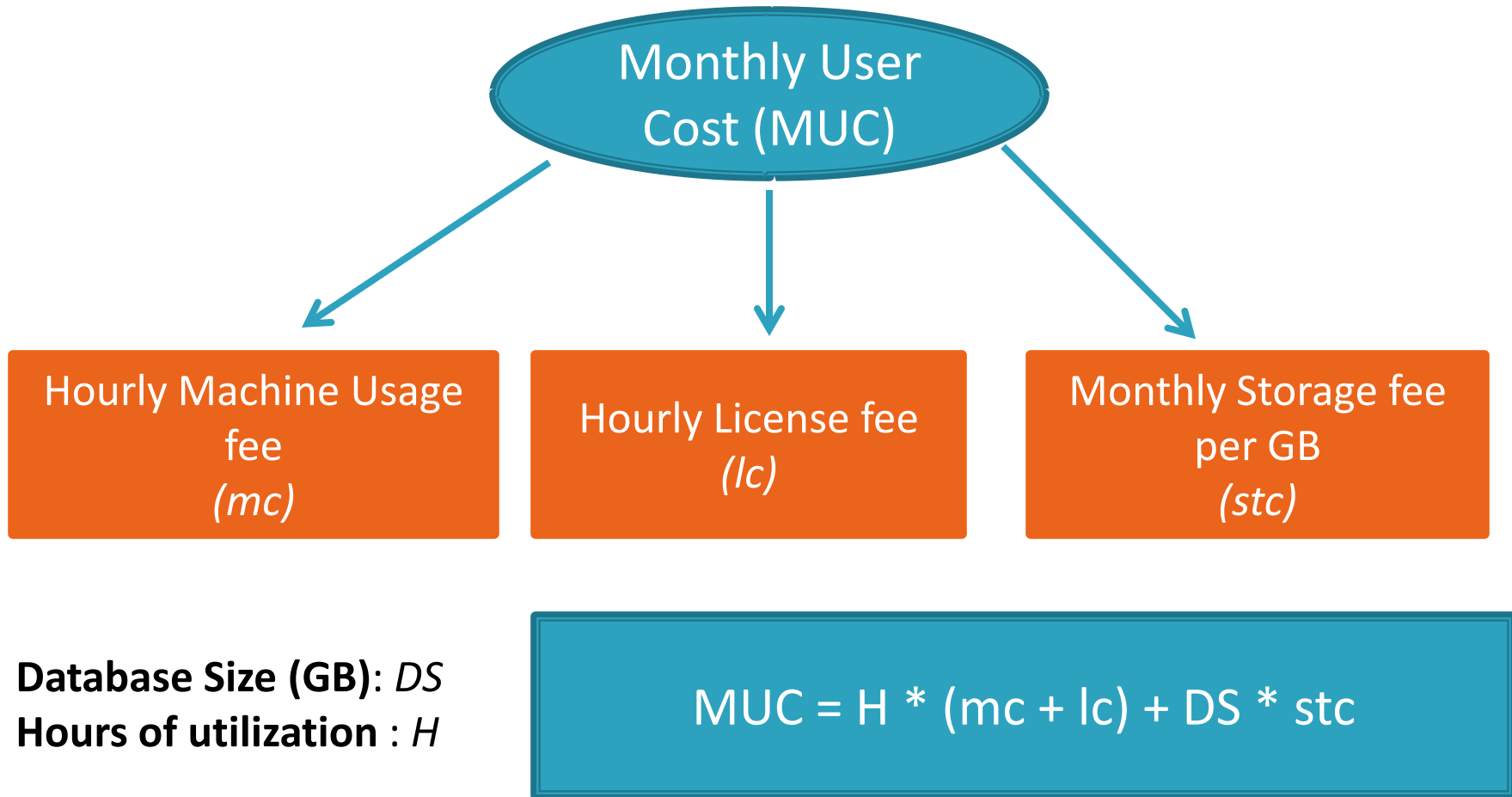
Example



#Repetitions/Month = 4000

Using a more expensive (per hour) DBMS, can save 54%!

Cost Model



Experimental Setting

- Server Configuration:
 - 2 quad-core processors @2.13GHz
 - 32 GB of RAM
 - 12 146 GB SAS drives
- DBMSs:
 - MySQL (Community Server 5.5.9)
64-bit Ubuntu Server 9.10
 - SQL Server 2008 R2 (Data Center Edition)
64-bit Windows Server 2008 R2 Enterprise Edition
- DBMS Configuration:
 - 24 GB buffer pool
 - 10 data disks, 1 log disk
 - InnoDB storage engine for MySQL

Workload

- ▶ Wisconsin Benchmark
 - Two tables of 80 GB each
 - One table of 8GB

- ▶ 6 different types of workloads
 - **OLTP** (HeapWorkload, IndexedWorkload)
 - **DSS** (HeapWorkload, IndexedWorkload)
 - **Mixed** (HeapWorkload, IndexedWorkload).

Cost Model Parameters

$$MUC = H * (1.30 + lc) + 25$$

mc = \$1.30

stc = \$0.10 per GB

DS= 250 GB

H : Hours of utilization

lc = {
\$0 , MySQL
\$0.65-\$3.90 , SQL Server

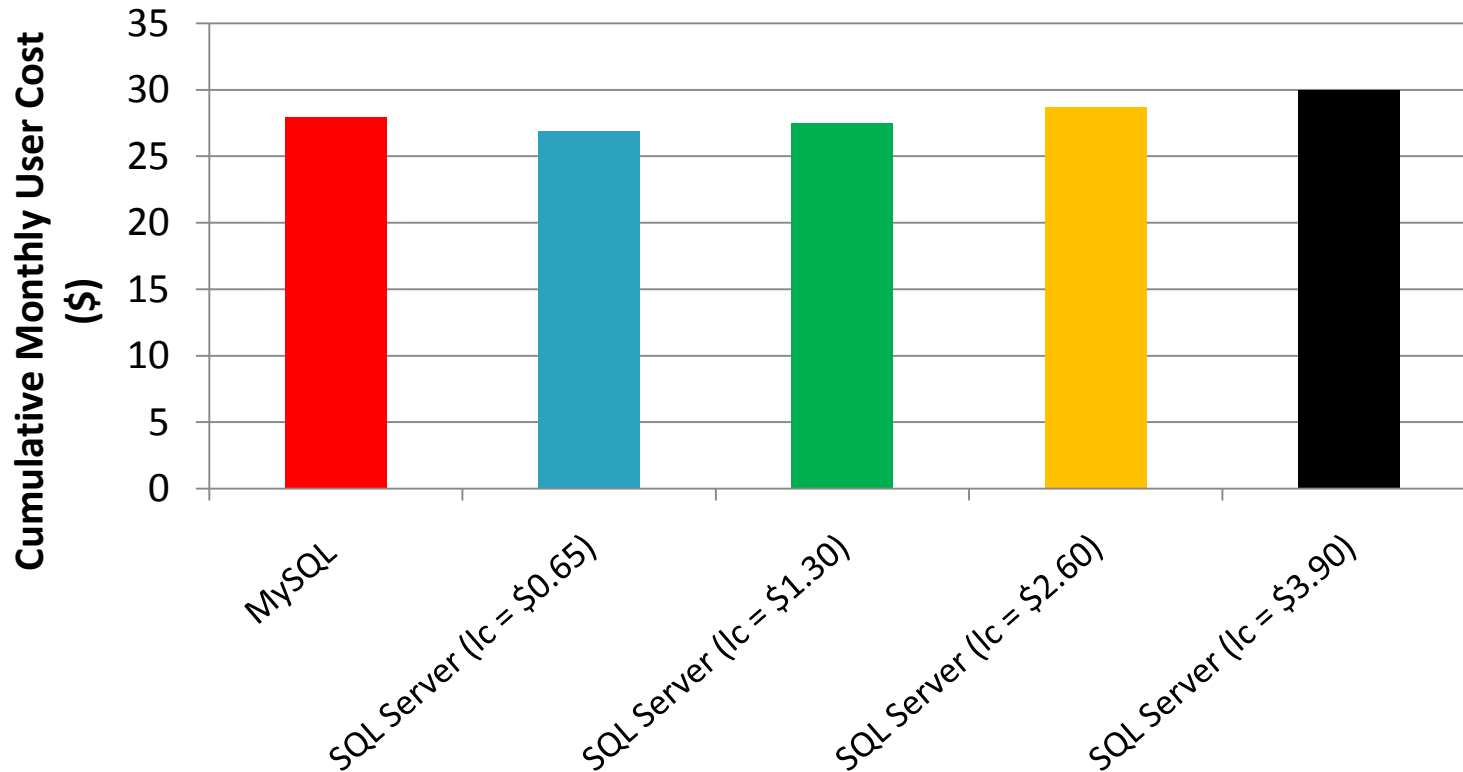
Mixed Indexed Workload

- Consists of 14 queries (DSS and OLTP)
- Queries in the workload are run sequentially
- Total Time : \sum Query time

SQL Server Time (sec)	MySQL Time (sec)
3441	8079

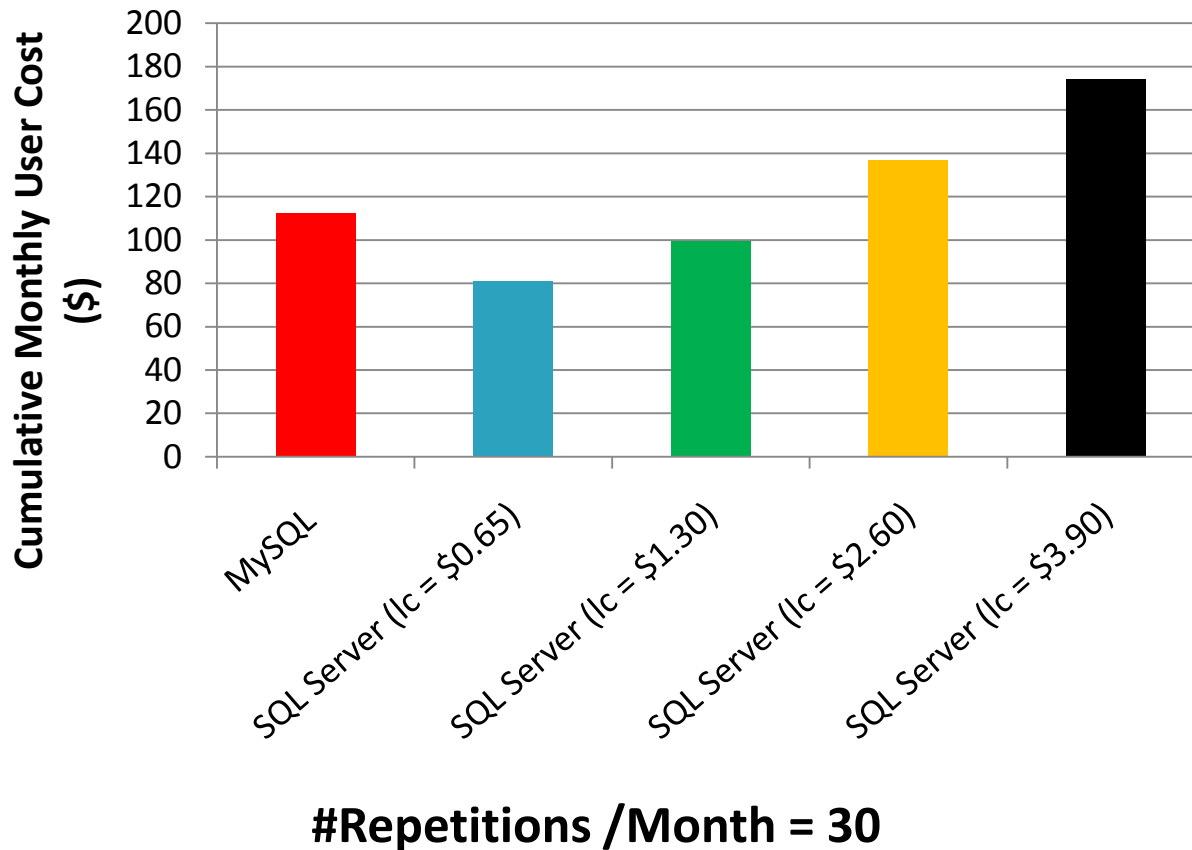
How does this 2.3X performance gap affect the monthly user cost for different query loads?

Mixed Indexed Workload



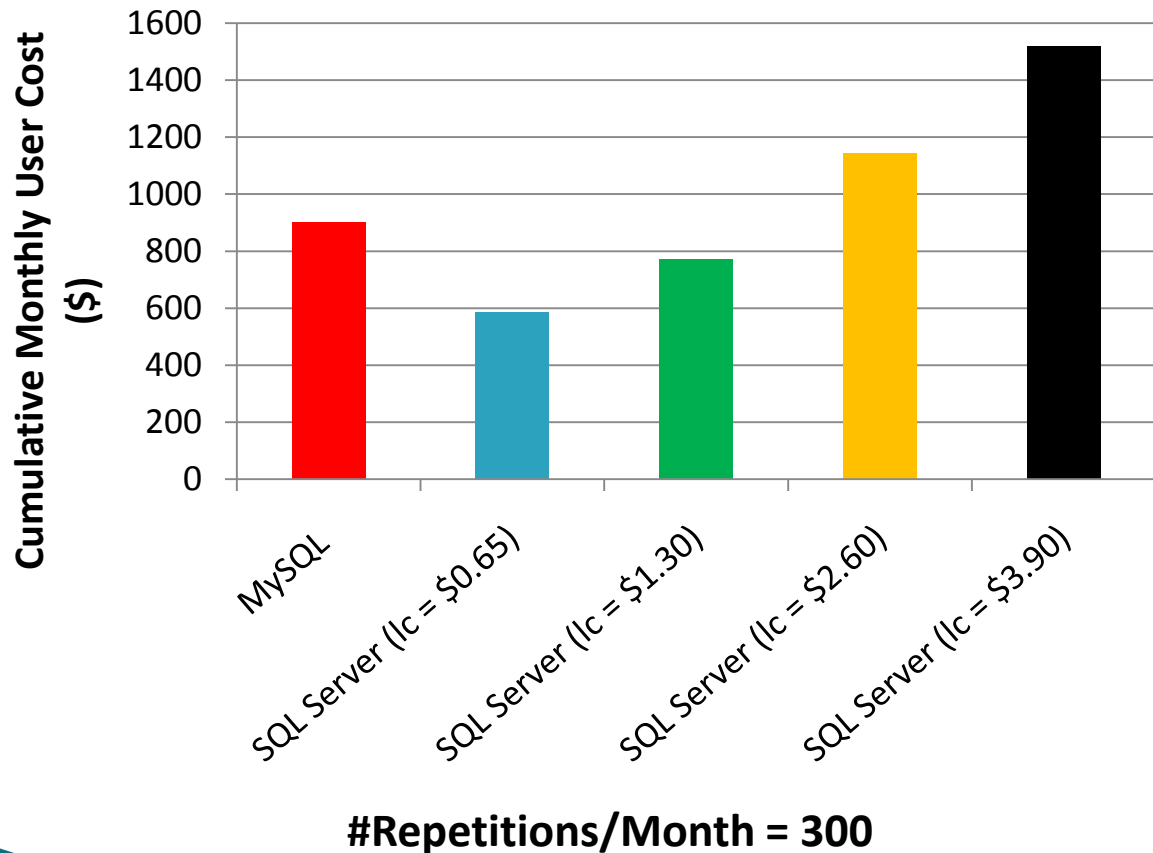
#Repetitions/Month = 1

Mixed Indexed Workload



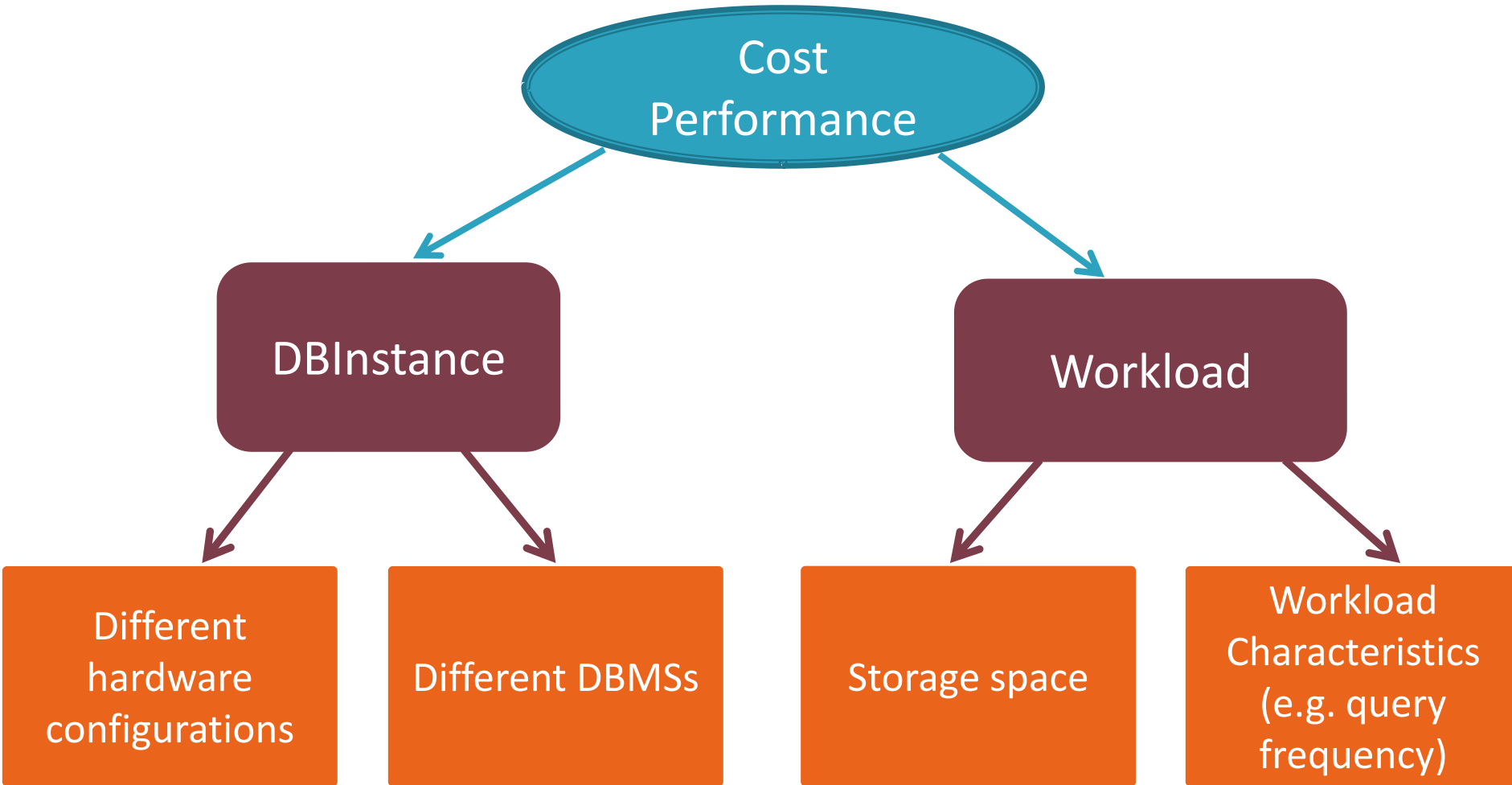
Using the commercial DBMS can save up to 28% (lc = \$0.65)

Mixed Indexed Workload

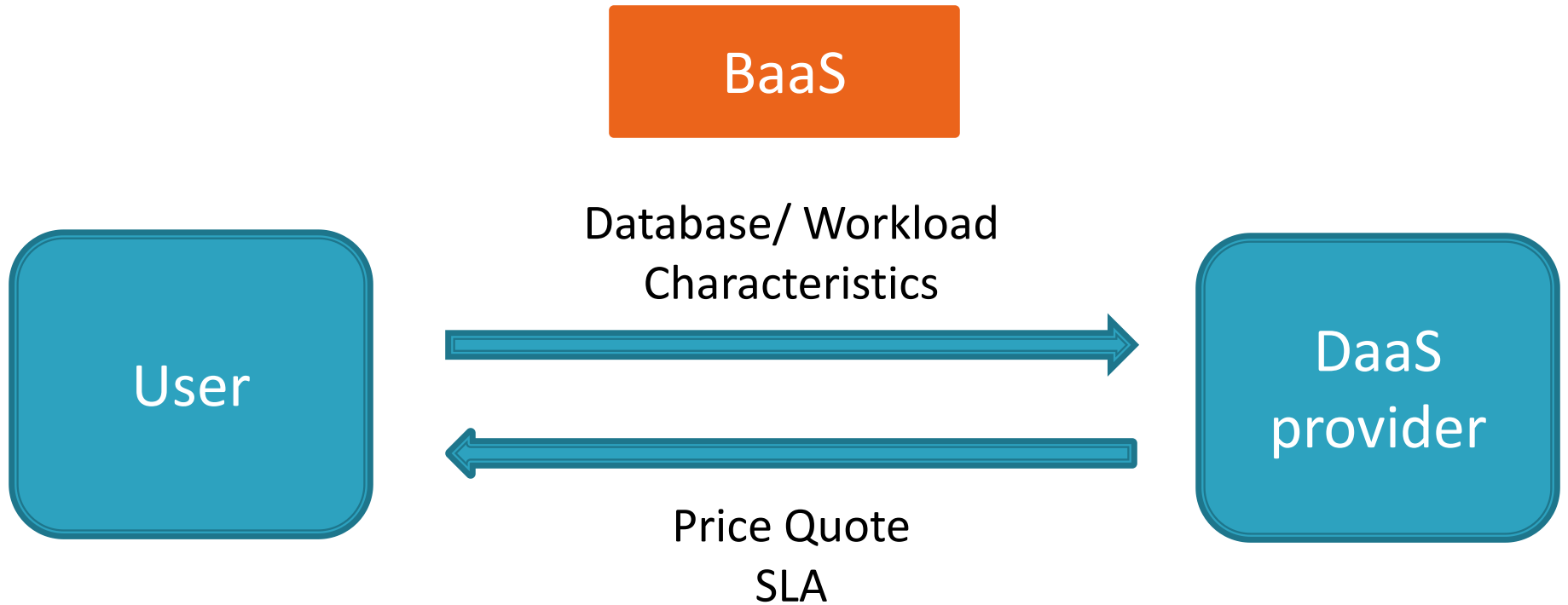


Using the commercial DBMS can save up to 35% and at the same time being 2.3X faster!

The User's Perspective



Towards Benchmark-as-a-Service



BaaS

- ▶ Closer to the “utility” model.
- ▶ Reduces complexity.

- ▶ Motivation to find the most optimal way to run the backend DBMS engine.
- ▶ Reduces the operational cost.
- ▶ The system operates more efficiently.



Directions for Future Work

- ▶ Setting up BaaS is challenging
 - How should the user describe the workload?
 - How should the provider run a mix of workloads that started with a BaaS?
 - How should the provider monitor changes in workloads that started with a price quote from the BaaS?

Conclusions

- ▶ Argue that existing cloud pricing models are too complex for the user.
- ▶ Influence of different types of DBMSs on performance/cost when running database workloads on the cloud.
- ▶ Propose a new type of service (BaaS) that will introduce transparency and clarity when pricing DaaS.

“We really don’t need any new benchmarks. Every DaaS customer has a benchmark – his/her workload! What we really need is Benchmark-as-a-Service, and not new benchmarks!”

-- Jignesh Patel