OSA CON 23

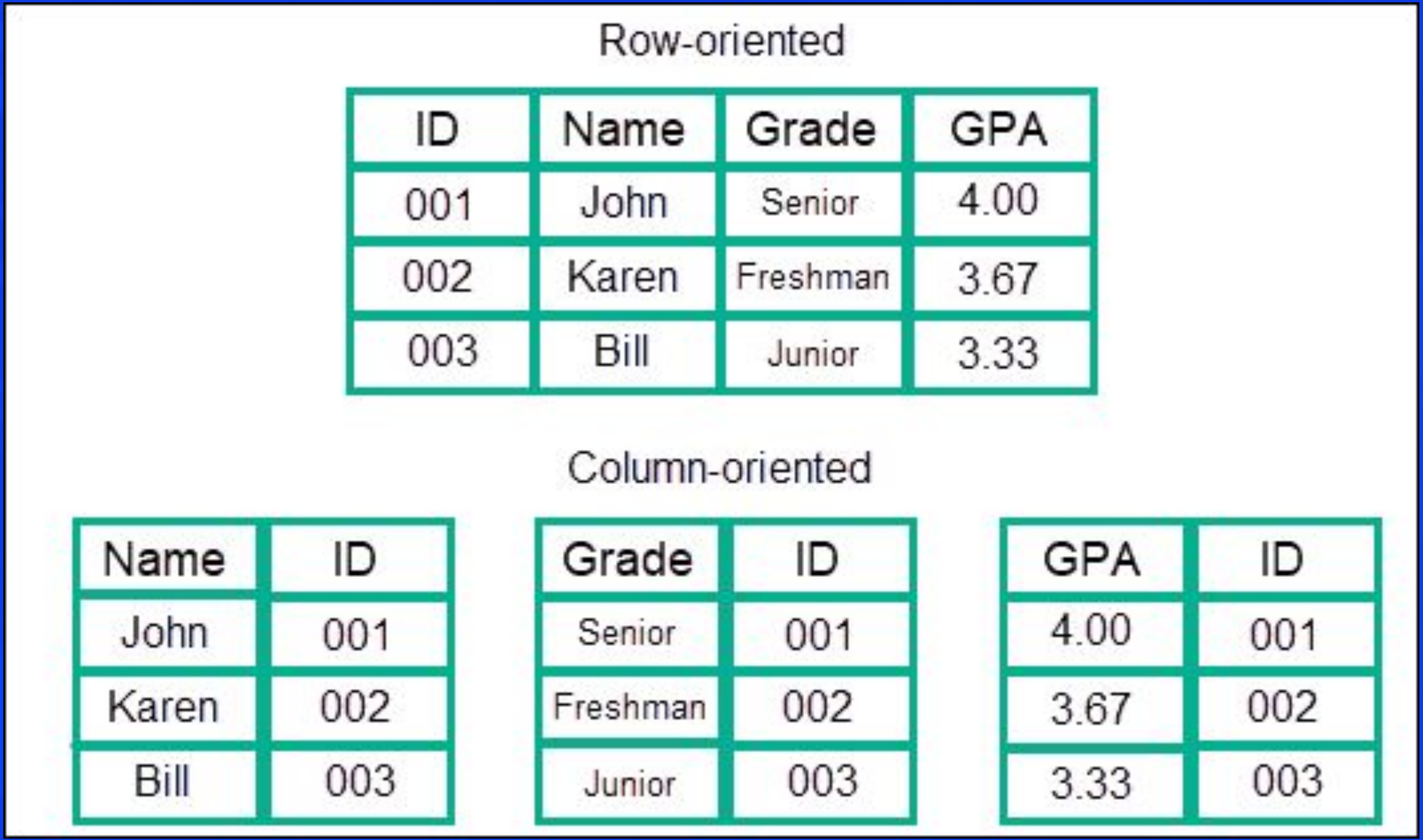
You put OLTP in my OLAP!

Analytics and Real-time Converged

Felipe Cardeneti Mendes
Solution Architect @ ScyllaDB
Find me on LinkedIn!

December 12-14 2023

The Basics



Row-oriented

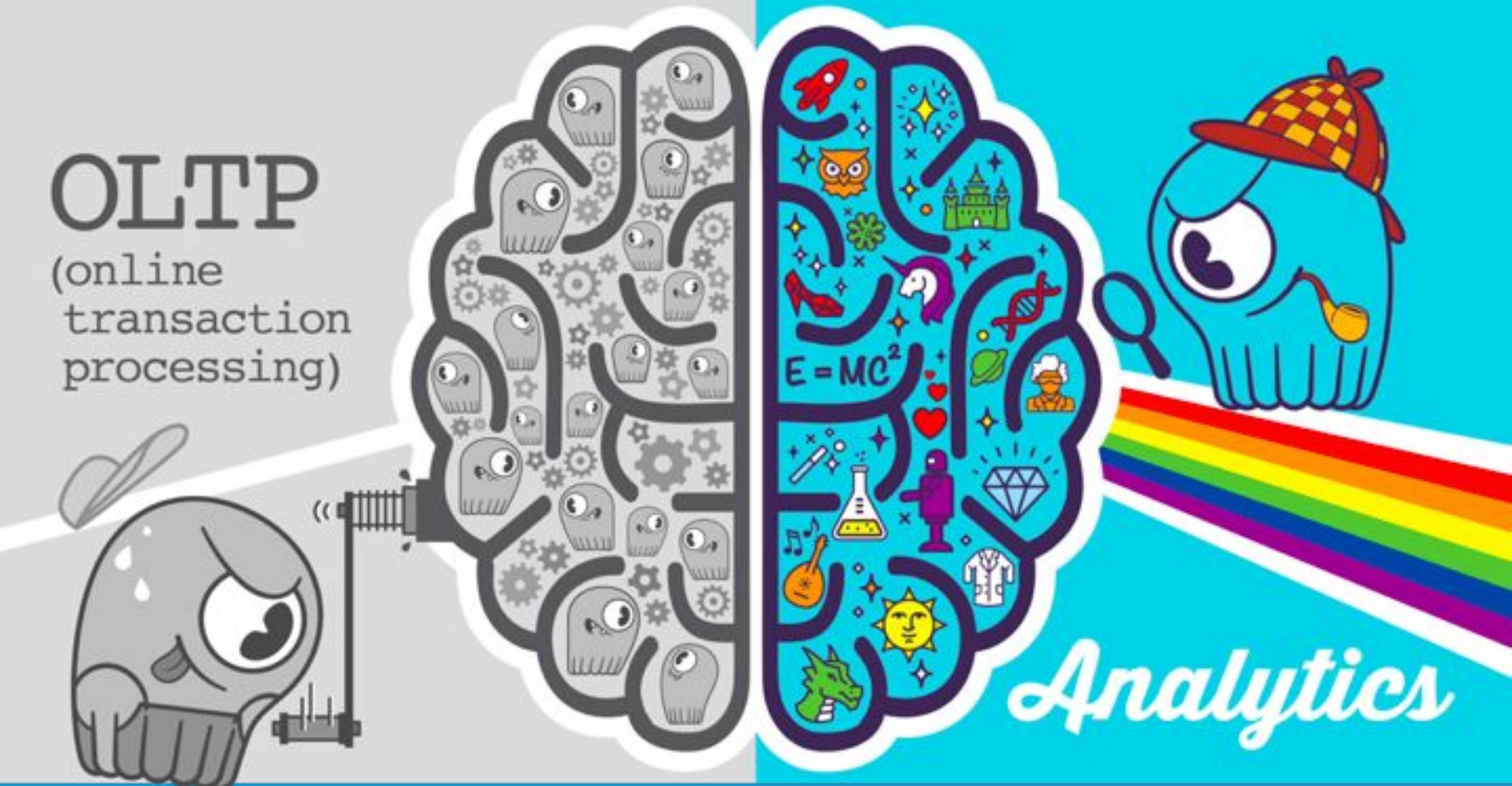
ID	Name	Grade	GPA
001	John	Senior	4.00
002	Karen	Freshman	3.67
003	Bill	Junior	3.33

Column-oriented

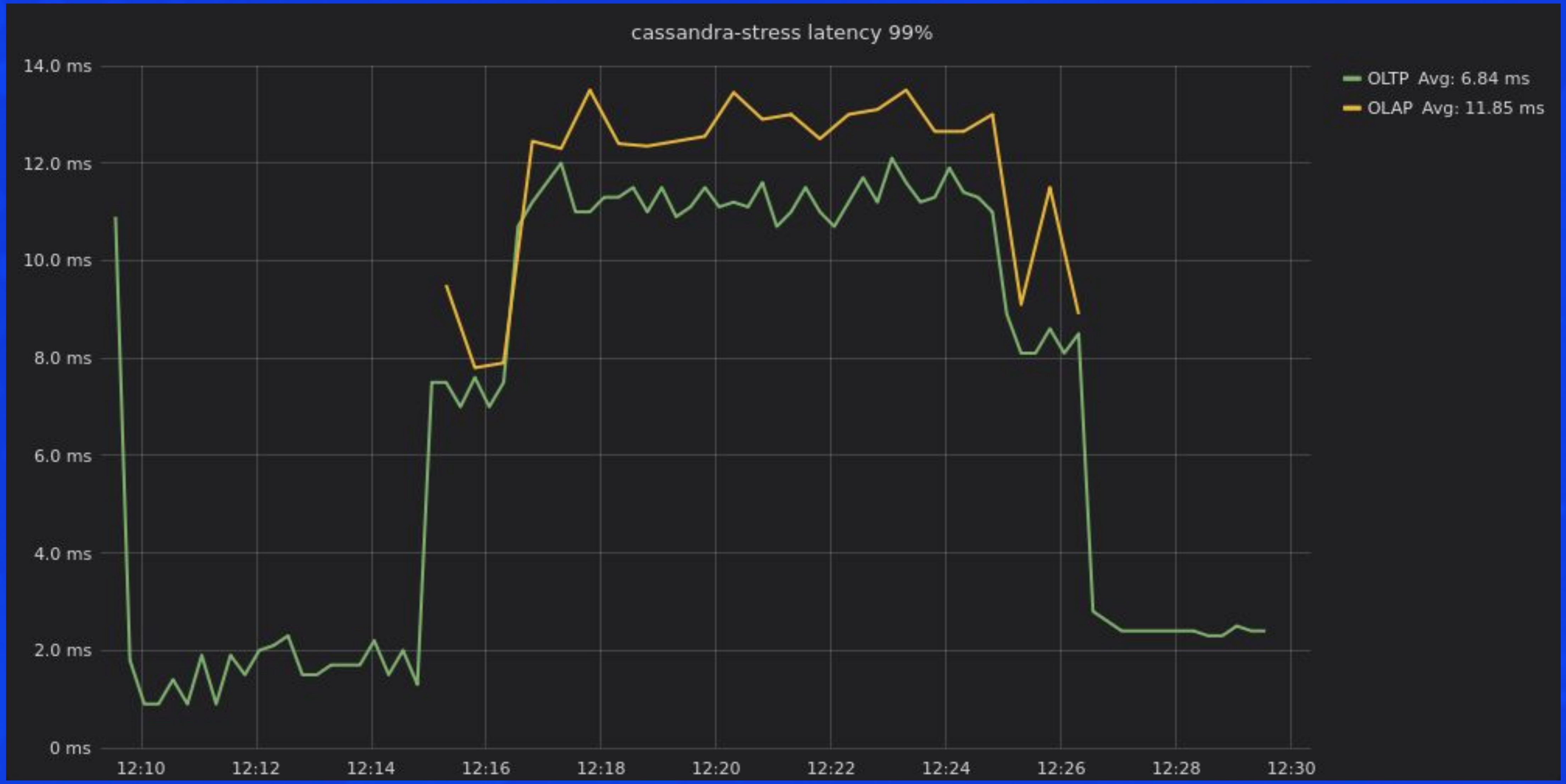
Name	ID
John	001
Karen	002
Bill	003

Grade	ID
Senior	001
Freshman	002
Junior	003

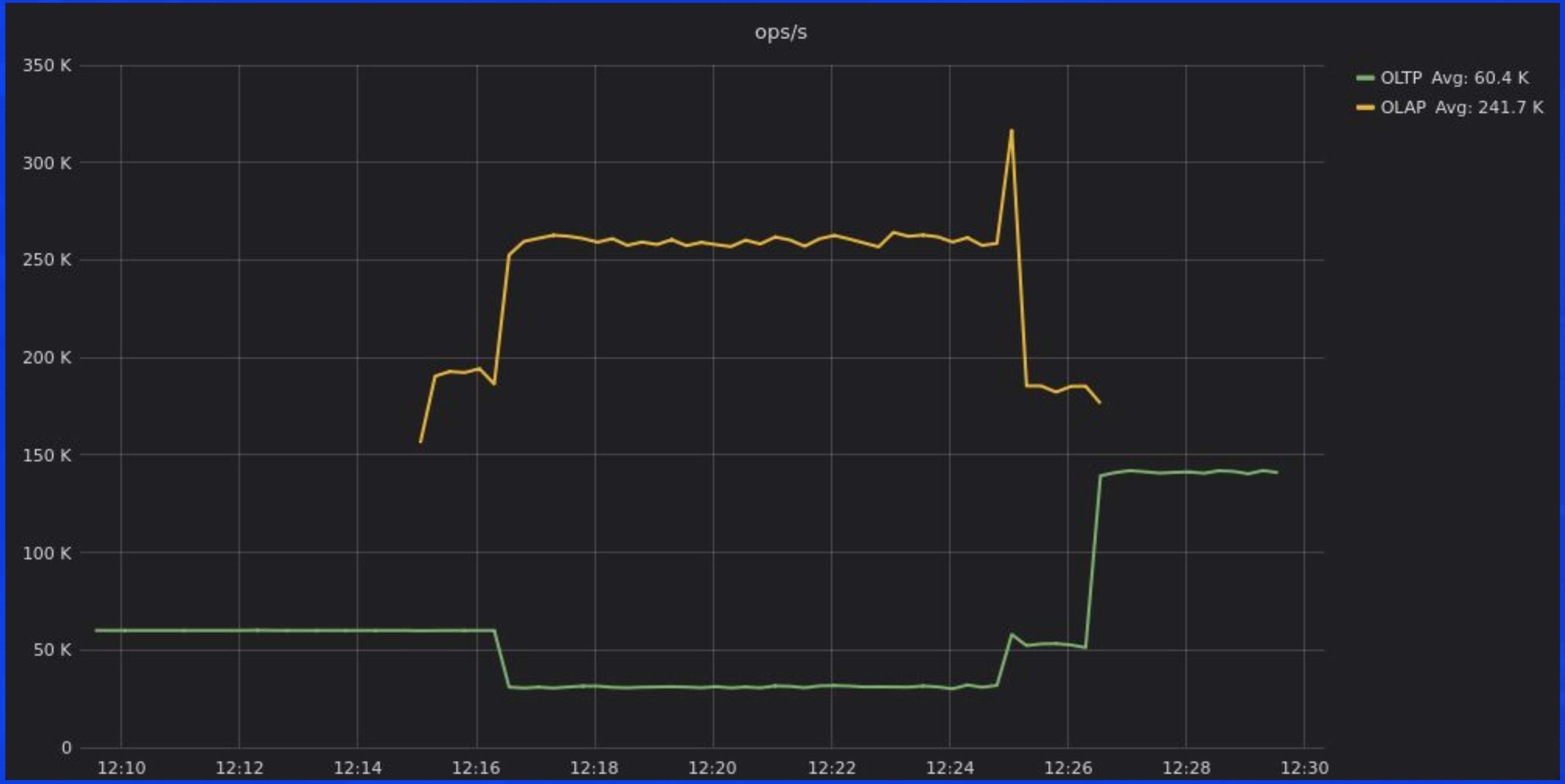
GPA	ID
4.00	001
3.67	002
3.33	003


OLTP
(online
transaction
processing)
$E = MC^2$
Analytics

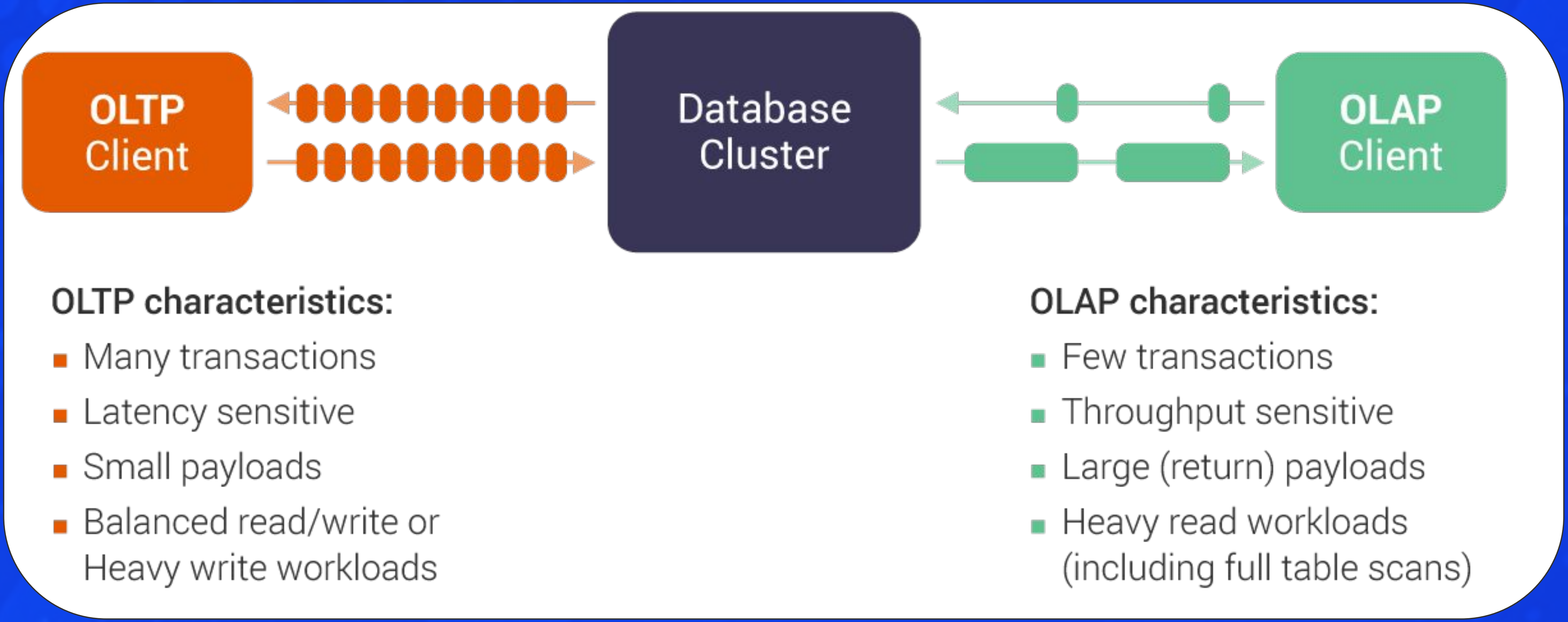
The Latency Problem



The Throughput Problem



Why Contention Happens?



OLTP characteristics:

- Many transactions
- Latency sensitive
- Small payloads
- Balanced read/write or Heavy write workloads

OLAP characteristics:

- Few transactions
- Throughput sensitive
- Large (return) payloads
- Heavy read workloads (including full table scans)

Physical Isolation

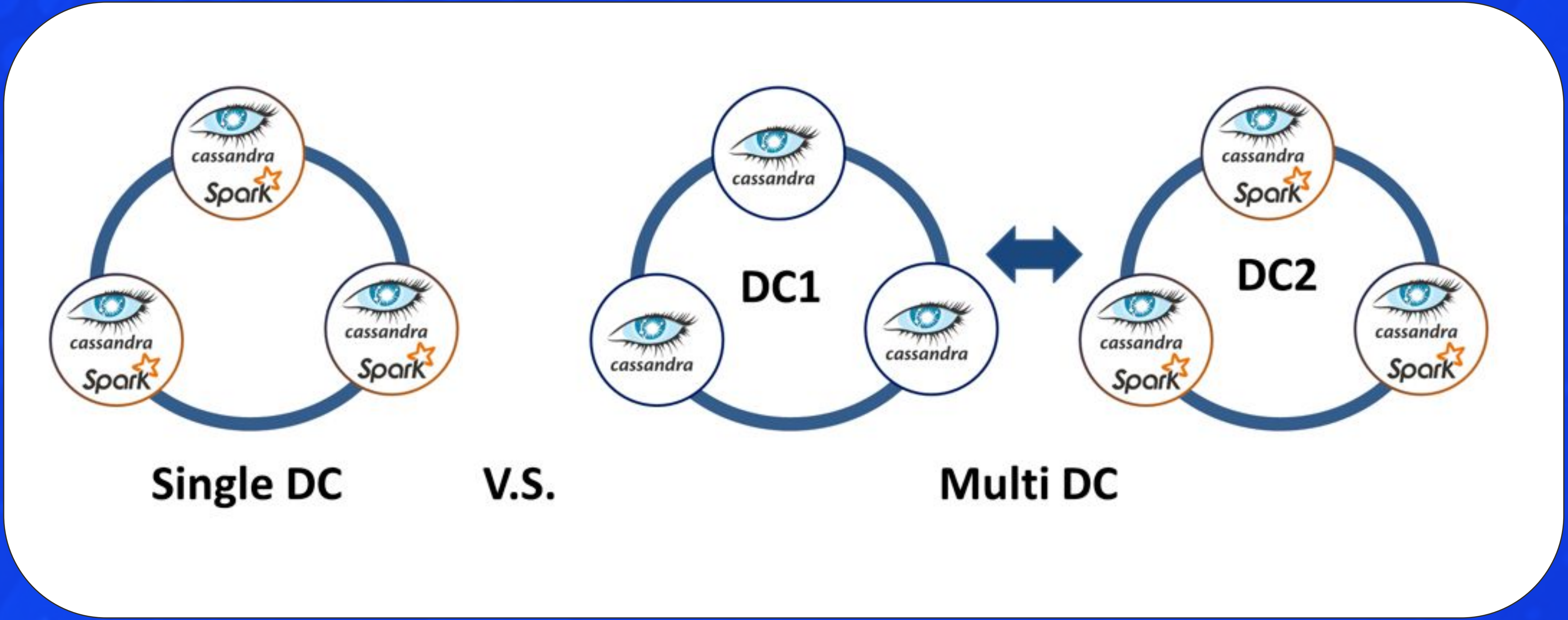


Image credits (and a bonus benchmark): [Instaclustr]

Also see: [GumGum: Multi Region Cassandra in AWS]

Scheduled Isolation

- Involve running Analytics during specific windows
  - eg: Off-peak hours
  - Contention still likely to manifest, but less critical
- Still requires concurrency tuning & fine-graining
- Over time jobs may take longer
  - Potential for overrun and end up in the same state…
  - … Or abort, and miss your SLOs

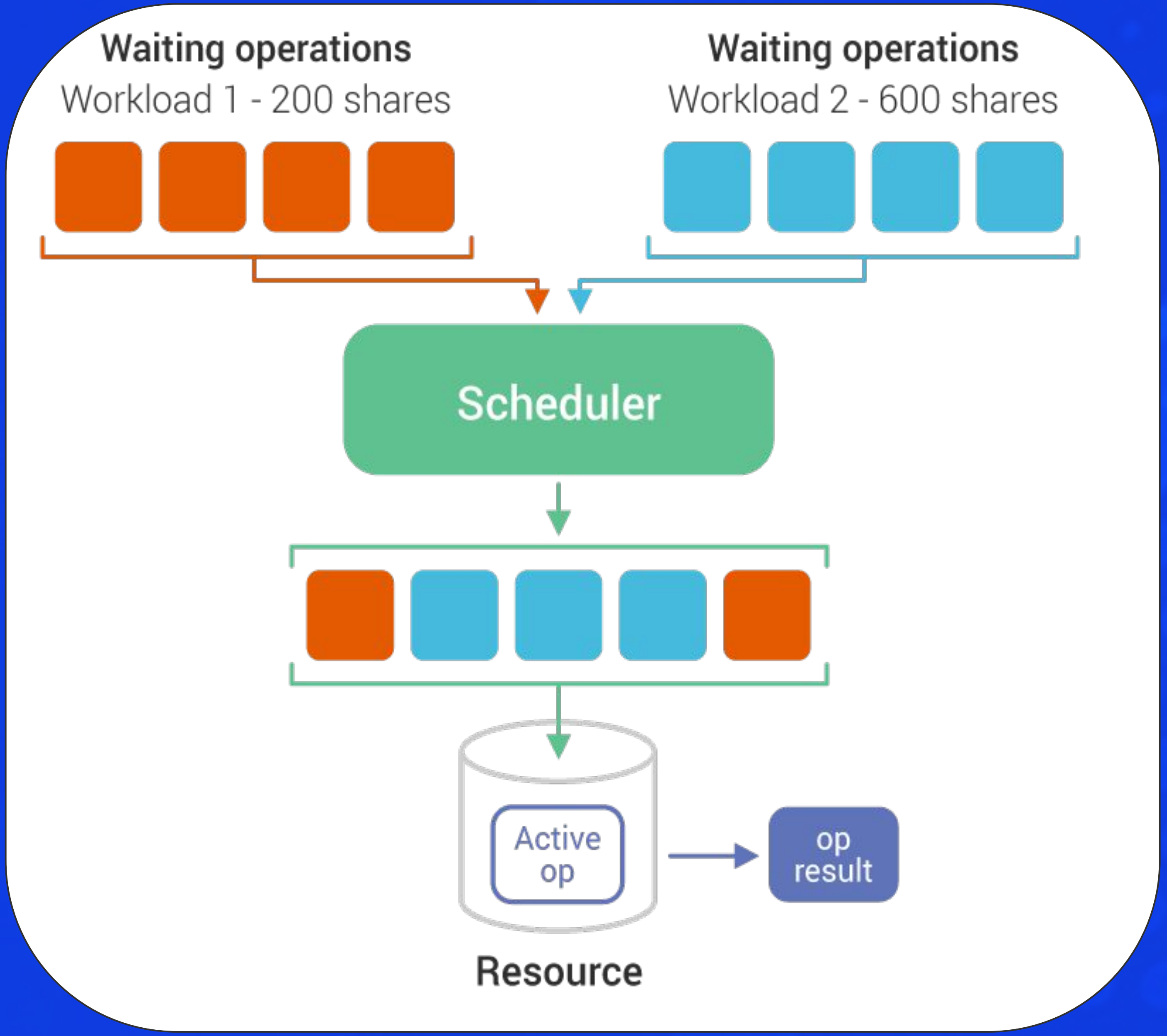
ScyllaDB Solution: Workload Prioritization

For a primary workload:

```
CREATE SERVICE LEVEL main WITH shares = 600
```

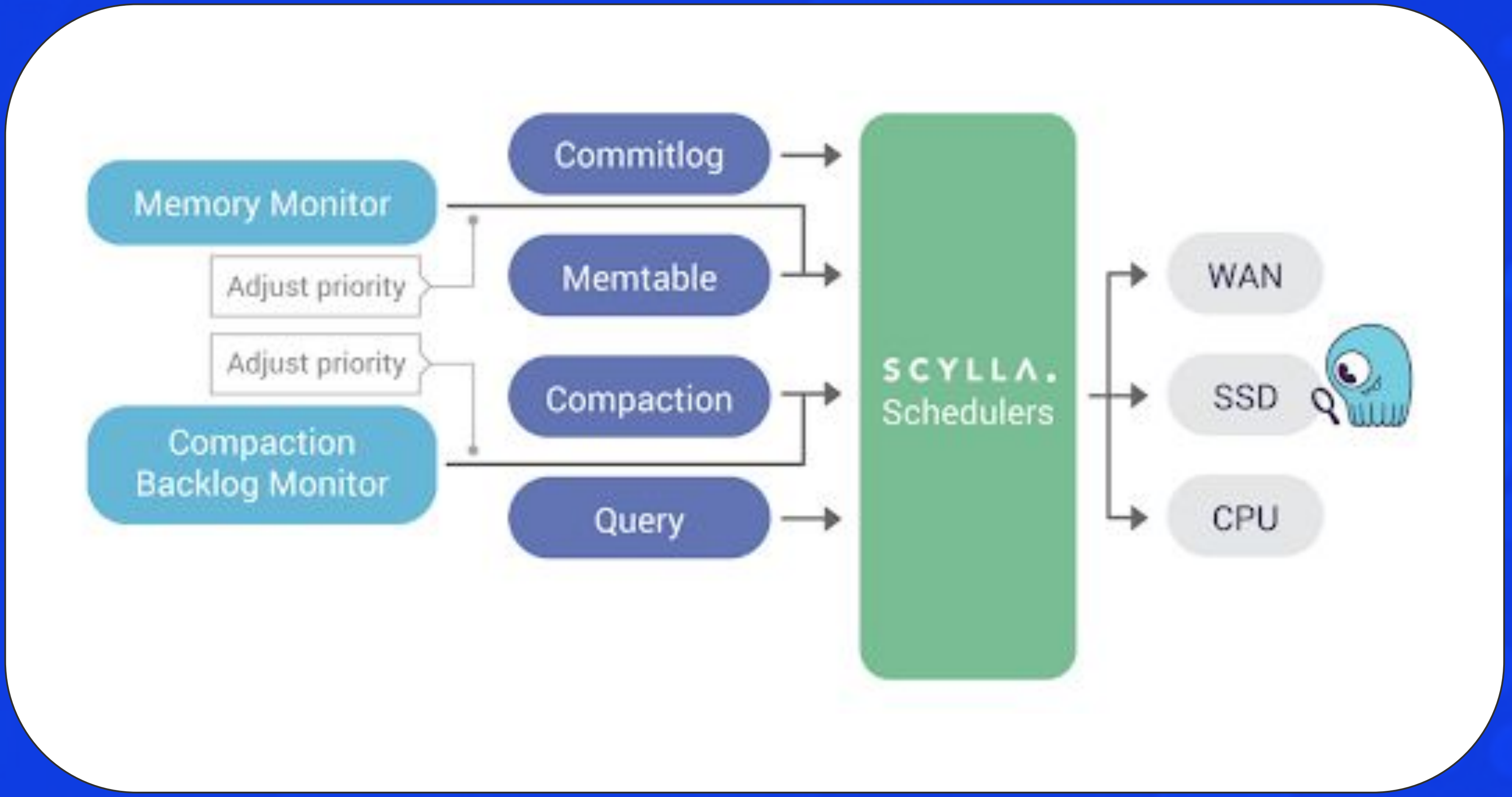
For a secondary one:

```
CREATE SERVICE LEVEL secondary WITH shares = 200
```



Under the Hood

- Resource isolation is not a new ScyllaDB concept
- Makes use of [Seastar Scheduling Groups]
- Auto-tuning
- Used by:
  - Compaction
  - Streaming
  - Memtables
  - User facing workload
  - Workload Prioritization



Prioritization and isolation is simply NOT enough

Workloads Characteristics: Time

The timeout dilemma:

1. Timeout should follow: $Tserver \leq Tclient$
2. For OLTP:
   - Can’t be too high
   - Incurs retries or dropped requests
   - Excessive retries result in wasted resources
3. For Analytics:
   - Can’t be too low
   - Otherwise Batch will likely fail
   - High throughput will typically increase latencies due to contention

Workloads Characteristics: Shedding

Overload response:

1. Interactive workload:
   - Throttling won't help
   - Delaying response to user A will not cause some user B to stop sending requests
   - Unbound concurrency
2. Batch workload:
   - Just throttle
   - Allow us to have a knob that controls the pace of the analytics workload
   - Bounded concurrency

Introducing Workload Characterization

Ideally – we want our database to behave differently:

- For Real-time:
  - Have low timeout
  - Load shedding (fail excessive requests), as the database can NOT slow down interactive workloads.
  - Dedicate most of the resources to this workload.
- For Batch:
  - Relatively higher timeout
  - Apply back-pressure via throttling
  - Use mostly unused resources

Introducing Workload Characterization

Why not just hint the database with specifics?

- For Real-Time:

Have low timeout (30ms)	timeout=30ms
Load shedding	AND workload_type=interactive
Dedicate most of the resources*	AND shares=800

- For Analytics:

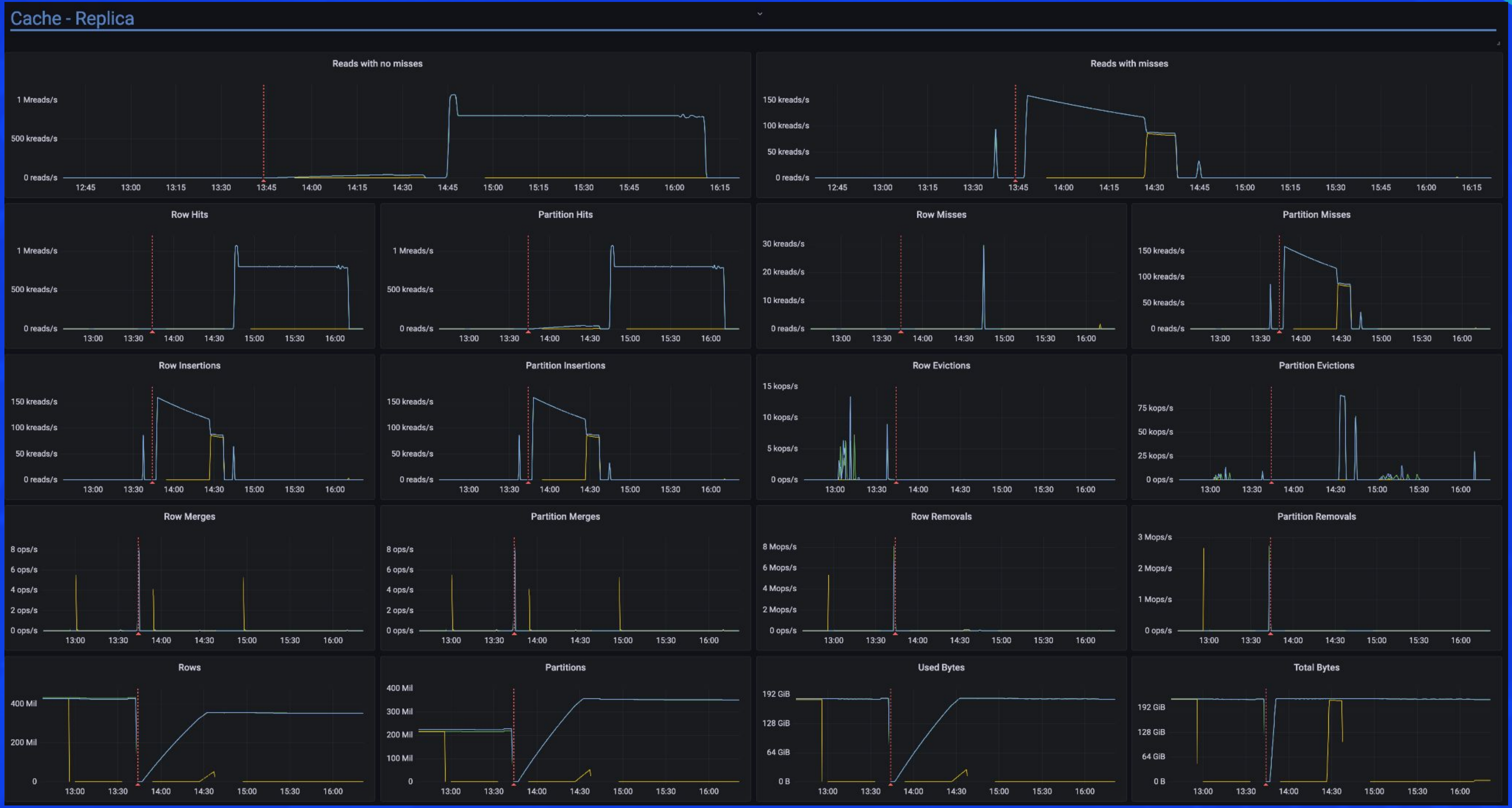
Have relatively high timeout (5s)	timeout=5s
Throttling	AND workload_type=batch
Use mostly unused resources*	AND shares=200

* ScyllaDB Enterprise only

Are we finally done?

… Not really! ;-)

Cache Pollution?



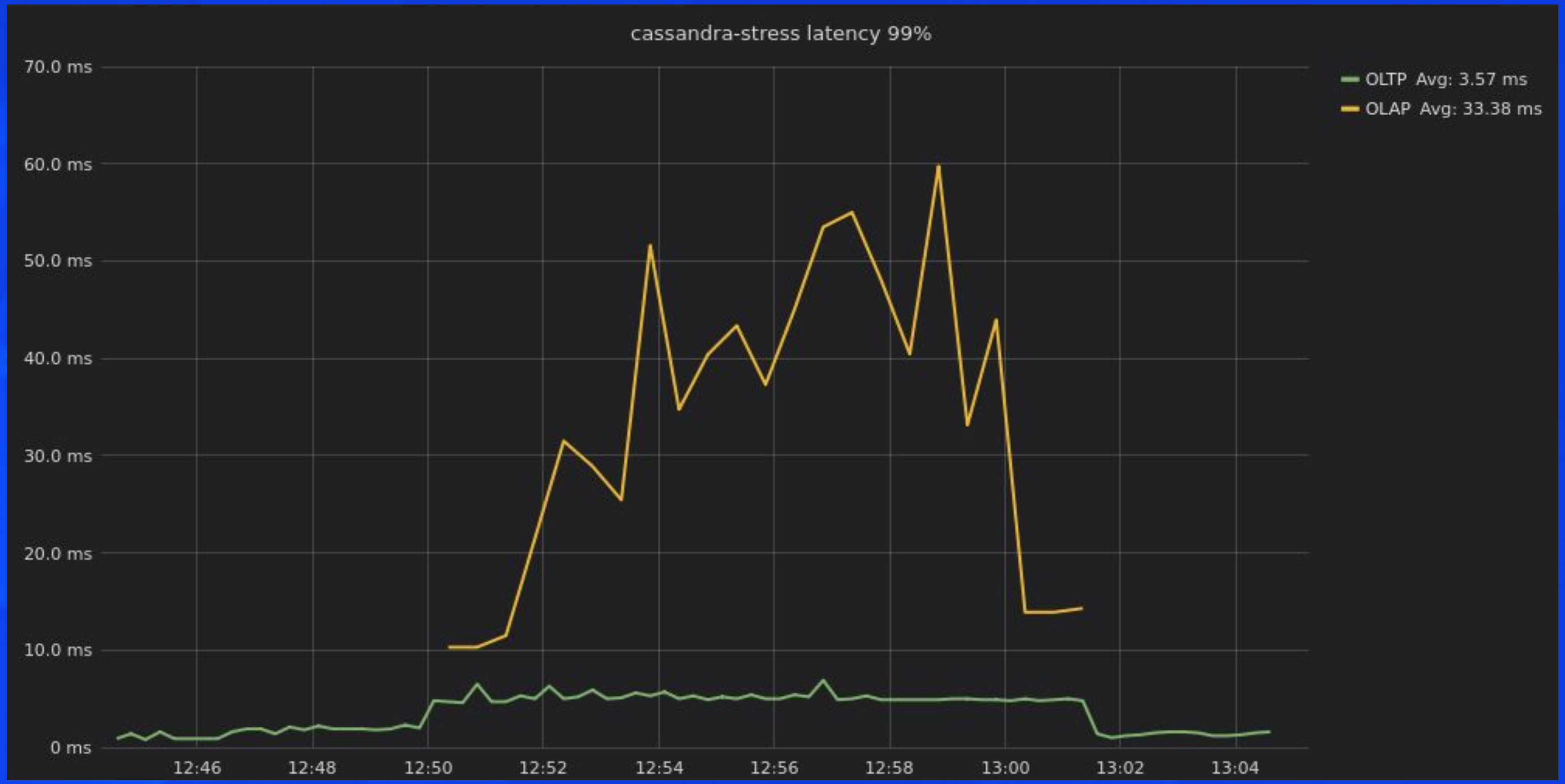
Bypass Cache!

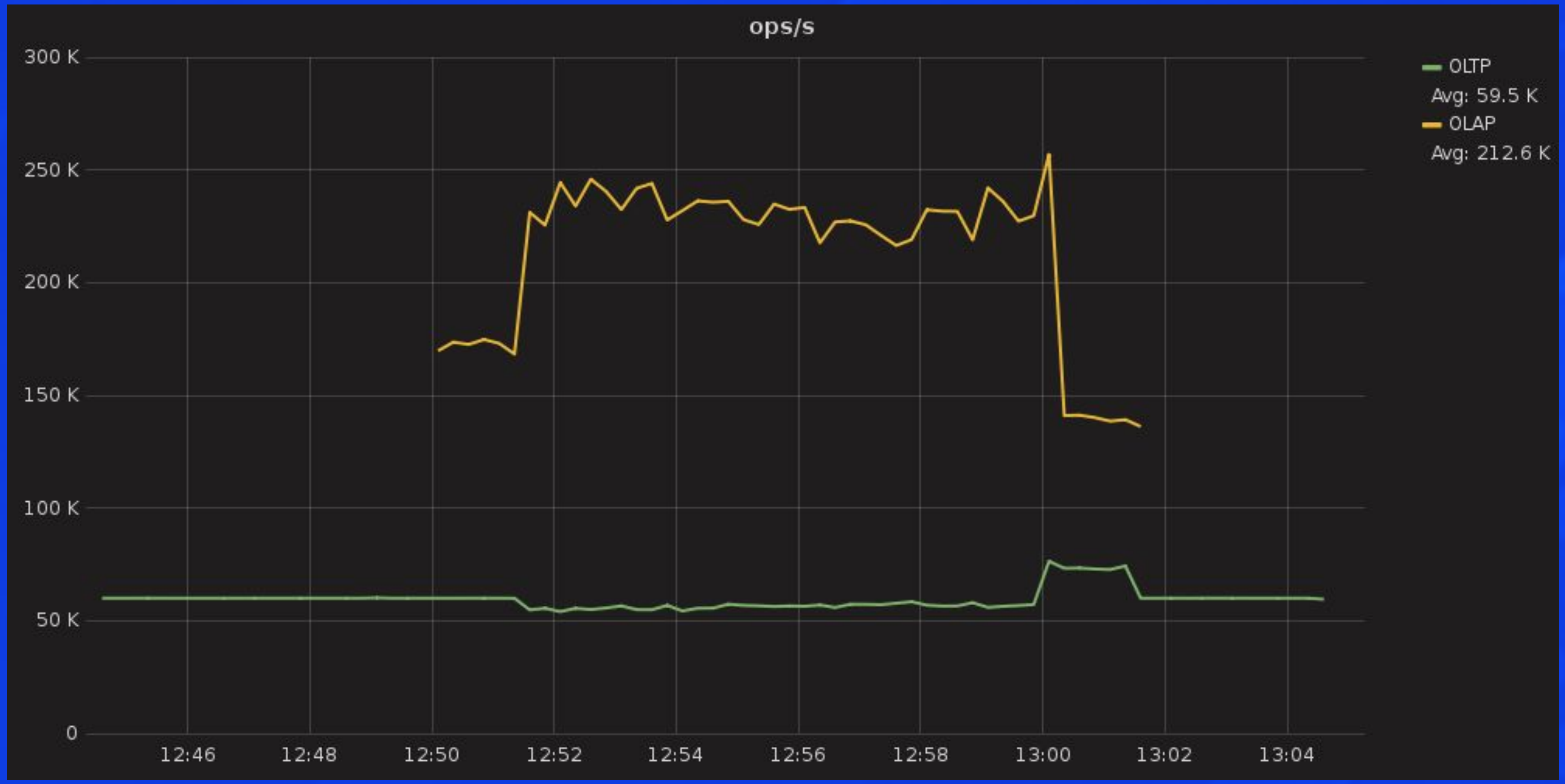
```
SELECT * FROM users BYPASS CACHE;
```

```
SELECT name, occupation FROM users WHERE userid IN (199, 200, 207)
BYPASS CACHE;
```

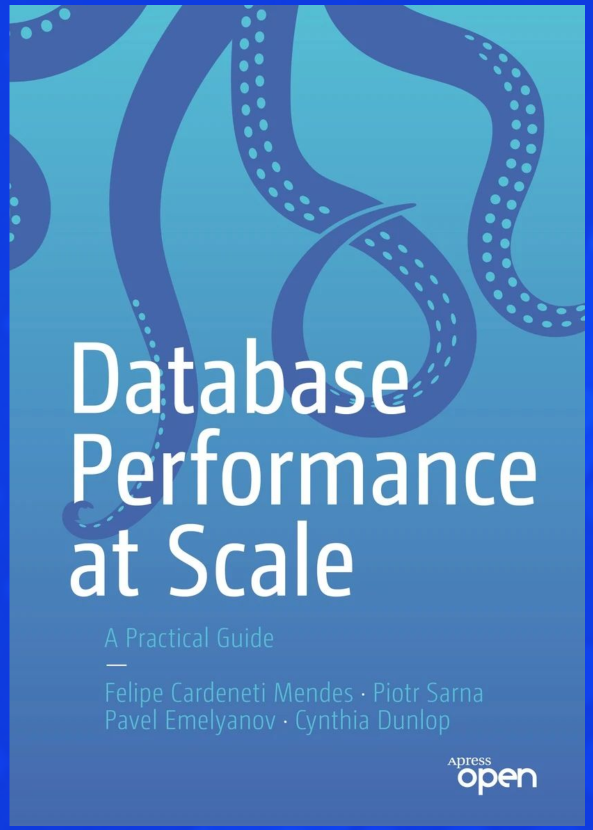
```
SELECT * FROM users WHERE birth_year = 1981 AND country = 'FR' ALLOW
FILTERING BYPASS CACHE;
```

Finally… The Results!





Q&A



Database Performance at Scale

A Practical Guide

Felipe Cardeneti Mendes · Piotr Sarna
Pavel Emelyanov · Cynthia Dunlop

Apress open

Thank You!

Find me on LinkedIn!

ScyllaDB

Database Performance at Scale: A Practical Guide