



TiDB as an HTAP Database

主讲人：PingCap CEO 刘奇



About me

- CEO and co-founder of PingCAP
- Open source hacker
- Infrastructure engineer
- Founder of TiDB/TiKV/Codis
- Infrastructure software engineer
- Wandoulabs/JD

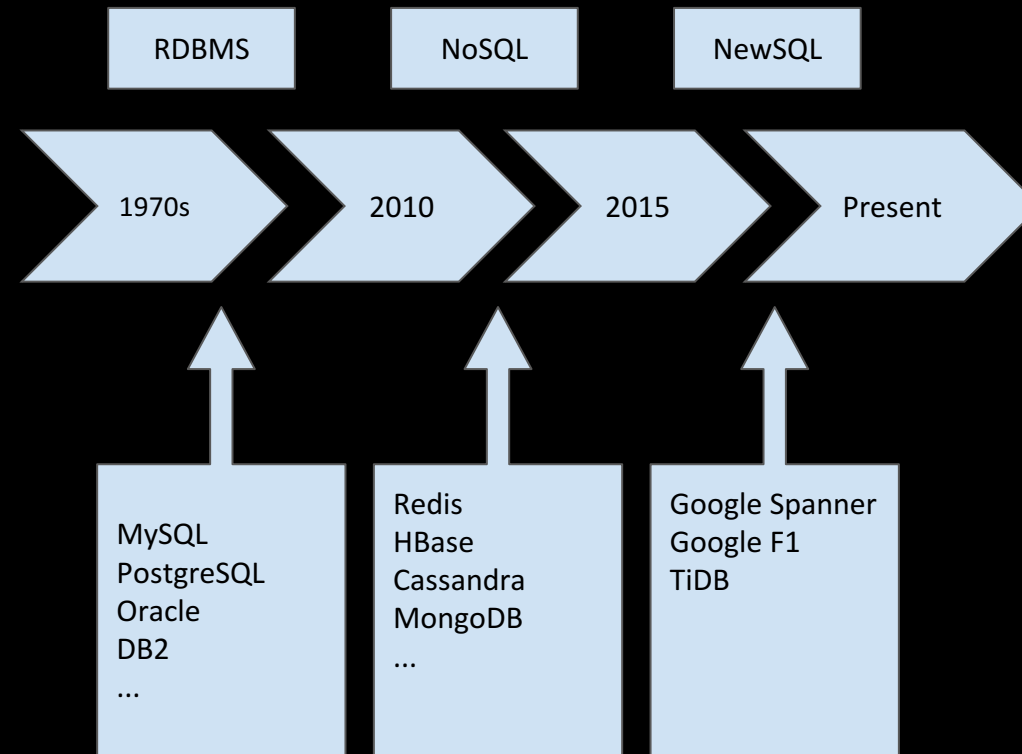


Why a new database?



Brief History

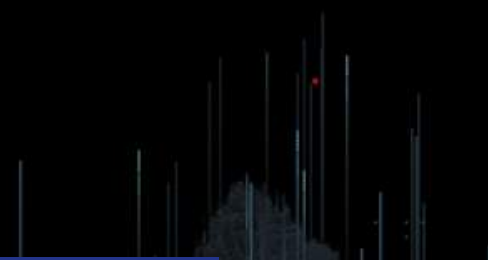
- Standalone RDBMS
- NoSQL
- Middleware & Proxy
- NewSQL





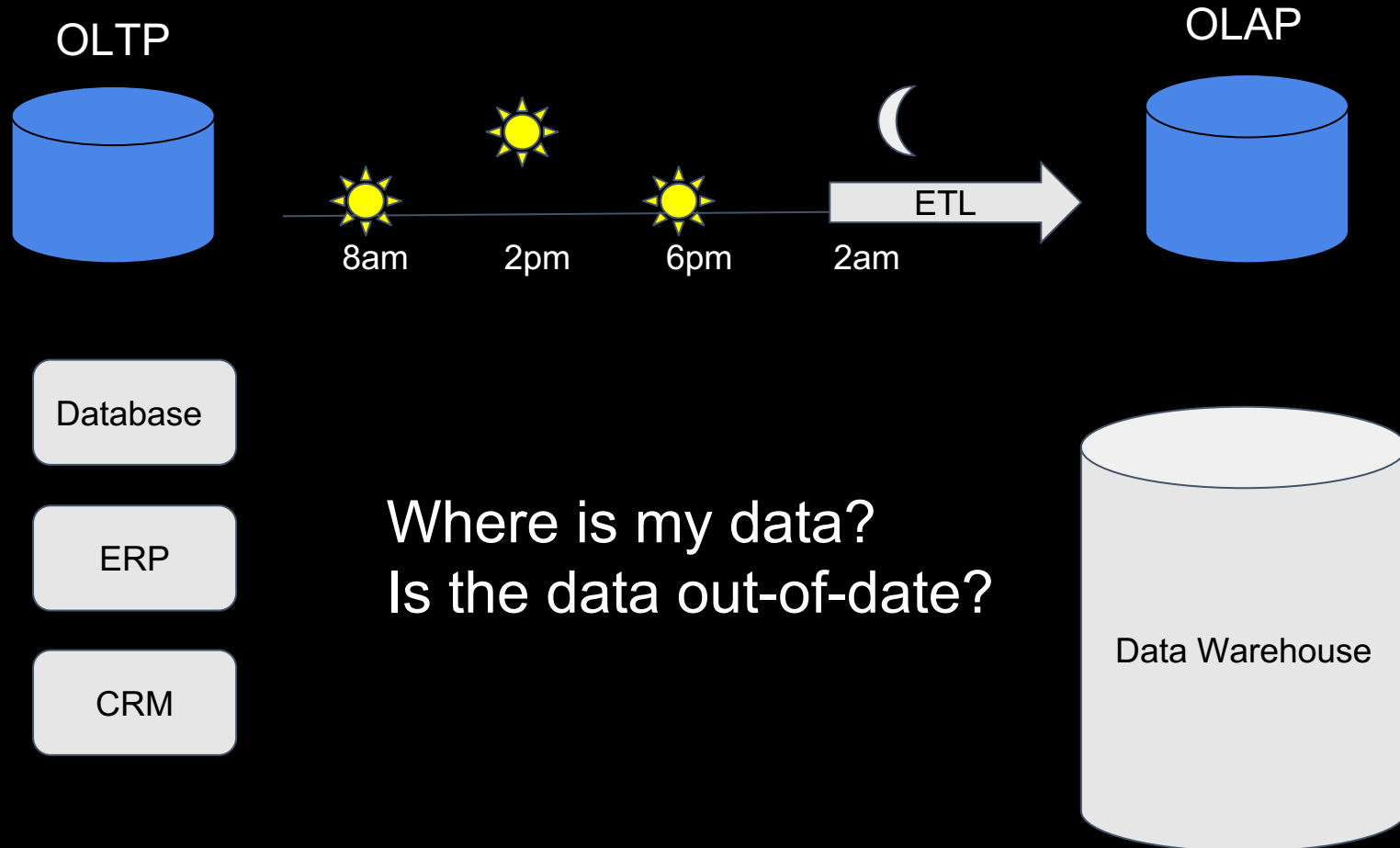
NewSQL Database

- Horizontal Scalability
- ACID Transaction
- High Availability
- SQL at Scale





OLTP & OLAP





Why two separate systems

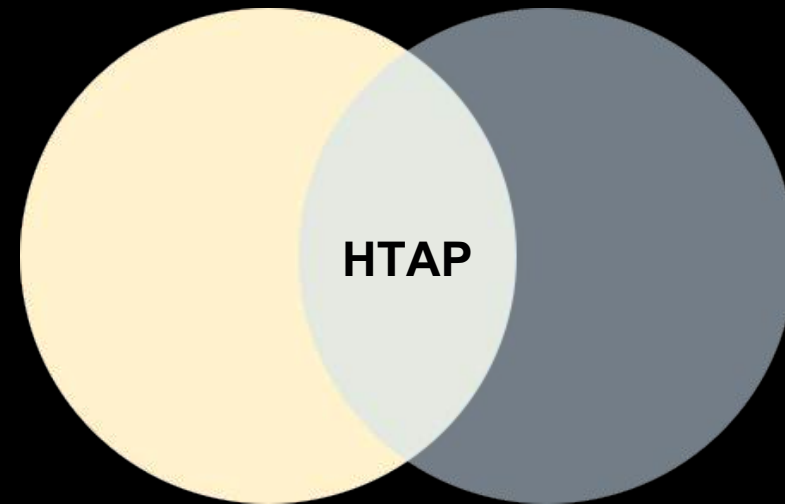
- Huge data size
- Complex query logic
- Latency VS Throughput
- Point query VS Full range scan
- Transaction & Isolation level



OLAP + OLTP = HTAP

Hybrid Transactional / Analytical Processing

- ACID Transaction
- Real-time analysis
- SQL





How do we build the new database

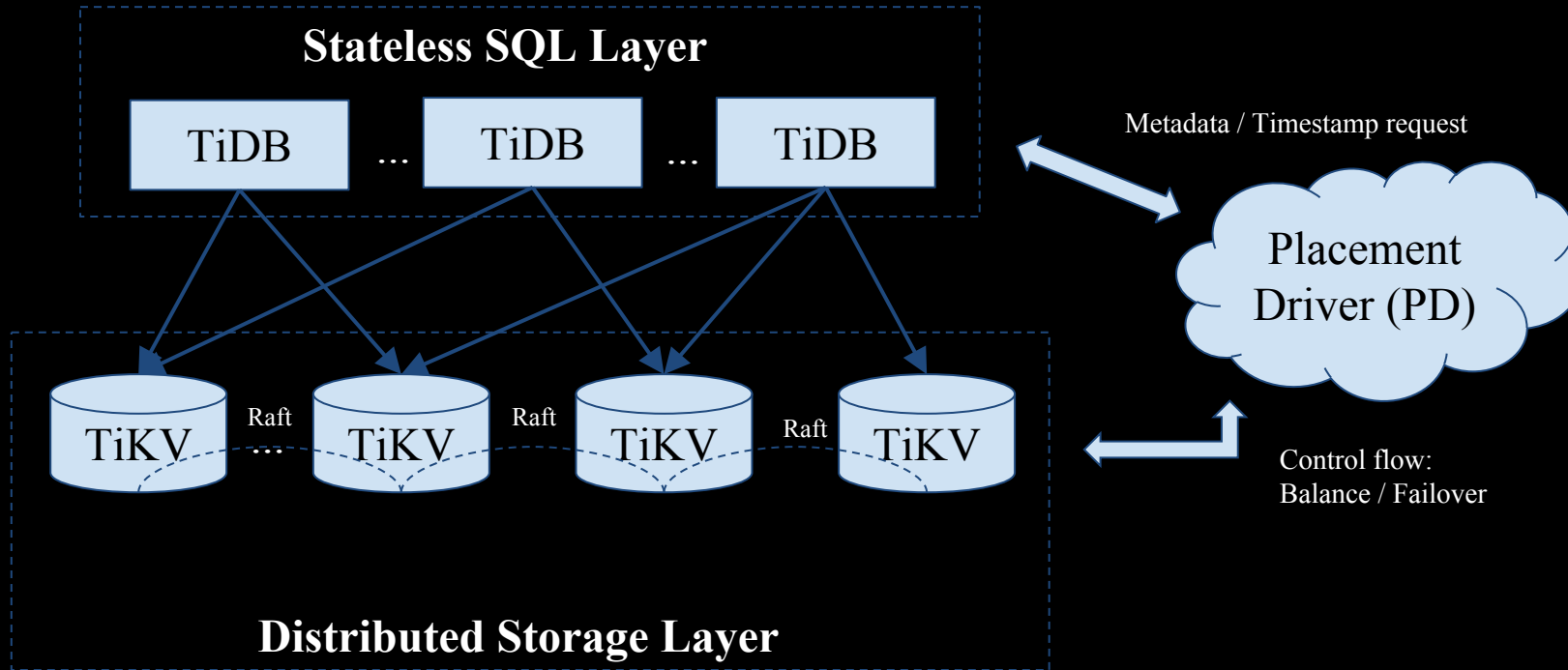


What is TiDB

- Scalability as the first class feature
- SQL is necessary
- Compatible with MySQL, in most cases
- OLTP + OLAP = HTAP (Hybrid Transactional/Analytical Processing)
- 24/7 availability, even in case of datacenter outages
- Open source, of course



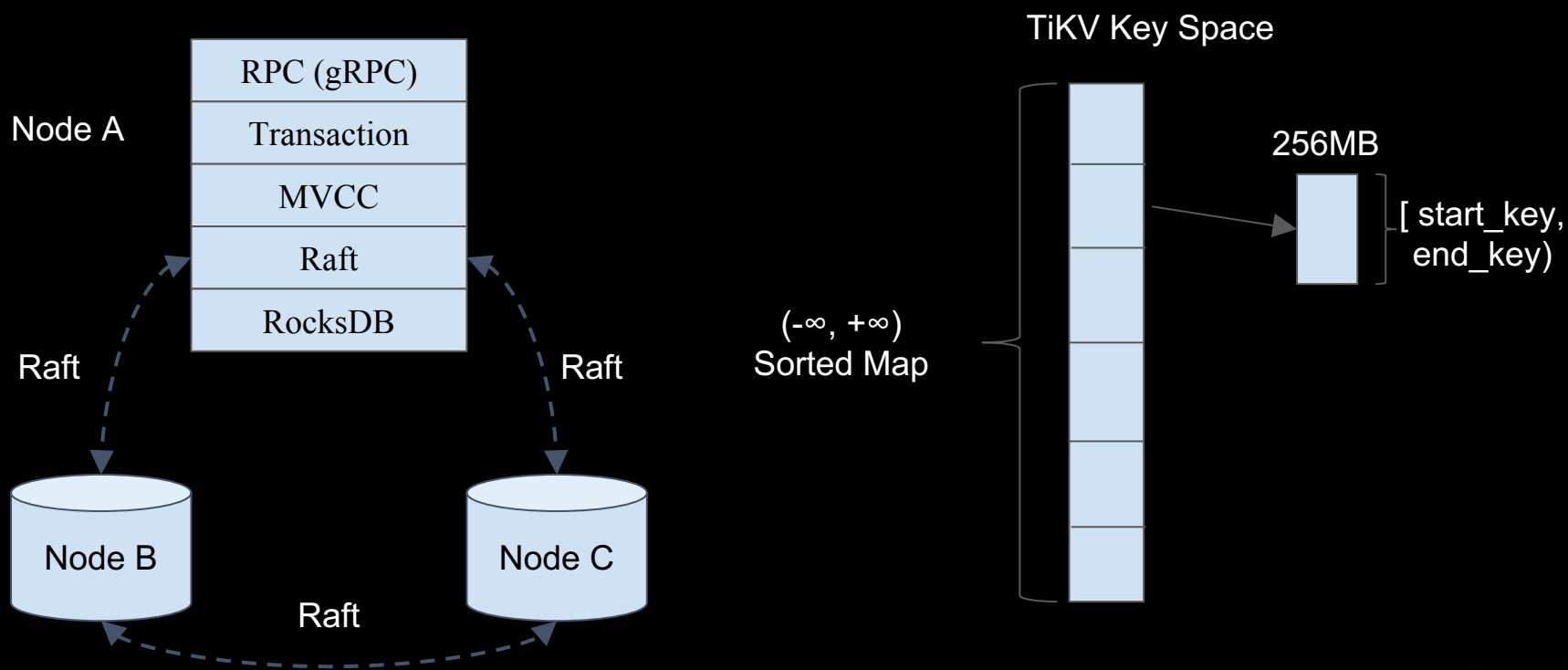
Architecture





TiKV - Overview

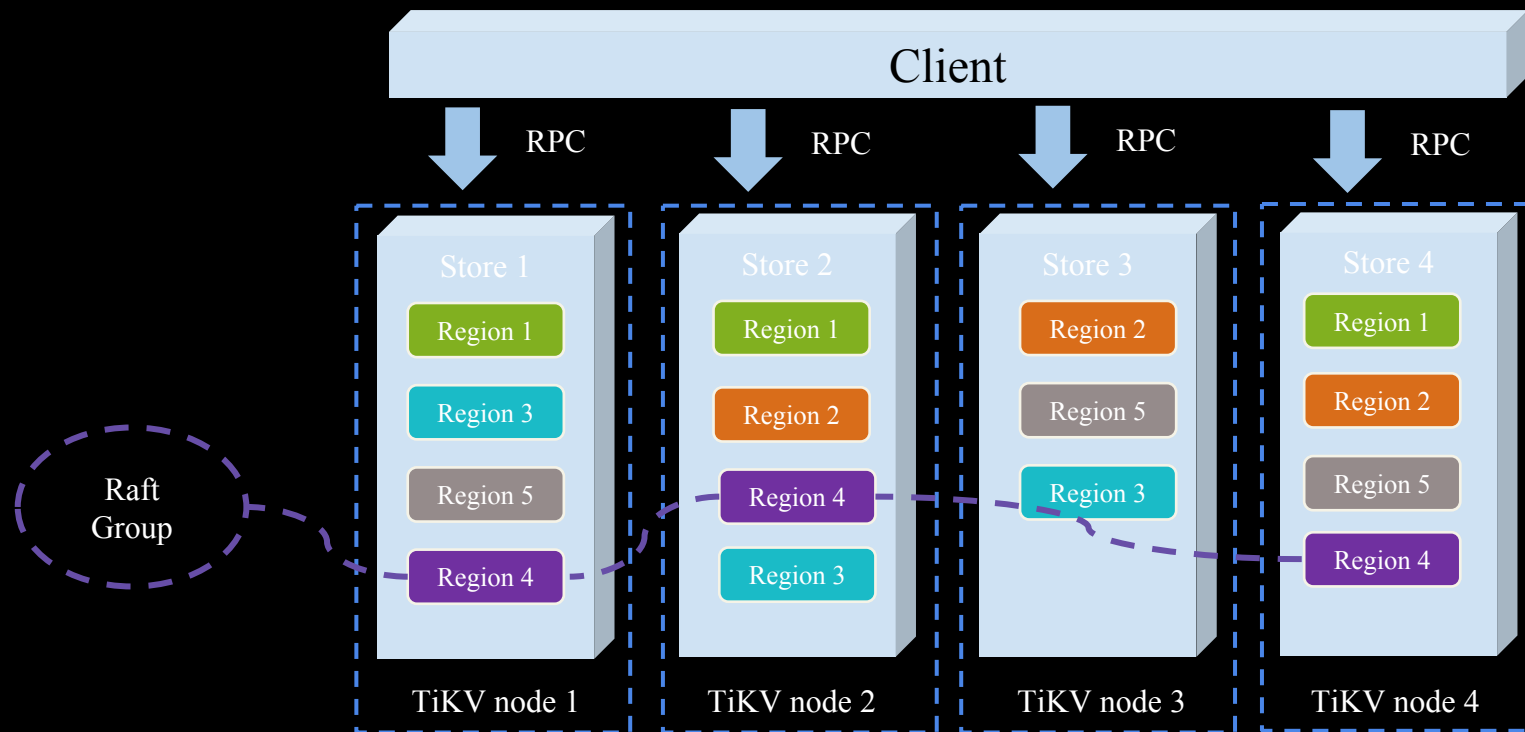
- Region: a set of continuous key-value pairs
- Data is organized/stored/replicated by Regions
- Highly layered





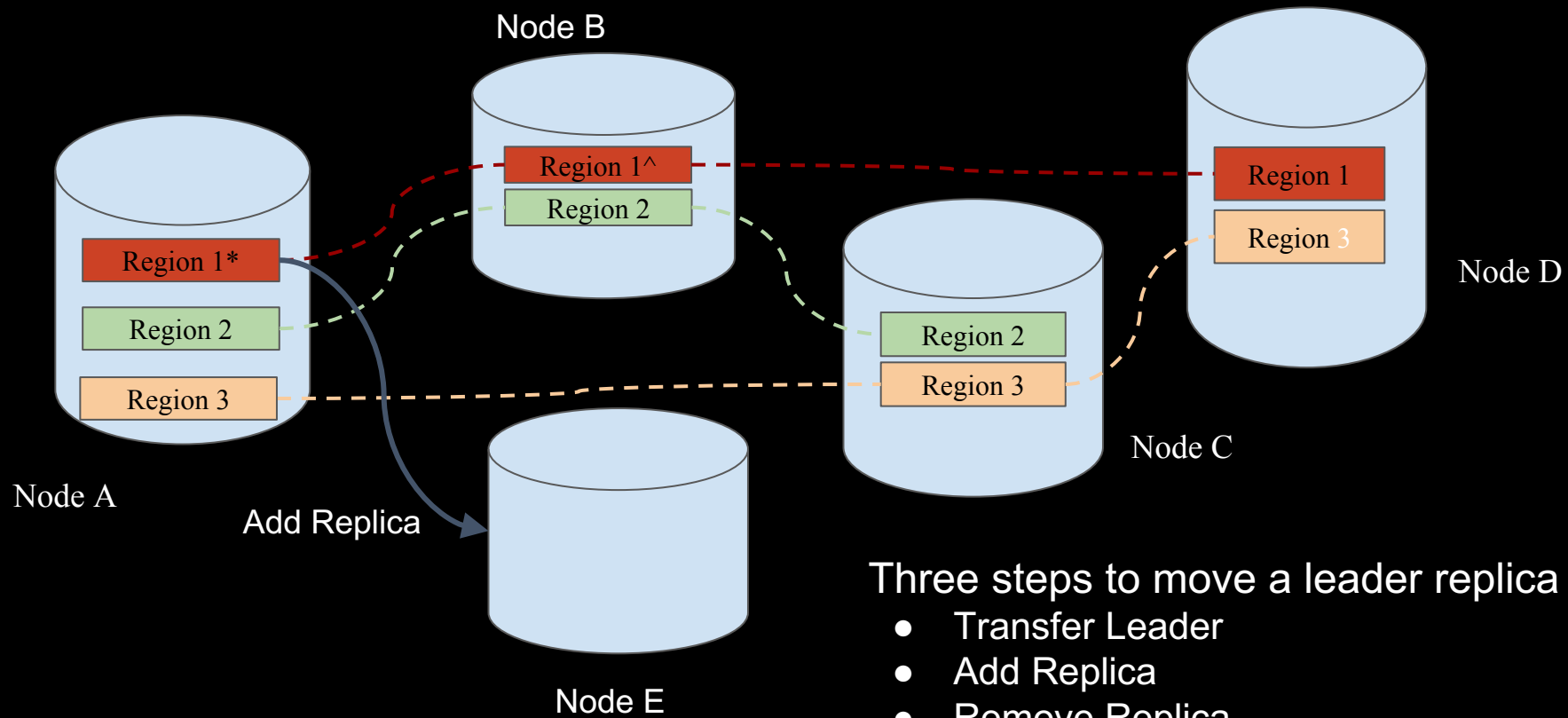
TiKV - Multi-Raft

Multiple raft groups in the cluster, one group for each region.





TiKV - Horizontal Scale



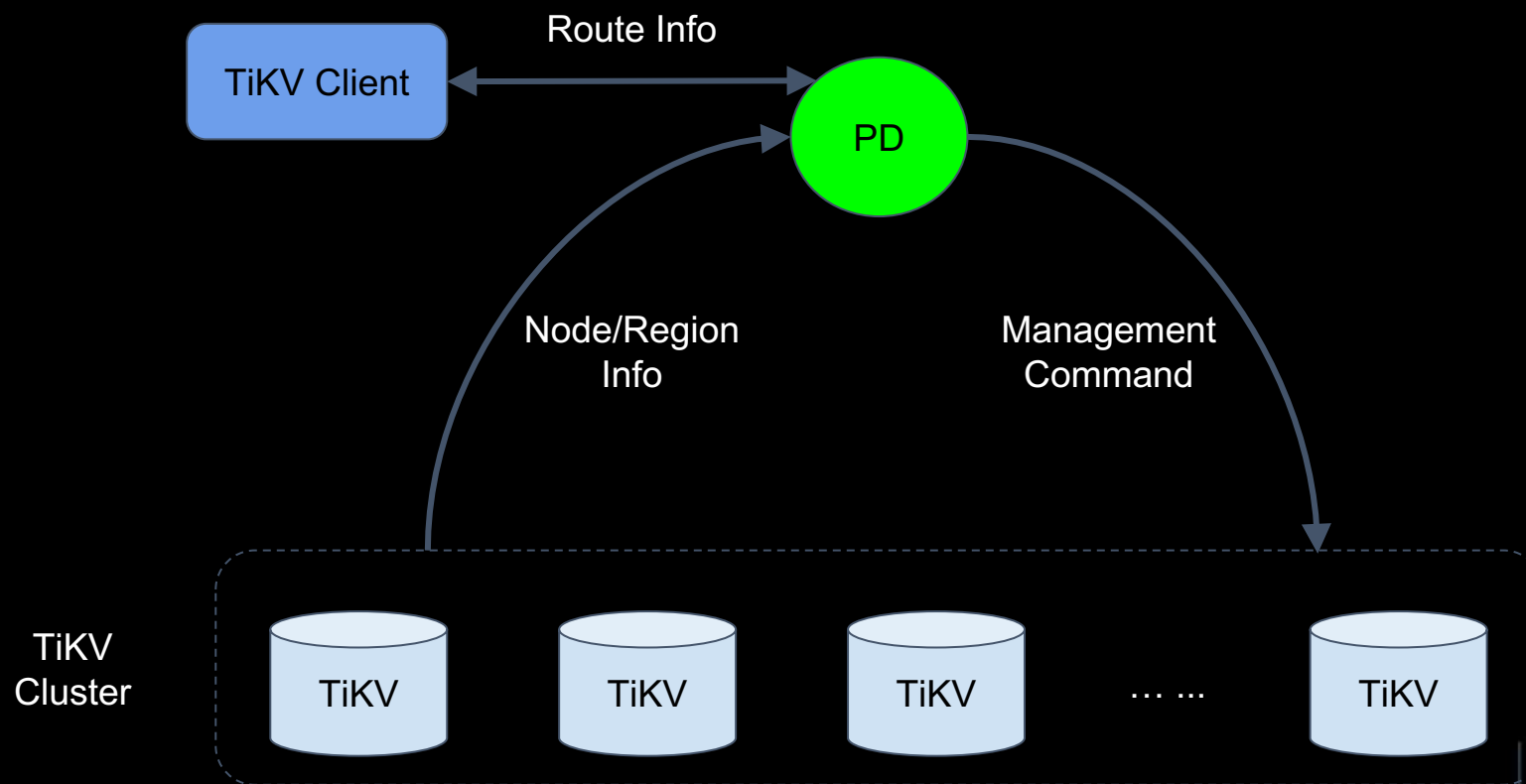
Three steps to move a leader replica

- Transfer Leader
- Add Replica
- Remove Replica



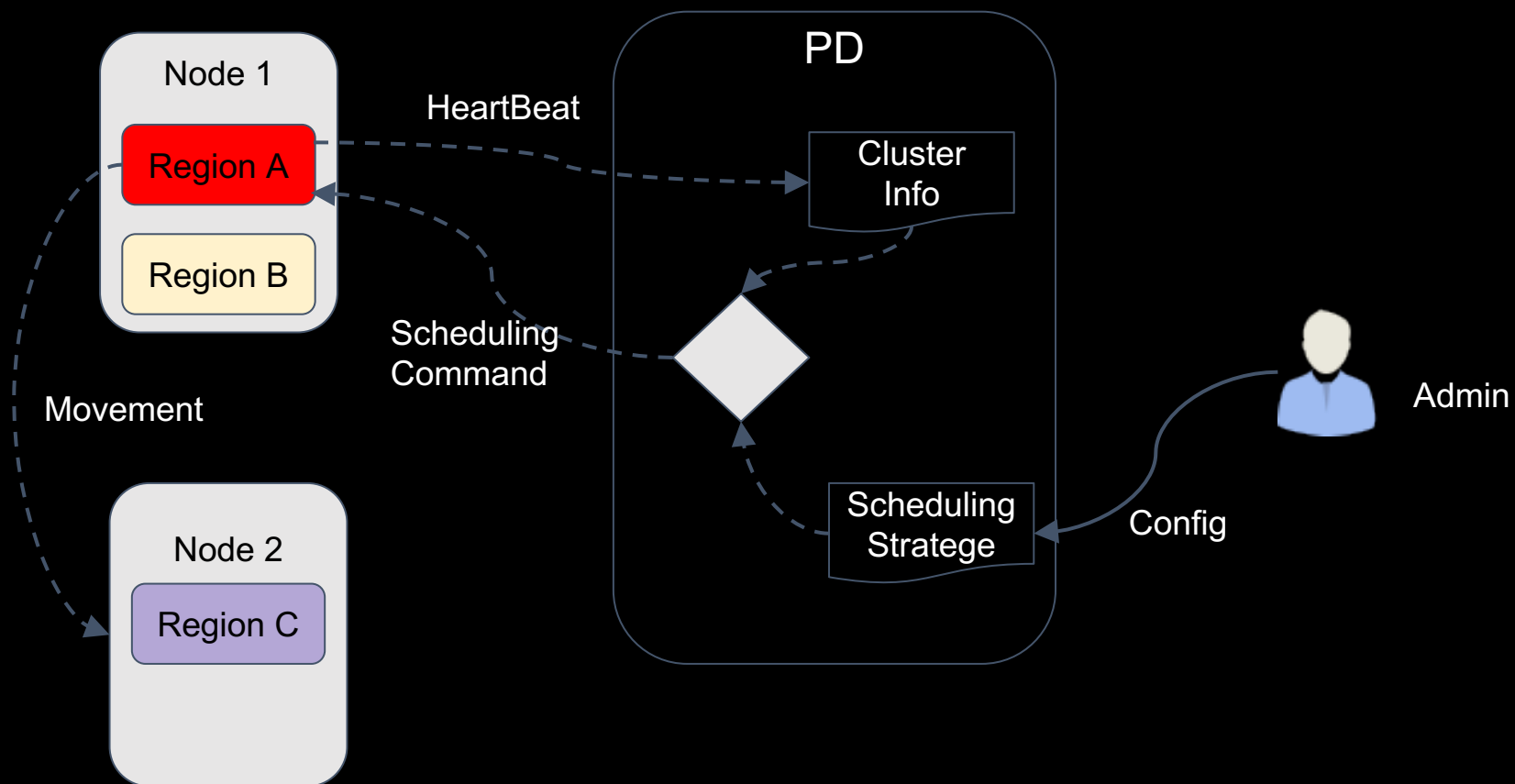
PD - Overview

- Meta data management
- Load balance management





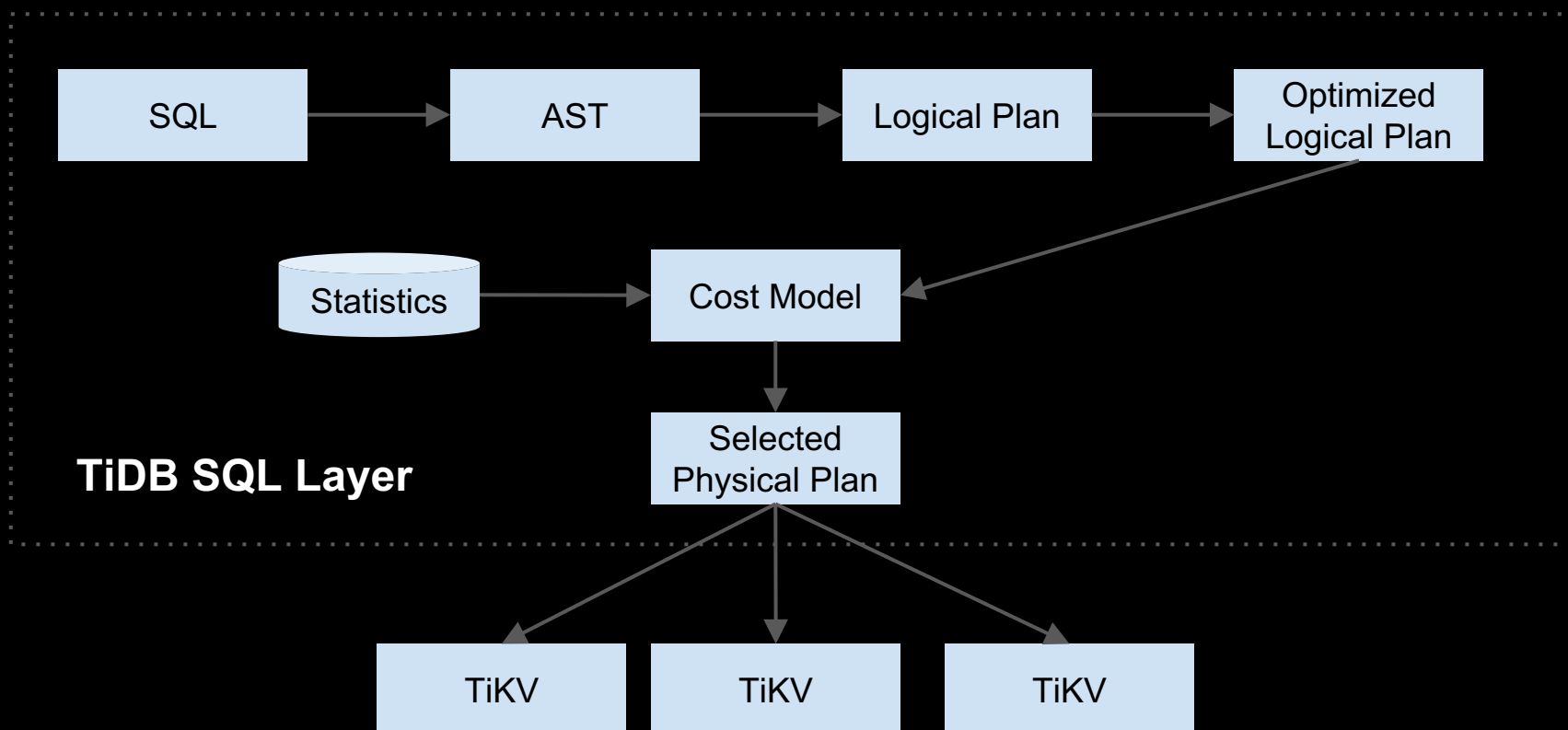
PD - TiKV Cluster Management





TiDB - Overview

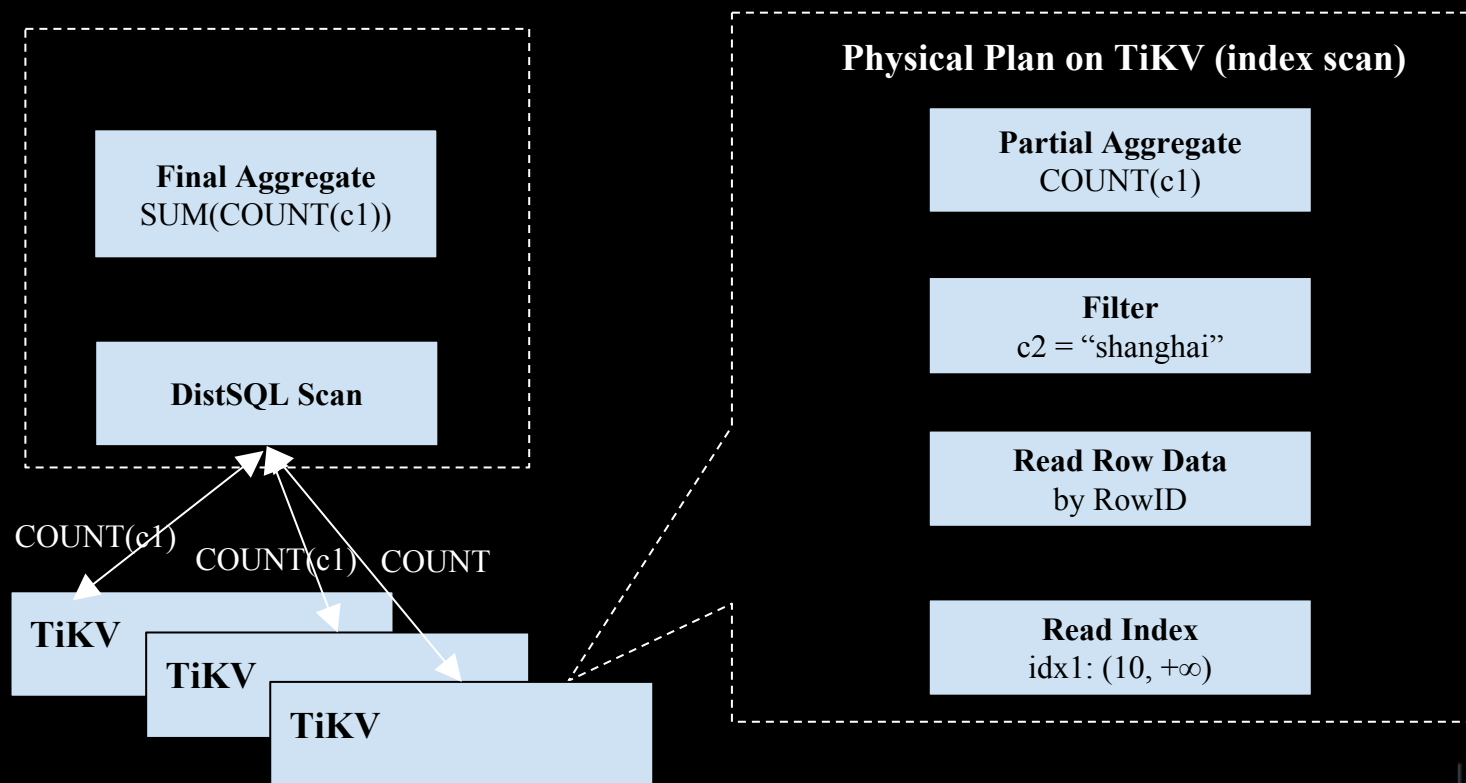
- The stateless SQL layer





TiDB - Distributed SQL

```
SELECT COUNT(c1) FROM t WHERE c1 > 10 AND c2 = 'shanghai';
```



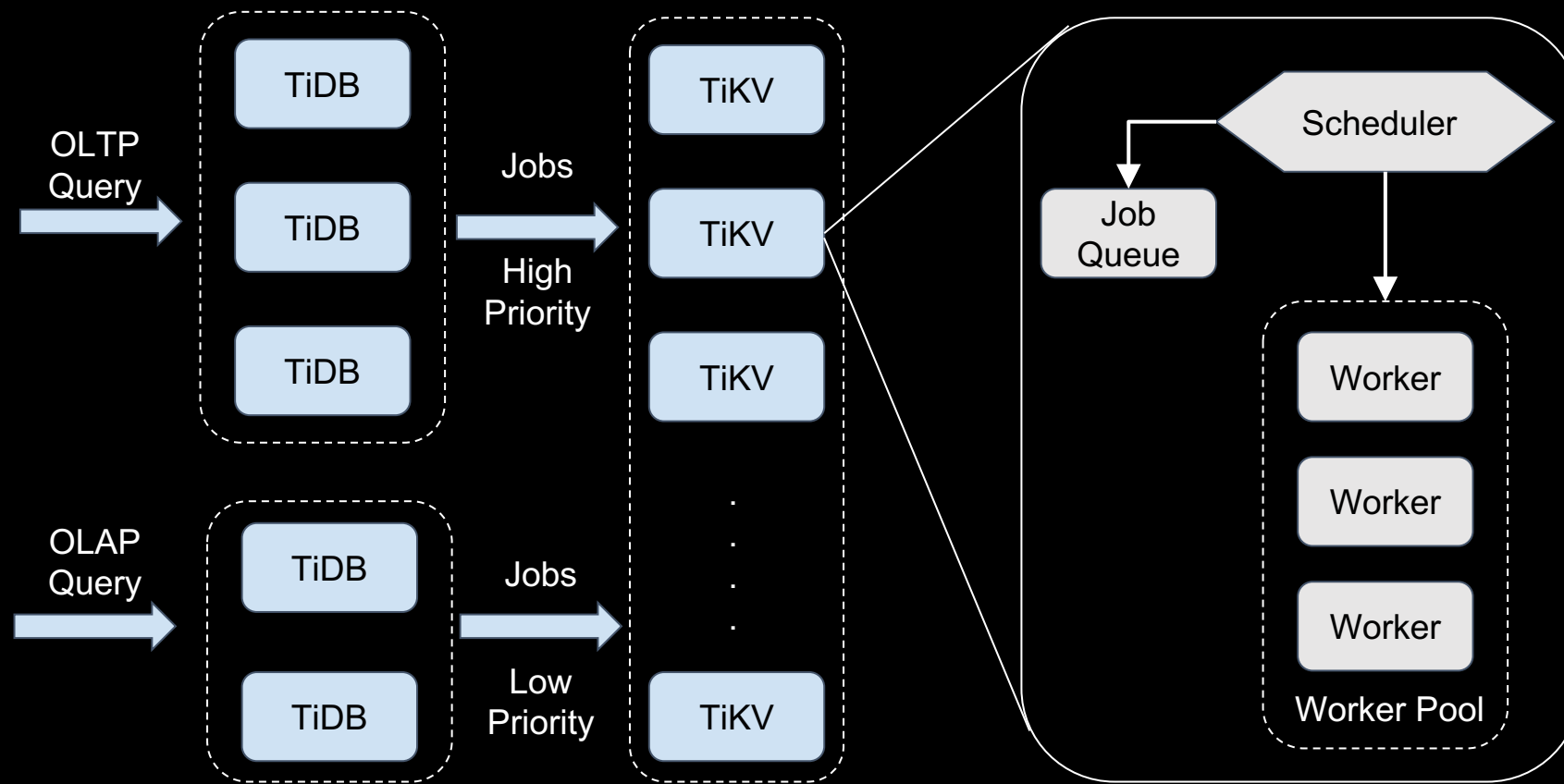


TiDB - Cost Based Optimizer

- Predicate Pushdown
- Column Pruning
- Eager Aggregate
- Convert Subquery to Join
- Statistics framework
- CBO Framework
 - Index Selection
 - Join Operator Selection
 - Hash join
 - Index lookup join
 - Sort-merge join
 - Stream Operators VS Hash Operators

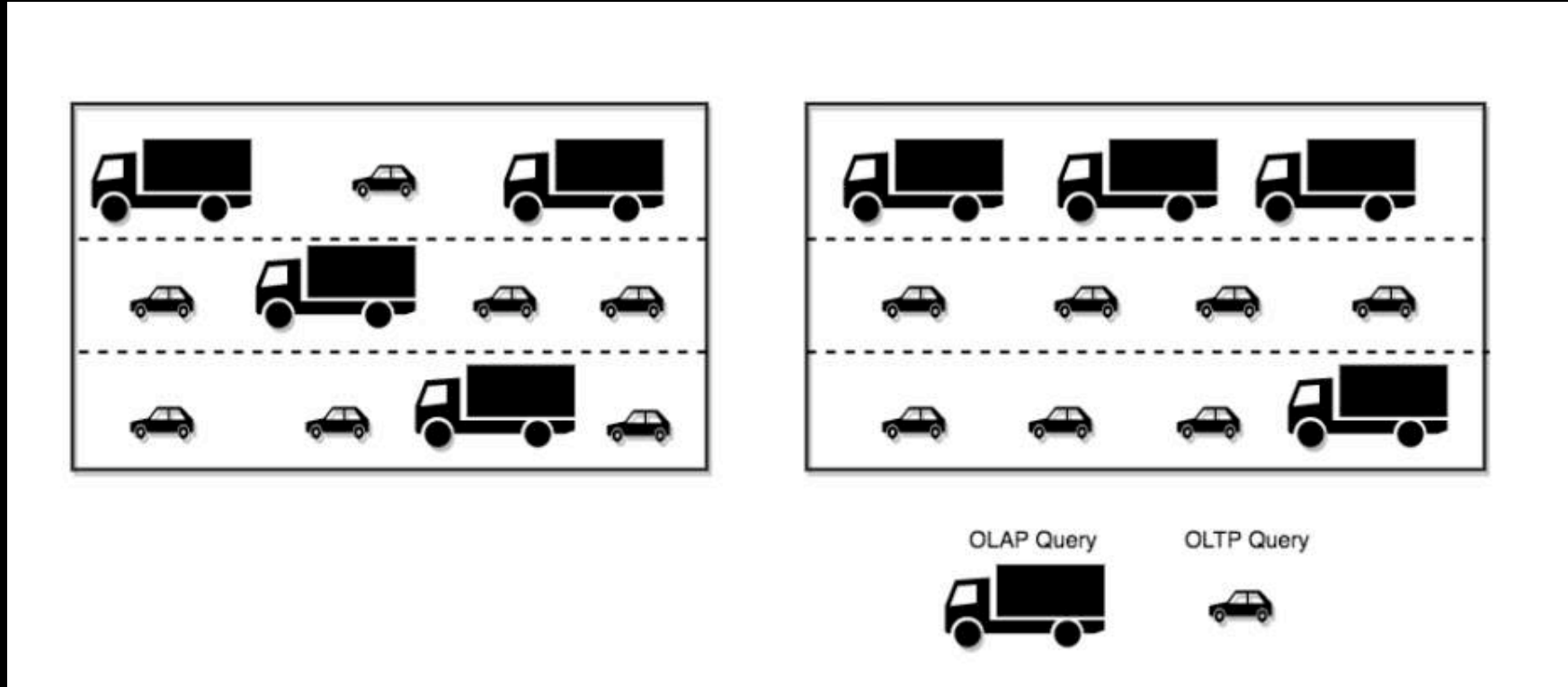


OLTP + OLAP





HTAP

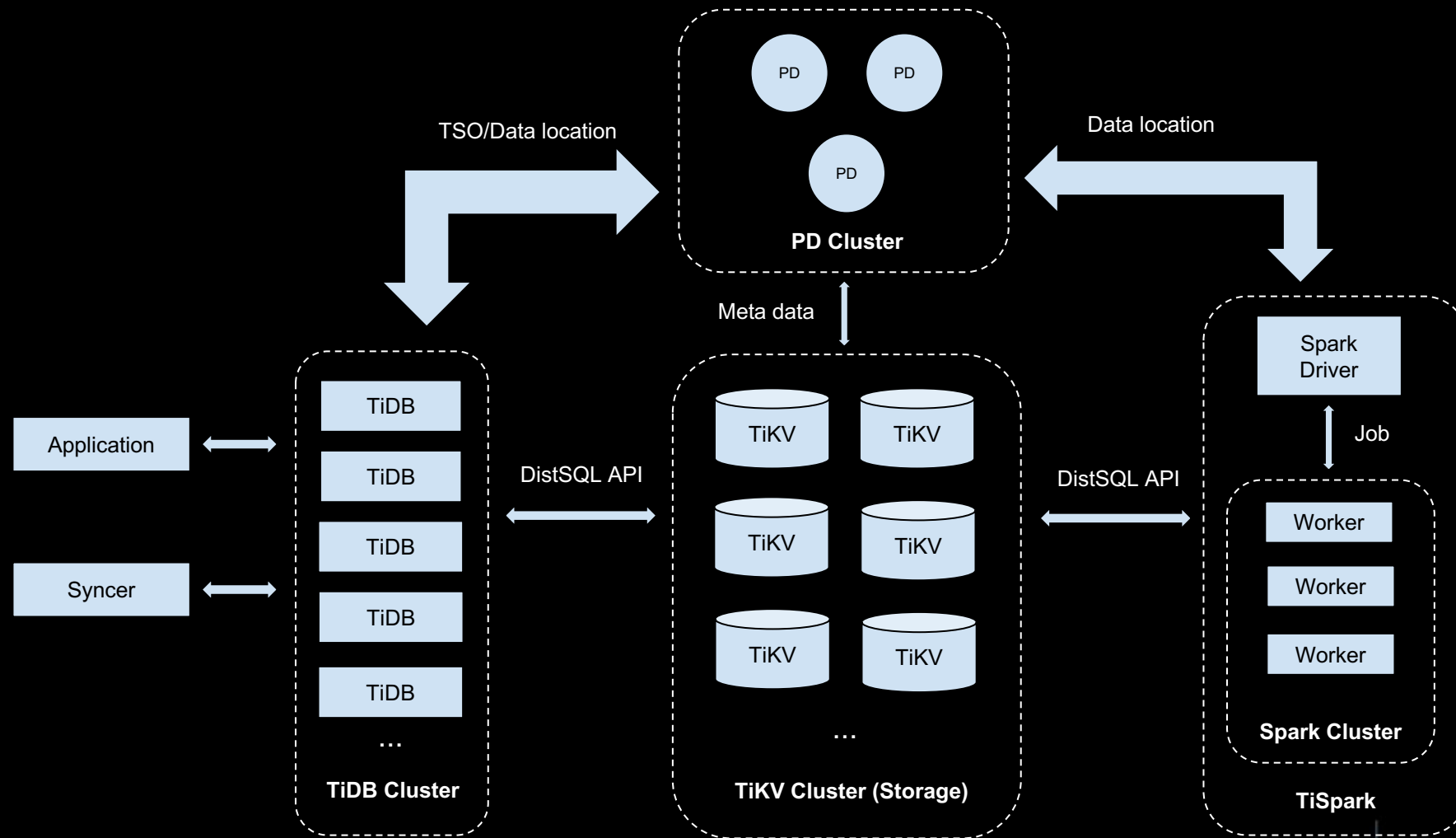




Ecosystem: Beyond TiDB and SQL



Spark on TiDB





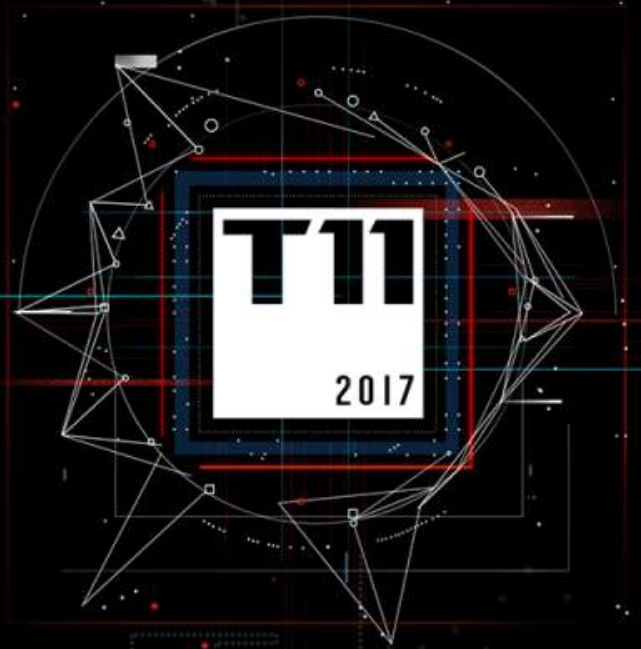
Spark on TiDB

- Spark ecosystem
- TiKV Connector is better than JDBC connector
- Index support
- Complex Calculation Pushdown
- CBO
 - Pick up right Access Path
 - Join Reorder
- Priority & Isolation Level



Future plan

- Code Generation
- MPP Engine
- Mixed storage engine (Columnar / Row-based)
- Heterogeneous computing (CPU/GPU/FPGA)



THANKS