MySQL + RocksDB
For Better Storage Efficiency Than InnoDB

XLDB 2015 Lightning Talk
Siying Dong, Software Engineer
Database Engineering Team @ Facebook
Facebook Website Architecture

- SSD Capacity is the bottleneck
- Also want to write fewer bytes to SSDs
MySQL + RocksDB (MyRocks)

- RocksDB
  - Embedded persistent key-value store
  - Log-Structure-Merge tree
  - Open-Sourced
- RocksDB Storage Engine in MySQL
  https://github.com/MySQLOnRocksDB/mysql-5.6
InnoDB vs. RocksDB Storage Engines

- **Workload:** Social Graph Data (Edges + Vertexes)
- **RocksDB:** zlib compression from L3
- **InnoDB:** edges use zlib, vertexes are pre-compressed.
Lower Space Amplification

InnoDB

Uncompressed 16KB

Row
Row
Row
Wasted
Row

Compressed to 5KB

Row
Row
Row
Wasted
Row

Using 8KB space on storage

Row
Row
Wasted
Row

4KB
4KB

RocksDB

Level 0

Level 1

Target 1GB

Level 2

Target 10 GB

Level 3

Target 100 GB

Level 4

Target 1000 GB

stale

Target 1GB

stale

stale

stale

stale

stale
Lower Write Amplification (Worst Case)

InnoDB

RocksDB

Write Amp = Page size / row size

Read

Modify

Write

Write Amp 1

Write Amp 10

Write Amp 10

Write Amp 10

Write Amp 10

Target 1GB

Target 10 GB

Target 100 GB

Target 1000 GB

flush

Level 0

Level 1

Level 2

Level 3

Level 4

Merge

Merge

Merge

Merge

Flush
Conclusion

- Current bottlenecks of our MySQL databases:
  - disk capacity
  - bytes written

- MySQL + RocksDB vs. InnoDB
  - compresses better
  - writes less data to storage devices

- We anticipate RocksDB will be able to replace many uses of InnoDB after some more work