RocksDB Festival

Supported by IITP, StarLab.

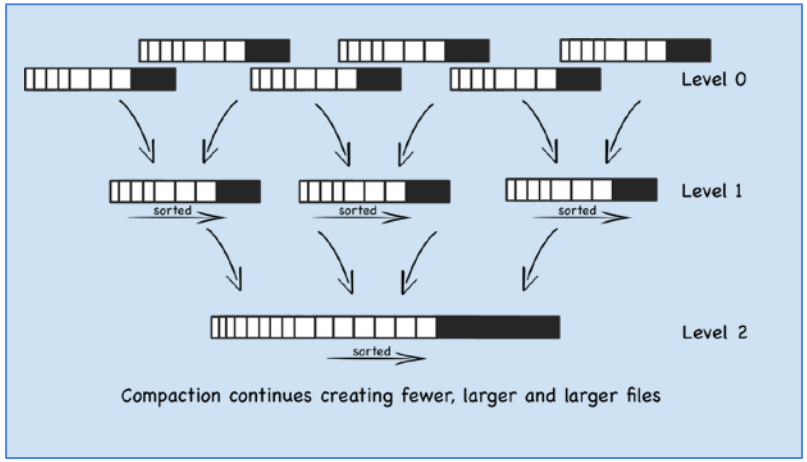
July 26, 2021

강정현, 김산, 허진
rilac1@naver.com, waterfog9580@gmail.com, jinh2352@gmail.com
Team: JSJ

RocksDB Festival

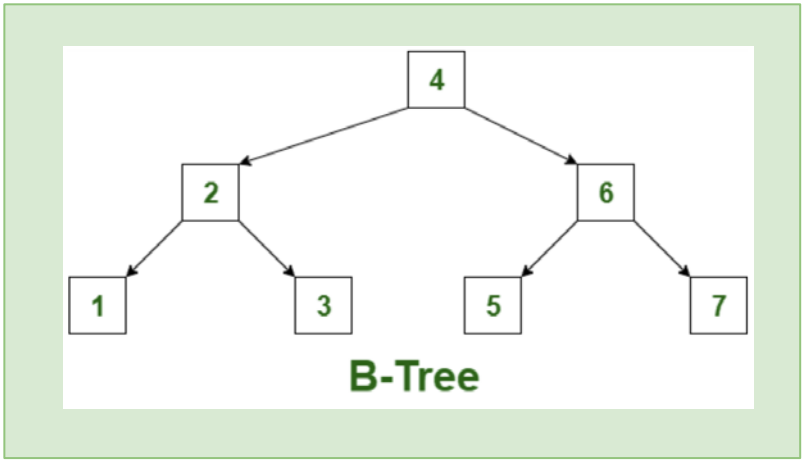
- RocksDB vs MongoDB
  - ✔ Data Structure

Rocks DB



LSM-Tree
(out-place update)
-> Write Good!

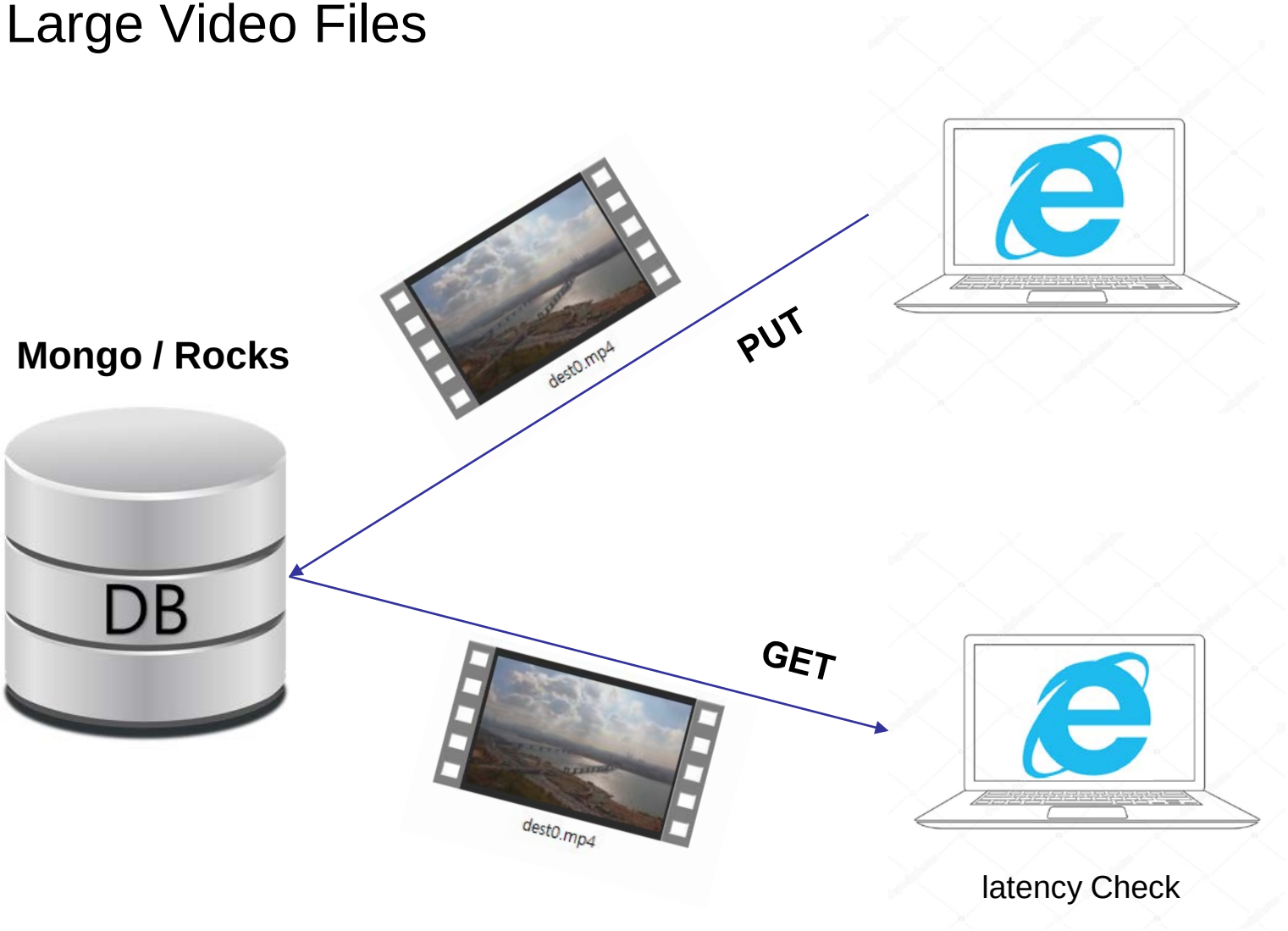
Mongo DB



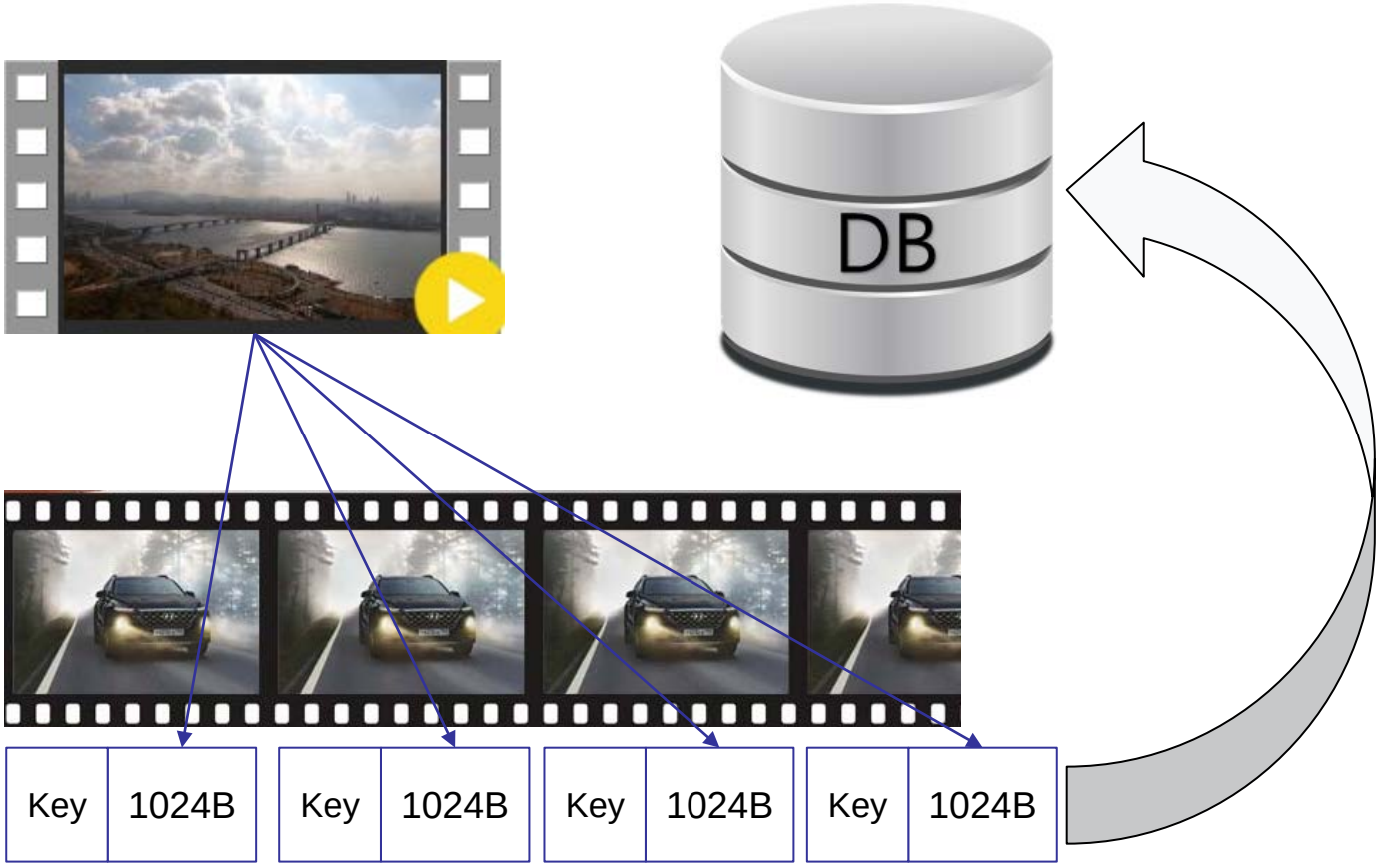
B-Tree
(in-place update)
-> Read Good!

RocksDB Festival

- Put Large Video Files



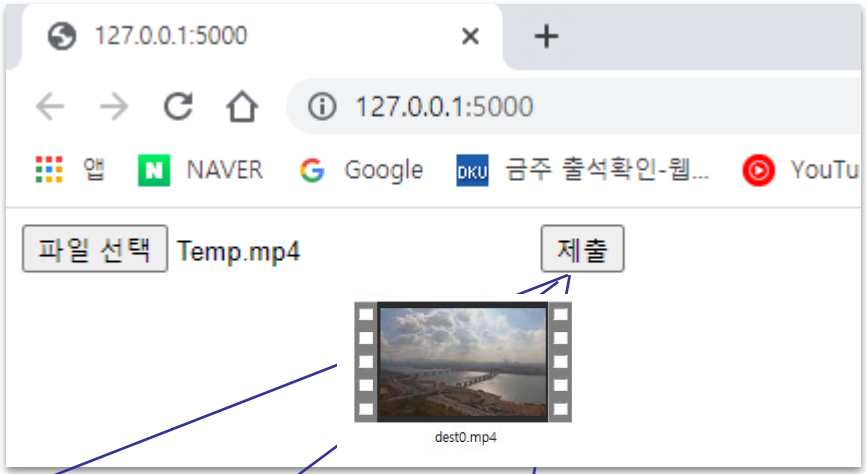
- Put Large Video Files



- MongoDB
  - ✔ GridFS

```python
@app.route('/upload', methods=['POST'])
def upload():
    file = request.files.get("file")
    file_name = request.form.get("file_name")
    data = file.read()
    content_type = file.content_type
    insertimg = gfs.put(data, content_type=content_type, filename=file_name)
    return str(insertimg)

```



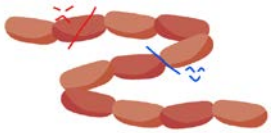
- RocksDB for Video Files Store

Issue

✔ RocksDB API parameter issue, only supports std::string types.

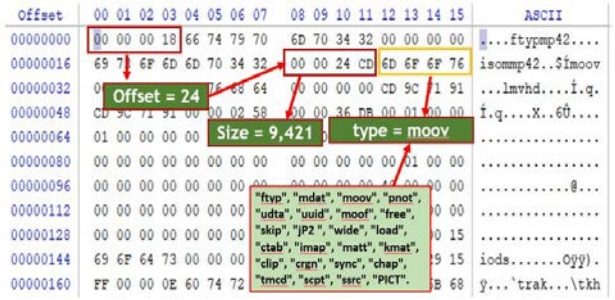
-> Simply parsing by interval causes problems with metadata access

✔ How to cut it in frames? (O) (X)

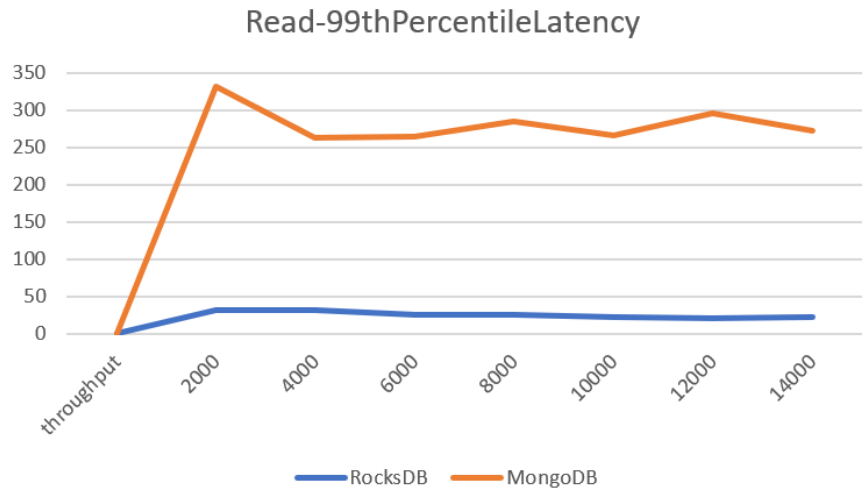


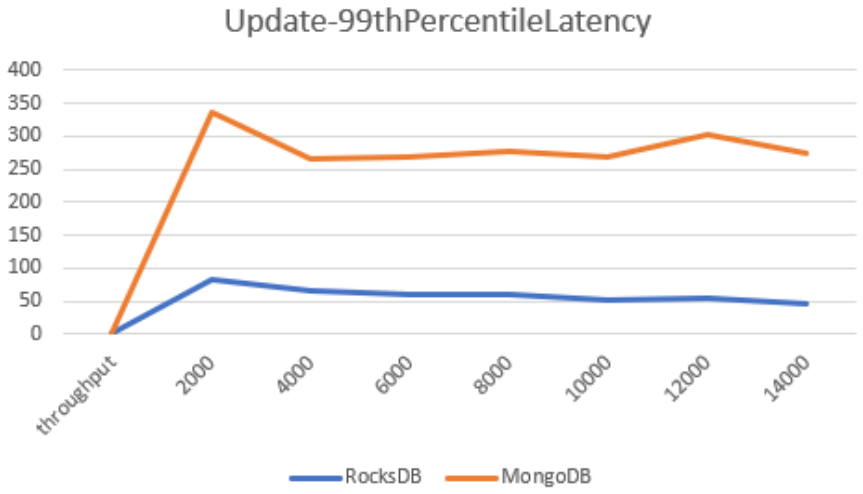
✔ What about metadata access?

-> Simply parsing by interval causes problems with metadata access



- RocksDB vs MongoDB
  - ✔ YCSB Bench

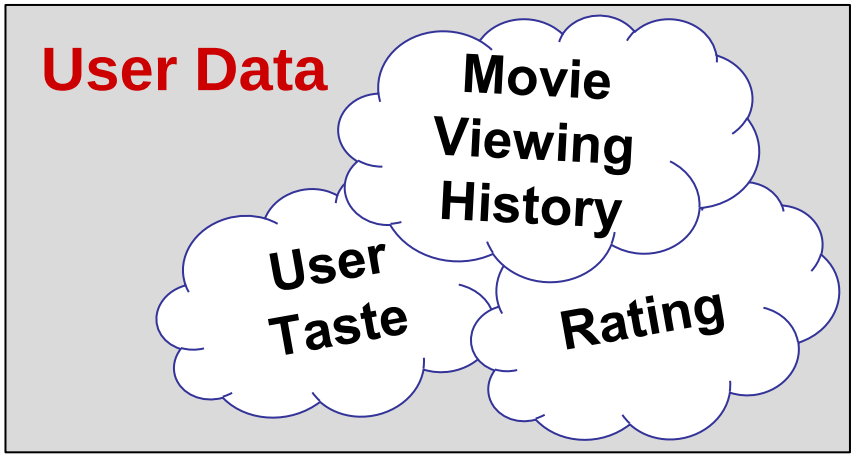




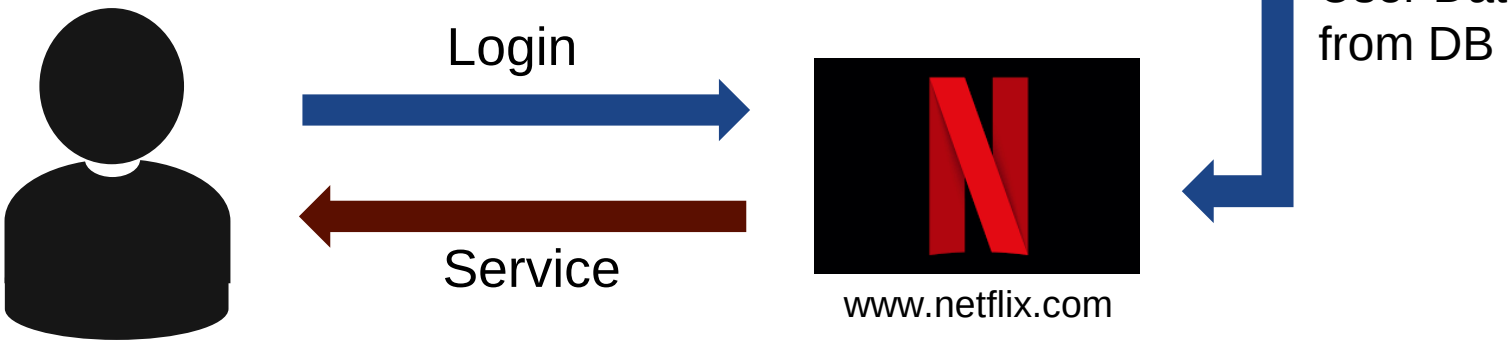
Latency	: 99thPercentileLatency(us)
Throughput	: Operation/sec (YCSB의 target 옵션을 통해 제한하면서 실험 진행)
Record	: 100,000
Workload	: Read(50%), Write(50%)

- Who Uses RocksDB
  - Distributed Database
    - Embed RocksDB as Storage Engine
      - Apache Cassandra
      - CockroachDB
      - MySQL (MyRocks)
      - Rockset
  - Outside of Distributed Database
      - **Kafka Streams**
      - Apache Samza
      - **Netflix**
      - Stander UK
      - **Uber**

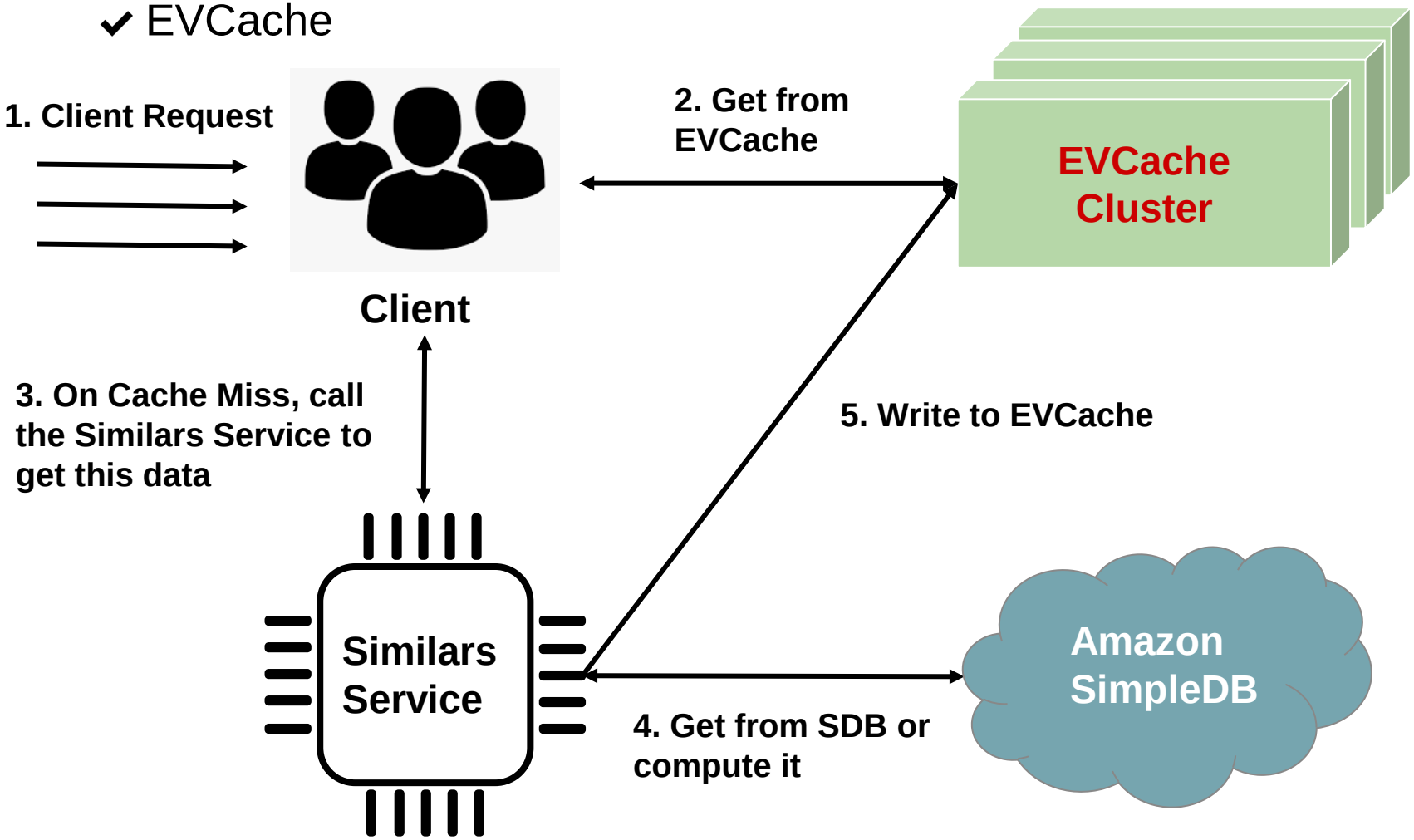
RocksDB Festival



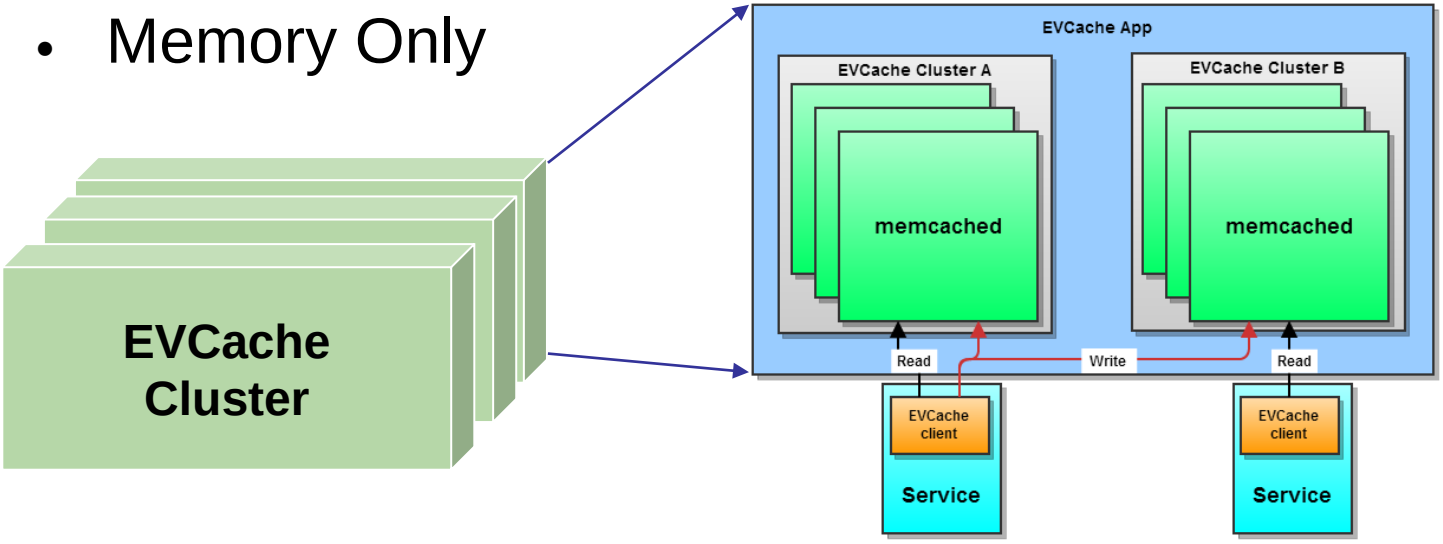
- Provide **personalized recommendations** services



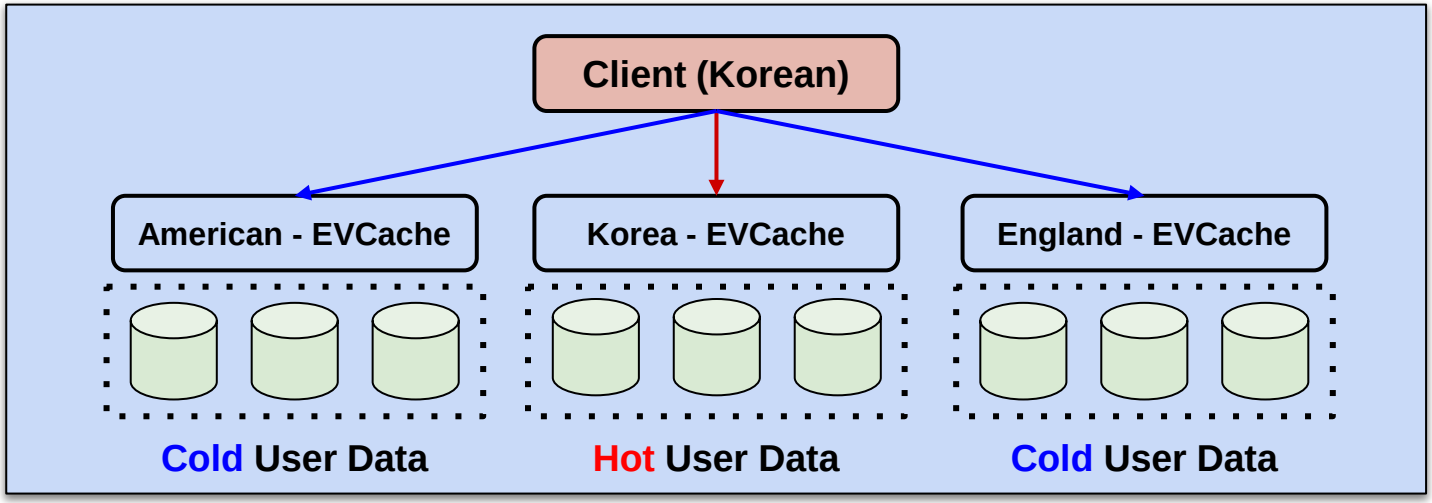
- Netflix
  - ✔ EVCache



- Memory Only

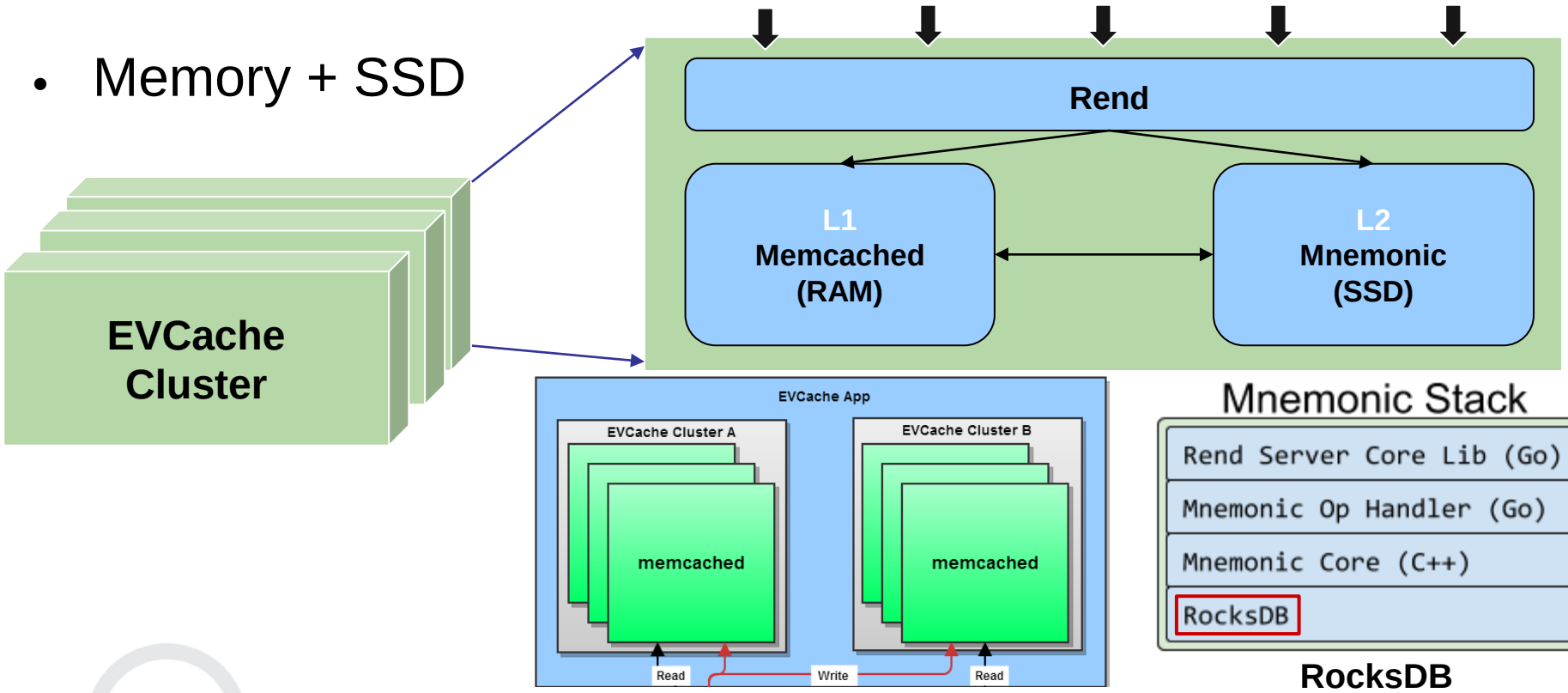


Ketama consistent hashing algorithm



RocksDB Festival

- Memory + SSD



Ketama consistent hashing algorithm

RocksDB

amazon web services™

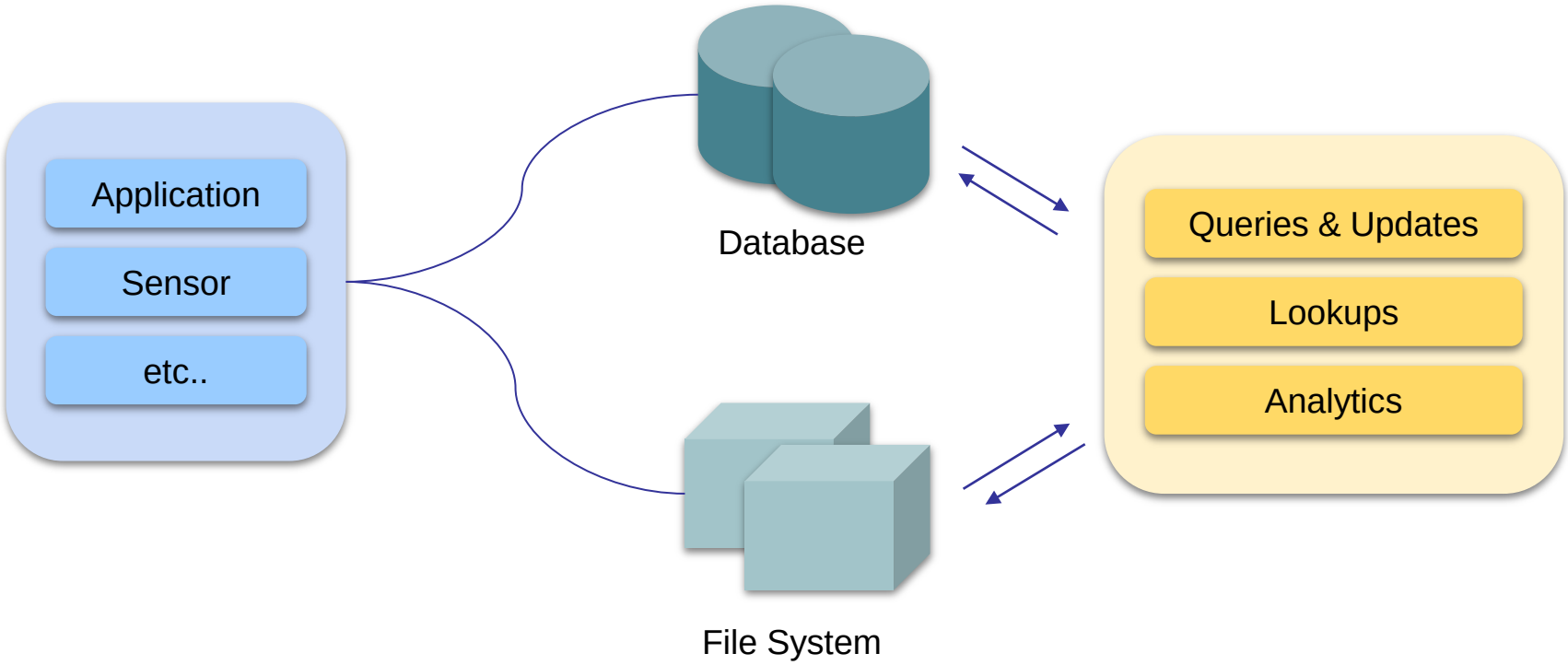
* Linux 기준

* 1 GiB = 1.07 GB

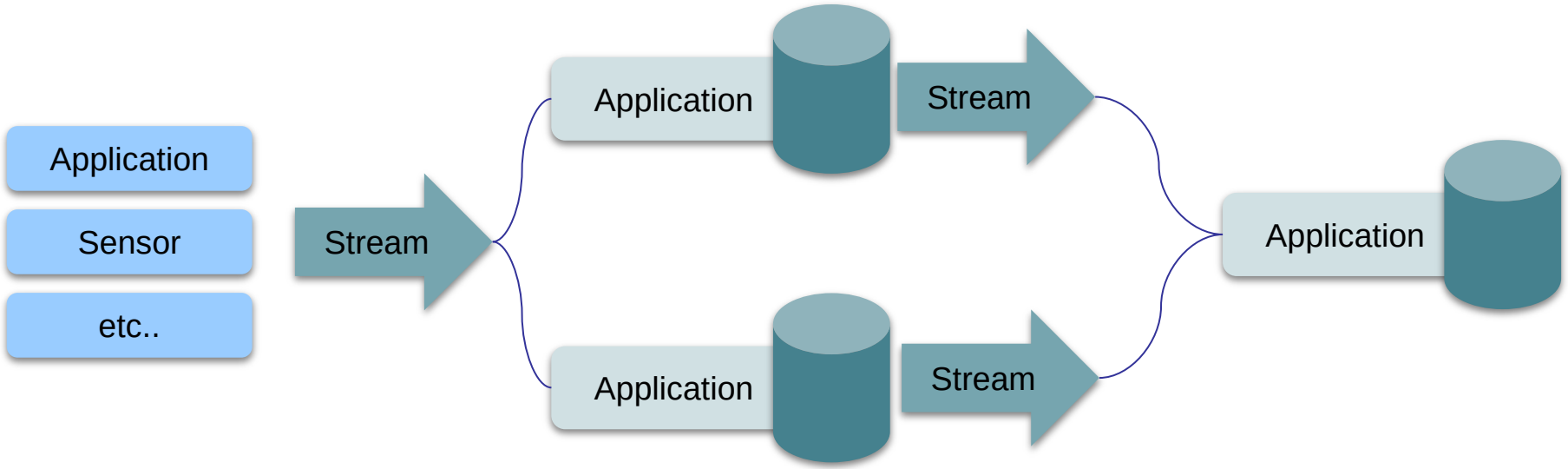
인스턴스 패밀리	인스턴스 유형	vCPU	ECU	메모리	인스턴스 스토리지 (SSD)	시간 당 가격
메모리 최적화	**r3**.xlarge	4	13	30.5 GiB	80 GB	**0.332 USD**
스토리지 최적화	**i2**.xlarge	4	14	**30.5 GiB**	**800 GB** x 100	**0.853 USD**

x 2.5

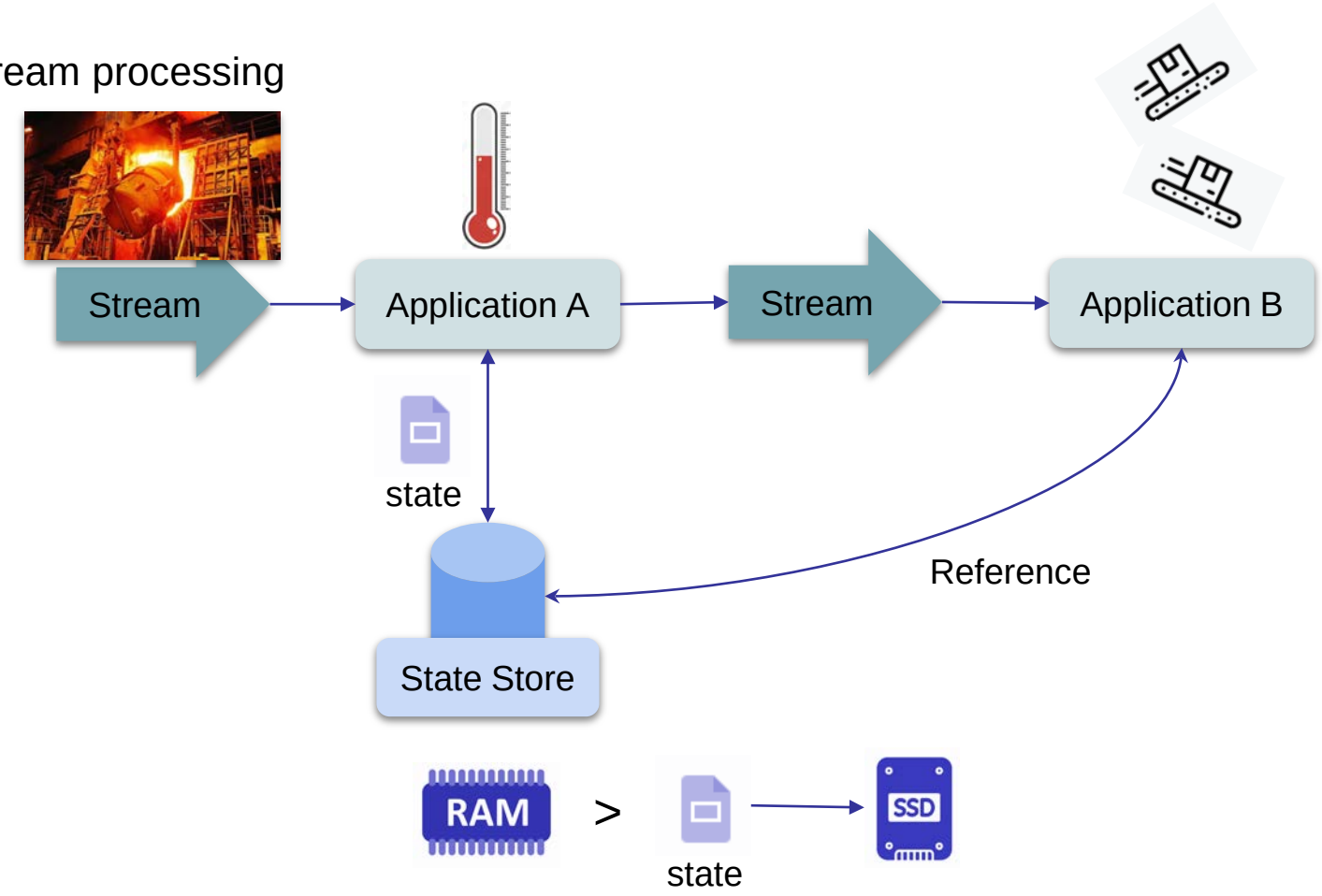
- Kafka Streams
  - Before Stream processing



- Kafka Streams
  - Stream processing

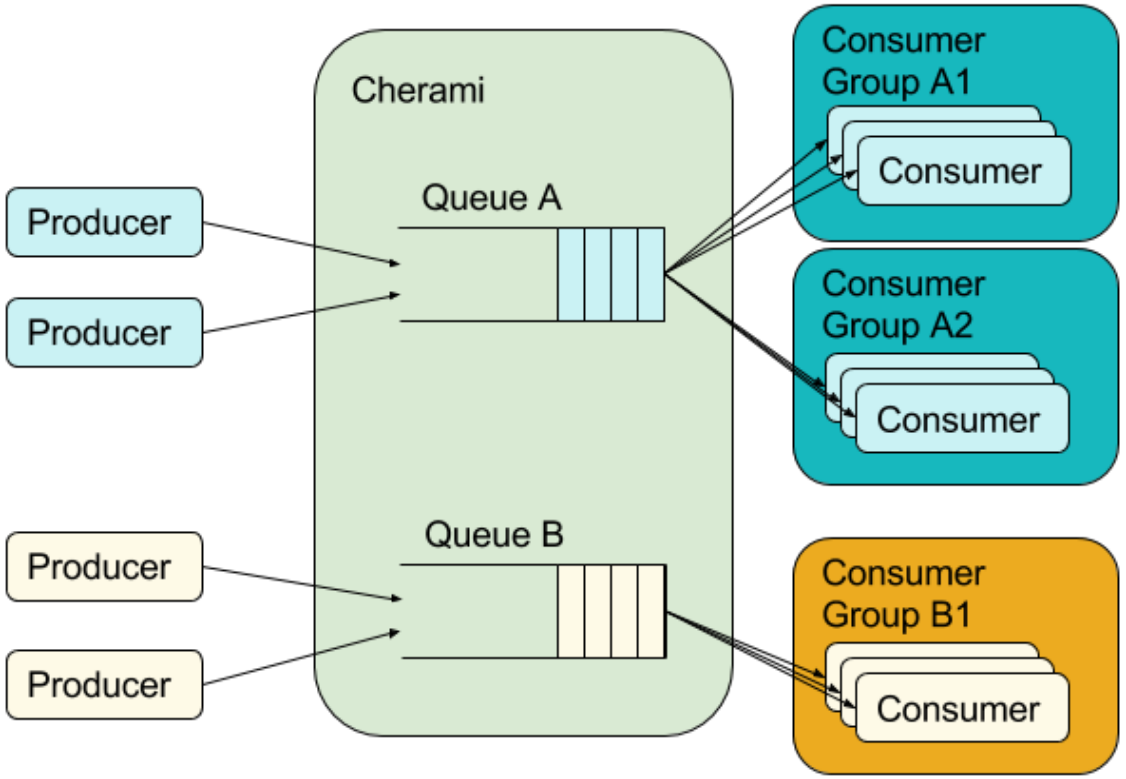


- Kafka Streams
  - Stream processing

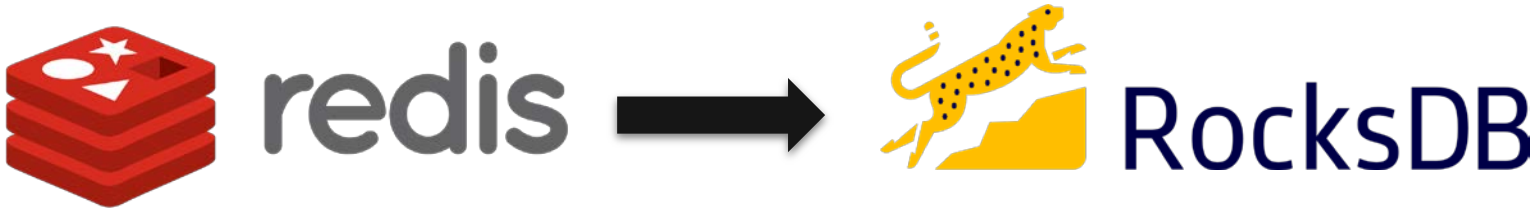


state larger than available main memory can be supported

- Uber : Cherami (Message Queue System)
  - Message Queue System

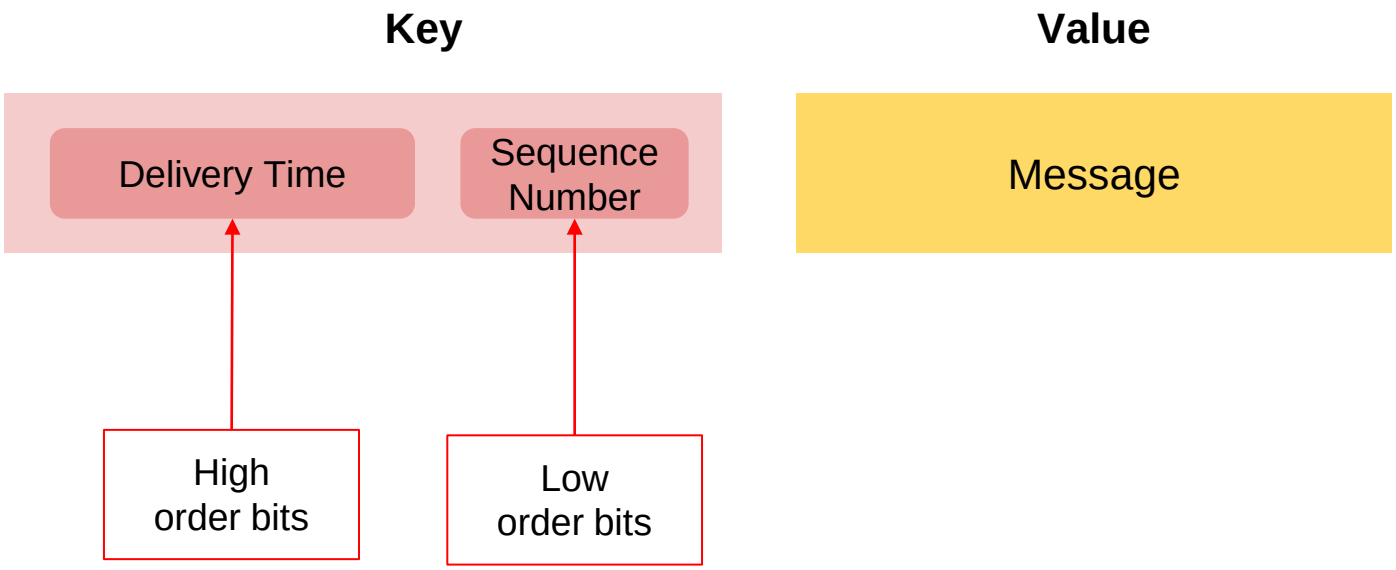


- Uber : Cherami (Message Queue System)
  - Redis to RocksDB



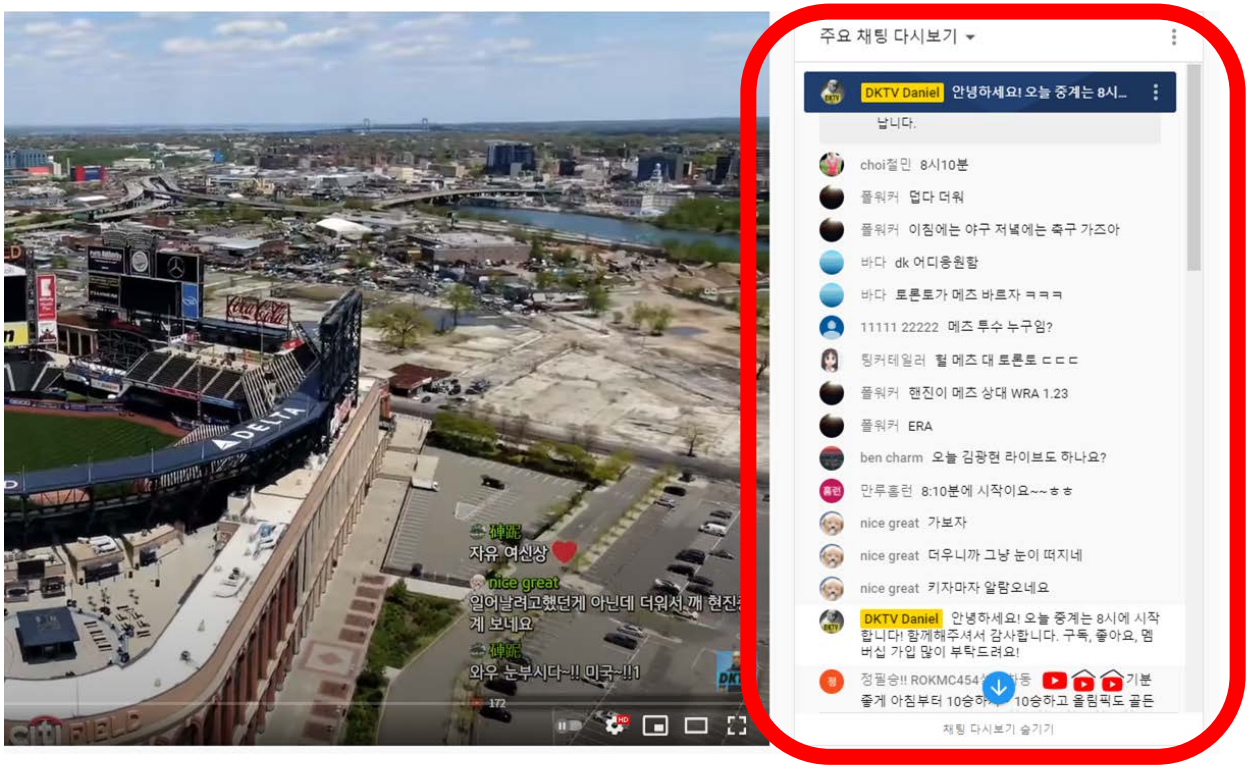
In-memory Key-Value Store
Isn't as **durable** or **scalable** as Uber needed.

- Uber : Cherami (Message Queue System)
  - Implement Message Queue by RocksDB



RocksDB Festival

- Real workload(live chatting data) - interface implement



Discussion