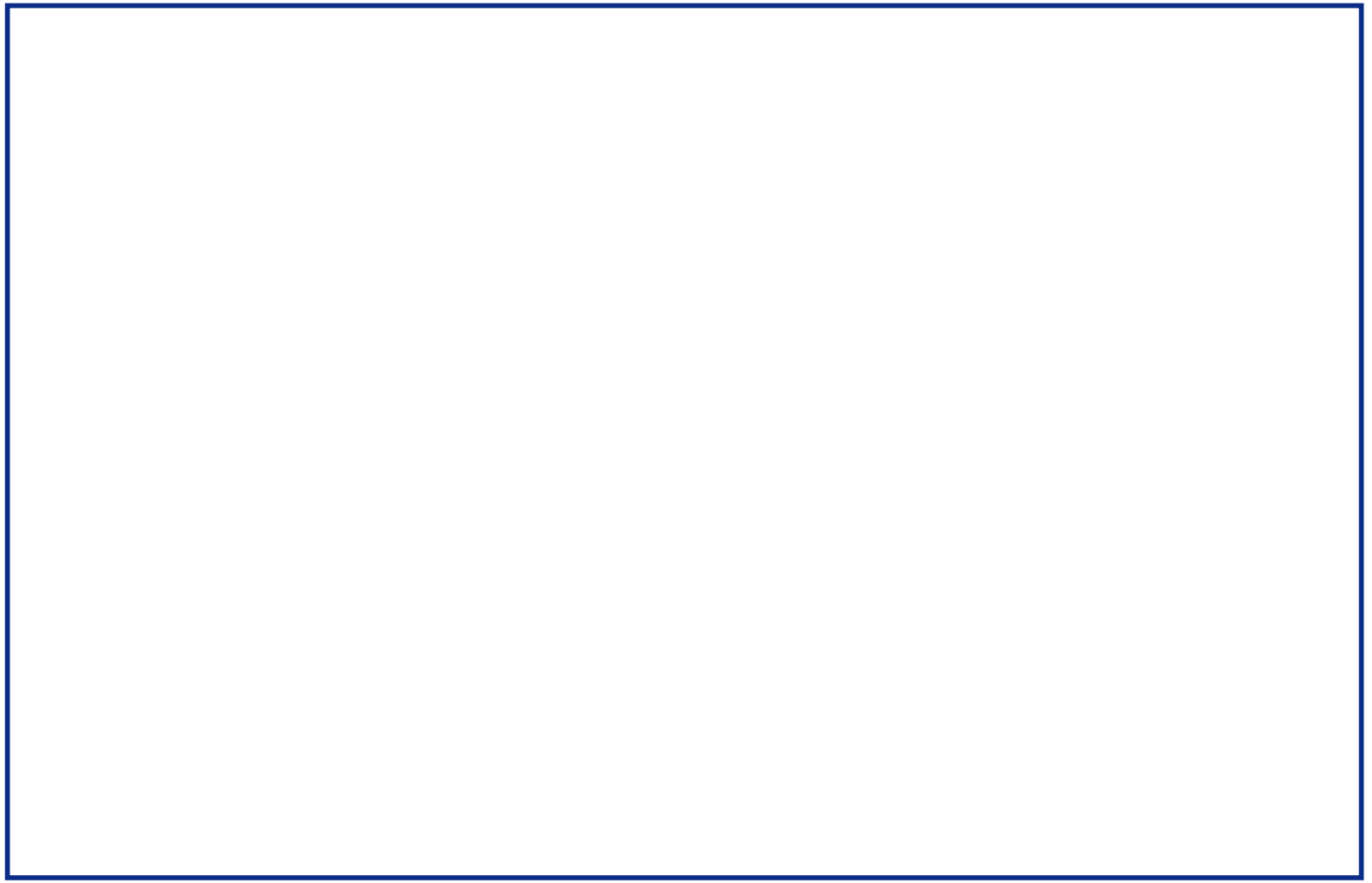
SSTable Read - overview

Sanghyun Cho, Jongki Park

E-Mail: 98shcho@naver.com

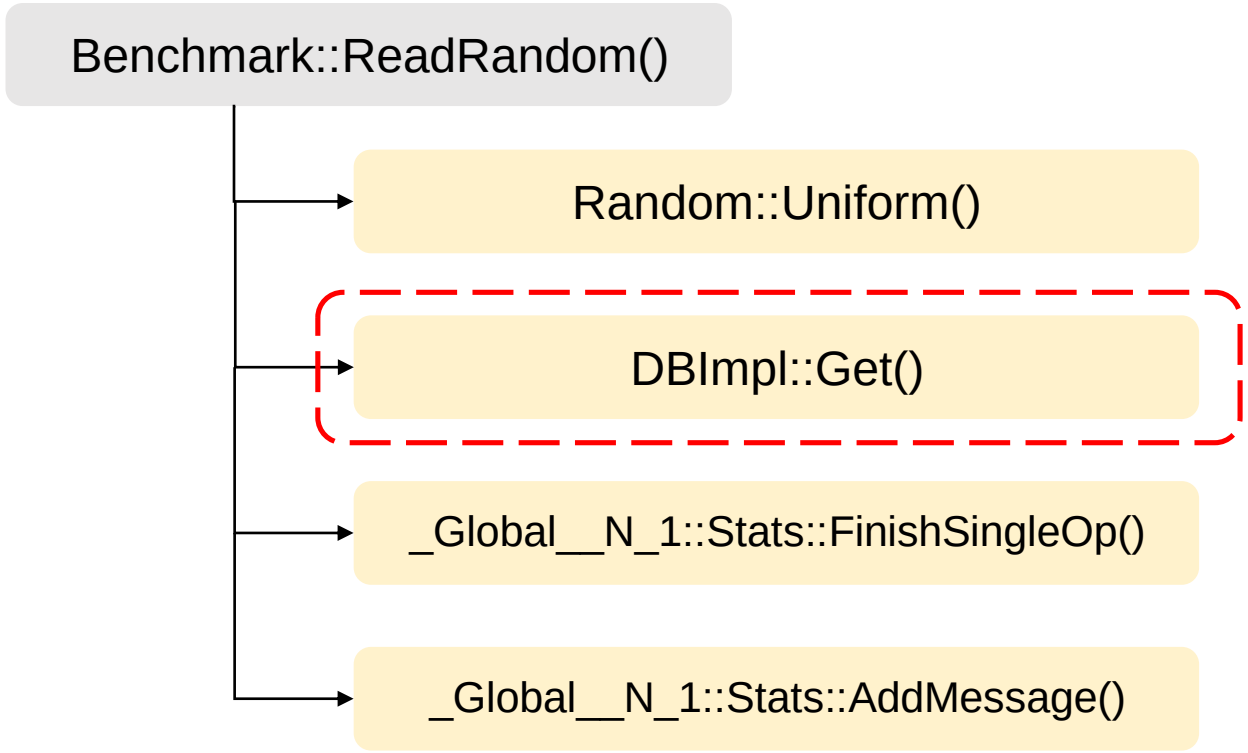
jkipark@dankook.ac.kr

1. Process of readrandom
2. Overall Get process
3. Finding candidates
4. Finding value from SSTable
5. References



1. Process of read random

- From Level db benchmark




```
readrandom   :      13.990 micros/op; (31440 of 50000 found)
readrandom   :      13.954 micros/op; (31440 of 50000 found)
readrandom   :      13.943 micros/op; (31440 of 50000 found)
```

Process of read random

- Level db doesn’t seem to generate a truly random number

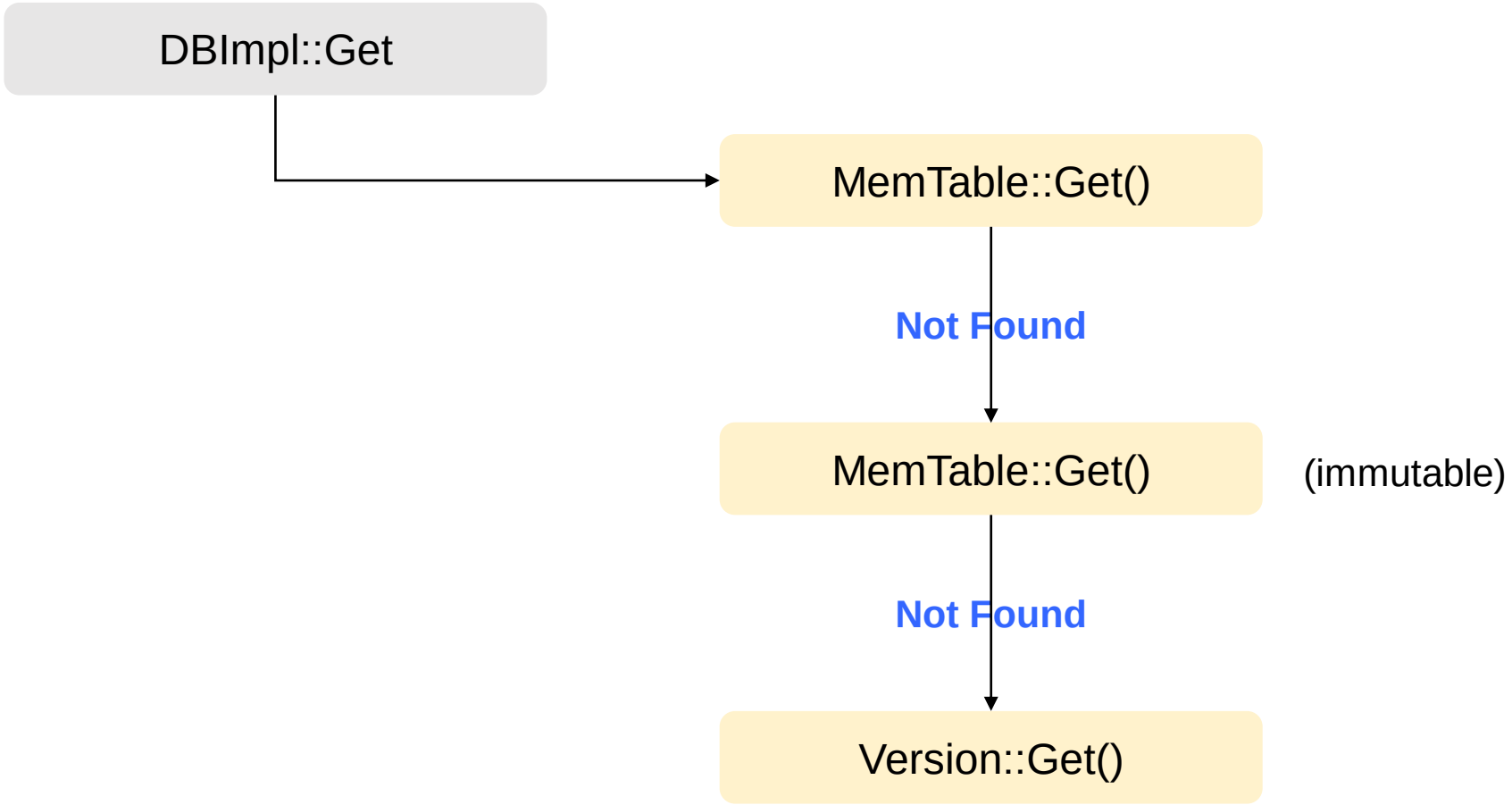
```
readrandom   :      23.352 micros/op; (50579 of 80000 found)
readrandom   :      23.936 micros/op; (50579 of 80000 found)

readrandom   :      14.543 micros/op; (25152 of 40000 found)
readrandom   :      14.554 micros/op; (25152 of 40000 found)

readrandom   :      26.051 micros/op; (43887 of 70000 found)
readrandom   :      25.187 micros/op; (43887 of 70000 found)
```

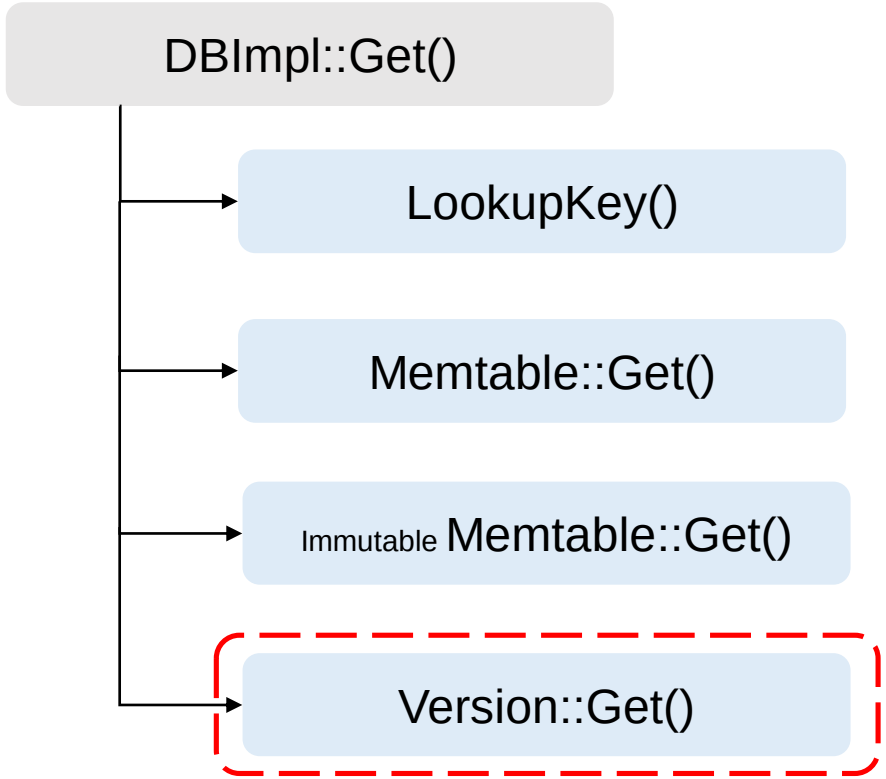
2. Overall Get process

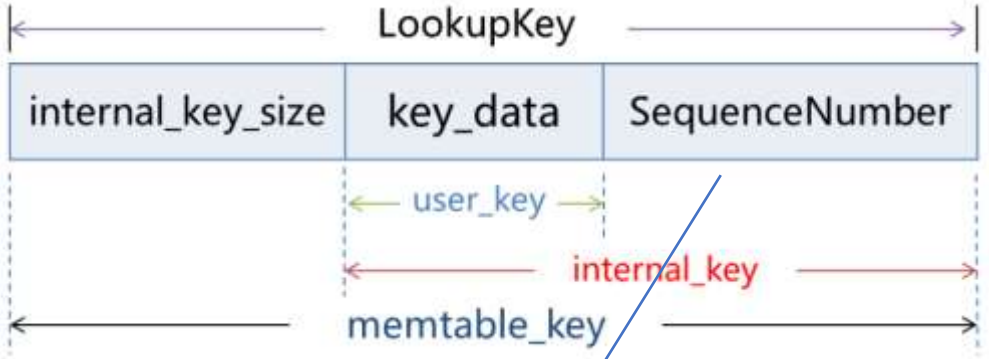
- Overall process of DBImpl::Get()



Overall Get process

- In Level db , DBImpl::Get() generates LookupKey from user key



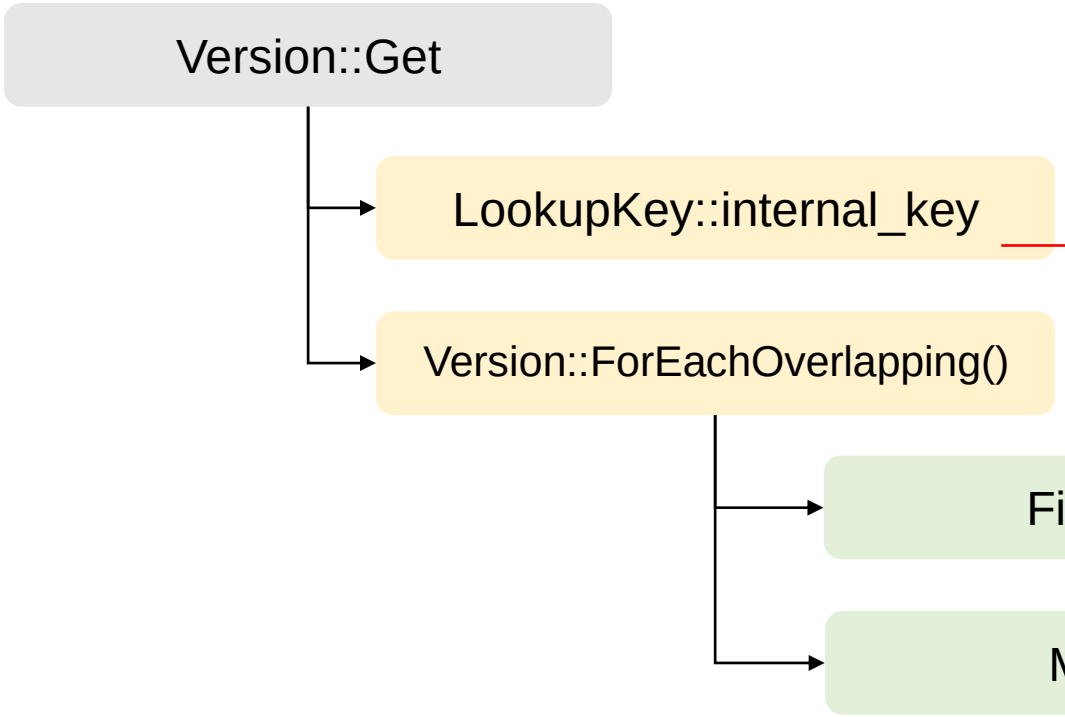


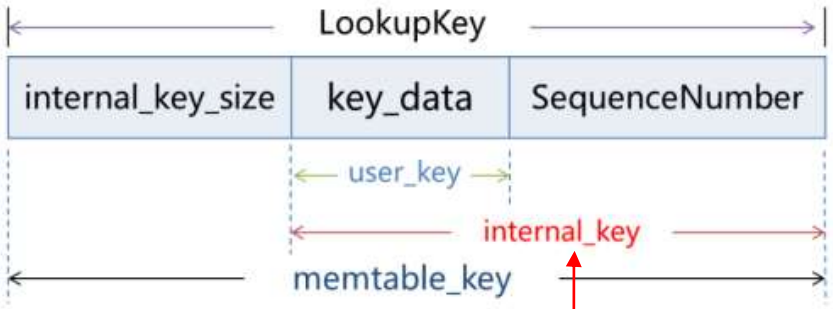
Source : https://bean-li.github.io/leveldb-memtable/

For compaction and snapshot
(every put and delete has sequence number)

3. Finding candidates (SSTable)

- Find candidate with metadata





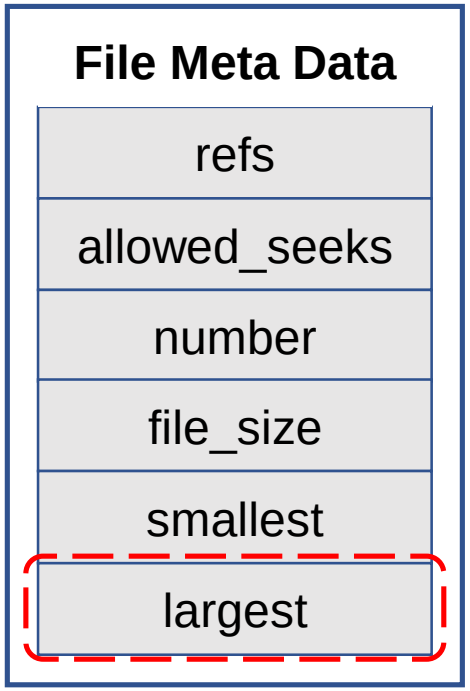
Finding candidates (SSTable)

- In ForEachOverlapping() …
  - Level 0

    Linear search in order from newest to oldest (The key ranges of each table overlap at level 0)

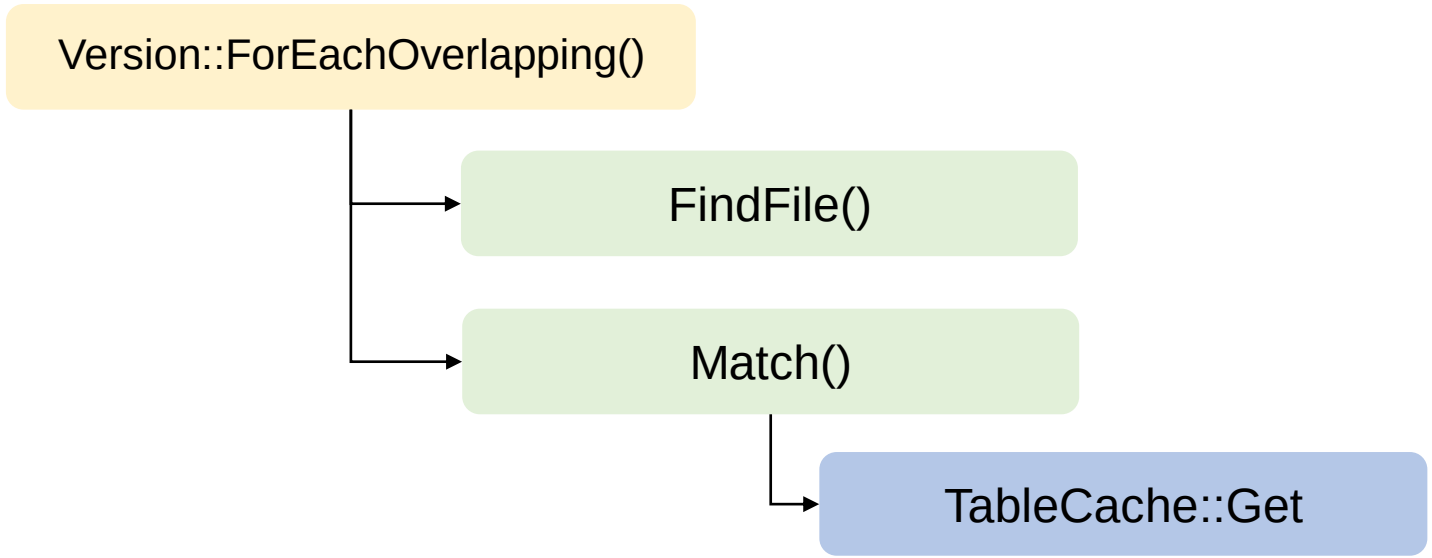
  - Level 1 ~ (each level)

    FindFile() **Binary searchs** using largest key in each sstable



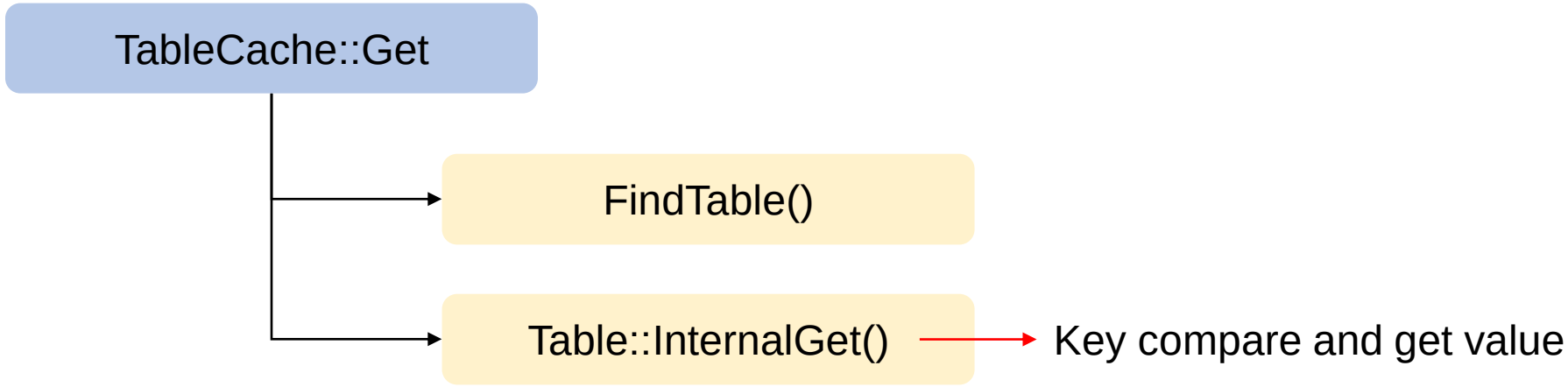
4. Finding value from SSTable

- Getting sstable from disk



Finding value from SSTable

- First look in cache before looking directly on disk



5. References

[1] https://bean-li.github.io/leveldb-memtable/

[2] https://sukill.tistory.com/

Thank you