Performance Analysis
RockDB 분석

컴퓨터소프트웨어학부
김현정 박정호

CONTENTS

Perf

시행착오

사용한 옵션

분석 결과

시행착오 – 실험 방식

1. 100초, 300초, 600초 실행 후 전체 실행에 대한 performance sampling

   너무 실험시간이 길어서 세 실험 모두 5개의 level을 다 사용하고 있었고, 따라서 각 실행의 결과가 거의 같았다. (이전의 latency 측정 실험 기준으로 대략 30초 이내로 모든 레벨을 사용하게 되었다.)

2. 25, 50, 75, 100초 실행 후 전체 실행에 대한 performance sampling

   전체 실행에 대해서 샘플링한 것으로 인해 유의미한 performance pattern 이 가려졌고, 이로 인해 각 함수의 CPU Cycle 비율의 차이가 거의 없었다.

3. 100초 실행 중, 초반 10초, 후반 10초의 performance sampling

   특정 구간의 performance pattern 을 볼 수 있으리라 기대했으나, 구간의 길이가 너무 길어서 이 역시 두 결과 간의 CPU Cycle 비율 차이가 거의 없었다.

사용한 옵션

1. 초반 측정
   sudo perf record –g sleep 2
   sudo perf record –g sleep 10

2. 후반 90~100초 측정
   sudo perf record –g –d 90000 sleep 10

3. 특정 symbol filter
   sudo perf report –symbol-filter=[symbol_name]

이후 실험 결과는 총 100초 실행 중 초반 2초, 후반 2초 간의 performance pattern 을 측정한 기록이다.

0초~2초 --no-children

9.19%	read_bottleneck	libc-2.27.so	[.]	__random_r
7.78%	read_bottleneck	libc-2.27.so	[.]	__random
4.77%	read_bottleneck	read_bottleneck_breakdown_test	[.]	main
4.21%	rocksdb:low0	[kernel.kallsyms]	[k]	copy_user_enhanced_fast_string
1.93%	rocksdb:low0	libc-2.27.so	[.]	__memmove_avx_unaligned_erms
1.81%	read_bottleneck	libc-2.27.so	[.]	rand
1.70%	rocksdb:low0	libc-2.27.so	[.]	__memset_avx2_erms
1.69%	read_bottleneck	[kernel.kallsyms]	[k]	copy_user_enhanced_fast_string
1.45%	read_bottleneck	read_bottleneck_breakdown_test	[.]	rand@plt
1.36%	rocksdb:low0	libsnappy.so.1.1.7	[.]	snappy::internal::CompressFragment
1.16%	rocksdb:low0	read_bottleneck_breakdown_test	[.]	rocksdb::crc32c::crc32c_3way
1.09%	read_bottleneck	libc-2.27.so	[.]	cfree@GLIBC_2.2.5
1.08%	rocksdb:low0	libc-2.27.so	[.]	cfree@GLIBC_2.2.5
0.85%	read_bottleneck	[kernel.kallsyms]	[k]	do_syscall_64
0.84%	read_bottleneck	libc-2.27.so	[.]	__memmove_avx_unaligned_erms
0.81%	read_bottleneck	[vdso]	[.]	__vdso_clock_gettime
0.76%	read_bottleneck	read_bottleneck_breakdown_test	[.]	rocksdb::IndexBlockIter::SeekImpl
0.72%	rocksdb:low0	[kernel.kallsyms]	[k]	do_syscall_64
0.72%	rocksdb:low0	read_bottleneck_breakdown_test	[.]	rocksdb::CompactionIterator::NextFromInput
0.66%	read_bottleneck	read_bottleneck_breakdown_test	[.]	rocksdb::MemTable::KeyComparator::operator()
0.61%	read_bottleneck	libc-2.27.so	[.]	malloc
0.61%	read_bottleneck	read_bottleneck_breakdown_test	[.]	rocksdb::crc32c::crc32c_3way
0.60%	rocksdb:low0	libc-2.27.so	[.]	malloc
0.53%	read_bottleneck	read_bottleneck_breakdown_test	[.]	rocksdb::(anonymous namespace)::BytewiseComparatorImpl::Compare
0.51%	read_bottleneck	libc-2.27.so	[.]	__memcmp_avx2_movbe
0.48%	rocksdb:low0	read_bottleneck_breakdown_test	[.]	rocksdb::MergingIterator::Next
0.47%	read_bottleneck	read_bottleneck_breakdown_test	[.]	rocksdb::MemTable::KeyComparator::operator()
0.43%	rocksdb:low0	[kernel.kallsyms]	[k]	__block_commit_write.isra.34
0.43%	rocksdb:low0	libc-2.27.so	[.]	_int_malloc
0.42%	rocksdb:low0	[kernel.kallsyms]	[k]	radix_tree_next_chunk
0.42%	read_bottleneck	read_bottleneck_breakdown_test	[.]	rocksdb::BlockIter<rocksdb::IndexValue>::CompareCurrentKey
0.40%	read_bottleneck	libc-2.27.so	[.]	_int_malloc
0.39%	read_bottleneck	[kernel.kallsyms]	[k]	__radix_tree_lookup
0.38%	rocksdb:low0	libstdc++.so.6.0.25	[.]	std::__cxx11::basic_string<char, std::char_traits<char>, std::allocator<char>
0.38%	read_bottleneck	[kernel.kallsyms]	[k]	find_get_entry
0.36%	rocksdb:low0	read_bottleneck_breakdown_test	[.]	rocksdb::BlockBasedTableIterator::status
0.35%	read_bottleneck	[kernel.kallsyms]	[k]	__entry_trampoline_start
0.34%	rocksdb:low0	read_bottleneck_breakdown_test	[.]	rocksdb::LRUHandleTable::FindPointer

0초~2초_1 --children

Children	Self	Command	Shared Object	Symbol
26.13%	4.77%	read_bottleneck	read_bottleneck_breakdown_test	[.] main
18.31%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BackgroundCallCompaction
17.72%	0.00%	rocksdb:low0	libstdc++.so.6.0.25	[.] 0x00000000000bd6df
17.72%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::ThreadPoolImpl::Impl::BGThreadWrapper
17.72%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BGWorkCompaction
17.72%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::ThreadPoolImpl::Impl::BGThread
16.81%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BackgroundCompaction
16.78%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionJob::Run
16.60%	0.29%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionJob::ProcessKeyValueCompaction
15.07%	0.07%	rocksdb:low0	[kernel.kallsyms]	[k] entry_SYSCALL_64_after_hwframe
14.99%	0.72%	rocksdb:low0	[kernel.kallsyms]	[k] do_syscall_64
11.93%	0.00%	read_bottleneck	[unknown]	[.] 0xb46ab06000000000
11.82%	0.02%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.80%	0.03%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.77%	0.02%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.74%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
11.72%	0.04%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
11.55%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::Get
11.55%	0.00%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::Get
11.47%	0.09%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::GetImpl
9.83%	0.22%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::Version::Get
9.54%	0.00%	read_bottleneck	[unknown]	[.] 0x5541c6894d5641cf
9.46%	0.07%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::PessimisticTransactionDB::Put
9.21%	9.19%	read_bottleneck	libc-2.27.so	[.] __random_r
9.13%	0.09%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionIterator::Next
9.00%	0.20%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TableCache::Get
8.64%	0.33%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::Get
8.32%	0.10%	read_bottleneck	[kernel.kallsyms]	[k] entry_SYSCALL_64_after_hwframe
8.20%	0.85%	read_bottleneck	[kernel.kallsyms]	[k] do_syscall_64
7.89%	7.78%	read_bottleneck	libc-2.27.so	[.] __random
7.77%	0.00%	rocksdb:low0	libpthread-2.27.so	[.] __libc_write
7.77%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] sys_write
7.76%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] vfs_write
7.76%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] __vfs_write
7.76%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] new_sync_write
7.76%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] ext4_file_write_iter
7.76%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] __generic_file_write_iter
7.68%	0.05%	rocksdb:low0	[kernel.kallsyms]	[k] generic_perform_write

0초~2초_2 --children

```
7.68%     0.05%  rocksdb:low0     [kernel.kallsyms]              [k] generic_perform_write
7.42%     0.48%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::MergingIterator::Next
7.32%     0.07%  read_bottleneck  read_bottleneck_breakdown_test [.] rocksdb::PessimisticTransaction::Commit
6.24%     0.20%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTableIterator::NextAndGetResult
6.18%     0.04%  read_bottleneck  read_bottleneck_breakdown_test [.] rocksdb::WriteCommittedTxn::CommitWithoutPrepareInternal
5.98%     0.10%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTableIterator::Next
5.97%     0.26%  read_bottleneck  read_bottleneck_breakdown_test [.] rocksdb::DBImpl::WriteImpl
5.95%     0.16%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::(anonymous namespace)::LevelIterator::NextAndGetResult
5.55%     0.13%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTableIterator::FindBlockForward
4.99%     0.08%  rocksdb:low0     libpthread-2.27.so             [.] __libc_pread64
4.98%     0.13%  read_bottleneck  read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTable::NewDataBlockIterator<rocksdb::DataBlockIter>
4.72%     0.04%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTableIterator::InitDataBlock
4.69%     0.04%  read_bottleneck  read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTable::RetrieveBlock<rocksdb::Block>
4.63%     0.25%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTableBuilder::Add
4.62%     0.15%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTable::NewDataBlockIterator<rocksdb::DataBlockIter>
4.60%     0.16%  read_bottleneck  read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTable::MaybeReadBlockAndLoadToCache<rocksdb::Block>
4.53%     0.06%  read_bottleneck  libpthread-2.27.so             [.] __libc_pread64
4.33%     0.05%  read_bottleneck  libpthread-2.27.so             [.] __libc_write
4.21%     4.21%  rocksdb:low0     [kernel.kallsyms]              [k] copy_user_enhanced_fast_string
3.66%     0.10%  rocksdb:low0     read_bottleneck_breakdown_test [.] rocksdb::BlockBasedTable::RetrieveBlock<rocksdb::Block>
3.63%     0.05%  read_bottleneck  [kernel.kallsyms]              [k] sys_write
3.57%     0.04%  rocksdb:low0     [kernel.kallsyms]              [k] sys_pread64
3.48%     0.04%  read_bottleneck  [kernel.kallsyms]              [k] sys_pread64
3.46%     0.05%  read_bottleneck  [kernel.kallsyms]              [k] vfs_write
3.39%     0.04%  rocksdb:low0     [kernel.kallsyms]              [k] vfs_read
3.30%     0.04%  read_bottleneck  [kernel.kallsyms]              [k] vfs_read
3.26%     0.00%  read_bottleneck  [kernel.kallsyms]              [k] __vfs_write
3.25%     0.03%  read_bottleneck  [kernel.kallsyms]              [k] new_sync_write
3.21%     0.04%  read_bottleneck  [kernel.kallsyms]              [k] ext4_file_write_iter
3.15%     0.00%  rocksdb:low0     [kernel.kallsyms]              [k] __vfs_read
3.14%     0.03%  rocksdb:low0     [kernel.kallsyms]              [k] new_sync_read
3.13%     0.02%  read_bottleneck  [kernel.kallsyms]              [k] __generic_file_write_iter
```

90초~100초 --no-children

Overhead	Command	Shared Object	Symbol
8.88%	read_bottleneck	libc-2.27.so	[.] __random_r
7.66%	read_bottleneck	libc-2.27.so	[.] __random
4.80%	read_bottleneck	read_bottleneck_breakdown_test	[.] main
4.40%	rocksdb:low0	[kernel.kallsyms]	[k] copy_user_enhanced_fast_string
1.92%	rocksdb:low0	libc-2.27.so	[.] __memmove_avx_unaligned_erms
1.82%	read_bottleneck	libc-2.27.so	[.] rand
1.68%	rocksdb:low0	libc-2.27.so	[.] __memset_avx2_erms
1.60%	read_bottleneck	[kernel.kallsyms]	[k] copy_user_enhanced_fast_string
1.47%	read_bottleneck	read_bottleneck_breakdown_test	[.] rand@plt
1.31%	rocksdb:low0	libsnappy.so.1.1.7	[.] snappy::internal::CompressFragment
1.17%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::crc32c::crc32c_3way
1.04%	rocksdb:low0	libc-2.27.so	[.] cfree@GLIBC_2.2.5
0.97%	read_bottleneck	libc-2.27.so	[.] cfree@GLIBC_2.2.5
0.80%	read_bottleneck	[kernel.kallsyms]	[k] do_syscall_64
0.78%	read_bottleneck	[vdso]	[.] __vdso_clock_gettime
0.75%	read_bottleneck	libc-2.27.so	[.] __memmove_avx_unaligned_erms
0.75%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::IndexBlockIter::SeekImpl
0.74%	rocksdb:low0	[kernel.kallsyms]	[k] do_syscall_64
0.68%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionIterator::NextFromInput
0.68%	rocksdb:low0	libc-2.27.so	[.] malloc
0.64%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTable::KeyComparator::operator()
0.58%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::crc32c::crc32c_3way
0.55%	read_bottleneck	libc-2.27.so	[.] malloc
0.55%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::MergingIterator::Next
0.54%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::(anonymous namespace)::BytewiseComparatorImpl::Compare
0.50%	read_bottleneck	libc-2.27.so	[.] __memcmp_avx2_movbe
0.49%	rocksdb:low0	[kernel.kallsyms]	[k] radix_tree_next_chunk
0.48%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTable::KeyComparator::operator()
0.46%	rocksdb:low0	libstdc++.so.6.0.25	[.] std::__cxx11::basic_string<char, std::char_traits<char>, std::allocator<char> >:
0.46%	rocksdb:low0	[kernel.kallsyms]	[k] __block_commit_write.isra.34
0.46%	rocksdb:low0	libc-2.27.so	[.] _int_malloc
0.43%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTableIterator::status
0.41%	read_bottleneck	libc-2.27.so	[.] _int_malloc
0.40%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBuilder::Add
0.40%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockIter<rocksdb::IndexValue>::CompareCurrentKey
0.38%	read_bottleneck	[kernel.kallsyms]	[k] __radix_tree_lookup
0.38%	read_bottleneck	[kernel.kallsyms]	[k] find_get_entry

90초~100초_1 --children

Children	Self	Command	Shared Object	Symbol
25.14%	4.80%	read_bottleneck	read_bottleneck_breakdown_test	[.] main
19.62%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BackgroundCallCompaction
18.82%	0.00%	rocksdb:low0	libstdc++.so.6.0.25	[.] 0x00000000000bd6df
18.82%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::ThreadPoolImpl::Impl::BGThreadWrapper
18.82%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::ThreadPoolImpl::Impl::BGThread
18.82%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BGWorkCompaction
18.08%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BackgroundCompaction
18.05%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionJob::Run
17.86%	0.28%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionJob::ProcessKeyValueCompaction
15.48%	0.06%	rocksdb:low0	[kernel.kallsyms]	[k] entry_SYSCALL_64_after_hwframe
15.40%	0.74%	rocksdb:low0	[kernel.kallsyms]	[k] do_syscall_64
11.49%	0.00%	read_bottleneck	[unknown]	[.] 0x46e5106000000000
11.38%	0.02%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.36%	0.02%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.33%	0.02%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.30%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
11.29%	0.07%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
11.11%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::Get
11.10%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::Get
11.02%	0.09%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::GetImpl
9.89%	0.10%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionIterator::Next
9.40%	0.22%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::Version::Get
9.02%	0.00%	read_bottleneck	[unknown]	[.] 0x5541c6894d5641cf
8.90%	8.88%	read_bottleneck	libc-2.27.so	[.] __random_r
8.90%	0.06%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::PessimisticTransactionDB::Put
8.61%	0.17%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TableCache::Get
8.26%	0.29%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::Get
8.17%	0.55%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::MergingIterator::Next
7.93%	0.10%	read_bottleneck	[kernel.kallsyms]	[k] entry_SYSCALL_64_after_hwframe
7.91%	0.00%	rocksdb:low0	libpthread-2.27.so	[.] __libc_write
7.91%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] sys_write
7.91%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] vfs_write
7.91%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] __vfs_write
7.91%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] new_sync_write
7.90%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] ext4_file_write_iter
7.90%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] __generic_file_write_iter
7.82%	0.80%	read_bottleneck	[kernel.kallsyms]	[k] do_syscall_64

90초~100초_2 --children

7.82%	0.80%	read_bottleneck	[kernel.kallsyms]	[k] do_syscall_64
7.82%	0.05%	rocksdb:low0	[kernel.kallsyms]	[k] generic_perform_write
7.75%	7.66%	read_bottleneck	libc-2.27.so	[.] __random
6.96%	0.06%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::PessimisticTransaction::Commit
6.86%	0.23%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTableIterator::NextAndGetResult
6.57%	0.10%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTableIterator::Next
6.52%	0.15%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::(anonymous namespace)::LevelIterator::NextAndGetResult
6.15%	0.17%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTableIterator::FindBlockForward
5.88%	0.04%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::WriteCommittedTxn::CommitWithoutPrepareInternal
5.65%	0.26%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::WriteImpl
5.28%	0.08%	rocksdb:low0	libpthread-2.27.so	[.] __libc_pread64
5.16%	0.08%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTableIterator::InitDataBlock
5.10%	0.28%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTableBuilder::Add
5.01%	0.23%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::NewDataBlockIterator<rocksdb::DataBlockIter>
4.77%	0.12%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::NewDataBlockIterator<rocksdb::DataBlockIter>
4.49%	0.04%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::RetrieveBlock<rocksdb::Block>
4.40%	4.40%	rocksdb:low0	[kernel.kallsyms]	[k] copy_user_enhanced_fast_string
4.40%	0.18%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::MaybeReadBlockAndLoadToCache<rocksdb::Block>
4.35%	0.05%	read_bottleneck	libpthread-2.27.so	[.] __libc_pread64
4.09%	0.03%	read_bottleneck	libpthread-2.27.so	[.] __libc_write
3.87%	0.13%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::RetrieveBlock<rocksdb::Block>
3.82%	0.07%	rocksdb:low0	[kernel.kallsyms]	[k] sys_pread64
3.60%	0.04%	rocksdb:low0	[kernel.kallsyms]	[k] vfs_read
3.45%	0.04%	read_bottleneck	[kernel.kallsyms]	[k] sys_write
3.34%	0.04%	read_bottleneck	[kernel.kallsyms]	[k] sys_pread64
3.34%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] __vfs_read
3.33%	0.04%	rocksdb:low0	[kernel.kallsyms]	[k] new_sync_read
3.29%	0.04%	read_bottleneck	[kernel.kallsyms]	[k] vfs_write
3.29%	0.03%	rocksdb:low0	[kernel.kallsyms]	[k] ext4_file_read_iter
3.23%	0.23%	rocksdb:low0	[kernel.kallsyms]	[k] generic_file_read_iter
3.16%	0.03%	read_bottleneck	[kernel.kallsyms]	[k] vfs_read

90초~92초_1 --no-children

Overhead	Command	Shared Object	Symbol
9.13%	read_bottleneck	libc-2.27.so	[.] __random_r
7.82%	read_bottleneck	libc-2.27.so	[.] __random
4.86%	read_bottleneck	read_bottleneck_breakdown_test	[.] main
4.52%	rocksdb:low0	[kernel.kallsyms]	[k] copy_user_enhanced_fast_string
1.95%	rocksdb:low0	libc-2.27.so	[.] __memmove_avx_unaligned_erms
1.84%	read_bottleneck	libc-2.27.so	[.] rand
1.66%	rocksdb:low0	libc-2.27.so	[.] __memset_avx2_erms
1.66%	read_bottleneck	[kernel.kallsyms]	[k] copy_user_enhanced_fast_string
1.48%	read_bottleneck	read_bottleneck_breakdown_test	[.] rand@plt
1.38%	rocksdb:low0	libsnappy.so.1.1.7	[.] snappy::internal::CompressFragment
1.24%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::crc32c::crc32c_3way
1.07%	rocksdb:low0	libc-2.27.so	[.] cfree@GLIBC_2.2.5
0.99%	read_bottleneck	libc-2.27.so	[.] cfree@GLIBC_2.2.5
0.85%	read_bottleneck	[kernel.kallsyms]	[k] do_syscall_64
0.80%	read_bottleneck	[vdso]	[.] __vdso_clock_gettime
0.79%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::IndexBlockIter::SeekImpl
0.77%	read_bottleneck	libc-2.27.so	[.] __memmove_avx_unaligned_erms
0.75%	rocksdb:low0	[kernel.kallsyms]	[k] do_syscall_64
0.68%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionIterator::NextFromInput
0.66%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTable::KeyComparator::operator()
0.64%	rocksdb:low0	libc-2.27.so	[.] malloc
0.59%	read_bottleneck	libc-2.27.so	[.] malloc
0.56%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::crc32c::crc32c_3way
0.55%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::(anonymous namespace)::BytewiseComparatorImpl::Compare
0.51%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::MergingIterator::Next
0.51%	read_bottleneck	libc-2.27.so	[.] __memcmp_avx2_movbe
0.49%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTable::KeyComparator::operator()
0.45%	rocksdb:low0	[kernel.kallsyms]	[k] __block_commit_write.isra.34
0.44%	rocksdb:low0	libc-2.27.so	[.] _int_malloc
0.42%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockIter<rocksdb::IndexValue>::CompareCurrentKey
0.41%	read_bottleneck	libc-2.27.so	[.] _int_malloc
0.40%	read_bottleneck	[kernel.kallsyms]	[k] __radix_tree_lookup
0.39%	rocksdb:low0	libstdc++.so.6.0.25	[.] std::__cxx11::basic_string<char, std::char_traits<char>, std::allocator<char> >::_M_replace
0.38%	rocksdb:low0	[kernel.kallsyms]	[k] radix_tree_next_chunk
0.38%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::LRUHandleTable::FindPointer
0.38%	read_bottleneck	[kernel.kallsyms]	[k] find_get_entry
0.37%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTableIterator::status

90초~92초_2 --no-children

```
0.34%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::BlockFetcher::ReadBlockContents
0.33%  rocksdb:low0     read_bottleneck_breakdown_test  [.] rocksdb::BlockBuilder::Add
0.32%  read_bottleneck  [kernel.kallsyms]               [k] __entry_trampoline_start
0.32%  rocksdb:low0     read_bottleneck_breakdown_test  [.] rocksdb::CompactionJob::ProcessKeyValueCompaction
0.31%  read_bottleneck  [kernel.kallsyms]               [k] syscall_return_via_sysret
0.31%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::BlockBasedTable::Get
0.30%  rocksdb:low0     [kernel.kallsyms]               [k] crc32c_pcl_intel_update
0.30%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::Status::operator=
0.30%  rocksdb:low0     [kernel.kallsyms]               [k] __entry_trampoline_start
0.30%  rocksdb:low0     [kernel.kallsyms]               [k] get_page_from_freelist
0.28%  rocksdb:low0     [kernel.kallsyms]               [k] find_get_entry
0.27%  read_bottleneck  [vdso]                          [.] __vdso_gettimeofday
0.26%  rocksdb:low0     read_bottleneck_breakdown_test  [.] rocksdb::BlockBasedTableBuilder::Add
0.26%  rocksdb:low0     read_bottleneck_breakdown_test  [.] rocksdb::BlockFetcher::ReadBlockContents
0.26%  rocksdb:high0    libc-2.27.so                    [.] __memmove_avx_unaligned_erms
0.26%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::InlineSkipList<rocksdb::MemTableRep::KeyComparator const&>::RecomputeSpliceLevels
0.26%  rocksdb:low0     [vdso]                          [.] __vdso_gettimeofday
0.25%  rocksdb:low0     [kernel.kallsyms]               [k] syscall_return_via_sysret
0.25%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::DBImpl::WriteImpl
0.24%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::HistogramBucketMapper::IndexForValue
0.24%  rocksdb:low0     read_bottleneck_breakdown_test  [.] rocksdb::HistogramBucketMapper::IndexForValue
0.24%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::Env::Default
0.24%  rocksdb:low0     libstdc++.so.6.0.25             [.] std::__cxx11::basic_string<char, std::char_traits<char>, std::allocator<char> >::_M_append
0.24%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::Version::Get
0.23%  read_bottleneck  libc-2.27.so                    [.] __clock_gettime
0.23%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::(anonymous namespace)::PosixEnv::NowNanos
0.22%  rocksdb:low0     [kernel.kallsyms]               [k] generic_file_read_iter
0.22%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::LRUHandleTable::FindPointer
0.22%  rocksdb:low0     read_bottleneck_breakdown_test  [.] rocksdb::RandomAccessFileReader::Read
0.21%  read_bottleneck  libpthread-2.27.so              [.] __pthread_mutex_lock
0.21%  rocksdb:low0     [kernel.kallsyms]               [k] release_pages
0.21%  rocksdb:low0     read_bottleneck_breakdown_test  [.] rocksdb::BlockBasedTableIterator::NextAndGetResult
0.20%  rocksdb:low0     [kernel.kallsyms]               [k] __radix_tree_lookup
```

90초~92초_1 --children

Children	Self	Command	Shared Object	Symbol
25.53%	4.86%	read_bottleneck	read_bottleneck_breakdown_test	[.] main
18.83%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BackgroundCallCompaction
18.11%	0.00%	rocksdb:low0	libstdc++.so.6.0.25	[.] 0x00000000000bd6df
18.11%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::ThreadPoolImpl::Impl::BGThreadWrapper
18.11%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::ThreadPoolImpl::Impl::BGThread
18.11%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BGWorkCompaction
17.26%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::BackgroundCompaction
17.23%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionJob::Run
17.03%	0.32%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionJob::ProcessKeyValueCompaction
15.62%	0.07%	rocksdb:low0	[kernel.kallsyms]	[k] entry_SYSCALL_64_after_hwframe
15.54%	0.75%	rocksdb:low0	[kernel.kallsyms]	[k] do_syscall_64
11.73%	0.00%	read_bottleneck	[unknown]	[.] 0x81bcc06000000000
11.60%	0.02%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.58%	0.03%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.55%	0.03%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TransactionBaseImpl::Get
11.52%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
11.50%	0.06%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
11.34%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::Get
11.32%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::Get
11.25%	0.09%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::DBImpl::GetImpl
9.69%	0.24%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::Version::Get
9.46%	0.08%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::CompactionIterator::Next
9.15%	9.13%	read_bottleneck	libc-2.27.so	[.] __random_r
9.11%	0.00%	read_bottleneck	[unknown]	[.] 0x5541c6894d5641cf
8.96%	0.06%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::PessimisticTransactionDB::Put
8.87%	0.16%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::TableCache::Get
8.56%	0.31%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::BlockBasedTable::Get
8.06%	0.09%	read_bottleneck	[kernel.kallsyms]	[k] entry_SYSCALL_64_after_hwframe
8.01%	0.00%	rocksdb:low0	libpthread-2.27.so	[.] __libc_write
8.01%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] sys_write
8.00%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] vfs_write
8.00%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] __vfs_write
8.00%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] new_sync_write
8.00%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] ext4_file_write_iter
8.00%	0.00%	rocksdb:low0	[kernel.kallsyms]	[k] __generic_file_write_iter
7.96%	0.85%	read_bottleneck	[kernel.kallsyms]	[k] do_syscall_64
7.92%	7.82%	read_bottleneck	libc-2.27.so	[.] __random

90초~92초_2 --children

```
7.92%     0.05%  rocksdb:low0     [kernel.kallsyms]                [k] generic_perform_write
7.86%     0.51%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::MergingIterator::Next
6.93%     0.06%  read_bottleneck  read_bottleneck_breakdown_test   [.] rocksdb::PessimisticTransaction::Commit
6.67%     0.21%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTableIterator::NextAndGetResult
6.40%     0.11%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTableIterator::Next
6.31%     0.13%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::(anonymous namespace)::LevelIterator::NextAndGetResult
5.93%     0.15%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTableIterator::FindBlockForward
5.84%     0.04%  read_bottleneck  read_bottleneck_breakdown_test   [.] rocksdb::WriteCommittedTxn::CommitWithoutPrepareInternal
5.64%     0.25%  read_bottleneck  read_bottleneck_breakdown_test   [.] rocksdb::DBImpl::WriteImpl
5.27%     0.08%  rocksdb:low0     libpthread-2.27.so               [.] __libc_pread64
5.04%     0.05%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTableIterator::InitDataBlock
4.91%     0.17%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTable::NewDataBlockIterator<rocksdb::DataBlockIter>
4.88%     0.11%  read_bottleneck  read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTable::NewDataBlockIterator<rocksdb::DataBlockIter>
4.64%     0.26%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTableBuilder::Add
4.62%     0.05%  read_bottleneck  read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTable::RetrieveBlock<rocksdb::Block>
4.53%     4.52%  rocksdb:low0     [kernel.kallsyms]                [k] copy_user_enhanced_fast_string
4.52%     0.19%  read_bottleneck  read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTable::MaybeReadBlockAndLoadToCache<rocksdb::Block>
4.47%     0.04%  read_bottleneck  libpthread-2.27.so               [.] __libc_pread64
4.08%     0.03%  read_bottleneck  libpthread-2.27.so               [.] __libc_write
3.94%     0.11%  rocksdb:low0     read_bottleneck_breakdown_test   [.] rocksdb::BlockBasedTable::RetrieveBlock<rocksdb::Block>
3.78%     0.06%  rocksdb:low0     [kernel.kallsyms]                [k] sys_pread64
3.57%     0.04%  rocksdb:low0     [kernel.kallsyms]                [k] vfs_read
3.45%     0.04%  read_bottleneck  [kernel.kallsyms]                [k] sys_write
3.43%     0.04%  read_bottleneck  [kernel.kallsyms]                [k] sys_pread64
3.34%     0.00%  rocksdb:low0     [kernel.kallsyms]                [k] __vfs_read
3.33%     0.03%  rocksdb:low0     [kernel.kallsyms]                [k] new_sync_read
3.30%     0.03%  rocksdb:low0     [kernel.kallsyms]                [k] ext4_file_read_iter
3.30%     0.03%  read_bottleneck  [kernel.kallsyms]                [k] vfs_write
3.25%     0.22%  rocksdb:low0     [kernel.kallsyms]                [k] generic_file_read_iter
3.25%     0.04%  read_bottleneck  [kernel.kallsyms]                [k] vfs_read
3.15%     0.00%  read_bottleneck  [kernel.kallsyms]                [k] __vfs_write
3.15%     0.01%  read_bottleneck  [kernel.kallsyms]                [k] new_sync_write
```

Symbol 분석

Children	Self	Command	Shared Object	Symbol
9.97%	0.22%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::Version::Get
0.33%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTableListVersion::Get
0.33%	0.02%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTableListVersion::GetFromList

Children	Self	Command	Shared Object	Symbol
9.99%	0.24%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::Version::Get
0.31%	0.00%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTableListVersion::Get
0.31%	0.01%	read_bottleneck	read_bottleneck_breakdown_test	[.] rocksdb::MemTableListVersion::GetFromList
0.00%	0.00%	rocksdb:high0	read_bottleneck_breakdown_test	[.] rocksdb::Version::GetTableProperties
0.00%	0.00%	rocksdb:low0	read_bottleneck_breakdown_test	[.] rocksdb::MemTableListVersion::GetEarliestSequenceNumber

Symbol 분석

```
Children      Self  Command          Shared Object                Symbol
........  ........  ...............  ...........................  ..........................................

   1.78%     0.80%  read_bottleneck  read_bottleneck_breakdown_test  [.] rocksdb::IndexBlockIter::SeekImpl
          |
          |--0.98%--rocksdb::IndexBlockIter::SeekImpl
          |          |
          |           --0.62%--rocksdb::BlockIter<rocksdb::IndexValue>::CompareCurrentKey
          |
           --0.80%--0x409d106000000000
                     main
                     rocksdb::TransactionBaseImpl::Get
                     rocksdb::TransactionBaseImpl::Get
                     rocksdb::TransactionBaseImpl::Get
                     rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
                     rocksdb::WriteBatchWithIndex::GetFromBatchAndDB
                     rocksdb::DBImpl::Get
                     rocksdb::DBImpl::Get
                     rocksdb::DBImpl::GetImpl
                     rocksdb::Version::Get
                     rocksdb::TableCache::Get
                     rocksdb::BlockBasedTable::Get
                     |
                      --0.79%--rocksdb::BlockIter<rocksdb::IndexValue>::Seek
                                rocksdb::IndexBlockIter::SeekImpl
```

분석 결과

1. 초반 N초와 후반 N초의 기록 간의 차이는 compaction의 유무 이외에는 크게 발견되지 않았다.

   총 실행의 CPU Cycle에서 차지하는 비율이 각 함수마다 크게 달라지지 않았고, 비율의 순위도 크게 변동되지 않았다.

2. GET 함수의 내용 중 가장 높은 비율을 차지하는 부분은 SST File에서 탐색하는 부분이었으며, 이는 SST File 에서 탐색하면서 참조하는 Data Block의 수가 많기 때문인 것으로 추측된다.

   또한, Data Block을 탐색하면서 sequence number는 사용하지 않고, user key만을 사용하는 것으로 파악되었다. Sequence number는 key value를 찾은 후 이를 validate하는 과정에서 사용되고 있었다.

To Do

해결해야 할 문제

질문

해결해야 할 문제 및 질문

1. 여러 방식으로 실험을 했음에도 유의미한 차이가 나타나지 않았는데 저희가 놓친 부분이 있나요?
   혹시 CPU Cycle 비율이 아닌 다른 척도를 확인해야 하나요?

2. 지금의 저희의 가설으로는 sequence number 를 더 활용해서 데이터 블록을 탐색하는 횟수를 줄이는 것이 성능 개선의 방법일 것이라고 생각합니다. 혹시 이런 방향성이 맞을까요?

THANK YOU!

감사합니다!