

Contents

1	TODO check	1
1.1	TODO src code	1
1.2	DONE default config [2/2]	1
2	RocksDB settings	1
2.1	Terminology	1
3	WiredTiger settings	2
4	Bench	2
4.1	General	2
4.2	Load	3
4.3	Run Workerload A (UPDATE heavy UPDATE:READ=0.95:0.05)	3
4.4	Run Workerload B (READ most READ:UPDATE=0.95:0.05)	3
4.5	Run Workerload C (READ only)	4
4.6	Run Workerload D (READ lastest READ:INSERT=0.95:0.05)	4
4.7	Run Workerload E (SCAN heavy SCAN:INSERT=0.95:0.05)	4
4.8	Run Workerload F(Read-Modify-Update READ:WRITE=1.5:1)	5

1 TODO check

1.1 TODO src code

1.2 DONE default config [2/2]

- State "DONE" from "TODO" [2023-02-22 Wed 11:18]
rocksdb: bloom bits 10 is recommended in both header and wiki.
wiredtiger: block compressor snappy

☒ rocksdb

☒ wiredtiger

2 RocksDB settings

Properties	Value	Remark
------------	-------	--------

Continued on next page

Continued from previous page

Properties	Value	Remark
rocksdb.compression	snappy	
rocksdb.max_background_jobs	2	
rocksdb.target_file_size_base	67108864	64MB
rocksdb.target_file_size_multiplier	1	
rocksdb.max_bytes_for_level_base	268435456	256MB
rocksdb.write_buffer_size	67108864	64MB
rocksdb.max_open_files	-1	files opened are always kept open
rocksdb.max_write_buffer_number	2	
rocksdb.use_direct_io_for_flush_compaction	false	
rocksdb.use_direct_reads	false	
rocksdb.allow_mmap_writes	false	
rocksdb.allow_mmap_reads	false	
rocksdb.cache_size	8388608	8MB
rocksdb.compressed_cache_size	0	
rocksdb.bloom_bits	10	
rocksdb.increase_parallelism	false	
rocksdb.optimize_level_style_compaction	false	

2.1 Terminology

Term	Description
block cache	in-memory data structure that cache the hot data blocks from the SST
target_file_size_base	target_file_size_base is per-file size for level1, in bytes
max_bytes_for_level_base	max size for level1, in bytes
write buffer	memtable

3 WiredTiger settings

Properties	Value	Remark
wiredtiger.cache_size	100MB	For BTree ONLY, including associated in-mem data structures
wiredtiger.direct_io	[]	

Continued on next page

Continued from previous page

Properties	Value	Remark
wiredtiger.in_memory	false	
wiredtiger.lsm_mgr.merge	true	
wiredtiger.lsm_mgr.max_workers	4	# of threads doing background merging
wiredtiger.blk_mgr.allocation_size	4KB	OS PAGE_SIZE
wiredtiger.blk_mgr.bloom_bit_count	16	per key
wiredtiger.blk_mgr.bloom_hash_count	8	
wiredtiger.blk_mgr.chunk_max	5GB	
wiredtiger.blk_mgr.chunk_size	10MB	
wiredtiger.blk_mgr.compressor	snappy	
wiredtiger.blk_mgr.btree.internal_page_max	4KB	
wiredtiger.blk_mgr.btree.leaf_key_max	0	1/10 the size of a newly split page
wiredtiger.blk_mgr.btree.leaf_value_max	0	1/2 the size a newly split page
wiredtiger.blk_mgr.btree.leaf_page_max	32KB	

4 Bench

4.1 General

Key size	Value size	Record count	Operation count	thread count
8B	24B	1G (10M for wt)	100M	8

```
./ycsb -run -db rocksdb -P workloads/workloadf -P rocksdb/rocksdb.properties -s \
-p fixedkey8b=true -p fixedfieldlen=true -p fieldcount=1 -p fieldlength=24 \
-p recordcount=1000000000 -p operationcount=100000000 -p threadcount=8
```

4.2 Load

DB	Load runtime(sec)	Load through-put(ops/sec)	Max	Min
Avg(us)	90	99	99.9	99.99
RocksDB-INSERT	5669.34	176,387	32079.87	4.40
44.93	20.02	31.90	14254.08	15114.24

Continued on next page

Continued from previous page

DB	Load runtime(sec)	Load throughput(ops/sec)	Max	Min
Avg(us)	90	99	99.9	99.99
WiredTiger-INSERT	676.592	14,780	56492.03	0.57
532.24	974.85	7712.77	16039.93	24297.47

4.3 Run Workerload A (UPDATE heavy | UPDATE:READ=0.95:0.05)

DB	Run runtime(sec)	Run throughput(ops/sec)
RocksDB	1267.29	78,908
WiredTiger	373.14	26,799

DB	Max	Min	Avg(us)	90	99	99.9	99.99
RocksDB-UPDATE	201719.81	4.50	105.02	273.41	908.29	2920.45	5509.12
RocksDB-READ	203030.53	0.30	95.30	261.12	896.00	2902.01	5414.91
Wiredtiger-UPDATE	54329.34	0.95	297.49	522.24	3448.83	14852.09	24461.31
WiredTiger-READ	52527.10	0.26	293.46	514.05	3403.78	14745.60	24379.39

4.4 Run Workerload B (READ most | READ:UPDATE=0.95:0.05)

DB	Run runtime(sec)	Run throughput(ops/sec)
RocksDB	1113.47	89,810
WiredTiger	152.114	65,740

DB	Max	Min	Avg(us)	90	99	99.9	99.99
RocksDB-UPDATE	18022.40	5.00	99.50	276.99	603.65	2217.98	3205.12
RocksDB-READ	32063.49	0.30	87.75	263.94	591.87	2207.74	3170.30
Wiredtiger-UPDATE	19251.20	1.05	130.26	225.79	575.49	3772.41	10911.74
WiredTiger-READ	24903.68	0.22	118.51	219.90	374.01	2748.79	8749.06

4.5 Run Workerload C (READ only)

DB	Run runtime(sec)	Run throughput(ops/sec)
RocksDB	1113.11	89,837
WiredTiger	127.54	78,407

DB	Max	Min	Avg(us)	90	99	99.9	99.99
RocksDB-READ	93061.12	1.20	88.31	264.70	515.58	1478.65	4313.09
WiredTiger-READ	14417.92	0.20	100.05	214.53	256.13	673.79	1638.40

4.6 Run Workerload D (READ lastest | READ:INSERT=0.95:0.05)

DB	Run runtime(sec)	Run throughput(ops/sec)
RocksDB	446.869	223,779
WiredTiger	133.523	74,893

DB	Max	Min	Avg(us)	90	99	99.9	99.99
RocksDB-INSERT	6078.46	4.20	11.75	15.61	22.63	29.50	78.91
RocksDB-READ	30605.31	0.20	36.29	123.71	409.34	2026.49	2951.17
Wiredtiger-INSERT	19218.43	2.20	172.26	245.25	914.94	3348.48	10280.96
WiredTiger-READ	21200.90	0.24	101.55	220.29	407.55	2729.98	8699.90

4.7 Run Workerload E (SCAN heavy | SCAN:INSERT=0.95:0.05)

DB	Run runtime(sec)	Run throughput(ops/sec)
RocksDB	1000.00	49,308
WiredTiger	158.148	63,232

DB	Max	Min	Avg(us)	90	99	99.9	99.99
RocksDB-INSERT	3143.68	4.30	14.34	17.50	24.21	31.61	81.86
RocksDB-SCAN	41123.84	5.70	167.98	380.67	708.61	2022.40	3336.19
Wiredtiger-INSERT	20234.24	2.15	162.22	239.49	807.42	4093.95	13197.31
WiredTiger-SCAN	30064.62	0.25	121.98	231.55	329.21	2684.93	10616.83

4.8 Run Workerload F(Read-Modify-Update | READ:WRITE=1.5:1)

DB	Run runtime(sec)	Run throughput(ops/sec)
RocksDB(actually 1.5G records)	1192.88	125,843
WiredTiger(actually 15M records)	355.461	42,204

DB	Max	Min	Avg(us)	90	99	99.9	99.99
RocksDB-READ	62226.43	0.30	87.69	242.69	685.57	2836.48	5316.61
RocksDB-UPDATE	11034.62	4.30	13.26	20.91	28.51	37.50	91.84
Wiredtiger-READ	50200.57	0.22	279.12	270.33	4042.75	15335.42	24887.29
WiredTiger-UPDATE	27394.05	0.72	4.32	4.51	7.17	35.42	2942.97