

RAFT based Key-Value Store with Transaction Support

CS244B: DISTRIBUTED SYSTEMS COURSE PROJECT

Chen Chen; Varun Kulkarni; Supriya Premkumar; Renga Srinivasan

June 9 2020

Motivation

- INCREASING TREND TOWARDS STRONG CONSISTENT, FAULT-TOLERANT KV STORES
- MULTI-SHARD TRANSACTIONS
- EXPLORE IMPLEMENTATION CHALLENGES

Properties

Properties	Our system	etcd
Consensus	hashicorp/raft	etcd/raft
Linearizable Reads/Writes/Deletes	Yes	Yes
Serializable (concurrent) Transactions	Yes	Yes
Multi-shard Transaction	Yes (two-phase commit)	No
Multi-Versioned Keys	No	Yes

Operation Support

SIMPLE OPERATION

- GET

- › GET BALANCE

- SET

- › SET BALANCE 100

COMPOUND OPERATION

- TRANSACTION

- › TXN

- › SET X 100

- › SET Y 100

- › DEL Z

- › END

CONDITIONAL OPERATION

- ADD

- › ADD BALANCE 50

- SUB

- › SUB BALANCE 25

- TRANSFER

- › XFER BALANCE_X BALANCE_Y 20

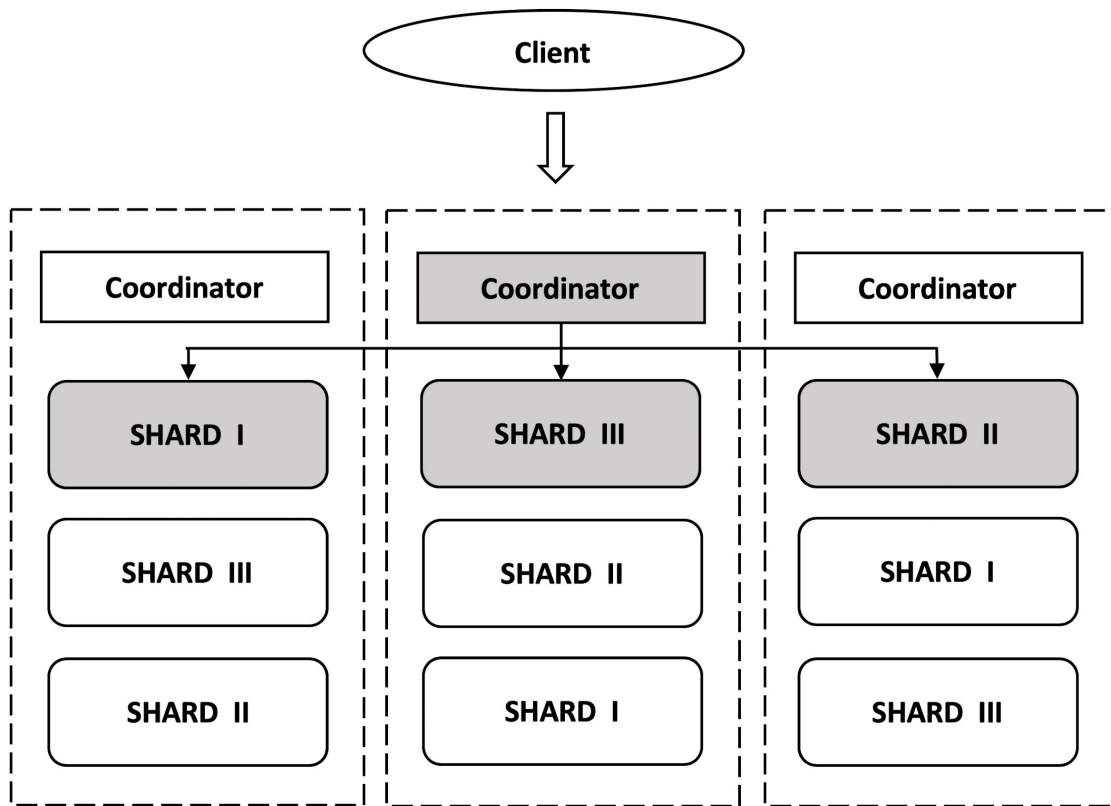
Architecture

MULTIPLE RAFT GROUPS

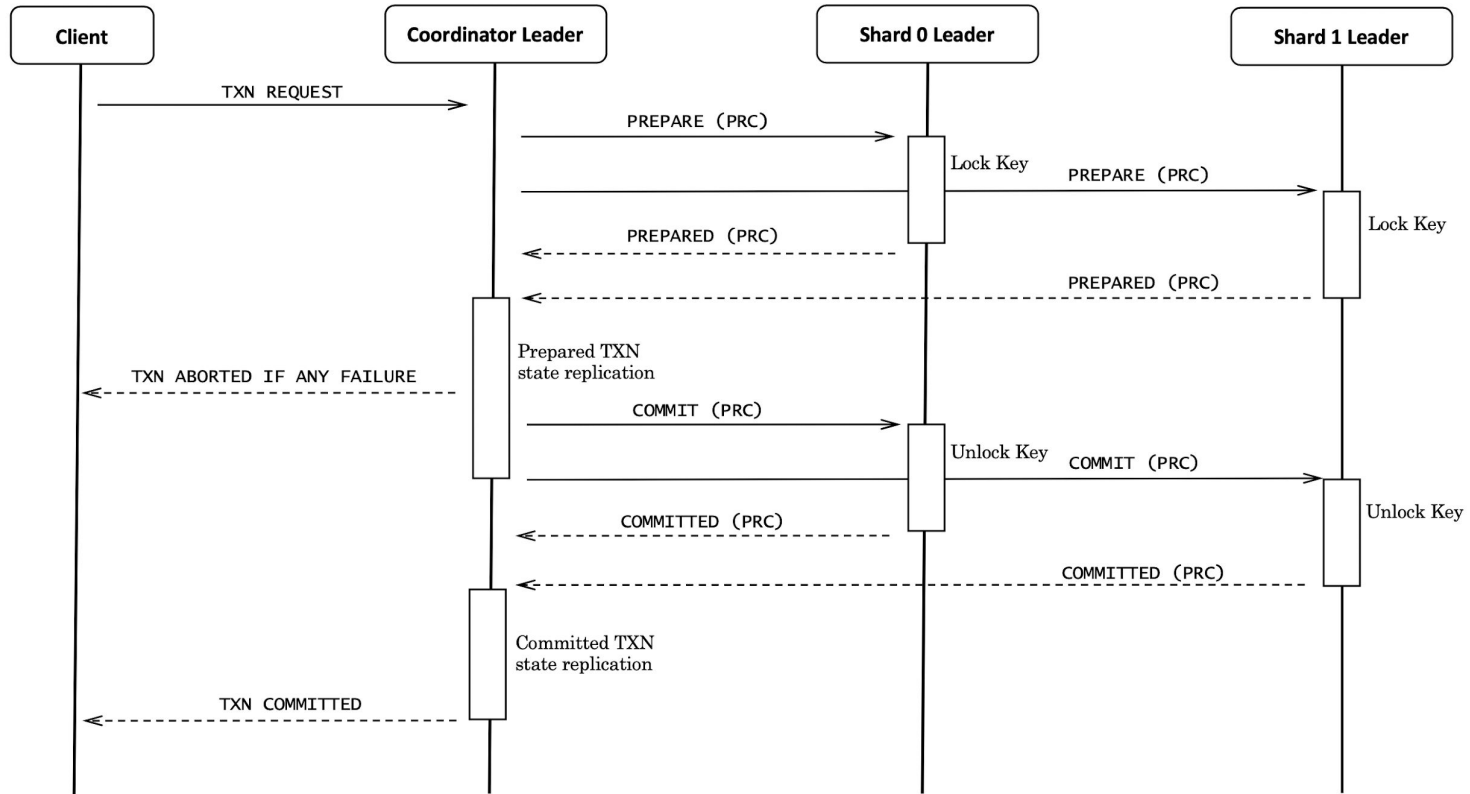
- COORDINATOR
- STORE SHARD

TRANSACTION

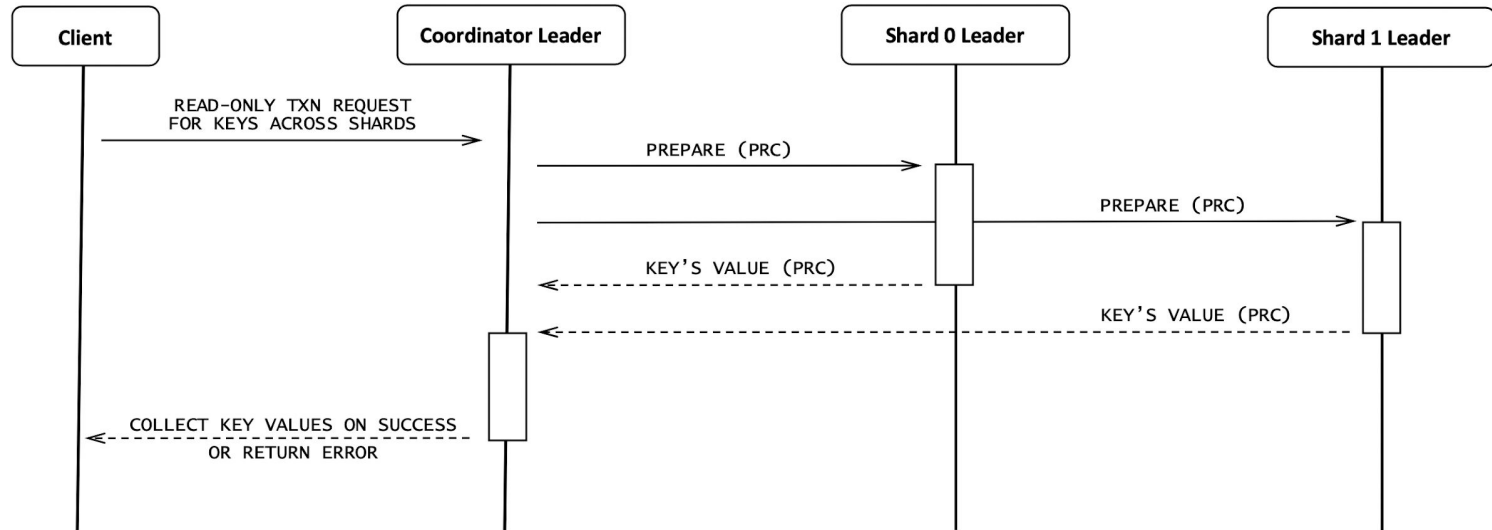
- TWO-PHASE LOCKING
- TWO-PHASE COMMIT
- TRANSACTION RECOVERY



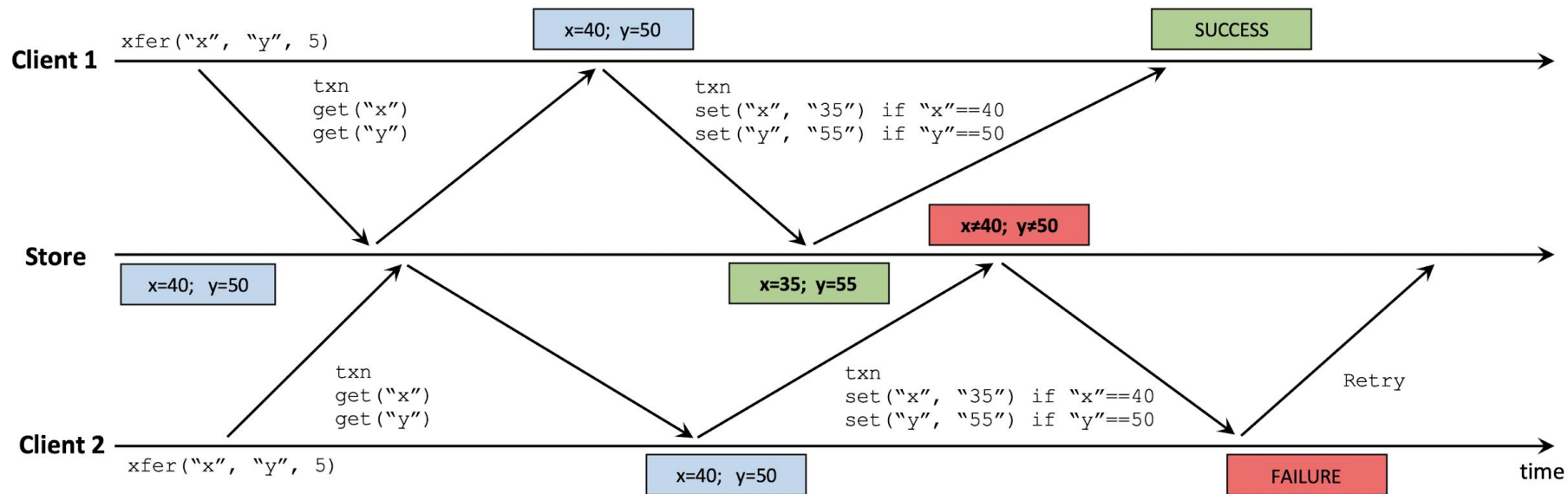
Two-Phase Commit Transaction



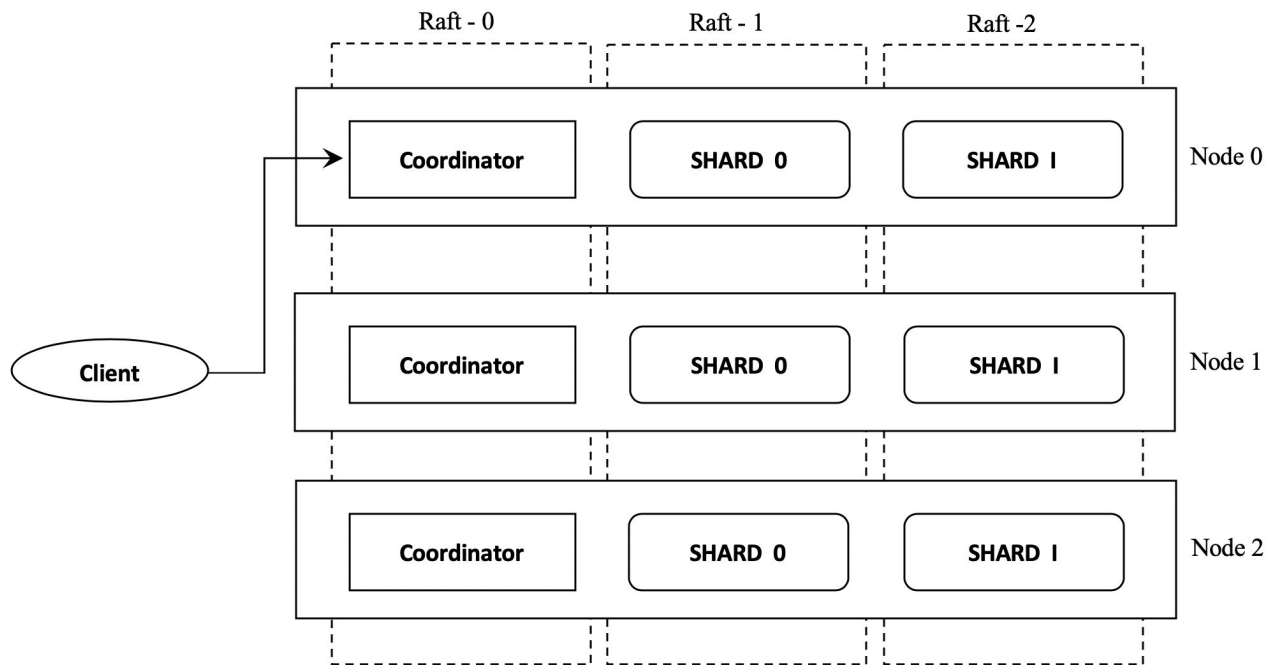
Read-Only Transaction



Optimistic Locking



Demo Setup



- Operations
- Fault-Tolerant
- Transaction Recovery



Demo Video

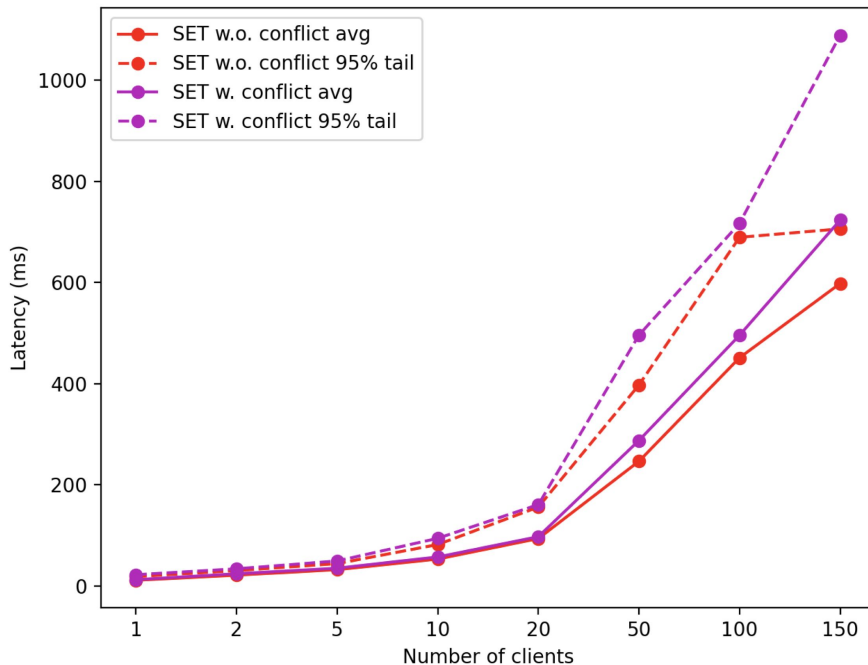
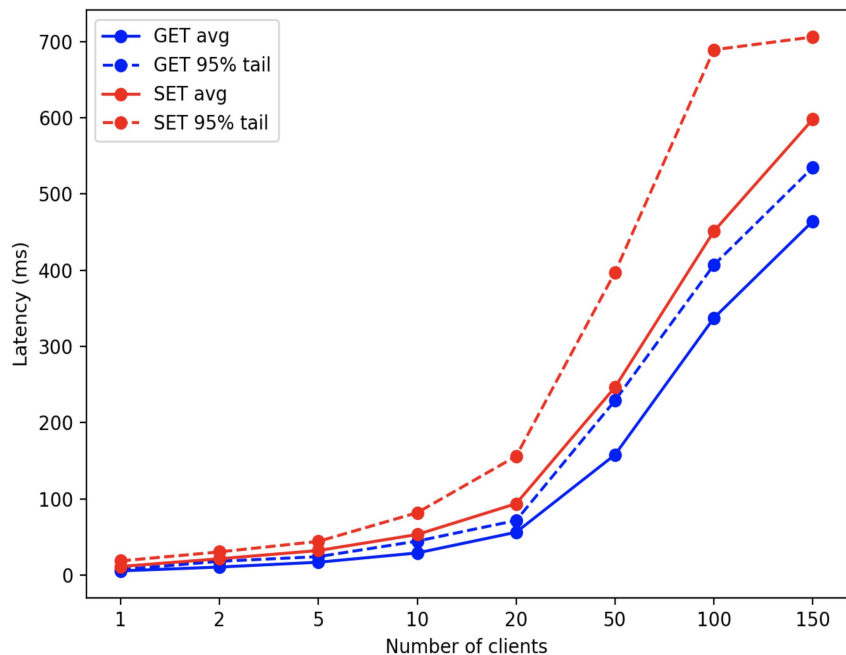
```
#M1 Retrying with alternate server:http://node1:17000
#M2 set txn failed with err: 10.10.10.2:18000
#M3 Retrying 1 times...
Submitting [method:"get" key:"x" method:"get" key:"y"]
get txn failed with err: 10.10.10.2:18000
Retrying 2 times...
Submitting [method:"get" key:"x" method:"get" key:"y"]
get txn failed with err: 10.10.10.2:18000
Retrying 3 times...
Submitting [method:"get" key:"x" method:"get" key:"y"]
get txn failed with err: 10.10.10.2:18000
Retrying 4 times...
Submitting [method:"get" key:"x" method:"get" key:"y"]
get txn failed with err: 10.10.10.2:18000
Retrying 5 times...
XFER Successful
>get x
redirecting request to leader: http://node2:17000
OK
>get x
Key=x, Value=40
>get y
Key=y, Value=60
>

# raft-kv-store git:(recovery) # tail -f node0/logs/coord.log | grep coordinator
Jun  8 02:19:47.279 [api.go:382][JoinC] [INFO] [coordinator] received join request f
on remote node node2 at node2:18000
Jun  8 02:19:47.284 [api.go:332][JoinC] [INFO] [coordinator] node node2 at node2:180
00 joined successfully
Jun  8 02:20:45.023 [api.go:51][SetC] [INFO] [coordinator] Processing Set request: K
ey=x Value=45
Jun  8 02:20:47.597 [api.go:51][SetC] [INFO] [coordinator] Processing Set request: K
ey=y Value=55
Jun  8 02:22:47.087 [api.go:118][TransactionC] [INFO] [coordinator] Processing Trans
action
Jun  8 02:22:47.087 [api.go:126][TransactionC] [INFO] [coordinator] Starting prepare
phase for txid: breqltpo8r1ka4cb2bg
Jun  8 02:22:47.088 [api.go:134][TransactionC] [INFO] [coordinator] [txid breqltpo8
r1ka4cb2bg] is read-only
Jun  8 02:22:47.105 [api.go:152][TransactionC] [INFO] [coordinator] [txid: breqltpo
8r1ka4cb2bg] read-only transaction, returning after prepare phase
Jun  8 02:22:47.108 [api.go:118][TransactionC] [INFO] [coordinator] Processing Trans
action
Jun  8 02:22:47.108 [api.go:126][TransactionC] [INFO] [coordinator] Starting prepare
phase for txid: breqltpo8r1ka4cb2bg
Jun  8 02:22:47.124 [api.go:156][TransactionC] [INFO] [coordinator] [txid breqltpo8
r1ka4cb2bg] Prepared sent
Jun  8 02:22:47.125 [api.go:194][TransactionC] [INFO] [coordinator] [txid: breqltpo
8r1ka4cb2bg] Prepared received: 2 Prepared Expected: 2
Jun  8 02:22:47.125 [coordinator.go:195][ReplicateC] [INFO] [coordinator] Committing
on coordinator
Jun  8 02:22:47.126 [api.go:207][TransactionC] [INFO] [coordinator] Simulating node
failure after prepared phase, kill the node

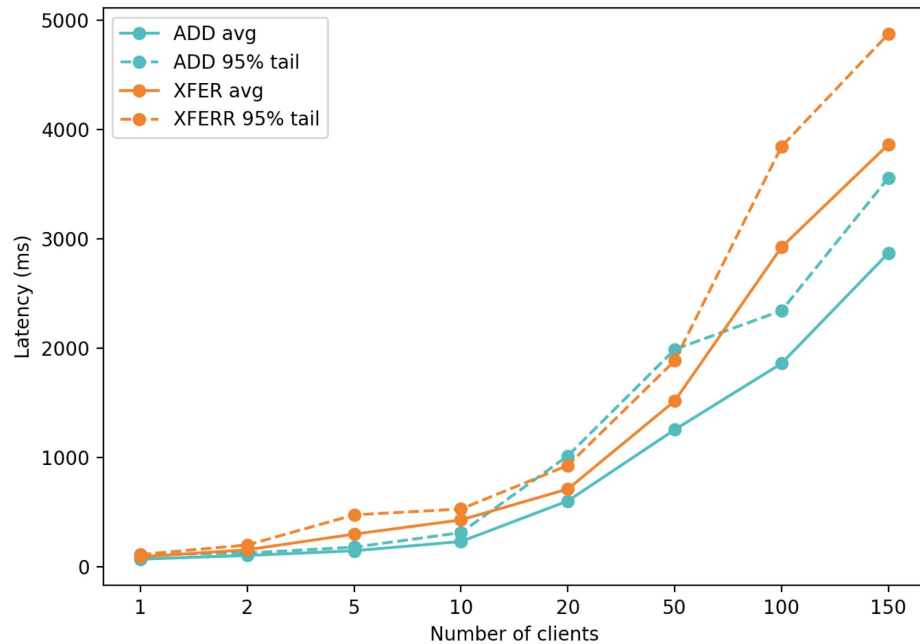
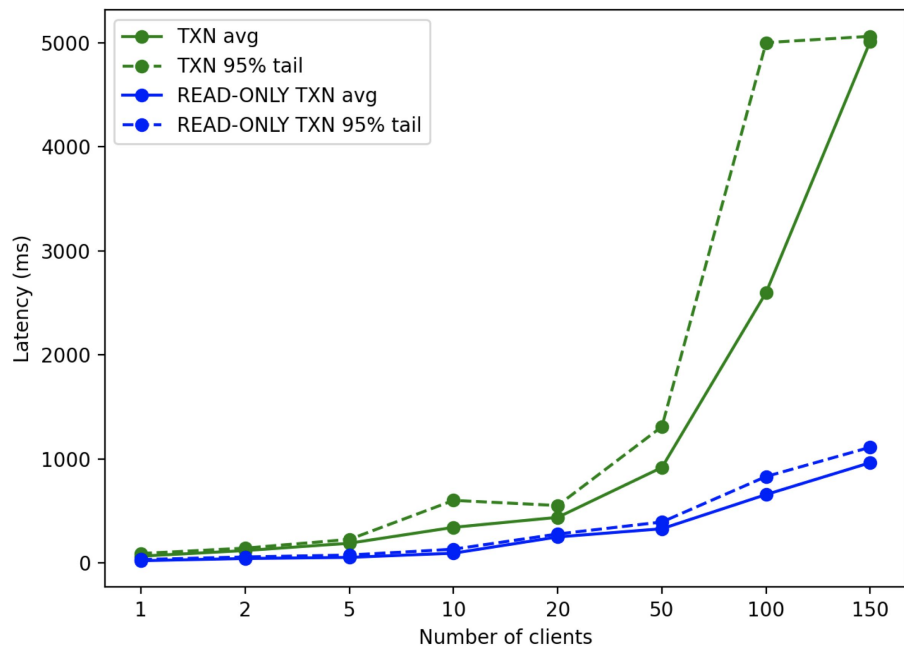
# raft-kv-store git:(recovery) # tail -f node0/logs/coord.log | grep coordinator
Jun  8 02:19:46.871 [coordinator.go:68][NewCoordinatorC] [INFO] [coordinator] Prepar
ing node-model with persistent directory /Users/varun.kulkarni/RAFT-KV-STORE/node0/co
ords, raftAddress node1:18000
Jun  8 02:19:46.884 [coordinator.go:99][NewCoordinatorC] [INFO] [coordinator] Starti
ng coordinator
Jun  8 02:19:46.892 [main.go:87][mainC] [INFO] [main] coordinator started successf
ully

# raft-kv-store git:(recovery) # docker exec node ps
Error: No such container: node
# raft-kv-store git:(recovery) # docker exec node0 ps
PID USER TIME COMMAND
1 root 0:00 bash /bootstrap.sh
8 root 0:01 kv -i node0 -l :17001 -r node0:18001 -d node0/shard-0
9 root 0:01 kv -i node0 -l :17002 -r node0:18002 -d node0/shard-1
11 root 0:00 kv -i node0 -c -l :17000 -r node0:18000 -d node0 -t co
mmit
69 root 0:00 ps
74 root 0:00 sleep 1
# raft-kv-store git:(recovery) # docker exec node0 kill 11
# raft-kv-store git:(recovery) #
```

GET and SET Latency



TXN and XFER Latency



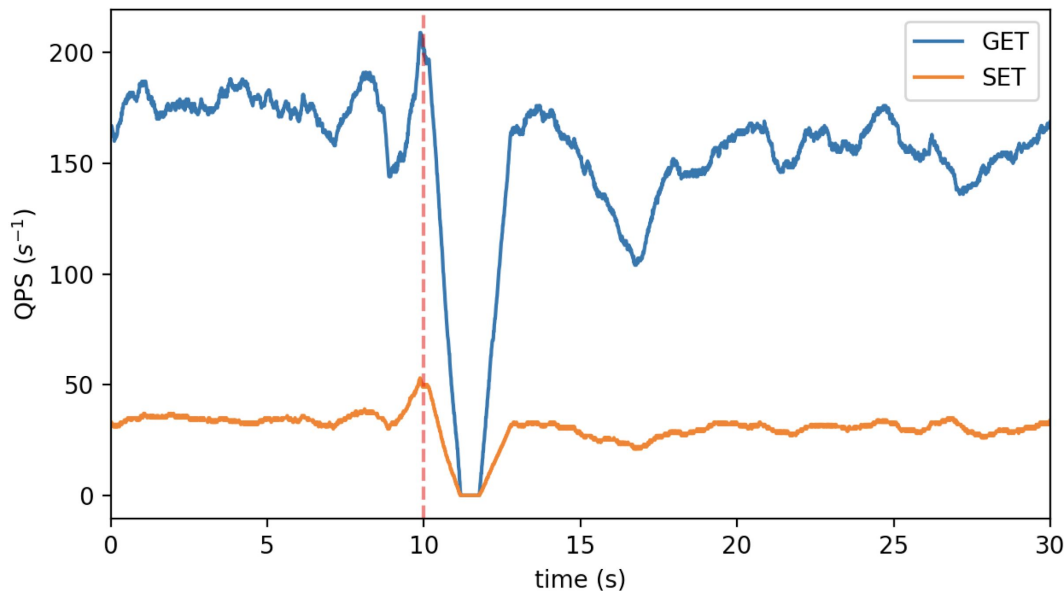
Fault Tolerance Test

REQUEST RATE

- 80% GET
- 20% SET

NODE FAILURE

- NODE WITH COORDINATOR LEA
FAILS AT 10s



Questions?

Thank You