DiemBFT Documentation and Test Report

Team Name: RVP

Raj Patel, 114363611 Prince Kumar Maurya, 114354075 Vishal Singh, 114708875

Setup

Test case Template

Generating Test cases

Logging files and Ledger Info

Implementation

Pending Block Tree

MemPool

Persistent Ledger

Speculative Ledger

Initiation of Diem Process and chain processing

Genesis Block

Client Request

Replica Info

Chain Termination

Test Case Report

Normal Execution Flow

Normal Replicas with multiple clients

Client requests timeout, re-submission, and handling request

de-duplication

Omission Failure

Forge Signature

Delay Failures

Setup

Test case Template

The following is the test case template used for all the test cases. In the figure shown above, descriptions of all fields are provided.

Generating Test cases

The test cases are generated using the testdiem.da file. The following code fetches the appropriate replica object.

```
def getReplicafromConfiguration(scenario):
    if scenario == "omission":
     return new(diem_replica_omission.Replica_Omission), self.specialArguments
    elif scenario == "normal":
     return new(diem_replica.Replica), {}
    elif scenario == "forge_signature":
     return new(diem_replica_forge.Replica), self.specialArguments
    elif scenario == "delay":
     return new(diem_replica_delay.Replica_Delay), self.specialArguments
def run:
   replica, special = self.getReplicafromConfiguration(self.scenario)
   specialArgs[i] = special
   replicas.append(replica)
 while i < replicas_required:</pre>
    replica, special = self.getReplicafromConfiguration("normal")
    specialArgs[i] = special
    replicas.append(replica)
```

The test cases are generated using the testdiem.da file. The following code fetches the appropriate replica object. The run method uses getReplicaConfiguration to use the appropriate replica based on the scenario specified. It can run multiple test cases one after another if the at-line position 202 in the code return value is an array.

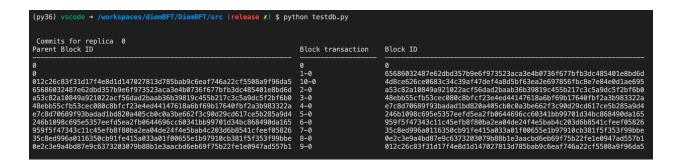
The non-faulty replicas which are equivalent to 2*f + 1 are invoked using the normal scenario.

Logging files and Ledger Info

The logging file is generated based on the file specified during the command. using the command specified in the readme will generate a log in out.log file.

The ledger is stored on a flat file in the /tmp directory corresponding to diemLedger_*. (Here * corresponds to the replica/validator ID).

To obtain the ledger after the test run. You can run the `testdb.py` file. The following is a sample output :



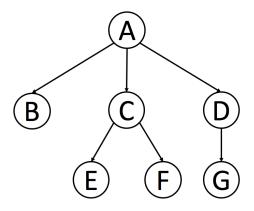
Note: After running the script, it clears the ledger. Also, the Block transaction is a string that is sorted. So although 10-0 appears first it s the last block getting committed. This can be verified by looking at the parent block ID.

Implementation

(This section is just to provide detailed explanation about our design decisions and reasoning. For actual pseudo-codes refer to the pseudo code document).

Pending Block Tree

The pending block tree is essentially a <u>trie</u>-like / n-ary tree data structure with the ability to perform pruning.



In the figure shown above A, B, C, etc are essentially block-ids. Each node contains a dictionary/map for its children. For example, A will contain B, C, D as its children, likewise C will contain E, F and D will contain G.

For making the lookup / adding / pruning transactions quick we have used a <u>cache</u> on top of this. The cache has a pointer to the reference in the tree.

The following is the pseudo-code for adding a transaction into the tree.

```
def add(prev_node_id, block):
    node = self.get_node(prev_node_id)
    if node == null:
        node = root

    node.childNodes[block.id]=Node(prev_node_id,block)
    cache[block.id]=node.childNodes[block.id]
```

The following is the pseudo-code for pruning a node

```
def prune(self,id):
    curr_node = self.get_node(id)

if curr_node == null:
    return
    self.root = curr_node
    self.cache_cleanup(id)

def cache_cleanup(id):
    cache = {}
    prune_helper(root)

def prune_helper(node):
    if node is None:
        return

    cache[node.block.id]=node

for block_id in node.childNodes.keys():
        cache[block_id] = node.childNodes[block_id]
        prune_helper(node.childNodes[block_id])
```

MemPool

The mempool uses a queue for holding the incoming transactions. It also uses a dictionary known as a locator for storing the client until the transaction is committed. If in case a transaction is chosen by a replica for proposal it uses a set (state) for indicating the transaction is in use.

The state and locator together are used for preventing request-duplication whenever client timeout is reached a new request is given.

```
Module MemPool:
    init():
        self.queue = queue()
        self.locator = {}
        self.state = set()
```

```
def get_transactions():
    # currently only sends one transaction
    if queue has elements:
        command = queue.deque()
        # command is present in locator and command not currently processing
        if command in locator and command not in state:
            self.state.add(command)
            return command
        else:
            return self.get_transactions()
    else:
        return None
```

Pseudo-code for inserting a command

```
def insert_command(command, client):
    # command not present in locator and command is not processing
    if command not in self.locator and command not in self.state:
        self.queue.append(command)
        self.locator[command] = client
    else:
        print("Command already present in mempool")
```

Persistent Ledger

For ledger persistent, each replica is storing its ledger in a flat-file inside the temp directory. The API used for storing and retrieving blocks is using <u>LevelDB</u>. For storage, we are using block_id as the key and block as the value.

While writing we make use of the sync flag which ensures flush happens before execution is returned back to the validator.

Only the committed blocks are stored inside the persistent ledger.

Speculative Ledger

The speculative ledger is similar to the persistent ledger. The only difference is that it stores the pending blocks as well. Thus in some way, it becomes storage of block to the block ids in the pending block tree.

Instead of keeping the speculative ledger in memory, we are using a flat file similar to a persistent ledger. This ensures the process does not run out of memory.

Initiation of Diem Process and chain processing

The initiation happens with round -1. Leader election chooses replica 0 as the first leader just for this round and advances_round_qc. Afterwhich block 1 is proposed and a new QC gets generated after receiving required vote messages. After generating a new QC, replica 0 again makes a proposal, and the next leader, replica 1 waits for vote messages.

The reason why we use current_round as -1 as opposed to round 0 as mentioned in the paper is to start with an even length cycle. If we start with 0, the vote messages for the proposal will be sent to replica 1 which is the next leader. (round 0 -> advance round to 1 -> broadcast proposal message -> send vote to next leader (current round 1 + 1) / 2 which will be replica 1).

Genesis Block

The genesis block is the very first block in the ledger. It contains 0 as its block ID and references parent ID as itself. It is portrayed as being formed at round -1, likewise, its parent round is portrayed as -1 as well.

Similar to the genesis block, genesis QC is formed as well. Which contains no votes as signatures but has been authored as 0.

<u>During leader election</u>, there are conditions in place such that the genesis block or genesis QC is not used for electing reputation leaders.

Client Request

The client request object holds the transaction as payload as well as the signature of the payload.

When the transaction is processed the client can request a replica to provide the committed block from the ledger and then use the payload to verify the signature with the original payload.

```
class ClientRequest:
    def init(payload, source, pvt_key):
        payload = payload
        source = source
        signature = sign(payload, pvt_key)
```

Replica Info

The replica info object is used to provide the metadata about a replica to other replicas/clients. It contains the public key of the replica, the process ID used to send messages, and its replica ID.

```
class ReplicaInfo:
    def init(process, public_key, replicaID):
        process = process
        public_key = public_key
        replicaID = replicaID
```

Chain Termination

Diem follows a 2-chain protocol where the Block proposed at round N will be committed at round N + 2 at all replicas. In order to ensure the termination of all transactions sent by the client, we make use of 2 additional dummy blocks. These dummy (empty) blocks serve as placeholders so that whatever client transaction has been sent gets committed. The only flaw with this approach is that the first dummy block will get committed at any one of the replicas. As QC for dummy block, 2 gets formed but not propagated as there are no further transactions in mempool.

As these dummy blocks are empty their presence <u>has no impact on the state of the blockchain</u>. For example: In a monetary blockchain, these blocks can be replaced with 0 value denominations. So a transaction like that will have no effect on the final state.

<u>Another way</u> of providing this functionality is to generate blocks whenever get_transactions is called. (This is the way the real-world Blockchain system maintains liveness). Another way is to generate the blocks and see if the last two generated blocks are dummy or empty blocks and terminate based on that behaviour.

In case no transactions are available for processing, the replicas go into an await where they wait for a transaction to appear to continue processing again. Round numbers will get incremented. This can be further optimized by not setting a timer for a round when the mempool is empty.

Test Case Report

All configuration files are in testdiem.da. Logging is common as specified by the run command. Check test setup section above for more details.

Normal Execution Flow

The format is as defined in the test template above. This case does not have any byzantine behavior. The log file is as defined in the test setup above.

The expected output is all 10 transactions to be executed and committed into the ledger. The resulting ledger state should look similar to the image below.

Normal Replicas with multiple clients

```
'Name' : "Normal Replicas with multiple clients",
'faultyReplicas': 1,
'timeoutDelta' : 2500,
'clients' : 2,
'requests' : 10,
'clientTimeout' : 5,
'testcase' : {
   "type" : "normal",
   "specialArguments": {}
}
},
```

This is similar to normal execution flow, instead, here there are two clients which are provided.

Client requests timeout, re-submission, and handling request de-duplication

```
'Name' : "Client small timeout with request resubmission and handling de-deuplication",
    'faultyReplicas': 1,
    'timeoutDelta' : 2500,
    'clients' : 1,
    'requests' : 10,
    'clientTimeout' : 0.5,
    'testcase' : {
        "type" : "normal",
        "specialArguments": {}
    }
},
```

In this test case we set the client timeout to be very low i.e 500ms due to which, the client keeps resending the request on timeout. On receiving a duplicate client request the replicas should handle them and only commit the unique ones.

Omission Failure

In the following test case the faulty replicas cause omission failures but not providing a Vote Message. This indicates the tolerance of DiemBFT to delays in the network.

Forge Signature

In the following test case, the faulty validator tries to forge the QC signature. Which causes a timeout and later leads the chain to recover. In our case, due to the design of mempool, new requests sent by clients do not get added to mempool as they were trying to be processed earlier. This is further discussed in bugs and limitations. (Note: the process does not terminate by itself as the client has not received all acks for all requests which were requested). In the figure, the client does try to re-transmit request 1, 2 but th mempool does not insert them.

```
'Name' : "Forge signature",
'faultyReplicas': 1,
'timeoutDelta' : 500,
'clients' : 1,
'requests' : 5,
'clientTimeout' : 2,
'testcase' : {
   "type" : "forge_signature",
   "specialArguments": {}
},
},
```

Delay Failures

In delay failure, we simulate network delay using a random seed to timeout a message which is to be sent. This has a similar issue to the former test case where it fails due to the design of mempool. (Note: the process does not terminate by itself as the client has not received all acks for all requests which were requested).