
MODULE *TunableMongoDB_Repl*

EXTENDS *Naturals, FiniteSets, Sequences, TLC*

constants and variables

CONSTANTS	<i>Client, Server,</i>	the set of clients and servers
	<i>Key, Value,</i>	the set of keys and values
	<i>Nil,</i>	model value, place holder
	<i>PtStop</i>	max physical time
VARIABLES	<i>Primary,</i>	Primary node
	<i>Secondary,</i>	secondary nodes
	<i>Oplog,</i>	<i>oplog[s]</i> : <i>oplog</i> at <i>server[s]</i>
	<i>Store,</i>	<i>store[s]</i> : data stored at <i>server[s]</i>
	<i>Ct,</i>	<i>Ct[s]</i> : cluster time at node <i>s</i>
	<i>Ot,</i>	<i>Ot[s]</i> : the last applied operation time at server <i>s</i>
	<i>ServerMsg,</i>	<i>ServerMsg[s]</i> : the channel of heartbeat <i>msgs</i> at server <i>s</i>
	<i>Pt,</i>	<i>Pt[s]</i> : physical time at server <i>s</i>
	<i>Cp,</i>	<i>Cp[s]</i> : majority commit point at server <i>s</i>
	<i>State,</i>	<i>State[s]</i> : the latest <i>Ot</i> of all servers that server <i>s</i> knows
	<i>CalState,</i>	<i>CalState</i> : sorted <i>State[Primary]</i>
	<i>CurrentTerm,</i>	<i>CurrentTerm[s]</i> : current election term at server <i>s</i> → updated in <i>update_position</i> , heartbeat and replicate
	<i>ReadyToServe,</i>	equal to 0 before any primary is elected
	<i>SyncSource</i>	<i>SyncSource[s]</i> : sync source of server node <i>s</i>

ASSUME *Cardinality(Client)* ≥ 1 at least one client

ASSUME *Cardinality(Server)* ≥ 2 at least one primary and one secondary

ASSUME *Cardinality(Key)* ≥ 1 at least one object

ASSUME *Cardinality(Value)* ≥ 2 at least two values to update

Helpers

$HLCLt(x, y) \triangleq$ IF $x.p < y.p$
 THEN TRUE
 ELSE IF $x.p = y.p$
 THEN IF $x.l < y.l$
 THEN TRUE
 ELSE FALSE
 ELSE FALSE

$HLCMin(x, y) \triangleq$ IF $HLCLt(x, y)$ THEN x ELSE y
 $HLCMax(x, y) \triangleq$ IF $HLCLt(x, y)$ THEN y ELSE x
 $HLCType \triangleq [p : Nat, l : Nat]$
 $Min(x, y) \triangleq$ IF $x < y$ THEN x ELSE y
 $Max(x, y) \triangleq$ IF $x > y$ THEN x ELSE y

$vars \triangleq \langle Primary, Secondary, Oplog, Store, Ct, Ot, ServerMsg, \\ Pt, Cp, CalState, State, \\ CurrentTerm, ReadyToServe, SyncSource \rangle$

RECURSIVE $CreateState(-, -)$ **init state**

$CreateState(len, seq) \triangleq$
 IF $len = 0$ THEN seq
 ELSE $CreateState(len - 1, Append(seq, [p \mapsto 0, l \mapsto 0]))$

$LogTerm(i, index) \triangleq$ IF $index = 0$ THEN 0 ELSE $Oplog[i][index].term$
 $LastTerm(i) \triangleq CurrentTerm[i]$

Is node i ahead of node j

$NotBehind(i, j) \triangleq \vee LastTerm(i) > LastTerm(j) \\ \vee \wedge LastTerm(i) = LastTerm(j) \\ \wedge Len(Oplog[i]) \geq Len(Oplog[j])$

$IsMajority(servers) \triangleq Cardinality(servers) * 2 > Cardinality(Server)$

Return the maximum value from a set, or undefined if the set is empty.

$MaxVal(s) \triangleq CHOOSE x \in s : \forall y \in s : x \geq y$

commit point

RECURSIVE $AddState(-, -, -)$

$AddState(new, state, index) \triangleq$
 IF $index = 1 \wedge HLCLt(new, state[1])$
 THEN $\langle new \rangle \circ state$ **less than the first**
 ELSE IF $index = Len(state) + 1$
 THEN $state \circ \langle new \rangle$
 ELSE IF $HLCLt(new, state[index])$
 THEN $SubSeq(state, 1, index - 1) \circ \langle new \rangle \circ SubSeq(state, index, Len(state))$
 ELSE $AddState(new, state, index + 1)$

RECURSIVE $RemoveState(-, -, -)$

$RemoveState(old, state, index) \triangleq$
 IF $state[index] = old$
 THEN $SubSeq(state, 1, index - 1) \circ SubSeq(state, index + 1, Len(state))$
 ELSE $RemoveState(old, state, index + 1)$

$AdvanceState(new, old, state) \triangleq AddState(new, RemoveState(old, state, 1), 1)$

clock

$MaxPt \triangleq LET x \triangleq CHOOSE s \in Server : \forall s1 \in Server \setminus \{s\} : \\ Pt[s] \geq Pt[s1] IN Pt[x]$

$Tick(s) \triangleq Ct' = IF Ct[s].p \geq Pt[s] THEN [Ct EXCEPT ![s] = [p \mapsto @.p, l \mapsto @.l + 1]] \\ ELSE [Ct EXCEPT ![s] = [p \mapsto Pt[s], l \mapsto 0]]$

heartbeat

Only *Primary* node sends heartbeat once advance pt

$BroadcastHeartbeat(s) \triangleq$
 LET $msg \triangleq [type \mapsto \text{"heartbeat"}, s \mapsto s, aot \mapsto Ot[s],$
 $ct \mapsto Ct[s], cp \mapsto Cp[s], term \mapsto CurrentTerm[s]]$
 IN $ServerMsg' = [x \in Server \mapsto \text{IF } x = s \text{ THEN } ServerMsg[x]$
 $\text{ELSE } Append(ServerMsg[x], msg)]$

Can node i sync from node j ?

$CanSyncFrom(i, j) \triangleq$
 $\wedge Len(Oplog[i]) < Len(Oplog[j])$
 $\wedge LastTerm(i) = LogTerm(j, Len(Oplog[i]))$

$Oplog$ entries needed to replicate from j to i

$ReplicateOplog(i, j) \triangleq$
 LET $len_i \triangleq Len(Oplog[i])$
 $len_j \triangleq Len(Oplog[j])$
 IN IF $i \neq Primary \wedge len_i < len_j$
 THEN $SubSeq(Oplog[j], len_i + 1, len_j)$
 ELSE $\langle \rangle$

Can node i rollback its log based on j 's log

$CanRollback(i, j) \triangleq$
 $\wedge Len(Oplog[i]) > 0$
 $\wedge Len(Oplog[j]) > 0$
 $\wedge CurrentTerm[i] < CurrentTerm[j]$
 \wedge
 $\vee Len(Oplog[i]) > Len(Oplog[j])$
 $\vee \wedge Len(Oplog[i]) \leq Len(Oplog[j])$
 $\wedge CurrentTerm[i] \neq LogTerm(j, Len(Oplog[i]))$

Returns the highest common index between two divergent logs.

If there is no common index between the logs, returns 0.

$RollbackCommonPoint(i, j) \triangleq$
 LET $commonIndices \triangleq \{k \in \text{DOMAIN } Oplog[i] :$
 $\wedge k \leq Len(Oplog[j])$
 $\wedge Oplog[i][k] = Oplog[j][k]\}$ IN
 IF $commonIndices = \{\}$ THEN 0 ELSE $MaxVal(commonIndices)$

Init Part

$InitPrimary \triangleq Primary = \text{CHOOSE } s \in Server : \text{TRUE}$
 $InitSecondary \triangleq Secondary = Server \setminus \{Primary\}$
 $InitOplog \triangleq Oplog = [s \in Server \mapsto \langle \rangle]$
 $InitStore \triangleq Store = [n \in Server \cup Client \mapsto [k \in Key \mapsto Nil]]$
 $InitCt \triangleq Ct = [n \in Server \cup Client \mapsto [p \mapsto 0, l \mapsto 0]]$
 $InitOt \triangleq Ot = [n \in Server \cup Client \mapsto [p \mapsto 0, l \mapsto 0]]$

$$\begin{aligned}
InitServerMsg &\triangleq ServerMsg = [s \in Server \mapsto \langle \rangle] \\
InitPt &\triangleq Pt = [s \in Server \mapsto 1] \\
InitCp &\triangleq Cp = [n \in Server \cup Client \mapsto [p \mapsto 0, l \mapsto 0]] \\
InitCalState &\triangleq CalState = CreateState(Cardinality(Server), \langle \rangle) \\
&\quad \text{create initial state(for calculate)} \\
InitState &\triangleq State = [s \in Server \mapsto [s0 \in Server \mapsto \\
&\quad [p \mapsto 0, l \mapsto 0]]] \\
InitCurrentTerm &\triangleq CurrentTerm = [s \in Server \mapsto 0] \\
InitReadyToServe &\triangleq ReadyToServe = 0 \\
InitSyncSource &\triangleq SyncSource = [s \in Server \mapsto Nil] \\
Init &\triangleq \\
&\quad \wedge InitPrimary \wedge InitSecondary \wedge InitOplog \wedge InitStore \wedge InitCt \\
&\quad \wedge InitOt \wedge InitPt \wedge InitCp \wedge InitCalState \\
&\quad \wedge InitServerMsg \\
&\quad \wedge InitState \wedge InitCurrentTerm \wedge InitReadyToServe \\
&\quad \wedge InitSyncSource
\end{aligned}$$

Next State Actions

Replication Protocol: possible actions

$$\begin{aligned}
TurnOnReadyToServe &\triangleq \\
&\quad \wedge ReadyToServe = 0 \\
&\quad \wedge \exists s \in Primary : \\
&\quad \quad \wedge CurrentTerm' = [CurrentTerm \text{ EXCEPT } ![s] = CurrentTerm[s] + 1] \\
&\quad \quad \wedge ReadyToServe' = ReadyToServe + 1 \\
&\quad \wedge \text{UNCHANGED } \langle Primary, Secondary, Oplog, Store, Ct, Ot, \\
&\quad \quad ServerMsg, Pt, Cp, \\
&\quad \quad State, CalState, SyncSource \rangle \\
Stepdown &\triangleq \\
&\quad \wedge ReadyToServe > 0 \\
&\quad \wedge \exists s \in Primary : \\
&\quad \quad \wedge Primary' = Primary \setminus \{s\} \\
&\quad \quad \wedge Secondary' = Secondary \cup \{s\} \\
&\quad \wedge \text{UNCHANGED } \langle Oplog, Store, Ct, Ot, ServerMsg, \\
&\quad \quad Pt, Cp, State, CalState, CurrentTerm, \\
&\quad \quad ReadyToServe, SyncSource \rangle
\end{aligned}$$

There are majority nodes agree to elect node i to become primary

$$\begin{aligned}
ElectPrimary &\triangleq \\
&\quad \wedge ReadyToServe > 0 \\
&\quad \wedge \exists i \in Server : \exists majorNodes \in \text{SUBSET}(Server) : \\
&\quad \quad \wedge \forall j \in majorNodes : \wedge NotBehind(i, j) \\
&\quad \quad \wedge CurrentTerm[i] \geq CurrentTerm[j]
\end{aligned}$$

$$\begin{aligned}
& \wedge \text{IsMajority}(\text{majorNodes}) \\
& \text{voted nodes for } i \text{ cannot be primary anymore} \\
& \wedge \text{Primary}' = \text{LET } \text{possiblePrimary} \triangleq \text{Primary} \setminus \text{majorNodes} \\
& \quad \text{IN } \text{possiblePrimary} \cup \{i\} \\
& \text{add voted nodes into secondaries} \\
& \wedge \text{Secondary}' = \text{LET } \text{possibleSecondary} \triangleq \text{Secondary} \cup \text{majorNodes} \\
& \quad \text{IN } \text{possibleSecondary} \setminus \{i\} \\
& \wedge \text{CurrentTerm}' = [\text{index} \in \text{Server} \mapsto \text{IF } \text{index} \in (\text{majorNodes} \cup \{i\}) \\
& \quad \quad \text{THEN } \text{CurrentTerm}[\text{i}] + 1 \\
& \quad \quad \text{ELSE } \text{CurrentTerm}[\text{index}]] \\
& \wedge \text{UNCHANGED } \langle \text{Oplog}, \text{Store}, \text{Ct}, \text{Ot}, \text{ServerMsg}, \text{Pt}, \text{Cp}, \text{State}, \text{CalState}, \\
& \quad \text{CurrentTerm}, \text{ReadyToServe}, \text{SyncSource} \rangle \\
\\
\text{AdvanceCp} & \triangleq \\
& \wedge \text{ReadyToServe} > 0 \\
& \wedge \text{Cp}' = [\text{Cp} \text{ EXCEPT } ![\text{Primary}] = \text{CalState}[\text{Cardinality}(\text{Server}) \div 2 + 1]] \\
& \wedge \text{UNCHANGED } \langle \text{Primary}, \text{Secondary}, \text{Oplog}, \text{Store}, \text{Ct}, \text{Ot}, \\
& \quad \text{ServerMsg}, \text{Pt}, \text{CalState}, \\
& \quad \text{State}, \text{CurrentTerm}, \text{ReadyToServe}, \text{SyncSource} \rangle \\
\\
\text{heartbeatoplogOtstore} & \triangleq \\
\text{ServerTakeHeartbeat} & \triangleq \\
& \wedge \text{ReadyToServe} > 0 \\
& \wedge \exists s \in \text{Server} : \\
& \quad \wedge \text{Len}(\text{ServerMsg}[s]) \neq 0 \quad \text{message channel is not empty} \\
& \quad \wedge \text{ServerMsg}[s].\text{type} = \text{"heartbeat"} \\
& \quad \wedge \text{Ct}' = [\text{Ct} \text{ EXCEPT } ![s] = \text{HLCMax}(\text{Ct}[s], \text{ServerMsg}[s][1].\text{ct})] \\
& \quad \wedge \text{State}' = \\
& \quad \quad \text{LET } \text{SubHbState} \triangleq \text{State}[s] \\
& \quad \quad \text{hb} \triangleq [\text{SubHbState} \text{ EXCEPT } ![\text{ServerMsg}[s][1].s] = \\
& \quad \quad \quad \text{ServerMsg}[s][1].\text{aot}] \\
& \quad \quad \text{IN } [\text{State} \text{ EXCEPT } ![s] = \text{hb}] \\
& \quad \wedge \text{CalState}' = \text{LET } \text{newcal} \triangleq \\
& \quad \quad \text{IF } s \in \text{Primary} \quad \text{primary node: update CalState} \\
& \quad \quad \quad \text{THEN } \text{AdvanceState}(\text{ServerMsg}[s][1].\text{aot}, \\
& \quad \quad \quad \quad \text{State}[s][\text{ServerMsg}[s][1].s], \text{CalState}) \\
& \quad \quad \quad \text{ELSE } \text{CalState} \\
& \quad \quad \text{IN } \text{newcal} \\
& \quad \wedge \text{Cp}' = \text{LET } \text{newcp} \triangleq \\
& \quad \quad \text{primary node: compute new mcp} \\
& \quad \quad \text{IF } s \in \text{Primary} \text{ THEN } \text{CalState}'[\text{Cardinality}(\text{Server}) \div 2 + 1] \\
& \quad \quad \text{secondary node: update mcp} \\
& \quad \quad \text{ELSE IF } \neg \text{HLCLt}(\text{ServerMsg}[s][1].\text{cp}, \text{Cp}[s]) \\
& \quad \quad \quad \wedge \neg \text{HLCLt}(\text{Ot}[s], \text{ServerMsg}[s][1].\text{cp}) \\
& \quad \quad \quad \text{THEN } \text{ServerMsg}[s][1].\text{cp}
\end{aligned}$$

$$\begin{aligned}
& \text{ELSE } Cp[s] \\
& \text{IN } [Cp \text{ EXCEPT } ![s] = newcp] \\
& \wedge ServerMsg' = [ServerMsg \text{ EXCEPT } ![s] = Tail(@)] \\
& \wedge CurrentTerm' = [CurrentTerm \text{ EXCEPT } ![s] = Max(CurrentTerm[s], ServerMsg[s][1].term)] \\
& \wedge \text{UNCHANGED } \langle Primary, Secondary, Oplog, Store, Ot, Pt, \\
& \quad ReadyToServe, SyncSource \rangle \\
ServerTakeUpdatePosition & \triangleq \\
& \wedge ReadyToServe > 0 \\
& \wedge \exists s \in Server : \\
& \quad \wedge Len(ServerMsg[s]) \neq 0 \quad \text{message channel is not empty} \\
& \quad \wedge ServerMsg[s].type = \text{"update_position"} \\
& \quad \wedge Ct' = [Ct \text{ EXCEPT } ![s] = HLCMax(Ct[s], ServerMsg[s][1].ct)] \quad \text{update } ct \text{ accordingly} \\
& \quad \wedge State' = \\
& \quad \quad \text{LET } SubHbState \triangleq State[s] \\
& \quad \quad \quad hb \triangleq [SubHbState \text{ EXCEPT } ![ServerMsg[s][1].s] = \\
& \quad \quad \quad \quad ServerMsg[s][1].aot] \\
& \quad \quad \text{IN } [State \text{ EXCEPT } ![s] = hb] \\
& \quad \wedge CalState' = \text{LET } newcal \triangleq \\
& \quad \quad \text{IF } s \in Primary \quad \text{primary node: update } CalState \\
& \quad \quad \quad \text{THEN } AdvanceState(ServerMsg[s][1].aot, \\
& \quad \quad \quad \quad State[s][ServerMsg[s][1].s], CalState) \\
& \quad \quad \quad \text{ELSE } CalState \text{ IN } newcal \\
& \quad \wedge Cp' = \text{LET } newcp \triangleq \\
& \quad \quad \text{primary node: compute new mcp} \\
& \quad \quad \text{IF } s \in Primary \text{ THEN } CalState'[Cardinality(Server) \div 2 + 1] \\
& \quad \quad \quad \text{secondary node: update mcp} \\
& \quad \quad \text{ELSE IF } \neg HLClt(ServerMsg[s][1].cp, Cp[s]) \\
& \quad \quad \quad \quad \wedge \neg HLClt(Ot[s], ServerMsg[s][1].cp) \\
& \quad \quad \quad \text{THEN } ServerMsg[s][1].cp \\
& \quad \quad \quad \text{ELSE } Cp[s] \\
& \quad \quad \text{IN } [Cp \text{ EXCEPT } ![s] = newcp] \\
& \quad \wedge ServerMsg' = \text{LET } newServerMsg \triangleq [ServerMsg \text{ EXCEPT } ![s] = Tail(@)] \\
& \quad \quad \quad appendMsg \triangleq [type \mapsto \text{"update_position"}, s \mapsto ServerMsg[s][1].s, aot \mapsto ServerMs \\
& \quad \quad \quad \quad ct \mapsto ServerMsg[s][1].ct, cp \mapsto ServerMsg[s][1].cp, term \mapsto Server \\
& \quad \quad \quad newMsg \triangleq \text{IF } s \in Primary \\
& \quad \quad \quad \quad \text{THEN } newServerMsg \quad \text{If } s \text{ is primary, accept the } msg, \text{ else forward it to its sy} \\
& \quad \quad \quad \quad \text{ELSE } [newServerMsg \text{ EXCEPT } ![SyncSource[s]] = Append(newServer \\
& \quad \quad \quad \quad \text{IN } newMsg \\
& \quad \quad \wedge CurrentTerm' = [CurrentTerm \text{ EXCEPT } ![s] = Max(CurrentTerm[s], ServerMsg[s][1].term)] \\
& \quad \wedge \text{UNCHANGED } \langle Primary, Secondary, Oplog, Store, Ot, \\
& \quad \quad Pt, ReadyToServe, SyncSource \rangle \\
NTPSync & \triangleq \quad \text{simplify } NTP \text{ protocol} \\
& \wedge ReadyToServe > 0
\end{aligned}$$

$\wedge Pt' = [s \in Server \mapsto MaxPt]$
 $\wedge \text{UNCHANGED } \langle Primary, Secondary, Oplog, Store, Ct, Ot, ServerMsg, Cp, CalState, State, CurrentTerm, ReadyToServe, SyncSource \rangle$

$AdvancePt \triangleq$

$\wedge ReadyToServe > 0$
 $\wedge \exists s \in Server :$
 $\quad \wedge s \in Primary$ for simplicity
 $\quad \wedge Pt[s] \leq PtStop$
 $\quad \wedge Pt' = [Pt \text{ EXCEPT } ![s] = @ + 1]$ advance physical time
 $\quad \wedge BroadcastHeartbeat(s)$ broadcast heartbeat periodically
 $\wedge \text{UNCHANGED } \langle Primary, Secondary, Oplog, Store, Ct, Ot, State, Cp, CalState, CurrentTerm, ReadyToServe, SyncSource \rangle$

Replicate oplog from node j to node i , and update related structures accordingly

$Replicate \triangleq$

$\wedge ReadyToServe > 0$
 $\wedge \exists i, j \in Server :$
 $\quad \wedge CanSyncFrom(i, j)$ i can sync from j only need not to rollback
 $\quad \wedge i \in Secondary$
 $\quad \wedge ReplicateOplog(i, j) \neq \langle \rangle$
 $\quad \wedge Oplog' = [Oplog \text{ EXCEPT } ![i] = @ \circ ReplicateOplog(i, j)]$
 $\quad \wedge Store' = [Store \text{ EXCEPT } ![i] = Store[j]]$
 $\quad \wedge Ct' = [Ct \text{ EXCEPT } ![i] = HLCMax(Ct[i], Ct[j])]$ update $Ct[i]$
 $\quad \wedge Ot' = [Ot \text{ EXCEPT } ![i] = HLCMax(Ot[i], Ot[j])]$ update $Ot[i]$
 $\quad \wedge Cp' = [Cp \text{ EXCEPT } ![i] = HLCMax(Cp[i], Cp[j])]$ update $Cp[i]$
 $\quad \wedge CurrentTerm' = [CurrentTerm \text{ EXCEPT } ![i] = Max(CurrentTerm[i], CurrentTerm[j])]$ update $CurrentTerm$
 $\quad \wedge State' =$
 $\quad \quad \text{LET } SubHbState \triangleq State[i]$
 $\quad \quad \quad hb \triangleq [SubHbState \text{ EXCEPT } ![j] = Ot[j]]$
 $\quad \quad \text{IN } [State \text{ EXCEPT } ![i] = hb]$ update j 's state i knows
 $\quad \wedge \text{LET } msg \triangleq [type \mapsto \text{"update_position"}, s \mapsto i, aot \mapsto Ot'[i], ct \mapsto Ct'[i], cp \mapsto Cp'[i], term \mapsto CurrentTerm[j]]$
 $\quad \quad \text{IN } ServerMsg' = [ServerMsg \text{ EXCEPT } ![j] = Append(ServerMsg[j], msg)]$
 $\quad \wedge SyncSource' = [SyncSource \text{ EXCEPT } ![i] = j]$
 $\quad \wedge \text{UNCHANGED } \langle Primary, Secondary, Pt, CalState, ReadyToServe \rangle$

Rollback i 's oplog and recover it to j 's state

Recover to j 's state immediately to prevent internal client request

$RollbackAndRecover \triangleq$

$\wedge ReadyToServe > 0$
 $\wedge \exists i, j \in Server :$
 $\quad \wedge i \in Secondary$
 $\quad \wedge CanRollback(i, j)$
 $\quad \wedge \text{LET } cmp \triangleq RollbackCommonPoint(i, j) \text{ IN}$

$$\begin{aligned}
& \text{LET } \text{commonLog} \triangleq \text{SubSeq}(\text{Oplog}[i], 1, \text{cmp}) \\
& \quad \text{appendLog} \triangleq \text{SubSeq}(\text{Oplog}[j], \text{cmp} + 1, \text{Len}(\text{Oplog}[j])) \\
& \text{IN } \text{Oplog}' = [\text{Oplog} \text{ EXCEPT } ![i] = \text{commonLog} \circ \text{appendLog}] \\
& \wedge \text{CurrentTerm}' = [\text{CurrentTerm} \text{ EXCEPT } ![i] = \text{Max}(\text{CurrentTerm}[i], \text{CurrentTerm}[j])] \quad \text{update CurrentTerm} \\
& \wedge \text{Store}' = [\text{Store} \text{ EXCEPT } ![i] = \text{Store}[j]] \\
& \wedge \text{Ct}' = [\text{Ct} \text{ EXCEPT } ![i] = \text{HLCMax}(\text{Ct}[i], \text{Ct}[j])] \quad \text{update Ct}[i] \\
& \wedge \text{Ot}' = [\text{Ot} \text{ EXCEPT } ![i] = \text{HLCMax}(\text{Ot}[i], \text{Ot}[j])] \quad \text{update Ot}[i] \\
& \wedge \text{Cp}' = [\text{Cp} \text{ EXCEPT } ![i] = \text{HLCMax}(\text{Cp}[i], \text{Cp}[j])] \quad \text{update Cp}[i] \\
& \wedge \text{State}' = \text{LET } \text{SubHbState} \triangleq \text{State}[i] \\
& \quad \text{hb} \triangleq [\text{SubHbState} \text{ EXCEPT } ![j] = \text{Ot}[j]] \\
& \quad \text{IN } [\text{State} \text{ EXCEPT } ![i] = \text{hb}] \quad \text{update j's state i knows} \\
& \wedge \text{LET } \text{msg} \triangleq [\text{type} \mapsto \text{"update_position"}, s \mapsto i, \text{aot} \mapsto \text{Ot}'[i], \text{ct} \mapsto \text{Ct}'[i], \text{cp} \mapsto \text{Cp}'[i]] \\
& \quad \text{IN } \text{ServerMsg}' = [\text{ServerMsg} \text{ EXCEPT } ![j] = \text{Append}(\text{ServerMsg}[j], \text{msg})] \\
& \wedge \text{SyncSource}' = [\text{SyncSource} \text{ EXCEPT } ![i] = j] \\
& \wedge \text{UNCHANGED } \langle \text{Primary}, \text{Secondary}, \text{Pt}, \text{CalState}, \\
& \quad \text{ReadyToServe} \rangle
\end{aligned}$$

$$\begin{aligned}
& \text{ClientRequest} \triangleq \\
& \wedge \text{ReadyToServe} > 0 \\
& \wedge \exists s \in \text{Server}, k \in \text{Key}, v \in \text{Value} : \\
& \quad \wedge s \in \text{Primary} \\
& \quad \wedge \text{Tick}(s) \\
& \quad \wedge \text{Ot}' = [\text{Ot} \text{ EXCEPT } ![s] = \text{Ct}'[s]] \\
& \quad \wedge \text{Store}' = [\text{Store} \text{ EXCEPT } ![s][k] = v] \\
& \quad \wedge \text{Oplog}' = \text{LET } \text{entry} \triangleq [k \mapsto k, v \mapsto v, \text{ot} \mapsto \text{Ot}'[s], \text{term} \mapsto \text{CurrentTerm}[s]] \\
& \quad \quad \text{newLog} \triangleq \text{Append}(\text{Oplog}[s], \text{entry}) \\
& \quad \quad \text{IN } [\text{Oplog} \text{ EXCEPT } ![s] = \text{newLog}] \\
& \quad \wedge \text{State}' = \text{LET } \text{SubHbState} \triangleq \text{State}[s] \\
& \quad \quad \text{hb} \triangleq [\text{SubHbState} \text{ EXCEPT } ![s] = \text{Ot}'[s]] \\
& \quad \quad \text{IN } [\text{State} \text{ EXCEPT } ![s] = \text{hb}] \\
& \quad \wedge \text{CalState}' = \text{AdvanceState}(\text{Ot}'[s], \text{Ot}[s], \text{CalState}) \\
& \quad \wedge \text{UNCHANGED } \langle \text{Primary}, \text{Secondary}, \text{ServerMsg}, \\
& \quad \quad \text{Pt}, \text{Cp}, \\
& \quad \quad \text{CurrentTerm}, \text{ReadyToServe}, \text{SyncSource} \rangle
\end{aligned}$$

$$\begin{aligned}
& \text{Next state for all configurations} \\
& \text{Next} \triangleq \vee \text{Replicate} \\
& \quad \vee \text{AdvancePt} \\
& \quad \vee \text{ServerTakeHeartbeat} \\
& \quad \vee \text{ServerTakeUpdatePosition} \\
& \quad \vee \text{Stepdown} \\
& \quad \vee \text{RollbackAndRecover} \\
& \quad \vee \text{TurnOnReadyToServe} \\
& \quad \vee \text{ElectPrimary} \\
& \quad \vee \text{ClientRequest}
\end{aligned}$$

$$\vee NTPSync$$

$$Spec \triangleq Init \wedge \Box[Next]_{vars}$$

Properties to check?

$$IsLogPrefix(i, j) \triangleq$$

$$\wedge Len(Oplog[i]) \leq Len(Oplog[j])$$

$$\wedge Oplog[i] = SubSeq(Oplog[j], 1, Len(Oplog[i]))$$

If two logs have the same last *log* entry term, then one is a prefix of the other (from Will)

$$LastTermsEquivalentImplyPrefixes \triangleq$$

$$\forall i, j \in Server :$$

$$LogTerm(i, Len(Oplog[i])) = LogTerm(j, Len(Oplog[j])) \Rightarrow$$

$$IsLogPrefix(i, j) \vee IsLogPrefix(j, i)$$

$$TermsMonotonic \triangleq$$

$$\Box[\forall s \in Server : CurrentTerm'[s] \geq CurrentTerm[s]]_{vars}$$

\ * Modification *History*
 \ * Last modified *Tue Apr 19 22:17:02 CST 2022* by *dh*
 \ * Created *Mon Apr 18 11:38:53 CST 2022* by *dh*