EXTENDS Naturals, FiniteSets, Sequences, TLC

The set of all ONOS nodes CONSTANTS Nodes

Stream states
CONSTANTS Open, Closed

 $\begin{array}{c} {\rm Master~arbitration~message~types} \\ {\rm CONSTANTS~} Master Arbitration Update \end{array}$

Write message types
CONSTANTS WriteRequest, WriteResponse

Response status constants CONSTANTS Ok, AlreadyExists, PermissionDenied

Empty value
CONSTANT Nil

The current state of mastership elections VARIABLES term, master, backups

The current mastership event queue for each node VARIABLE events

The current mastership state for each node VARIABLE masterships

Whether the node has received a MasterArbitrationUpdate indicating it is the current master VARIABLE isMaster

The state of all streams and their requests and responses VARIABLE streams, requests, responses

The current set of elections for the switch, the greatest of which is the current master VARIABLE elections

Counting variables used to enforce state constraints ${\tt VARIABLES}\ master ship Changes,\ stream Changes,\ message Count$

A sequence of successful writes to the switch used for model checking VARIABLE writes

Mastership/consensus related variables $mastershipVars \triangleq \langle term, master, backups, mastershipChanges \rangle$

```
Node related variables node Vars \triangleq \langle events, masterships, isMaster \rangle

Stream related variables stream Vars \triangleq \langle streams, stream Changes \rangle

Message related variables message Vars \triangleq \langle requests, responses, message Count \rangle

Device related variables device Vars \triangleq \langle elections, writes \rangle

A sequence of all variables vars \triangleq \langle mastership Vars, node Vars, stream Vars, message Vars, device Vars \rangle
```

Helpers

```
Returns a sequence with the head removed Pop(q) \stackrel{\triangle}{=} SubSeq(q, 2, Len(q))
```

Returns a sequences with the element at the given index removed $Drop(q, i) \stackrel{\Delta}{=} SubSeq(q, 1, i-1) \circ SubSeq(q, i+1, Len(q))$

Returns the set of values in f $Range(f) \stackrel{\triangle}{=} \{f[x] : x \in \text{DOMAIN } f\}$

Returns the maximum value from a set or undefined if the set is empty $Max(s) \stackrel{\Delta}{=} \text{CHOOSE } x \in s : \forall y \in s : x \geq y$

Messaging between the Nodes and the device are modelled on TCP. For each node, a request and response sequence provides ordered messaging between the two points. Requests and responses are always received from the head of the queue and are never duplicated or reordered, and request and response queues only last the lifetime of the stream. When a stream is closed, all that stream's requests and responses are lost.

This section models mastership arbitration on the controller side. Mastership election occurs in two disctinct types of state changes. One state change occurs to change the mastership in the consensus layer, and the other occurs when a node actually learns of the mastership change. Nodes will always learn of mastership changes in the order in which they occur, and nodes will always learn of a mastership change. This, of course, is not representative of practice but is sufficient for modelling the mastership election algorithm.

```
Adds a node to the mastership election
JoinMastershipElection(n) \stackrel{\Delta}{=}
    \land \lor \land master = Nil
           \wedge term' = term + 1
           \wedge master' = n
           \wedge backups' = \langle \rangle
           \land events' = [i \in Nodes \mapsto Append(events[i], [
                                                 term \mapsto term',
                                                 master \mapsto master',
                                                 backups \mapsto backups'[)]
        \lor \land master \neq Nil
           \land master \neq n
           \land n \notin Range(backups)
           \wedge backups' = Append(backups, n)
           \land events' = [i \in Nodes \mapsto Append(events[i], ]
                                                 term \mapsto term,
                                                 master \mapsto master,
                                                 backups \mapsto backups'[)]
           \land UNCHANGED \langle term, master \rangle
    \land mastershipChanges' = mastershipChanges + 1
    ∧ UNCHANGED ⟨masterships, isMaster, stream Vars, message Vars, device Vars⟩
 Removes a node from the mastership election
LeaveMastershipElection(n) \triangleq
    \land \lor \land master = n
           \land \lor \land Len(backups) > 0
```

```
\wedge term' = term + 1
                 \land master' = backups[1]
                 \land \ backups' = Pop(backups)
                 \land events' = [i \in Nodes \mapsto Append(events[i], [
                                                         term \mapsto term',
                                                         master \mapsto master',
                                                         backups \mapsto backups'[)]
              \lor \land Len(backups) = 0
                 \wedge master' = Nil
                 \land UNCHANGED \langle term, backups, events \rangle
        \lor \land n \in Range(backups)
           \land backups' = Drop(backups, CHOOSE j \in DOMAIN backups : backups[j] = n)
           \land UNCHANGED \langle term, master, events \rangle
    \land mastershipChanges' = mastershipChanges + 1
    \land UNCHANGED \langle masterships, isMaster, stream Vars, message Vars, device Vars <math>\rangle
Sets the current master to node 'n' if it's not already set
SetMastership(n) \triangleq
    \lor \land master = n
        \land UNCHANGED \langle mastership Vars \rangle
    \lor \land master \neq n
        \wedge term' = term + 1
        \wedge master' = n
        \land \lor \land n \in Range(backups)
              \land backups' = Drop(backups, CHOOSE \ j \in DOMAIN \ backups : backups[j] = n)
           \vee \wedge n \notin Range(backups)
              \land UNCHANGED \langle backups \rangle
        \land mastershipChanges' = mastershipChanges + 1
 Receives a mastership change event from the consensus layer on node 'n'
LearnMastership(n) \stackrel{\triangle}{=}
    \wedge Len(events[n]) > 0
    \wedge \text{ LET } e \stackrel{\triangle}{=} events[n][1]
             m \triangleq masterships[n]
       IN
            \lor \land e.term > m.term
               \land masterships' = [masterships \ EXCEPT \ ! [n] = [
                                           term \mapsto e.term,
                                           master \mapsto e.master,
                                           backups \mapsto e.backups,
                                                   \mapsto \text{FALSE}]]
            \lor \land e.term = m.term
               \land masterships' = [masterships \ EXCEPT \ ! [n] = [
                                           term
                                                     \mapsto e.term,
                                           master \mapsto e.master,
```

```
backups \mapsto e.backups,
                                                                                           sent \mapsto m.sent]
          \land events' = [events \ EXCEPT \ ![n] = Pop(events[n])]
          \land UNCHANGED \langle mastership Vars, is Master, stream Vars, message Vars, device Vars <math>\rangle
  Notifies the device of node 'n' mastership info if it hasn't already been sent
SendMasterArbitrationUpdateRequest(n) \stackrel{\Delta}{=}
          \land streams[n] = Open
          \wedge LET m \stackrel{\Delta}{=} masterships[n]
                           \land m.term > 0
                           \land \neg m.sent
                           \land \lor \land m.master = n
                                        \land SendRequest(n, [
                                                                                    \mapsto MasterArbitrationUpdate,
                                                       election\_id \mapsto m.term + Cardinality(Nodes),
                                                       term
                                                                                    \mapsto m.term)
                                 \lor \land m.master \neq n
                                        \land n \in Range(m.backups)
                                        \land SendRequest(n, [
                                                                                     \mapsto MasterArbitrationUpdate,
                                                       election\_id \mapsto m.term + Cardinality(Nodes) - CHOOSE \ i \in DOMAIN \ m.backups : m.backups 
                                                                                    \mapsto m.term
          \land masterships' = [masterships \ EXCEPT \ ![n].sent = TRUE]
          \land UNCHANGED \langle mastership Vars, events, is Master, device Vars, stream Vars, responses <math>\rangle
  Receives a master arbitration update response on node 'n'
ReceiveMasterArbitrationUpdateResponse(n) \triangleq
          \land streams[n] = Open
          \land \textit{HasResponse}(n, \textit{MasterArbitrationUpdate})
          \wedge LET m \stackrel{\triangle}{=} NextResponse(n)
               IN
                           \vee \ \wedge m.status = \mathit{Ok}
                                 \land isMaster' = [isMaster \ EXCEPT \ ![n] = TRUE]
                                 \wedge SetMastership(n)
                           \lor \land m.status = AlreadyExists
                                 \land isMaster' = [isMaster \ EXCEPT \ ![n] = FALSE]
                                 \land UNCHANGED \langle mastership Vars \rangle
          \land DiscardResponse(n)
          ∧ UNCHANGED ⟨events, masterships, device Vars, stream Vars, requests, message Count⟩
 Sends a write request to the device from node 'n'
SendWriteRequest(n) \triangleq
         \land streams[n] = Open
 \land \text{LET } m \stackrel{\triangle}{=} masterships[n]
```

```
\land m.term > 0
            \land m.master = n
            \wedge isMaster[n]
            \land SendRequest(n, [
                   type
                                \mapsto WriteRequest,
                   election\_id \mapsto m.term + Cardinality(Nodes),
                               \mapsto m.term)
    ∧ UNCHANGED ⟨mastership Vars, node Vars, device Vars, stream Vars, responses⟩
 Receives a write response on node 'n'
ReceiveWriteResponse(n) \stackrel{\Delta}{=}
    \land streams[n] = Open
    \land HasResponse(n, WriteResponse)
    \wedge \text{ LET } m \stackrel{\triangle}{=} NextResponse(n)
      IN
            TODO: This should be used to determine whether writes from old masters are allowed
            \vee m.status = Ok
            \lor m.status = PermissionDenied
    \land DiscardResponse(n)
    \land UNCHANGED \land mastership Vars, node Vars, device Vars, stream Vars, requests, message Count\land
```

This section models the P4 switch. The switch side manages stream states between the device and the controller. Streams are opened and closed in a single state transition for the purposes of this model. Switches can handle two types of messages from the controller nodes: MasterArbitrationUpdate and Write.

```
Returns the highest election ID for the given elections ElectionId(e) \stackrel{\triangle}{=} Max(Range(e))
```

```
Returns the master for the given elections  \begin{aligned} Master(e) &\triangleq \\ \text{IF } Cardinality(\{i \in Range(e) : i > 0\}) > 0 \text{ THEN} \\ \text{CHOOSE } n \in \text{DOMAIN } e : e[n] = ElectionId(e) \\ \text{ELSE} \\ Nil \end{aligned}
```

Opens a new stream between node 'n' and the device

When a new stream is opened, the 'requests' and 'responses' queues for the node are cleared and the 'streams' state is set to 'Open'.

 $ConnectStream(n) \stackrel{\triangle}{=}$

```
 \begin{array}{l} \textit{InectStream}(n) = \\ \land \textit{streams}[n] = \textit{Closed} \\ \land \textit{streams'} = [\textit{streams} \ \texttt{EXCEPT} \ ![n] = \textit{Open}] \\ \land \textit{streamChanges'} = \textit{streamChanges} + 1 \\ \land \texttt{UNCHANGED} \ \langle \textit{mastershipVars}, \ \textit{nodeVars}, \ \textit{deviceVars}, \ \textit{messageVars} \rangle \\ \end{array}
```

Closes the open stream between node 'n' and the device

```
When the stream is closed, the 'requests' and 'responses' queues for the node are
 cleared and a 'MasterArbitrationUpdate' is sent to all remaining connected nodes
 to notify them of a mastership change if necessary.
CloseStream(n) \triangleq
    \land streams[n] = Open
    \land elections' = [elections \ EXCEPT \ ![n] = 0]
    \land streams' = [streams \ EXCEPT \ ![n] = Closed]
    \land requests' = [requests \ EXCEPT \ ![n] = \langle \rangle]
    \wedge LET oldMaster \triangleq Master(elections)
             newMaster \triangleq Master(elections')
       IN
            \lor \land oldMaster \neq newMaster
               \land responses' = [i \in DOMAIN \ streams' \mapsto
                                     IF streams'[i] = Open THEN
                                         If i = newMaster then
                                              Append(responses[i], [
                                                                \mapsto MasterArbitrationUpdate,
                                                   type
                                                                \mapsto Ok,
                                                   election\_id \mapsto ElectionId(elections')])
                                           ELSE
                                              Append(responses[i], [
                                                                \mapsto MasterArbitrationUpdate,
                                                   type
                                                                \mapsto AlreadyExists,
                                                   status
                                                   election\_id \mapsto ElectionId(elections')])
                                      ELSE
                                         \langle \rangle ]
               \land messageCount' = messageCount + 1
            \lor \land oldMaster = newMaster
               \land responses' = [responses \ EXCEPT \ ! [n] = \langle \rangle]
               \land UNCHANGED \langle messageCount \rangle
    \land streamChanges' = streamChanges + 1
    \land UNCHANGED \langle mastership Vars, node Vars, writes \rangle
 Handles a master arbitration update on the device
 If the election_id is already present in the 'elections', send an 'AlreadyExists'
 response to the node. Otherwise,
HandleMasterArbitrationUpdate(n) \triangleq
    \land streams[n] = Open
    \land HasRequest(n, MasterArbitrationUpdate)
    \wedge \text{ LET } m \stackrel{\triangle}{=} NextRequest(n)
            \lor \land m.election\_id \in Range(elections)
               \land elections[n] \neq m.election\_id
               \land streams' = [streams \ EXCEPT \ ![n] = Closed]
               \land requests' = [requests \ EXCEPT \ ![n] = \langle \rangle]
```

```
\land responses' = [responses \ EXCEPT \ ![n] = \langle \rangle]
                \land UNCHANGED \langle deviceVars, streamChanges, messageCount \rangle
             \lor \land m.election\_id \notin Range(elections)
                \land elections' = [elections \ EXCEPT \ ![n] = m.election\_id]
                 \land \text{ LET } oldMaster \stackrel{\triangle}{=} Master(elections) \\ newMaster \stackrel{\triangle}{=} Master(elections') 
                  IN
                        \lor \land oldMaster \neq newMaster
                           \land responses' = [i \in DOMAIN \ streams \mapsto
                                                  If streams[i] = Open then
                                                      If i = newMaster then
                                                           Append(responses[i], [
                                                                              \mapsto MasterArbitrationUpdate,
                                                                type
                                                                status
                                                                              \mapsto Ok,
                                                                election\_id \mapsto ElectionId(elections'))
                                                       ELSE
                                                           Append(responses[i], [
                                                                              \mapsto MasterArbitrationUpdate,
                                                                type
                                                                              \mapsto AlreadyExists,
                                                                status
                                                                election\_id \mapsto ElectionId(elections'))
                                                   ELSE
                                                       responses[i]]
                           \land messageCount' = messageCount + 1
                        \lor \land oldMaster = newMaster
                           \land SendResponse(n, [
                                   type
                                                 \mapsto MasterArbitrationUpdate,
                                                 \mapsto Ok,
                                   status
                                   election\_id \mapsto ElectionId(elections'))
                \land UNCHANGED \langle stream Vars \rangle
     \land DiscardRequest(n)
     \land UNCHANGED \langle mastership Vars, node Vars, writes \rangle
 Handles a write request on the device
HandleWrite(n) \triangleq
     \land streams[n] = Open
     \land HasRequest(n, WriteRequest)
     \wedge \text{ LET } m \stackrel{\triangle}{=} NextRequest(n)
             \lor \land elections[n] = m.election\_id
                \wedge Master(elections) = n
                \land writes' = Append(writes, [node \mapsto n, term \mapsto m.term])
                \land SendResponse(n,
                       type \mapsto WriteResponse,
                       status \mapsto Ok
             \lor \land \lor elections[n] \neq m.election\_id
```

```
\vee Master(elections) \neq n
                 \land SendResponse(n, [
                        type \mapsto WriteResponse,
                        status \mapsto PermissionDenied)
                 \land UNCHANGED \langle writes \rangle
     \land DiscardRequest(n)
     \land UNCHANGED \langle mastership Vars, node Vars, elections, stream Vars <math>\rangle
 The invariant asserts that no master can write to the switch after the switch
 has been notified of a newer master
TypeInvariant \stackrel{\triangle}{=} \forall i \in \text{DOMAIN} \ writes: i = 1 \lor writes[i-1].term \le writes[i].term
Init \triangleq
     \wedge term = 0
     \wedge master = Nil
     \land backups = \langle \rangle
     \land events = [n \in Nodes \mapsto \langle \rangle]
     \land masterships = [n \in Nodes \mapsto [term \mapsto 0, master \mapsto 0, backups \mapsto \langle \rangle, sent \mapsto FALSE]]
     \land isMaster = [n \in Nodes \mapsto false]
     \land streams = [n \in Nodes \mapsto Closed]
     \land requests = [n \in Nodes \mapsto \langle \rangle]
     \land responses = [n \in Nodes \mapsto \langle \rangle]
     \land elections = [n \in Nodes \mapsto 0]
     \land mastershipChanges = 0
     \wedge streamChanges = 0
     \land \ messageCount = 0
     \land writes = \langle \rangle
Next \triangleq
     \vee \exists n \in Nodes : ConnectStream(n)
     \vee \exists n \in Nodes : CloseStream(n)
     \vee \exists n \in Nodes : JoinMastershipElection(n)
     \lor \exists n \in Nodes : LeaveMastershipElection(n)
     \vee \exists n \in Nodes : LearnMastership(n)
     \vee \exists n \in Nodes : SendMasterArbitrationUpdateRequest(n)
     \vee \exists n \in Nodes : HandleMasterArbitrationUpdate(n)
     \vee \exists n \in Nodes : ReceiveMasterArbitrationUpdateResponse(n)
     \vee \exists n \in Nodes : SendWriteRequest(n)
     \vee \exists n \in Nodes : HandleWrite(n)
     \vee \exists n \in Nodes : ReceiveWriteResponse(n)
```

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