MODULE *Config*

EXTENDS *Naturals*, *FiniteSets*, *Sequences*, *TLC*

Indicates that a configuration change is waiting to be applied to the network
CONSTANT *Pending*

Indicates that a configuration change is being applied to the network
CONSTANT *Applying*

Indicates that a configuration change has been applied to the network
CONSTANT *Complete*

Indicates that a configuration change was successful
CONSTANT *Succeeded*

Indicates that a configuration change failed
CONSTANT *Failed*

Indicates a change is a configuration
CONSTANT *Change*

Indicates a change is a rollback
CONSTANT *Rollback*

Indicates a device is available
CONSTANT *Available*

Indicates a device is unavailable
CONSTANT *Unavailable*

Indicates that an error occurred when applying a change
CONSTANT *Error*

The set of all nodes
CONSTANT *Node*

The set of all devices
CONSTANT *Device*

An empty constant
CONSTANT *Nil*

Per-node election state
VARIABLE *leadership*

Per-node per-device election state
VARIABLE *mastership*

A sequence of network-wide configuration changes

Each change contains a record of 'changes' for each device
VARIABLE *networkChange*

A record of sequences of device configuration changes
Each sequence is a list of changes in the order in which they
are to be applied to the device
VARIABLE *deviceChange*

A record of sequences of pending configuration changes to each device.
VARIABLE *deviceQueue*

A record of device configurations derived from configuration chagnes pushed
to each device.
VARIABLE *deviceConfig*

A record of device states - either *Available* or *Unavailable*
VARIABLE *deviceState*

A count of leader changes to serve as a state constraint
VARIABLE *electionCount*

A count of configuration changes to serve as a state constraint
VARIABLE *configCount*

A count of device availability changes to serve as a state constraint
VARIABLE *availabilityCount*

Node variables
$nodeVars \triangleq \langle leadership, mastership \rangle$

Configuration variables
$configVars \triangleq \langle networkChange, deviceChange \rangle$

Device variables
$deviceVars \triangleq \langle deviceQueue, deviceConfig, deviceState \rangle$

State constraint variables
$constraintVars \triangleq \langle electionCount, configCount, availabilityCount \rangle$

$vars \triangleq \langle leadership, mastership, networkChange, deviceChange, deviceConfig \rangle$

The invariant asserts that any configuration applied to a device implies that all prior configurations of the same device have been applied to all associated devices.

$TypeInvariant \triangleq$
$\quad \land \forall\, d \in \text{DOMAIN } deviceConfig :$
$\quad\quad deviceConfig[d] \neq 0 \Rightarrow$

$$Cardinality(\text{UNION } \{\{y \in \text{DOMAIN } deviceChange[x] : \\ \land\ deviceChange[x][y].network < deviceConfig[x] \\ \land\ deviceChange[x][y].state \neq Complete\} : \\ x \in \text{DOMAIN } deviceChange\}) = 0$$

This section models leader election for control loops and for devices. Leader election is modelled as a simple boolean indicating whether each node is the leader for the cluster and for each device. This model implies the ordering of leadership changes is irrelevant to the correctness of the spec.

Set the leader for node n to l

$$SetNodeLeader(n,\ l) \triangleq \\ \land\ leadership' = [leadership \text{ EXCEPT } ![n] = n = l] \\ \land\ electionCount' = electionCount + 1 \\ \land\ \text{UNCHANGED } \langle mastership,\ configVars,\ deviceVars,\ configCount,\ availabilityCount\rangle$$

Set the master for device d on node n to l

$$SetDeviceMaster(n,\ d,\ l) \triangleq \\ \land\ mastership' = [mastership \text{ EXCEPT } ![n] = [mastership[n] \text{ EXCEPT } ![d] = n = l]] \\ \land\ electionCount' = electionCount + 1 \\ \land\ \text{UNCHANGED } \langle leadership,\ configVars,\ deviceVars,\ configCount,\ availabilityCount\rangle$$

This section models the northbound API for the configuration service.

Enqueue network configuration change c

$$SubmitChange(c) \triangleq \\ \land\ networkChange' = Append(networkChange, [\\ \qquad phase \mapsto Change, \\ \qquad changes \mapsto c, \\ \qquad value \mapsto Len(networkChange), \\ \qquad state \mapsto Pending, \\ \qquad result \mapsto Nil]) \\ \land\ configCount' = configCount + 1 \\ \land\ \text{UNCHANGED } \langle nodeVars,\ deviceChange,\ deviceVars,\ electionCount,\ availabilityCount\rangle$$

Roll back configuration change c

$$SubmitRollback(c) \triangleq \\ \land\ networkChange' = Append(networkChange, [\\ \qquad phase \mapsto Rollback, \\ \qquad changes \mapsto networkChange[c].changes, \\ \qquad value \mapsto c, \\ \qquad state \mapsto Pending, \\ \qquad result \mapsto Nil]) \\ \land\ configCount' = configCount + 1 \\ \land\ \text{UNCHANGED } \langle nodeVars,\ deviceChange,\ deviceVars,\ electionCount,\ availabilityCount\rangle$$

This section models a configuration change scheduler. The role of the scheduler is to determine when network changes can be applied and enqueue the relevant changes for application by changing their state from *Pending* to *Applying*. The scheduler supports concurrent application of non-overlapping configuration changes (changes that do not impact intersecting sets of devices) by comparing *Pending* changes with Applying changes.

Return the set of all network changes prior to the given change

$$PriorNetworkChanges(c) \triangleq \{n \in \text{DOMAIN } networkChange : n < c\}$$

Return the set of all completed device changes for network change c

$$\begin{aligned} NetworkCompletedChanges(c) \triangleq\ & \{d \in \text{DOMAIN } networkChange[c].changes : \\ & \quad \land Cardinality(\{x \in \text{DOMAIN } deviceChange[d] : deviceChange[d][x].network = c\}) \neq 0 \\ & \quad \land deviceChange[d][\text{CHOOSE } x \in \text{DOMAIN } deviceChange[d] \\ & \qquad : deviceChange[d][x].network = c].state = Complete\} \end{aligned}$$

Return a boolean indicating whether all device changes are complete for the given network change

$$NetworkChangesComplete(c) \triangleq Cardinality(NetworkCompletedChanges(c)) = Cardinality(\text{DOMAIN } networkChange[c].changes)$$

Return the set of all incomplete device changes prior to network change c

$$\begin{aligned} PriorIncompleteDevices(c) \triangleq\ & \text{UNION } \{\text{DOMAIN } networkChange[n].changes : \\ & \qquad n \in \{n \in PriorNetworkChanges(c) : \neg NetworkChangesComplete(n)\}\} \end{aligned}$$

Return the set of all devices configured by network change c

$$NetworkChangeDevices(c) \triangleq \text{DOMAIN } networkChange[c].changes$$

Return a boolean indicating whether network change c can be applied
A change can be applied if its devices do not intersect with past device changes that have not been applied

$$CanApply(c) \triangleq Cardinality(NetworkChangeDevices(c) \cap PriorIncompleteDevices(c)) = 0$$

Node n handles a network configuration change event c

$$\begin{aligned} NetworkSchedulerNetworkChange(n,\ c) \triangleq\ & \land leadership[n] = \text{TRUE} \\ & \land networkChange[c].state = Pending \\ & \land CanApply(c) \\ & \land networkChange' = [networkChange \text{ EXCEPT } ![c].state = Applying] \\ & \land \text{UNCHANGED } \langle nodeVars,\ deviceChange,\ deviceVars,\ constraintVars \rangle \end{aligned}$$

This section models the network-level change controller. The network control loop reacts to both network and device changes. The network controller runs on each node in the cluster, and the control loop can only be executed on a node that believes itself to be the leader. Note, however, that the model does not require a single leader.

When a network change is received:

- If the network change state is *Pending*
  - Add device changes for each configured device
- If the network change state is *Applying*
  - Update device change states to *Applying*

When a device change is received:
- If all device change states for the network are *Complete*
  - Mark the network change *Complete* with a *Succeeded* result if all device changes succeeded
  - Otherwise mark the network change *Complete* with a *Failed* result
- If all device changes failed due to *Unavailable* devices
  - Set the network change back to *Pending*
- If at least one device change succeeded but a subset failed
  - Fail the network change

Updates to network and device changes are atomic, and real-world implementations of the spec must provide for atomic updates for network and device changes as well. This can be done using either optimistic or pessimistic concurrency control.

Return a boolean indicating whether change c on device d already exists

$$HasDeviceChange(d,\ c) \triangleq$$
$$\quad Cardinality(\{x \in \text{DOMAIN } deviceChange[d] : deviceChange[d][x].network = c\}) \neq 0$$

Return the index of the device change for network change c

$$DeviceChange(d,\ c) \triangleq$$
$$\quad \text{CHOOSE } x \in \text{DOMAIN } deviceChange[d] : deviceChange[d][x].network = c$$

Return a boolean indicating whether the device change for network change c has state s

$$HasDeviceState(d,\ c,\ s) \triangleq$$
$$\quad HasDeviceChange(d,\ c) \land deviceChange[d][DeviceChange(d,\ c)].state = s$$

Add change c on device s

$$AddDeviceChange(d,\ c) \triangleq$$
$$\begin{array}{l}
\quad \text{IF } d \in \text{DOMAIN } networkChange[c].changes \land \neg HasDeviceChange(d,\ c) \text{ THEN} \\
\qquad Append(deviceChange[d],\ [\\
\qquad\quad network \mapsto c, \\
\qquad\quad phase \quad \mapsto networkChange[c].phase, \\
\qquad\quad state \quad \mapsto Pending, \\
\qquad\quad value \quad \mapsto networkChange[c].value, \\
\qquad\quad result \quad \mapsto Nil, \\
\qquad\quad reason \quad \mapsto Nil]) \\
\quad \text{ELSE} \\
\qquad deviceChange[d]
\end{array}$$

Change the state of change c on device s from *Pending* to *Applying*

$$\begin{array}{l}
ApplyDeviceChange(d,\ c) \triangleq \\
\quad \text{IF } d \in \text{DOMAIN } networkChange[c].changes \text{ THEN} \\
\qquad \text{IF } HasDeviceChange(d,\ c) \text{ THEN} \\
\qquad\quad \text{IF } HasDeviceState(d,\ c,\ Pending) \text{ THEN} \\
\qquad\qquad [deviceChange[d] \text{ EXCEPT } ![DeviceChange(d,\ c)].state = Applying] \\
\qquad\quad \text{ELSE}
\end{array}$$

```
                deviceChange[d]
        ELSE
            Append(deviceChange[d], [
                network ↦ c,
                phase   ↦ networkChange[c].phase,
                state   ↦ Applying,
                value   ↦ networkChange[c].value,
                result  ↦ Nil,
                reason  ↦ Nil])
    ELSE
        deviceChange[d]

 Change the state of change c on device s to Pending
PendDeviceChange(d, c) ≜
    IF d ∈ DOMAIN networkChange[c].changes THEN
        [deviceChange[d] EXCEPT ![DeviceChange(d, c)].state = Pending]
    ELSE
        deviceChange[d]

 Return the set of all device changes for network change c
DeviceChanges(c) ≜
    {deviceChange[d][DeviceChange(d, c)] :
        d ∈ {d ∈ DOMAIN networkChange[c].changes : HasDeviceChange(d, c)}}

 Return a boolean indicating whether all device changes for network change c are complete
DeviceChangesComplete(c) ≜
    Cardinality({x ∈ DeviceChanges(c) : x.state = Complete}) = Cardinality(DeviceChanges(c))

 Return a boolean indicating whether all device changes for network change c were successful
DeviceChangesSucceeded(c) ≜
    Cardinality({x ∈ DeviceChanges(c) : x.result = Succeeded}) = Cardinality(DeviceChanges(c))

 Return a boolean indicating whether all device changes failed due to Unavailable devices
DeviceChangesUnavailable(c) ≜
    Cardinality({x ∈ DeviceChanges(c) : x.result = Failed ∧ x.reason = Unavailable}) =
        Cardinality(DeviceChanges(c))

 Node n handles a network configuration change c
NetworkControllerNetworkChange(n, c) ≜
    ∧ leadership[n] = TRUE
    ∧ LET change ≜ networkChange[c]IN
        ∨ ∧ change.state = Pending
          ∧ Cardinality({d ∈ DOMAIN networkChange[c].changes :
                ¬HasDeviceState(d, c, Pending)}) > 0
          ∧ deviceChange′ = [d ∈ Device ↦ AddDeviceChange(d, c)]
        ∨ ∧ change.state = Applying
          ∧ Cardinality({d ∈ DOMAIN networkChange[c].changes :
```

$\neg HasDeviceState(d,\ c,\ Applying)\}) > 0$
$\land\ deviceChange' = [d \in Device \mapsto ApplyDeviceChange(d,\ c)]$
$\land$ UNCHANGED $\langle nodeVars,\ networkChange,\ deviceVars,\ constraintVars\rangle$

Node n handles a device configuration change c

$NetworkControllerDeviceChange(n,\ d,\ c) \triangleq$
$\land\ leadership[n] =$ TRUE
$\land$ LET $change \triangleq deviceChange[d][c]$
IN
$\land\ change.state = Complete$
$\land\ networkChange[change.network].state \neq Complete$
$\land\ \lor\ \land\ DeviceChangesUnavailable(change.network)$
$\land\ networkChange' = [networkChange$ EXCEPT $![change.network] = [$
$networkChange[change.network]$ EXCEPT
$!.state = Pending]]$
$\lor\ \land\ DeviceChangesComplete(change.network)$
$\land\ DeviceChangesSucceeded(change.network)$
$\land\ networkChange' = [networkChange$ EXCEPT $![change.network] = [$
$networkChange[change.network]$ EXCEPT
$!.state = Complete,\ !.result = Succeeded]]$
$\lor\ \land\ DeviceChangesComplete(change.network)$
$\land\ \neg DeviceChangesSucceeded(change.network)$
$\land\ networkChange' = [networkChange$ EXCEPT $![change.network] = [$
$networkChange[change.network]$ EXCEPT
$!.state = Complete,\ !.result = Failed]]$
$\land$ UNCHANGED $\langle nodeVars,\ deviceChange,\ deviceVars,\ constraintVars\rangle$

This section models the device-level change controller. The device control loop reacts to device changes and applies changes to devices. The device controller runs on each node in the cluster. A master is elected for each device, and the control loop can only be executed on the master for the device targeted by a change. Note, however, that the model does not require a single master per device. Multiple masters may exist for a device without violating safety properties.

When a device change is received: - If the node believes itself to be the master for the device and the change state is *Applying*, apply the change
- Set the change state to *Complete*
- If the change was applied successfully, set the change result to *Succeeded*
- If the change failed, set the change result to *Failed*

Note: the above is modelled in two separate steps to allow the model checker to succeed and fail device changes.

Updates to network device changes are atomic, and real-world implementations of the spec must provide for atomic updates for network and device changes as well. This can be done using either optimistic or pessimistic concurrency control.

Node n handles a device configuration change event c

$DeviceControllerDeviceChange(n,\ d,\ c) \triangleq$
$\land\ mastership[n][d] =$ TRUE

```
    ∧ LET change ≜ deviceChange[d][c]
      IN
          ∧ change.state = Applying
          ∧ Cardinality({i ∈ DOMAIN deviceQueue[n][d] : deviceQueue[n][d][i] = c}) = 0
          ∧ ∨ ∧ change.phase = Change
              ∧ deviceQueue' = [deviceQueue EXCEPT ![n] = [
                                  deviceQueue[n] EXCEPT ![d] = Append(deviceQueue[n][d], c)]]
            ∨ ∧ change.phase = Rollback
              ∧ networkChange[deviceChange[change.value]].state = Complete
              ∧ networkChange[deviceChange[change.value]].result = Succeeded
              ∧ deviceQueue' = [deviceQueue EXCEPT ![n] = [
                                  deviceQueue[n] EXCEPT ![d] = Append(deviceQueue[n][d], c)]]
    ∧ UNCHANGED ⟨nodeVars, configVars, deviceConfig, deviceState, constraintVars⟩

 Return a sequence with the head removed
Pop(q) ≜ SubSeq(q, 2, Len(q))

 Mark change c on device d succeeded
SucceedChange(n, d) ≜
    ∧ Len(deviceQueue[n][d]) > 0
    ∧ deviceState[d] = Available
    ∧ deviceChange' = [deviceChange EXCEPT ![d] = [
                        deviceChange[d] EXCEPT ![deviceQueue[n][d][1]] = [
                            deviceChange[d][deviceQueue[n][d][1]] EXCEPT
                                !.state = Complete,
                                !.result = Succeeded]]]
    ∧ deviceConfig' = [deviceConfig EXCEPT ![d] =
                        deviceChange[d][deviceQueue[n][d][1]].value]
    ∧ deviceQueue' = [deviceQueue EXCEPT ![n] = [
                        deviceQueue[n] EXCEPT ![d] = Pop(deviceQueue[n][d])]]
    ∧ UNCHANGED ⟨nodeVars, networkChange, deviceState, constraintVars⟩

 Mark change c on device d failed
 A change can be failed if the device rejects the change or the device is not reachable.
FailChange(n, d) ≜
    ∧ Len(deviceQueue[n][d]) > 0
    ∧ deviceChange' = [deviceChange EXCEPT ![d] = [
                        deviceChange[d] EXCEPT ![deviceQueue[n][d][1]] = [
                            deviceChange[d][deviceQueue[n][d][1]] EXCEPT
                                !.state = Complete,
                                !.result = Failed]]]
    ∧ deviceQueue' = [deviceQueue EXCEPT ![n] = [
                        deviceQueue[n] EXCEPT ![d] = Pop(deviceQueue[n][d])]]
    ∧ UNCHANGED ⟨nodeVars, networkChange, deviceConfig, deviceState, constraintVars⟩
```

This section models device states. Devices begin in the *Unavailable* state and can only be configured while in the *Available* state.

Set device d state to *Available*

$$
\begin{array}{l}
ActivateDevice(d) \triangleq \\
\quad \land deviceState' = [deviceState \text{ EXCEPT } ![d] = Available] \\
\quad \land availabilityCount' = availabilityCount + 1 \\
\quad \land \text{UNCHANGED } \langle nodeVars,\ configVars,\ deviceQueue,\ deviceConfig,\ electionCount,\ configCount \rangle
\end{array}
$$

Set device d state to *Unavailable*

$$
\begin{array}{l}
DeactivateDevice(d) \triangleq \\
\quad \land deviceState' = [deviceState \text{ EXCEPT } ![d] = Unavailable] \\
\quad \land availabilityCount' = availabilityCount + 1 \\
\quad \land \text{UNCHANGED } \langle nodeVars,\ configVars,\ deviceQueue,\ deviceConfig,\ electionCount,\ configCount \rangle
\end{array}
$$

Init and next state predicates

$$
\begin{array}{l}
Init \triangleq \\
\quad \land leadership = [n \in Node \mapsto \text{FALSE}] \\
\quad \land mastership = [n \in Node \mapsto [d \in Device \mapsto \text{FALSE}]] \\
\quad \land networkChange = \langle\rangle \\
\quad \land deviceChange = [d \in Device \mapsto \langle\rangle] \\
\quad \land deviceQueue = [n \in Node \mapsto [d \in Device \mapsto \langle\rangle]] \\
\quad \land deviceConfig = [d \in Device \mapsto 0] \\
\quad \land deviceState = [d \in Device \mapsto Unavailable] \\
\quad \land electionCount = 0 \\
\quad \land configCount = 0 \\
\quad \land availabilityCount = 0
\end{array}
$$

$$
\begin{array}{l}
Next \triangleq \\
\quad \lor \exists\, d \in \text{SUBSET } Device : \\
\qquad SubmitChange([x \in d \mapsto 1]) \\
\quad \lor \exists\, c \in \text{DOMAIN } networkChange : \\
\qquad SubmitRollback(c) \\
\quad \lor \exists\, n \in Node : \\
\qquad \exists\, l \in Node : \\
\qquad\quad SetNodeLeader(n,\ l) \\
\quad \lor \exists\, n \in Node : \\
\qquad \exists\, d \in Device : \\
\qquad\quad \exists\, l \in Node : \\
\qquad\qquad SetDeviceMaster(n,\ d,\ l) \\
\quad \lor \exists\, n \in Node : \\
\qquad \exists\, c \in \text{DOMAIN } networkChange : \\
\qquad\quad NetworkSchedulerNetworkChange(n,\ c) \\
\quad \lor \exists\, n \in Node : \\
\qquad \exists\, c \in \text{DOMAIN } networkChange : \\
\qquad\quad NetworkControllerNetworkChange(n,\ c)
\end{array}
$$

$\lor \exists\, n \in Node :$
$\quad \exists\, d \in Device :$
$\quad\quad \exists\, c \in \text{DOMAIN } deviceChange[d] :$
$\quad\quad\quad NetworkControllerDeviceChange(n,\ d,\ c)$
$\lor \exists\, n \in Node :$
$\quad \exists\, d \in Device :$
$\quad\quad \exists\, c \in \text{DOMAIN } deviceChange[d] :$
$\quad\quad\quad DeviceControllerDeviceChange(n,\ d,\ c)$
$\lor \exists\, n \in Node :$
$\quad \exists\, d \in Device :$
$\quad\quad SucceedChange(n,\ d)$
$\lor \exists\, n \in Node :$
$\quad \exists\, d \in Device :$
$\quad\quad FailChange(n,\ d)$
$\lor \exists\, d \in Device :$
$\quad ActivateDevice(d)$
$\lor \exists\, d \in Device :$
$\quad DeactivateDevice(d)$

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

* Modification History
* Last modified *Thu Dec* 12 12:13:07 *PST* 2019 by *jordanhalterman*
* Created *Fri Sep* 27 13:14:24 *PDT* 2019 by *jordanhalterman*