

---

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*

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 states derived from configuration changes pushed  
to each device.  
VARIABLE *deviceState*

---

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

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

Device variables  
 $deviceVars \triangleq \langle deviceQueue, deviceState \rangle$   
 $vars \triangleq \langle leadership, mastership, networkChange, deviceChange, deviceState \rangle$

---

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.

$TypeInvariant \triangleq$   
 $\wedge \forall d \in \text{DOMAIN } deviceState :$   
 $deviceState[d] \neq Nil \Rightarrow$   
 $Cardinality(\text{UNION } \{ \{y \in \text{DOMAIN } deviceChange[x] :$   
 $\wedge deviceChange[x][y].network < deviceState[x].network$   
 $\wedge deviceChange[x][y].status \neq Complete\} :$   
 $x \in \text{DOMAIN } deviceChange\}) = 0$

---

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  $l$   
 $SetNodeLeader(n, l) \triangleq$   
 $\wedge leadership' = [leadership \text{ EXCEPT } ![n] = n = l]$   
 $\wedge \text{UNCHANGED } \langle mastership, configVars, deviceVars \rangle$

Set the master for device  $d$  on node  $n$  to  $l$   
 $SetDeviceMaster(n, d, l) \triangleq$   
 $\wedge mastership' = [mastership \text{ EXCEPT } ![n] = [mastership[n] \text{ EXCEPT } ![d] = n = l]]$   
 $\wedge \text{UNCHANGED } \langle leadership, configVars, deviceVars \rangle$

---

This section models the northbound *API* for the configuration service. The *API* exposes a single step to enqueue a configuration change. Rollback is not explicitly modelled as it can be implemented in an additional *Configure* step performing the inverse of the change being rolled back. When a configuration change is enqueued, it's simply added to network change for control loops to handle.

Enqueue network configuration change  $c$   
 $Configure(c) \triangleq$   
 $\wedge networkChange' = Append(networkChange, [changes \mapsto c, status \mapsto Pending])$   
 $\wedge \text{UNCHANGED } \langle nodeVars, deviceChange, deviceVars \rangle$

---

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 status 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 incomplete device changes prior to network change  $c$   
 $IncompleteChanges(c) \triangleq$

UNION  $\{\{d : i \in \{x \in \text{DOMAIN } deviceChange[d] :$   
 $\quad \wedge deviceChange[d][x].network < c$   
 $\quad \wedge deviceChange[d][x].status \neq Complete\} \} :$   
 $\quad d \in \text{DOMAIN } deviceChange\}$

Return the set of all devices configured by network change  $c$   
 $NetworkChangeDevices(c) \triangleq \text{DOMAIN } networkChange[c].changes$

Node  $n$  handles a network configuration change event  $c$   
 $NetworkSchedulerNetworkChange(n, c) \triangleq$   
 $\quad \wedge leadership[n] = \text{TRUE}$   
 $\quad \wedge \text{IF } Cardinality(NetworkChangeDevices(c) \cap IncompleteChanges(c)) = 0 \text{ THEN}$   
 $\quad \quad \wedge networkChange' = [networkChange \text{ EXCEPT } ![c].status = Applying]$   
 $\quad \text{ELSE}$   
 $\quad \quad \wedge \text{UNCHANGED } \langle networkChange \rangle$   
 $\quad \wedge \text{UNCHANGED } \langle nodeVars, deviceChange, deviceVars \rangle$

---

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 status is *Pending*, add device changes for each configured device
- If the network change status is *Applying*, update device change statuses to *Applying*

When a device change is received:

- If all device change statuses 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

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

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

Add change  $c$  on device  $s$  with status  $s$

$AddDeviceChange(d, c, s) \triangleq$   
 $\quad Append(deviceChange[d], [network \mapsto c, status \mapsto s, value \mapsto networkChange[c].changes[d]])$

Update change  $c$  on device  $d$  to status  $s$   
 $UpdateDeviceChange(d, c, s) \triangleq$   
 $[deviceChange[d] \text{ EXCEPT } ![CHOOSE x \in \text{DOMAIN } deviceChange[d] :$   
 $deviceChange[d][x].network = c].status = s]$

Set change  $c$  on device  $d$  to status  $s$   
 $SetDeviceChange(d, c, s) \triangleq$   
 IF  $d \in \text{DOMAIN } networkChange[c].changes$  THEN  
 IF  $HasDeviceChange(d, c)$  THEN  
 $UpdateDeviceChange(d, c, s)$   
 ELSE  
 $AddDeviceChange(d, c, s)$   
 ELSE  
 $deviceChange[d]$

Return the set of all device changes for network change  $c$   
 $DeviceChanges(c) \triangleq$   
 $\{d \in \text{DOMAIN } networkChange[c].changes :$   
 $\{x \in deviceChange[d] : deviceChange[d][x].network = c\}\}$

Return a boolean indicating whether all device changes for network change  $c$  are complete  
 $DeviceChangesComplete(c) \triangleq$   
 $Cardinality(\{x \in DeviceChanges(c) : x.status = Complete\}) = Cardinality(DeviceChanges(c))$

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

Node  $n$  handles a network configuration change  $c$   
 $NetworkControllerNetworkChange(n, c) \triangleq$   
 $\wedge leadership[n] = \text{TRUE}$   
 $\wedge \text{LET } change \triangleq networkChange[c] \text{ IN}$   
 $\vee \wedge change.status = Pending$   
 $\wedge deviceChange' = [d \in Device \mapsto SetDeviceChange(d, c, Pending)]$   
 $\vee \wedge change.status = Applying$   
 $\wedge deviceChange' = [d \in Device \mapsto SetDeviceChange(d, c, Applying)]$   
 $\vee \wedge change.status = Complete$   
 $\wedge \text{UNCHANGED } \langle deviceChange \rangle$   
 $\wedge \text{UNCHANGED } \langle nodeVars, networkChange, deviceVars \rangle$

Node  $n$  handles a device configuration change  $c$   
 $NetworkControllerDeviceChange(n, d, c) \triangleq$   
 $\wedge leadership[n] = \text{TRUE}$   
 $\wedge \text{LET } change \triangleq deviceChange[d][c]$   
 IN  
 $\vee \wedge change.status = Complete$   
 $\wedge \vee \wedge DeviceChangesComplete(change.network)$

$$\begin{aligned}
& \wedge \text{DeviceChangesSucceeded}(\text{change.network}) \\
& \wedge \text{networkChange}' = [\text{networkChange} \text{ EXCEPT } ![\text{change.network}] = [ \\
& \quad \text{networkChange}[\text{change.network}] \text{ EXCEPT} \\
& \quad \quad !.status = \text{Complete}, !.result = \text{Succeeded}]] \\
\vee & \wedge \text{DeviceChangesComplete}(\text{change.network}) \\
& \wedge \neg \text{DeviceChangesSucceeded}(\text{change.network}) \\
& \wedge \text{networkChange}' = [\text{networkChange} \text{ EXCEPT } ![\text{change.network}] = [ \\
& \quad \text{networkChange}[\text{change.network}] \text{ EXCEPT} \\
& \quad \quad !.status = \text{Complete}, !.result = \text{Failed}]] \\
\vee & \wedge \neg \text{DeviceChangesComplete}(\text{change.network}) \\
& \wedge \text{UNCHANGED } \langle \text{networkChange} \rangle \\
\vee & \wedge \text{change.status} \neq \text{Complete} \\
& \wedge \text{UNCHANGED } \langle \text{networkChange} \rangle \\
& \wedge \text{UNCHANGED } \langle \text{nodeVars}, \text{deviceChange}, \text{deviceVars} \rangle
\end{aligned}$$

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 status

- is *Applying*, apply the change
- Set the change status 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  $c$

$$\begin{aligned}
& \text{DeviceControllerDeviceChange}(n, d, c) \triangleq \\
& \quad \wedge \text{mastership}[n][d] = \text{TRUE} \\
& \quad \wedge \text{LET } \text{change} \triangleq \text{deviceChange}[d][c] \\
& \quad \text{IN} \\
& \quad \quad \vee \wedge \text{change.status} = \text{Applying} \\
& \quad \quad \quad \wedge \text{deviceQueue}' = [\text{deviceQueue} \text{ EXCEPT } ![d] = \text{Append}(\text{deviceQueue}[d], \text{deviceChange}[d][c])] \\
& \quad \quad \vee \wedge \text{change.status} \neq \text{Applying} \\
& \quad \quad \quad \wedge \text{UNCHANGED } \langle \text{deviceQueue} \rangle \\
& \quad \wedge \text{UNCHANGED } \langle \text{nodeVars}, \text{configVars}, \text{deviceState} \rangle
\end{aligned}$$

Return a sequence with the head removed

$$\text{Pop}(q) \triangleq \text{SubSeq}(q, 2, \text{Len}(q))$$

Mark change  $c$  on device  $d$  succeeded

$$\begin{aligned}
\text{SucceedChange}(d, c) &\triangleq \\
&\wedge \text{deviceChange}' = [\text{deviceChange} \text{ EXCEPT } ![d] = [\text{deviceChange}[d] \text{ EXCEPT } ![c] = [ \\
&\quad \text{deviceChange}[d][c] \text{ EXCEPT } !.status = \text{Complete}, !.result = \text{Succeeded}]]] \\
&\wedge \text{deviceState}' = [\text{deviceState} \text{ EXCEPT } ![d] = \text{deviceQueue}[d]] \\
&\wedge \text{deviceQueue}' = [\text{deviceQueue} \text{ EXCEPT } ![d] = \text{Pop}(\text{deviceQueue}[d])] \\
&\wedge \text{UNCHANGED } \langle \text{nodeVars}, \text{networkChange} \rangle
\end{aligned}$$

Mark change  $c$  on device  $d$  failed

$$\begin{aligned}
\text{FailChange}(d, c) &\triangleq \\
&\wedge \text{deviceChange}' = [\text{deviceChange} \text{ EXCEPT } ![d] = [\text{deviceChange}[d] \text{ EXCEPT } ![c] = [ \\
&\quad \text{deviceChange}[d][c] \text{ EXCEPT } !.status = \text{Complete}, !.result = \text{Failed}]]] \\
&\wedge \text{deviceQueue}' = [\text{deviceQueue} \text{ EXCEPT } ![d] = \text{Pop}(\text{deviceQueue}[d])] \\
&\wedge \text{UNCHANGED } \langle \text{nodeVars}, \text{networkChange}, \text{deviceState} \rangle
\end{aligned}$$

---

Init and next state predicates

$$\begin{aligned}
\text{Init} &\triangleq \\
&\wedge \text{leadership} = [n \in \text{Node} \mapsto \text{FALSE}] \\
&\wedge \text{mastership} = [n \in \text{Node} \mapsto [d \in \text{Device} \mapsto \text{FALSE}]] \\
&\wedge \text{networkChange} = \langle \rangle \\
&\wedge \text{deviceChange} = [d \in \text{Device} \mapsto \langle \rangle] \\
&\wedge \text{deviceQueue} = [d \in \text{Device} \mapsto \langle \rangle] \\
&\wedge \text{deviceState} = [d \in \text{Device} \mapsto \text{Nil}]
\end{aligned}$$

$$\begin{aligned}
\text{Next} &\triangleq \\
&\vee \exists d \in \text{SUBSET } \text{Device} : \\
&\quad \text{Configure}([x \in d \mapsto 1]) \\
&\vee \exists n \in \text{Node} : \\
&\quad \exists l \in \text{Node} : \\
&\quad \quad \text{SetNodeLeader}(n, l) \\
&\vee \exists n \in \text{Node} : \\
&\quad \exists d \in \text{Device} : \\
&\quad \exists l \in \text{Node} : \\
&\quad \quad \text{SetDeviceMaster}(n, d, l) \\
&\vee \exists n \in \text{Node} : \\
&\quad \exists c \in \text{DOMAIN } \text{networkChange} : \\
&\quad \quad \text{NetworkSchedulerNetworkChange}(n, c) \\
&\vee \exists n \in \text{Node} : \\
&\quad \exists c \in \text{DOMAIN } \text{networkChange} : \\
&\quad \quad \text{NetworkControllerNetworkChange}(n, c) \\
&\vee \exists n \in \text{Node} : \\
&\quad \exists d \in \text{Device} : \\
&\quad \exists c \in \text{DOMAIN } \text{deviceChange}[d] : \\
&\quad \quad \text{NetworkControllerDeviceChange}(n, d, c) \\
&\vee \exists n \in \text{Node} : \\
&\quad \exists d \in \text{Device} :
\end{aligned}$$

$$\begin{aligned}
& \exists c \in \text{DOMAIN } deviceChange[d] : \\
& \quad DeviceControllerDeviceChange(n, d, c) \\
\vee & \exists d \in Device : \\
& \quad \exists c \in \text{DOMAIN } deviceQueue[d] : \\
& \quad \quad SucceedChange(d, c) \\
\vee & \exists d \in Device : \\
& \quad \exists c \in \text{DOMAIN } deviceQueue[d] : \\
& \quad \quad FailChange(d, c) \\
Spec & \triangleq Init \wedge \Box[Next]_{vars}
\end{aligned}$$


---

\ \* Modification History  
\ \* Last modified Sun Sep 29 01:50:46 PDT 2019 by jordanhalterman  
\ \* Created Fri Sep 27 13:14:24 PDT 2019 by jordanhalterman