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 chagnes 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.

$$
\begin{aligned}
&TypeInvariant \triangleq \\
&\quad \land \forall\, d \in \text{DOMAIN } deviceState : \\
&\qquad deviceState[d] \neq Nil \Rightarrow Cardinality(\{x \in deviceChange : \\
&\qquad\qquad\qquad \{y \in deviceChange[d] : \\
&\qquad\qquad\qquad\quad \land\, y < deviceState[d].network \\
&\qquad\qquad\qquad\quad \land\, deviceChange[d][y].status \neq Complete\}\}) = 0
\end{aligned}
$$

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

$$
\begin{aligned}
&SetNodeLeader(n,\, l) \triangleq \\
&\quad \land\, leadership' = [leadership \text{ EXCEPT } ![n] = n = l] \\
&\quad \land \text{UNCHANGED } \langle mastership,\, configVars,\, deviceVars \rangle
\end{aligned}
$$

Set the master for device d on node n to l

$$
\begin{aligned}
&SetDeviceMaster(n,\, d,\, l) \triangleq \\
&\quad \land\, mastership' = [mastership \text{ EXCEPT } ![n] = [mastership[n] \text{ EXCEPT } ![d] = n = l]] \\
&\quad \land \text{UNCHANGED } \langle leadership,\, configVars,\, deviceVars \rangle
\end{aligned}
$$

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

$$
\begin{aligned}
&Configure(c) \triangleq \\
&\quad \land\, networkChange' = Append(networkChange, [changes \mapsto c,\, status \mapsto Pending]) \\
&\quad \land \text{UNCHANGED } \langle nodeVars,\, deviceChange,\, deviceVars \rangle
\end{aligned}
$$

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 device changes prior to the given change c

$$\begin{aligned}
&PrevChanges(c) \triangleq \\
&\quad \text{UNION } \{\{deviceChange[d][i] : \\
&\qquad\qquad i \in \{x \in \text{DOMAIN } deviceChange[d] : x < c\}\} : \\
&\qquad\qquad\quad d \in \text{DOMAIN } deviceChange\}
\end{aligned}$$

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

$$\begin{aligned}
&IncompleteChanges(c) \triangleq \\
&\quad \text{UNION } \{\{d : i \in \{x \in \text{DOMAIN } deviceChange[d] : x < c \land deviceChange[d][x].status \neq Complete\}\} : \\
&\qquad\qquad d \in \text{DOMAIN } deviceChange\}
\end{aligned}$$

Return the set of all devices configured by network change c

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

Node n handles a network configuration change event c

$$\begin{aligned}
&NetworkSchedulerNetworkChange(n,\ c) \triangleq \\
&\quad \land\ leadership[n] = \text{TRUE} \\
&\quad \land\ \text{IF } Cardinality(NetworkDevices(c) \cap IncompleteChanges(c)) = 0 \text{ THEN} \\
&\qquad \land\ networkChange' = [networkChange \text{ EXCEPT } ![c].status = Applying] \\
&\qquad \text{ELSE} \\
&\qquad \land\ \text{UNCHANGED } \langle networkChange \rangle \\
&\quad \land\ \text{UNCHANGED } \langle nodeVars,\ deviceChange,\ deviceVars \rangle
\end{aligned}$$

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 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 result of 'Succeeded' if all device changes succeeded, otherwise *Failed*

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 the given device change already exists

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

Add change c on device s with status s

```
AddDeviceChange(d, c, s) ≜
    Append(deviceChange[d], [network ↦ c, status ↦ s, value ↦ networkChange[c].changes[d]])

 Update change c on device d to status s
UpdateDeviceChange(d, c, s) ≜
    [deviceChange[d] EXCEPT ![CHOOSE x ∈ DOMAIN deviceChange[d] :
                    deviceChange[d][x].network = c].status = s]

 Set change c on device d to status s
SetDeviceChange(d, c, s) ≜
    IF d ∈ 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) ≜
    {d ∈ DOMAIN networkChange[c].changes :
        {x ∈ deviceChange[d] : deviceChange[d][x].network = c}}

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

 Node n handles a network configuration change c
NetworkControllerNetworkChange(n, c) ≜
    ∧ leadership[n] = TRUE
    ∧ LET change ≜ networkChange[c]IN
            ∨ ∧ change.status = Pending
              ∧ deviceChange′ = [d ∈ Device ↦ SetDeviceChange(d, c, Pending)]
            ∨ ∧ change.status = Applying
              ∧ deviceChange′ = [d ∈ Device ↦ SetDeviceChange(d, c, Applying)]
            ∨ ∧ change.status = Complete
              ∧ UNCHANGED ⟨deviceChange⟩
    ∧ UNCHANGED ⟨nodeVars, networkChange, deviceVars⟩

 Node n handles a device configuration change c
NetworkControllerDeviceChange(n, d, c) ≜
    ∧ leadership[n] = TRUE
    ∧ LET change ≜ deviceChange[d][c]
```

```
    IN
        ∨ ∧ change.status = Complete
          ∧ ∨ ∧ DeviceChangesComplete(change.network)
                ∧ DeviceChangesSucceeded(change.network)
                ∧ networkChange′ = [networkChange EXCEPT ![change.network] = [
                                       networkChange[change.network] EXCEPT
                                           !.status = Complete, !.result = Succeeded]]
            ∨ ∧ DeviceChangesComplete(change.network)
              ∧ ¬DeviceChangesSucceeded(change.network)
              ∧ networkChange′ = [networkChange EXCEPT ![change.network] = [
                                       networkChange[change.network] EXCEPT
                                           !.status = Complete, !.result = Failed]]
            ∨ ∧ ¬DeviceChangesComplete(change.network)
              ∧ UNCHANGED ⟨networkChange⟩
        ∨ ∧ change.status ≠ Complete
          ∧ UNCHANGED ⟨networkChange⟩
    ∧ UNCHANGED ⟨nodeVars, deviceChange, deviceVars⟩
```

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

```
DeviceControllerDeviceChange(n, d, c) ≜
    ∧ mastership[n][d] = TRUE
    ∧ LET change ≜ deviceChange[d][c]
      IN
          ∨ ∧ change.status = Applying
            ∧ deviceQueue′ = [deviceQueue EXCEPT ![d] = Append(deviceQueue[d], deviceChange[d][c])]
          ∨ ∧ change.status ≠ Applying
            ∧ UNCHANGED ⟨deviceQueue⟩
    ∧ UNCHANGED ⟨nodeVars, configVars, deviceState⟩
```

Return a sequence with the head removed

$Pop(q) \triangleq SubSeq(q, 2, Len(q))$

Mark change c on device d succeeded

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

Mark change c on device d failed

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

Init and next state predicates

$Init \triangleq$
$\quad \land leadership = [n \in Node \mapsto \text{FALSE}]$
$\quad \land mastership = [n \in Node \mapsto [d \in Device \mapsto \text{FALSE}]]$
$\quad \land networkChange = \langle\rangle$
$\quad \land deviceChange = [d \in Device \mapsto \langle\rangle]$
$\quad \land deviceQueue = [d \in Device \mapsto \langle\rangle]$
$\quad \land deviceState = [d \in Device \mapsto Nil]$

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

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

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

* Modification History
* Last modified Sun *Sep* 29 00:43:51 *PDT* 2019 by *jordanhalterman*
* Created *Fri Sep* 27 13:14:24 *PDT* 2019 by *jordanhalterman*