YANG models for ACTN TE Performance Monitoring Telemetry and Network Autonomics

draft-lee-teas-actn-pm-telemetry-autonomics-03

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Lee</td>
<td>Huawei</td>
</tr>
<tr>
<td>Dhruv Dhody</td>
<td></td>
</tr>
<tr>
<td>Satish K</td>
<td></td>
</tr>
<tr>
<td>Ricard Vilalta</td>
<td>CTTC</td>
</tr>
<tr>
<td>Daniel King</td>
<td>Lancaster University</td>
</tr>
<tr>
<td>Daniele Ceccarelli</td>
<td>Ericsson</td>
</tr>
</tbody>
</table>
Overview Re-cap

• YANG data models that describe
  • Key Performance Indicator (KPI) telemetry
  • Network autonomies for TE-tunnels and ACTN VNs.

• Requirement 7 [ACTN-Requirement] and [I-D.xu-actn-perf-dynamic-service-control-03] provide the operator’s requirements for:
  • Performance Monitoring
  • Dynamic control in ACTN – creation, modification, optimization etc.
  • Monitor Network Traffic, Detects traffic imbalance, Initiate optimization!
  • Measure customer SLA, take dynamic action to make sure you meet them at all times
  • Scalability of Performance data

• Support for
  • Performance telemetry data
  • Scaling Intent
Interactions

1. CNC configures the grouping-operation for delay and bandwidth utilization.
2. CNC subscribes to the VN or VN member level telemetry for delay and bandwidth utilization streaming data.

3. MDSC configures the grouping-operation for delay and bandwidth utilization.
4. MDSC subscribes to the Tunnel level telemetry for delay and bandwidth utilization on a periodic basis.

5. MDSC receives the Domain tunnel telemetry data:
   - Domain Tunnel Bandwidth Utilization as $X(i)$ (Minimum across domain TE-links);
   - Domain Tunnel Delay as $Y(i)$ (Maximum across domain TE-links).

6. CNC receives the VN telemetry or per VN member data:
   - VN Bandwidth Utilization as $X$ (Minimum across VN Members);
   - VN Delay as $Y$ (Maximum across VN-Members);
   - VN member B/W Utilization as $W$ (Minimum across Domain Tunnels and inter-domain links);
   - VN member Delay as $Z$ (Maximum across Domain Tunnels and Inter-domain links).

Grouping operation (e.g. MAX) is a way to tell how to consolidate the underlying telemetry information.
Yang Model Relationships

TE Tunnel → Augments → TE KPI Telemetry

- TE KPI Telemetry model provides the TE tunnel level performance monitoring.
- Augment the TE tunnel State with performance attributes
  - Use the notification subscription mechanism to subscribe to telemetry (YANG PUSH)
  - Scaling Intent configurations for auto scaling in/out based on the performance monitored attributes

ACTN VN → Augments → ACTN TE Telemetry

- ACTN TE KPI Telemetry model provides the VN level aggregated performance monitoring.
- Augment the VN state as well as individual VN-member state with performance attributes.
  - Use notification subscription (YANG PUSH)
  - Scaling Intent configurations at the VN level to reach to the monitored performance KPI
  - Allow configuration of aggregation mechanism from the lower level telemetry details (max, mean etc.)
  - From VN-Member to VN
  - From per-domain tunnel to E2E VN-Member

Enable auto-scaling by configuring the condition when to scale out or in automatically!
Updates for this version (03)

- Removed packet-loss related data from the models to make the draft/models technology-agonostic.
  - one-way-packet-loss
  - two-way-packet-loss
- Removed packet-loss from grouping operation for network scaling autonomies mechanism.
- Imported ietf-te-types and corrected *utilized-bandwidth* type to `te-types:te-bandwidth`.
- NMDA Compliancy Status
  - ietf-te-kpi-telemetry (as this module augments TE-Tunnel module, it depends on that)
  - ietf-actn-te-kpi-telemetry (NMDA complaint)
Models

module: ietf-te-kpi-telemetry
augment /actn-vn:actn/actn-vn/vn/actn-vn:vn-list:
  +++rw te-scaling-intent
  | | rw scale-in
  | | | rw scale-in-operation-type? uint32
  | | | | scaling-criteria-operation
  | | | rw threshold-time? uint32
  | | | rw scale-out-condition? [performance-type]
  | | | rw performance-type identityref
  | | | rw performance-data? binary
  | rw scale-down
  | | rw cool-down-time? uint32
  | rw scale-out-operation-type? uint32
  | | scaling-criteria-operation
  | | rw scale-out-condition? [performance-type]
  | | rw performance-type identityref
  | | rw performance-data? binary
augment /te:te/te:tunnels/te:tunnel/te:config:
  +++rw te-scaling-intent
  | rw data
  | | ro one-way-delay? uint32
  | | ro two-way-delay? uint32
  | | ro one-way-delay-min? uint32
  | | ro one-way-delay-max? uint32
  | | ro two-way-delay-min? uint32
  | | ro two-way-delay-max? uint32
  | | ro one-way-delay-variation? uint32
  | | ro two-way-delay-variation? uint32
  | | ro utilized-bandwidth? te-types:te-bandwidth
Next Steps

• Continue to enhance the model...& Comments are welcome!

• Is this work a good base for WG adoption?
THANK YOU!