Why NRI and how?

• Network-resource isolation (NRI) is necessary for network virtualization.
  – It is important to avoid resource interference between slices (virtual networks) so that a single slice may not disrupt the whole infrastructure.

• Two methods of NRI based on shaping and policing (QoS mechanisms) have been developed.
  – Per-slice shaping (PSS)
  – Per-link policing (PLP)
2 Network Virtualization Platform and VNode

- **Virtualization node project (VNP)**
  - is a project in Japan aims to build virtualization-platform technology and a high-performance virtualization testbed.

- **Virtualization platform in VNP**
  - consists of a domain controller (DC), VNodes, and Gateways.

- **VNode (virtualization node): physical node**
  - forwards packets on the platform. Each packet on the platform contains a virtualized packet on a slice.
  - are connected by tunnels using a protocol such as GRE.

- **VNode consists of**
  - **Programmer**
    - is a programmable component that processes packets on the slices.
  - **Redirector**
    - is a component that can forward or route packets on the platform.
    - forward (redirect) packets from another VNode to a programmer and forward packets from a programmer to another VNode.
  - **VNode Manager**
    - a software component that manages the VNode.
Internal Structure of Redirector in VNode

- Redirector contains a high-end switch (or router) and a packet encoder/decoder (such as a GRE encoder/decoder).
- Packet encoder/decoder exists before/after the programmer.
  - Decoder converts the VNode-external data format to the internal format.
  - Encoder converts vice versa.
- High-end switch has **policers** and **shapers** that can be used for implementing NRI.
  - Redirectors play the most important role in NRI.
Slices (Virtual Networks) in VNP

- **Two types of slice components**
  - **Node Sliver** (or virtual node)
    - represents computational resources that exist in a VNode.
    - is used for control or protocol processing with an arbitrary packet format.
    - is generated by slicing physical computational resources.
  - **Link Sliver** (or virtual link)
    - represents resources of a virtual link that connects two node slivers.
    - is generated by slicing physical network resources such as bandwidth.

- **Slice definition**
  - is written by a (human) slice developer writes in XML.
  - is sent to DC, distributed to each VNode Manager, and sent to the programmer and the redirector.
Specifications for NRI

- Bandwidth (and the burst size) is specified in link slivers.

- **Example of link sliver specification:**
  
<table>
<thead>
<tr>
<th>port0</th>
<th>port1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth = 30 Mbps, Burst size = 10 kB</td>
<td></td>
</tr>
</tbody>
</table>

  ```xml
  <linkSliver type="link" subtype="GRE" name="LinkSliver1">
    <vports><vport name="port0" /><vport name="port1" /></vports>
    <resources>
      <resource key="bandwidth" value="30M" />
      <resource key="burstSize" value="10k" />
    </resources>
  </linkSliver>
  ```
Traffic control functions used for NRI

- **Shaping**
  - queues packets, and limits and schedules the egress traffic.
  - delays the packet, and drops it when the queue is filled.
  - is *more expensive and less scalable* than policing (i.e., requires more memory and scheduling overhead).

- **Policing**
  - measures network traffic without accumulating packets and drops packets when the bandwidth (or the burst size) exceeds a limit.
  - can be used for guaranteeing bandwidth of link slivers that shares a queue (i.e., divides bandwidth reserved for a queue to slices).
  - is *less expensive and more scalable* than shaping.
Methods of NRI

• PSS (Per-slice shaping)
  – isolates slices by shaping traffic per-slice instead of per-link-sliver (i.e., per-link shaping).
  – does not drop packets (does not use policing).
  – is **sufficient for NRI** between slices but does not guarantee per-link bandwidth.
  – is **more scalable than per-link shaping** (because using 80–90% less queues).
• **PLP (Per-link policing)**
  - isolates slices by policing traffic per link-sliver.
  - guarantees per-link bandwidth by measuring and dropping packets per link-sliver.
  - uses shaping per slice-class (by per-class shaping).
  - is **more scalable than per-link shaping** (is applicable to tens or hundreds of slices).
  - may be influenced more by other slices than PSS (may be worse in delay and jitter).
• **Combined method (PSS with PLP)**
  – isolates slices by policing traffic per sliver and shaping traffic per slice.
  – is as strict as PSS in isolation from other slices (is good in delay and jitter).
  – guarantees per-link bandwidth.

**Combined method**

<table>
<thead>
<tr>
<th>Policing per link</th>
<th>Shaping per slice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slice s1 (Class c1)</td>
<td>Slice s1 (Class c1)</td>
</tr>
<tr>
<td>Slice s2 (Class c1)</td>
<td>Slice s2 (Class c1)</td>
</tr>
<tr>
<td>Slice s3 (Class c2)</td>
<td>Slice s3 (Class c2)</td>
</tr>
</tbody>
</table>
Implementation and Evaluation

- **Implementation using high-end L3 switches**
  - Three methods for NRI, i.e., PLP, PSS, and the combined method (PSS with PLP) have been implemented.

- **Evaluation of slow-path and fast-path node slivers**
  - Method: Three slices are used: one for foreground traffic to be measured and two for background cross traffic.
  - Result: Slow-path (Linux VM) node slivers
  - Slices can be isolated when the foreground traffic is 4.0 Gbps or less. (The link bandwidth is 10 Gbps.)

<table>
<thead>
<tr>
<th>Isolation type</th>
<th>Delay (mS)</th>
<th>Jitter (mS)</th>
<th>Drop ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Std dev</td>
<td>Average</td>
</tr>
<tr>
<td>PLP</td>
<td>1.60</td>
<td>0.12</td>
<td>0.10</td>
</tr>
<tr>
<td>PSS</td>
<td>1.30</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>Combined</td>
<td>1.33</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>No isolation</td>
<td>12.08</td>
<td>4.28</td>
<td>0.12</td>
</tr>
<tr>
<td>(Congestion-less)</td>
<td>1.31</td>
<td>0.15</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Conditions: Link sliver bandwidth = 100 Mbps, traffic = 90 Mbps. Cross traffic fills the link.

- Result: Fast-path node sliver (using a network processor)
  - Slices can be isolated when the foreground traffic is 4.0 Gbps or less. (The link bandwidth is 10 Gbps.)
Conclusion

• Two methods of NRI for virtualization networks are proposed.
  – PSS enables NRI with 80–90% less queues compared to the per-link shaping.
  – PLP enables less strict isolation between tens or hundreds of slices using only one queue.

• Evaluations: PSS performs slightly better in terms of delay and packet-drop ratio.

• Applications of PSS and PLP:
  – PSS and the combined method are effective for delay-sensitive services.
  – PLP may be sufficiently used for the other types of services.
Network-resource Isolation for Virtualization Nodes

Yasusi Kanada, Hitachi, Ltd.
Kei Shiraishi, Hitachi, Ltd.
Akihiro Nakao, University of Tokyo

Introduction

What are VNP, VNodes, and Slices?

Methods of NRI