MidoNet Scalability
Report
Abstract:

This whitepaper describes the testing of MidoNet 5.0 in a Rackspace/Intel environment. We focused on validating the availability, scalability, and agility of MidoNet as a network virtualization overlay solution.
The Challenges

Network and security administrators spend a considerable amount of time implementing policies for their network and also debugging issues pertaining to their network. The recent trend indicates that the team’s focus has now shifted to making strategic improvements to deploy the companies workloads, on-premise, public or hybrid Clouds.

Many organizations aspire to emulate Google’s approach to infrastructure. Many organizations are adopting the “Google’s Infrastructure for Everyone Else” or GIFEE when building out new hyper scale data centers. This means essentially deploying application and analytic workloads on systems built to run at Internet scale.

In a recent ESG survey, respondents were asked to identify their top IT priorities for 2016. The top cited priorities include business intelligence and data analytics (cited by 23% of the respondents), data growth (22%) and integration (21%), improving backup and recovery (18%).

Transformative application and data analytic workloads require significant agility and place incredible demands on the data center network. These types of workloads pose challenges to using traditional networking devices and tools that were designed for physical, relatively static network infrastructures. Traditional core switches and independent top of rack switches were well suited for traditional data center network topologies with lots of North-South traffic.

In the never-ending quest for operation efficiency and rapid delivery of services, a growing number of organizations are evaluating the benefits of software-defined networking (SDN) as they make their digital transformation. The digital enterprise brings with it the need for new Cloud network solutions that meet the requisite flexibility, scalability and performance. The answer is MidoNet.

The Solution:
Open Source MidoNet, MidoNet Enterprise MidoNet (MEM), MEM Insights

MidoNet is an open source network virtualization overlay software that enables operators to build, operate and manage virtual networks at scale. With its distributed architecture, operators can overlay MidoNet on top of their existing hardware and hypervisor software. This means a single network for any platform.
Midokura Enterprise MidoNet (MEM)

The commercial subscription or Midokura Enterprise MidoNet enables enterprises to leverage MidoNet to achieve network flexibility, faster service provisioning, and faster time to production for cloud or internal business solutions. While the GUI-based management features and big data analytics are part of the commercial version, the features enabling core network virtualization capabilities will always be part of the MidoNet.

Built on open source MidoNet, MEM adds value in the form of management and analytic tools along with enterprise-class support. MEM can be thought of as the finely tuned package of advanced software, support, and services that is designed for the modern enterprise and web services companies.

MEM Insights

MEM Insights made its debut in MEM 5.0. This innovative new technology offers end-to-end operational tools, including advanced analytics and dynamic visualization. OpenStack operators benefit from rich, dynamic data visualizations that visually communicate relevant, up-to-date details of the state of the virtual network for any operational facets of interest. The benefit is more simplicity, security and agility over networks.

How MidoNet Works

The overlay network model in MidoNet describes the state of the virtual network and its relationships to the virtual objects (routers, bridges, tunnel zones, tenants, BGP). MidoNet packet-processing workflow is handled by the MidoNet agent on each host with network simulation. It works by having the MidoNet agent on each host "simulate" the overlay topology by retrieving the virtual topology data from the network state database. The output of the simulation is installed as a flow rule into the Linux kernel of the host.

Network trace files logs for every packet and every event that occurred in the simulation and for analysis. Purpose-built, analytical tools like MEM Insights help operators visualize the flow through the virtual network devices traversed, examine the security groups applied to the virtual ports and show the flow path across hosts.

1 Network Simulation is a technique where a software program models the behavior of a network either by calculating the interaction between the different network entities (hosts/packets, etc).
Environment Details

The validation was performed at Rackspace, San Antonio on 132 HP DL380 servers:
- **Model:** HP DL380 Gen9
- **Processor:** 2x 12-core Intel E5-2680 v3 @ 2.50GHz
- **RAM:** 256GB RAM
- **Disk:** 12x 600GB 15K SAS - RAID10
- **NICS:** 2x Intel X710 Dual Port 10 GbE

The setup consisted of 12 total TORs switches. Each switch has 4x40GB connections that uplink to the aggregation switches (2x connections to Aggr A and 2x connections to Aggr B).

- **Aggregation Network:** Cisco Nexus 9508 with 12x X9636PQ cards
- **ToR Network Switches:** Cisco Nexus 3172-PQ

Software Details

The software used for the validation was OpenStack Liberty Release (projects Neutron, Keystone) and Open Source MidoNet release 5.0.2. The server OS was Ubuntu 14.04.4 with Linux Kernel version 4.2. Servers have 10G interfaces.

Setup

All 132 HP servers were divided into 6 racks. Each rack had redundant TOR switches (or a total of 12 TOR switches). 2 aggregation switches connected the racks.

MidoNet was installed on all compute nodes. The servers with the 10G interfaces were configured with a txqueuelen of 10000 and MTU of 8192.
Two networks were used in the cloud setup, one for the data plane traffic and the other one for the management plane (monitoring and Network State Database NSDB cluster). There were six gateways for North South traffic that provided a theoretical maximum throughput of 6X10GB.

**Legend:**
- **CTL:** Controller (MidoNet cluster + Neutron + Keystone)
- **MEM:** Midokura Enterprise MidoNet Node with Insights
- **ES:** Elastic Search
- **NSDB:** MidoNet’s Network State Database nodes (based on Apache Zookeeper and Cassandra)

**Test Cases**

We performed two types of tests. First type was for Scalability and the second type was for performance. For the performance testing, several Netperf benchmarks (including TCP_stream, TCP_MAERTS, TCP_RR) were used to test end-to-end latency.
Part 1: Scalability Testing
Scalability testing is trying to demonstrate that MidoNet can satisfy the following conditions:

1. Validate that MidoNet can support 1000 physical hosts, 10,000 VMs
2. Validate that an individual MidoNet agent can support 100 VMs
3. Validate that MidoNet can support 6 Gateways with 10G uplinks

Table 1: MidoNet Performance in Scalability Tests

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Results Achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>MidoNet System can Support 1000 physical hosts, 10K VMs</td>
<td>YES</td>
</tr>
<tr>
<td>MidoNet System can support 6 Gateways with 10G uplinks</td>
<td>YES</td>
</tr>
<tr>
<td>Each MidoNet agent can support 100 VMs</td>
<td>YES</td>
</tr>
</tbody>
</table>

Why This Matters

The results of this testing show that MidoNet can support large-scale, high-performance clouds through multi-host distributed architectures.

The MidoNet distributed architecture enables significantly more cost-effective implementations. MidoNet allows data centers to distribute workload across a number of less-expensive, entry-level or blade infrastructure rather than requiring an investment in large expensive SMP systems.

Part 2: Performance Testing
These tests are trying to show that MidoNet can handle the simulations for large-scale deployments without performance overhead.

1. Validate that MidoNet delivers the same data transfer performance when deployed on bare metal versus virtual machines
2. Validate that MidoNet can deliver the same or acceptable latency on bare metal servers when compared to virtual machines
3. Validate that MidoNet delivers the same or equivalent transaction [request/response] rate on bare metal servers versus virtual machines
Table 2: Midonet Delivers Similar Performance on Bare Metal as on Virtual Machines

<table>
<thead>
<tr>
<th>Test</th>
<th>Bare Metal</th>
<th>VMs with Floating IP</th>
<th>VMs without Floating IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP_STREAM Mbps</td>
<td>9410.18</td>
<td>6186.18*</td>
<td>6191.45*</td>
</tr>
<tr>
<td>TCP_MAERTS Mbps</td>
<td>9408.77</td>
<td>6062.78</td>
<td>6329.5</td>
</tr>
<tr>
<td>TCP_RR TransRate/Sec</td>
<td>16637.17</td>
<td>17947.43</td>
<td>17846.75</td>
</tr>
</tbody>
</table>

*The TCP_STREAM demonstrates a result of 9060 Mbps for both cases (Container with FIP and Container without FIP) if running multiple TCP streams in parallel.

Why This Matters

These tests show that MidoNet delivers the same latency and throughput in performance in virtual environments as in bare metal. Virtualized infrastructure can generally operate more efficiency and lower admin to server ratios than physical. Since there is no performance penalty when running on virtualized servers, MidoNet provides flexibility and lets operators choose the environment most suited for their applications without any restrictions.

Tools Used

Netperf is an industry benchmark that can be used to measure various aspects of networking performance. Netperf is distributed in source forms. Ansible was used for the installation and automation of the tests. Anyone can access Netperf via the web page and replicate the tests using same test configurations.

Closing Thoughts

The advent of IoT, Big Data and containers has dramatically increased the amount of data managed in a modern enterprise. These businesses require a highly scalable, agile, and secure networking layer to handle their virtual environments.

Midokura's testing demonstrates how well MidoNet handles production cloud use cases, with performance in virtual environments essentially equivalent to bare metal.