ONOS
Towards an Open, Distributed SDN OS

Pankaj Berde, Matteo Gerola, Jonathan Hart, Yuta Higuchi, Masayoshi Kobayashi, Toshio Koide, Bob Lantz, Brian O’Connor, Pavlin Radoslavov, William Snow, Guru Parulkar

Open Networking Laboratory, NEC Corporation of America, Create-Net, Stanford
WAN Networks Today

Core
~250 routers
5K - 10K ports

Metro
10K - 50K routers
3M+ ports

Access
~50K devices
1M+ ports

High Availability:
99.99%

High Throughput:
500K – 1M
ops/sec

Low Latency:
10 – 100ms

ONOS
Global Network View/State
(200GB – 1TB+)

Application

Core Network
Distributed, SDN OS

- Scale-out
- High Availability
- Programming Abstractions
- High Performance
- Applications

Network OS for WAN and Service Provider networks
Clean separation of Control Plane from Data Plane
Distributed, SDN OS

ONOS

- High Availability
- Programming Abstractions
- Scale-out
- High Performance
- Applications

Network OS for WAN and Service Provider networks
Clean separation of Control Plane from Data Plane
Distributed Architecture

- Distributed Registry (Zookeeper)
- Distributed Data Store (RAMCloud)
- Event Notifications (Hazelcast)
- Application
  - Global Context Modules
  - OpenFlow Manager (Floodlight)
  - Network View API
  - Global Context Modules
  - OpenFlow Manager (Floodlight)
  - Global Context Modules
  - OpenFlow Manager (Floodlight)
Network OS for WAN and Service Provider networks

Clean separation of Control Plane from Data Plane

Distributed, SDN OS

ONOS

- Scale-out
- High Performance
- Applications
- Programming Abstractions
- High Availability
Improving Latency

• Initial system performance was terrible

• Reduce number of remote operations

<table>
<thead>
<tr>
<th>Adding a Switch</th>
<th>Reads</th>
<th>Writes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic Graph Data Model</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Custom Data Model</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Topology State

- Sometimes remote reads/writes are too slow

Topology Replicas

- Exploit read-heavy access pattern by storing a copy on each instance
- Build indices in-memory to improve lookup time
- Apply updates atomically to maintain integrity
Evaluation

6 node ONOS cluster, Mininet topology, 1,000 affected flows, 6 hop path

Reaction Time:
45.2 ms (median)
75.8 ms (99th percentile)

Total Time to Reroute:
71.2 ms (median)
116 ms (99th percentile)
Distributed, SDN OS

ONOS

Scale-out

High Performance

Applications

High Availability

Programming Abstractions

Network OS for WAN and Service Provider networks
Clean separation of Control Plane from Data Plane
Global Network View

Network State
- Topology
  (Switch, Port, Link, ...)
- Network Events
  (Link down, Packet-In, ...)
- Flow state
  (Flow-tables, connectivity paths, ...)

Applications

<table>
<thead>
<tr>
<th></th>
<th>Switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>Port</td>
</tr>
<tr>
<td>-</td>
<td>Link</td>
</tr>
<tr>
<td>△</td>
<td>Host</td>
</tr>
<tr>
<td>![Intent]</td>
<td>Intent</td>
</tr>
<tr>
<td>![FlowPath]</td>
<td>FlowPath</td>
</tr>
<tr>
<td>![FlowEntry]</td>
<td>FlowEntry</td>
</tr>
</tbody>
</table>
Distributed, SDN OS

- Scale-out
- High Availability
- Programming Abstractions
- High Performance
- Applications

Network OS for WAN and Service Provider networks
Clean separation of Control Plane from Data Plane
Use Cases

- SDN-IP: BGP peering and prefix routing (deployment with Internet)

- Traffic Engineering on converged Packet/Optical core network

- Segment Routing using MPLS labels (in collaboration with ONF)

- Virtual Central Offices (SDN + NFV)
Looking Ahead

• Open Source by the end 2014
• Improvements to HA and performance
• Better and more general abstractions
• Isolation and Security
• Resource Scheduling
• Hierarchical or Peer-to-Peer coordination
• More use cases and deployments
Learn more at: http://onlab.us

Thanks!