SONiC Enable Fast Evolution of Cloud Networking

Xin Liu | Principal PM @ Microsoft
Wei Bai | Researcher @ Microsoft
25th June 2019
Recap: What is SONiC

configuration and management tools

Jenkins
ANSIBLE
kubernetes
puppet
CHEF

1st party

New
More apps
SNMP
BGP
DHCP
IPv6
New

Database
Platform
SWSS
Utility

TeamD
LLDP
Redis DB
SYNCD

SONiC

Switch Abstraction Interface (SAI)

Linux
<table>
<thead>
<tr>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linux</td>
<td>Linux</td>
<td>Linux</td>
<td>Linux</td>
</tr>
<tr>
<td>Basic L2/L3</td>
<td>RDMA/QoS</td>
<td>Streaming Telemetry</td>
<td>Richer Features</td>
</tr>
<tr>
<td>Containerized</td>
<td>IPv6</td>
<td>Config DB</td>
<td>Advanced Mgmt</td>
</tr>
<tr>
<td>Redis DB</td>
<td>Mgmt. via Swarm</td>
<td>Support Virtualization</td>
<td>Stringent Tests</td>
</tr>
<tr>
<td>Fast Reboot(&lt;30s)</td>
<td>ARM based</td>
<td>Warm Reboot (&lt;1s)</td>
<td>Development Tools</td>
</tr>
<tr>
<td>40G</td>
<td>100G</td>
<td>ARM based</td>
<td>Chassis Support</td>
</tr>
<tr>
<td>ASIC</td>
<td>ASIC</td>
<td>Lower end</td>
<td>ASIC</td>
</tr>
<tr>
<td>BRCM: Trident 2</td>
<td>BRCM: Tomahawk/ Tomahawk2</td>
<td>ASIC</td>
<td>BRCM: DNX</td>
</tr>
<tr>
<td>MLNX: Spectrum</td>
<td>Marvell: Prestera</td>
<td>Nephos: Taurus</td>
<td>Innovium: Teralynx</td>
</tr>
<tr>
<td>Cavium: Xpliant</td>
<td>Barefoot: Tofino</td>
<td>Barefoot: Tofino</td>
<td>Marvell: Falcon</td>
</tr>
<tr>
<td>Centec: Goldengate</td>
<td>16 platforms</td>
<td>BRCM: TD2/TH3, Helix4</td>
<td>MLNX: Spectrum II</td>
</tr>
<tr>
<td>5 platforms</td>
<td></td>
<td>Cisco: Lacrosse</td>
<td></td>
</tr>
</tbody>
</table>
SONiC Support for Disaggregated Chassis
SONiC Is Powering Microsoft At Cloud Scale
Enabling SONiC Beyond Tier 1?
Chassis – the Challenges

- Power efficiency
- Port density
- Low table scale on backend ASICs
- No standard topology/connectivity
- Proprietary ports/packet format
- Proprietary switching/load balancing
SONiC Support for Disaggregated Chassis

- CLOS Topology with Ethernet ports
- Routing: BGP-EVPN
  - One SONiC/BGP instance per ASIC
  - Frontend SONiC directly redistribute routes using EVPN
- Forwarding: VXLAN-based switching
  - Each frontend chip is a VXLAN Tunnel End Point (VTEP)
  - Packets inside the chassis are encapsulated with VXLAN headers

Routing Table of VTEP6:

<table>
<thead>
<tr>
<th>Destination</th>
<th>Next Hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.1.0/24</td>
<td>VTEP1</td>
</tr>
</tbody>
</table>
Growing Ecosystem
Newly Joined Members since Last Year
Open Invitation

• Inviting contributions in all areas
  • SONiC/SAI
  • Hardware platform
  • New features, applications, tests and tools
  • Download, test, Deploy!

• Website: https://azure.github.io/SONiC/
• Source code: https://github.com/Azure/SONiC/blob/gh-pages/sourcecode.md
Thank you