100 Gbps Open-Source Software Router? It's Here.

Jim Thompson, CTO, Netgate

@gonzopancho
Agenda

• Edge Router Use Cases – ‘Need for Speed’
• Cost, Flexibility, Control, Evolution
• The Engineering Challenge
• Solution Components
• Test Configuration & Results
• Productization
• Value Proposition
• Vendor Value Add
Need

- Campuses equipped with 100 Gbps white box h/w and fast NICs
- Secure routing software cannot keep up
- Limits campus interconnect to <10 Gbps of IMIX traffic
100 Gbps Edge Router Use Case #2

Need

- 1 – 10 Gbps secure routing needed for e-commerce, content access and developer productivity
- Big brand solution forces unnecessary features, and is expensive
- Customer wants API-based ACL mgmt
**100 Gbps Edge Router Use Case #3**

**Need**
- 100 Gbps large packet, IPSec content distribution
- Fast adaptation to network bandwidth costs
- Reduced per customer delivery costs to drive profit from competitive MRR
Evolution of Cost, Flexibility, & Control

2005 - 2014
- Vendor Proprietary H/W
- Specialty Silicon
- Vendor Proprietary S/W
- Vendor Proprietary Mgmt

2016-2017
- Whitebox H/W
- x86 Silicon
- Vendor Proprietary S/W
- Vendor Proprietary Mgmt

2018+
- Whitebox H/W
- x86 Silicon
- Open Source High Performance S/W
- Open Source Orchestration Mgmt

Vendor Control

User Control
Drivers

• Rapid movement of workloads and data to cloud
• Virtualization of everything
• Buyer demand for network architecture freedom

Enablers

• High powered, inexpensive commodity silicon
• Open source Vector Packet Processing (VPP)
• Open source Restful API + orchestration management
The Engineering Challenge

How do you fill a 100 Gbps pipe with small packets using s/w?

10 Gbps Line Rate

- Shortest possible Ethernet packet
  - 46 byte payload
  - 18 byte Ethernet header & CRC
  - 20 byte preamble, start-of-frame delimiter + inter-frame gap (IFG)
  - 84 byte total (672 bits)
- True line-rate is 10,000,000,000 bits per second, 672 bits at a time
- **14,880,952 packets per second**
- 67.2 ns to process one packet
- A 2 GHz clock CPU core → 1 core clock cycle of 0.5 ns
- 67.2/.5 = 134 CPU clock cycles/packet (CPP)

40 Gbps Line Rate

- Per packet processing budget is 67.2/4 = 16.7 ns
- Equals 134/4 = 33.5 CPP
  - Receive the packet on a given interface
  - Process the packet
  - Transmit the packet out a (presumably) different interface

100 Gbps Line Rate

- Per packet processing budget is 67.2/10 = 6.7 ns
  - Equals 134/10 = 13 CPP
  - **148,809,523 packets per second**
  - Secure routing s/w cannot get there with kernel-based, single packet per time processing
  - Enter Vector Packet Processing (VPP)
Key Solution Components

White box

High-performance open source router s/w
Key Software Components

TNSR: Open Source-based Secure Networking Software Platform

Routing Stack
Intelligence as routing protocols to forward packets as well as a control plane

Forwarding Plane
Raw muscle to get packets through the system

Discrete Mgmt
Orchestration Management Interface enables external policy configuration and visibility

CONTROL PLANE
- Free Range Routing (FRR)
- Strongswan IPSec Management

MANAGEMENT PLANE
- CLIXON
  - CLI
  - NETCONF API
  - RESTCONF API
  - YANG Parser
  - Embedded DB

DATA PLANE
- FD.io Vector Packet Processing
- DPDK Libraries and drivers for fast packet processing
Test Configuration

- Xeon Gold 6130
- 32GB RAM
- Mellanox 100 Gbps NIC
- 1U server

- I7-6950X CPU, overclocked to 3.5GHz (10C)
- ASUS X99 board
- Water-cooled to avoid thermal throttling
- 32GB RAM
- Mellanox 100 Gbps NIC
- Intel DH8950 (Coletto Creek) QAT
## Packet Processing Test Results

### Mpps

<table>
<thead>
<tr>
<th>Stream Type</th>
<th># Streams / Cores</th>
<th>1/1</th>
<th>2/2</th>
<th>8/8</th>
<th>8/256</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 byte</td>
<td>1/1</td>
<td>14.1</td>
<td>28.3</td>
<td>67.6</td>
<td>56.3</td>
</tr>
<tr>
<td>64 byte; AES-CBC-128</td>
<td>1/1</td>
<td>4.5</td>
<td>10.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64 byte; AES-GCM</td>
<td>1/1</td>
<td>4.3</td>
<td>9.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 byte</td>
<td>1/1</td>
<td>12.6</td>
<td>22.6</td>
<td>58.0</td>
<td>48.0</td>
</tr>
<tr>
<td>128 byte; AES-CBC-128</td>
<td>1/1</td>
<td>4.5</td>
<td>10.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 byte; AES-GCM</td>
<td>1/1</td>
<td>4.2</td>
<td>8.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 byte</td>
<td>1/1</td>
<td>6.3</td>
<td>11.2</td>
<td>30.7</td>
<td>34.1</td>
</tr>
<tr>
<td>256 byte; AES-CBC-128</td>
<td>1/1</td>
<td>3.7</td>
<td>10.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 byte; AES-GCM</td>
<td>1/1</td>
<td>3.5</td>
<td>8.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>512 byte</td>
<td>1/1</td>
<td>6.0</td>
<td>12.0</td>
<td>23.5</td>
<td>21.7</td>
</tr>
<tr>
<td>512 byte; AES-CBC-128</td>
<td>1/1</td>
<td>3.5</td>
<td>7.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>512 byte; AES-GCM</td>
<td>1/1</td>
<td>3.4</td>
<td>8.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1380 byte; AES-CBC-128</td>
<td>1/1</td>
<td>3.5</td>
<td>3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1380 byte; AES-GCM</td>
<td>1/1</td>
<td>2.8</td>
<td>3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 Gbps at 512 byte frames</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Gbps

<table>
<thead>
<tr>
<th>Stream Type</th>
<th># Streams / Cores</th>
<th>1/1</th>
<th>2/2</th>
<th>8/8</th>
<th>8/256</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 byte</td>
<td>1/1</td>
<td>9.48</td>
<td>19.02</td>
<td>45.43</td>
<td>37.83</td>
</tr>
<tr>
<td>64 byte; AES-CBC-128</td>
<td>1/1</td>
<td>3.02</td>
<td>7.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64 byte; AES-GCM</td>
<td>1/1</td>
<td>2.89</td>
<td>6.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 byte</td>
<td>1/1</td>
<td>14.92</td>
<td>26.76</td>
<td>68.67</td>
<td>56.83</td>
</tr>
<tr>
<td>128 byte; AES-CBC-128</td>
<td>1/1</td>
<td>5.33</td>
<td>12.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 byte; AES-GCM</td>
<td>1/1</td>
<td>4.97</td>
<td>10.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 byte</td>
<td>1/1</td>
<td>13.91</td>
<td>24.73</td>
<td>67.79</td>
<td>75.29</td>
</tr>
<tr>
<td>256 byte; AES-CBC-128</td>
<td>1/1</td>
<td>8.17</td>
<td>22.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 byte; AES-GCM</td>
<td>1/1</td>
<td>7.73</td>
<td>18.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>512 byte</td>
<td>1/1</td>
<td>25.54</td>
<td>51.07</td>
<td>100.02</td>
<td>92.36</td>
</tr>
<tr>
<td>512 byte; AES-CBC-128</td>
<td>1/1</td>
<td>14.90</td>
<td>32.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>512 byte; AES-GCM</td>
<td>1/1</td>
<td>14.47</td>
<td>36.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1380 byte; AES-CBC-128</td>
<td>1/1</td>
<td>39.20</td>
<td>39.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1380 byte; AES-GCM</td>
<td>1/1</td>
<td>31.36</td>
<td>42.56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 100 Gbps at 512 byte frames
- With these limits...
  - 10 cores
  - 40 Gbps QAT card
  - Single socket i7
- Lab cost (S/W not incl.)...
  - $2,000 box (mostly CPU)
  - $700 NIC
  - $650 CPIC (QAT) card
- ...and the H/W can be boosted
  - 2/4 socket boxes w/ NUMA
  - More cores
  - Faster NICs
Productization

White Box Edge Router  Cloud VM Router  Cloud Container Interconnect

Speed, Scale, Flexibility – when and where you need it.
Value Proposition

- Proprietary Support
- Proprietary Management
- Proprietary Software
- Proprietary HW w/ Custom Silicon

Significant CAPEX & OPEX Reduction

- Community / Vendor Support
- Open Source Management
- Open Source Software
- COTS H/W

$ + Freedom

- 40 Gbps IPsec
- 100 Gbps Routing
- $7,500 / instance
- Including support

$$ + Lock-in
Who Needs an Open Source Vendor?

• ‘Free like a puppy’
• All software evolves, has issues
• Requires understanding, software integration development, productization, testing, verification, packaging, distribution, support
• Is it worth your time?

“You can go your own way…”

• Open source software productization vendor
• Moved from project to product
• Development, productization, testing, verification, packaging, distribution, support provided
• Fraction of the cost of proprietary vendor model

“Won’t you please, please help me…”

“You can go your own way…”
Summary

- Superfast, highly scalable, open-source based router
- Fast, easy and cheap to reconfigure or add new services
- At a fraction of today’s prices
- Utility-based secure networking affordable for all
- Open source vendors make it ready for prime time