Welcome to the Dawn of Open-Source Networking™

Linux IP Routers

Bob Gilligan
gilligan@vyatta.com
Outline

- About Vyatta: Open source project, and software product
- Areas we’re working on or interested in working on
- Some of our performance testing results
- Conclusions
Vyatta - The Service Router, Redefined
Integrated, Yet Open

Web GUI & CLI

Rich Networking and Security Features

Industry-Standard 32-bit x86
Vyatta: Scalable Software Performance

Twice The Performance

<table>
<thead>
<tr>
<th>Frame Size</th>
<th>Line Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>0%</td>
</tr>
<tr>
<td>128</td>
<td>25%</td>
</tr>
<tr>
<td>256</td>
<td>50%</td>
</tr>
<tr>
<td>512</td>
<td>75%</td>
</tr>
<tr>
<td>1024</td>
<td>100%</td>
</tr>
<tr>
<td>1280</td>
<td>100%</td>
</tr>
<tr>
<td>1518</td>
<td>100%</td>
</tr>
</tbody>
</table>

- **Vyatta**: HW: Dell PE860, SW: Vyatta
- **Cisco**: HW: 2821, SW: Cisco IOS

Half The Price

<table>
<thead>
<tr>
<th>Frame Size</th>
<th>Retail Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>$0</td>
</tr>
<tr>
<td>128</td>
<td>$1,000</td>
</tr>
<tr>
<td>256</td>
<td>$2,000</td>
</tr>
<tr>
<td>512</td>
<td>$3,000</td>
</tr>
<tr>
<td>1024</td>
<td>$4,000</td>
</tr>
</tbody>
</table>
Target Markets: Mid-range Router / Firewall / VPN

- Linksys
- CRS-1

Footprint

Price vs. Performance

- SOHO
- Enterprise branch
- SMB office
- Carrier CPE
- Enterprise edge
- WAN core
- Carrier core
Vyatta Product

- Linux Software Distribution
- Open Source Project and Product
  - Community (free) and Supported (pay) versions available
  - “Appliance” version also available.
  - Open bugzilla bug database, wiki, user group lists, docs
  - Open Git source repository
- Getting Community Version:
  - Start at: http://www.vyatta.com/community/
  - Download and burn live CD: http://www.vyatta.com/download/
  - Full source at: http://suva.vyatta.com/git/
Software Focus Areas

- Interested in working with the community on features relevant to running Linux as a router
  - Router issues not necessarily the same as server or desktop issues

- Routing protocol performance: XORP Package
  - Fast convergence large routing tables
  - Software optimization
  - MP scaling

- IP Forwarding performance
  - Performance with large routing tables (> 200,000 routes)
    - Kernel routing table (FIB) hash vs. TRIE tree implementation
  - Performance forwarding min-size (64 byte) packets
  - MP scaling: Efficiently take advantage of dual/quad core processors
    - Most new machines will be dual/quad core

- Scheduling IP forwarding and user-level routing protocols
  - Router runs both; Both are CPU intensive
  - Need to ensure both get adequate CPU under heavy load
  - Efficiently and fairly on MPs
Hardware Focus Areas

- Features NICs should support:
  - PCI-e, especially for serial cards
    - Some new machines support only PCI-e
    - Older serial/WAN NICs are still PCI-X
  - Multiple MAC addresses
    - For MAC-address takeover
    - Used by Virtual Router Redundancy Protocol (VRRP) to provide High Availability
  - NAPI support
  - VLAN/Tagging support

- IPsec performance
  - Raw encryption performance
  - Hw encryption engine performance vs. more cores
Testing results: IP Forwarding Performance

Two standardized router forwarding performance tests:
- Zero-Loss Throughput Test
  - Reduce offered rate until all packets get through (higher is better)
- Packet Loss Test
  - At 100 % offered rate, measure packet loss rate at various packet sizes (lower is better)
- Both tests defined in the IETF Benchmark Methodology Spec (RFC 2544)
- Both measure at range of packet sizes (64 bytes – 1518 bytes)

Test configuration
- SmartBits 600B network traffic generator
- 2 GbE links – Bidirectional test
Testing results: IP Forwarding Performance

Platforms tested:

1. Dell PE860:
   - Celeron 336: 2.8 GHz CPU, 256 KB L2 cache
   - 533 MHz FSB
   - 2 x On-board BCM 5721 NIC
   - PCI-e x1 lane interconnect to each NIC

2. SuperMicro PDSM4+ motherboard:
   - Dual-core Pentium-D 935: 3.2 GHz CPUs, 2 MB L2 cache
   - 800 MHz FSB
   - Off-board 2-port Intel 82571 NIC
   - PCI-e x4 lane interconnect to NIC

Linux 2.6.20 kernel
- No firewall rules
- No NAT
Zero-Loss Throughput Comparison

Throughput (% of line rate)

Packet Size (Bytes)

- CPU Bottlenecked
- Link Bottlenecked
- CPU not pegged

Supermicro Throughput
Dell Throughput
Loss Rate Test Comparison

- **CPU Bottlenecked**
  - Low Interrupt Rate (NAPI)
  - Higher Performance
  - CPU = Lower Loss Rate

- **Link Bottlenecked**
  - CPU Not Pegged
  - High Interrupt Rate

Packet size (bytes) vs. Loss rate (% of line rate)
Observations

- “Low end” server platforms deliver excellent IP forwarding performance
- Forwarding performance correlates with CPU performance
  - Higher performance CPU $\Rightarrow$ higher throughput rate, lower loss rate
  - At small packet sizes, when CPU is pegged
- NAPI appears to be working
  - Interrupts moderated when CPU is pegged
- One issue to be investigated:
  - Only one CPU utilized on dual-core Pentium-D platform
Conclusion

- Linux on x86 server platform makes a great IP router!
  - “Twice the performance at half the price”.

- Vyatta is interested in working with the community to improve features relevant to IP routing