



ISP Design – Using MikroTik CHR as a BGP edge router

PRESENTED BY:

KEVIN MYERS,
NETWORK ARCHITECT

Background:

- 19+ years in Networking
- Designed/Built Networks on 6 continents
- MikroTik Certified Trainer
- MikroTik, Cisco and Microsoft Certified



Community Involvement:



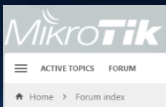
Packet Pushers (Podcast Guest / Blogger)



Group Contributor (RouterOS / WISP Talk and others)



Delegate/Roundtable contributor (NFD14)



MT Forum (Forum Veteran – Member since 2012)



Network Collective (Podcast Guest)



Profile: About IP ArchiTechs



Expert Networking

Whitebox | ISP | Data Center | Enterprise

- ✓ Global Consulting
- ✓ Managed Networks
- ✓ Monitoring
- ✓ Load Testing
- ✓ Development

Locations in: US | Canada | South America

Call us at: +1 855-645-7684

E-mail: consulting@iparchitech.com

Web: www.iparchitech.com

Goal of this presentation: When the presentation is finished, hopefully you will have walked away with a few key concepts:



- Performance characteristics of the CHR on multiple hypervisors
- Best practices when deploying the CHR
- Benefits of using the CHR vs CCR as a BGP edge router

- Which platform is better?
- Throughput capabilities?
- x86 CPU vs. Tileria?
- BGP Updates, Set BGP Attributes (Community, Localpref)

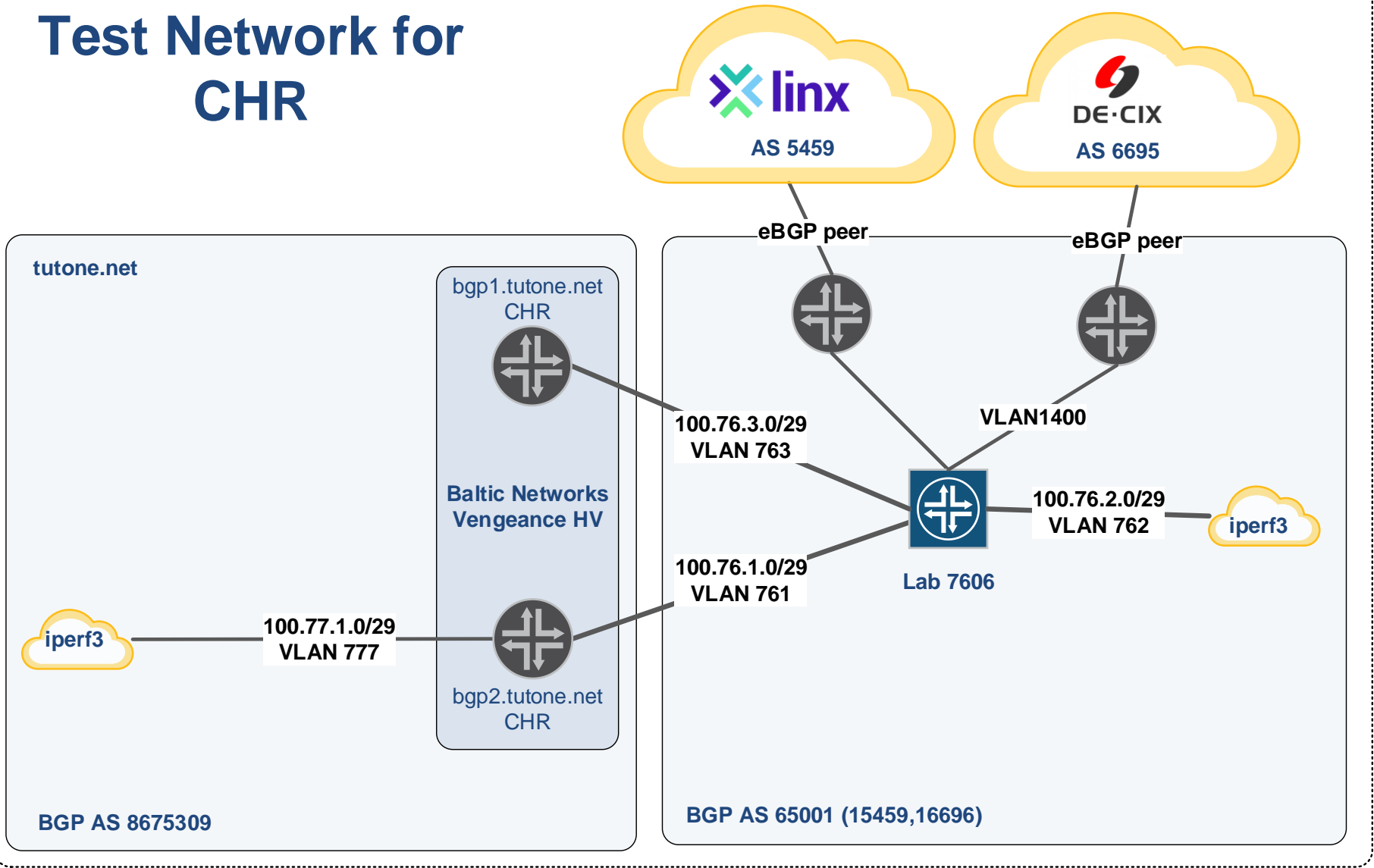


VS.

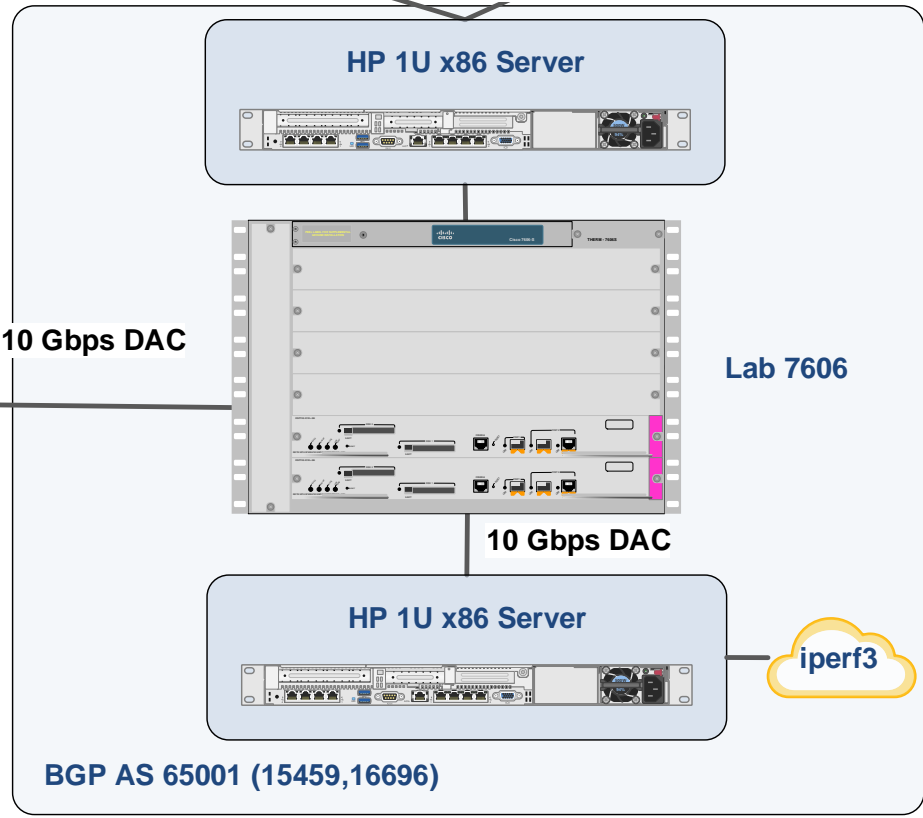
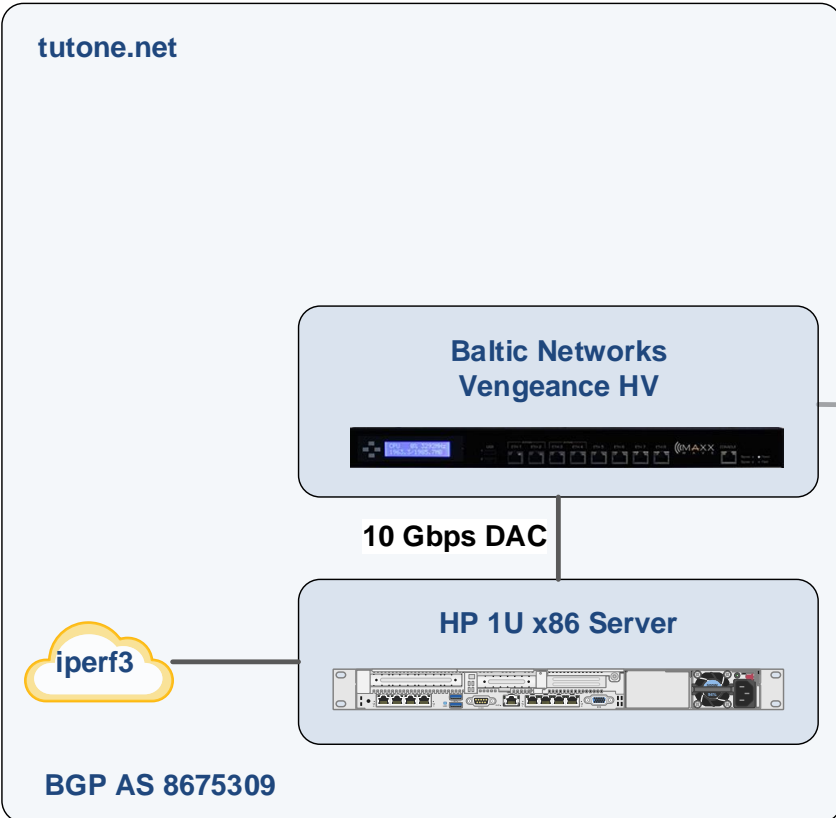
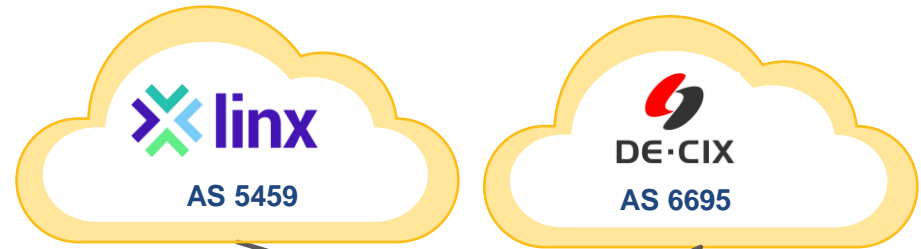


Platform		
CPU BGP routers with full tables have a high computational requirement.	x86 Better for heavy computational work. Higher power draw.	Tilera Optimized for packet transfer. Designed to be low power draw.
Throughput At 1500 bytes, 512 bytes and 64 bytes	x86 More CPU and power is required to move data at the same speed as a CCR	Tilera Handles throughput at different frame sizes slightly better than x86
Routing table size Impact of multiple tables	x86 Not limited by ASIC, x86 can process multiple tables more quickly than CCR if the HV host has a high clock speed CPU.	Tilera Not currently limited by ASIC. Limited to fixed speed single core performance

Test Network for CHR



Physical Test Network for CHR





Design: **CHR testing lab setup**

- Hypervisor details – VM provisioning
- 8 vCPUs Total
 - Performance was better when using more sockets vs. more cores with a single socket
 - With BGP – single core performance is critical. You need to use a CPU with the most powerful single core performance that you can afford for the hypervisor host.
- 4096 MB RAM

- **Concept of testing**

- Performance on 1 BGP full table
- Performance on 2 BGP full tables
- Performance at 1500 bytes, 512 bytes and 96 bytes
- Throughput performance with convergence test

- **Performance settings**

- Sockets vs CPUs – More sockets is better
- TSO/LSO – Disable for best performance
- Clock speed – Highest speed possible

Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

DE-CIX March 2018 Routing Table: Time - 26 seconds

BGP

Instances	VRFs	Peers	Networks	Aggregates	VPN4 Routes	Advertisements			
						Refresh	Refresh All	Resend	Resend All
Name	Remote Address	Remote AS	Uptime	Prefix Count	State				
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:00:26	680122	established				
bgp1.tutone.net	10.255.236.160	8675309			idle				



Design: CHR performance on VMWARE ESXi (2 Full Tables)

Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

DE-CIX and LINX March 2018 Routing Table: Time – 4:46 seconds

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ 📁 🗑️ Refresh Refresh All Resend Resend All

Name	Remote Address	Remote AS	Uptime	Prefix Co...	State
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	00:04:46	527926	established
bgp2.tutone.net	10.255.236.150	8675309	00:04:34	678597	established

50 (bgp2.tutone.net (ESXi)) - WinBox v6.41.3 on CHR (x86_64)

hboard

Session: 10.255.236.150

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ 📁 🗑️ Refresh Refresh All Resend Resend All

Name	Remote Address	Remote AS	Uptime	Prefix Count	State
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:04:35	680122	established
bgp1.tutone.net	10.255.236.160	8675309	00:04:35	100150	established



Design: CHR performance on VMWARE ESXi (1500 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

Throughput: **5.3 Gbps** Peak VM CPU: **28%**

Session: 10.255.236.160 CPU: 28%

Interface <ether2>

General | Ethernet | Loop Protect | Status | Traffic

Tx/Rx Rate: 40.8 Mbps / **5.3 Gbps**

Tx/Rx Packet Rate: 76 051 p/s / 438 428 p/s

FP Tx/Rx Rate: 0 bps / 0 bps

FP Tx/Rx Packet Rate: 0 p/s / 0 p/s

Tx/Rx Bytes: 35.0 GiB / 3937.0 GiB

Tx/Rx Packets: 559 158 356 / 2795 529 962

Tx/Rx Drops: 0 / 0

Tx/Rx Errors: 0 / 0

Legend:
Tx: 40.8 Mbps
Rx: 5.3 Gbps

Legend:
Tx Packet: 76 051 p/s
Rx Packet: 438 428 p/s

Buttons: OK, Cancel, Apply, Disable, Comment, Torch, Cable Test, Blink, Reset MAC Address



Design: CHR performance on VMWARE ESXi (512 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

Throughput: **3.0 Gbps** Peak VM CPU: **38%**

Session: 10.255.236.160 CPU: 38%

Interface <ether2>

General Ethernet Loop Protect Status Traffic

Tx/Rx Rate: 49.1 Mbps / **3.0 Gbps**

Tx/Rx Packet Rate: 90 859 p/s / 669 883 p/s

FP Tx/Rx Rate: 0 bps / 0 bps

FP Tx/Rx Packet Rate: 0 p/s / 0 p/s

Tx/Rx Bytes: 38.2 GiB / 4207.0 GiB

Tx/Rx Packets: 609 492 425 / 3092 913 865

Tx/Rx Drops: 0 / 0

Tx/Rx Errors: 0 / 0

OK
Cancel
Apply
Disable
Comment
Torch
Cable Test
Blink
Reset MAC Address

Tx: 49.1 Mbps
Rx: 3.0 Gbps

Tx Packet: 90 859 p/s
Rx Packet: 669 883 p/s

Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

Throughput: **1.1 Gbps** Peak VM CPU: **43%**

Session: 10.255.236.160 CPU: 43%

Interface <ether2>

General | Ethernet | Loop Protect | Status | Traffic

Tx/Rx Rate: 41.0 Mbps / **1107.8 Mbps**

Tx/Rx Packet Rate: 76 037 p/s / 923 181 p/s

FP Tx/Rx Rate: 0 bps / 0 bps

FP Tx/Rx Packet Rate: 0 p/s / 0 p/s

Tx/Rx Bytes: 40.8 GiB / 4289.7 GiB

Tx/Rx Packets: 650 829 762 / 3494 941 812

Tx/Rx Drops: 0 / 0

Tx/Rx Errors: 0 / 0

Legend:
■ Tx: 41.0 Mbps
■ Rx: 1107.8 Mbps

Legend:
■ Tx Packet: 76 037 p/s
■ Rx Packet: 923 181 p/s

Buttons: OK, Cancel, Apply, Disable, Comment, Torch, Cable Test, Blink, Reset MAC Address



Design: CHR performance on VMWARE ESXi (full table + bw load)

Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

5 Gbps load plus reset of both upstreams: **Convergence Time – 11:05**

Session: 10.255.236.160 CPU: 37%

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ [Filter] Refresh Refresh All Resend Resend All Find

Name	Remote Address	Remote AS	Uptime	Prefix Co...	State
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	00:11:05	527926	established
bgp2.tutone.net	10.255.236.150	8675309	00:11:06	680122	established

150 (bgp2.tutone.net (ESXi)) - WinBox v6.41.3 on CHR (x86_64)

shboard

Session: 10.255.236.150 CPU: 12%

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ [Filter] Refresh Refresh All Resend Resend All Find

Name	Remote Address	Remote AS	Uptime	Prefix Count	State
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:11:06	680122	established
bgp1.tutone.net	10.255.236.160	8675309	00:11:06	99048	established

Platform	Hypervisor	CHR
Baltic Vengeance	ProxMox (KVM) 5.1	6.41.3

DE-CIX March 2018 Routing Table: **Time - 26 seconds**

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ 📁 🗑️ Refresh Refresh All Resend Resend All

Name	Remote Address	Remote AS	Uptime	Prefix Count	State
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:00:25	680122	established
bgp1.tutone.net	10.255.236.157	8675309			idle



Design: CHR performance on Proxmox (KVM) (2 Full Tables)

Platform	Hypervisor	CHR
Baltic Vengeance	ProxMox (KVM) 5.1	6.41.3

DE-CIX and LINX March 2018 Routing Table: Time – 1:34

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ [Filter] Refresh Refresh All Resend Resend All

Name	Remote Address	Remote AS	Uptime	Prefix Count	State
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:01:34	680122	established
bgp1.tutone.net	10.255.236.157	8675309	00:01:33	97225	established

157 (bgp1.tutone.net (ProxMox)) - WinBox v6.41.3 on CHR (x86_64)

Dashboard

Session: 10.255.236.157

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ [Filter] Refresh Refresh All Resend Resend All

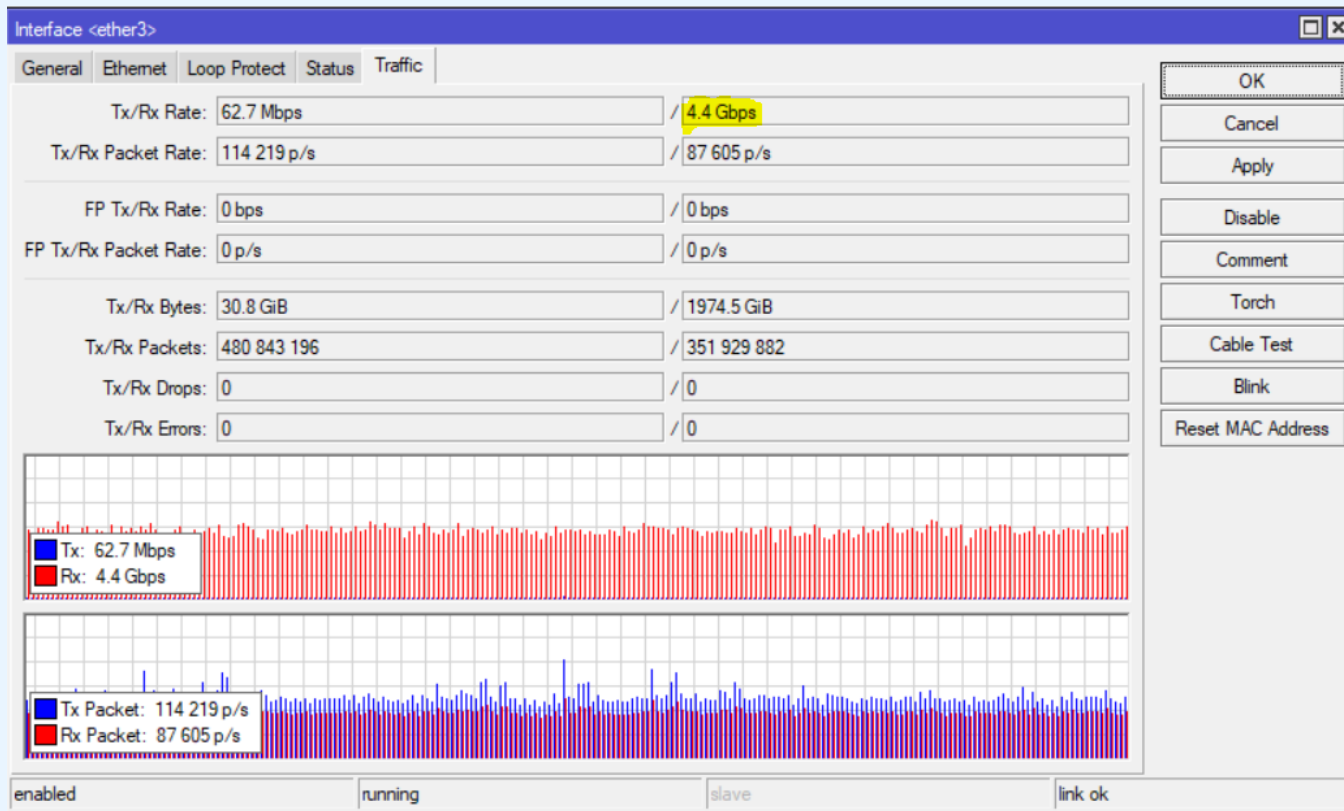
Name	Remote Address	Remote AS	Uptime	Prefix Co...	State
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	00:01:38	527926	established
bgp2.tutone.net	10.255.236.162	8675309	00:01:34	680122	established



Design: CHR performance on Proxmox (KVM) (1500 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	ProxMox (KVM) 5.1	6.41.3

Throughput: **4.4 Gbps** Peak VM CPU: **28%**





Design: CHR performance on Proxmox (KVM) (512 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	ProxMox (KVM) 5.1	6.41.3

Throughput: **1.8 Gbps** Peak VM CPU: **29%**





Design: CHR performance on Proxmox (KVM) (96 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	ProxMox (KVM) 5.1	6.41.3

Throughput: **423 Mbps** Peak VM CPU: **28%**



Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

5 Gbps load plus reset of both upstreams: **Convergence Time – 9:03**

The image shows two screenshots of the BGP configuration interface. The top screenshot shows the BGP configuration for 'bgp2.tutone.net' with two peers: LINX (IPA-JXN-7606-01) and bgp2.tutone.net. Both peers have a convergence time of 00:09:03. The bottom screenshot shows the BGP configuration for 'bgp1.tutone.net' with two peers: DE-CIX (IPA-JXN-7606-01) and bgp1.tutone.net. Both peers have a convergence time of 00:09:06.

BGP Configuration 1 (bgp2.tutone.net):

Name	Remote Address	Remote AS	Uptime	Prefix Co...	State
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	00:09:03	527926	established
bgp2.tutone.net	10.255.236.162	8675309	00:09:03	680122	established

BGP Configuration 2 (bgp1.tutone.net):

Name	Remote Address	Remote AS	Uptime	Prefix Count	State
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:09:06	680122	established
bgp1.tutone.net	10.255.236.157	8675309	00:09:03	99048	established



Design: CHR performance on Hyper V (Microsoft)(1 full table)

Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

DE-CIX March 2018 Routing Table: **Time - 12 seconds**

Session: 10.255.227.36

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - [check] [x] [trash] [filter] Refresh Refresh All Resend Resend All

Name	Remote Address	Remote AS	Remote ID	Uptime	Prefix Count	State
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	100.76.2.1	00:00:12	527926	established
bgp2.tutone.net	10.255.227.37	8675309				idle

2 items (1 selected)



Design: CHR performance on Hyper V (Microsoft)(2 full tables)

Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

DE-CIX and LINX March 2018 Routing Table: **Time - 41 seconds**

The image shows two screenshots of the WinBox BGP configuration interface. The top screenshot shows the BGP configuration for session 10.255.227.36, with memory usage at 3371.6 MB and CPU at 12%. It displays two BGP peers: LINX (100.76.1.1) with an uptime of 00:00:41 and 527926 prefixes, and bgp2.tutone.net (10.255.227.37) with an uptime of 00:00:42 and 680122 prefixes. The bottom screenshot shows the BGP configuration for session 10.255.227.37, with memory usage at 3456.2 MB and CPU at 0%. It displays two BGP peers: DE-CIX (100.76.3.1) with an uptime of 00:00:43 and 680122 prefixes, and bgp1.tutone.net (10.255.227.36) with an uptime of 00:00:42 and 98884 prefixes.

Name	Remote Address	Remote AS	Remote ID	Uptime	Prefix Count	State
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	100.76.2.1	00:00:41	527926	established
bgp2.tutone.net	10.255.227.37	8675309	100.76.3.2	00:00:42	680122	established

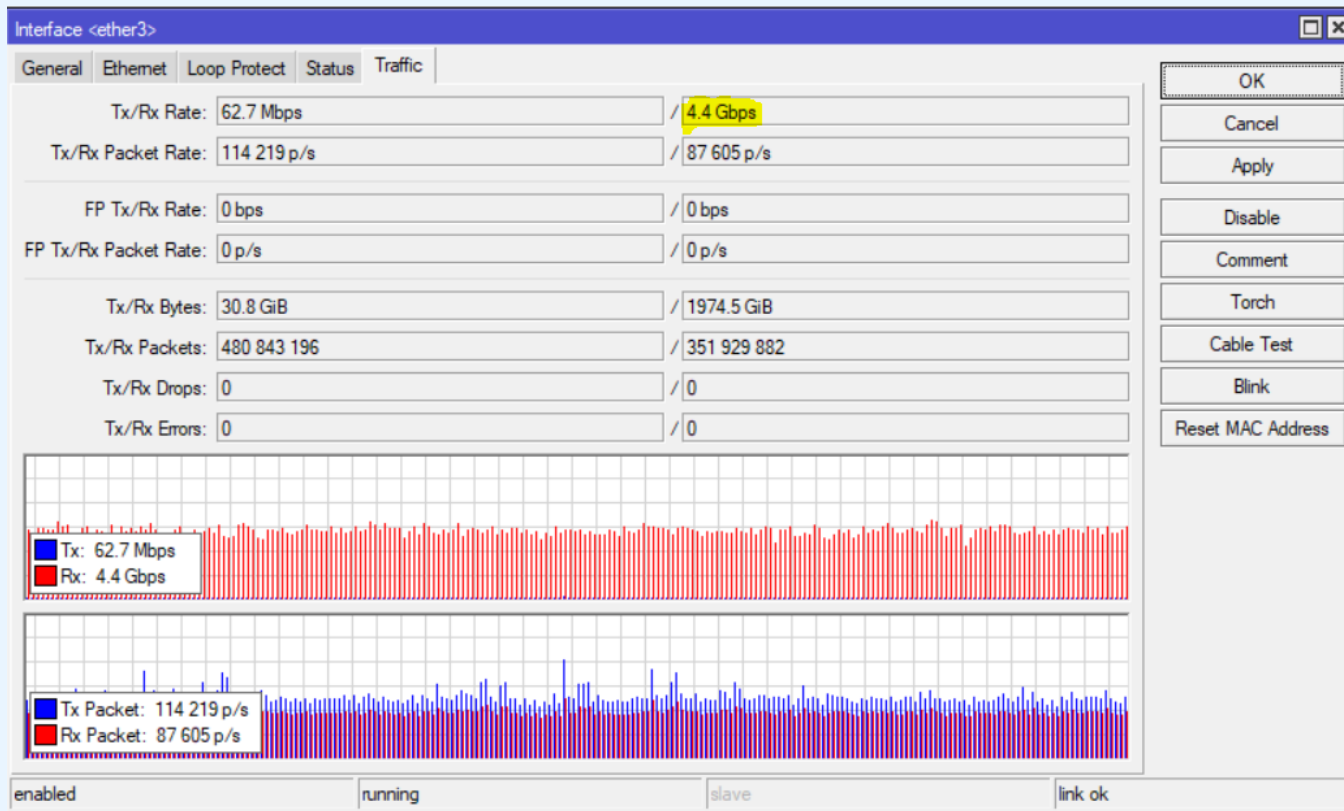
Name	Remote Address	Remote AS	Remote ID	Uptime	Prefix Count	State
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	100.76.2.1	00:00:43	680122	established
bgp1.tutone.net	10.255.227.36	8675309	100.76.1.2	00:00:42	98884	established



Design: CHR performance on Hyper-V (Microsoft)(1500 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

Throughput: **4.4 Gbps** Peak VM CPU: **32%**

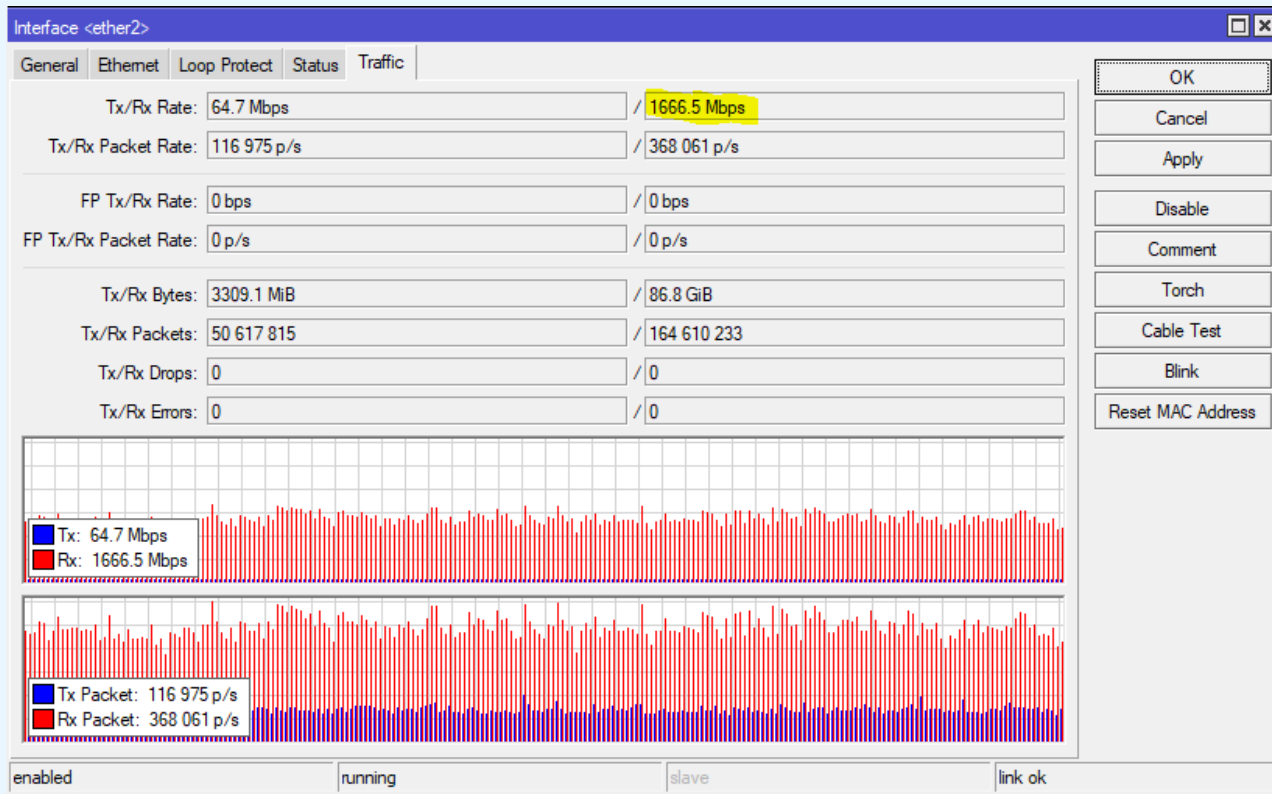




Design: CHR performance on Hyper-V (Microsoft)(512 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

Throughput: **1.6 Gbps** Peak VM CPU: **33%**





Design: CHR performance on Hyper-V (Microsoft)(96 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

Throughput: **646 Mbps** Peak VM CPU: **37%**





Design: CHR performance on Hyper-V (Microsoft)(96 byte MTU)

Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

5 Gbps load plus reset of both upstreams: **Convergence Time – :43**

BGP configuration window showing peers:

Name	Remote Address	Remote AS	Uptime	Prefix Co...	State
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	00:00:42	527925	established
bgp2.tutone.net	10.255.227.37	8675309	00:00:41	679894	established

7 (bgp2.tutone.net (HyperV)) - WinBox v6.41.3 on CHR (x86_64)

hboard

Session: 10.255.227.37

BGP configuration window showing peers with convergence time highlighted:

Name	Remote Address	Remote AS	Uptime	Prefix Co...	State
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:00:43	679894	established
bgp1.tutone.net	10.255.227.36	8675309	00:00:41	99012	established



Your local internet source for the Pacific Northwest

Residential

Business





Design: CHR real world performance example - PogoZone

US WISP-Throughput: 1.6 Gbps Peak VM CPU: 32%



- WinBox v6.41 on CHR (x86_64)

Session Settings Dashboard

Safe Mode Session:

Interface List

Name	Type	Actual MTU	L2 MTU	Tx	Rx	Tx Packet (p/s)	Rx Packet (p/s)	FP Tx	FP Rx	FP Tx Packet (p/s)	FP Rx Packet (p/s)	Comment
R	ETH1-Mgmt-NOC	1500	9014	0 bps	512 bps	0	1	0 bps	0 bps	0	0	
R	eth0	1500	65535	0 bps	0 bps	0	0	0 bps	0 bps	0	0	
R	v2001-Core-to-Edge	1500		1675.0 Mbps	223.1 Mbps	157 063	108 400	0 bps	0 bps	0	0	
R		1500		6.8 kbps	3.4 kbps	4	6	0 bps	0 bps	0	0	
R		1500		138.6 Mbps	319.6 Mbps	41 860	37 235	0 bps	0 bps	0	0	
R		1500		0 bps	1920 bps	0	4	0 bps	0 bps	0	0	
R		1500		79.3 Mbps	1351.6 Mbps	63 520	116 909	0 bps	0 bps	0	0	
R		1476	65535	3.8 Mbps	4.3 kbps	3 009	5	0 bps	0 bps	0	0	
R		1480	65535	0 bps	0 bps	0	0	0 bps	0 bps	0	0	

Interface <v2001-Core-to-Edge>

General Ethernet Loop Protect Status Traffic

Tx/Rx Rate: 1675.0 Mbps / 223.1 Mbps

Tx/Rx Packet Rate: 157 063 p/s / 108 400 p/s

FP Tx/Rx Rate: 0 bps / 0 bps

FP Tx/Rx Packet Rate: 0 p/s / 0 p/s

Tx/Rx Bytes: 188530.3 GiB / 39659.1 GiB

Tx/Rx Packets: 169379 233 465 / 113474 766 689

Tx/Rx Drops: 0 / 207 380

Tx/Rx Errors: 0 / 0

CPU

CPU	Load (%)	IRQ (%)	Disk (%)
cpu0	11	11	0
cpu1	20	20	0
cpu2	28	28	0
cpu3	66	31	0

Resources

Uptime: 12d 17:33:51

Free Memory: 2501.9 MB

Total Memory: 3931.3 MB

CPU: Intel(R)

CPU Count: 4

CPU Frequency: 4200 MHz

CPU Load: 32 %

Free HDD Space: 4014.4 MB

Total HDD Size: 4040.2 MB

Sector Writes Since Reboot: 250 296

Total Sector Writes: 250 297

Architecture Name: x86_64

Board Name: CHR

Version: 6.41 (stable)

Build Time: Dec/22/2017 11:55:15

- **Which Hypervisor is the best?**
 - **#1 Hyper-V** was faster than ESXi and ProxMox all the way around and the most consistent with results in testing. It was by far the clear winner in routing convergence times
 - **#2 ProxMox KVM** delivered much better convergence speeds than ESXi but not quite as good as Hyper-V
 - **#3 VM Ware ESXi** was the slowest in everything but raw throughput and it only has a marginal edge over Hyper-V in that area.



Design: **Questions?**

Questions??