



ISP Design – Using MikroTik CHR as a BGP edge router

PRESENTED BY:

KEVIN MYERS,
NETWORK ARCHITECT

Profile: About Kevin Myers

Background:

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



Community Involvement:



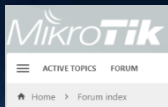
Packet Pushers (Podcast Guest / Blogger)



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



Design: **CHR vs. CCR** for full BGP tables?

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

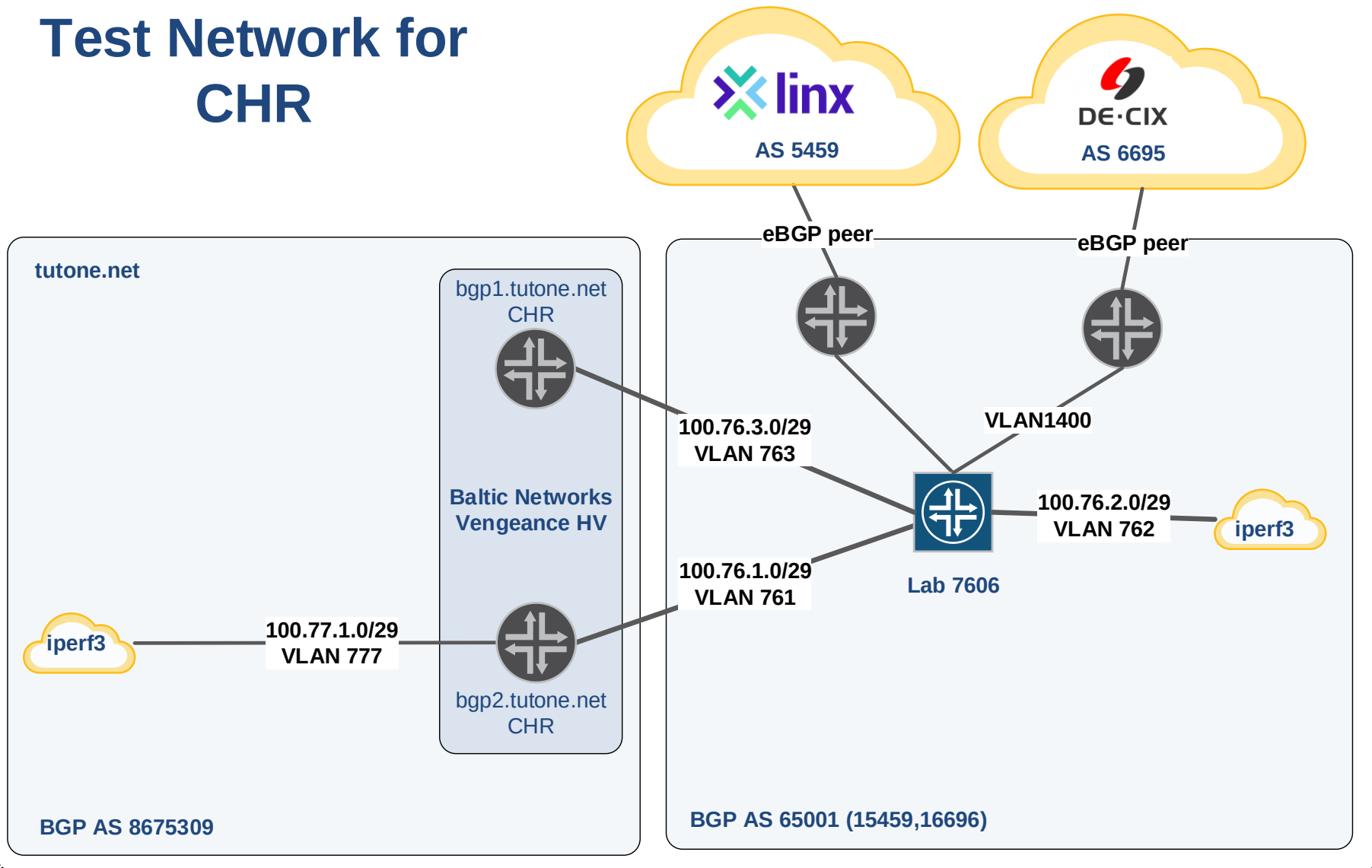


VS.



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

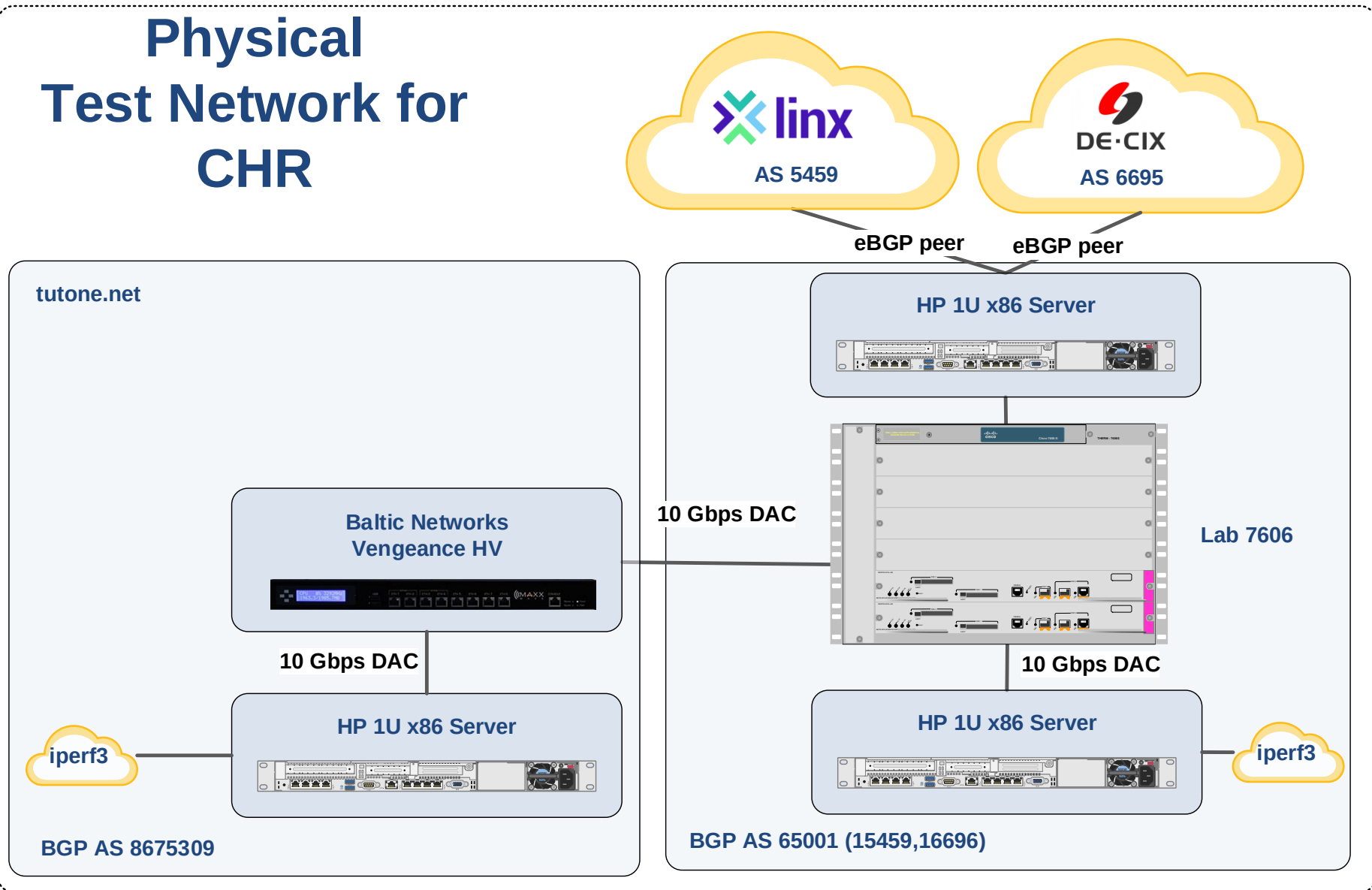
Test Network for CHR





Design: CHR testing lab setup

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

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
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

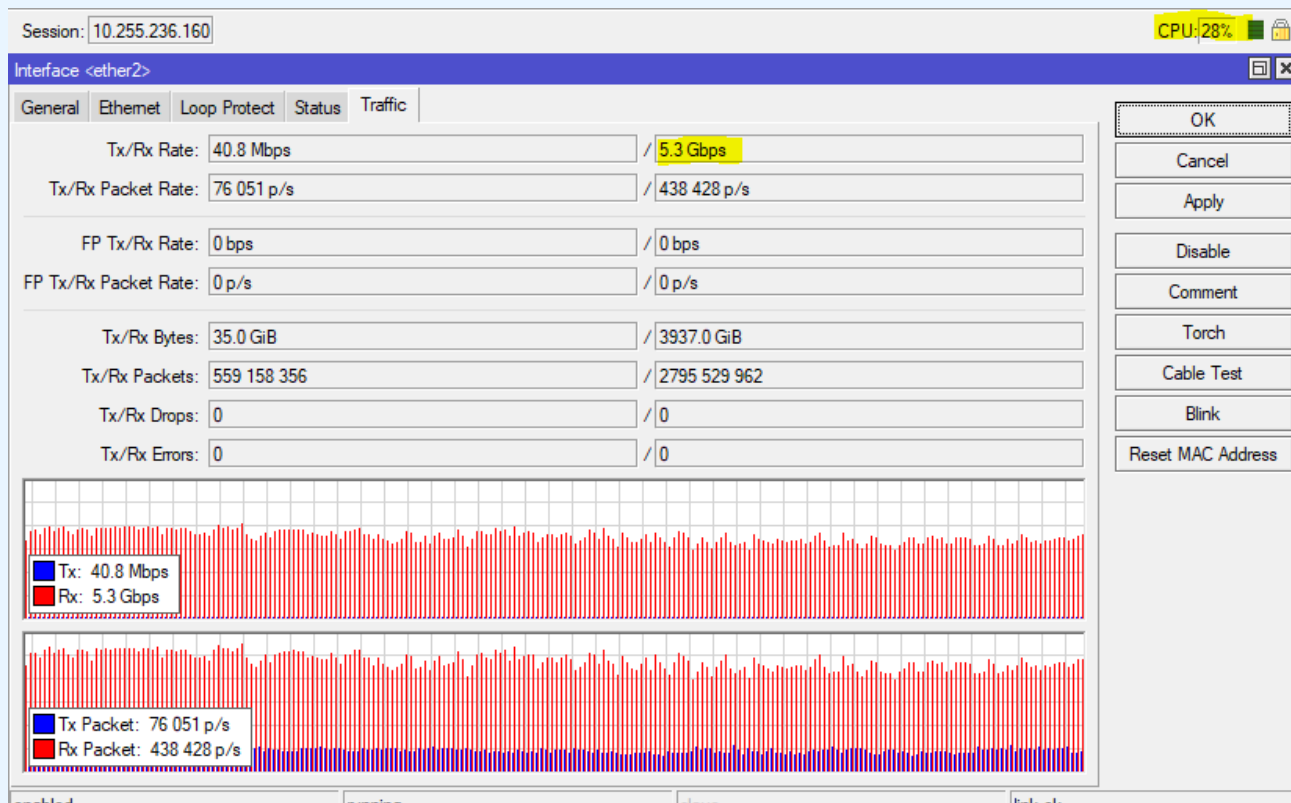
BGP						
Instances	VRFs	Peers	Networks	Aggregates	VPN4 Routes	Advertisements
Name	Remote Address	Remote AS	Uptime	Prefix Count	State	
DE-CIX (IPA-JXN-7606-01)	100.76.3.1	16695	00:04:35	680122	established	
bgp1.tutone.net	10.255.236.160	8675309	00:04:35	100150	established	



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%**

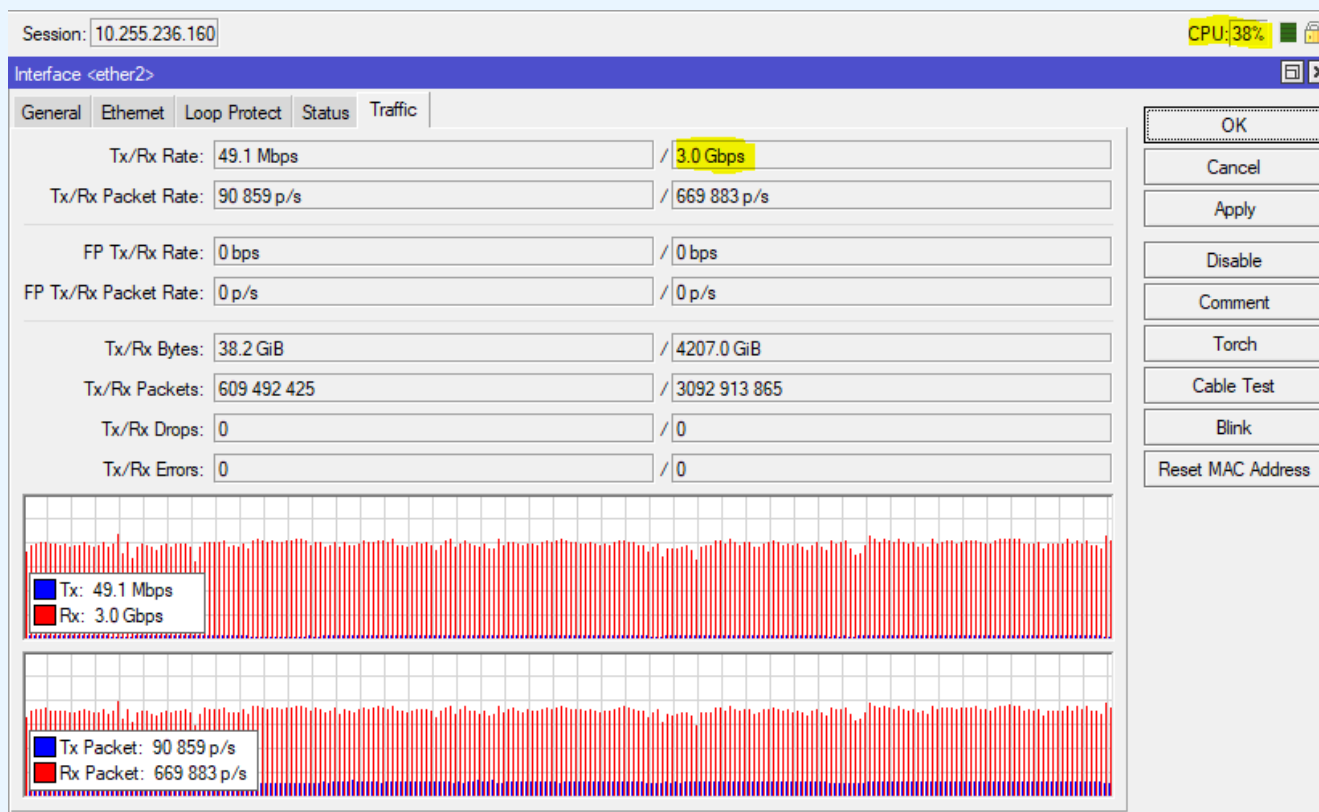




Design: CHR performance on VMWARE ESXi (512 byte MTU)

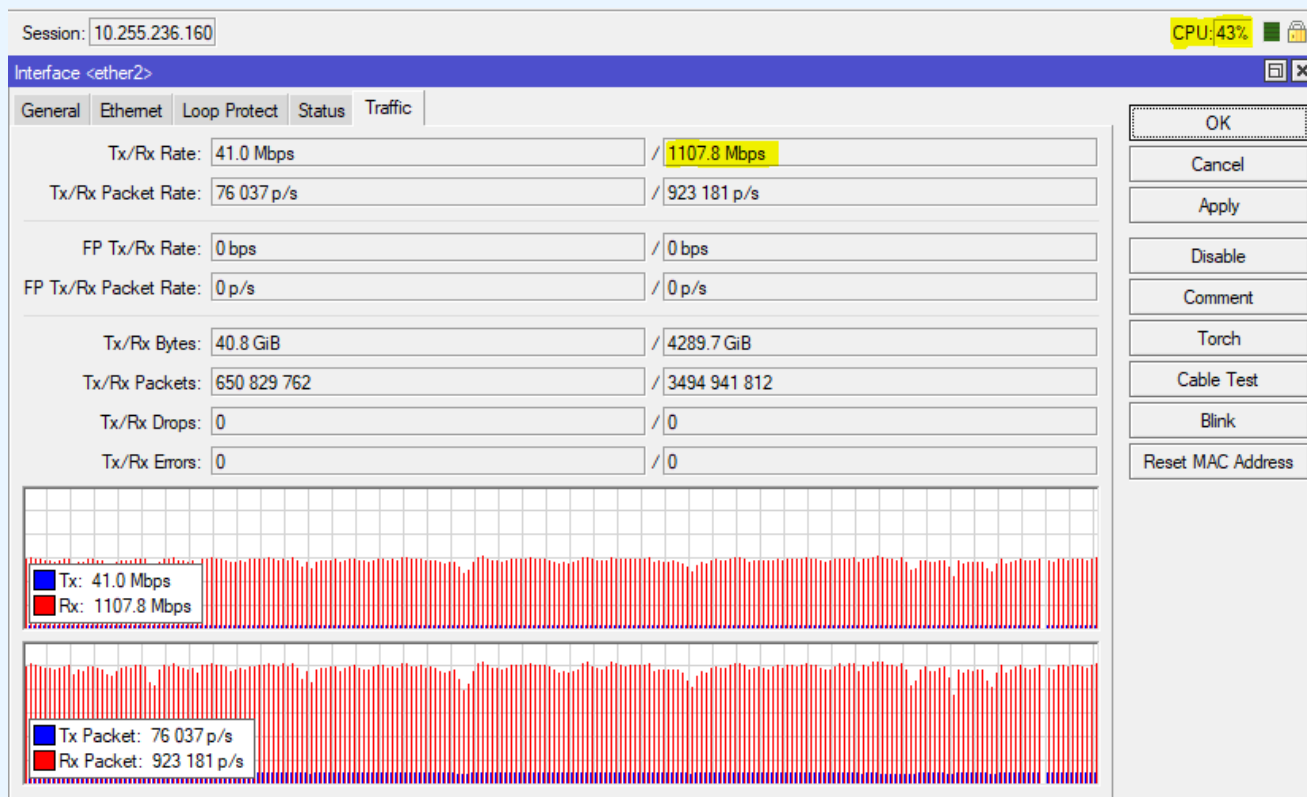
Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

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



Platform	Hypervisor	CHR
Baltic Vengeance	VM Ware ESXi 6.5	6.41.3

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





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

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		0	idle

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						
+ - ✓ ✗ 📄 🔍 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

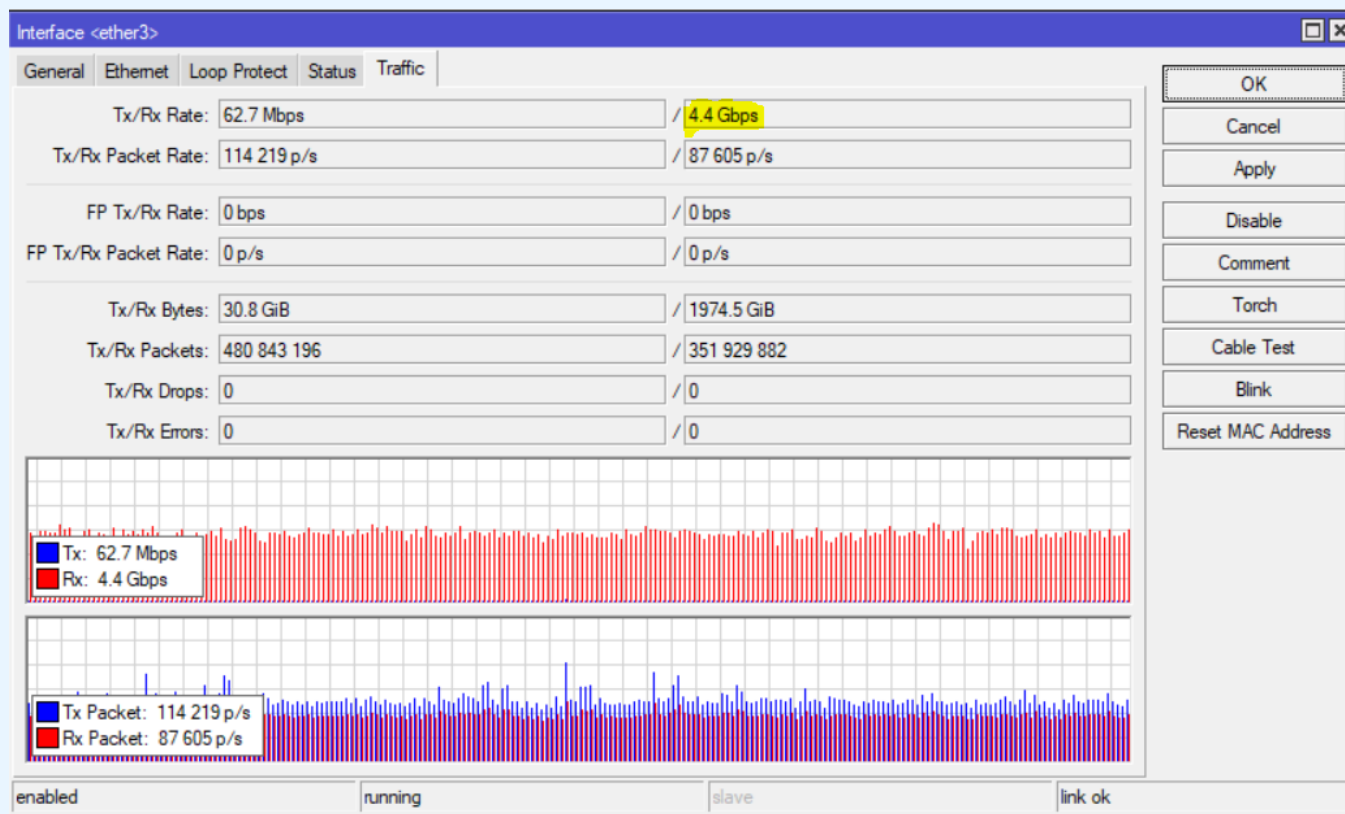
BGP						
Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements						
+ - ✓ ✗ 📄 🔍 Refresh Refresh All Resend Resend All						
Name	Remote Address	Remote AS	Uptime	Prefix Co...	State	
LINX (IPA-JXN-7606-01)	100.76.1.1	65001	00:01:38	527926	established	
bgp2.tutone.net	10.255.236.162	8675309	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%**

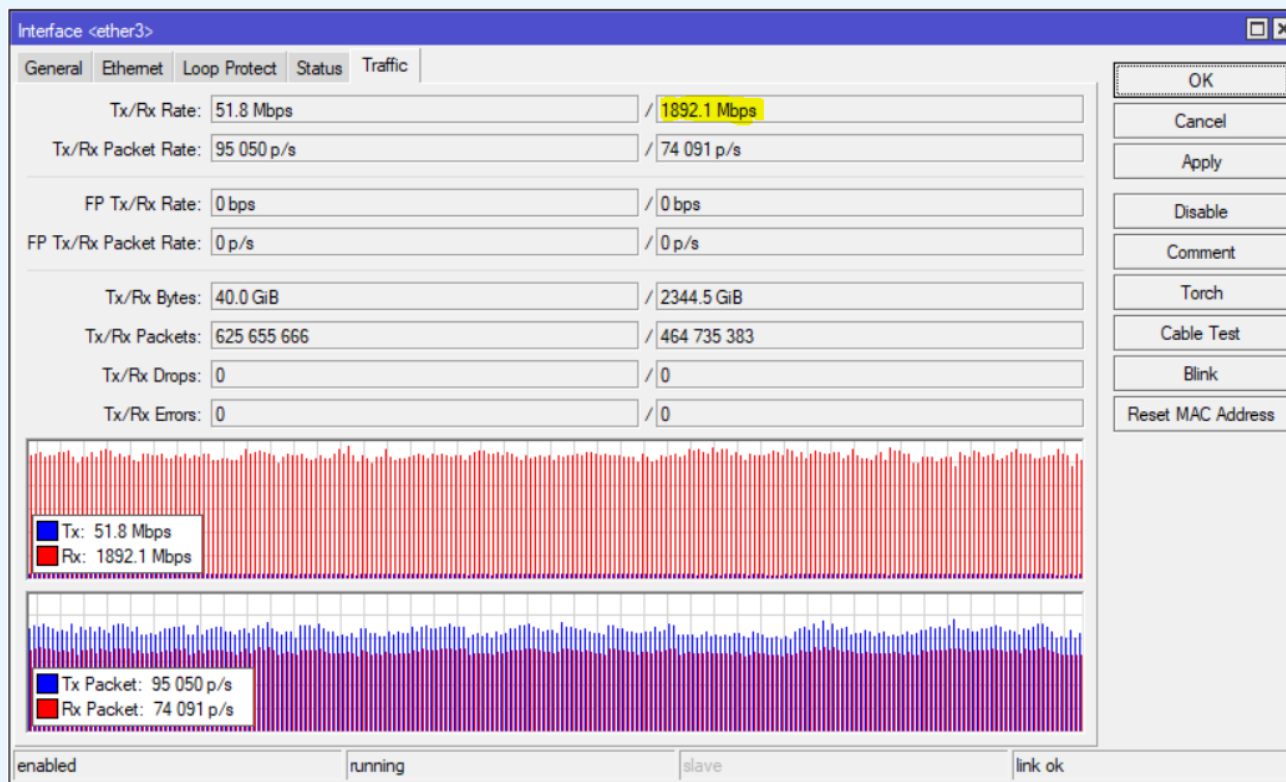




Design: CHR performance on Proxmox (KVM) (512 byte MTU)

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 displays two screenshots of the BGP configuration window in WinBox v6.41.3 on CHR (x86_64). The top screenshot shows the BGP configuration for 'bgp2.tutone.net' with a convergence time of 00:09:03. The bottom screenshot shows the BGP configuration for 'bgp1.tutone.net' with a convergence time of 00:09:06.

BGP Configuration Window (Top Screenshot):

- Instances: VRFs Peers Networks Aggregates VPN4 Routes Advertisements
- Buttons: +, -, ✓, ✗, [icon], Refresh, Refresh All, Resend, Resend All, Find
- Table:

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

62 (bgp2.tutone.net (ProxMox)) - WinBox v6.41.3 on CHR (x86_64)

hboard

Session: 10.255.236.162 CPU: 0%

BGP Configuration Window (Bottom Screenshot):

- Instances: VRFs Peers Networks Aggregates VPN4 Routes Advertisements
- Buttons: +, -, ✓, ✗, [icon], Refresh, Refresh All, Resend, Resend All, Find
- Table:

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] [yellow] [funnel] 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)

Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

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

Session: 10.255.227.36 Memory: 3371.6 MiB CPU: 12%

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

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

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

2 items (2 selected)

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

Dashboard

Session: 10.255.227.37 Memory: 3456.2 MiB CPU: 0%

BGP

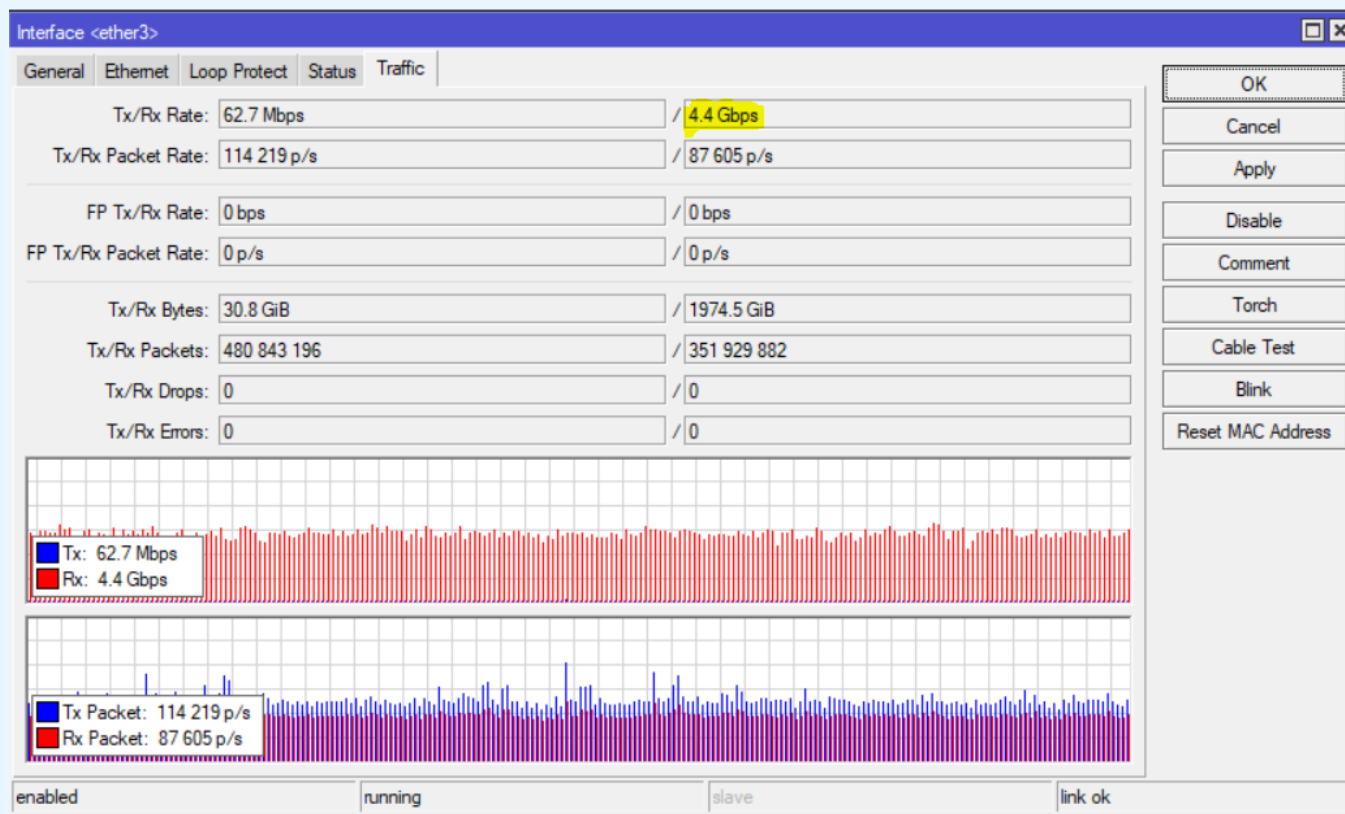
Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

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

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

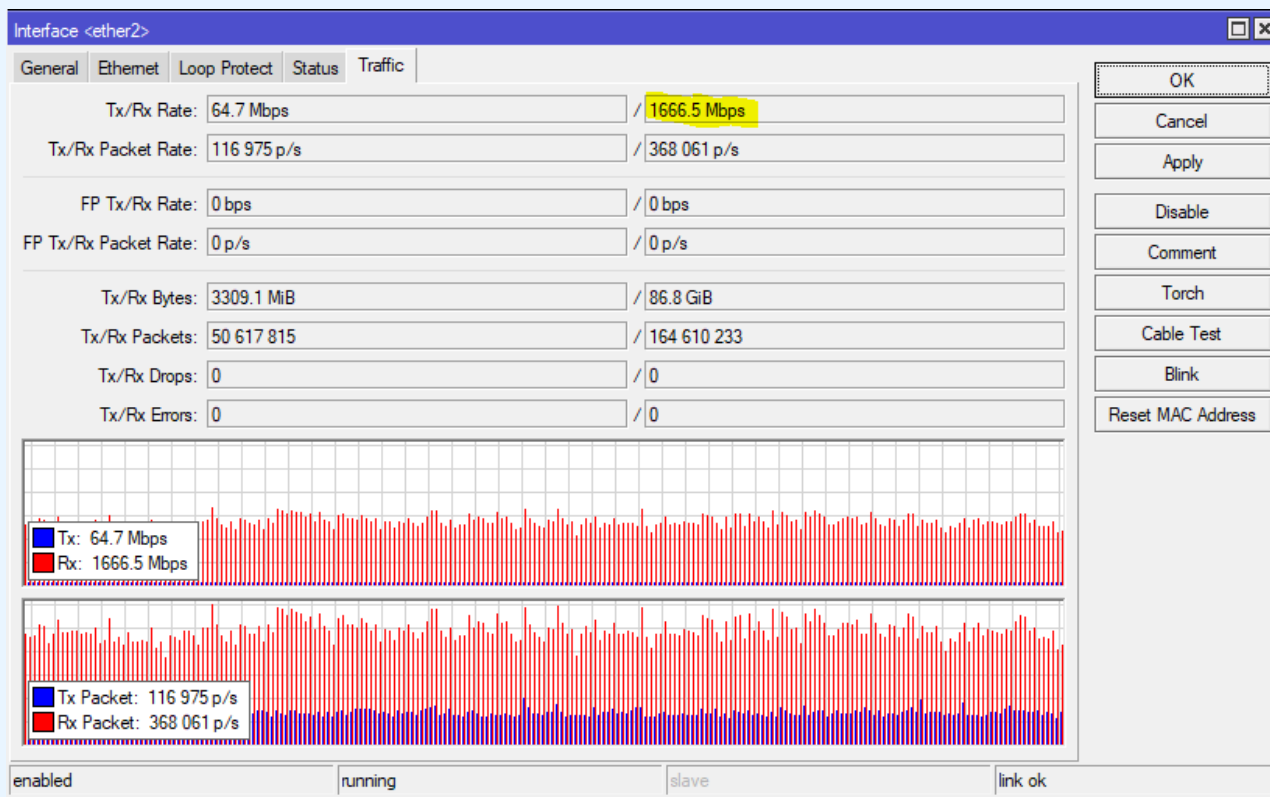
Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

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



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%**



Platform	Hypervisor	CHR
Baltic Vengeance	Hyper-V 2012	6.41.3

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

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	00:00:42	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

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
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



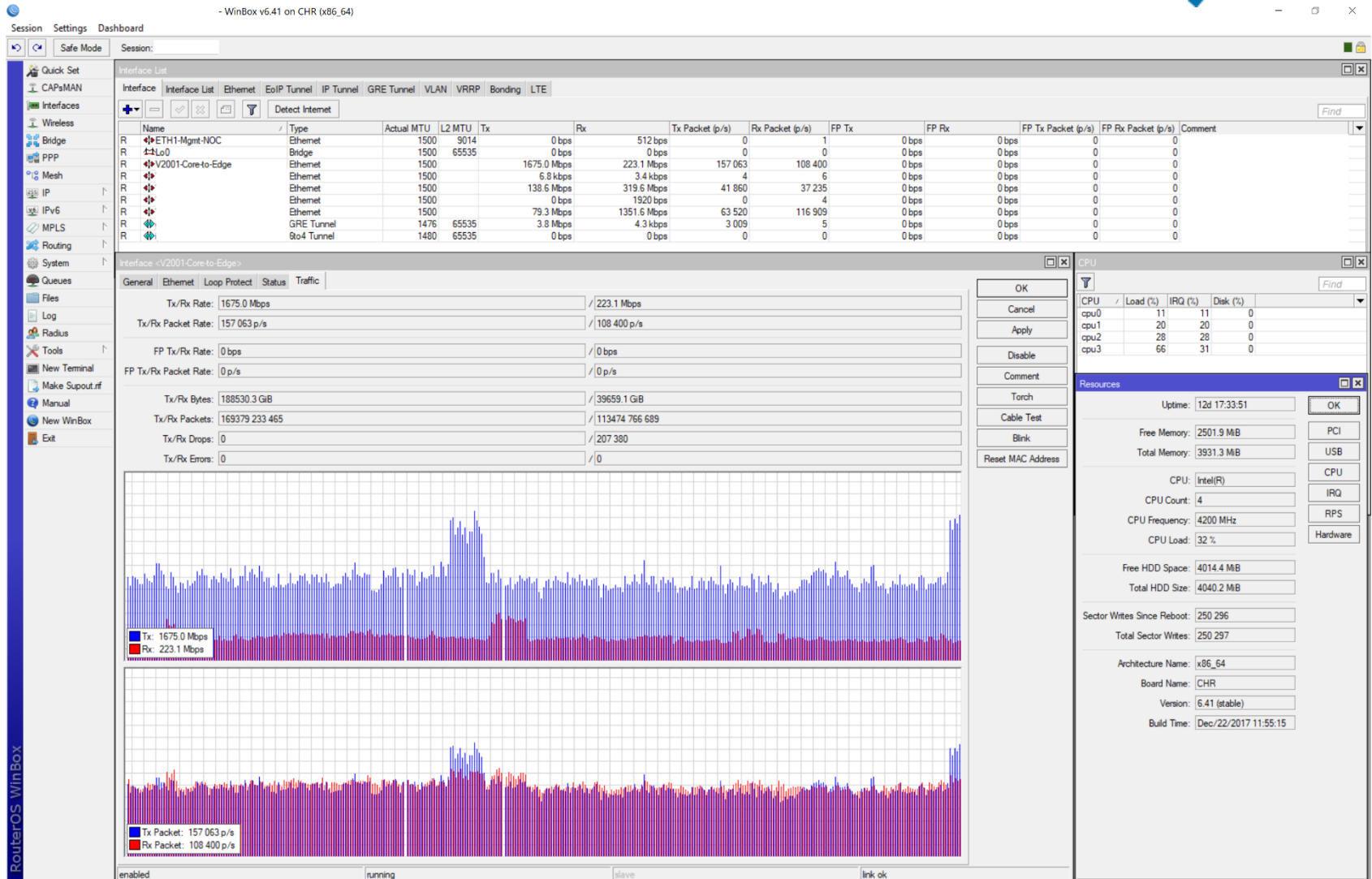
Your local internet source for the Pacific Northwest

[Residential](#)[Business](#)



Design: CHR real world performance example - PogoZone

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



- **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??