VIRTUAL SWITCH/ROUTER BENCHMARKING

Muhammad Durrani

Ramki Krishnan

Brocade Communications

Sarah Banks

Akamai

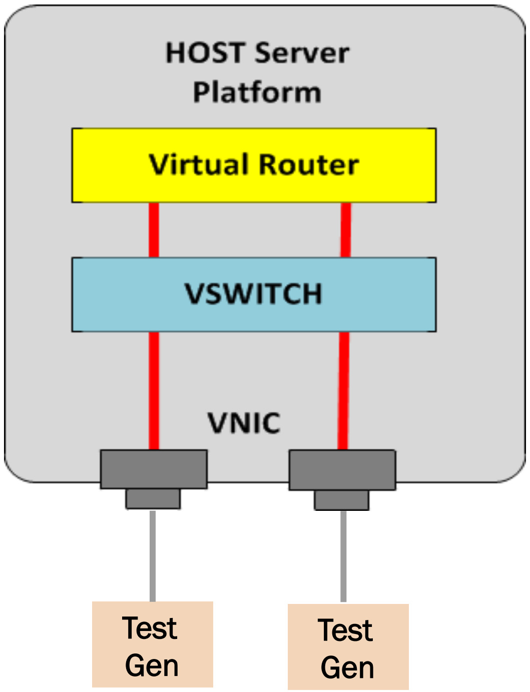
Objective

Develop comprehensive set of benchmarking tests for virtual switch/router.

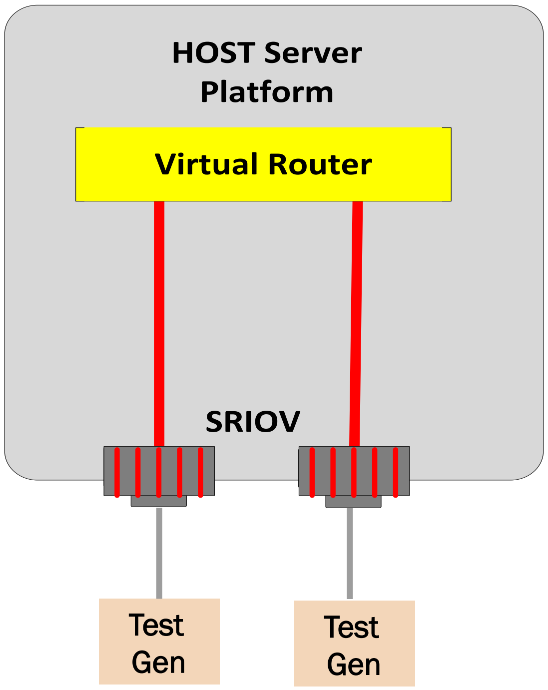
The results of the defined benchmark tests is intended to provide operators with standards based comparable data between evaluated virtual switches/routers.

Virtual Router Architectures

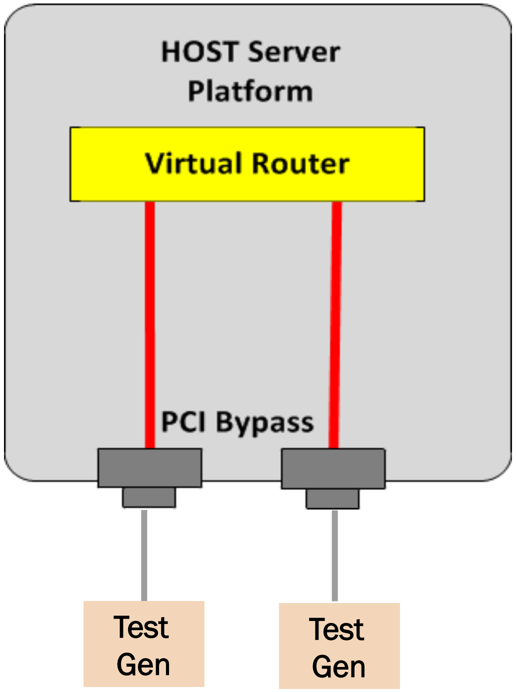
Topology 1: VNIC



Topology 2: SRIOV



Topology 3: PCI Bypass



Virtual Router Benchmark Categories

Benchmark Categories	High Level Description
Forwarding Performance	Standardized IPv4/IPv6 Unicast / Multicast forwarding performance (throughput, latency, delay variation) Test methodologies
High Availability	Benchmarking high availability in Virtual network functions (VNFs) using multiple VMs

Forwarding Performance

Discuss and clearly define a series of tests designed to determine L3 forwarding performance for virtual routers in terms of the following parameters across three Architectures:

- Maximum Throughput for min and IMIX packet sizes amendments to: RFC 2544/3918
- Forwarding Latency average for tested packet sizes
- Delay Variation [variation in latency of packet stream through the DUT] RFC 3393

NOTE: Tests will be designed to pull apart and isolate the s/w and h/w dependencies wherever possible to highlight the performance benefits and bottlenecks of a given Virtual Router to achieve this there are several amendments to existing RFCs for Virtual Routers:

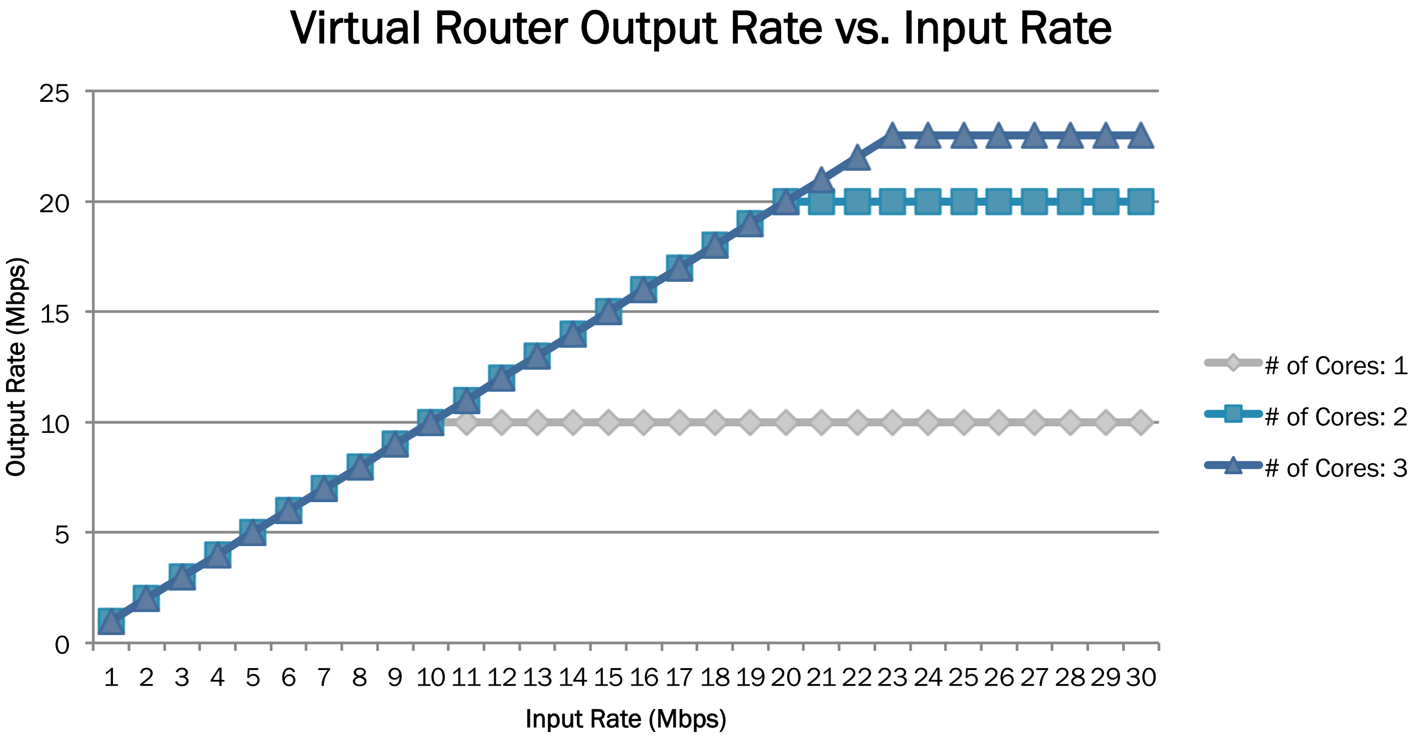
RFC 2544 / 3918 amendments:

a) IP destinations which are not present in the route lookup cache will force route lookup in extended data structures.

b) Incremental CPU core testing as capacity needs increase -- to identify and remove processor bottleneck in the data plane.

c) Effect of background router events e.g. route table churn / management polling on data plane performance

Throughput Example Results



High Availability

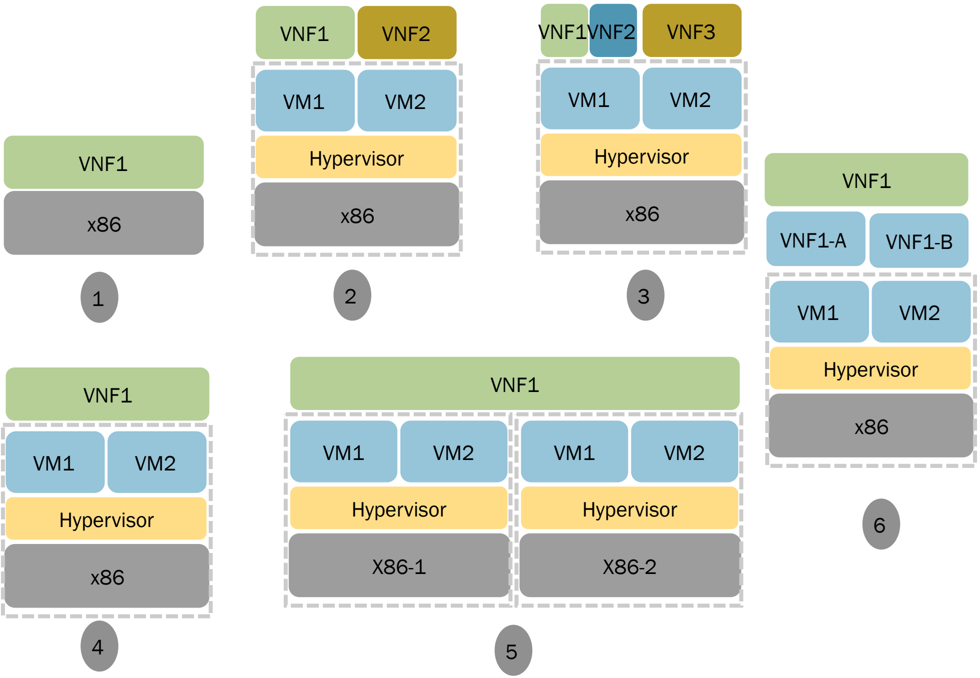
Discuss and clearly define a series of tests designed to determine recovery and convergence for virtual routers in clustered topology in terms of the following parameters

- VNF failure(Application crash/ Overload condition)
- VM failure(OS crash/overload condition)
- Hypervisor failure
- Server failure (OS crash / Overload condition)
- Physical infrastructure failure (power cycle)

Fault Domains to consider in different VNF Topologies

Simplex vs Clustered VNFs

- Depending on VNF Type (simplex vs composite), the impact of failure will vary.
  - Single instance topology (1,2 or 3 here)- VNF deployed on a single virtual machine.
  - Clustered or Composite Topology (4 and 5 here) – VNF deployed in a group of VMs which operate in a logical cluster to realize the network function.
  - VNF may also be composite of multiple distinct sub-functions (e.g. as in #6)
- As opposed to clustered VNF, Simplex VNF has no redundancy at VNF or VM layer.
- Failure at any lower layers may be cascaded and can affect multiple VNFs.
- Physical infrastructure and virtualization layers are expected to be fault tolerant.



Phased Approach

- Phase-1 – Virtual Switches/Routers
  - Single Thread versus Multi-thread implementations with Topologies covering three Architectures
    - Benchmark Unicast and Multicast Forwarding Performance (Latency/Throughput – RFC 2889,2554,3918 with Amendments)
    - Benchmark Control and Data Plane Convergence for Layer 1 ~ 3 routing Protocols
    - Benchmark service Availability by injecting infrastructure faults
- Phase-2 – other virtual appliances
  - Benchmarking Load Balancers and Firewall Appliances

THANK YOU

BACKUP

Forwarding Performance Summary Matrix

	Virtual Router Configuration & Resource Assignment									
Traffic Forwarding Performance: a) Maximum throughput for minimum & IMIX packet sizes b) Latency c) Delay Variation L3 forwarding verification: TTL decrement, CRC calculation, Egress interface selection, sequence	Virtual Router on Bare Metal: steady state	Virtual Router on Bare Metal: common background events (e.g. route churn)	Virtual Router VM: Single CPU core assigned - steady state	Virtual Router VM: Single CPU core assigned with common background events (e.g. route churn)	Virtual Router VM: N CPU cores assigned - steady state [incremental core increase to remove CPU bottleneck]	Virtual Router VM: N core assigned with common background events (e.g. route churn)	Virtual Router VNF: Single CPU core assigned - steady state	Virtual Router VNF: Single CPU core assigned with common background events (e.g. route churn)	Virtual Router VNF: N CPU cores assigned - steady state [incremental core increase to remove CPU bottleneck]	Virtual Router VNF: N core assigned with common background events (e.g. route churn)
Unicast Traffic Forwarding										
Unicast IPv4: Single source single destination (CPU cache lookup)										
Unicast IPv4: Single source scaled destination (Main memory lookup)										
Unicast IPv6: Single source single destination (CPU cache lookup)										
Unicast IPv6: Single source scaled destination (Main memory lookup)										
Unicast IPv4: Single Bi-directional flow (CPU cache lookup)										
Unicast IPv4: Scaled Bi-directional flows (Main memory lookup)										
Multicast Traffic Forwarding [Connected Hosts]										
Multicast IPv4 : Single source to single group with directly connected hosts [IGMP joins]										
Multicast IPv4 : Multiple source to single group with directly connected hosts [IGMP joins]										
Multicast IPv4 : Single source to scaled groups with directly connected hosts [IGMP joins]										
Multicast IPv4 : Multiple source to scaled groups with directly connected hosts [IGMP joins]										
Multicast IPv6 : Single source to single group with directly connected hosts [MLD joins]										
Multicast IPv6 : Multiple source to single group with directly connected hosts [MLD joins]										
Multicast IPv6 : Single source to scaled groups with directly connected hosts [MLD joins]										
Multicast IPv6 : Multiple source to scaled groups with directly connected hosts [MLD joins]										

Convergence

Discuss and clearly define series of tests designed to determine convergence performance of virtual routers'

- Control-plane Convergence tests:
  - L1 – L3 Protocols (UDLD , OSPF , ISIS , BGP, PIM-SM and PIM-SSM)
    - Best Path Selection with and without ECMP with various triggers (VNF / VM / Link etc. failure)
    - Route Reflector Convergence
    - Prefix distribution from real Feed (/8 , /19 , /23 , /24 , /29 , /32)
- Data-plane test for following traffics(Packet sizes 64 to 9k)
  - Unicast (Routes learned via OSPF , ISIS , BGP etc)
  - Multicast (Join and Leave Latency , OIF replication and SPT switch over)
  - Forwarding Performance with Features
    - QoS
    - ACLs