Juniper Networks MPC3E-NG and MPC5E EANTC Performance and Scalability Test

Juniper commissioned EANTC with an independent test of two Modular Port Concentrators, the MPC5E and MPC3E-NG for Juniper's MX Series 3D Universal Edge Routers. EANTC validated performance, scalability, energy efficiency and high availability capabilities in realistic test scenarios. Our team developed a detailed, reproducible test plan and executed the tests on site at Juniper's labs in Bangalore, India, in September 2015, and at Juniper labs in Sunnyvale, USA, in March 2016.

When the EANTC team tests new router line cards designed and manufactured by a leading vendor such as Juniper, we baseline data plane and control plane scalability, service and high availability support. We specifically focus on the integration of all the requirements in a stable solution supporting diverse multi-service scenarios.

Executive Summary

Juniper's new line cards showed flawless data throughput, excelled in control plane scalability for L2VPNs and L3VPNs, and impressed in the L3VPN service failover tests as shown in the table below.

Test Highlights

- ➔ **Up to 6,584,000 IPv4 VPN routes without impact to performance or scale**
- ➔ **Demonstrated full line rate performance (240 Gbps) for IMIX and packet sizes of 128 bytes or larger.**
- ➔ **Demonstrated impressive hitless failover of 16.3 ms for 8000 L3 VPNs and 6,584,000 routes**
- ➔ **Showed 128,000 VPWS instances**
- ➔ **FIB supported 10 Million IPv4 routes**
- ➔ **RIB supported 65 Million IPv4 entries**

Test Setup

Juniper's MPCs provide packet forwarding services for MX240, MX480, MX960, MX2010, and MX2020 routers. In the EANTC test, we used MX240s, MX480s and MX2010s to host the MPCs.

The MPC3E-NG utilizes up to two MICs (Modular Interface Cards) which provide the physical interfaces. The MPC5E is available in a variety of fixed configurations which combine packet forwarding and high density Ethernet interfaces on a single line card.

MPC3E-NG & MPC5E-40G10G

For our test, Juniper provided the following MPCs:

- **MPC3E-3D-NG-Q**
  Flexible configuration, here used with one 4-port 10-Gigabit Ethernet MIC (MIC-3D-4XGE-XFP)
- **MPC5EQ-40G10G**
  Fixed configuration MPC with six 40-Gigabit Ethernet ports and 24 10-Gigabit Ethernet ports and up to one million queues per port; in some cases, an MPC5E-40G10G was used instead (32,000 queues per port)
- **MPC5E-100G10G**
  Fixed configuration MPC with two 100-Gigabit Ethernet ports and four 10-Gigabit Ethernet ports; in the power efficiency test, one such module and one MPC5EQ-100G10G with more queues was configured in the system under test.

In our test, we used a number of test topologies per different test areas as shown in Figures 1–4.

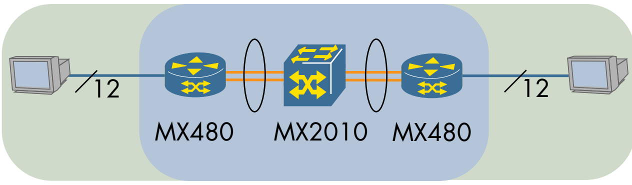


Figure 1: L3VPN Performance Test Setup

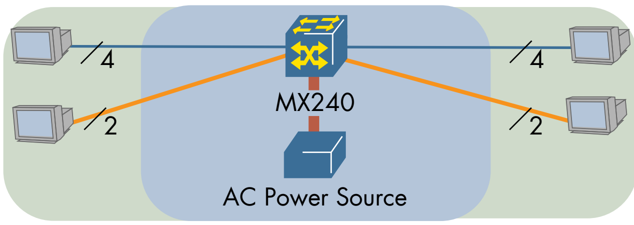


Figure 2: Power Efficiency Test Setup

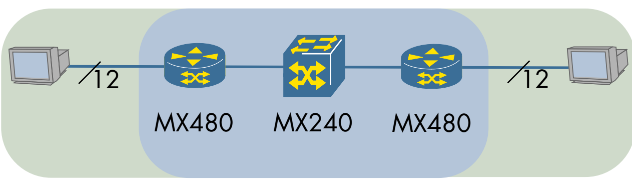


Figure 3: VPWS Scalability Test Setup

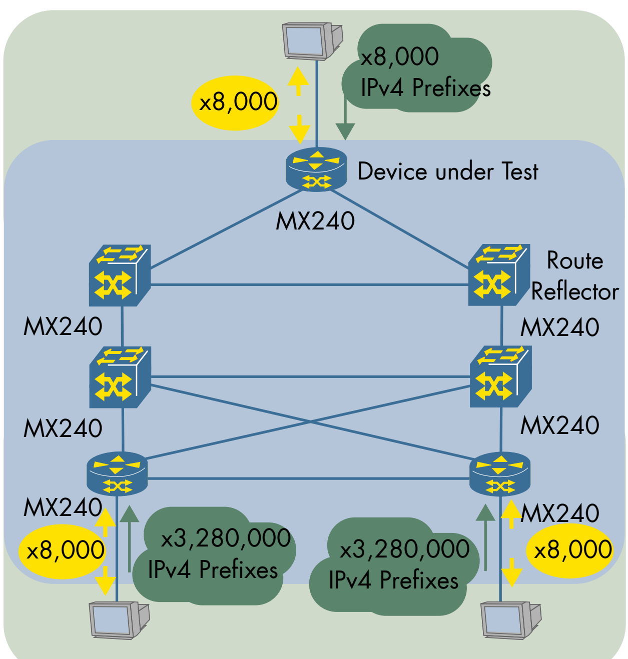
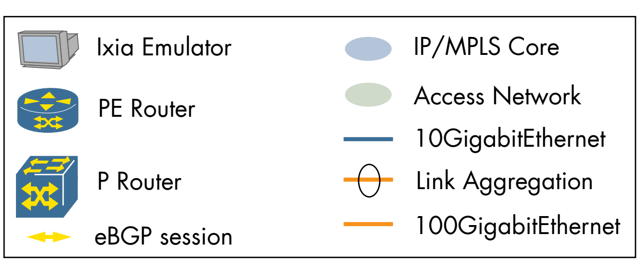


Figure 4: L3VPN Scalability VRF Instances Test Setup



EANTC defined an IPv4 and IPv6 traffic mix ("IMIX") with a range of packet sizes to ensure a realistic traffic load for the line cards as shown in the table:

Frame Size (Bytes)	Weight
64 (IPv4) or 78 (IPv6)	3
100	26
373	6
570	5
1300	6
1518	16
9000	1

Test Phase 1, Bangalore

Initially, we executed all tests using an Ixia XG12 test chassis and tester ports in a mix of 10GbE and 100GbE ports. IxOS version 6.30.850.30 and IxNetwork version 6.30 GA were used.

Junos OS version 15.1R2 was used for all tests.

Test Phase 2, Sunnyvale

In March 2016, EANTC conducted the L3VPN and IPv4-IPv6 forwarding performance tests for a second time with adjusted parameters (different frame sizes). The tests took place in Sunnyvale, USA. The same hardware models and software versions were used as in the first test phase with the following exceptions: In the L3PVN test the transit router was changed and an MX2010 router running Junos OS version 15.1F5.3 was used instead of an MX960 router. The PE routers under test remained the same type of MX480 routers. The line cards under tests remained the same type as well, running software version 15.1R2 as before. The Ixia test equipment software was updated to IxNetwork 8.00.1027.17EA and IxOS 8.00.1200.7EA.

For the second test phase, Juniper configured the MX Series 3D in Hyper Mode, which has been available since Junos OS Release 13.3R4. Enhanced MPCs such as the MPC3E, MPC4E, MPC5E, and MPC6E can be configured in Hyper Mode to support increased packet processing rates; this mode enables the router to provide better performance and throughput in common use cases. While this configuration can significantly improve performance, EANTC notes that a number of features are not supported, including creation of virtual chassis, interoperability with legacy DPCs or non-Ethernet interfaces, termination or tunneling of subscriber-based services.

Test Results: L3VPNs

VRF Instance and Route Scalability

MPC5E line cards successfully demonstrated up to 8,000 L3VPN instances with 821 IPv4 prefixes per instance.

L3VPNs, more precisely MPLS/BGP VPNs servicing IPv4 multi-point transport for business connectivity, are the most common managed packet transport service offered today. L3VPN service scalability is a critical performance metric for edge routers: The number of VPNs supported, and the number of routes per VPN (or routes in total) are key baseline figures characterizing router capabilities.

In this test, we used the topology shown in Figure 4. Juniper configured 8000 VPN instances across the three PE routers. Once these VPNs were all successfully established between the Juniper MX240 routers, we configured the Ixia emulator to advertise 821 unique IPv4 prefixes with three selected subnet masks (/16, /24 and /32) to the three PE routers. All prefixes of the same VPN instance were based on a single mask length.

VPNv4 Route Scalability

MPC5E line cards successfully demonstrated support for up to 6,584,000 IPv4 routes while operating at only 12 % CPU and 48 % memory resource usage.

Continuing with the same topology and configuration as in the previous test case, we generated traffic in all 8,000 VPNs and routes across the whole infrastructure to verify the prefixes had been successfully learned and distributed between the three PE routers.

The MX240 routers successfully learned all 6,568,000 advertised routes and 16,000 interface routes. Medium levels of traffic passed without packet loss, using the IMIX defined above.

In parallel, we monitored the CPU and memory usage of the PE routers to verify how stressful the L3VPN administration was. CPU loads remained very low (around 12 %). The PE routers used around 10 GB each to maintain all routes, resulting in maximum of 48 % memory use based on 32 GB or 73 % based on 16 GB total RAM per router. The results confirmed that there was still sufficient main memory left for other services.

L3VPN Forwarding Performance

MPC5E line cards demonstrated full line rate (240 Gbit/s) IPv4 L3VPN throughput without any packet drops, using IMIX and a variety of single packet sizes starting at 128 Bytes.

These results were achieved in the second test phase.

Once we had confirmed the control plane (service) scalability and performance of L3VPNs, we verified the data plane forwarding performance — i.e. the IPv4 throughput. Measuring in accordance with RFC2544's frame loss test, we tested the throughput and latency performance of the L3VPN traffic, using the test topology shown in Figure 1.

We forwarded IPv4 test traffic at 240 Gbit/s via 96 L3VPNs with 10,243 prefixes each, resulting in a total of 983,328 prefixes. The MPC5E line card reached full line rate on all interfaces at 128 bytes, 256 bytes, 512 bytes, 1024 bytes, 1280 bytes, 1518 bytes single frame sizes and IMIX packet sizes. During the official test runs of 600 seconds duration, the MPC5E line card did not experience any packet loss.

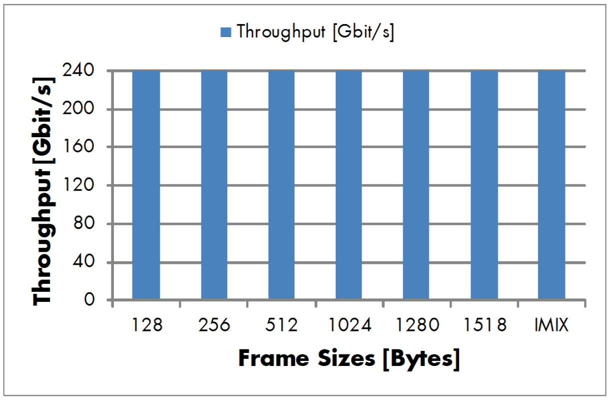


Figure 5: L3VPN Throughput

The maximum latency was measured at 70.0 μs for any fixed packet size and 83.6 μs for IMIX. The average latency was 49.0 μs for any fixed packet size and 67.3 μs for IMIX. The reported latency values represent the whole chain of three routers.

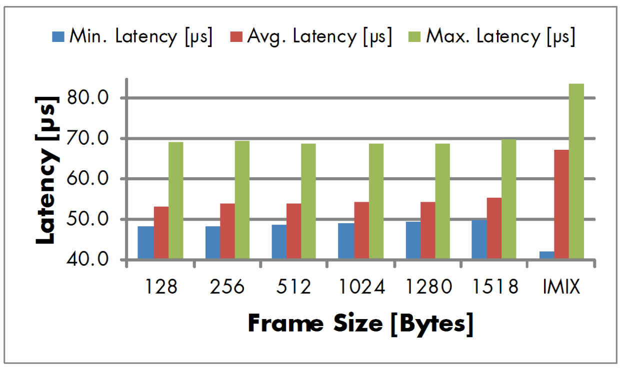


Figure 6: Latency L3VPN

IP/LDP Fast Reroute

As the next test in the L3VPN group, we measured the resiliency of VPN services in case of node or link failures in the MPLS transport network.

Using IP Fast Reroute, the MPC5E line card successfully restored service to 8000 L3VPNs with 6,584,000 routes in 16.3 ms or less.

Using the same test topology as in the first VPN test cases (Figure 4), Juniper configured Fast Reroute support to provide a loop-free alternate route (LFA) in the IP/MPLS test network.

We emulated a link failure between the top PE router and the router tagged "route reflector" in Figure 4. The emulation was achieved in a realistic way by adding a transparent shadow switch in between which interrupted only the Bidirectional Forwarding Detection (BFD) between the two MX240 routers. This failure emulated an underlying transport network issue more realistically, compared to the usual lab test way of pulling an optical fiber.

Prior to the failover event, we configured all services to pass across the primary path. Following the failover event, we verified that all services passed across the backup path. The out of service time was measured in three test repetitions, between 13.7–16.3 ms. After this time, traffic to all routes in all VPNs was fully restored.

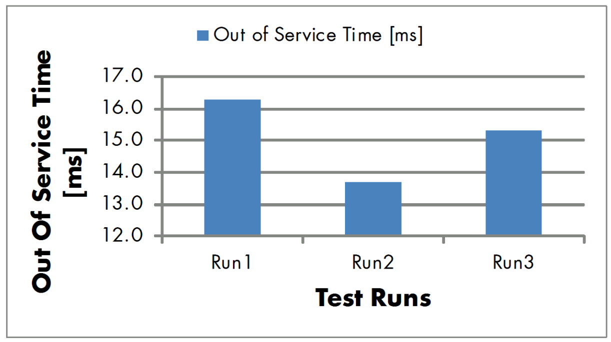


Figure 7: IP FRR Out of Service Time

We also tested recovery of the network, when the emulated BFD failure was switched off. All services returned to normal; recovery was hitless with zero packet loss.

Tail-End Protection

While LDP Fast Reroute (FRR) offers a local repair mechanism at the network level, it is largely ineffective when a failure occurs at the egress node. In a multi-homed egress node scenario, *tail-end-protection* offers a solution with repair times considerably lower than simple IP rerouting.

At the egress, the MPC5E line card demonstrated tail-end protection support, restoring 3.28 Million IPv4 routes across 8,000 VPNs in 14.1 ms or less. IP services restoration may take considerably longer depending on the individual BGP configuration.

Using the same topology (Figure 4) as previously, we emulated a node failure of the bottom right router. We configured the two bottom Ixia load generator ports to be part of an emulated multi-home CPE installation, so that the network could reach the destination connected to the bottom right PE router by routing to the bottom left PE router instead.

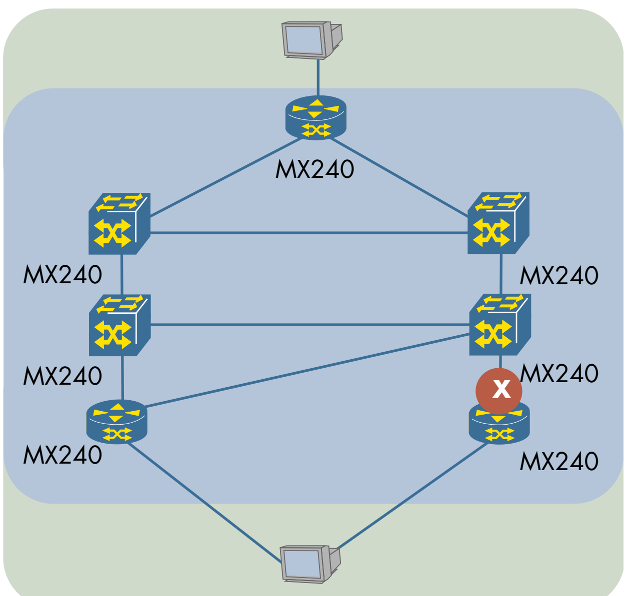


Figure 8: Tail-End Protection Test Setup

The failure itself was emulated similarly to before: An additional transparent shadow switch was inserted up from the bottom right router. Some redundant links were removed to enforce the node failure scenario. We configured two identical sets of 3,280,000 IPv4 prefixes for each of the bottom two PE routers, emulating the dual-homed CPE.

Prior to the failover event, we ensured that all routes took the primary path on the right hand side of the diagram. After the failover event (disabling BFD where marked in Figure 8), services on all 3.28 million IPv4 routes were recovered within 12.7–14.1 ms across three test runs.

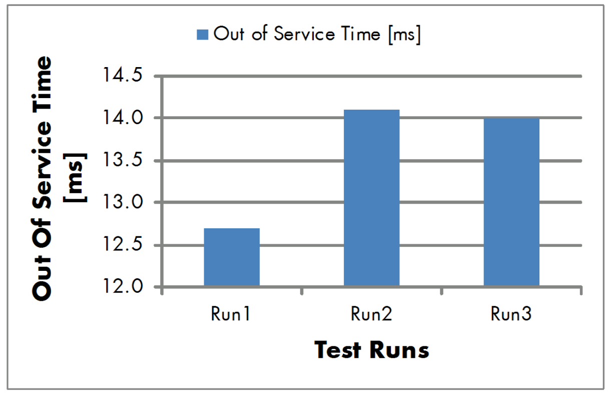


Figure 9: Tail-end Protection

We verified that the backup labels were in use in the forwarding tables and label information bases of the MPLS routers involved.

Following this fast failover, there was a subsequent service interruption of 28 s, caused by BGP global convergence within the entire test network. This was a glitch that could certainly be fixed with a proper network design; it shows that any failover designs need to take race conditions of parallel resiliency mechanisms into account.

When we switched back on the connection through the bottom right router, recovery of the service was hitless.

Test Results: L2 (Ethernet) VPNs

Ethernet VPNs are one of the most important transport services to be implemented by any service provider router; they are particularly important for wholesale access markets and some enterprise industry sectors.

VPWS Scalability

The MPC5E demonstrated 128,000 simultaneously active VPWS services. Only 6 % of the router's CPU resources and 12 % of its memory resources were used.

A point-to-point service (Virtual Private Wire Service, VPWS) has two endpoints and does not require MAC address learning. It is widely used due to its simple configuration and provisioning.

In the first test case in this area, we aimed to qualify the number of VPWS supported on a single MPC line card. We used the test scenario shown in Figure 3 with one MPC5E card, using 12 of its 10-GbE ports. This test scenario facilitated functional tests since the core router was attached with one 10-GbE port to each PE router.

Juniper configured 128,000 VPWS services on each of the two MX480 PE routers in the scenario. We validated that the services were up and running by sending Ethernet traffic across each VPWS.

In all test runs, the CPU usage observed via CLI remained below 5 % and memory usage below 28 %.

IP/LDP FRR

Orthogonal to the L3VPN failover case, we tested the failover performance for L2VPN services. The mechanisms in use are identical — IP Fast Reroute (IP FRR) provides resiliency solution to cope efficiently with node and link failures. Failing segments are repaired locally by a router detecting the failure by using pre-established backup routes, without the immediate need to inform other routers of the failure.

Using LDP Fast Reroute, the MPC5E demonstrated failover of 22.7 ms or less when configured with 8,188 VPLS instances and 511,124 MAC addresses.

We ran this test using a load of 5 Gbit/s per direction. Our team created a BFD-only failure (no loss of optical carrier) between the top PE router and the P router next to it on the bottom right. The failover scenario was staged exactly identical to the IP FRR failover test for L3VPNs.

The measurements showed between 15.1–22.7 ms recovery time for all services. This was well below our target for 50 ms. We validated that the routers adjacent to the link failure used repair routes and had updated their label information base.

We measured that the restoration was hitless and had no impact on traffic.

Test Results: Hardware and Operational Aspects

This test section covers a number of scenarios to highlight MPC5E capabilities that are important from the operational point of view, will be useful for specific usage scenarios, or are baseline figures important to qualify the router's performance.

Nonstop Active Routing (NSR)

Nonstop routing is a hardware feature that saves routing protocol information by running the routing protocol process on a backup routing engine in addition to the primary routing engine on an MX Series 3D router. The failover takes place purely internally; in contrast to Graceful Restart mechanisms, no alerts are sent to neighboring routers.

During failover of the primary routing engine, the MX Series 3D router continued to service L3VPNs, VPWS and VPLS services with IPv4-only traffic without any service disruption.

This test verified that the state of routing and signaling protocols IS-IS, BGP, and LDP would be fully maintained, and no traffic dropped in case of primary routing engine failure.

We ran this test in the topology according to Figure 4, failing the topmost PE router by pulling its active (primary) routing engine during operations. Juniper set up 1,000 L3VPNs with 500,000 unique IPv4 routes, 1,000 VPWS and 1,000 VPLS instances.

The device under test was populated with an MPC5E card with 24 GbE and 6 10GbE ports.

The MX Series 3D routers successfully showed nonstop active routing functionality: all test traffic consisting of 1,000 L3VPN traffic flows, 1,000 VPLS traffic flows and 1,000 VPWS traffic flows was forwarded hitless.

The MX Series 3D router demonstrated hitless service operation during the upgrade of an unused line card, no router reboot was required.

Power Efficiency

Network operators are sensitive to the electrical power consumption of their network devices. In this test we measured the power consumption of the MPC5E line card based on the test methodology defined in "Energy efficiency for telecommunication equipment: methodology for measurement and reporting for router and ethernet switch products", ATIS-0600015.03.2013.

In an initial baseline test, we measured the power consumption of a fully populated router configuration including the module under test (MX-MPC5E 2x100GbE+4x10GbE). In a second step, we removed the module under test from the router and measured the power consumption again. We calculated the modular power consumption of the line card under test (Pwi) as the difference between baseline power consumption and the power consumption without the module.

The Weighted Modular Energy Consumption of the MPC5E line card reached 390.4 W. The Energy Efficiency Rating for the MPC5E at full load was measured at 2.1 W/Gbit.

Junos Continuity

Service providers require maximum router uptime and rebooting a router is not a welcome activity. Juniper Networks Junos Continuity is a solution that reduces the frequency of MX Series 3D router reboots: it enables upgrades of line card software (drivers) without installing new Junos OS software or rebooting the router as a whole.

Part of this solution is a software called Juniper Agile Deployment (JAM) which manages the line card software management.

We performed the following three test steps:

- Baseline test: We sent IPv4 test traffic across the single router under test.
- Juniper installed an additional MPC3E-NG line card. "Unknown hardware status" was displayed via CLI, as expected, for the newly inserted line card. We then started the installation of JAM package (jam-mpc-2e-3e-ng64-14.1R6.2) to the base software version. (Junos version 14.1R6.2). During this operation, the traffic flows on the other line cards continued without packet loss.
- After the JAM package was installed, the correct hardware information was shown via CLI. We sent additional test traffic consisting of 1,000 L3VPN flows, 1,000 VPLS flows and 1,000

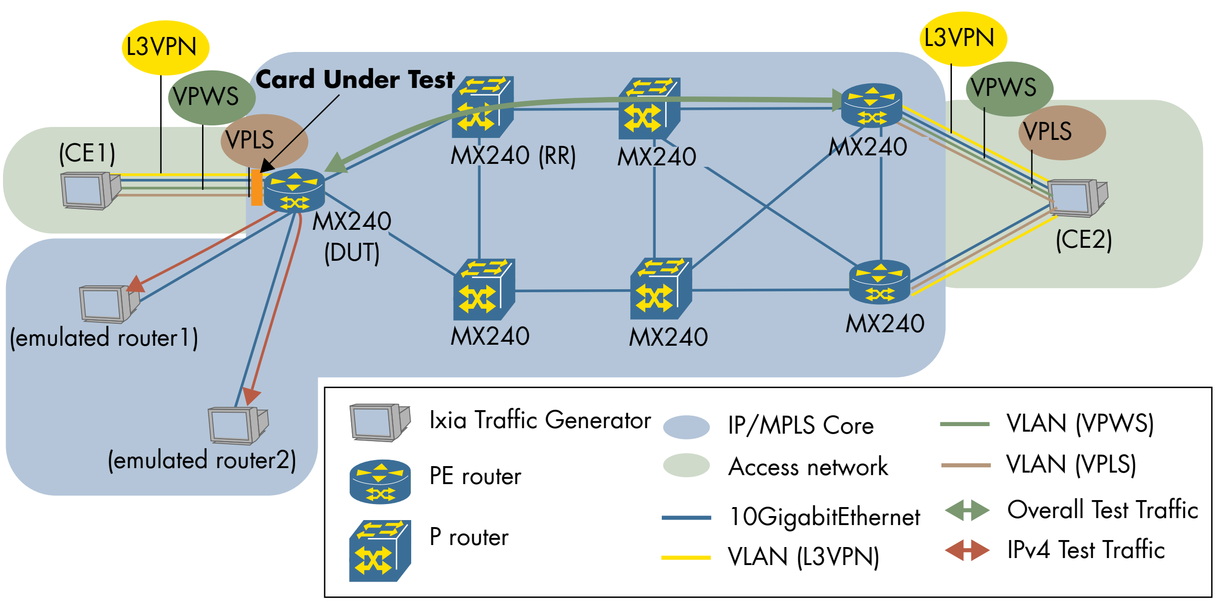


Figure 10: Juniper Continuity / Juniper Agile Deployment (JAM) Test Setup

VPWS flows on the newly detected line card under test. No packet loss was observed.

- JAM package deinstallation: as expected we did not observe any service interruption in IPv4 background traffic running on other line cards.

The router under test was able to install and uninstall the JAM package without any impact on existing IPv4 test traffic. In addition, we confirmed that the newly installed line card became usable after the upgrade by transferring VPN traffic.

FIB Scalability

The MX240, using the MPC5E line card, successfully installed 10 Million unique IPv4 routes and (separately) 10 Million unique IPv6 routes on a single port into the FIB.

Three additional baseline tests reviewed the control plane scalability aspects of the forwarding information base (FIB) and the routing information base (RIB) for IPv4 and IPv6. Edge routers must effectively support a large number of business services flows, Internet flows and broadband network gateway (BNG) flows in any combination without restriction. To ensure sufficient scale in all use cases, the maximum number of supported routes at the FIB and RIB level are a key router selection criteria.

In the first test with IPv4 routes only, the Ixia test equipment emulated a total of 10 million unique routes with different prefix lengths from /8 to /30. The large number of prefixes with lengths larger than 24 bits is not unusual for internal service provider networks. Most routes were adjacent.

EANTC validated that the route aggregation was disabled on the router. All IPv4 routes were advertised to the router on a single interface using BGP. Separately, Juniper asked us to run the same test with 10 Million IPv6 unique routes only, using an otherwise identical configuration.

Address Prefix IPv4	Prefix Length	Prefix Step IPv4	# IPv4 Prefixes
2.0.0.0	8	2.0.0.0	5
12.0.0.0	9	1.0.0.0	5
17.0.0.0	10	0.128.0.0	20
27.0.0.0	11	0.64.0.0	20
32.0.0.0	12	0.32.0.0	20
34.128.0.0	13	0.16.0.0	30
36.96.0.0	14	0.8.0.0	30
37.80.0.0	15	0.4.0.0	30
37.200.0.0	16	0.2.0.0	30
38.4.0.0	17	0.1.0.0	50
38.54.0.0	18	0.0.128.0	60
38.84.0.0	19	0.0.64.0	70
38.101.128.0	20	0.0.32.0	200
38.126.128.0	21	0.0.16.0	300
38.145.64.0	22	0.0.8.0	400
38.157.192.0	23	0.0.4.0	500
38.165.144.0	24	0.0.2.0	1k
38.173.96.0	25	0.0.1.0	1k
38.177.72.0	26	0.0.0.128	996k
46.74.152.0	27	0.0.0.64	8M
76.207.24.0	28	0.0.0.32	1M
78.183.96.0	30	0.0.0.8	230
Total			10M

Emulated BGP IPv4 Prefixes

In both cases, we verified the actual installation in the FIB by sending 5 Gbit/s data traffic using all routes across the line card under test.

IPv4 and IPv6 Mixed Performance

The results described below were achieved during the March 2016 test session.

Routers are faced with a growing fraction of IPv6 traffic as service providers and enterprises deploy new networks and services using IPv6 addresses. Currently, the average IPv6 to IPv4 ratio in service provider networks worldwide is 99:1 — but in certain markets the ratio is to two to three percent; enterprise values may be much higher. Therefore, router capability to concurrently handle both IPv4/IPv6 (dual stack) across the same physical interface is very important.

With IPv4 traffic, IPv6 traffic and with 80:20 IPv4:IPv6 traffic mix, the MPC5E line card exhibited line-rate performance for IMIX and single packet sizes of 128, 256, 512, 1024, 1280 and 1518 bytes.

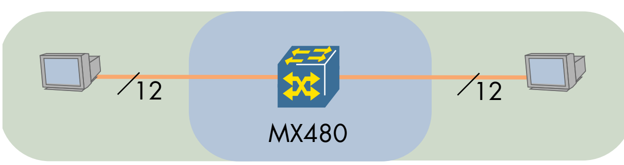


Figure 11: IPv4/IPv6 Performance Test Setup

We used the topology in Figure 12; the MX480 was configured for an IPv4/IPv6 eBGP routing scenario. We forwarded IPv4 traffic, IPv6 and a mix of IPv4 and IPv6 (80:20 proportion) test traffic at 240 Gbit/s. Packet sizes for all IPv4 and IPv6 streams were chosen to have identical sizes of 128 bytes, 256 bytes, and 512 bytes, 1024 bytes, 1280 bytes and 1518 bytes and the IMIX. In all cases, the MPC5E line card reached full line rate throughput without any packet loss.

With IPv4 traffic, the router showed average latency values between 17.5–20.2 µs and maximum latency values of 33.1–34.6 µs for fixed packet sizes. The IMIX average latency was measured as 29.4 µs and the maximum latency as 46.5 µs.

With IPv6 traffic, the router showed average latency values between 19.0–21.3 µs and maximum

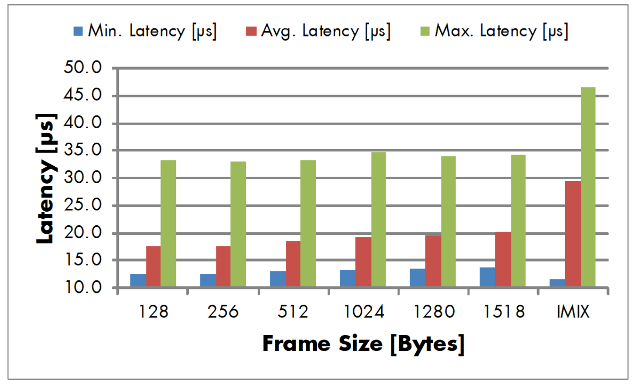


Figure 12: IPv4 Forwarding Delay

latency values of 24.4–27.7 µs for fixed packet sizes. The IMIX average latency was measured as 30.5 µs and the maximum latency as 38.7 µs.

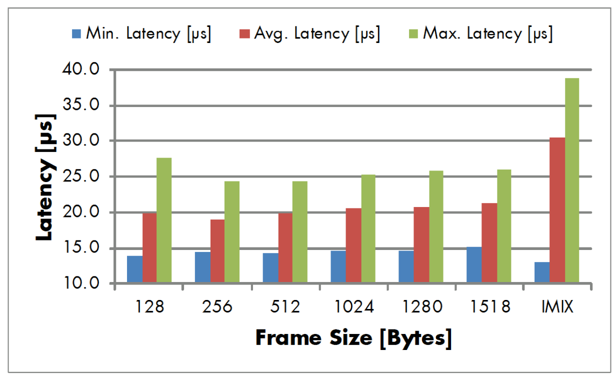


Figure 13: IPv6 Forwarding Delay

With IPv4:IPv6 traffic, the router showed average latency values between 18.7–21.0 µs and maximum latency values of 33.2–35.1 µs for fixed packet sizes. The IMIX average latency was measured as 29.7 µs and the maximum latency as 46.5 µs.

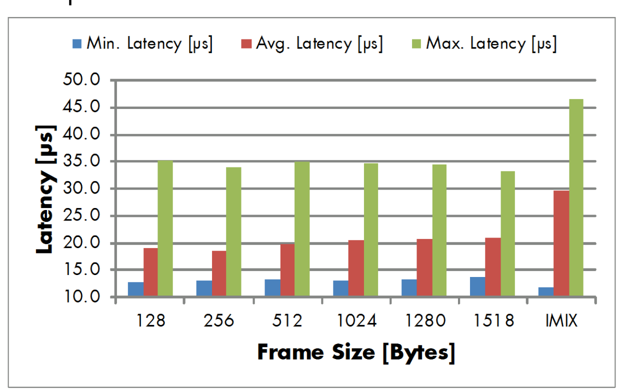


Figure 14: IPv4/IPv6 Forwarding Delay

RIB Scalability IPv4 and IPv6

The MPC5E line card demonstrated support for 65 million IPv4 routes installed in the RIB. In a separate test run, MPC5E demonstrated support for 65 million IPv6 routes installed in the RIB.

Often, routers have to process multiple routes per IPv4 prefix, since prefixes can usually be reached via multiple gateways. In this final baseline test, we evaluated the size of the routing information base (RIB) which needs to hold all these copies of routes.

We verified the maximum number of IPv4 BGP routes that the DUT can sustain in the Routing Information Base (RIB). Separately, we verified the same statement for IPv6 routes only.

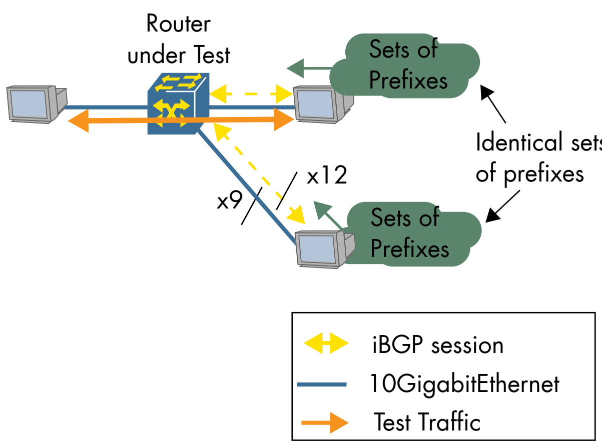


Figure 15: RIB Test Topology

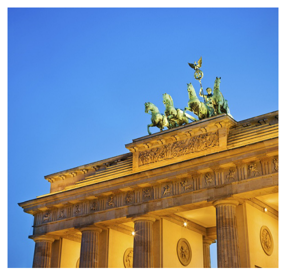
For this test, Juniper suggested another test topology where 13 identical copies of 5 million unique routes were advertised over 10 interfaces of an MPC5E card, resulting in a total of 65 million routes to be installed in the RIB. The eleventh port was used as a traffic source, and the twelfth port remained unused on Juniper's request.

In the IPv4 case, the Ixia emulator measured 21.2 minutes to populate the RIB and 5.8 minutes in addition to populate the FIB; the delay was probably related to internal processing times. We did not measure such a delay for IPv6 (both RIB and FIB population time were 22.7 minutes).

Juniper suggested to use 5 Gbit/s bidirectional throughput across the entire card (using two ports only). There was no frame loss monitored as expected.

The DUT (Juniper MX240) using the module under test: the MPC5E 3D 24XGE+6XLGE line card, successfully installed up to 65,000,044 IPv4 routes including 5,000,000 unique IPv4 prefixes in RIB table.

About EANTC



EANTC (European Advanced Networking Test Center) is internationally recognized as one of the world's leading independent test centers for telecommunication technologies. Based in Berlin, Germany, the company offers vendor-neutral consultancy and realistic, reproducible high-quality testing services since 1991. Customers include leading network equipment manufacturers, tier-1 service providers, large enterprises and governments worldwide. EANTC's proof of concept, acceptance tests and network audits cover established and next-generation fixed and mobile network technologies. http://www.eantc.com

EANTC AG, Salzufer 14, 10587 Berlin, Germany
info@eantc.com, http://www.eantc.com /
v2.5 20160628