Juniper Networks MPC9E: EANTC Performance, Scale and Power Test Report

MPC9E

Introduction

Juniper Networks commissioned EANTC to test the MPC9E and MPC8E, state of the art line cards for the MX 2000 Series 3D Universal Edge Router family. Our testing validated the performance, scale and energy efficiency of the MPC9E and MPC using a detailed and reproducible test plan. All tests were executed at Juniper's labs in Sunnyvale, California, US in April 2016.

Executive summary

The MPC9E exhibited impressive data throughput with 1,6Tbit/s forwarded in unicast and multicast scenarios. The MPC9E also showed highly-respectable average latencies in a unicast test scenario with IPv4, IPv6 and a mix of IPv4 and IPv6 traffic, as well as excellent FIB scale and performance with 10,000,000 routes in both IPv4 and IPv6 scenarios. Additionally, Juniper exhibited good power efficiency in a non-ATIS compliant test configuration, and demonstrated a queue monitor feature that enables operators to monitor queuing delays and identify the root cause of end-to-end delay issues without impacting routing or forwarding performance.

Tested devices and equipment

We tested the Juniper MX2020 3D Universal Edge Router with MPC8E and MPC9E line cards running Junos OS version 15.1I. These cards are both powered by four programmable Trio chip-set-based packet forwarding engines that provide a maximum bandwidth of 200 Gbit/s and 400 Gbit/s, respectively.

Both line cards support two Multi-Rate Modular Interface Cards (MICs). Each Multi-Rate MIC has 12 modular ports that support quad small form-factor pluggable plus (QSFP+) transceivers. On the MPC9E, all 12 ports can be configured to support 4x 10GE, 1x 40GE, or 8x 100GE[1]; in our tests we used 8 ports per MIC in 100GE mode.

Juniper configured the MPC9E in Hyper Mode, which provides the best performance and throughput for common use cases. While this configuration significantly improves performance, EANTC notes that a number of features are not supported, including Virtual Chassis, legacy DPC interoperability, non-Ethernet interfaces, and subscriber-based services.

Test Highlights

- **Demonstrated full line rate (1,600 Gbit/s) performance with IPv4 traffic for packet mix and packet sizes of 128 bytes or larger.**
- **Demonstrated full line rate (1,600 Gbit/s) performance with IPv6 traffic for packet mix and packet sizes of 256 bytes or larger.**
- **Demonstrated full line rate (1,600 Gbit/s) performance with IPv4/ IPv6 for packet mix and packet sizes of 128 bytes or larger.**
- **FIB supported 10 million IPv4 and IPv6 (separately) routes.**
- **Demonstrated full line rate (1,600 Gbit/s) performance with multicast traffic (PIM-SSM) for packet sizes of 512 bytes and 1518 bytes, 99,6% of line rate with 128 bytes frames.**

1. On the MPC8E, all 12 MIC ports can be configured for 4x 10GE and 1x 40GE; up to 4 ports can be configured for 100GE; in our tests we used 2 ports per MIC in 100GE mode. The MPC8E is also software upgradeable to the MPC9E.

An Ixia XG12 test chassis was used with 100GE ports connected to the three line cards: 16 ports to each MPC9E and 2 ports to an MPC8E. IxOS version 6.90.1150.9EA and IxNetwork version 7.50.1009.20EA were used. The Juniper MX2020 and the MPC8E and MPC9E were running Junos version 15.1I for all tests except for the telemetry demo.

Topology and test traffic

In our test, we used four test topologies per different test areas as shown in Figures 1–4. For this test we did not use our custom IMIX; instead, as requested by Juniper, we used a packet mix only containing frame sizes mentioned in RFC2544 plus one jumbo frame size. The distribution of frame sizes is shown in the table below and reflects realistic load conditions:

Frame Size (Bytes)	Weight
64 (IPv4) or 78 (IPv6)	3
128	26
256	6
512	5
1024	6
1518	16
9000	1

Test results

IPv4 and IPv6 Forwarding Performance

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 (*TBD: rewording based on actual references*). However, in certain markets the ratio is two to three percent or more; enterprise values may be much higher. Therefore, an edge router's ability to concurrently handle both IPv4/IPv6 (dual stack) addresses across the same physical interface is very important. Since 2013, EANTC's standard requirement is to run all throughput tests in dual-stack scenarios; we also tested performance figures for IPv4 only and then for IPv6 only.

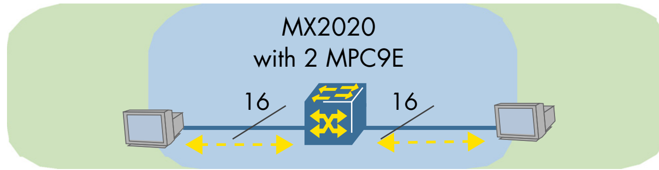


Figure 1: Forwarding Performance IPv4 & IPv6

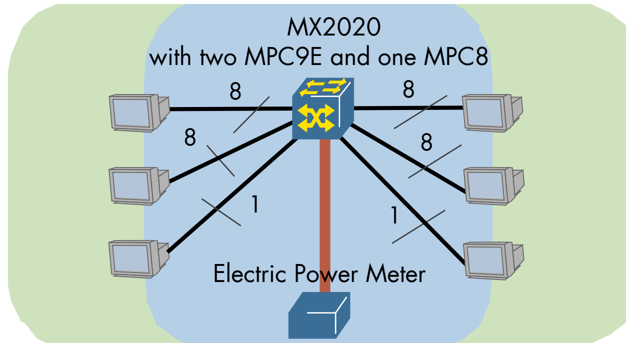


Figure 2: Power dissipation

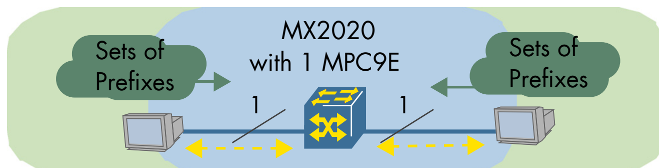


Figure 3: FIB Scalability IPv4 & IPv6

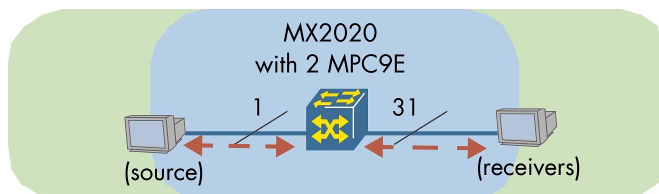
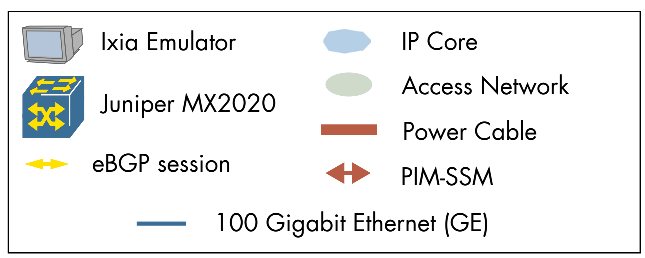


Figure 4: Multicast Performance



With IPv4 traffic, the MPC9E line card exhibited line rate performance for the packet mix and single packet sizes ranging from 128 to 9000 bytes.

The MX2020 was configured for an IPv4/IPv6 eBGP routing scenario using the topology depicted in Figure 1. As requested by Juniper, the prefix distribution was simple and used 800,000/24 prefixes for IPv4 and 32,000 /64 prefixes for IPv6. The focus for this test was to show the plain forwarding performance and not a realistic prefix distribution. The DUT forwarded IPv4 traffic, IPv6 and a mix of IPv4 and IPv6 (80:20 proportion) test traffic at full line rate of 1,600 Gbit/s.

We tested according to RFC2544 with the following standard frame sizes: 64 bytes (IPv4)/ 78 bytes (IPv6), 128 bytes, 256 bytes, 512 bytes, 1024 bytes, 1280 bytes and 1518 bytes. In addition, we used two jumbo frame sizes: 2048 bytes and 9000 bytes. We measured the throughput performance with full-mesh traffic; latency performance was measured with pair-mesh traffic.

For IPv4 traffic, the MPC9E reached full line rate throughput for frame sizes starting from 128 bytes up to 9000 bytes and for the packet mix. We measured 67.7 percent of line rate for 64 byte frames.

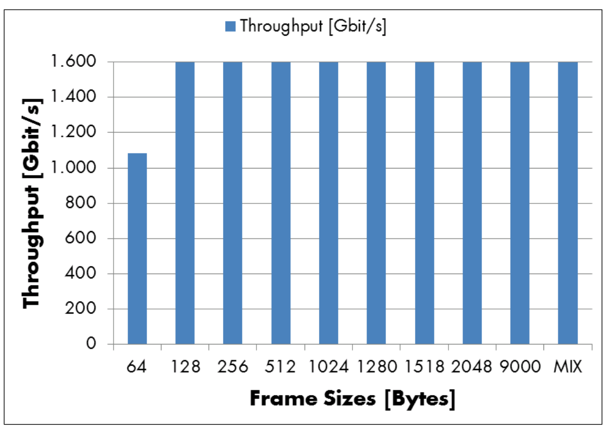


Figure 5: IPv4 Throughput

The MPC9E showed average latency values between 16 and 20 µs and maximum latency values between 28 and 57 µs for fixed packet sizes. The packet mix average latency was measured as 18 µs and the maximum latency as 52 µs.

With IPv6 traffic the MPC9E reached full line rate throughput for frame sizes starting from 256 bytes up to 9000 bytes and the packet mix. It reached 55.7 percent of line rate with 78 byte frames and 84.5 percent of line rate with 128 byte frames.

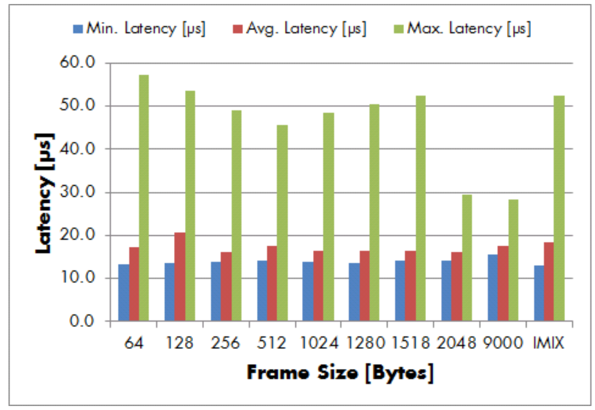


Figure 6: IPv4 Forwarding Delay

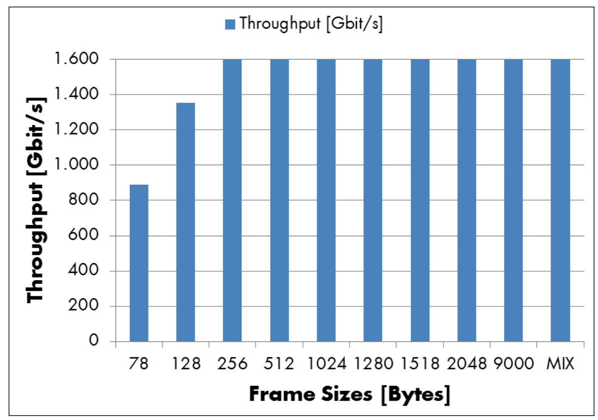


Figure 7: IPv6 Throughput

The MPC9E showed average latency values between 14 and 18 µs and maximum latency values of 21 and 54 µs for fixed packet sizes. The packet mix average latency was measured as 19 µs and the maximum latency as 56 µs.

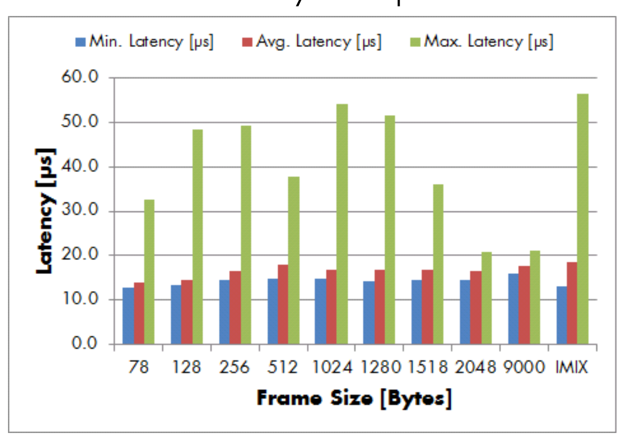


Figure 8: IPv6 Forwarding Delay

With the mix of IPv4/IPv6 traffic the MPC9E reached full line rate throughput for frame sizes starting from 128 bytes up to 9000 bytes and the packet mix. It reached 68.4 percent of line rate with the 64/78 bytes frames.

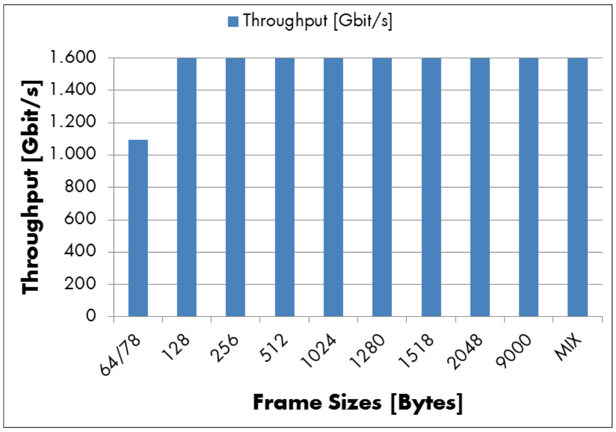


Figure 9: IPv4 & IPv6 Throughput

The MPC9E showed average latency values between 16 and 21 μs and maximum latency values of 30 and 54 μs for fixed packet sizes. The packet mix average latency was measured as 18 μs and the maximum latency as 51 μs.

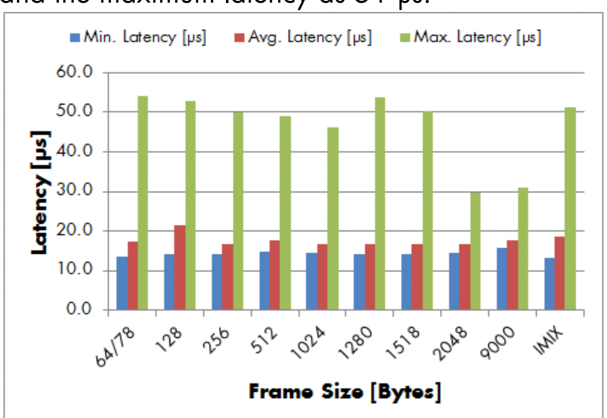


Figure 10: IPv4 & IPv6 Forwarding Delay

FIB Scalability: IPv4 and IPv6

We evaluated the scalability of the routing information base (RIB) and the forwarding information base (FIB) for IPv4 and IPv6. To ensure sufficient scale in all use cases, the maximum number of supported routes at the FIB level are a key router selection criteria. We used the topology in Figure 3; the MX2020 was configured for an IPv4/IPv6 eBGP routing scenario.

To test the scale of IPv4 route entries, the Ixia traffic generator emulated a total number of 10 million unique routes with different prefix lengths ranging from /8 to /30 as depicted in the table below. One part of the IPv4 prefixes was advertised to the router on one interface, the other part was advertised to the second interface using BGP. Memory and CPU utilization after route population was complete remained stable and were measured respectively at 40 percent and 28 percent (20 percent kernel, 6 percent user and 2 percent interrupt) during the last 5 minutes. Finally we tested the FIB limit feature, by advertising 10 percent extra prefixes (1,000,000) to the router. An alarm was properly raised by the system when the number of routes had been exceeded. As expected, the extra routes were not learned and previously learned routes were not impacted.

Start IPv4 Prefix	Prefix Length	Count
2.0.0.0	8	5
12.0.0.0	9	5
17.0.0.0	10	20
27.0.0.0	11	20
32.0.0.0	12	20
34.128.0.0	13	30
36.96.0.0	14	30
37.80.0.0	15	30
37.200.0.0	16	30
38.4.0.0	17	50
38.54.0.0	18	60
38.84.0.0	19	70
38.101.128.0	20	200
38.126.128.0	21	300
38.145.64.0	22	400
38.157.192.0	23	500
38.165.144.0	24	1,000
38.173.96.0	25	1,000
38.177.72.0	26	996,000
46.74.152.0	27	8,000,000
76.207.24.0	28	1,000,000
78.183.96.0	30	230

For the IPv6 test we used an identical configuration. The Ixia test equipment emulated a total number of 10 million unique routes with different prefix lengths from /29 to /48 as depicted in the table below. Memory and CPU utilization after FIB population was complete remained stable and were measured at 40 percent and 31 percent, respectively (21 percent kernel, 8 percent user and 2 percent interrupt) during the last 5 minutes.

Start IPv6 Prefix	Prefix Length	Count
2a00::	29	2,500,000
2c00::	32	2,500,000
2f00::	48	5,000,000

In both cases, we verified that all routes have been installed into the FIB by sending 200 Gbit/s data traffic using all routes across the MPC9E under test.

MX2020 using MPC9E successfully installed 10 million unique IPv4 routes or 10 million unique IPv6 routes on a single port into the FIB.

Multicast Performance: PIM-SSM

We tested the multicast performance of the MPC9E, as this is important for most customers. The protocol PIM-SSM is currently most deployed in multicast environments, especially because of its support for IGMPv3. This is the scenario that we tested.

PIM-SSM throughput performance with 1 source and 31 receivers over 500 groups: Line rate for 512 and 1518 byte frames, 99,6% of line rate for 128 byte frames.

We used the topology in Figure 4, with the MX2020 configured for IPv4 PIM-SSM routing. The Ixia test equipment was used to emulate IPv4 multicast from one source port to 31 receiving ports; 500 groups were configured and distributed over all receiving ports. We used frames of the following sizes: 128 bytes, 512 bytes and 1518 bytes. We generated IPv4 test traffic at 100 Gbit/s from the source port.

With PIM-SSM IPv4 traffic, the MPC9E reached full line rate throughput without any packet loss using 512 byte and 1518 byte frames; it reached 99.6 percent of line rate with128 byte frames. The router showed average latency values between 41 and 46 μs and maximum latency values between 93 and 109 μs.

Power Efficiency

We tested the power consumption of the system under test using two tests. In an initial baseline test, we measured the power consumption of a MX2020 router including the module under test (MPC9E with16x100GE) and two other line cards (MPC9E with 16x100GE and MPC8E with 2x100GE). The MPC8E was configured in snake mode with 2 ports connected to the Ixia test equipment and both MPC9Es had 16 ports connected to the Ixia test equipment. Each card had full mesh traffic configured within itself. We forwarded IPv4 test traffic at different loads and measured the power consumption for each step for 15 minutes, with measurements every second.

The test methodology was not compliant to ATIS-0600015.03.2013 which requires all tests be performed with a fully populated chassis. The MX2020 chassis has 20 slots. Juniper only had three line cards available for the test. By intention, ATIS-0600015.03.2013 specifies that these tests should be performed with a fully populated chassis as the influence of the power consumption of other components like route engine or switching fabric can not be predicted if only a few modules are used in the chassis.

In a second test, we removed the module under test from the MX2020 and measured power consumption again. Current and voltage was metered at the chassis power input; therefore measured power consumption includes all overheads, including power supply efficiency overheads. The following table summarizes the measured results. As requested by Juniper, the table also shows the delta values between both steps, the calculation does not represent the methodology defined in ATIS-0600015.03.2013.

	Power Dissipation [kW]		
Setup	0% Load	30% Load	100% Load
MX2020 with 2 MPC9E and 1 MPC8E	5.45	5.73	6.73
MX2020 with 1 MPC9E and 1 MPC8E	4.21	4.58	5.64
Delta after removing one MPC9E	1.25	1.15	1.09

NPU Telemetry, Queue Monitor Demonstration

End-to-end latency and packet loss rate are two key metrics of any packet-based connectivity service. Operators are using SLA monitoring and interface monitoring to observe those KPI parameters. Juniper's queue monitor feature enables operators to monitor queuing delays and helps them identify the root cause of end-to-end delay issues.

The queue monitor feature reports maximum queue utilization (high water mark) over configurable reporting intervals for up to 16,000 queues. One second reporting interval was used in the tests. According to Juniper, the feature is implemented completely in the packet forwarding engine itself without impacting router control plane.

Juniper demonstrated the queue monitor feature using a Juniper MX2020 3D Universal Edge Router which was equipped with two MPC9E line cards, one 100GE port on one line card was connected to the tester port emulating the access network and two 100GE ports were connected to the test ports emulating uplink connections.

The port facing the access side had 10 VLANs configured, each VLAN shaped to 100 Mbit/s, plus 8 queues per each VLAN, 95 percent of the transmit rate allocated to best effort and 5 percent to the network control traffic. IPv4 routing was used in each VLAN.

While sending traffic meshes through the device, micro bursts have been sent to produce temporary network congestion on one VLAN, but no packet loss. Each micro burst consisted only of 120 packets of 1500 bytes each, resulting in a 14.63 millisecond burst duration on the VLAN interface.

The queue monitor on the MPC9 line card sent the information about the queue depth to the collector, which was running on an open source platform called OpenNTI. OpenNTI visualized the queue depth live with a sample interval of 10 seconds, with micro bursts clearly identifiable.

The observed increase of the queue depth matched the calculated queue value based on the maximum latency measured for the burst.

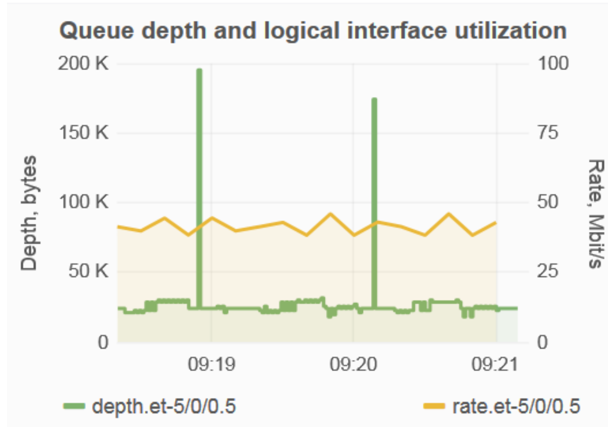
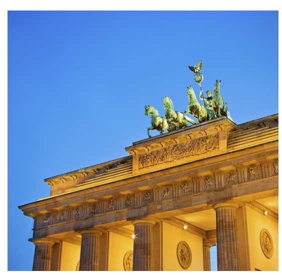


Figure 11: Queue Monitor Graph

During the demonstration we observed that neither routing engine nor line card CPU utilization changed when telemetry was enabled, as compared to the case where telemetry was disabled.

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/

v0.9 20160720