Scalability of IPv6 Transition Technologies for IPv4aaS

draft-lencse-v6ops-transition-scalability

Gábor LENCSE

lencse@sze.hu

Széchenyi István University

IETF 112, v6ops, November 11, 2021.

What is this new draft about?

- Scalability measurement results of **`iptables`** (stateful NAT44)
  - Primary goal: to serve as a sample for **stateful NAT64** tests
  - Secondary goal: they may be interesting for some network operators
- The types of measurements
  - Scalability against the number of CPU cores
  - Scalability against the number of concurrent sessions
- Using the methodology described in

  G. Lencse, K. Shima, "Benchmarking Methodology for Stateful NATxy Gateways using RFC 4814 Pseudorandom Port Numbers", Internet Draft, October 10, 2021, https://datatracker.ietf.org/doc/html/draft-lencse-bmwg-benchmarking-stateful

Scalability against the number of CPU cores

- The number of CPU cores were set to: 1, 2, 4, 8, 16
- Quantities measured:
  - Maximum connection establishment rate
  - Throughput (using bidirectional traffic), as recommended by RFC 8219
- Parameters
  - Dell PowerEdge R430 servers
    - 2.1GHz Intel Xeon E5-2683 v4 CPUs
    - 384GB 2400MHz DDR4 RAM
    - Intel 10G dual port X540 network adapters
  - Debian 9.13 Linux operating system with 4.9.0-16-amd64 kernel

Scalability against the number of CPU cores

num. cores	1	2	4	8	16
cps median	**223.5**	**371.1**	**708.7**	**1,341**	**2,383**
cps min	221.6	367.7	701.7	1,325	2,304
cps max	226.7	375.9	723.6	1,376	2,417
cps rel. scale up	*1*	*0.830*	*0.793*	*0.750*	*0.666*
throughput med.	**414.9**	**742.3**	**1,379**	**2,336**	**4,557**
throughput min	413.9	740.6	1,373	2,311	4,436
throughput max	416.1	746.9	1,395	2,361	4,627
tp. rel. scale up	*1*	*0.895*	*0.831*	*0.704*	*0.686*

Units: 1,000 connections per second, 1,000 frames per second

Scalability against the number of concurrent sessions

- The number of concurrent sessions was set with the source port number destination port number ranges
  - Usually increased fourfold, except the last case (due to memory limit)
  - The hash table size was usually increased proportionally, except the last two cases (due to memory limit)
  - NUMA issue influenced the last measurement
    - The connection tracking table did not fit into the NUMA local memory
- Quantities measured:
  - Maximum connection establishment rate
  - Throughput (using bidirectional traffic), as recommended by RFC 8219

Scalability against the number of concurrent sessions

num. conn.	1.56M	6.35M	25M	100M	400M	*800M*
hash table size	2^21	2^23	2^25	2^27	*2^28*	*2^28*
n.c./h.t.s.	0.745	0.745	0.745	0.745	*1.490*	*2.980*
cps median	**2.406**	**2.279**	**2.278**	**2.237**	**2.013**	**1.405**
cps min	2.358	2.226	2.226	2.124	1.983	1.390
cps max	2.505	2.315	2.317	2.29	2.05	1.440
throughput med.	**5.326**	**4.369**	**4.51**	**4.516**	**4.244**	**3.689**
throughput min	5.217	4.240	3.994	4.373	4.217	3.670
throughput max	5.533	4.408	4.572	4.537	4.342	3.709

Units: 1,000,000 connections per second, 1,000,000 frames per second

My questions

- Are these results useful and sufficient for the network operators?
- Will the same method be appropriate for the stateful NAT64 scalability measurements?
- Is Jool (https://jool.mx) a good choice of implementation to test?