Scalability of IPv6 Transition Technologies for IPv4aaS

draft-lencse-v6ops-transition-scalability

Gábor LENCSE

lencse@sze.hu

Széchenyi István University

IETF 113, v6ops, March 21, 2021.

What new results are in this draft?

- Scalability measurement results of **Jool** stateful NAT64 solution
  - Jool was recommended at IETF 112
  - Open source stateful NAT64 implementation: https://www.jool.mx/en/
- 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, March 4, 2022, 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 NIC)
  - Debian 10.11 Linux operating system with 4.19.0-18-amd64 kernel
  - Jool 4.1.6 (Release date: 2021-12-10)

Scalability against the number of CPU cores

num. cores	1	2	4	8	16
cps median	**228.6**	**358.5**	**537.4**	**569.9**	**602.6**
cps min	226.5	352.5	530.7	562.0	593.7
cps max	230.5	362.4	543	578.3	609.7
cps rel. scale up	1	0.784	0.375	0.133	0.066
throughput median	**251.8**	**405.7**	**582.4**	**604.1**	**612.3**
throughput min	249.8	402.9	573.2	587.3	599.8
throughput max	253.3	409.6	585.7	607.2	616.6
tp. rel. scale up	1	0.806	0.359	0.130	0.063

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
  - Increased fourfold (usually by doubling the size of both ranges)
  - Unlike previously with **`iptables`**, no tuning was done with Jool
- Quantities measured:
  - Maximum connection establishment rate
  - Throughput (using bidirectional traffic), as recommended by RFC 8219
  - Connection tear down rate -- NEW metric!

Scalability against the number of concurrent sessions

num. conn.	1.56M	6.35M	25M	100M	400M	**1600M**
src ports	2,500	5,000	10,000	20,000	40,000	**40,000**
dst ports	625	1,250	2,500	5,000	10,000	**40,000**
cps median	**480.2**	**394.8**	**328.6**	**273.0**	**243.0**	**232.0**
cps min	468.6	392.7	324.9	269.4	243.0	230.5
cps max	484.9	397.4	331.3	280.6	244.5	233.6
throughput med.	**511.5**	**423.9**	**350.0**	**286.5**	**257.8**	**198.4**
throughput min	509.2	420.3	348.2	284.2	257.8	195.3
throughput max	513.1	428.3	352.5	290.8	260.9	201.6

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

Connection tear down rate measurements

- Our Internet-Draft recommends an aggregate measurement:

$$Connection\ tear\ down\ rate = \frac{number\ of\ connections\ in\ the\ connection\ tracking\ table}{deletion\ time\ of\ the\ entire\ connection\ tracking\ table}$$

- To be measured using different number of connections
  - Details can be found in
    https://datatracker.ietf.org/doc/html/draft-lencse-bmwg-benchmarking-stateful-03#section-4.7
- Further technical refinement
  - Subtract the deletion time of an empty table
    - It counts when low number of connections are used
    - Also eliminates command execution overhead

Connection tear down rate of Jool

num. conn.	1.56M	6.35M	25M	100M	400M	**1600M**
src ports	2,500	5,000	10,000	20,000	40,000	**40,000**
dst ports	625	1,250	2,500	5,000	10,000	**40,000**
full cont. del med	0.87	2.05	7.84	36.38	126.09	474.68
full cont. del min	0.80	2.02	7.80	36.27	125.84	473.20
full cont. del max	0.91	2.09	7.94	36.80	127.54	481.38
emp. ct. del med	0.46	0.46	0.46	0.46	0.46	0.46
conn. del. time	0.41	1.59	7.38	35.92	125.63	474.22
conn. t. d. r. (M)	**3.811**	**3.931**	**3.388**	**2.784**	**3.184**	**3.374**

Units: *seconds* for time; *1,000,000connections/s* for connection tear down rate

Connection tear down rate of `iptables`

num. conn.	1.56M	6.35M	25M	100M	400M	**800M**
src ports	2,500	5,000	10,000	20,000	40,000	**40,000**
dst ports	625	1,250	2,500	5,000	10,000	20,000
conntrack t. s.	2^{21}	2^{23}	2^{25}	2^{27}	2^{29}	**2^{30}**
hash table size	2^{21}	2^{23}	2^{25}	2^{27}	**2^{28}**	**2^{28}**
full cont. del med	4.33	18.05	74.47	305.33	1178.3	2263.1
emp. ct. del med	0.55	1.28	4.17	15.74	31.2	31.2
conn. del time	3.78	16.77	70.30	289.59	1147.2	2232.0
conn. tear d. rate	**413.4**	**372.7**	**355.6**	**345.3**	**348.7**	**358.4**

Units: *seconds* for time; *1,000connections/s* for connection tear down rate

My questions

- Any comments for the connection tear down rate measurement method?
  - E.g. a better idea?
- Are these results <u>useful</u> and <u>sufficient</u> for the network operators?
  - What other metrics would be useful?
- Is there an interest in the working group for further results?
  - Potential adoption of the draft?