SZÉCHENYI EGYETEM
UNIVERSITY OF GYŐR

Doctoral School of Multidisciplinary Engineering Sciences (MMTDI)

Performance Analysis of ICT Systems

Lecture 4

Benchmarking Methodology for Network Interconnect Devices
RFC 2544, RFC 5180 and RFC 8219

https://www.tilb.sze.hu/cgi-bin/tilb.cgi?0=m&1=targyak&2=NGD_MDA64_1

Dr. Gábor Lencse

Professor

Dept. of Telecommunications, University of Győr

lencse@sze.hu

Agenda

- Introduction to the RFC-s
- Benchmarking Methodology for Network Interconnect Devices
  - Based on RFC 2544, RFC 5180 and RFC 8219
- Research and publication possibilities

What are the RFC-s?

- RFC: Request for Comments
  - Primary documents about networking protocols, application, measurement methods
  - Identified by their RFC numbers
  - Published by IETF (Internet Engineering Task Force) after a long process
    - All the details can be found at: https://www.ietf.org/standards/process/

Types of RFC-s

- Standards Track:
  - Old: three phases
    - Proposed Standard --> Draft Standard --> Standard
    - It proved to be too many
  - New: only two phases (RFC 6410)
    - Proposed Standard --> Internet Standard
- Also exist Informational, Experimental RFC-s
  - For example, the Benchmarking Working Group produces only Informational RFC-s
    - They are still ‘de facto’ standards.
  - And there are RFC-s with the date of April 1 ☺

How an RFC is born?

- Someone invents something and writes an *individual Internet Draft* (I-D)
  - There is a required structure, style and format
  - Upload: https://datatracker.ietf.org/submit/
  - Its filename looks like:
    - **`draft-author-intendedwg-some-words-00`**
  - An e-mail is sent to the mailing list of the intended working group
    - If someone is interested in: a discussion is started. ☺
    - If no one is interested in: the draft will be forgotten ☹
      - A draft expires in 6 months, new version can be submitted, etc.

How an RFC is born?

- There are IETF meetings 3 times a year
  - A draft may be presented in a working group if the WG chair gives prior permission
  - Responses during the meeting and on the mailing list
  - Once a working group adopts a draft, it becomes a *working group draft*, and its name looks like:
    - **`draft-ietf-wgname-some-words-00`**
    - The “-00” version must be approved by a WG chair
  - If there is a rough consensus in the working group, then there is a WGLC (WG Last Call)
    - It may happen multiple times! ;-)

How an RFC is born?

- A few further reviews follow…
  - Further new versions with minor modifications may be produced
  - When everything is finished, an RFC number is assigned, and no more changes are possible
    - An *Errata* may be added
    - May be *obsoleted by* or *updated by* other RFC-s
  - RFC-s are worth reading in HTML format, as it contains information, if the RFC was *obsolated* or *updated* by other RFC-s, or if *errata exists*

Personal experience

- Three examples
  - Marius Georgescu invited me as a co-author in 2015:
    - https://tools.ietf.org/html/draft-ietf-bmwg-ipv6-tran-tech-benchmarking-00
    - Published as RFC 8219 in 2017
  - We tried as strangers in 2017
    - https://tools.ietf.org/html/draft-lencse-tsvwg-mpt-00
  - I tried on the v6ops mailing list in 2018
    - https://tools.ietf.org/html/draft-lmhp-v6ops-transition-comparison-00
    - One new co-author, it may be a long lasting task…

Lessons learned

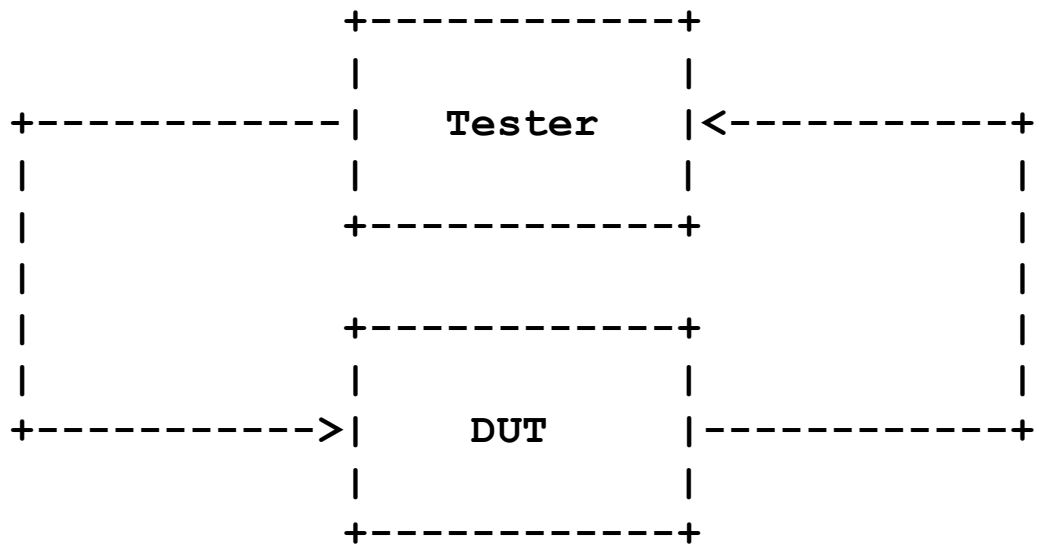
- Creating an RFC requires a lot of work an time
  - An individual I-D is not easily published as an RFC
    - Especially, if the authors are not known in the WG
  - If the WG adopts it, chances are much better
    - Key factor: the reaction of the prominent WG members
    - It is worth
      - Writing together with honored WG members
      - Being active in the WG, giving useful comments to others' drafts
      - Testing at the proper time, if there is support in the WG

RFC 2544

- Aim: to measure the performance of network interconnect devices in an objective way
  - Defines the most important aspects of the measurement to prevent gaming
    - Measurement setup
    - DUT (Device Under Test) settings
      - May NOT be optimized for the given task! ☺
    - Frame format, frame sizes
      - E.g. for Ethernet: 64, 128, 256, 512, 1024, 1280, 1518
    - Duration of the test (min. 60s)
    - … (and several others)

Measurement Setup

- Three measurement setups were defined
  - By default, the first one should be used



  - Although the arrows are unidirectional, bidirectional traffic should be used

Throughput measurement procedure

- DEFINTION: Througput
  - The maximum (constant) frame rate, at which there is no frame loss.
- Measurement procedure (Throughput)
  - The Tester transmits frames at a given frame rate for a given time (min 60s) through the DUT, and the Tester checks if the frames arrive back
    - If all frames arrived back, the frame rate is increased
    - If some frames were lost, the frame rate is decreased
  - In practice, it is worth using a binary search
    - There may be linear steps, some “hints” can be given, too.

Latency Measurement Procedure

- Measurement procedure (Latency)
  - Use the frame rate determined by the Througput test
  - Use an at least 120s long stream
  - Mark a frame with an identifying tag after 60s
  - Timestamp “A”: taken after the tagged frame is completely transmitted
  - Timestamp “B”: taken after the tagged frame is completely received
  - Latency: B-A
  - The measurement must be performed at least 20 times, and the final result is the *average* of the values

Frame Loss Rate Test

- Measurement Procedure (Frame Loss Rate)
  - Start at the maximum frame rate for the media
  - Decrease the frame rate in fixed steps (max. 10%)
  - Measure the frame loss at the given frame rate
    - What percent of the sent frames was lost?
  - Stop, if no frame loss in two consecutive measurements

Further Measurement Procedures

- Back-to-back frames
  - How long the DUT can operate at the maximum frame rate of the media?
- System recovery
  - After overload situation, how long time is needed until lossless operation is restored
- Reset
  - After hardware/software reset (or power outage), how long time is needed until lossless operation is restored?

RFC 2544 versus RFC 5180

- Az RFC 2544 is IP version independent (in principle)
  - But focuses on IPv4, e.g. IP addresses: 192.18.0.0/15
  - Mentioned media types show its age
- RFC 5180 is a kind of “update”
  - In principle, focuses on IPv6 specific issues
    - E.g. IP addresses: 2001:2::/48
  - Also technology update: newer media types
  - Declares IPv6 transition technologies *out of scope*

RFC 8219 – only in short

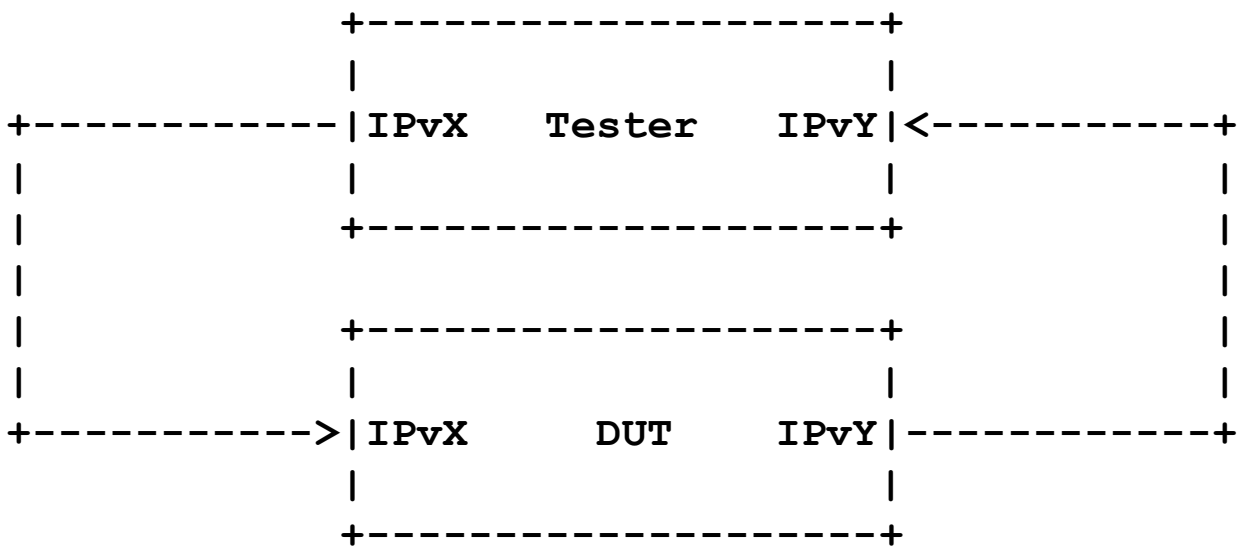
- There are a high number of IPv6 transition technologies
  - It is not practical to handle them one-by-one ☹
- Classifies them into small number of categories
  - Only the categories are to be dealt with ☺
  - Future technologies may also fit into them ☺
  - DNS64 does not fit into any of them
    - There is a separate section for DNS64

Categories of IPv6 trans. technologies

	Generic category	IPv6 Transition Technology
1	Dual-stack	Dual IP Layer Operations [RFC4213]
2	Single translation	NAT64 [RFC6146], IVI [RFC6219]
3	Double translation	464XLAT [RFC6877], MAP-T [RFC7599]
4	Encapsulation	DSLite[RFC6333], MAP-E [RFC7597]
		Lightweight 4over6 [RFC7596]
		6RD [RFC5569], 6PE [RFC4798], 6VPE
		6VPE [RFC4659]

Measurement Setup – 1

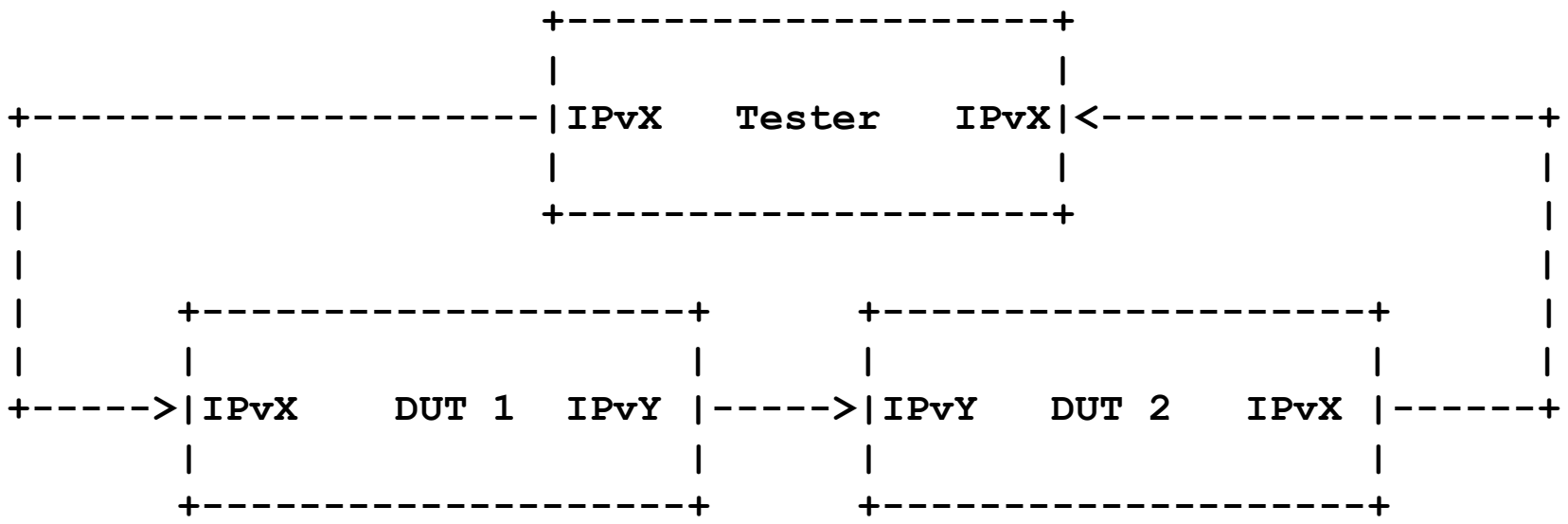
- For single translation technologies
  - RFC 2544 setup is modified as follows*:*



  - Remark: IPvX and IPvY: X=6 and Y=4 *or:* X=4 and Y=6
  - We call it as “single DUT test setup”

Measurement Setup – 2

- For double translation / encapsulation technologies
  - Two kinds of test setups can be used:
    - The before mentioned “single DUT test setup”
    - The below “dual DUT test setup”



    - Remark: the two devices are peers: the second one performs the “reverse transformation” of the first one

Measurement Procedures

- In part, it reuses the RFC 2544 benchmarking procedures with minor modifications
  - Througput, Frame Loss Rate
- In part, it defines better procedures
  - Latency
    - At least 500 tagged frames (instead of a single one)
    - Typical Latency: the median of the measured latencies
    - Worst Case Latency: the 99.9-th percentile of the latencies
- In part, new measurement procedures
  - PDV, IPDV

Research and Publication Possibilities

- DNS64 benchmarking methodology
  - G. Lencse, M. Georgescu, and Y. Kadobayashi, "Benchmarking Methodology for DNS64 Servers", *Computer Communications* (Elsevier), vol. 109, no. 1, pp. 162-175, September 1, 2017, DOI: 10.1016/j.comcom.2017.06.004
- DNS64 benchmarking tool
  - **`dns64perf++`**: free software, available: https://github.com/bakaid/dns64perfpp
  - G. Lencse, D. Bakai, "Design and implementation of a test program for benchmarking DNS64 servers", *IEICE Transactions on Communications*, vol. E100-B, no. 6. pp. 948-954, June 2017. DOI: 10.1587/transcom.2016EBN0007

Research and Publication Possibilities

- DNS64 measurements
  - G. Lencse and Y. Kadobayashi, "Benchmarking DNS64 Implementations: Theory and Practice", *Computer Communications* (Elsevier), vol. 127, no. 1, pp. 61-74, September 1, 2018, DOI: 10.1016/j.comcom.2018.05.005
- Benchmarking Authoritative DNS servers
  - G. Lencse, "It is worth upgrading form BIND: Performance comparison of authoritative DNS servers", IIJ Seminar, Tokyo, April 25, 2019. https://iijlab-seminars.connpass.com/event/127000/
  - G. Lencse, “Benchmarking Authoritative DNS Servers", *IEEE Access*, vol. 8. pp. 130224-130238, July 2020. DOI: 10.1109/ACCESS.2020.3009141

Research and Publication Possibilities

- Stateless NAT64 measurement tool
  - **`siitperf`**: free software, available: https://github.com/lencsegabor/siitperf
  - G. Lencse, "Design and Implementation of a Software Tester for Benchmarking Stateless NAT64 Gateways", *IEICE Transactions on Communications*, DOI: 10.1587/transcom.2019EBN0010
  - G. Lencse, "Adding RFC 4814 Random Port Feature to Siitperf: Design, Implementation and Performance Estimation", under review in *International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems*, vol 9, no 3, pp. 18-26, 2020, DOI: 10.11601/ijates.v9i3.291

Research and Publication Possibilities

- Stateless NAT64 measurements
  - G. Lencse and K. Shima, "Performance Analysis of SIIT Implementations: Testing and Improving the Methodology", *Computer Communications* (Elsevier), vol. 156, no. 1, pp. 54-67, April 15, 2020, DOI: 10.1016/j.comcom.2020.03.034
- Further publications available
  - http://www.hit.bme.hu/~lencse/publications/

Possiblities for doing research abroad

- I was a guest researcher
  - *Laboratory of Cyber Resilience*, NAIST, Ikoma, Japan
    June 15 – December 15, 2017.

  - *IIJ Innovation Institute*, Tokyo, Japan
    April 2 – July 2, 2019.

- I am glad to help, if someone is interested!

Thank you for your attention!

Questions?

Dr. Gábor Lencse

Professor

Dept. of Telecommunications, University of Győr

lencse@sze.hu