



**FAKULTA
ELEKTROTECHNICKÁ
ČVUT V PRAZE**

iPerf aneb Měřit pouze maximální propustnost je nesmysl, pojďme to dělat lépe

CSNOG 2024, 23.1.2024

Zbyněk Kocur
Ondřej Vondrouš
Ondřej Votava
Josef Šebánek
[f-tester@fel.cvut.cz]

iPerf

Server address: 192.168.1.210

Server port: 5201

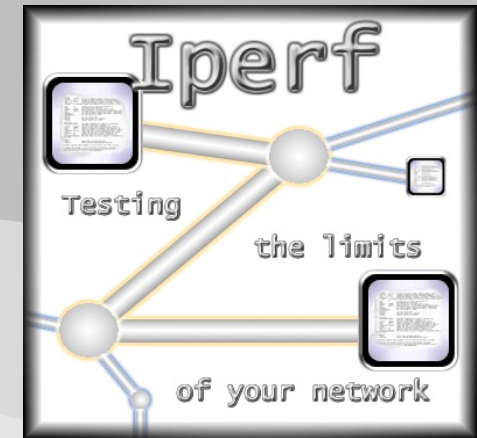
Transmit mode: Upload Download

Streams: 1 2 3 4 5

Test duration: 10s 30s 5 min

135 Mbits/s

avg: 123 min: 94 max: 135



18:04

iPerf

192.168.1.5 START

Intervall sek	Retr	Cwnd Bytes	Übertragung Bytes	Bandweite bit/sek
1	91	212.0K	10.6M	88.8M
2	34	124.0K	11.1M	92.8M
3	19	134.0K	10.3M	86.0M
4	144	74.9K	9.57M	80.3M
5	73	136.0K	9.32M	78.2M
Empfänger-Zusammenfassung 0.0-5.0			50.8M	85.2M
Sender-Zusammenfassung 0.0-5.04			49.5M	82.5M

iPerf

- Open Source projekt pro měření vlastností sítí
 - Windows, Linux, Android, MacOS X, FreeBSD, OpenBSD, NetBSD, VxWorks, Solaris, ...
- Existuje iPerf2 a iPerf3
- Architektura klient – server
- TCP, SCTP
 - MSS/MTU, velikost okna, dosažený datový tok
 - Nastavitelný congestion algoritmus, omezený datový tok, ...
- UDP
 - Počet datagramů, ztrátovost, jitter
 - Nastavitelný omezený datový tok, velikost datagramu, ...
- Dopředný a zpětný režim (upload, download)
- Paralelní toky

iPerf2 / iPerf3

- Jaké jsou hlavní rozdíly?
 - CSV / JSON
 - Server obslouží více klientů současně / server komunikuje pouze s jedním klientem
 - Parametry na straně klienta i serveru / dynamický server – výměna parametrů testu při zahájení spojení
 - Informace o spojení pouze z klientské strany / informace o spojení z klienta i serveru
 - Obousměrný test v jednom spojení / pro dva směry musí být dvě spojení
 - Měření TCP sestavení spojení / nezapočítání TCP slowstart do analýzy
 - Vlákna / procesy

Maximální propustnost TCP spojení – BDP

- maximální propustnost TCP spojení je dána vztahem

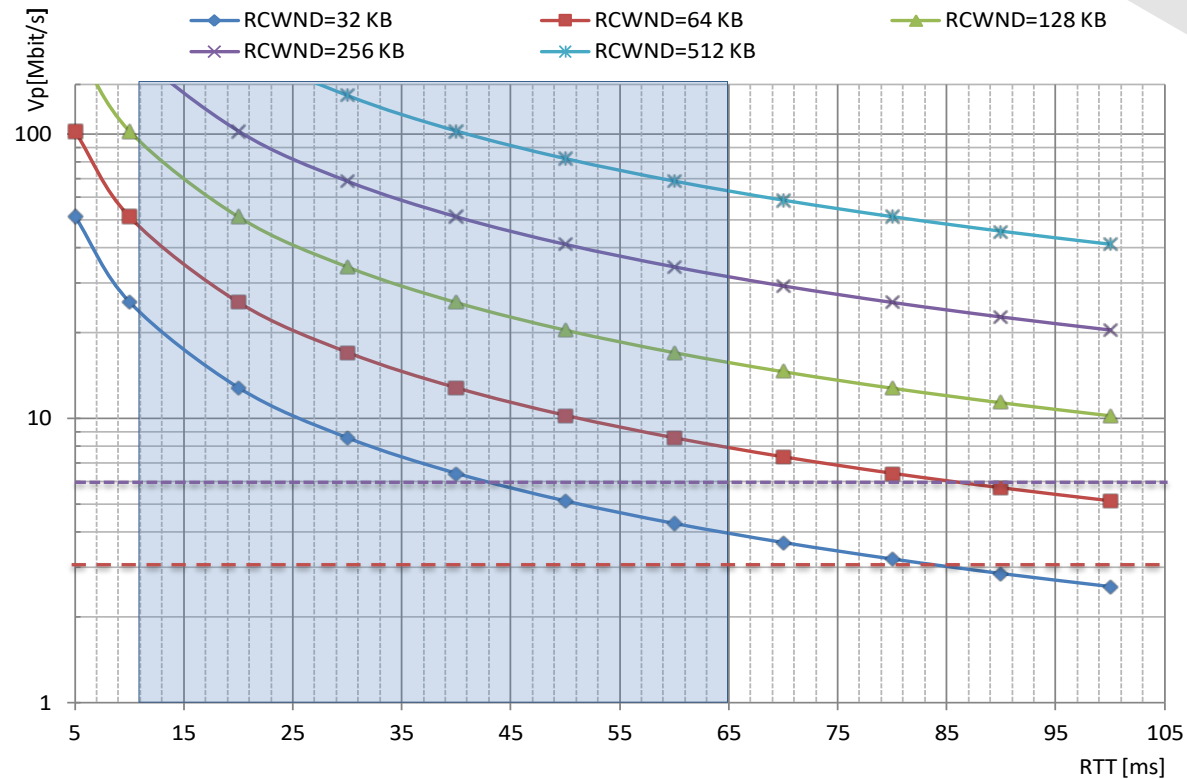
$$TCP_{TP} = \frac{8 * WS}{RTT} \quad [\text{bit/s}]$$

$$WS = TCP_{TP} * RTT \quad [B]$$

- kde: **WS** – velikost okna na straně vysílače a **RTT** – „Round Trip Time“ - čas, který uplyne od vyslání paketu k jeho příjmu a přijetí potvrzení, nazýváno také jako **Bandwidth Delay Product (BDP)**

Maximální propustnost TCP spojení – BDP

- Omezení propustnosti závislé na zpoždění (RTT) a velikosti okna pro potvrzování datových segmentů
- Navýšení rychlosti zvýšením počtů paralelních TCP spojení



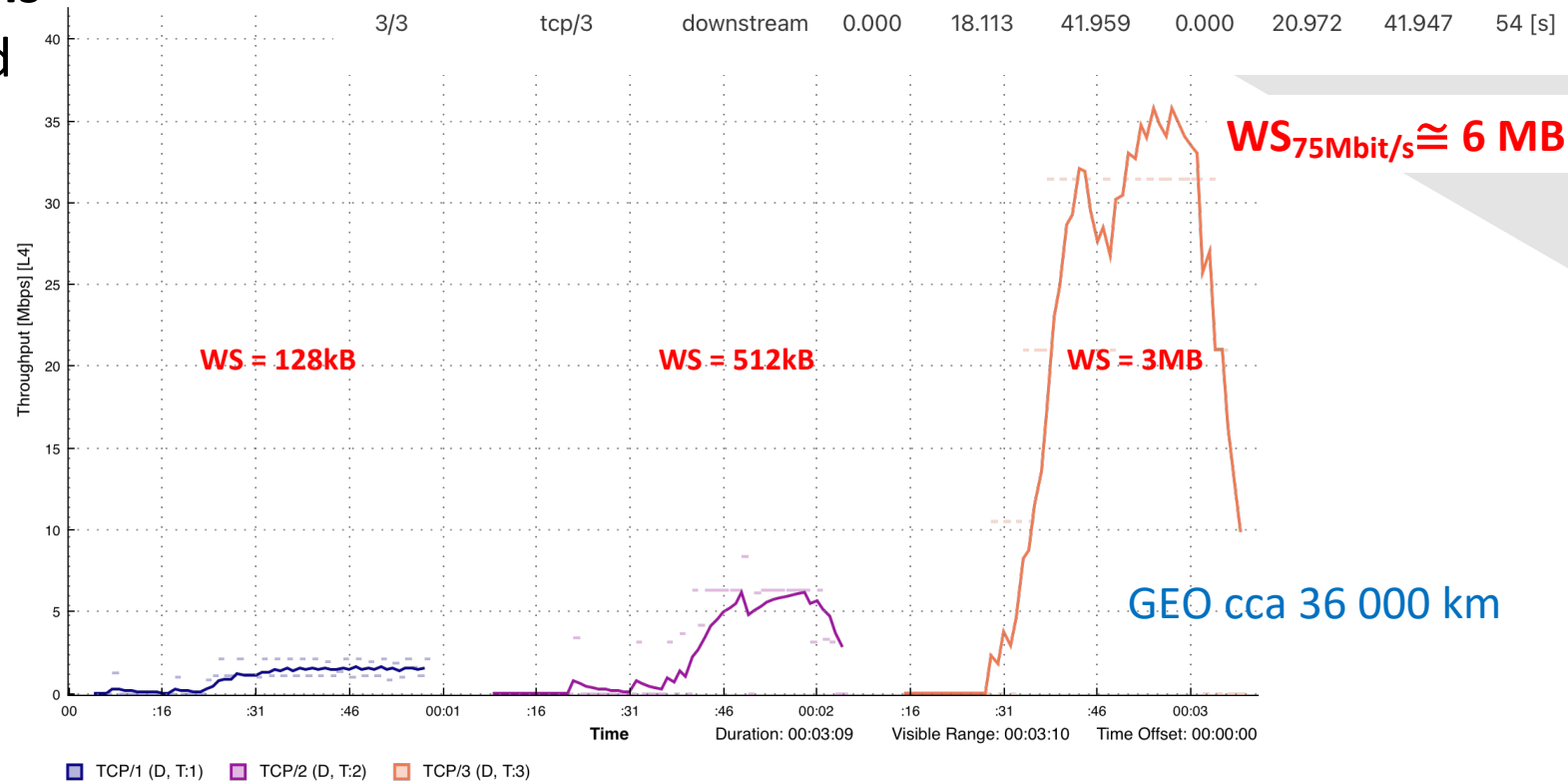
$$N \geq \frac{BDP}{8 \cdot RWND}$$

Maximální propustnost TCP spojení – BDP

- Satelitní služba Eutelsat
- Dostupná rychlost služby **75 Mbit/s**
- Průměrné RTT **600 ms**
- TCP Cubic, download

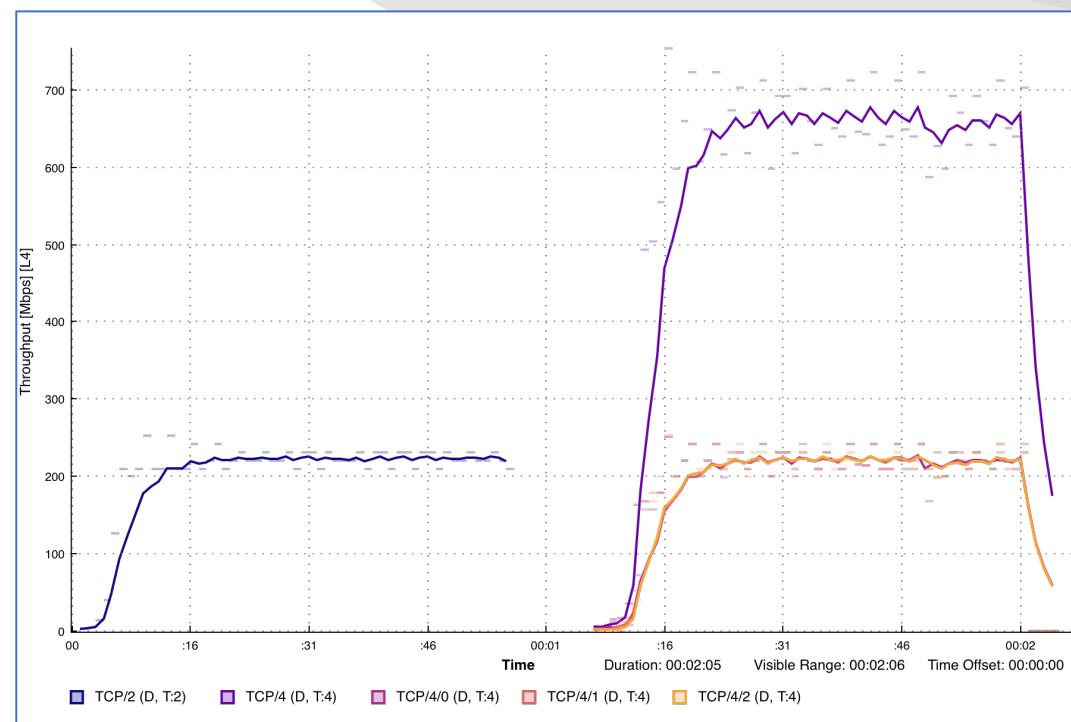
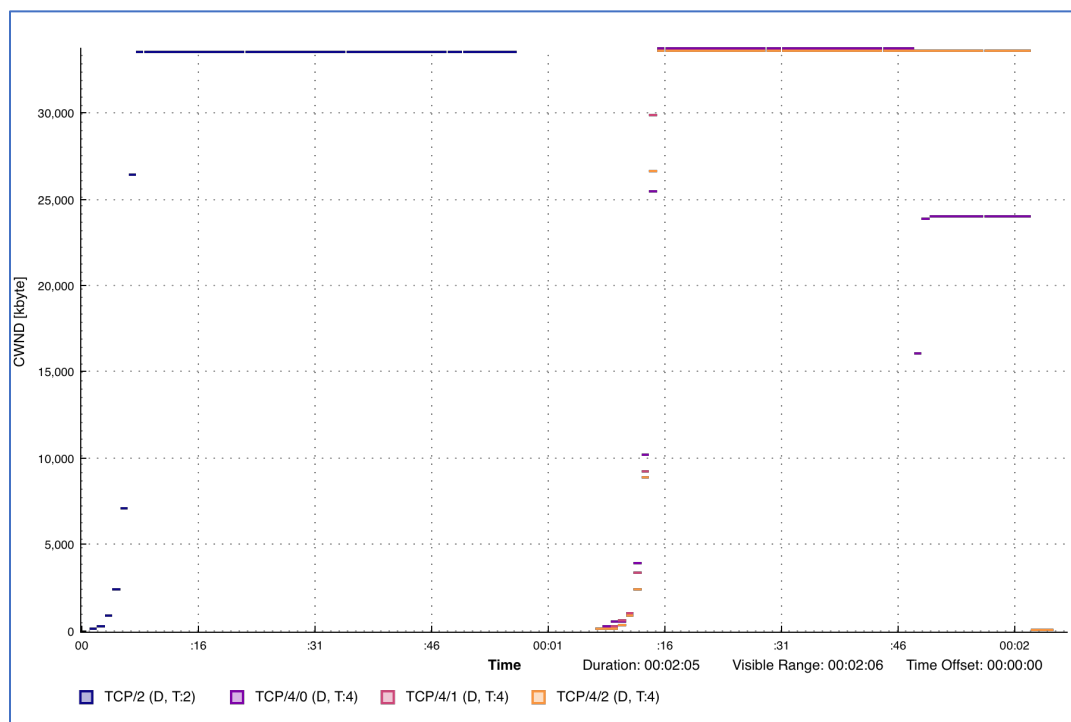
Throughput [Mbit/s] [L4]

Test / File ID	Stream ID	Direction	Min TP	Avg TP	Max TP	5%ile	50%ile	95%ile	Duration	Transferred
1/1	tcp/1	downstream	0.000	1.023	2.089	0.000	1.043	2.088	53 [s]	6.9 [MB]
2/2	tcp/2	downstream	0.000	2.499	8.370	0.000	0.000	6.309	56 [s]	17.8 [MB]
3/3	tcp/3	downstream	0.000	18.113	41.959	0.000	20.972	41.947	54 [s]	124.5 [MB]



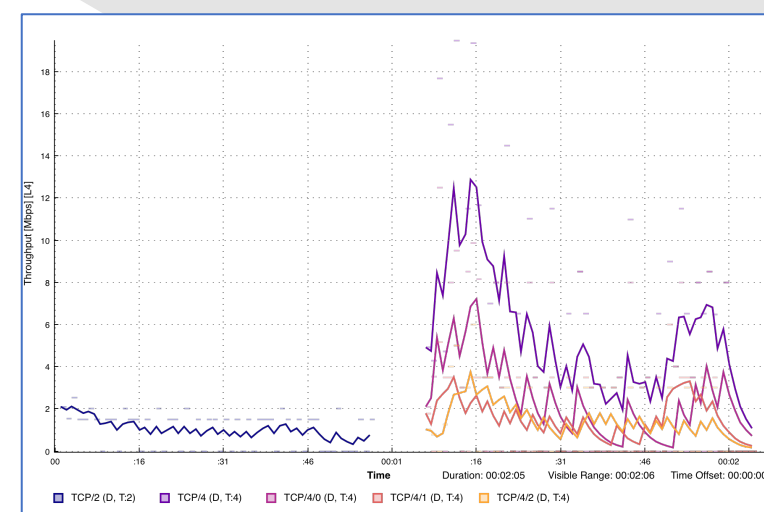
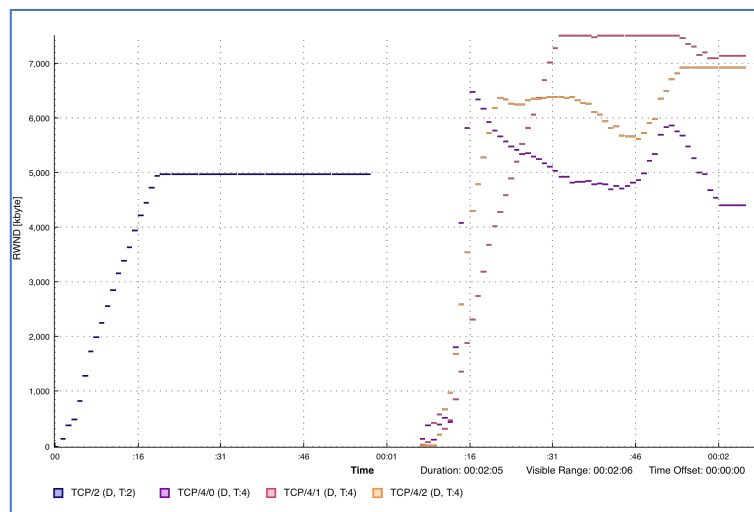
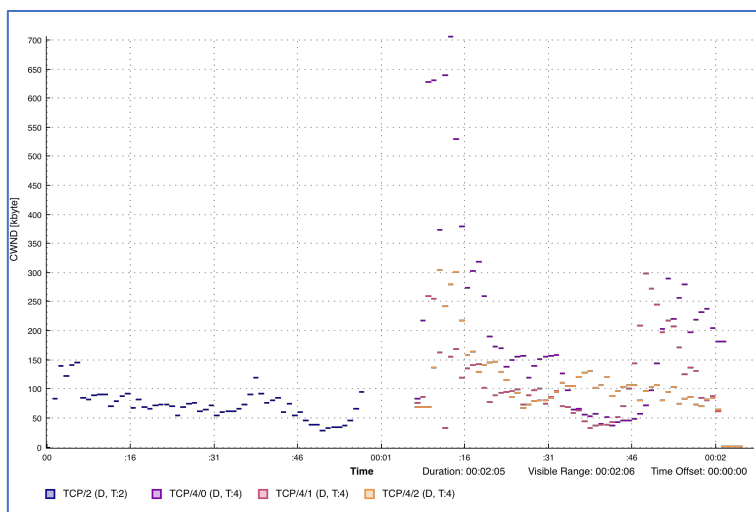
Maximální propustnost TCP spojení – BDP

- iPerf3, TCP Cubic, MSS 1400 B, Auto Window Size, RTT 600 ms, PLR 0 %
- **Limitem je paměť, pps, nebo RWND**



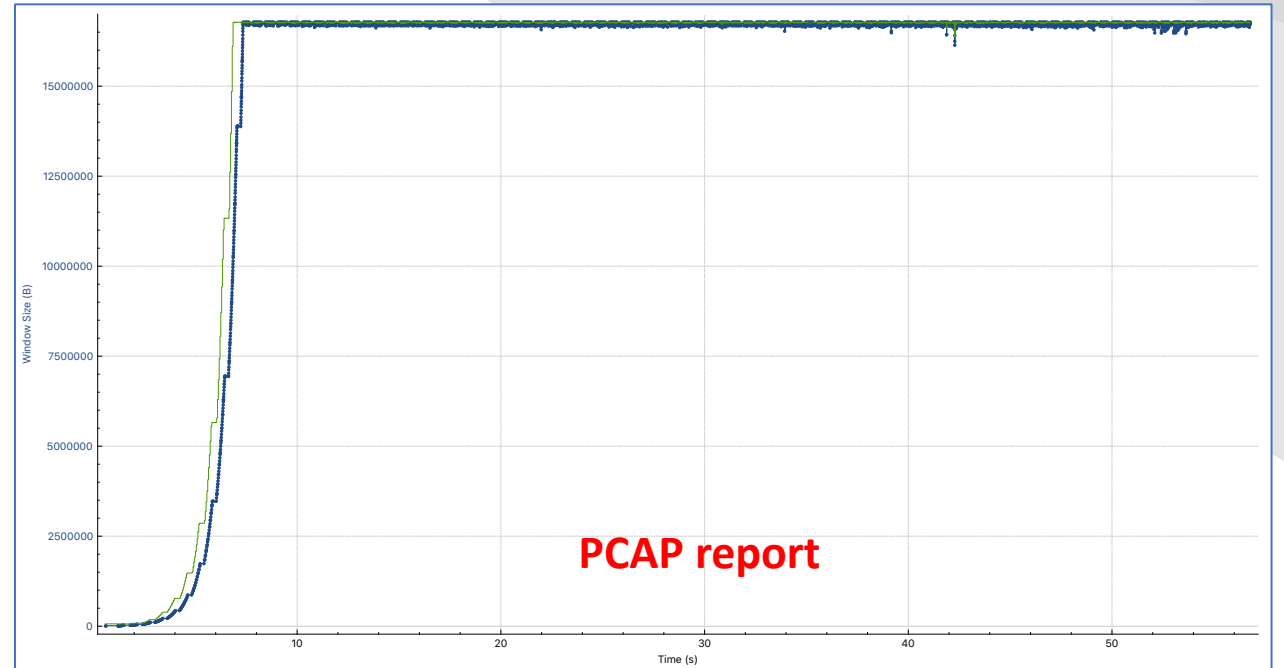
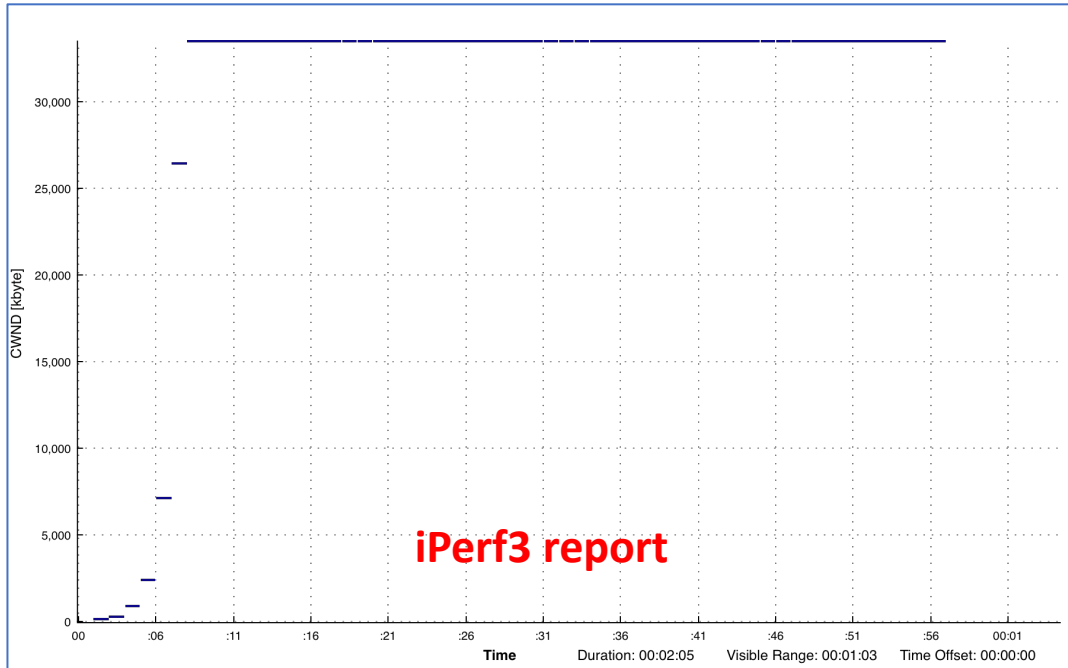
Maximální propustnost TCP spojení – BDP

- iPerf3, TCP Cubic, MSS 1400 B, Auto Window Size, RTT 600 ms, PLR 0.25 %
- **Limitem je síť, případně použitý algoritmus cubic**



iPerf3 CWND report problém

- iPerf3, TCP Cubic, MSS 1400 B, Auto Window Size, RTT 600 ms, PLR 0 %
- **Špatné reportování velikosti CWND**, RWND je uváděno v pořádku



iPerf3 – výstup CLI

- iPerf3, TCP Cubic, MSS 1448 B, Auto Window Size, RTT 600 ms, PLR 0 %

```
root@cli1:~# iperf3 -c 10.44.0.13
Connecting to host 10.44.0.13, port 5201
[ 5] local 10.44.0.12 port 38492 connected to 10.44.0.13 port 5201
[ ID] Interval           Transfer     Bitrate      Retr  Cwnd
[ 5]  0.00-1.00    sec      160 KBytes  1.31 Mbits/sec    0   31.1 KBytes
[ 5]  1.00-2.00    sec     93.3 KBytes  765 Kbits/sec    0   74.9 KBytes
[ 5]  2.00-3.00    sec     669 KBytes  5.48 Mbits/sec    0   199 KBytes
[ 5]  3.00-4.00    sec     1.93 MBytes 16.2 Mbits/sec    0   608 KBytes
[ 5]  4.00-5.00    sec     4.92 MBytes 41.3 Mbits/sec    0   1.88 MBytes
[ 5]  5.00-6.00    sec     16.2 MBytes 136 Mbits/sec    0   5.91 MBytes
[ 5]  6.00-7.00    sec     16.2 MBytes 136 Mbits/sec    0   18.7 MBytes
[ 5]  7.00-8.00    sec     26.2 MBytes 220 Mbits/sec    0   45.6 MBytes
[ 5]  8.00-9.00    sec     26.2 MBytes 220 Mbits/sec    0   46.4 MBytes
[ 5]  9.00-10.00   sec     26.2 MBytes 220 Mbits/sec    0   33.6 MBytes
-----
[ ID] Interval           Transfer     Bitrate      Retr
[ 5]  0.00-10.00    sec     119 MBytes 99.8 Mbits/sec    0      sender
[ 5]  0.00-10.61   sec     114 MBytes 90.1 Mbits/sec    receiver

iperf Done.
```

iPerf3 – výstup CLI

- iPerf3, TCP Cubic, MSS 1448 B, Auto Window Size, RTT 600 ms, PLR 0,1 %

```
root@cli1:~# iperf3 -c 10.44.0.13
Connecting to host 10.44.0.13, port 5201
[ 5] local 10.44.0.12 port 57276 connected to 10.44.0.13 port 5201
[ ID] Interval           Transfer     Bitrate      Retr  Cwnd
[ 5]  0.00-1.00    sec    129 KBytes  1.05 Mbits/sec    0   31.1 KBytes
[ 5]  1.00-2.00    sec    204 KBytes  1.67 Mbits/sec    0   74.9 KBytes
[ 5]  2.00-3.00    sec    789 KBytes  6.46 Mbits/sec    0   204 KBytes
[ 5]  3.00-4.00    sec   1.24 MBytes 10.4 Mbits/sec    0   618 KBytes
[ 5]  4.00-5.00    sec   2.29 MBytes 19.2 Mbits/sec    4   970 KBytes
[ 5]  5.00-6.00    sec   5.00 MBytes 41.9 Mbits/sec    0   2.19 MBytes
[ 5]  6.00-7.00    sec   3.75 MBytes 31.5 Mbits/sec    2   2.30 MBytes
[ 5]  7.00-8.00    sec  15.0 MBytes 126 Mbits/sec    3   7.27 MBytes
[ 5]  8.00-9.00    sec  10.0 MBytes 83.9 Mbits/sec   12  20.9 MBytes
[ 5]  9.00-10.00   sec   8.75 MBytes 73.4 Mbits/sec    9  27.7 MBytes
-----
[ ID] Interval           Transfer     Bitrate      Retr
[ 5]  0.00-10.00    sec   47.1 MBytes 39.5 Mbits/sec   30
[ 5]  0.00-10.75   sec   36.6 MBytes 28.6 Mbits/sec

iperf Done.
```

iPerf a měření dle RFC 6349

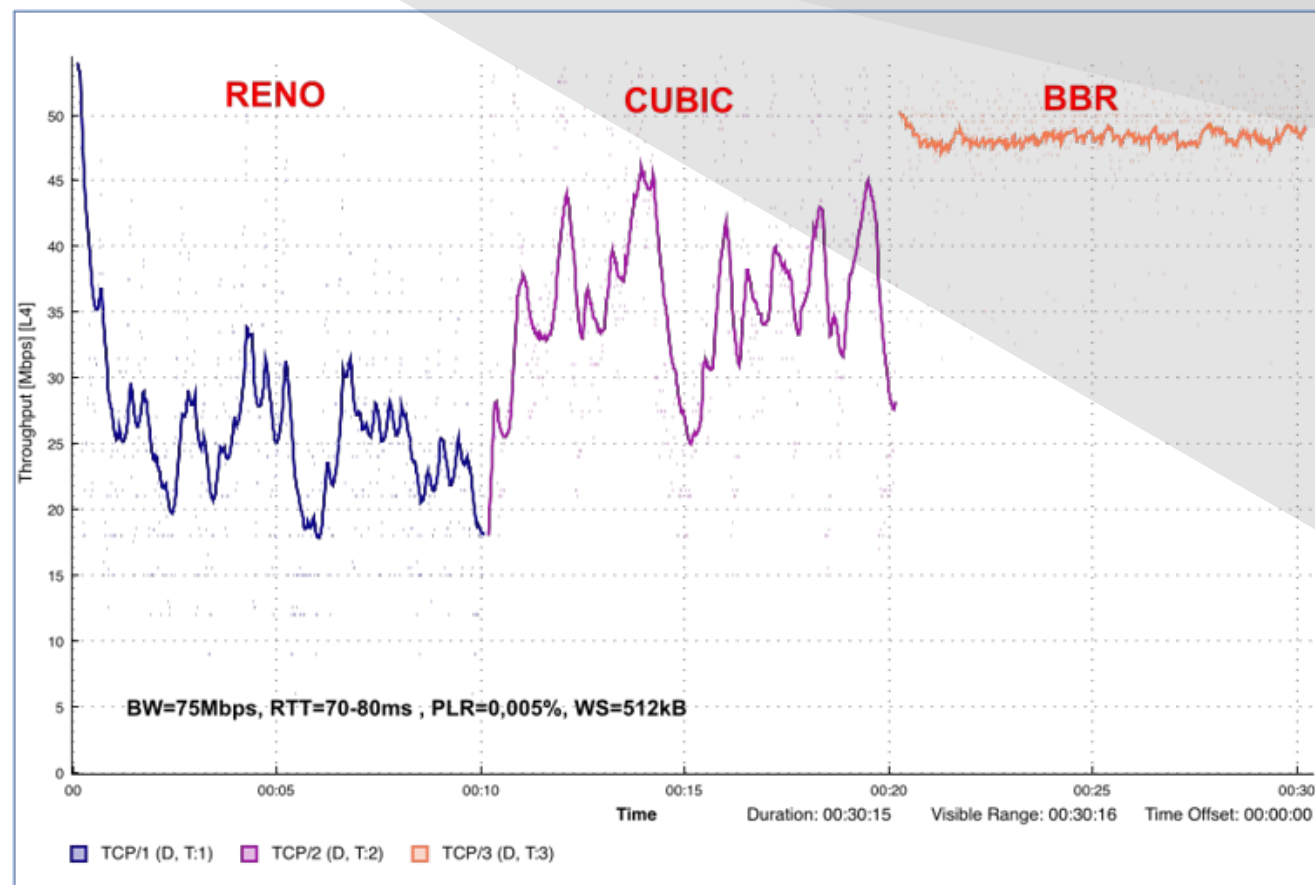
- Smyslem je měřit uživatelskou zkušenost při přenosu dat pomocí TCP
- Zavádí metriky: Transfer Time Ratio, TCP Efficiency, Buffer Delay
 - jsou dobře použitelné a jsou odolné vůči parametrizaci TCP
- RFC 6349 neříká úplně jak měřit propustnost sítě pomocí TCP!
 - **Jak škálovat WS a počet toků v případě ztrátovosti?**
- Naznačuje, že měření do cca 5% ztrátovosti u TCP je OK
 - V rámci zmiňovaných metrik to je v pořádku, ale ...
 - **I malá ztráta paketů (0.1%) má znatelný dopad na propustnost TCP** (u všech algoritmů, které v důsledku ztráty indikují zahlcení sítě)
- Různé CA algoritmy a využití dostupné kapacity? (reno, cubic, bbr, ...)

Throughput [Mbit/s] [L4]

Test / File ID	Stream ID	Direction	Min TP	Avg TP	Max TP	5%ile	50%ile	95%ile	Duration	Transferred
1/1	tcp/1	downstream	5.995	24.950	53.965	11.994	23.985	40.976	9m 58s	1.9 [GB]
2/2	tcp/2	downstream	14.988	35.947	54.461	20.985	34.480	52.468	9m 58s	2.7 [GB]
3/3	tcp/3	downstream	29.981	48.245	53.968	36.976	49.468	52.472	9m 58s	3.6 [GB]

Využití kapacity v závislosti na CA algoritmu

- Jak tedy METODICKY měřit propustnost TCP z uživatelského pohledu a má to řešení?
- Důsledkem příliš velké volnosti je, že měření nemusí správně reflektovat uživatelskou zkušenost.





**FAKULTA
ELEKTROTECHNICKÁ
ČVUT V PRAZE**

**Testování datových sítí efektivně,
spolehlivě a cíleně**

f-tester@fel.cvut.cz

<https://f-tester.fel.cvut.cz>

Děkujeme za pozornost