Performance Metrics & Basic Tools

Robert Stoy, DFN

EGI TF, Madrid

September 2013

Basic network performance metrics

- Packet Loss
- One Way Delay / OWD Variation
- Path-MTU
- TCP throughput

Basic Tools

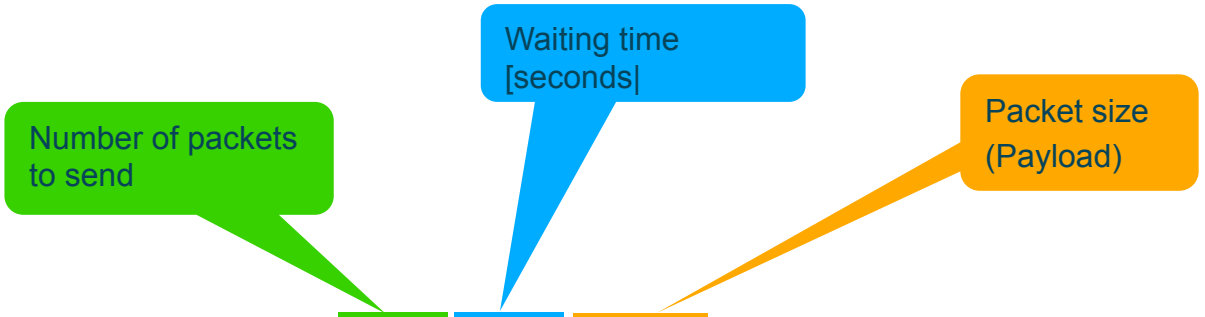
- Two-Way Delay (Round Trip Time) : Ping
- High Packet Loss (> 1 %) : Ping
- Path Finding : Traceroute
- Path-MTU : Ping, Iperf
- TCP Throughput , UDP Packet Loss : Iperf (or Nuttcp)
- Central control of Iperf (or Nuttcp) : BWCTL

Delay : Ping (1)

- How it works
  - Source host A sends *ICMP echo-request* packets to a target, waits for a given time, default 1 second.
  - Target host B replies with ICMP echo-reply packet
- Measurements using multiple packets provide statistics on RTT (min,mean,max) and Packet Loss (%)
  - *Round Trip Time of packets:*
    = Packet processing time in hosts A and B,
    + round trip time (RTT) on nework path on both directions A->B , B->A.
  - *Lost Packets:*
    Packets that do not arrive after timeout parameter

Delay : Ping (2)

- Command Line Switch, Examples
- Typical sequence in order to measure
- RTT and Loss statistics and pattern on path




```
stoy@host:~$ ping -c 100 -i 0.3 -s 1492 deneb
PING host.mydomain.de (10.10.76.9) 1492(1520) bytes of data.
1500 bytes from host.mydomain.de (10.10.76.9): icmp_req=1 ttl=245 time=22.6 ms
1500 bytes from host.mydomain.de (10.10.76.9): icmp_req=2 ttl=245 time=22.5 ms
1500 bytes from host.mydomain.de (10.10.76.9): icmp_req=3 ttl=245 time=22.6 ms
...
1480 bytes from host.mydomain.de (10.10.76.9): icmp_req=99 ttl=246 time=21.5 ms
1480 bytes from host.mydomain.de (10.10.76.9): icmp_req=100 ttl=246 time=21.5 ms

--- host.mydomain.de ping statistics ---
100 packets transmitted, 100 received, 0% packet loss, time 29778ms
rtt min/avg/max/mdev = 21.442/21.877/22.648/0.515 ms
```

Delay : Ping (3)

- Flood ping
  - get fast results on packet loss
  - requires root privileges

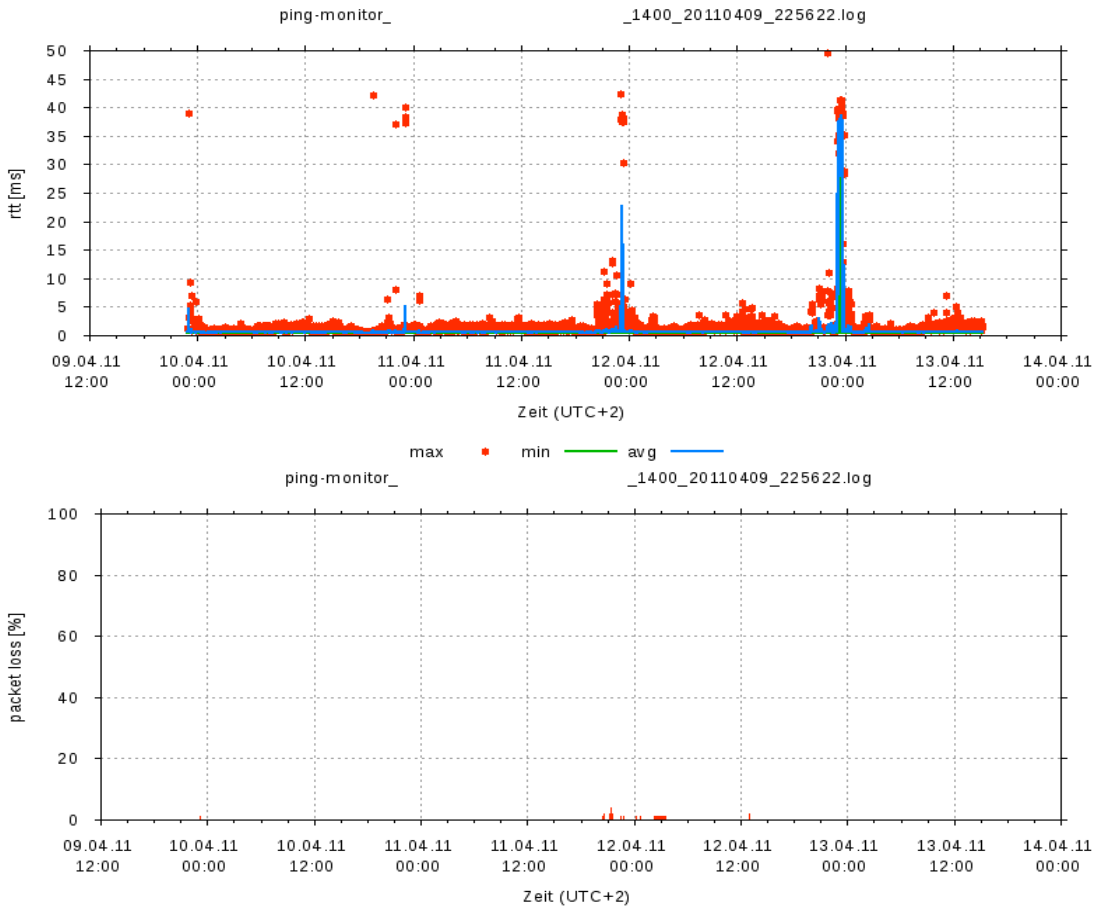

```
root@host:~# ping -c 100 -f -s 1492 deneb
PING host.mydomain.de (10.10.76.9) 1492(1520) bytes of data.

--- host.mydomain.de ping statistics ---
100 packets transmitted, 100 received, 0% packet loss, time 1107ms
rtt min/avg/max/mdev = 22.525/22.633/22.698/0.204 ms, pipe 3, ipg/ewma
11.186/22.628 ms
```

Delay : Ping, Usability

- Fast check of connectivity
- If remote endsystem is maintained on stable, low load level:
  - Detects packet loss on path from 1% upwards
  - Delay changes indicate path changes
  - Delay variation indicate load on the path
  - Scheduled measurements
    get statistics to be correlated over time with other statistics:
    - *path changes*
    - *load on network links*

Delay, Loss : Ping, Statistics Graph



- Scheduled pings
  Output as graph

Path Finding : Traceroute, How it works

- Source host sends UDP or ICMP Packets with increasing TTL values, starting at 1, to destination host.
- Routers on network do:
  - count TTL value down by 1 if packet comes in
    - *if TTL value > 0 : packet is forwarded,*
    - *if TTL value = 0 :discard packet, and send back to source host a ICMP time-to-live-exceed packet.*
- Traceroute program on source hosts analyses packets coming back, and displays result.

Path Finding : Traceroute, How it works

```
stoy@host:~$ traceroute host.mydomain.de
traceroute to host.mydomain.de (10.10.76.9), 30 hops max, 60 byte packets
 1  kr-sgs1.sgs.dfn.de (193.174.247.1)  0.412 ms  0.569 ms  0.727 ms
 2  xr-fzk1-ge8-2-301.x-win.dfn.de (188.1.159.249)  2.094 ms  2.236 ms  2.381 ms
 3  xr-stu1-te2-3.x-win.dfn.de (188.1.145.130)  20.263 ms  20.325 ms  20.324 ms
 4  xr-gar1-te2-4.x-win.dfn.de (188.1.144.229)  21.483 ms * *
 5  xr-aug1-te1-1.x-win.dfn.de (188.1.144.110)  21.641 ms  21.774 ms  21.901 ms
 6  cr-erl1-te0-0-0-7.x-win.dfn.de (188.1.144.149)  22.348 ms  22.809 ms  22.784 ms
 7  cr-tub1-hundredgige0-1-0-0-7.x-win.dfn.de (188.1.144.185)  22.951 ms  22.006 ms  22.034 ms
 8  xr-tub2-vlan50.x-win.dfn.de (188.1.144.158)  22.393 ms  21.681 ms  21.946 ms
 9  xr-hub1-ge3-11-1170.x-win.dfn.de (188.1.159.245)  22.630 ms  22.433 ms  22.560 ms
10  kr-dfnbln5-0.x-win.dfn.de (188.1.159.246)  22.764 ms  22.887 ms  22.746 ms
11  host.mydomain.de (10.10.76.9)  21.316 ms  21.313 ms  21.260 ms
stoy@host:~$
```

- Sender-Default: 3 packets on each TTL value
- Result:
  - List of Routers along the path
  - Round Trip Time between source host and router, 3 values on each hop

Path Finding : Traceroute, Points to consider (I)

- Path is shown only between Layer 3 devices,

  -> Path changes on lower layers not visible in hop list.

- Path is shown only in one direction.

  -> traceroute required from both end systems

- Measurement Timestamp not in output, but required

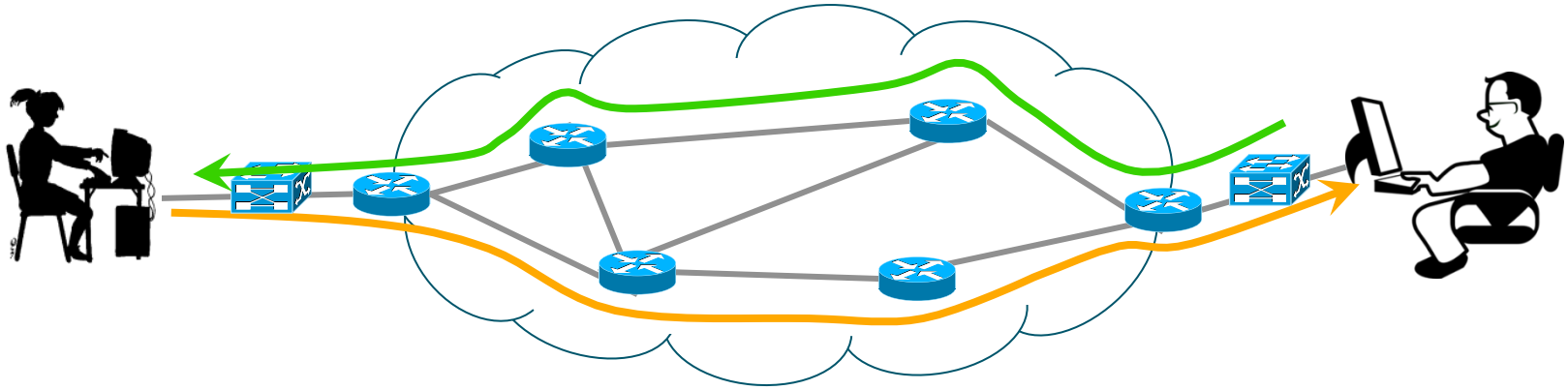
  -> do a ‘unix date’ separately.

Path Finding : Traceroute, Points to consider (II)

- Network providers support traceroute, but keep in mind:
  - Processing load and -time on routers is higher as with answering to pings (because of ICMP time-to-live-exeed generation)
  - Usually routers treat TTL=0 packets with low priority
  - Router protection mechanisms could prevent generation of ICMP time-to-live-exceed packets.
  - => * in traceroute output could occur although there's no problem on forwarding path
  - => RTT could show high values, although delay on path is lower.

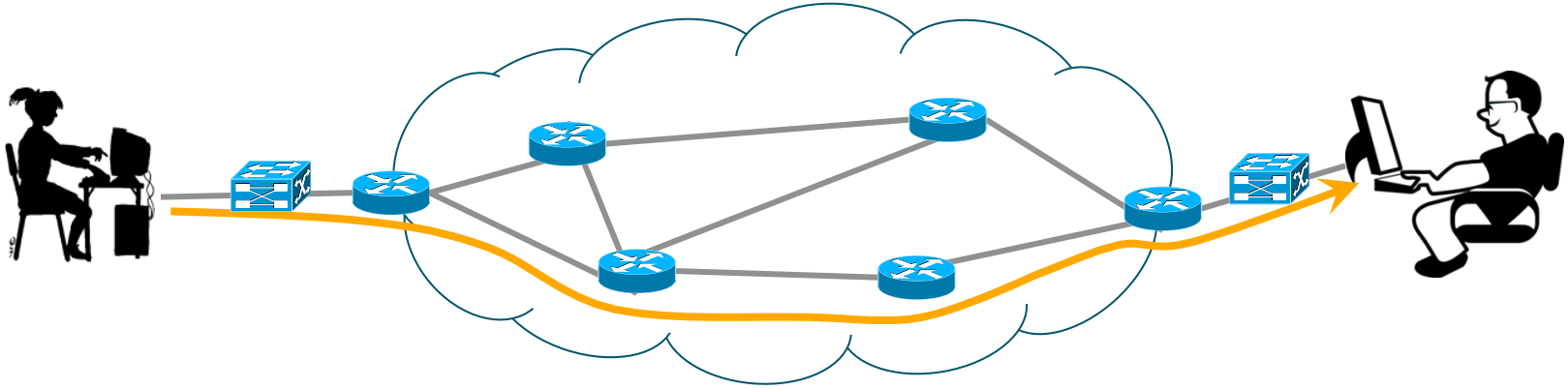
- Firewall administrators an WAN/LAN Borders sometimes prevent traceroute

TCP Throughput : Iperf
How it works



- Data Transfer between memory on local and remote host through end systems TCP/IP-Stack and network
- One program, two roles : Server (Receiver) & Client (Sender)
- Two modes: TCP, UDP
- Best practise:
  - Interactive session on both sides
  - Take result from Server (Receiver)

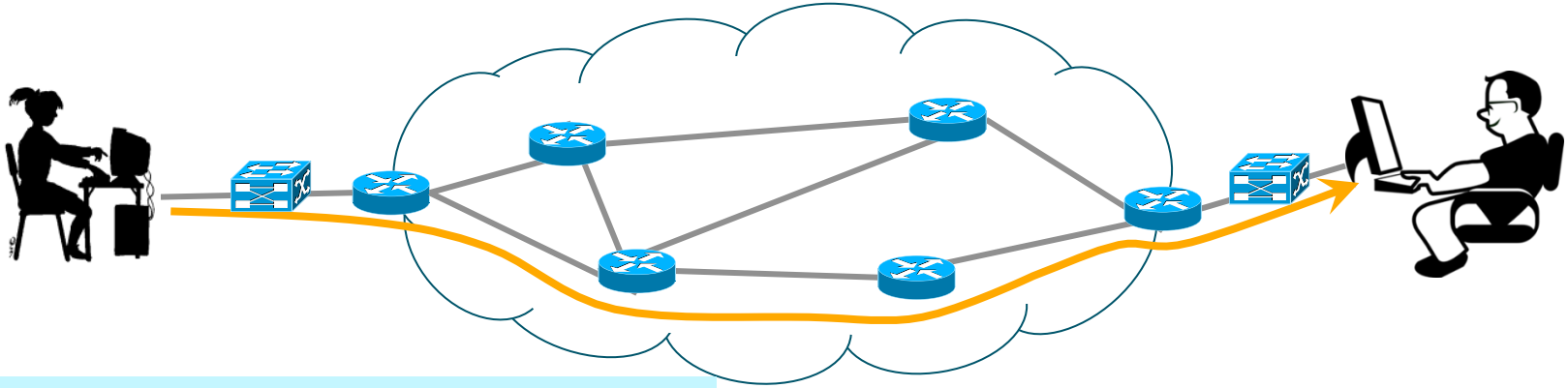
TCP Throughput : Iperf
Example TCP-Mode



```
stoy@mp-siteA:~$ iperf -B 10.10.12.186  -c  10.10.23.194 -m
-i 1 -t 10
------------------------------------------------------------
Client connecting to 10.10.23.194, TCP port 5001
Binding to local address 10.10.12.186
TCP window size:  128 KByte (default)
------------------------------------------------------------
[  3] local 10.10.12.186 port 5001 connected with
10.10.23.194 port 5001
[ ID] Interval       Transfer     Bandwidth
[  3]  0.0- 1.0 sec  81.4 MBytes   683 Mbits/sec
[  3]  1.0- 2.0 sec   114 MBytes   956 Mbits/sec
[  3]  2.0- 3.0 sec   111 MBytes   930 Mbits/sec
[  3]  3.0- 4.0 sec   111 MBytes   934 Mbits/sec
[  3]  4.0- 5.0 sec   111 MBytes   933 Mbits/sec
[  3]  5.0- 6.0 sec   111 MBytes   932 Mbits/sec
[  3]  6.0- 7.0 sec   113 MBytes   949 Mbits/sec
[  3]  7.0- 8.0 sec   115 MBytes   963 Mbits/sec
[  3]  8.0- 9.0 sec   111 MBytes   932 Mbits/sec
[  3]  9.0-10.0 sec   111 MBytes   933 Mbits/sec
[  3]  0.0-10.0 sec  1.06 GBytes   911 Mbits/sec
[  3] MSS size 1448 bytes (MTU 1500 bytes, ethernet)
```

```
stoy@mp-siteB:~$ iperf -s  -B 10.10.23.194 -m -i 1
------------------------------------------------------------
-
Server listening on TCP port 5001
Binding to local address 10.10.23.194
TCP window size:  128 KByte (default)
------------------------------------------------------------
-
[  4] local 10.10.23.194 port 5001 connected with
10.10.12.186 port 5001
[ ID] Interval       Transfer     Bandwidth
[  4]  0.0- 1.0 sec  72.8 MBytes   611 Mbits/sec
[  4]  1.0- 2.0 sec   112 MBytes   937 Mbits/sec
[  4]  2.0- 3.0 sec   112 MBytes   937 Mbits/sec
[  4]  3.0- 4.0 sec   112 MBytes   937 Mbits/sec
[  4]  4.0- 5.0 sec   112 MBytes   937 Mbits/sec
[  4]  5.0- 6.0 sec   112 MBytes   937 Mbits/sec
[  4]  6.0- 7.0 sec   112 MBytes   937 Mbits/sec
[  4]  7.0- 8.0 sec   112 MBytes   936 Mbits/sec
[  4]  8.0- 9.0 sec   112 MBytes   937 Mbits/sec
[  4]  9.0-10.0 sec   112 MBytes   937 Mbits/sec
[  4]  0.0-10.1 sec  1.06 GBytes   904 Mbits/sec
[  4] MSS size 14
```

TCP Throughput : Iperf Example UDP-Mode



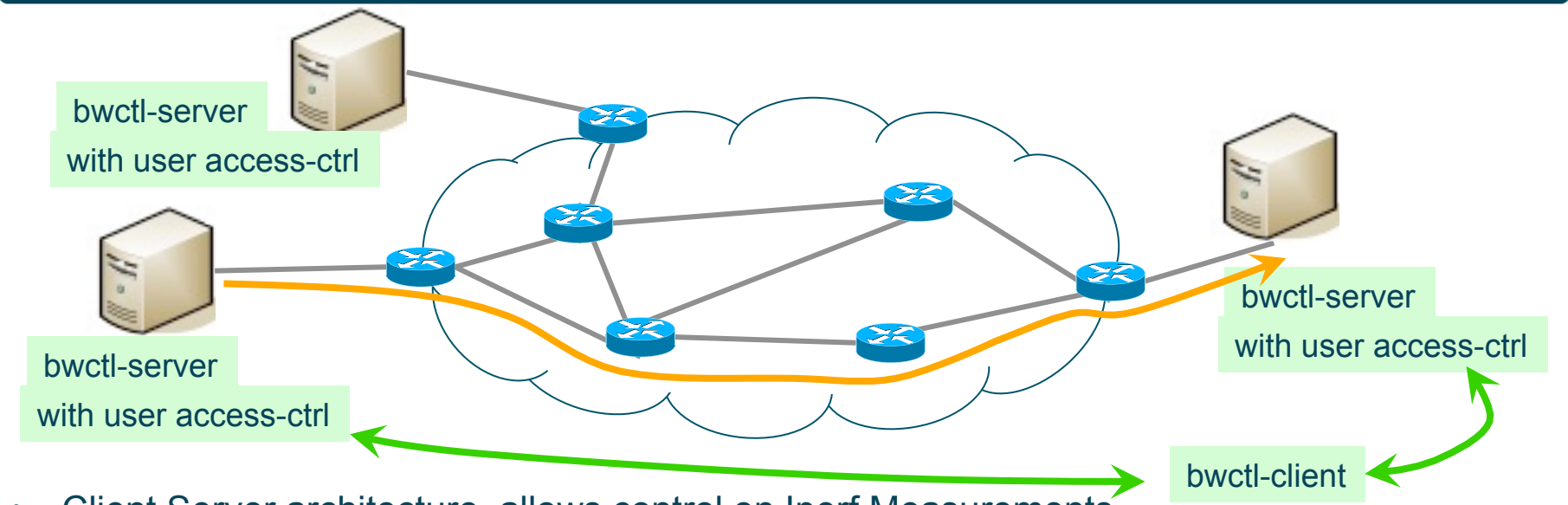
```
stoy@mp-siteA:~$ iperf -u -B 10.10.12.186  -c  10.10.23.194
-i 1 -t 10 -b 100M
------------------------------------------------------------
Client connecting to 10.10.23.194, UDP port 5001
Binding to local address 10.10.12.186
Sending 1470 byte datagrams
UDP buffer size:  160 KByte (default)
------------------------------------------------------------
[  3] local 10.10.12.186 port 5001 connected with 10.10.23.194
port 5001
[ ID] Interval       Transfer     Bandwidth
[  3]  0.0- 1.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  1.0- 2.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  2.0- 3.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  3.0- 4.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  4.0- 5.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  5.0- 6.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  6.0- 7.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  7.0- 8.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  8.0- 9.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  9.0-10.0 sec  12.0 MBytes   101 Mbits/sec
[  3]  0.0-10.0 sec   120 MBytes   101 Mbits/sec
[  3] Sent 85471 datagrams
[  3] Server Report:
[  3]  0.0-10.0 sec   120 MBytes   101 Mbits/sec   0.016 ms
0/85470 (0%)
[  3]  0.0-10.0 sec  1 datagrams received out-of-order
```

```
stoy@mp-siteB:~$ iperf -u -s  -B 10.10.23.194  -i 1
------------------------------------------------------------
Server listening on UDP port 5001
Binding to local address 10.10.23.194
Receiving 1470 byte datagrams
UDP buffer size:  160 KByte (default)
------------------------------------------------------------
[  3] local 10.10.23.194 port 5001 connected with 10.10.12.186 port 5001
[ ID] Interval       Transfer     Bandwidth        Jitter   Lost/Total
Datagrams
[  3]  0.0- 1.0 sec  12.0 MBytes   101 Mbits/sec   0.016 ms    0/ 8547 (0%)
[  3]  1.0- 2.0 sec  12.0 MBytes   101 Mbits/sec   0.018 ms    0/ 8547 (0%)
[  3]  2.0- 3.0 sec  12.0 MBytes   101 Mbits/sec   0.016 ms    0/ 8547 (0%)
[  3]  3.0- 4.0 sec  12.0 MBytes   101 Mbits/sec   0.022 ms    0/ 8547 (0%)
[  3]  4.0- 5.0 sec  12.0 MBytes   101 Mbits/sec   0.018 ms    0/ 8547 (0%)
[  3]  5.0- 6.0 sec  12.0 MBytes   101 Mbits/sec   0.016 ms    0/ 8547 (0%)
[  3]  6.0- 7.0 sec  12.0 MBytes   101 Mbits/sec   0.016 ms    0/ 8547 (0%)
[  3]  7.0- 8.0 sec  12.0 MBytes   101 Mbits/sec   0.016 ms    0/ 8547 (0%)
[  3]  8.0- 9.0 sec  12.0 MBytes   101 Mbits/sec   0.017 ms    0/ 8547 (0%)
[  3]  9.0-10.0 sec  12.0 MBytes   101 Mbits/sec   0.016 ms    0/ 8547 (0%)
[  3]  0.0-10.0 sec   120 MBytes   101 Mbits/sec   0.017 ms    0/85470 (0%)
[  3]  0.0-10.0 sec  1 datagrams received out-of-order
```

Iperf : Points to consider

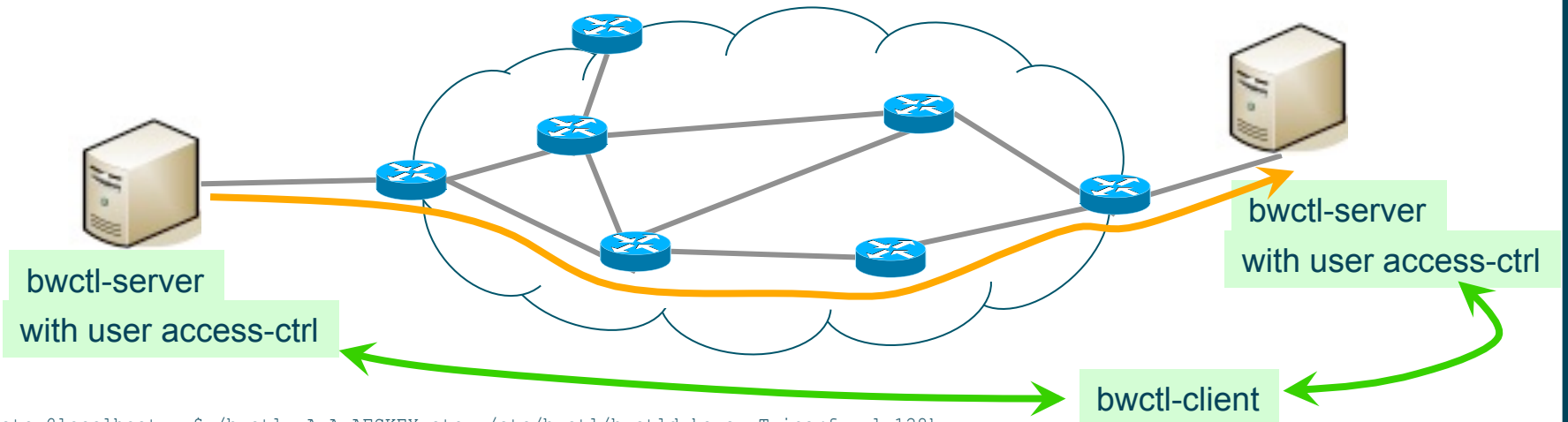
- UDP measurements:
  - Dangerous:
    Could cause packet loss on other data streams on the network
  - Use only in controlled environment
    keep as much as local, typical: Check a link between routers
    only small data rates
    only short duration
- Recommended Version: Iperf v 2.0.5
- Linux: Packages for Debian and Redhat based distributions

BWCTL , Centrally Controlled Iperf Measurements (I)



- Client Server architecture, allows control on Iperf Measurements
- Protocol relies on NTP synchronised servers and clients
- Allows enforcement of
  - client user access control
  - Max. number of simultaneous measurement flows
  - Limits on measurement parameters per user group
- Servers are started once with configured usage policy
- Logging on servers supported by posthook script interface
- On demand measurement, simply from a client CLI

BWCTL , Centrally Controlled Iperf Measurements (II)




```
stoy@localhost :~$ /bwctl -A A AESKEY stoy /etc/bwctl/bwctld.keys -T iperf  -l 128k
-i 1 -t 10 -s mp-tub -c mp-stu
bwctl: 18 seconds until test results available
RECEIVER START
bwctl: exec_line: /usr/bin/iperf -B mp-stu -s -f b -l 131072 -m -p 5004 -t 10 -i 1
bwctl: start_tool: 3588256214.259536
------------------------------------------------------------
Server listening on TCP port 5004
Binding to local address mp-stu
TCP window size: 131072 Byte (default)
------------------------------------------------------------
[ 15] local mp-stu port 5004 connected with mp-tub port 5004
[ ID] Interval       Transfer     Bandwidth
[ 15]  0.0- 1.0 sec  82052368 Bytes  656418944 bits/sec
[ 15]  1.0- 2.0 sec  117065008 Bytes  936520064 bits/sec
[ 15]  2.0- 3.0 sec  117096864 Bytes  936774912 bits/sec
[ 15]  3.0- 4.0 sec  117066456 Bytes  936531648 bits/sec
…
[ 15]  8.0- 9.0 sec  117066456 Bytes  936531648 bits/sec
[ 15]  9.0-10.0 sec  117095416 Bytes  936763328 bits/sec
[ 15]  0.0-10.1 sec  1146617856 Bytes  908849978 bits/sec
[ 15] MSS size 1448 bytes (MTU 1500 bytes, ethernet)
bwctl: stop_exec: 3588256230.789755
RECEIVER END
```

BWCTL , Centrally Controlled Iperf Measurements (III)

- Further option: control of Nuttcp-Measurements
  - Supports output of CPU load on measurement node
- Deployment
  - Linux: Server and Client Packages for Debian and RedHat based Systems available
  - Source Code and Handbook @ Internet2
- Room for improvement
  - Results are sent to the client, but not stored, -> no history
  - Used network path not known, traceroute feature is missing
  - User authentication and authorisation not really strong
  - No graphical output

=> Evolution towards PerfSonar