WP3

March 24, 2004

Cees de Laat,
UvA,
The Netherlands

DataTAG is a project funded by the European Commission under contract IST-2001-32459

Talk Outline

- Wait 15 minutes

WP3 objectives

- Task 3.1. Performance validation
  - Create, collect, test network-tools to cope with the extreme Lambda environment (high RTT, BW)
  - measure basic properties and establish a baseline performance benchmark
- Task 3.2 End user performance validation/monitoring/optimization
  - Use "out of band" tools to measure and monitor what performance a user in principle should be able to reach
- Task 3.3 Application performance validation, monitoring and optimization
  - Use diagnostic libraries and tools to monitor and optimize real applications to compare their performance with task 3.2 outcome.

WP3 achievements t3.1

- Performance Test Tools
  - http://carol.science.uva.nl/~jblom/datatag/wp3_1/tools/test_tools.html
- Lambda's are a Layer 1-2(-3) environment
  - use end to end traffic generating tools
  - not ICMP based tools since those need routers in the path
  - tools must work in Gigabit environment
- Netperf
  - TCP bulk transfer, UDP unusable (no shaping)
  - modifications made to enhance IPv6 support
- Iperf
  - TCP and UDP shaped bulk transfer
  - lightweight multiple streams test support
  - IPv6 support
  - modifications made to introduce TCP shaping, debugging

Understand Lightpath

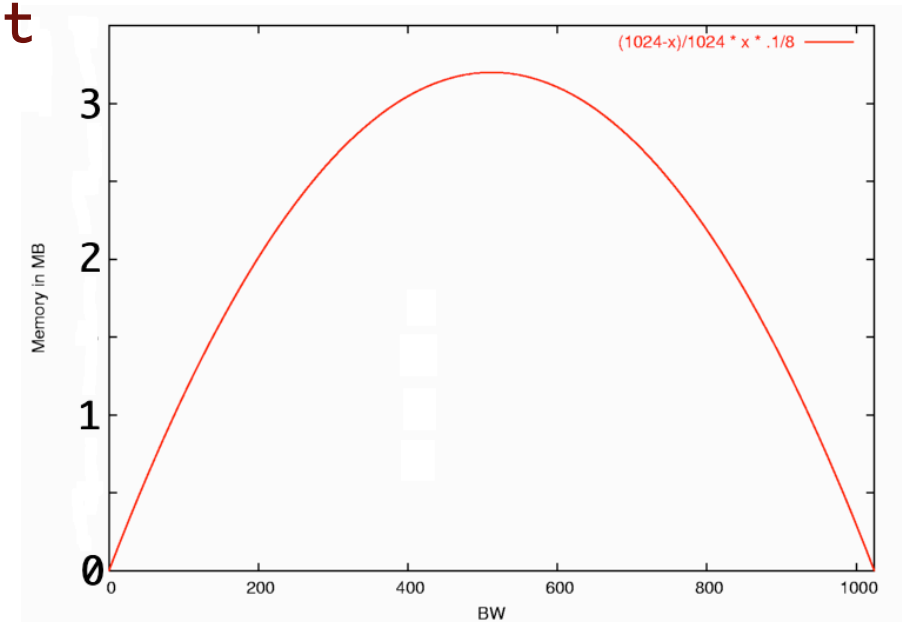
TCP is bursty due to sliding window protocol and slow start algorithm.

```
Window = BandWidth * RTT     &    BW == slow
```

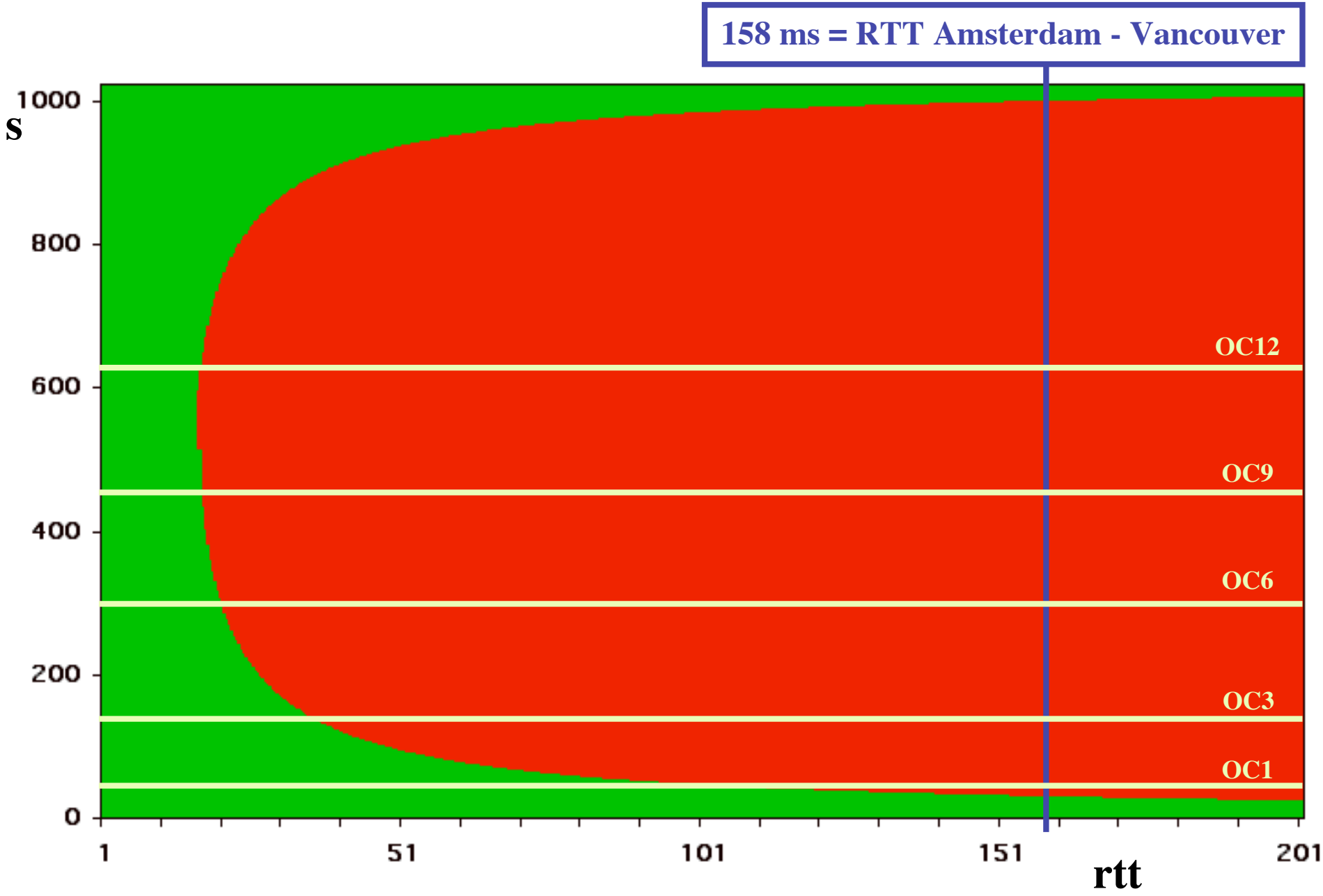
$$\text{Memory-at-bottleneck} = \frac{\text{fast} - \text{slow}}{\text{fast}} * \text{slow} * \text{RTT}$$

So pick from menu:

- *Flow control*
- *Traffic Shaping*
- *RED (Random Early Discard)*
- *Self clocking in TCP*
- *Deep memory*



Forbidden area, solutions for s when f = 1 Gb/s, M = 0.5 Mbyte AND NOT USING FLOWCONTROL



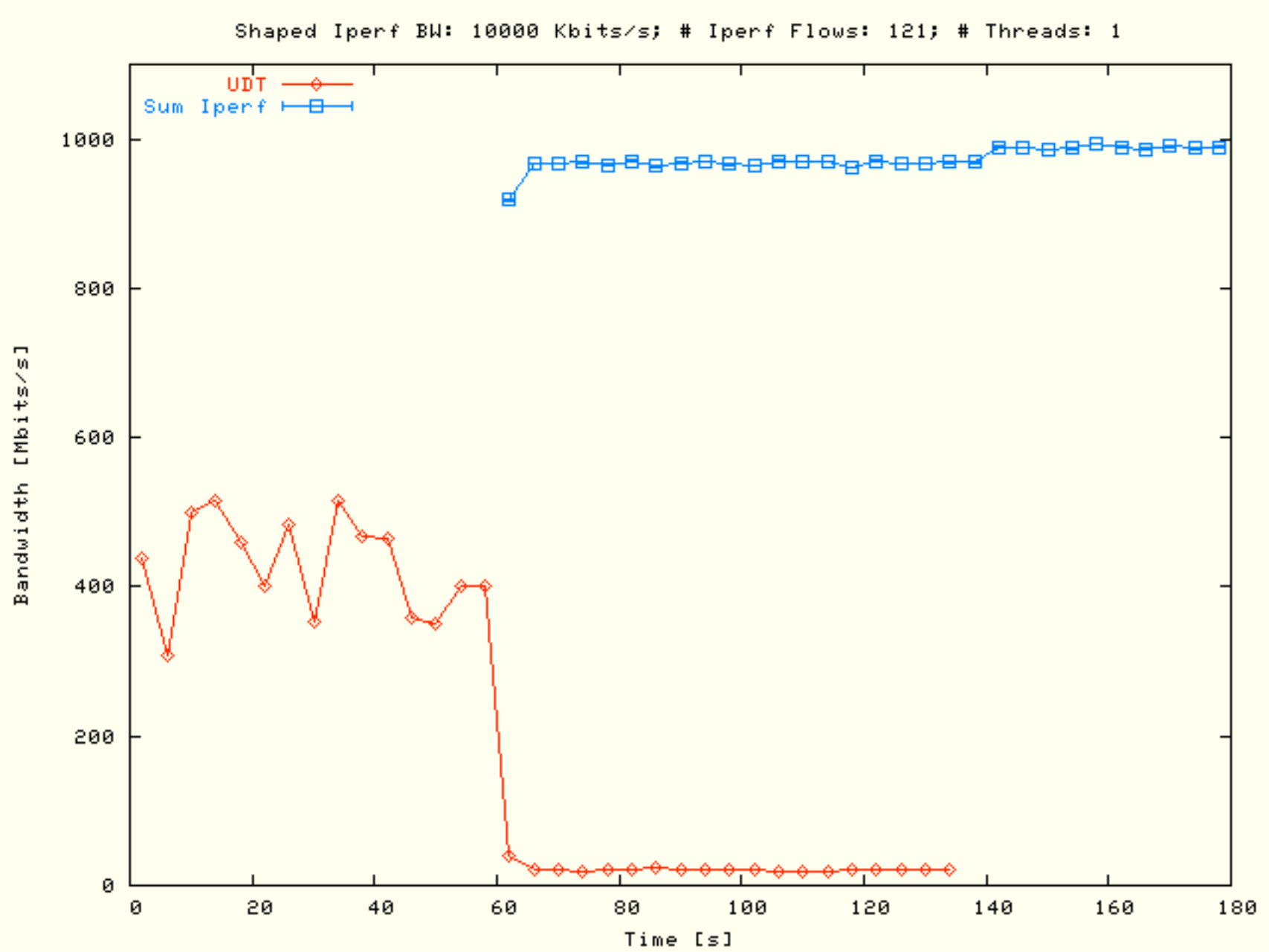
SCTP Performance

- Stream Control Transmission Protocol:
  - TCP-lite
- Few versions support Linux
- LK-SCTP not optimized for performance:
  - Not currently usable for Grids
- Feedback to SCTP protocol designers
- Results put on the Web:
  - one of the pages with the greatest hit rate

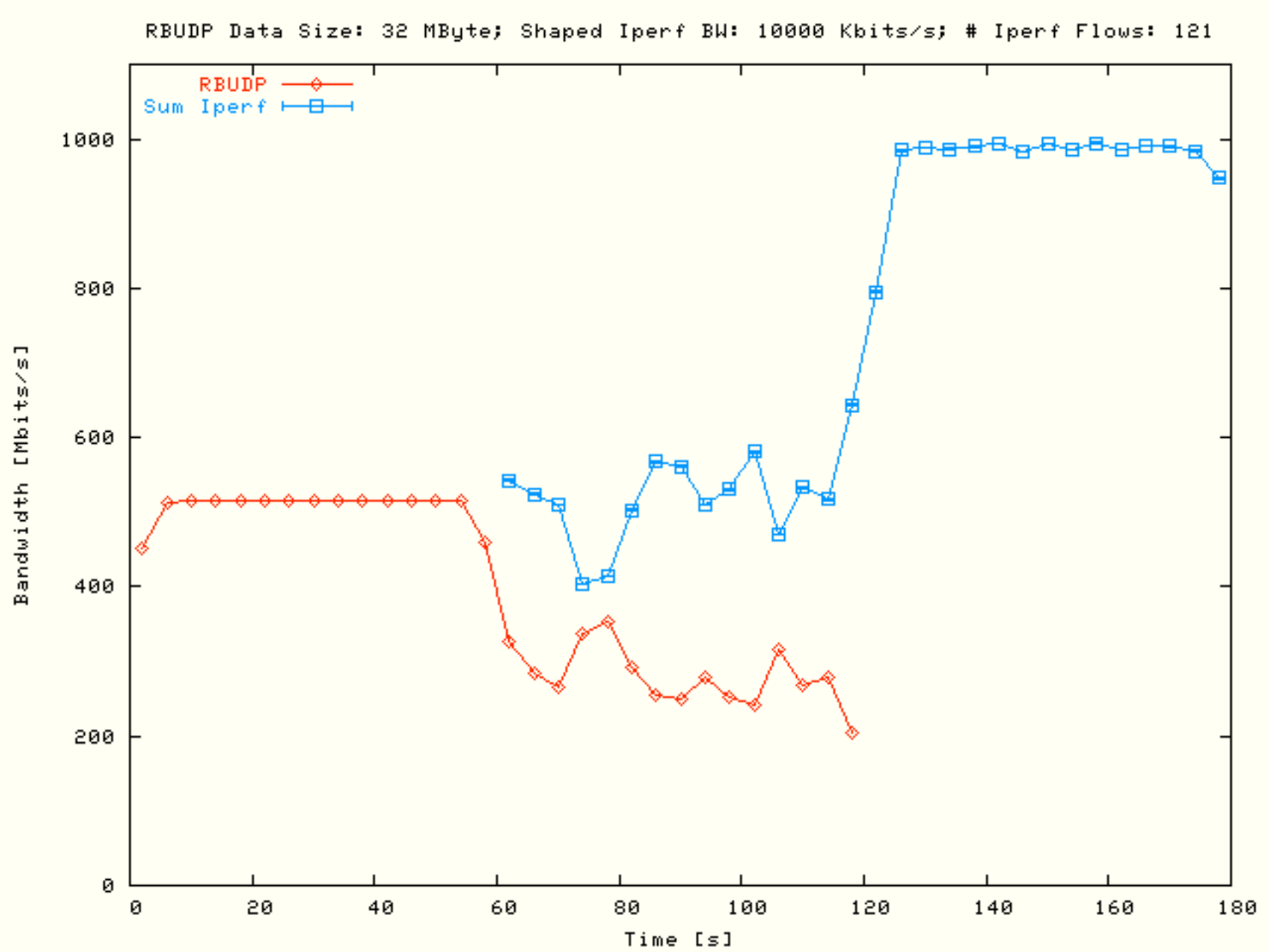
Protocol Testbed

- Non Goal
  - To come up with yet another protocol
- Goal
  - To test new and supposedly aggressive protocols mixed with many vanilla streams in controlled environments
- How
  - Use gridcluster to generate #100^{ths} of shaped flows background and monitor those small flows
  - Start special protocol stream
  - As function of RTT!

Test of UDT


Shaped Iperf BW: 10000 Kbits/s; # Iperf Flows: 121; # Threads: 1
UDT
Sum Iperf
Bandwidth [Mbits/s]
0
200
400
600
800
1000
0
20
40
60
80
100
120
140
160
180
Time [s]

Test of RBUDP



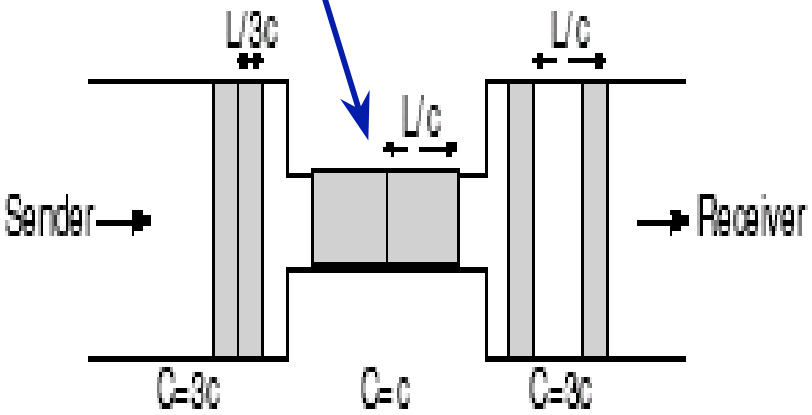
End user performance validation / monitoring / optimization t3.2

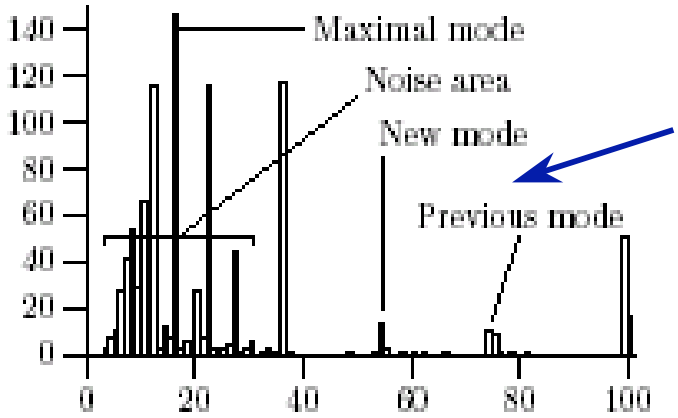
- Overview of tools (continuously updated)
  - http://carol.science.uva.nl/~jblom/datatag/wp3_1/tools/monitor_tools.html
- Homebrew:
  - rTPL(end user performance monitoring):
  - http://carol.science.uva.nl/~jblom/rtpl/
- Current monitors (Amsterdam):
  - http://195.169.124.34/monitor/index.html
- TraceRate
  - next slides in more detail

Tools: Trace-Rate Hop by hop measurements

- A method to measure the hop-by-hop capacity, delay, and loss up to the path bottleneck
  - Not intrusive
  - Operates in a high-performance environment
  - Does not need cooperation of the destination
- Based on Packet Pair Method
  - Send sets of b2b packets with increasing time to live
  - For each set filter "noise" from rtt
  - Calculate spacing – hence bottleneck BW
- Robust regarding the presence of invisible nodes

Effect of the bottleneck on a packet pair.
L is a packet size
C is the capacity





Examples of parameters that are iteratively analysed to extract the capacity mode

Tools: Trace-Rate Some Results

- Capacity measurements as function of load in Mbit/s from tests on the DataTAG Link:

	tracerate			pathchar			pathrate		
	0%	25%	50%	0%	25%	50%	0%	25%	50%
Hop #1	165	170	165	92	92	93	N/A	N/A	N/A
Hop #2	*162	*165	*162	996	977	832	N/A	N/A	N/A
Hop #3	933	862	862	N/A	N/A	N/A	981-986	760-776	927-947
Duration	2'40	2'40	2'40	N/A	N/A	N/A	25"	5'30	5'40

- Comparison of the number of packets required

	pathchar	clink	pchar	nettimer	tracerate
4 hops path	11,562	6,002	11,732	982	4,000
11 hops path	31,782	16,400	32,417	6,663	11,000

- Validated by simulations in NS-2
- Linux implementations, working in a high-performance environment
- Research report: http://www.inria.fr/rrrt/rr-4959.html
- Research Paper: ICC2004 : International Conference on Communications, Paris, France, June 2004. IEEE Communication Society.

Gigabit Network Monitoring

- Tested non-intrusive monitoring techniques:
  - Complement 2002 work on intrusive techniques
- Use performance data available in network devices and hosts:
  - SNMP MIBs
- Investigated usefulness of Web Services:
  - UDDI is not appropriate for service discovery

Grid Application Monitoring t3.3

- Tested integration of NetLogger (LBNL) and GridFTP (Argonne):
  - Detailed information about the performance of Grid applications
- Integration is difficult
- Few versions interoperate smoothly
- Detailed how-to on the Web

Q&A