

TCP Performance over Gigabit Networks

Arne Øslebø

NORDUnet2002

Modern Network Interface Cards

- Programmable network processors
- Hardware checksumming
 - Offloads main CPU
- Interrupt coalescing
 - Less interrupts
- On board memory
- Zero-copy
 - Mainly operating system issue

Test tools

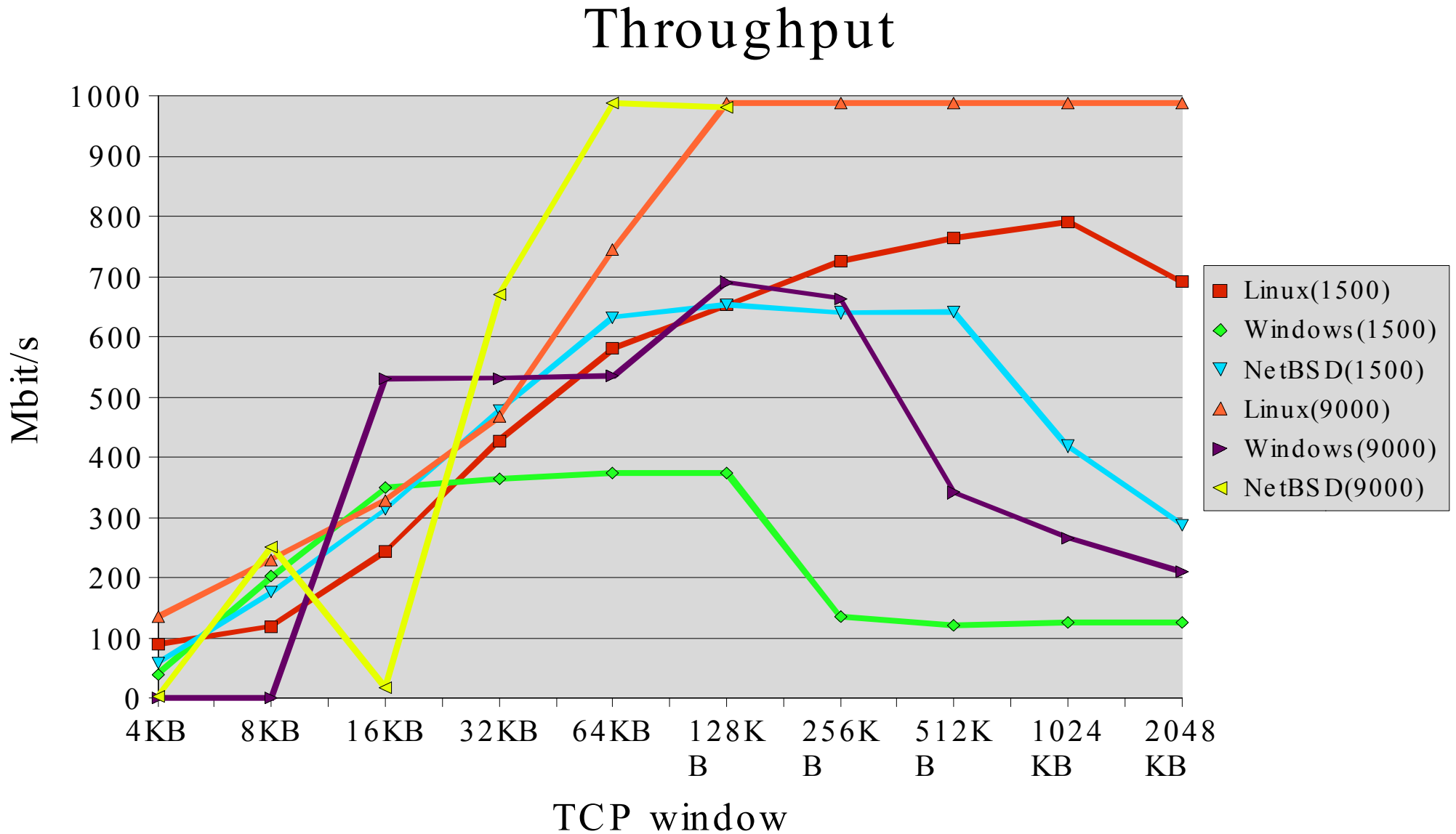
- netperf
 - used for measuring throughput between two systems.
 - <http://www.netperf.org/>
- vmstat / Microsoft Management Console
 - gives information about CPU load and interrupts/s.
- netstat
 - gives information about packet loss, retransmissions etc.
- One testrun consists of:
 - running netperf using various TCP window sizes.
 - for each TCP window size, netperf is run three or more times. Results presented are average values.
 - Before running netperf, information is collected from netstat
 - While running netperf, information is collected from vmstat.

Laboratory testing

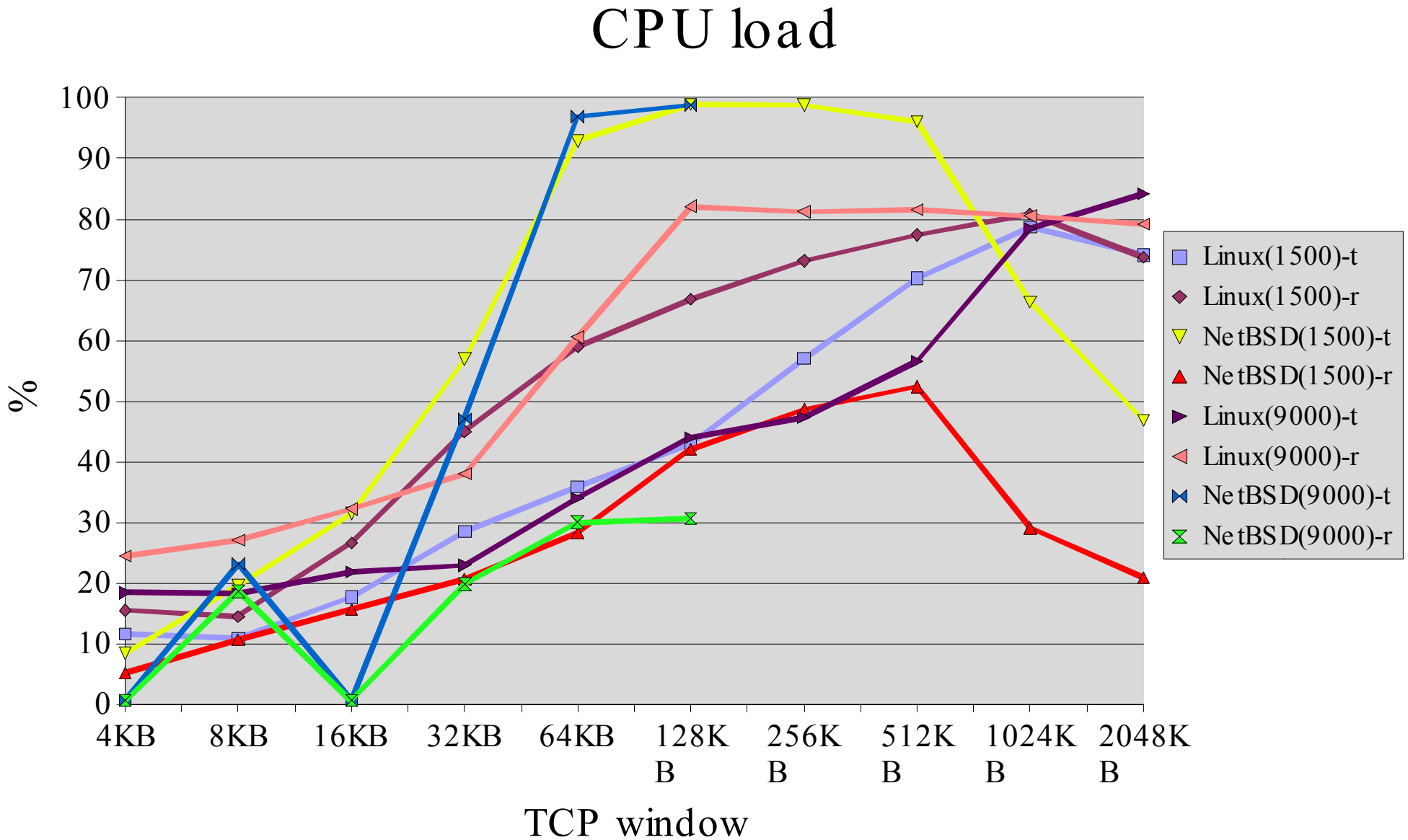
- Hardware:
 - Motherboard: Supermicro 370DE6
 - Memory: 1GB SDRAM
 - CPU: 1GHz Pentium III
 - NIC: 3COM 3C985B-SX
- Operating systems:
 - Linux: 2.4.9
 - NetBSD: current
 - Windows 2000: service pack 2
- $RTT=0.227ms$



Point to point - throughput

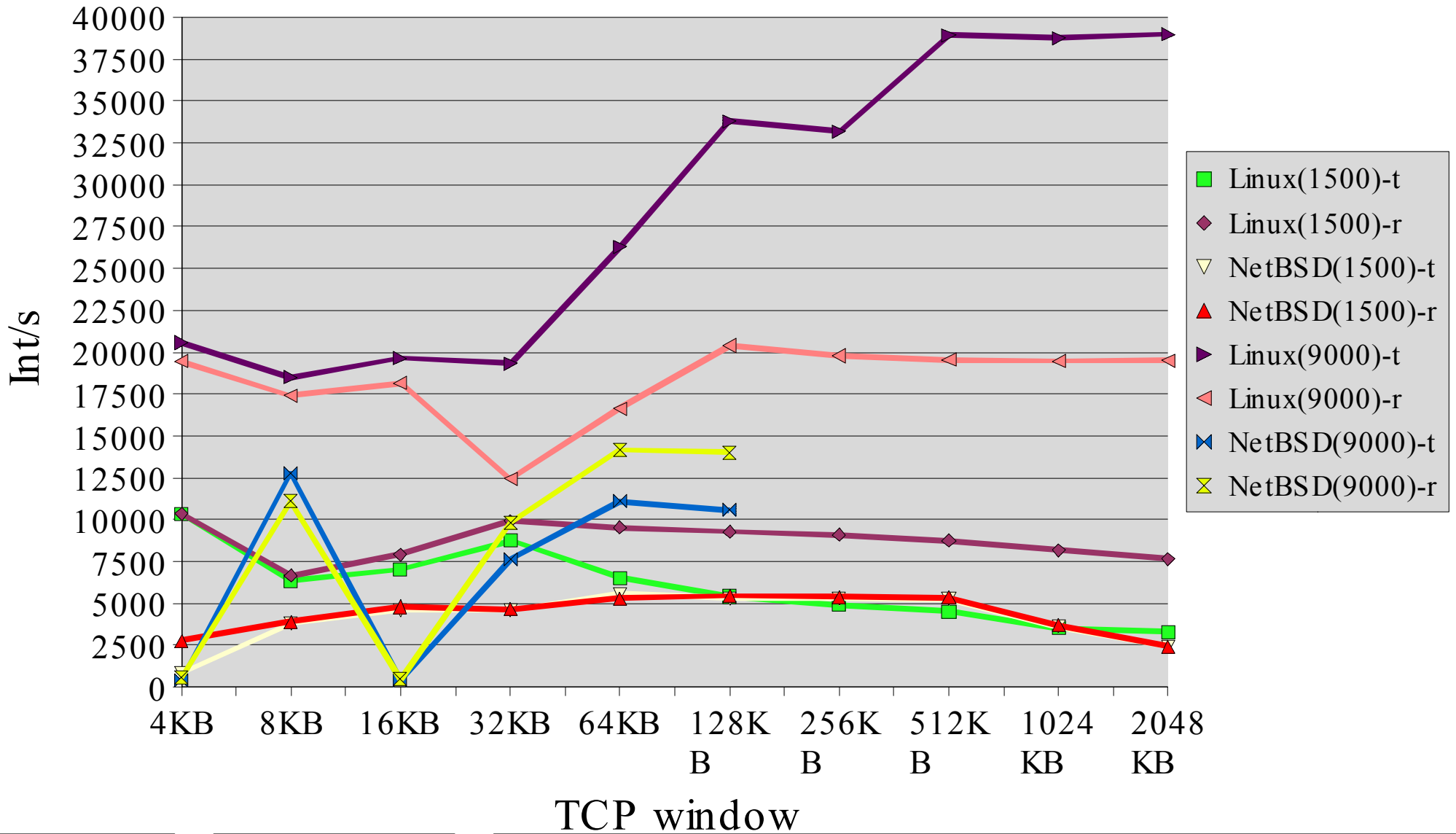


Point to point - CPU load



Point to point - interrupts/s

Interrupts

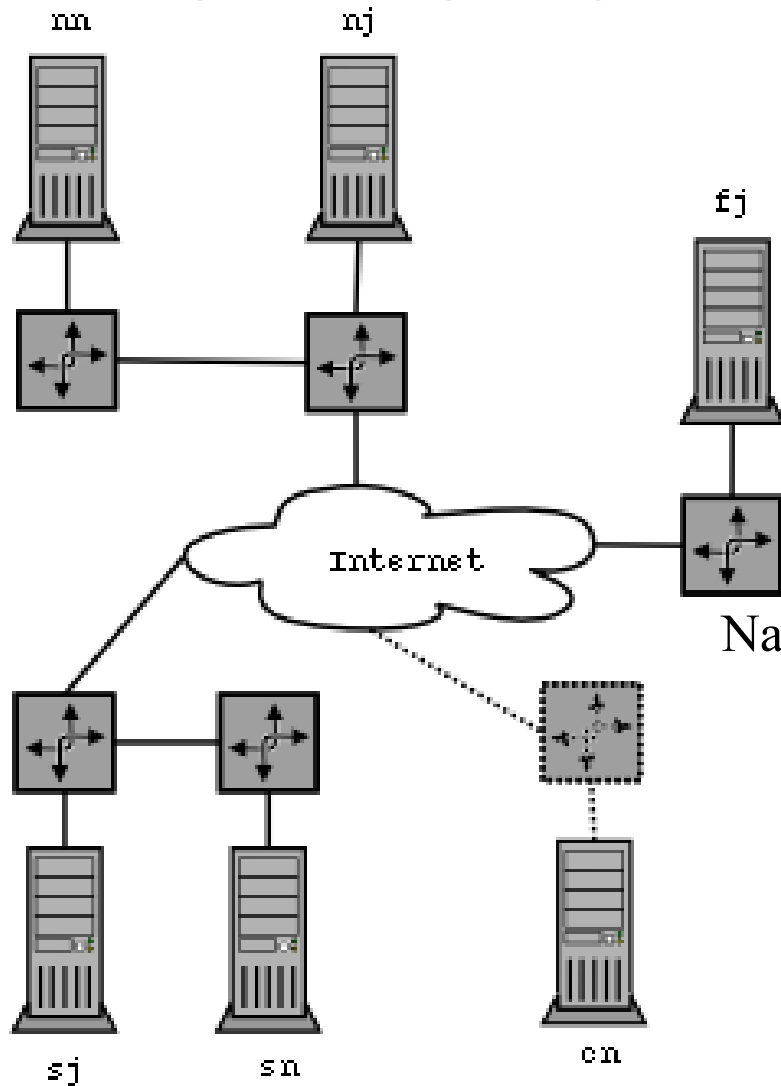


Analyzing results

M T U	L i n u x	N e t B S D	W i n d o w s
1500	792	640	375
9000	988	988	692

- Problems with using tcpdump
 - Negative effect on throughput
 - Drops many packets
 - Interrupt Coalescing
 - Gives inaccurate time stamps
- Low performance using Windows 2000 probably related to driver.

Real world experiments



These tests have been possible thanks to the good help from:

- Jari Miettinen, CSC
- Tuomas Toropainen, CSC
- Simon Leinen, SWITCH
- Sven Ubik, CESNET

Naming: **nn**

MTU size: n=normal(1500)
j=jumboframe(4470)

country: c=Czech republic

f=Finland

n=Norway

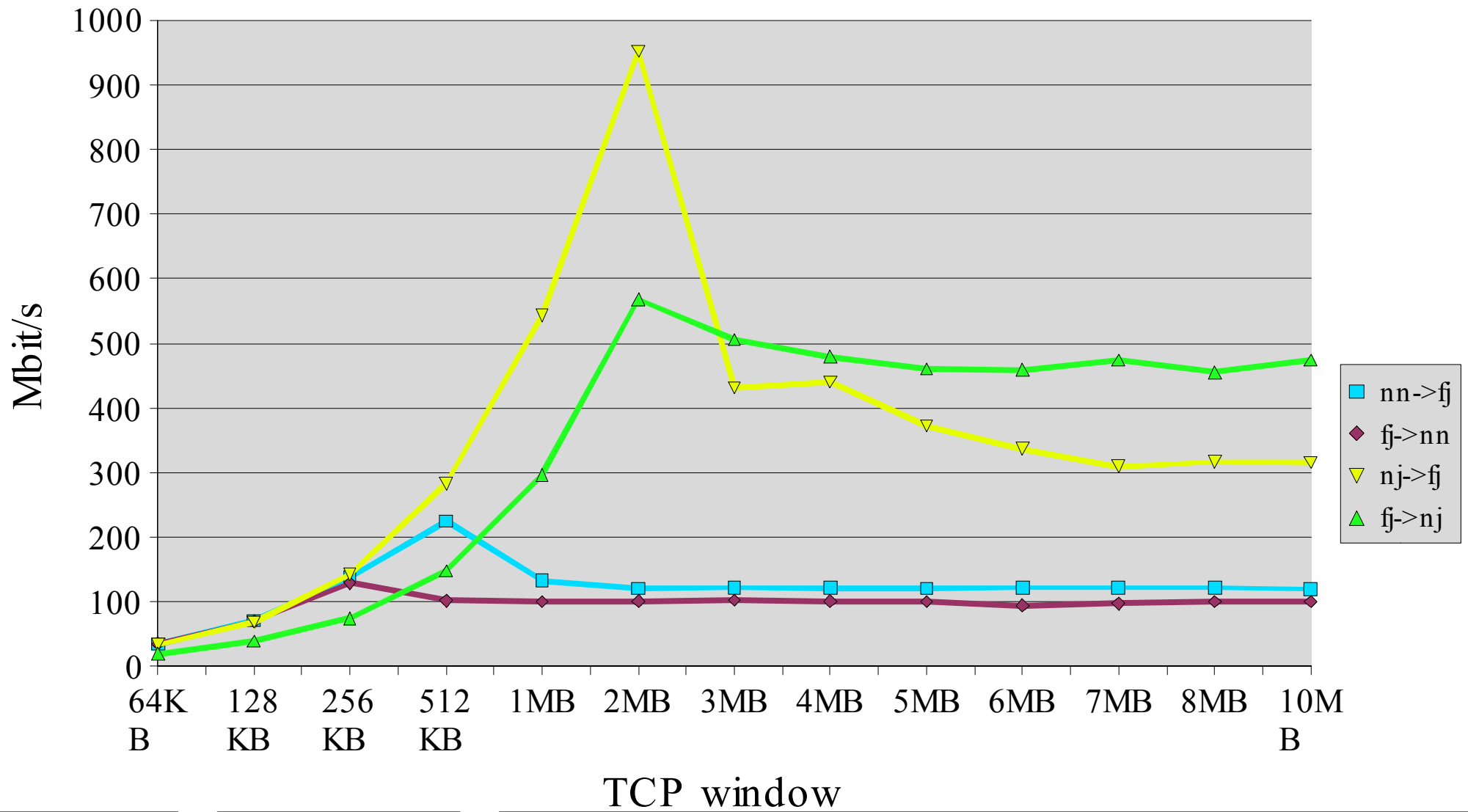
s=Switzerland

RTT:

	Czech	Finland	Norway	Switzerland	NIC
Czech		43.2	51.1	31.1	Allied Telesyn AT2970SX/SC
Finland	43.2		20.5	57.4	3COM 3C996
Norway	51.1	20.5		65.5	3COM 3C985B
Switzerland	31.1	57.4	65.5		SysKonnnect SK-9843

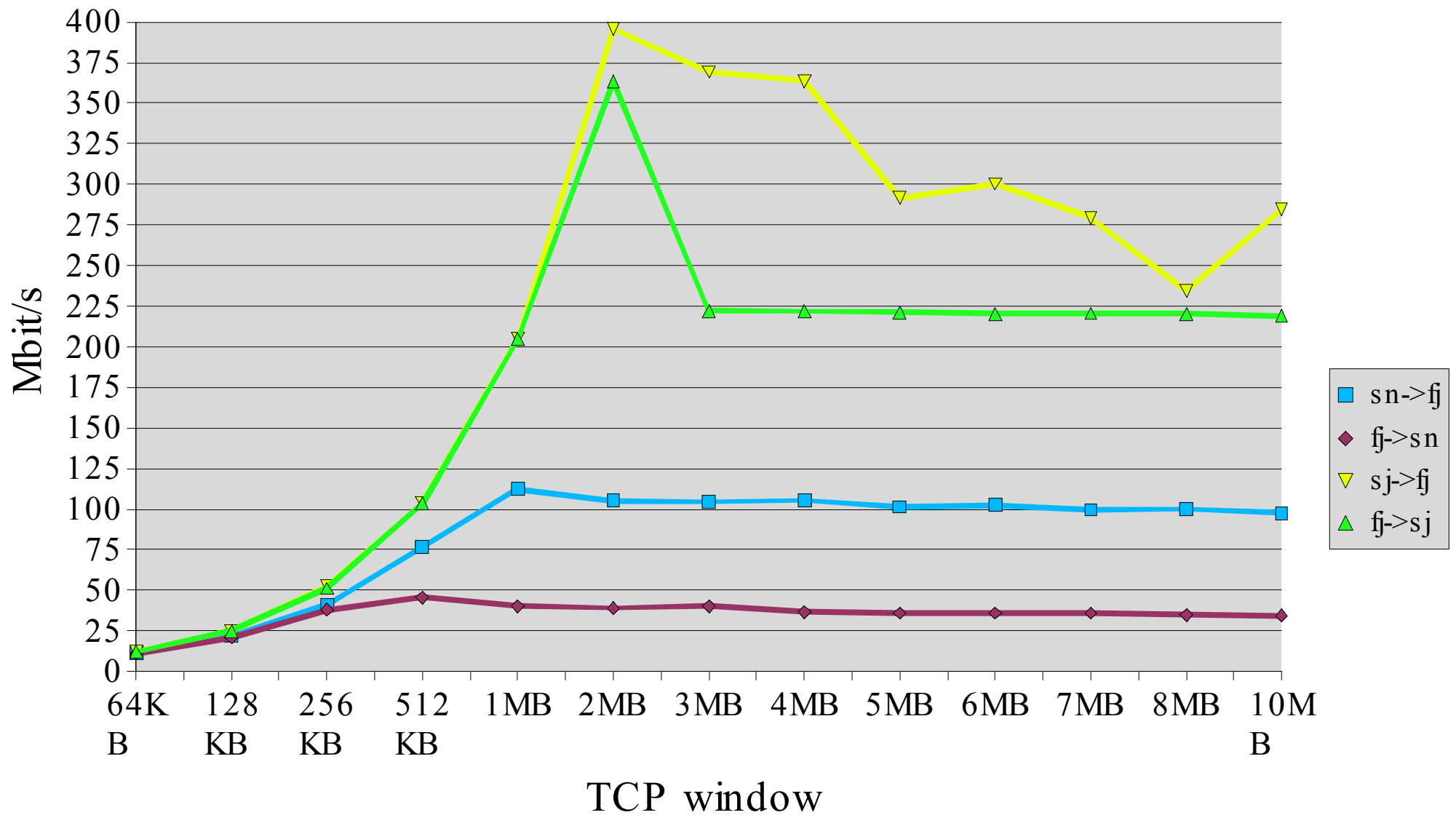
Finland - Norway

Throughput



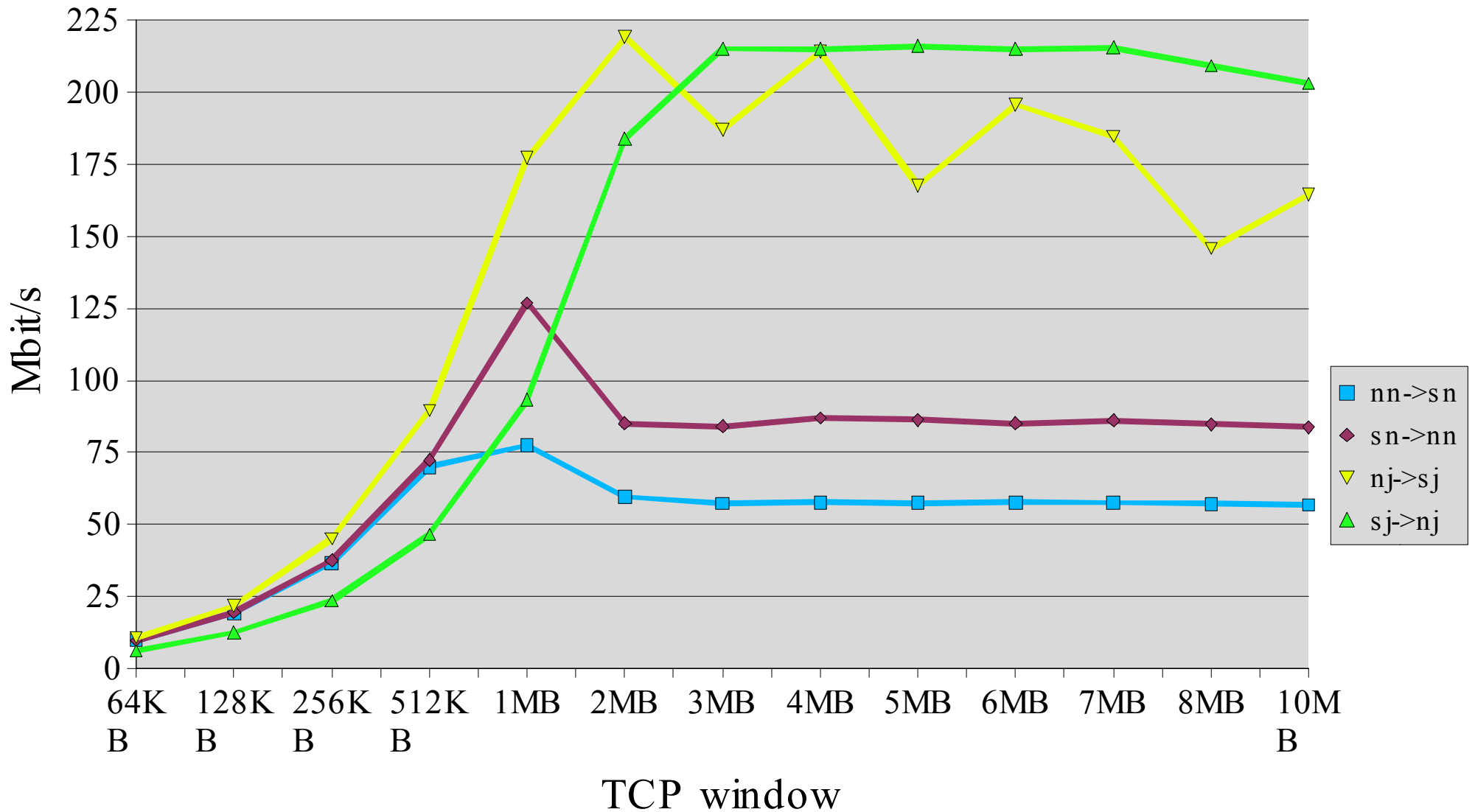
Finland - Switzerland

Throughput



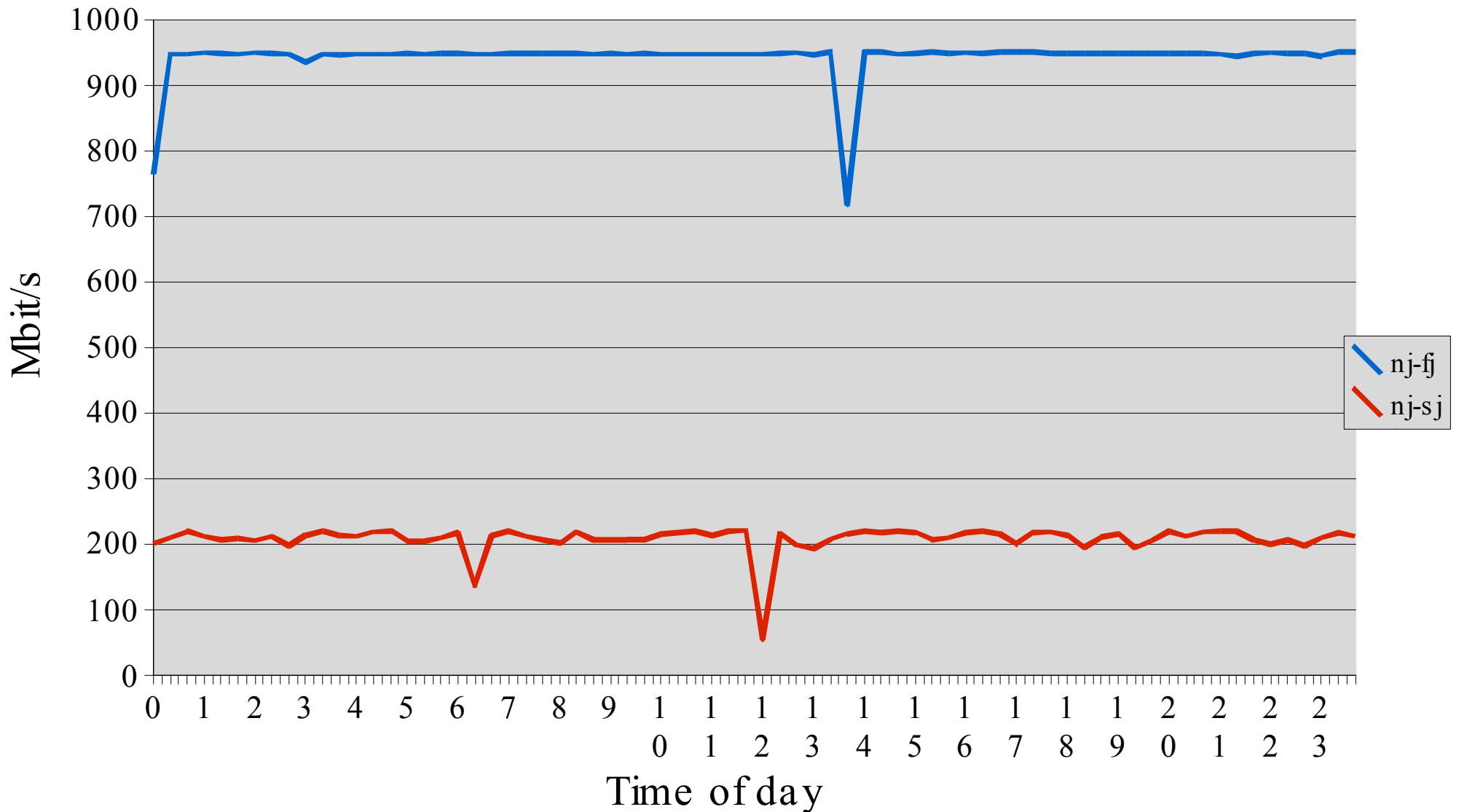
Norway - Switzerland

Throughput



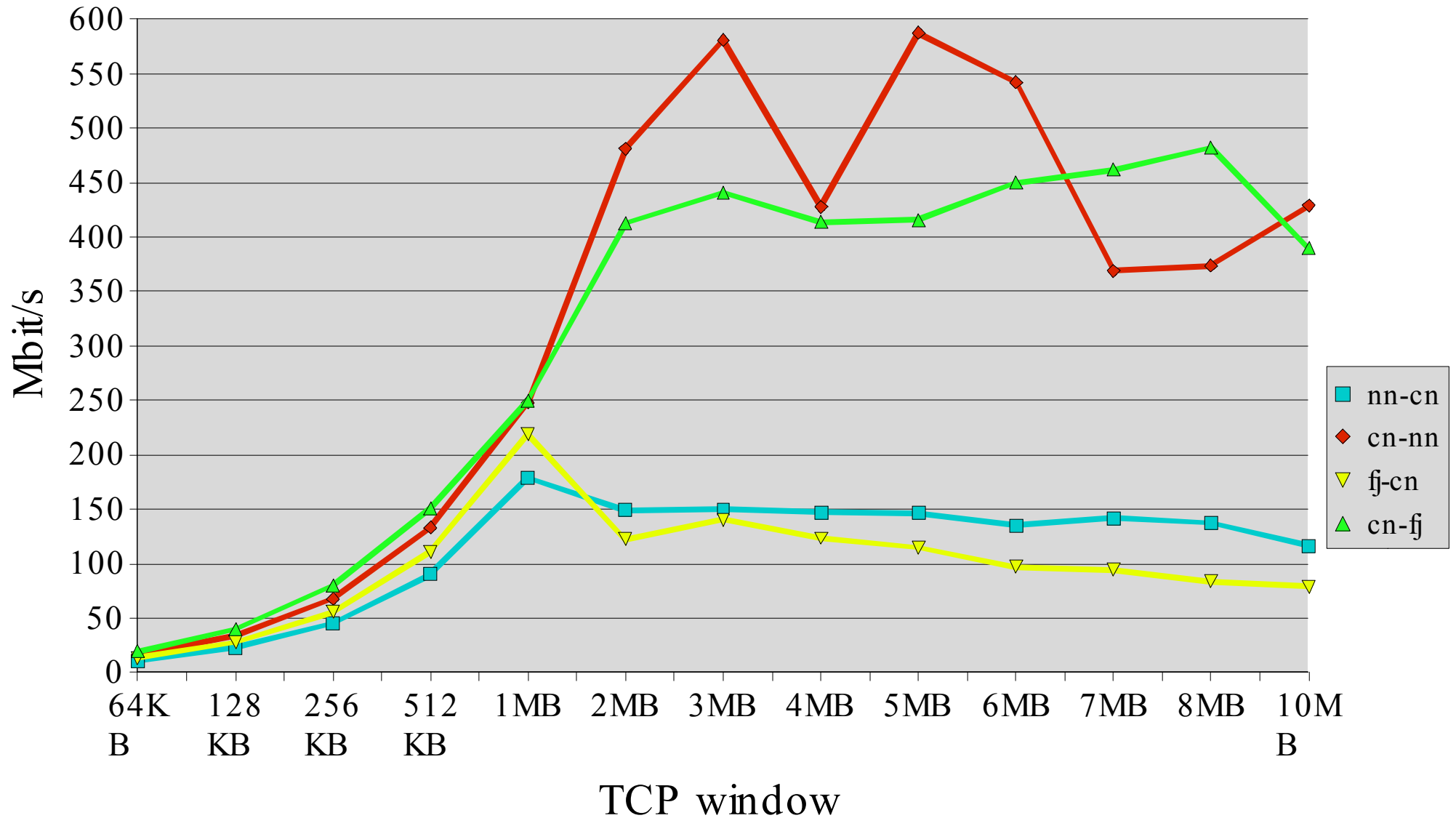
Variations over a day

Throughput



Czech-Finland-Norway

Throughput



Analyzing results

MTU=1500

	Czech	Finland	Norway	Switzerland
Czech		482	601	-
Finland	219		130	48
Norway	181	230		79
Switzerland	-	107	123	

To

MTU=4470

	Czech	Finland	Norway	Switzerland
Czech		-	-	-
Finland	-		568	364
Norway	-	953		219
Switzerland	-	396	215	

To

- Good performance using jumboframes.
- Problems with large TCP window sizes
- Asymmetric results
 - Probably driver and hardware related
- Variance in results to Switzerland caused by SACKs and retransmits

Conclusions and further work

- **Conclusions:**
 - Difficult to achieve high performance
 - Still possible if you use:
 - A good PC
 - A good network card
 - A good driver for the card
 - Large TCP windows leads to performance degradation.
 - Default maximum TCP window size too small for high speed networks.
- **Further work:**
 - Will investigate asymmetric results in more details.
 - Use DAG cards for packet capture to get accurate TCP traces.
- **Full report will be posted on**
<http://www.uninett.no/prosjekt/tcp-revisited/>