

# Poznan Supercomputing and Networking Center

## Testing ATRIUM with demanding applications

*Application of the network traffic from the computation and storage grid environment for network performance evaluation from end-user perspective*

Michal Przybylski (michalp@man.poznan.pl),  
Artur Binczewski (artur@man.poznan.pl),  
Maciej Brzezniak (maciekb@man.poznan.pl)  
Miroslaw Czyrnek (majrek@man.poznan.pl)  
Tomasz Rajtar (ritter@man.poznan.pl)  
Maciej Stroinski (stroins@man.poznan.pl),  
Szymon Trocha (szymon.trocha@man.poznan.pl),  
PSNC, POLAND  
Wim Barbaix (wim.barbaix@alcatel.be),  
ALCATEL BELL NV, BELGIUM

# Poznan Supercomputing and Networking Center

## ATRIUM (IST 1999-20675)

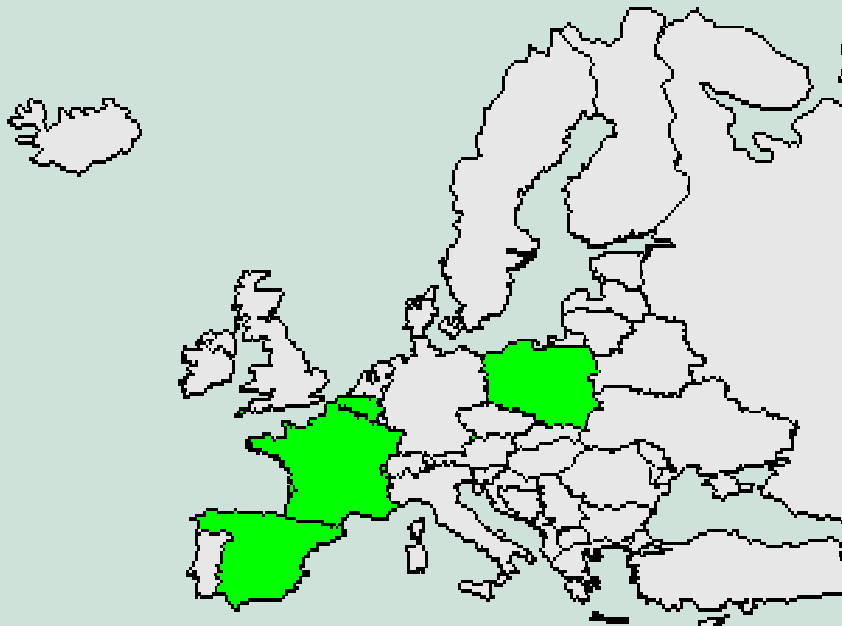
**„A testbed of terabit routers running MPLS over DWDM”**

*Key action: KA IV, Action Line: VII.1.2. RN2,*

***„testbeds for advanced networking and application experiments”.***

*June 2002 - June 2003 (Extended to December 2003), budget 8+ M€*

- ALCATEL
- FT R&D
- FUNDP
- PSNC
- TELEFONICA
- ULG



# Poznan Supercomputing and Networking Center

## New terabit router architecture

- redundant hardware  
(4 SF, 2 RS, 2MS)
- modular software (SOLARIS)
- high availability routing
- scalable platform
- scalable links
- scalable performance
- automated and fast provisioning
- cross layer traffic engineering
- coordinated protection and recovery



**The Alcatel 7770 RCP is a scalable carrier grade core platform, delivering differentiated IP and MPLS services**

# European Testbed



# Poznan Supercomputing and Networking Center

## General approach

- All partners provide PC. PSNC provides necessary servers, HPCs, storage systems
-  provides CCC connection for PSNC and TID; Necessary software is freely available or is provided by PSNC
- Different modes and different network parameters examined
- TE and QOS modifications
- Real applications do not necessarily mean bandwidth consuming
- Subjective/relative QoS metrics allowed

# Poznan Supercomputing and Networking Center

## Expected results



- Technology issues fully understood
  - Router
  - Network
- PSNC and partners can understand GRID applications behaviour with terabit router and large network
  - Delay / jitter
  - Buffering
  - Reordering
- Applications tuning to long fat pipe
- GÉANT can understand how to provision, monitor and debug the CCC service

# Poznan Supercomputing and Networking Center

## Demanding applications...

|         | GRID FTP                          | GAMESS | VIDEO      |
|---------|-----------------------------------|--------|------------|
| Max bw. | <250 Mbit/s                       | ?      | <1.5Mbit/s |
| Delay   | important for lots of small files | ?      | ?          |
| Loss    | low enough to sustain TCP session | ?      | ?          |
| Jitter  |                                   | ?      | ?          |

## ... in different network scenarios

- Empty network
- Congested network
- QoS enabled network
- Faults and recovery

# Poznan Supercomputing and Networking Center

## GRID FTP

- Globus Toolkit v.2.0
- FTP based protocol working in 4 modes:
  - Stream mode (raw data in the data connection + ctrl stream)
  - Extended block mode (file divided into blocks)
  - Parallel data transfer ( $1 \cdot \text{ctrl} + n \cdot \text{data}$ , increase efficiency in long latency networks)
  - Stripped data transfer ( $n \cdot \text{ctrl} + n \cdot \text{data} + n \cdot \text{server}$ , increase efficiency in long latency networks)
- manual control of TCP buffer size

# Poznan Supercomputing and Networking Center

## GRID FTP measurements

- Total experiment time
- Total „data transmission” time
- Avg., max., min. transmission bandwidth
- Time necessary to perform control operations
- Application usability

# Poznan Supercomputing and Networking Center

## **GAMESS - The General Atomic and Molecular Electronic Structure System**

- Wavefunctions computations
- Designed for parallel computers & clusters
- Able to run in distributed environment (TCP)
- Massive communication (peer to peer)
  - input data
  - computation results
  - computation synchronisation

### **Measurements:**

- time necessary to finish computations
- application usability
- validity of the results

# Poznan Supercomputing and Networking Center

## Multimedia streaming

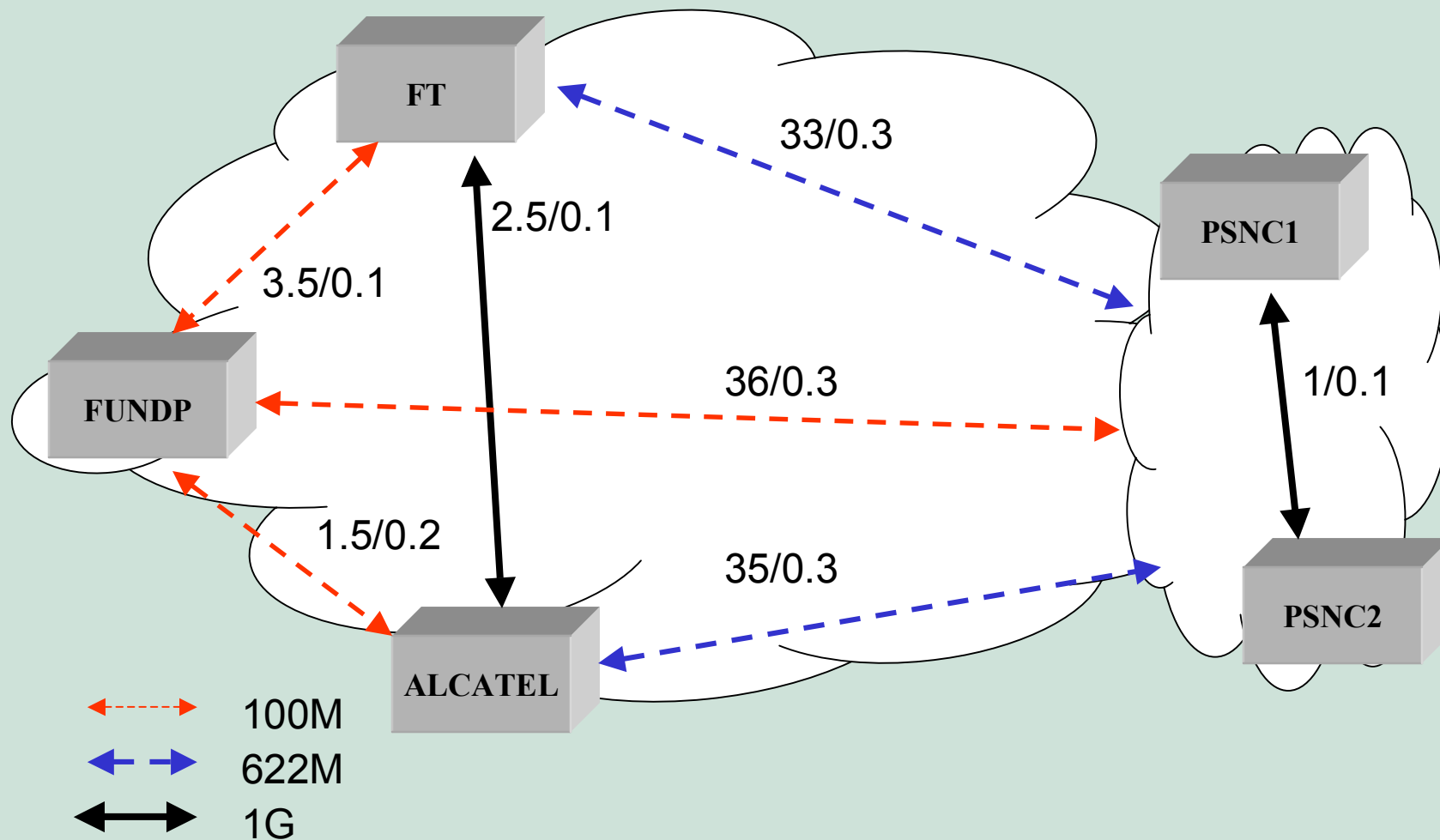
- Proprietary multimedia distribution system
- Based on the independent node concept
- Windows Media Services
- Live streaming
- On-demand content
- Multicast transmission

### Measurements

- Subjective video quality
- Available/used bandwidth
- Packet loss
- Frame loss
- Retransmissions

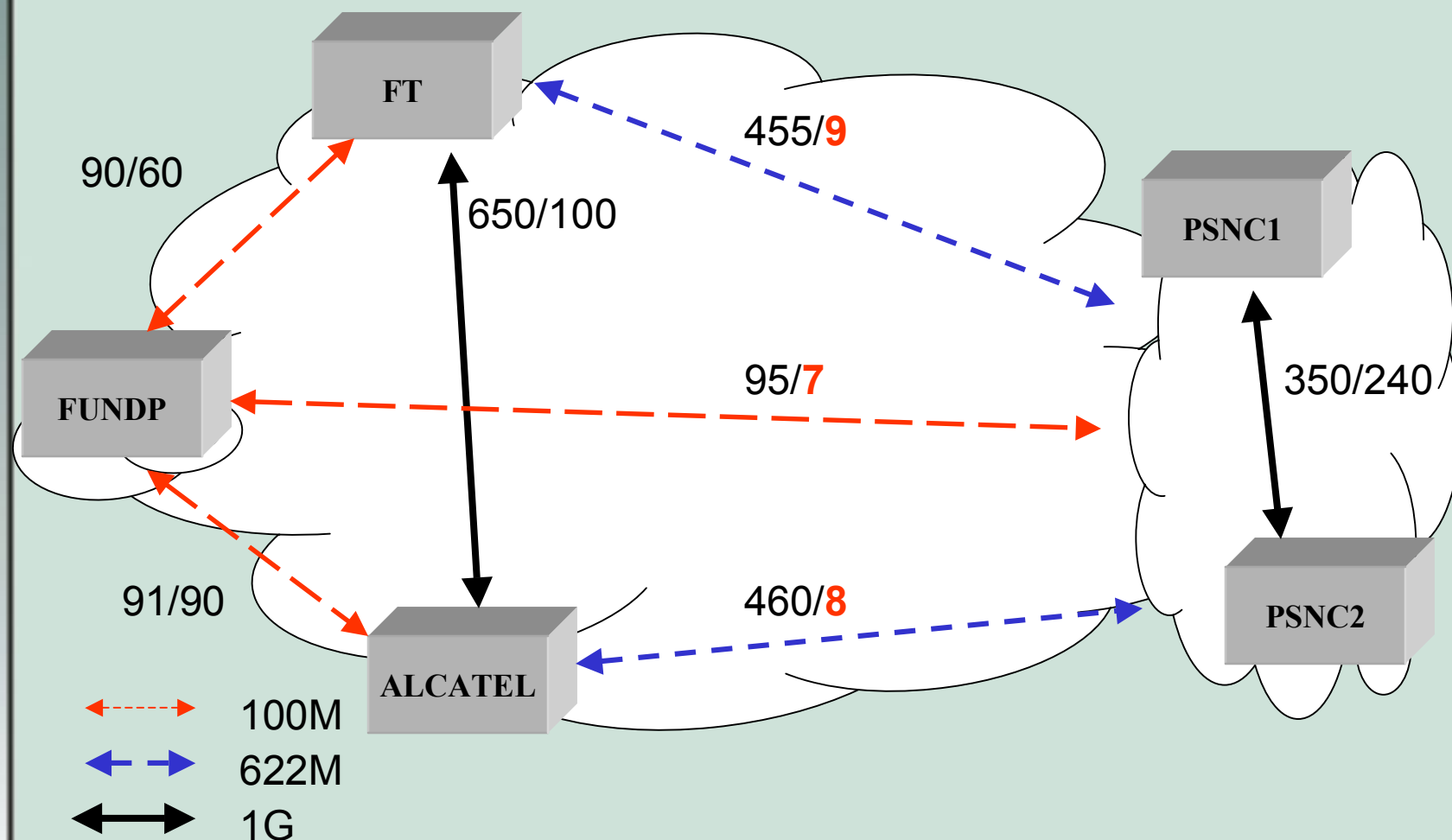
# Poznan Supercomputing and Networking Center

## Delay / jitter on empty network [ms]



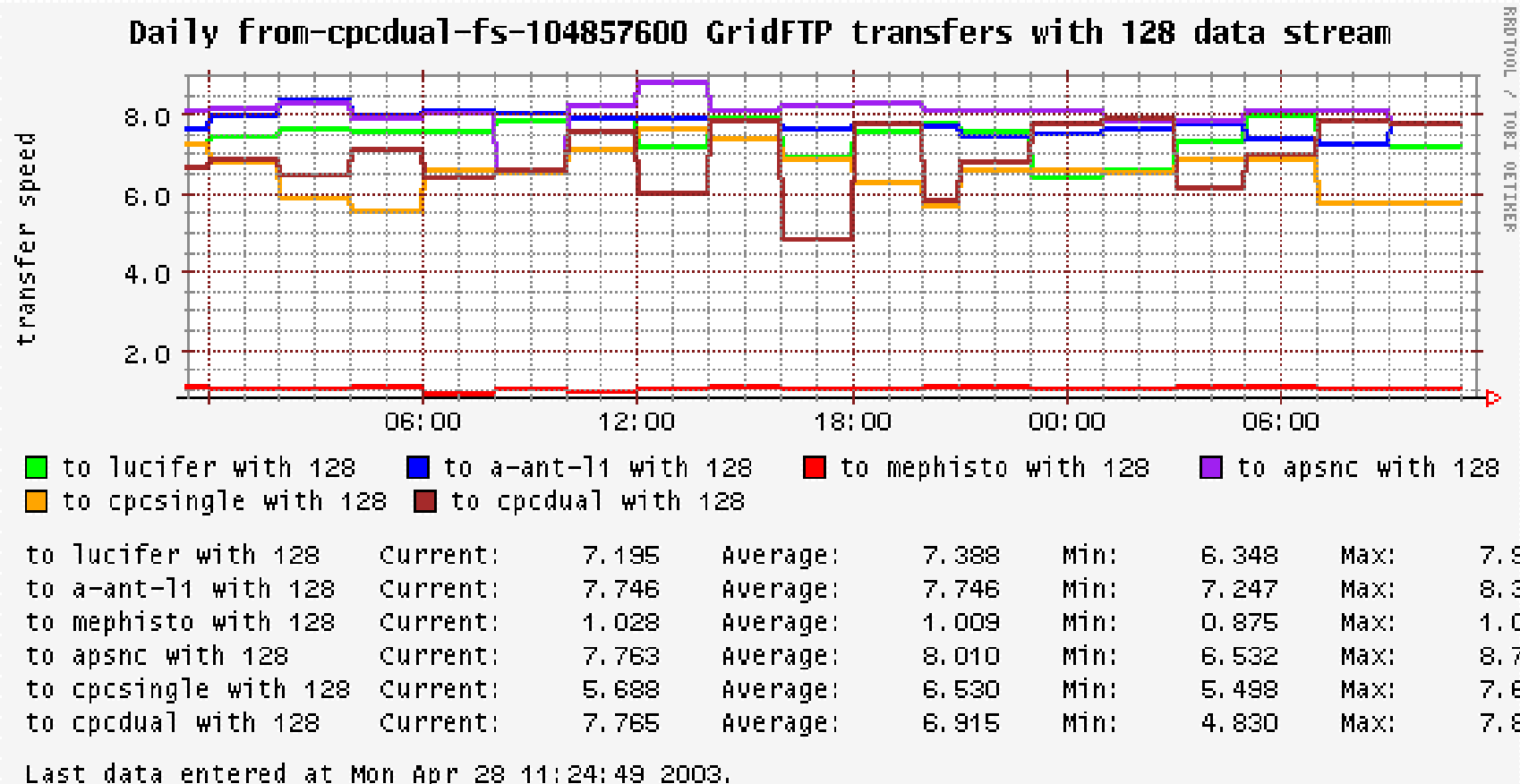
# Poznan Supercomputing and Networking Center

## Empty network: UDP/TCP bw [Mbit/s]



# Poznan Supercomputing and Networking Center

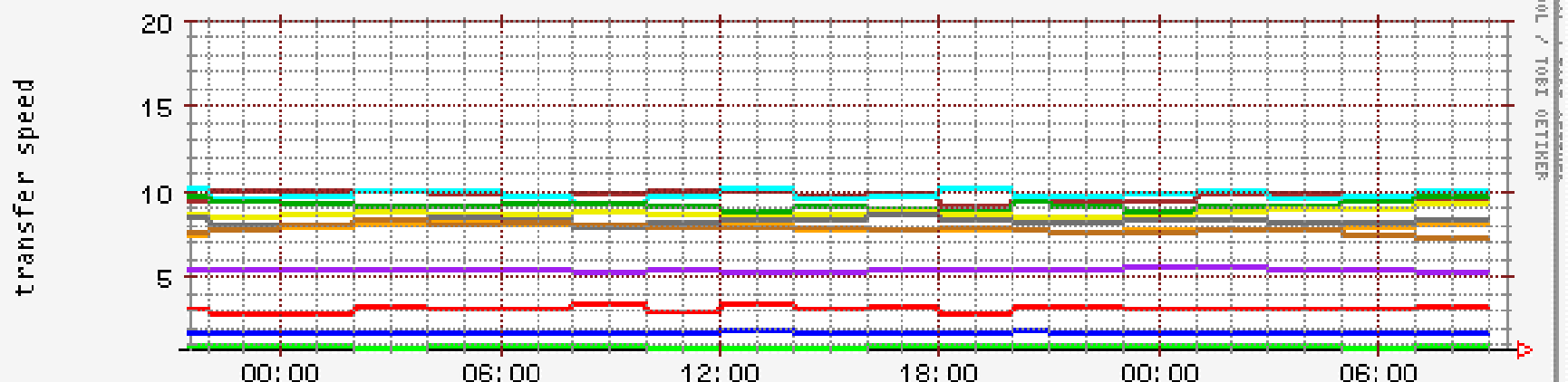
## GRID FTP – parallel transmission



# Poznan Supercomputing and Networking Center

## GRID FTP – parallel transmission...(1)

from-cpcdual-fs-104857600 GridFTP transfers to a-ant-l1 with different numbers of streams

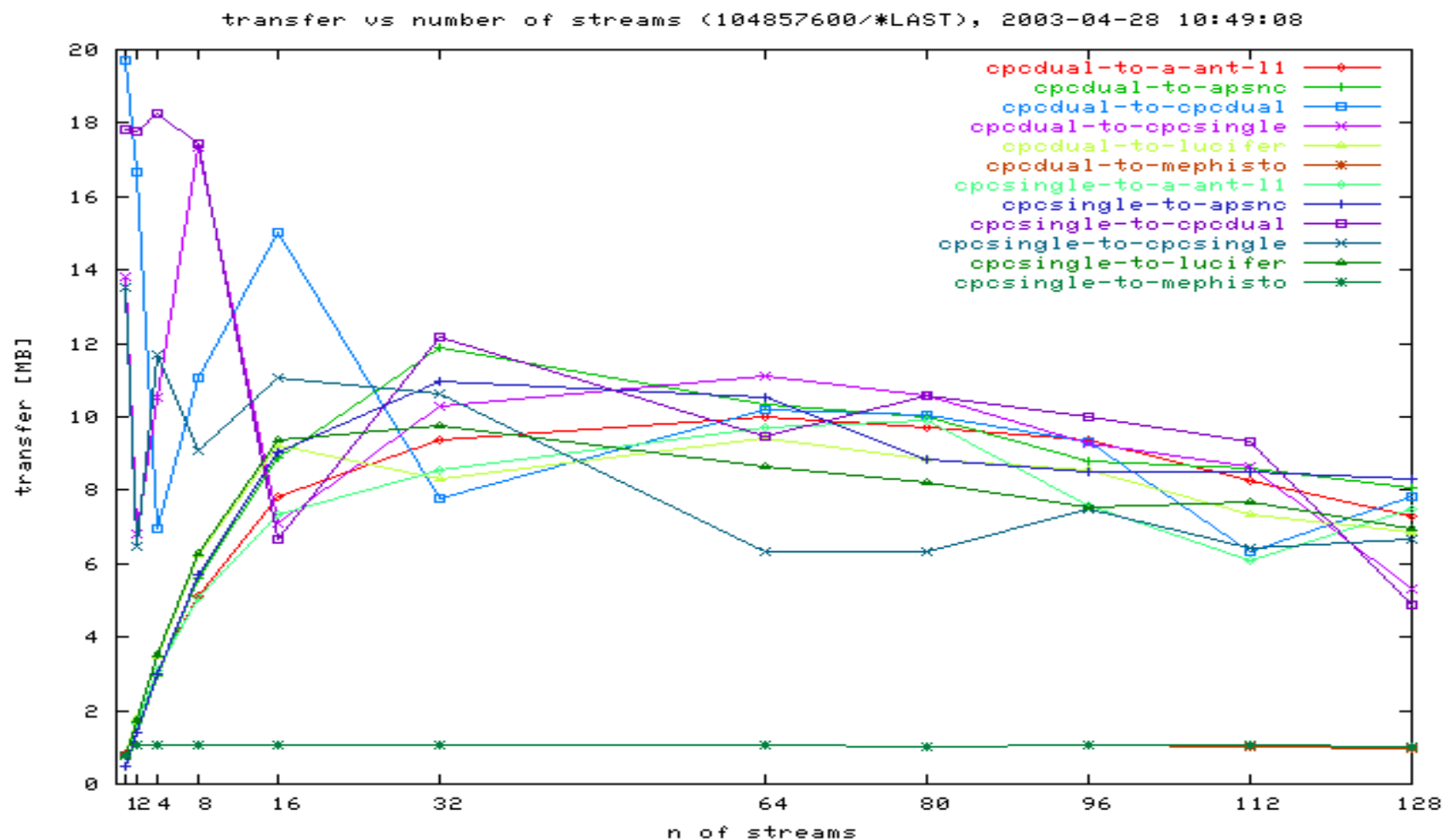


■ to a-ant-l1 with 1    ■ to a-ant-l1 with 2    ■ to a-ant-l1 with 4    ■ to a-ant-l1 with 8  
■ to a-ant-l1 with 16    ■ to a-ant-l1 with 32    ■ to a-ant-l1 with 64    ■ to a-ant-l1 with 80  
■ to a-ant-l1 with 96    ■ to a-ant-l1 with 112    ■ to a-ant-l1 with 128

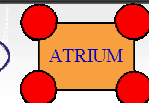
|                         |          |            |          |            |      |            |      |            |
|-------------------------|----------|------------|----------|------------|------|------------|------|------------|
| to a-ant-l1 with 1      | Current: | 783.242 m  | Average: | 779.162 m  | Min: | 735.587 m  | Max: | 799.464 m  |
| to a-ant-l1 with 2      | Current: | 1608.828 m | Average: | 1619.747 m | Min: | 1555.022 m | Max: | 1773.430 m |
| to a-ant-l1 with 4      | Current: | 3094.414 m | Average: | 3024.879 m | Min: | 2668.222 m | Max: | 3318.451 m |
| to a-ant-l1 with 8      | Current: | 5187.033 m | Average: | 5326.799 m | Min: | 5103.000 m | Max: | 5464.431 m |
| to a-ant-l1 with 16     | Current: | 7880.724 m | Average: | 7774.629 m | Min: | 7299.994 m | Max: | 8124.200 m |
| to a-ant-l1 with 32     | Current: | 9350.000 m | Average: | 9666.179 m | Min: | 8982.228 m | Max: | 9990.000 m |
| to cpcdual-fs-104857600 | Current: | 40.040 m   | Average: | 40.700 m   | Min: | 40.000 m   | Max: | 44.000 m   |

# Poznan Supercomputing and Networking Center

## GRID FTP – parallel transmission...(2)



T NC 2003 Zagreb, May 19-24



# Poznan Supercomputing and Networking Center

## Conclusions

- No routing/packet loss/delay/jitter problems on the network composed of 7770RCP router, dedicated optical paths and GEANT CCC connection
- For maximum performance GRID have to use parallel TCP transfers; Optimal application configuration depends on local conditions – file size, link delay, TCP window
- GEANT CCC connection keeps jitter low and does not increase significantly the delay – a sufficient equivalent of dedicated channel (for high cost!)
  - Service monitoring issue
- Testbeds must be hack-protected – 1 month delay due to testbed hack – need to reinstall GRID software and firewall systems