ATL-D-TN-0001

Native 10 Gigabit Ethernet Experiments between Amsterdam and Geneva

Catalin Meirosu[1,2], Piotr Golonka[1,3], Andreas Hirstius[4], Stefan Stancu[1,2], Bob Dobinson[1], Erik Radius[5], Antony Antony[6], Freek Dijkstra[7], Johan Blom[7], Cees de Laat[7]

Submitted for publication in Elsevier's Future Generations Computer Systems Journal

Abstract

The article reports on the experience of building the first trans-European native 10 Gigabit Ethernet testbed. We explain the operation of the 10 Gigabit Ethernet WAN PHY and why validation in the field is important. Tests on a 1700km-long WAN PHY connection using commercial traffic generators are described. Measurements were also made using servers equipped with 10 Gigabit Ethernet network interface cards and attached to the network. Sustained single stream TCP transfer rates of more then 5.4 Gbps for over 14 hours have been achieved.

---

[1] CERN, the European Organization for Nuclear Research, Geneva, Switzerland
[2] The "Politehnica" University, Bucuresti, Romania
[3] Institute of Nuclear Physics, Polish Academy of Sciences, Krakow, Poland
[4] CERN openlab for DataGrid applications, Geneva, Switzerland
[5] SURFnet bv, The Hague, The Netherlands
[6] NIKHEF, Amsterdam, The Netherlands
[7] University of Amsterdam, The Netherlands

1 Introduction

Historically, SONET/SDH (Synchronous Optical Network/Synchronous Digital Hierarchy) is the technology of choice for the wide area networks - WAN. Packet over SONET/SDH is the dominant technology in today's core routers [1]. High reliability and extensive operation and management features required by the telecom operator environment make SONET/SDH an expensive technology to deploy and exploit. The prices of very high-speed external connectivity (especially over the WAN) continue to remain out of the reach for a large majority of users in the research and commercial areas.

Today's universities and research centres are in the process of upgrading their internal networks to support traffic at 10 Gbps. The demand is mainly driven by high performance computing, various GRID applications and storage area networks (SAN). New paradigms are being developed to include network resources in the optimization process for distributing computing power worldwide [2].

The IEEE together with industry have acknowledged the increasing demands for bandwidth. The broad adoption of the Ethernet standard led to mass production and commoditization of Local Area Networks (LAN), hence reduced prices and a technology affordable by everyone. Now the same technology that has dominated the LAN is moving into the WAN, however there is still an open debate as to whether native Ethernet is suitable for deployment. Our work indicates that Ethernet can provide large bandwidth for high-performance networking applications at a trans-European level. It should be noted that line costs are still likely to dominate the price for external connections for the near future.

1.1 The 10 Gigabit Ethernet WAN PHY

Ethernet is a communication protocol operating at the first two layers of the Open Systems Interconnect (OSI) model. The IEEE 802.3 standard defines both the Layer 2 communication protocol (Media Access Control - MAC) and the Layer 1 transmission methods for Ethernet.

In June 2002, IEEE adopted the 10 Gigabit Ethernet (10 GE) 802.3ae standard [3]. The most obvious improvement over previous Ethernet standards is the increase in transmission speed to 10 Gbps. In addition, IEEE 802.3ae lays the foundation of a new kind of Ethernet, one capable of spanning worldwide distances.

All the previous Ethernet standards included MAC-level support for the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol supporting half-duplex communications. The use of CSMA/CD practically confined Ethernet to the LAN by limiting the maximum length of the connection in order to detect a collision on the shared transmission medium. The 10 GE standard frees Ethernet of the CSMA/CD legacy by only supporting full-duplex communications and optical fibre as transmission media. The length of a point-to-point Ethernet connection is only limited by the characteristics of the optical components used for transmitting and propagating the signal.

A transceiver is a physical layer device that is directly attached to the transmission media. The 10 GE standard specifies two major types of transceivers for Ethernet connections: the LAN PHY and the WAN PHY. The standard defines a 40km reach for both PHYs when using 1550nm lasers over single mode fibre. The LAN PHY transceivers have a data rate of 10.3125 Gbps. It is targeted mainly at local area networks, but recent developments [4] have demonstrated a reach of over 250km using optical amplifiers. It is generally accepted that 600 km is the limit after which the signal has to be regenerated due to attenuation and dispersion on the optical fibre. The signal regenerators deployed in the WAN assume SONET/SDH rate and framing of the incoming signal. However, the transmission rate and framing of the LAN PHY is incompatible with the current signal regenerators installed on the WAN infrastructure.

The WAN PHY has been defined to be compatible with SONET/SDH in terms of data rate and encapsulation method. It uses the same transmission rate, 9.95328 Gbps, for a payload capacity of 9.58464 Gbps, using the STS-192c / VC-4-64c frame format. The WAN PHY therefore enables the transport of native Ethernet frames over legacy long haul network infrastructure.

It should be noted that the 10 GE standard does not guarantee the strict interoperability of the WAN PHY with SONET/SDH equipment. "A 10GBASE-W interface is not intended to interoperate directly with interfaces that comply with SONET or SDH standards, or other synchronous networks. Such interoperation would require full conformance to the optical, electrical, and logical requirements specified by SONET or SDH, and is outside the scope and intent of this standard."[3] The optical characteristics of the WAN PHY lasers are relaxed by comparison to the SONET/SDH standard. The clock that is used as a reference for the WAN PHY transmission is allowed to be less accurate (20 ppm instead of 4.6 ppm). Certain bits of the SONET/SDH management overhead are unused or have default values [3, 5]. The IEEE 802.3ae standard specifies that a SONET frame carrying WAN PHY payload must be identified by setting the C2 byte (of the SONET management overhead) to 0x1A. The SONET/SDH ring topology and 50ms restoration are explicitly excluded from the WAN PHY specification.

The IEEE defines [3] a specific piece of equipment (the Ethernet Line Terminating Equipment - ELTE) for connecting the WAN PHY to SONET/SDH networks. However, no manufacturer has, to date, built an ELTE. The direct attachment of the WAN PHY to legacy infrastructure is the only solution currently available for transmitting native 10 GE frames long haul. Experiments performed at CANARIE [6] have shown successful operation in the laboratory.

The WAN PHY provides, for the first time, a native gateway from the LAN to the WAN. The 10 Gigabit Ethernet Alliance sets three goals for the WAN PHY [7, 8]:

- Direct attachment of the 10GE WAN PHY to today's SONET/SDH transponders, providing access to the installed base of DWDM (Dense Wavelength Division Multiplexing)
- Direct attachment of the 10GE WAN PHY to an OC-192 tributary interface
- Direct attachment of the 10GE WAN PHY to emerging ITU OTN transponders, providing access to the next era of wide area infrastructure

We have demonstrated that the first two goals can be met and the 10 Gigabit Ethernet WAN PHY is ready for deployment in the field. The WAN PHY enables a long haul point-to-point connection operating at more then 9 Gbps usable data rate. In addition to demonstrating the technology, we used this opportunity to determine the behaviour of higher layer transport protocols at this bandwidth. Most previous studies of TCP protocols were, until recently, only theoretical for bandwidths higher then 2.5 Gbps (OC-48 equivalent).

2 Test setups

SURFnet, the Dutch research network, is operating an OC-192 lightpath between CERN, Geneva and the NIKHEF institute in Amsterdam. Global Crossing delivers the lightpath. This entire bandwidth was allocated to our experiments for more then 6 weeks, from mid-August to October 2003. SURFnet uses two Cisco ONS 15454 devices [10] at the connection's endpoints. Each ONS contains two SONET OC-192 cards: one interfacing to the carrier and one interfacing a 10 GE WAN PHY port in the laboratory.

We defined two major test scenarios for connectivity tests:

- WAN PHY over DWDM
- WAN PHY over a SONET circuit

In the first scenario (figure 1) we have directly attached the WAN PHY to the DWDM equipment at the carrier Point of Presence (POP) in Geneva and Amsterdam.

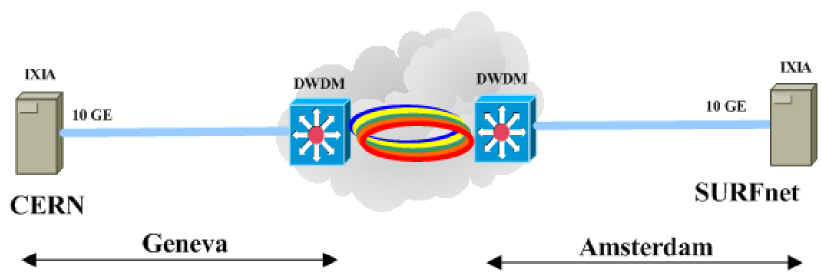

Figure 1: Traffic generator WAN PHY connected to DWDM equipment

The second scenario (Figure 2) attached the WAN PHY to the OC192 port in the ONS and the other OC192 port to the DWDM equipment.

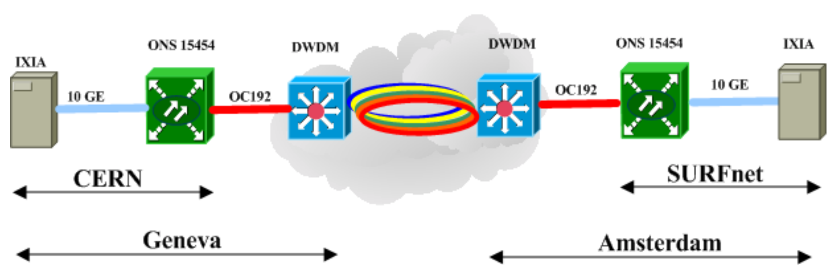

Figure 2: Traffic generator WAN PHY connected to a SONET OC-192 circuit

Ixia traffic generators [11] were used for basic characterization of the line. They enabled us to load the line at full capacity. They are also capable of measuring one-way latency with 100 ns accuracy due to the GPS synchronization of their timestamps.

For testing higher layer protocols we have interconnected two small LANs (Figure 3). Tests were also performed with the Force10 switch connected directly to the DWDM equipment.

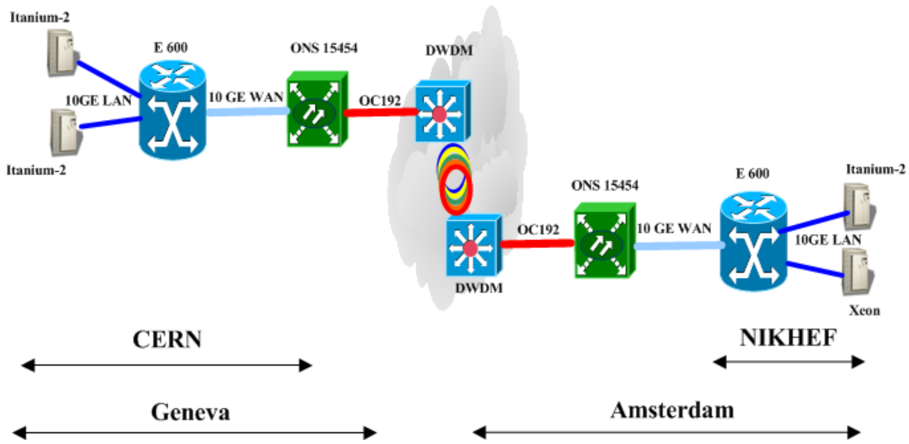

Figure 3: Local Area Networks interconnected through the WAN PHY

The LAN at CERN was composed of two HP RX2600 systems [12] configured as dual processor Itanium-2 with 4 GB of RAM and using Intel Pro/10Gb adapters [13] to connect to a Force10 E600 [14] switch. The LAN at NIKHEF was composed of one HP RX2600 system having a configuration similar to the CERN system and one Intel Xeon-based server. All computers were running the Linux operating system, kernel 2.4.21 with and without the net100 patch [15].

3 Results of line characterization experiments

We started by characterizing the connection in terms of latency, throughput and packet loss. Traffic generators were used to send and receive traffic at full line speed. To determine the packet loss, we have applied constant 100% load for long time intervals (more then 90 hours for each network configuration). The traffic was a mix of 1518 bytes frames (90%) and 64 bytes frames (10%). The payload of the 1518 bytes packets was populated with the CJPAT pattern [3]. The payload of the 64 bytes packets was populated with the CRPAT pattern [3]. Both patterns are related to the streams used in Bit Error Rate (BER) measurements. The traffic generators reported no packet loss: 3.6572 * 10^14 bytes transmitted in 91 hours and 2.47 seconds. This is equivalent to a bit error rate lower then 10^-15, three orders of magnitude better then the specifications of the 802.3ae standard.

The success of the experiments with traffic generators enabled us to proceed to the next step and connect the switches to the network. The new network topology is presented in Figure 3. Using the same mix of traffic, we have obtained comparable results: 3.9524 * 10^14 bytes transmitted in 98 hours, 48 minutes and 25.3 seconds with no packet loss. The average round trip time, measured on a looped back connection at a load of 100%, was 17.079 ms. A separate measurement of the one-way latency yielded 8.6 ms, in perfect agreement with the previously determined round trip time.

Once the basic performance parameters of the connection were understood we could proceed to test the network with traffic generated by PCs.

4 Basic network protocol tests

The theoretical bandwidth available on the PCI-X bus (133 MHz, 64 bits) is 8.512 Gbps. Hence this is the upper limit for traffic generated by a network adapter connected to the fastest PCI slot in today's commercially available servers. Since we knew that we could not saturate the WAN PHY connection from a single machine, the single stream tests represent more the performance of the server than the quality at the network connection. However, tests running over extended time intervals provide an additional validation of the connection.

For generating raw Ethernet frames from a server PC we have developed a simple program that writes (or reads) data to (from) a Linux socket as fast as it can. In figure 4 we show the throughput measured by the program. The data was transmitted from the computers in NIKHEF and received at CERN. Only one stream was active on the connection at any moment during these tests. We have observed no packet loss during these tests.

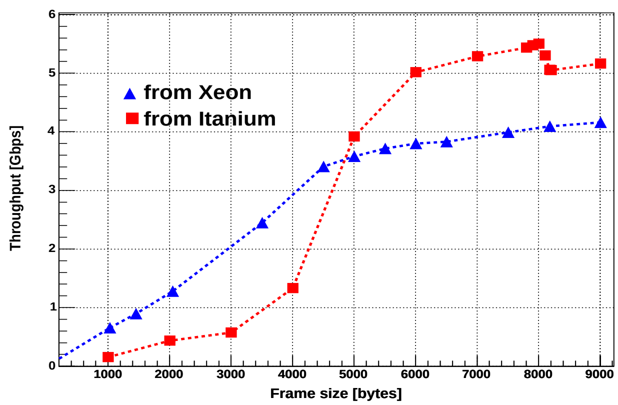

Figure 4: Raw Ethernet streaming results

As reported in [16], a message size around 8000 bytes gives an optimal throughput when sending from the Itanium-2 server. The Xeon-based computer achieves maximum bandwidth for a message size of 9000 bytes. The E600 switch supports jumbo frames up to 9252 bytes. However, increasing the MTU from 9000 to 9252 bytes would not make a significant difference. The throughput achieved by the Itanium-2 machine was 5.163 Gbps for 9000 bytes MTU. The Xeon achieved 4.130 Gbps. The combined throughput of the two streams rises above the capacity of the line. The outgoing port of the switch drops the traffic that exceeds the capacity of the line. Figure 5 shows a peak of around 9.2 Gbps for a 20 minutes interval when both streams were active at the same time.

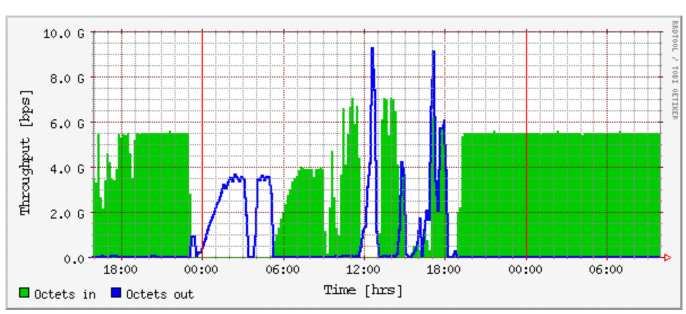

Figure 5: Bandwidth graph showing the rate at the output port of the switch in Amsterdam on September 23

The basic protocol tests have shown that the network connection is no longer the bottleneck. The PCI bus and the software overhead in the networking protocol stack of the operating system limit the traffic that can be generated by a single computer.

5 TCP experiments

TCP carries the bulk of today's Internet traffic. Several authors have shown that the commonly met implementations of the protocol have poor performance on high bandwidth, long latency connections when packet loss occurs [17, 18, 19]. However, the studies were based more on mathematical models than high-speed measurements. Providing solutions for this problem is currently a topical research area.

In addition to the standard TCP implementation in the Linux operating system, we have used the modified TCP stack included in the net100 distribution. Net100 implements the High Speed TCP (HSTCP) draft proposed by IETF [20]. The traffic was generated and measured by the Iperf program, a quasi-standard in the community. We have used Iperf version 1.6.5 with the modifications described on the University of Amsterdam DataTAG project website [21]. The patch was required in order for Iperf to work on the 64-bit architecture of the Itanium-2 processors. All the Iperf transfers were memory-to-memory transfers.

The aims of the TCP tests were:

- to study the scalability of the protocol with the number of parallel streams originating in a single computer
- to study the behaviour of a single stream when background traffic is applied in a way to induce packet loss
- to compare the reactive response to packet loss of standard TCP and HSTCP

5.1 Single TCP stream

We have started by determining the system performance for a single TCP stream between the Itanium-2 server at NIKHEF and the Itanium-2 server at CERN. The standard TCP implementation in the Linux 2.4.21 kernel was used for taking these measurements.

The round trip time of the Geneva-Amsterdam connection was 17ms. The bandwidth-delay product (BDP) is 17ms * 9.29Gbps = 19.74 MB. If we consider 6 Gbps the maximum achievable bandwidth, the BDP decreases to 12.75 MB. Therefore, the maximum throughput should be achieved for a TCP window size between 13 and 20 MB. Figure 6 presents the results of the tests for different MTU and TCP window sizes.

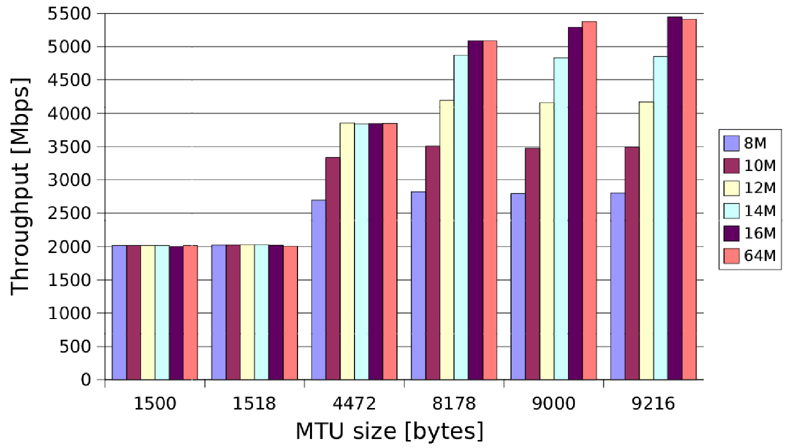

Figure 6: Single stream TCP throughput

Each bar in Figure 6 represents a different trial. Each trial lasted for 120 seconds. The maximum throughput of 5.447 Gbps was achieved for an MTU of 9216 bytes and a TCP window size of 16 MB. The values we have chosen for the TCP window size had no effect on the throughput for an MTU corresponding to the maximum frame size allowed by the Ethernet standard. The system limited the throughput at a much lower value (2.057 Gbps) than the one allowed by the window size corresponding to the BDP. This illustrates the importance of jumbo frames in the high speed networking applications area.

The standard TCP protocol reduces the transfer rate even if a single packet is lost. In the Internet, a packet may be dropped due to congestion in one of the computers at the end nodes or because of congestion on the network. A packet may also be dropped due to a transmission error detected by one of the CRC sequences embedded in the packet. Transmitting a TCP stream for a long time interval and recording the Iperf measurement of the throughput at the receiver provides an indication of the reliability of the connection. We have kept a single TCP stream running for almost 15 hours. The throughput history is presented in the Figure 7. The vertical axis starts at 4900 Mbps.

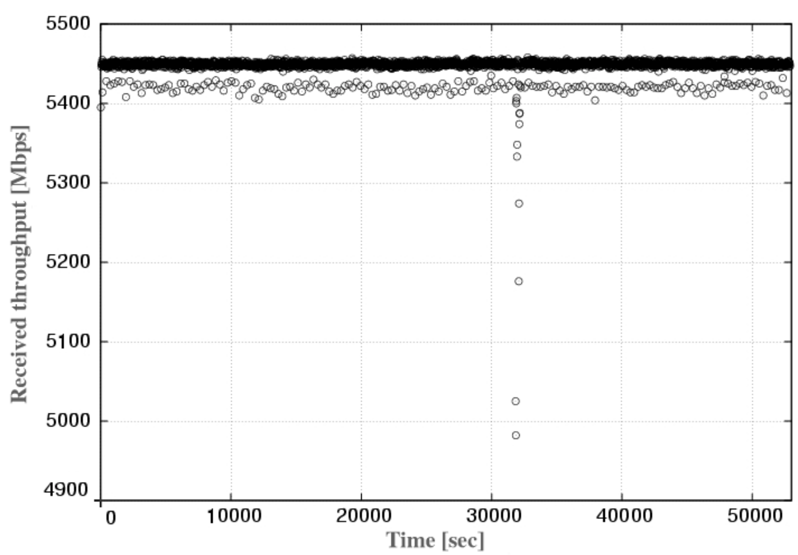

Figure 7: Sustained single stream TCP throughput

The average throughput over this time interval was 5.448 Gbps. Out of 2700 data points collected every 10 seconds, about 93% correspond to the throughput between 5.440 and 5.460 Gbps; about 7% had the throughput between 5.400 and 5.440 Gbps; only 10 points had the throughput below 5.400 Gbps, with minimum at 4.982 Gbps, which is around 90% of the average throughput.

These results indicate that the point-to-point long haul connection based on the WAN PHY is able to support real data transfers between servers at the maximum speed allowed by the PC hardware. The servers provided a stable transfer rate for an extended time interval.

In today's production networks it is unlikely that a single stream will be allowed to claim 5 Gbps of the bandwidth. We have tried to subject the TCP implementation to a scenario closer to what may occur in a real network. A single stream was sent by the Xeon server in NIKHEF and received by the Itanium-2 machine at CERN. A stable throughput was achieved at around 1900 Mbps for a TCP window size of 16 MB and a MTU of 8000 bytes. Each trial consisted in a different background traffic load applied for a short time interval by the traffic generator while the TCP connection was in stable condition. Figure 8 presents the results of the experiments.

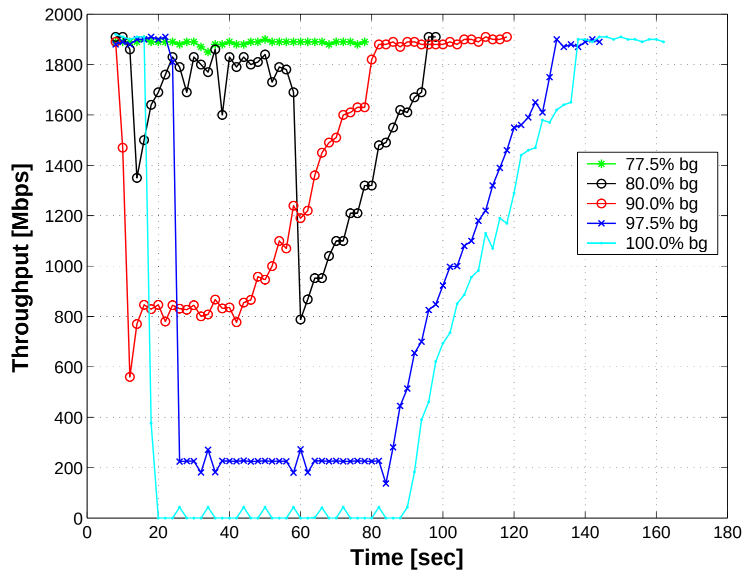

Figure 8: Behaviour of a single TCP stream in the presence of background traffic

The bandwidth of the TCP stream represents roughly 20% of the line capacity during each trial. The two second average rate of the TCP stream seems constant, but in practice the packets are sent on the network in bursts. The traffic generator sends constant bitrate traffic (having a constant time interval between consecutive packets) at increasing loads for each trial. Therefore even if the average rate of the two streams does not surpass the capacity of the line, temporary congestion is possible. The background traffic was switched on around the 10 seconds mark (variable for each measurement), and than switched back off again somewhere between the 50 and 90 seconds mark.

The temporary congestion is handled well by the switch: no packet loss occurs until the background traffic increases above 80% of the line speed. The throughput of the TCP stream remains constant during the induced congestion interval and occupies the available bandwidth.

5.2 Multiple concurrent TCP streams

The objective of the multiple TCP stream test was to determine whether the use of more then one TCP stream (at the same time) has an impact on the aggregated total throughput. The MTU was set to 9216 bytes. The TCP window size was set to 16 MB.

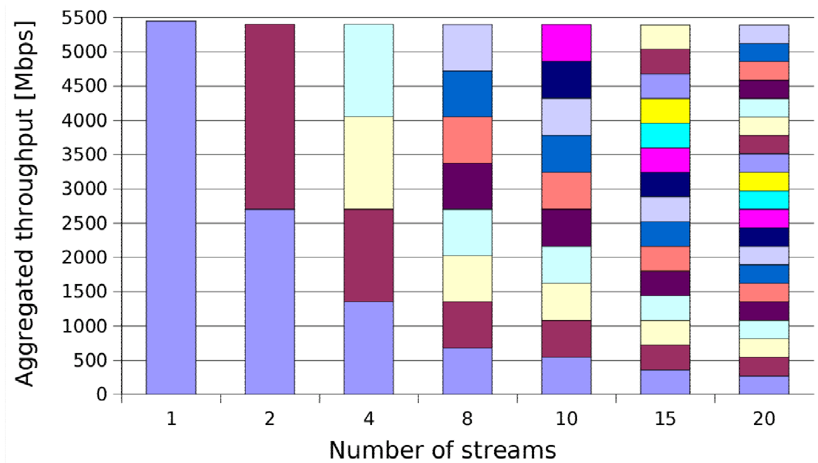

Figure 9: Throughput of multiple TCP streams

Figure 9 presents the throughput achieved by up to 20 TCP concurrent streams. The number of streams has little influence on the aggregated throughput. This indicates that the operating system adds little per-stream overhead. The TCP implementation scales well for a relatively small number of streams.

6 New transfer protocols

The behaviour of the TCP protocol in the occurrence of packet loss has been demonstrated to be inadequate for high speed, high latency networks. This is a topical research area and several authors have proposed solutions that shorten the recovery time. High Speed TCP (HSTCP) is the solution that is on the road of standardization by IETF [20]. HSTCP proposes a modification to TCP's congestion control mechanism for use with TCP connections with large congestion windows.

In Figure 10 we compare the responsiveness of the HSTCP implementation in Net100 to the standard TCP NewReno from the 2.4.21 version of the Linux kernel. We start a TCP stream between two hosts using Iperf. Then we create congestion by sending a UDP traffic burst for 2 seconds, from a different source to the same receiver. The switch that connects both traffic sources is forced to drop packets on the WAN interconnect port because the combined bandwidth of the TCP and UDP streams exceeds the available bandwidth.

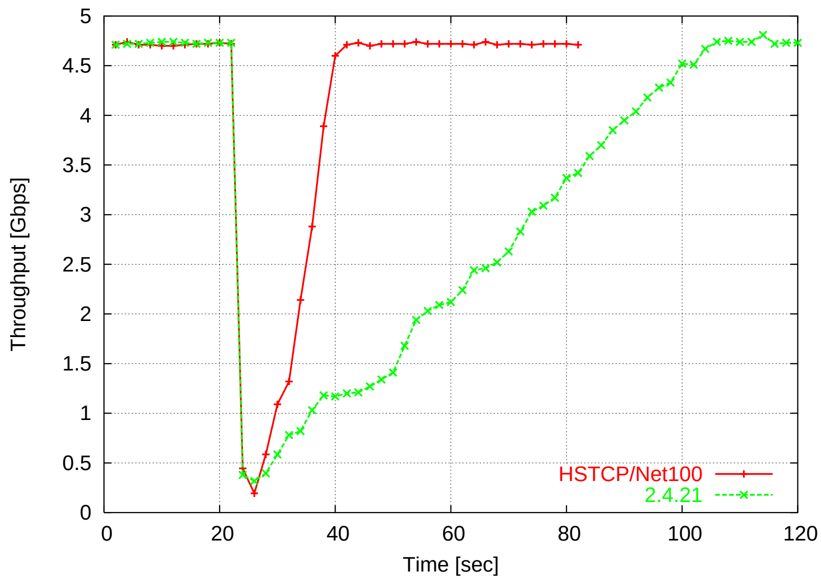

Figure 10: HSTCP versus stock TCP recovery time

The recovery time of the HSTCP implementation is more than 9 times faster compared to standard TCP.

The UDP based Data Transfer protocol (UDT) [22] was developed by the University of Illinois at Chicago to address the TCP problems over high BDP networks. UDT implements an end-to-end congestion control mechanism that combines rate based and window based control mechanisms to enable high performance transfers with intra-protocol fairness and TCP-friendliness [23]. The rate limiting mechanism is based on controlling the inter-packet time. Packets are sent on the network at equal intervals, once the available bandwidth has been estimated during the slow start process. A timer set to a constant interval of 0.01 seconds triggers the rate control mechanism. The number of unacknowledged packets is limited by the flow control mechanism. The flow control uses a sliding window of size calculated using the following formula [23]: W=W*alpha + AS***(**RTT + SYN**)***(1-alpha), where AS is the packet arrival speed calculated by the receiver and transmitted to the sender in the ACK packets, alpha has a constant value of 0.875, RTT is the round trip time and SYN is the timer interval for the rate control mechanism, constant to 0.01 seconds

We have tested the UDT to determine whether the performance of the implementation makes it comparable to TCP-based transfers on our uncongested 9.2 Gbps throughput connection.

Figure 11 shows the UDT transfer rate between Itanium 2 systems located at CERN and NIKHEF.

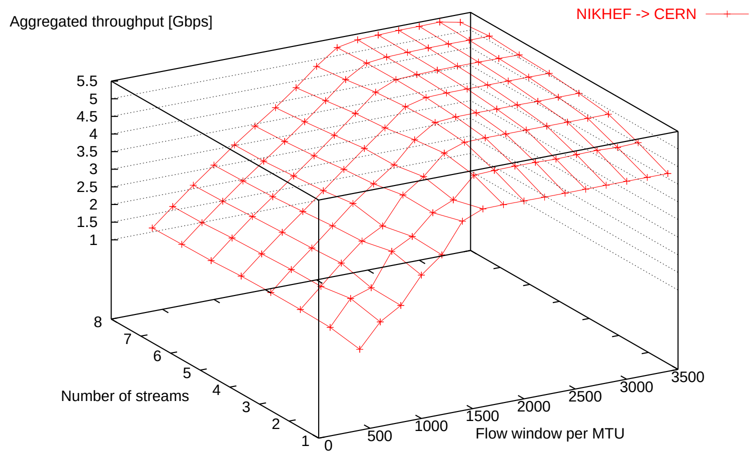

Figure 11: Throughput of UDT transfers

The throughput is about 5.4 Gbps, which makes it comparable in terms of bandwidth to the TCP transfers on the same computers. Since there was no packet loss during the transfers, we could not assess its advantage over TCP.

7 Conclusions and future work

Our experiments have demonstrated that the Ixia and Force10 implementation of the 10 Gigabit Ethernet WAN PHY interoperates successfully with the legacy long haul infrastructure used. However, we do not want to generalize this statement because the optics of the WAN PHY transceiver we have used exceed the specifications of the IEEE 802.3ae standard. The WAN PHY will be soon made available as XENPAK [24] pluggable optical devices. We are currently planning to field test these devices.

Steady TCP transfers at more than 5.4 Gbps over long time intervals have been achieved over an uncongested network. New variants of TCP have been demonstrated to have faster recovery time hence largely improving the throughput in the occurrence of packet loss. Multiple TCP streams originating on the same computer do not improve the overall transfer rate, which is currently limited by the hardware and software architecture of the computer.

Acknowledgements

We acknowledge the funding support of the EU projects ESTA (IST-2001-33182) and DataTAG (IST-2001-32459). We also acknowledge the generous contribution in equipment and technical assistance from Force10, Ixia, Cisco, Hewlett-Packard and Intel. SARA provided useful technical assistance at the Amsterdam end of the connection. Global Crossing provided technical assistance during our frequent change of network configuration. René Hatem from CANARIE and Wade Hong from the Carleton University in Ottawa gave the CERN team the opportunity to experiment with the WAN PHY in the CANARIE lab prior to the field tests described in this article.

References

[1] – David Ward – Building Core Networks and Routers in the 2002 Economy, presented at the 25th NANOG meeting, June 9-11 2002, Richmod Hill, Ontario, Canada, http://www.nanog.org/mtg-0206/ppt/David_Ward.pdf

[2] – J.M. Schopf - Ten Actions When SuperScheduling, Global GRID Forum document GFD.04, http://forge.gridforum.org/projects/ggf-editor/document/GFD-I.4/en/1

[3] – IEEE Std. 802.3ae-2002 Amendment: Media Access Control (MAC) Parameters, Physical Layers and Management for 10Gb/s Operation

[4] – M. Olesen - The use of dark fibres in Forskningsnettet, the 21st NORDUnet Network Conference, August 24-27 2003, Reykjavik, Iceland

[5] – Force10 Networks - OC192c/STM-64c and 10 Gigabit Ethernet WAN PHY, http://www.force10networks.com/products/w_oc192wan1.asp

[6] - René Hatem, Wade Hong, Catalin Meirosu - 10 GE WAN PHY to SONET Interoperability Tests, internal report, September 2003

[7] Val Oliva, Ethernet - The Next Generation WAN Transport Technology, http://www.10gea.org/SP0502NextGenerationWAN_Final.pdf

[8] Richard Brand - 10 Gigabit Ethernet Interconnection with Wide Area Networks, http://www.10gea.org/10GbE%20Interconnection%20with%20WAN_0302.pdf

[9] Sylvain Ravot - Internet2 Land Speed Record: 5.44 Gbps from Geneva to Chicago, http://sravot.home.cern.ch/sravot/Networking/10GbE/LSR.htm

[10] Cisco ONS 15454 Multiservice Transport Platform, http://www.cisco.com/en/US/products/hw/optical/ps2006/ps5320/index.html

[11] IXIA 400T Chassis, http://www.ixiacom.com/products/chassis/IXIAChassis.php

[12] Hewlett-Packard Integrity RX2600 Server, http://www.hp.com/products1/servers/integrity/entry_level/rx2600/

[13] Intel PRO/10Gbe Server Adapter, http://www.intel.com/support/network/adapter/pro10gbe/pro10gbelr/index.htm

[14] Force10 E600 switch, http://www.force10networks.com/products/products1.asp

[15] Net100, http://www.net100.org

[16] Wu-chun Feng et al. - Optimizing 10-Gigabit Ethernet for Networks of Workstations, Clusters, and Grids: A Case Study, Supercomm 2003, Phoenix, Arizona, November 15 – 21, 2003; http://www.sc-conference.org/sc2003/paperpdfs/pap293.pdf

[17] Tom Kelly – Scalable TCP: Improving Performance on Highspeed Wide Area Networks, PFLDNET 2003, CERN, Geneva, February 3 – 4, 2003,
http://datatag.web.cern.ch/datatag/pfldnet2003/papers/kelly.pdf

[18] Eitan Altman, Konstantin Avrachenkov, Chadi Barakat – TCP Network Calculus: The case of large delay-bandwidth product, IEEE INFOCON 2002

[19] Antony Antony, Johan Blom, Cees de Laat, Jason Lee, Wim Sjouw – "Microscopic examination of TCP flows over transatlantic links, iGrid2002 Future Generations Computer Systems, Volume 19, Issue 6, August 2003, pp 1017-1029

[20] Sally Floyd – HighSpeed TCP for Large Congestion Windows, Internet Draft <draft-floyd-tcp-highspeed-02.txt>, February 2003, Work in progress

[21] Modifications to Iperf –
http://www.science.uva.nl/research/air/projects/datatag/task3.1/tools/test/#Iperf-Mods-Sec

[22] UDT, The UDP-based Data Transfer Protocol – http://www.dataspaceweb.net

[23] Yunghong Gu, Robert L. Grossman – End-to-End Congestion Control for High Performance Data Transfer, submitted to IEEE/ACM Transactions on Networking, August. 2003

[24] The XENPAK Multisource Agreement website – http://www.xenpak.org