

Windows2000 Delivers 1.6 Gigabit per Second Desktop-To-Desktop Over a WAN

Introduction

The Next Generation Internet (NGI) initiative aims to increase end-to-end bandwidth 1,000 fold, to enable whole new classes of applications. Several advanced technology groups¹ have been cooperating to demonstrate that end-to-end gigabit networking has arrived.

The first public demonstration of this work was the record setting real-time transmission of a super high quality High Definition TV stream over Internet2 between Stanford University and the University of Washington in September 1999. Using custom high performance software written at the UW, a WindowsNT workstation on the Stanford campus packetized a studio-quality HDTV feed and sent it to a WindowsNT workstation at the University of Washington in Seattle where it was fed into Sony HDCAM decoders for display to a live audience in Meany Theater on a state of the art Sony HDTV projector. That unprecedented demonstration delivered more than 200 Mbps (mega-bits per second) of a number of hours worth of nearly flawless, film quality HDTV content. In November the UW, ResearchTV, and Sony extended that demonstration to deliver five HDTV streams from five WindowsNT computers in Seattle to five machines at SC99 in Portland. That demonstrated an aggregate bandwidth for the application of over 1 Gbps (gigabits per second) over a 300 km distance.

NCSA, UW, NTON, NLANR, Nortel, Qwest, and Microsoft have been cooperating to demonstrate gigabit-per-second desktop-to-desktop transfers -- first on a LAN (local area network), then on a MAN (metropolitan area network) and then on a WAN (wide area network). This note outlines progress so far and sketches the next steps. It shows that Windows2000 is gigabit ready -- end-users can get up to 1.3 Gbps between two workstation-class machines 30 kilometers apart. This is the sustained data rate, peak rates are substantially higher. The next step is to demonstrate this for a transcontinental network.

Local Area Network (LAN) Results.

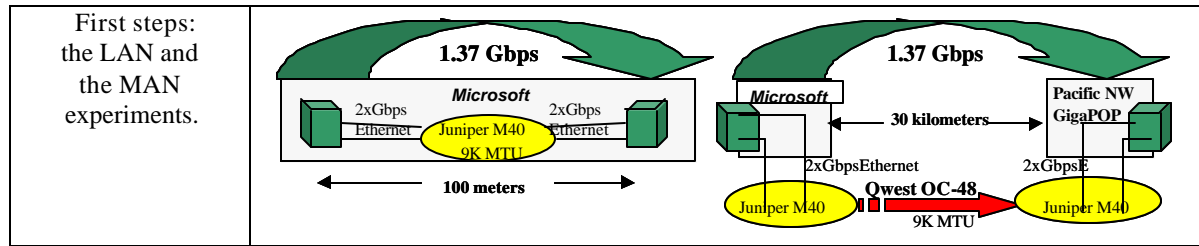
The first step in the progression was to send data between computers in a local area network. We used the standard `ttcp` program, modified to work with Windows2000 multi-threading and asynchronous I/O facilities. The sending program has 12-threads each of which drives one of 12 streams. The receiving program consumes the 12 streams with 12 threads, and continuously reports the observed data arrival rate. Threads in both programs issue up to 2 asynchronous requests per socket. In a LAN configuration Microsoft obtained **1.37 Gbps** between two desktops connected via dual 1 Gbps Ethernets and an Alteon 180 Gigabit Ethernet switch. At this rate the sender is 35% utilized and the receiver is cpu saturated. A single 1Gbps Ethernet delivered **930 Mbps**. These are user data rates; the remaining bandwidth is used by packet headers and by Ethernet overhead. NCSA replicated these results in their testing at Urbana.

In another set of experiments, an 6-way PIII 500MHz Xeon sender with 6 Alteon Ethernet cards delivered **3.92 Gbps** to eight dual PII Xeon 450Mhz receivers via an Alteon 180 switch. One application of this LAN technology is that a Windows2000 DELL system delivered over 25,303 SpecWeb96 operations per second and over 3.0 Gbps in a recent benchmark (<http://www.spec.org/osg/web96/results/res99q4/web96-19991004-03339.html>).

Metropolitan Area Network (MAN) Results.

The next step was to extend these results off the Microsoft/NCSA campus. Working with Qwest, the Pacific Northwest Gigapop, Juniper Networks, and U. Washington an OC48 link was configured from the Microsoft Research lab to the PNW GigaPOP. This created a 30-kilometer path. Data was sent 3 hops: desktop-to-local router, then to the GigaPOP router, then and then to a second desktop. The data rate averaged **1.37Gbps**, with peaks of **1.65 Gbps**. We believe this set a new speed record for a MAN TCP/IP desktop transmission.

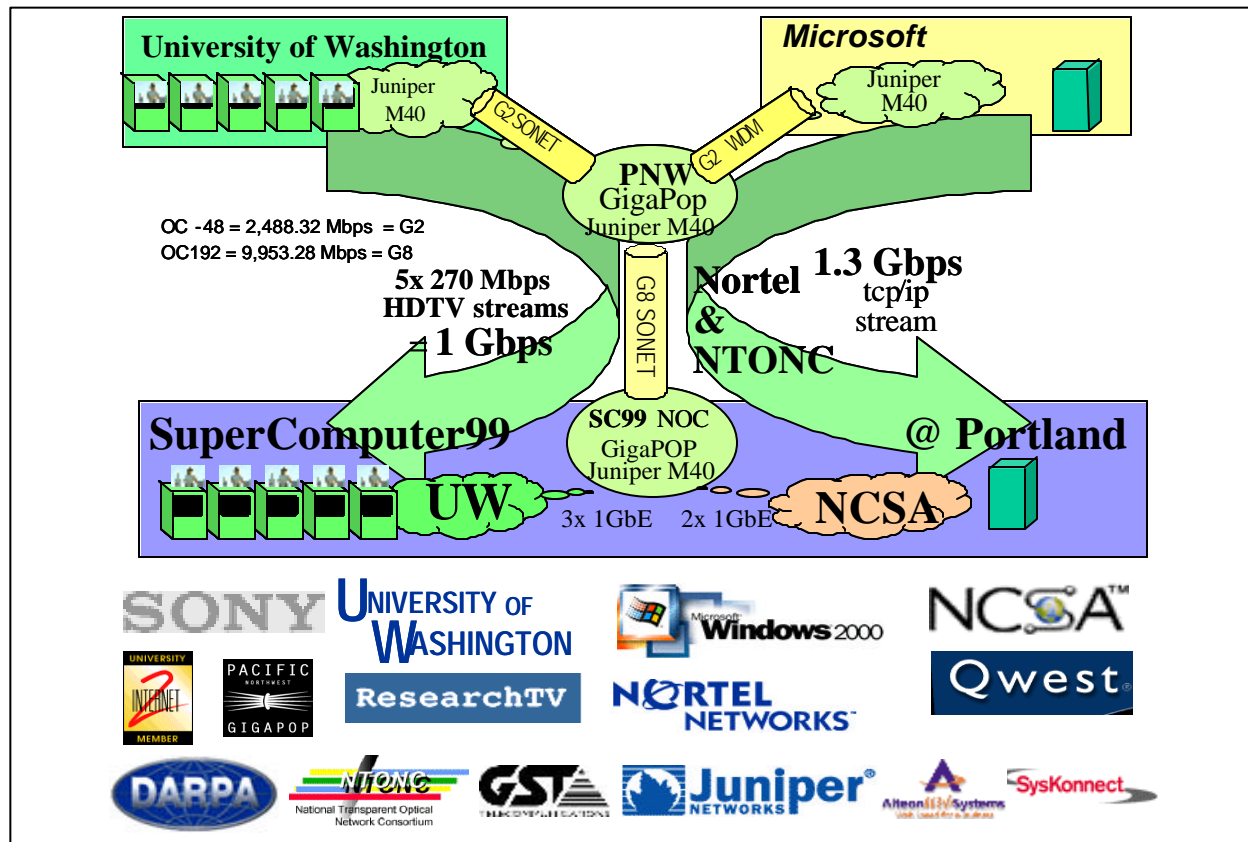
¹ The National Center for Supercomputing Applications (NCSA), the University of Washington (UW), ResearchTV, the National Transparent Optical Network Consortium (NTONC), the Defense Advanced Research Projects Agency (DARPA), Nortel Networks, GST Telecommunications (GST), Qwest Communications, National Laboratory for Applied Network Research (NLANR), Juniper Networks, Alteon Web Systems, Sony, and Microsoft.



Wide Area Network (WAN) Results.

Nortel and NTONC provisioned an OC48 link from the Pacific Northwest Gigapop to the Portland Oregon GigaPOP, and then to a Juniper Networks switch at SC99 show floor. This created a 300-kilometer 4-hop path from Microsoft in Redmond Washington to the Window2000 desktop on the show floor. The data rate averaged **1.3 Gbps**. Again, these are the sustained rates; peak rates are **1.6 Gbps**.

Concurrently, the University of Washington sent five HDTV streams from 5 workstations on the Seattle campus across the same backbone via TCP/IP to five WindowsNT systems on the SC99 show floor. So, the NTONC backbone was carrying in excess of 2.5 Gbps of TCP/IP traffic.



The Next Steps: Transcontinental Gigabit Networking.

NCSA and NLANR, working with NTONC, DARPA, GST, Qwest, and Nortel plan to configure a cross-country multi-gigabit research network. These demonstrations will be among the test applications for this demonstration. In addition, these research organizations are working on the problems of error handling, congestion control, quality of service, and system management for these ultra-speed networks.

Technical Details

The Workstation Hardware

- ?? DELL 6300 class workstations. (cost without Gigabit adapters: \$6,000)
- ?? Sender has two 400 MHz PII Xeon processors with 1Meg L2 Cache, 1GB ram.
Two Alteon ACEnic™ 1 Gbps Ethernet adapters on a single 64-bit 33 MHz PCI bus
- ?? Receiver has two 500 MHz PII Xeon processors with 2 MB L2 cache , 2 GB memory subsystem.
Two SysKconnect SX 1 Gbps adapters on the receiver's system 64-bit 33 MHz PCI bus.

The Software configuration

- ?? Windows2000 build 2174.
- ?? TCP/IP Window size set to 256KB in Registry on both sender and receiver systems.
- ?? Adapters are interrupt bound to each processor respectively, hence no interrupt sharing.
- ?? Socket receive buffer set to 2MB to accommodate in-bound traffic from "fast" sender.
- ?? Maximized interrupt moderation on NICs
- ?? Message Transfer Unit (MTU) was set to 9KB ("jumbo frames") on the NICs.

NT-TTCP was used to send and receive the data. Twelve threads and twelve streams were configured on each end. Windows2000 distributes the traffic across the two NICs. NT-TTCP reports the aggregate sustained user-level data transfer rate of these twelve streams.

Network Configuration

- ?? Microsoft, Redmond: Juniper 4000 switch with 4 x 1Gbps + 1xOC48 NICs
- ?? Qwest OC48 WEDM link from Redmond to Seattle- GigaPOP.
- ?? Seattle- GigaPOP: Juniper 4000 OC48
- ?? Nortel Networks (NTON) OC48 SONET link on OC192
thru Portland GigaPOP
to NOC on SC99 show floor.

- ??
- ?? SC99, Portland Juniper 4000 switch with 7 x 1Gbps + 1xOC48 NICs
- ?? Message Transfer Unit (MTU) was set to 9KB ("jumbo frames").
- ?? The white paper at
http://www.microsoft.com/Windows/server/technical/networking/tcpip_implem.asp
describes the many new features tcp/ip features in Windows2000.
- ?? A description of the U.W. and Research TV demo is at:
<http://www.washington.edu/hdtv/sc99>

Traffic description is at <http://staff.washington.edu/corbato/i2/pnw-gp/sc99/>