PROCEEDINGS OF SCIENCE

Trans-Atlantic UDP and TCP network tests

Anthony Rushton*, **Paul Burgess, Richard Hughes-Jones, Stephen Kershaw, Ralph Spencer and Matthew Strong**
The University of Manchester
Jodrell Bank Observatory
Macclesfield
Cheshire SK11 9DL
UK
E-mail: Anthony.Rushton@postgrad.manchester.ac.uk

A VLBI trans-Atlantic connection would greatly extend the resolution of capabilities of eVLBI. So far igrid 2005 and SC 2005 saw the first UKlight connection to the US via Chicago. We report on UDP and TCP network tests performed between Jodrell Bank Observatory, UK, and Haystack Observatory, USA, utilising the UKLight dedicated lightpath, provided by the ESLEA project, to the Starlight connecting node in Chicago. We show near linerate instantaneous UDP throughput over this lightpath, and IPerf TCP bandwidths in excess of 900 Mbps.

Lighting the Blue Touchpaper for UK e-Science - Closing Conference of ESLEA Project
March 26-28 2007
The George Hotel, Edinburgh, UK

*Speaker.

PoS(ESLEA)007

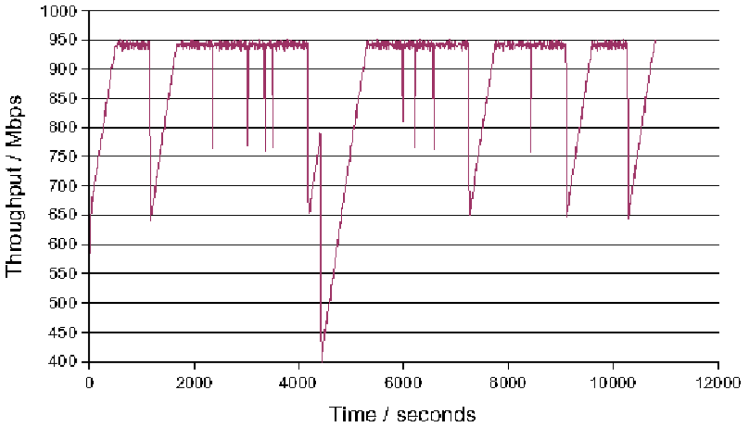

Figure 1: A TCP network bandwidth test between Manchester Computing and a server located in Chicago. The network application tool, Iperf was used to measure the maximum throughput for a period of over 3 hours.

1. Testing the trans-Atlantic link

Very long baseline interferometry (VLBI) generates large rates of data from many telescopes simultaneously observing a source in the sky. This can require the information to traverse intercontinental distances from each telescope to a correlator in order to systhesise high resolution images. With the dramatic development of the Internet and high bandwidth networks, it is becoming possible to transmit the data over large area networks. This allows the correlation of radio astronomical data to be done in real-time, whereas this process would take weeks using conventional disk based recording.

Jodrell Bank Observatory has two 1 Gbps dark fibres from the MERLIN telescope array to the University of Manchester's campus. This local network connects to UKLight at Manchester computing [1] via a Cisco 7600 switch. UKLight is a network of dedicated optical light paths, provided by UKERNA [2]. A guaranteed 1 Gbps bandwidth connects between UKLight and StarLight [3] to a server located in Chicago.

1.1 TCP bandwidth tests with Iperf

TCP throughput rates were measured with the software package, Iperf [4]. Fig. 1 shows the results of a test lasting 3.3 hours, with multiple packet losses observed throughout the test. Despite the network utilising a dedicated lightpath, packet losses are observed and as a result TCP goes into a congestion avoidance phase and reduces the transmitted bandwidth.

1.2 UDP network tests

In order to better understand the network behavior, the UDP protocol was used. Unlike TCP, UDP is not consider a 'fair' protocol and therefore is not widely used on production networks. If a packet is lost in the network, the transmission rate is not reduced, nor is the packet resent. This makes UDP an excellent diagnostic tool for troubleshooting dedicated network paths.

The network analysis software package, UDPmon [5], was used to investigate the link between Manchester and Chicago by transmitting packets at a constant rate, and reporting packet loss; see

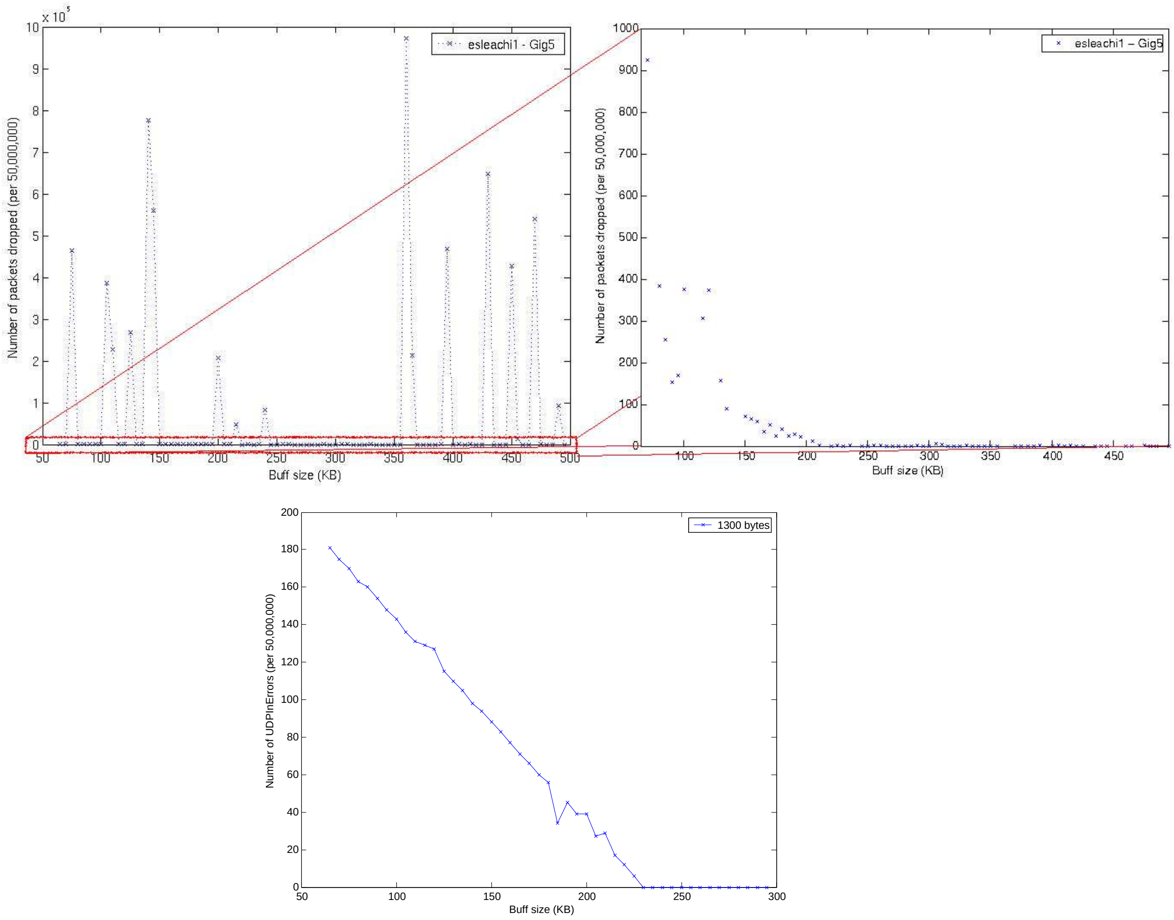

Figure 2: (top left) a) The number of packets lost in a UDPmon test at 940 Mbps from Chicago to Manchester Computing as a function of the receive buffer size. Large intermittent packet losses of >1 % were observed. (top right) b) The same UDPmon data with the packet loss axis limited to 1000 counts. At low receive buffer size constant packet loss is seen. (bottom) c) The number of packets lost between two computers directly connected between network interface cards (i.e. no network). At low receive buffer sizes there is a linear relationship between packets lost per test and buffer size.

results in Fig. 2. In order to emulate the transmission of eVLBI science data, long periods of time were measured, transferring 50 million packets every test. The receive buffer of the application was varied to observe the effects of packet loss.

The application reported large intermittent packet losses, as seen in Fig. 2a. Increasing the receive buffer size at the application layer did not stop this large intermittent packet loss of > 1 %. However varying the buffer does have a small constant effect on the packet loss. Fig. 2b shows during the large transmission of packets, if the receive buffer size in the application is too small then packets will also be lost.

We therefore tested the UDP performance by linking servers directly in a back-to-back test (i.e. without the network). The application transmitted UDP at line rate (940 Mbps). The buffer size of the receive host was varied from 65 KB to 300 KB. Packet loss was observed when the receive buffer was less than $\sim$ 230 KB.

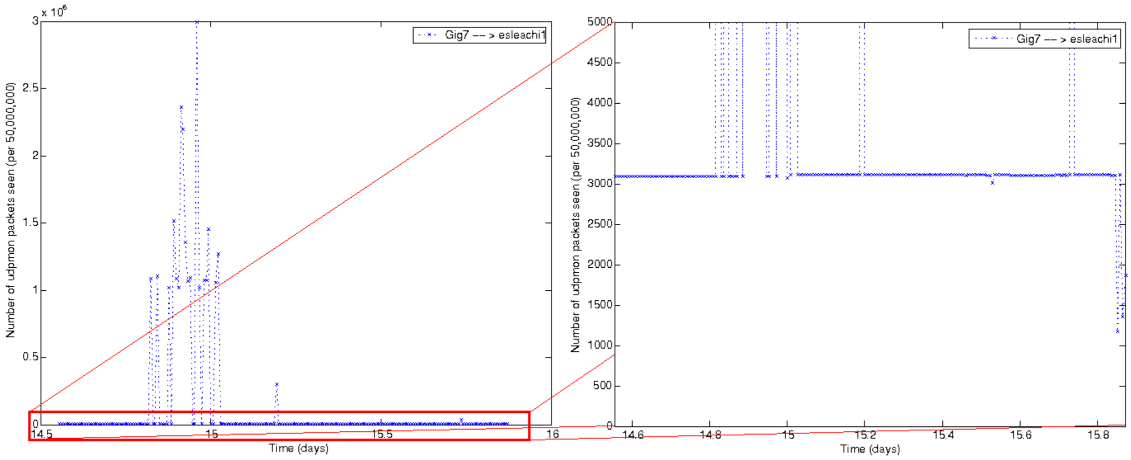

Figure 3: (left) a) The number of packets lost in a UDPmon test at 940 Mbps from Manchester Computing to Chicago as a function of time. Once again large intermittent packet losses of >1 % were observed. (right) b) The same UDPmon data with the packet loss axis limited to 5000 counts. A constant loss of at least 3000 packets per test is observed.

1.3 Constant packet loss when running at line rate

The network tests with UDPmon was repeated in the opposite direction (from Manchester to Chicago) in Fig. 3a. The receive buffer was set to 300 KB and the test was performed for a longer period of time ($\sim$ 30 hours). Once again the application was losing a large fraction of packets (>1%). However this time, as seen Fig. 3b. a constant loss of at least 3000 packets per 50,000,000 sent (0.001%) occurred in every test.

2. Network isolation

In order to have characterised this network link, it was important to examine where the data packets were dropped at the lowest point of the OSI model, i.e. layer 2. To do this we examined the SNMP (simple network management protocol) counters of the network inferace cards and the Cisco 7600 switch connecting to UKLight. The results showed the constant packet loss observed in section 1.3 were within the Cisco 7600. All 50,000,000 packets were received by the switch throughout each test. However if the transmission rate was reduced to 800 Mbps, the switch could transmit all 50,000,000 packets without loss.

We tested the switch's inferace from Jodrell Bank to the Cisco 7600 using a different connection to the computer in the University of Manchester campus. Both the switch's SNMP counts and UDPmon's reports showed the switch transmitted every packet at line rate without packet loss in this configuration.

2.1 UKERNA's maintenance

We concluded through process of elimination that the large intermittent packet loss of > 1% was therefore within UKLight. After private communication with UKERNA it is believed the network issues were due to a broken fiber on the Manchester Computing to StarLight connection. After multiple maintenance tickets were issued by UKLight, we repeated the UDPmon tests.

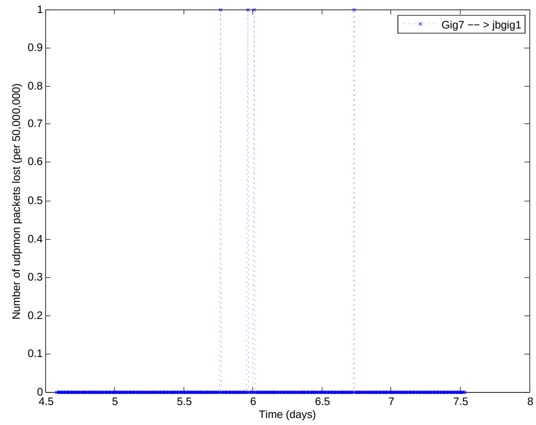

Figure 4: UDPmon tests from Jodrell Bank to Chicago sending UDP packets at 940 Mbps. Over 200 TB of data was transfered over a period of over 60 hours. Only 4 packets were lost throughout the experiment.

The configuration of the local network from Jodrell Bank into UKLight did not give the constant packet loss seen in section 1.3 even at line rate. The results in Fig. 4 show that when running at line rate (940 Mbps), very few packets were lost over (the observed) period of 2.5 days. Over 20 Billion packets were transmitted ($\sim$ 200 TB) with the loss of only four packets in the network.

3. Conclusion

This work demonstrates some of the challenges encounted when using high bandwidth networks. Software application tools have simulated the date rates required by eVLBI science by continually sending large numbers of packets for many hours. This has shown the need for the receive buffers of the applications to be capable enough to collect data at these rates for long periods of time.

Issues have arisen with the Cisco 7600 switch showing, that under certain circumstances the instrument does not perform to the manufacturers specifications. This highlights the requirement to identify and test equipment to maximum abilities. Problems with the link were isolated by eliminating the client, end-host servers and local network by inspecting level 2 SNMP packet counts. This led us to confidently conclude that large packet losses were within UKERNA's UKLight network. After maintenance on UKLight, our tests were repeated for a large time period ($\sim$ 3 days). This successfully showed it was possible to transmit packets between Manchester Computing and Chicago at 940 Mbps without losing a significant number of packets.

References

[1] Manchester Computing, http://www.mc.manchester.ac.uk

[2] United Kingdom Education and Research Networking Association, http://www.ukerna.ac.uk

[3] StarLight, http://www.startap.net/starlight

[4] Iperf home page, http://dast.nlanr.net/projects/Iperf

[5] UDPmon home page, http://www.hep.man.ac.uk/u/rich/net