PtsCase14

Below expected throughput, Budapest to New York

PTS Case 14

Synopsis

DANTE contacted the PERT to report they were experiencing below expected TCP throughput between the GN2 Measurement Point (MP) in Budapest and the MP in New York. Because the MPs are co-located with the GN2 routers in the GN2 PoPs, then on the non-lossy GN2 backbone TCP rates of 900Mbps should be consistently achievable, even over long distances. However, the TCP iperf tests run consistently achieved only around 600Mbps.

This ticket was open for a long time, and during its life there were several issues (across several MPs) that were identified and either accepted or resolved. Particularly poor transfer rates (5-20Mbps) were the result of duplex mismatches, and some improvement cold also be got by setting the TCP re-ordering value to '20' instead of the default '3' (to counter the M160 habit of packet re-ordering). Advice from researchers at North Carolina State University (NCSU) was to change the ssthresh_initial default value (so that there is no limit on the exponential growth of traffic in the absence of congestion) - as expected this improved the initial growth rate of TCP on long connections. However the biggest issue was determined to be losses on the POP LANs. Sometimes this was because packets were dropped by the server NICs (shown by the increasing error counts on the NICs), but in other situations with poor throughputs there was no recorded packet loss, on either server or switch LAN port. In these cases changing switch LAN ports made a notable difference, sometimes completely rectifying the problem. In summary, it seems that the throughput problems can be localised to the PoP LANs (not the backbones), and that if all else is configured properly then the limitation appears to come from a sub-standard switch (or switch port).

– Main.TobyRodwell - 31 Jan 2007