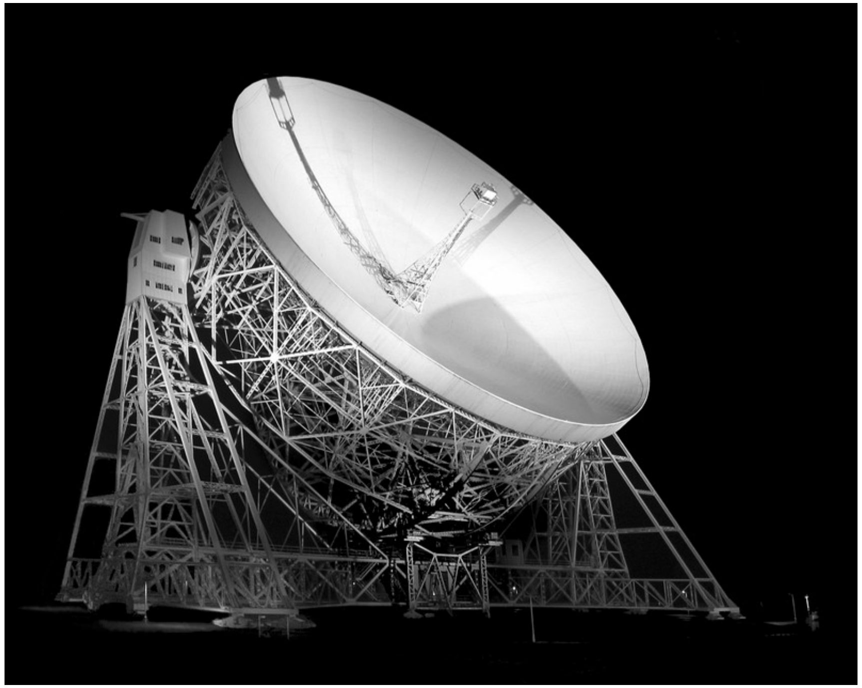


Packet Loss in High Data Rate Internet Data Transfer for eVLBI

Ralph Spencer
Richard Hughes-Jones
Adam Mathews
Stephen O'Toole
The University of Manchester

Why eVLBI?

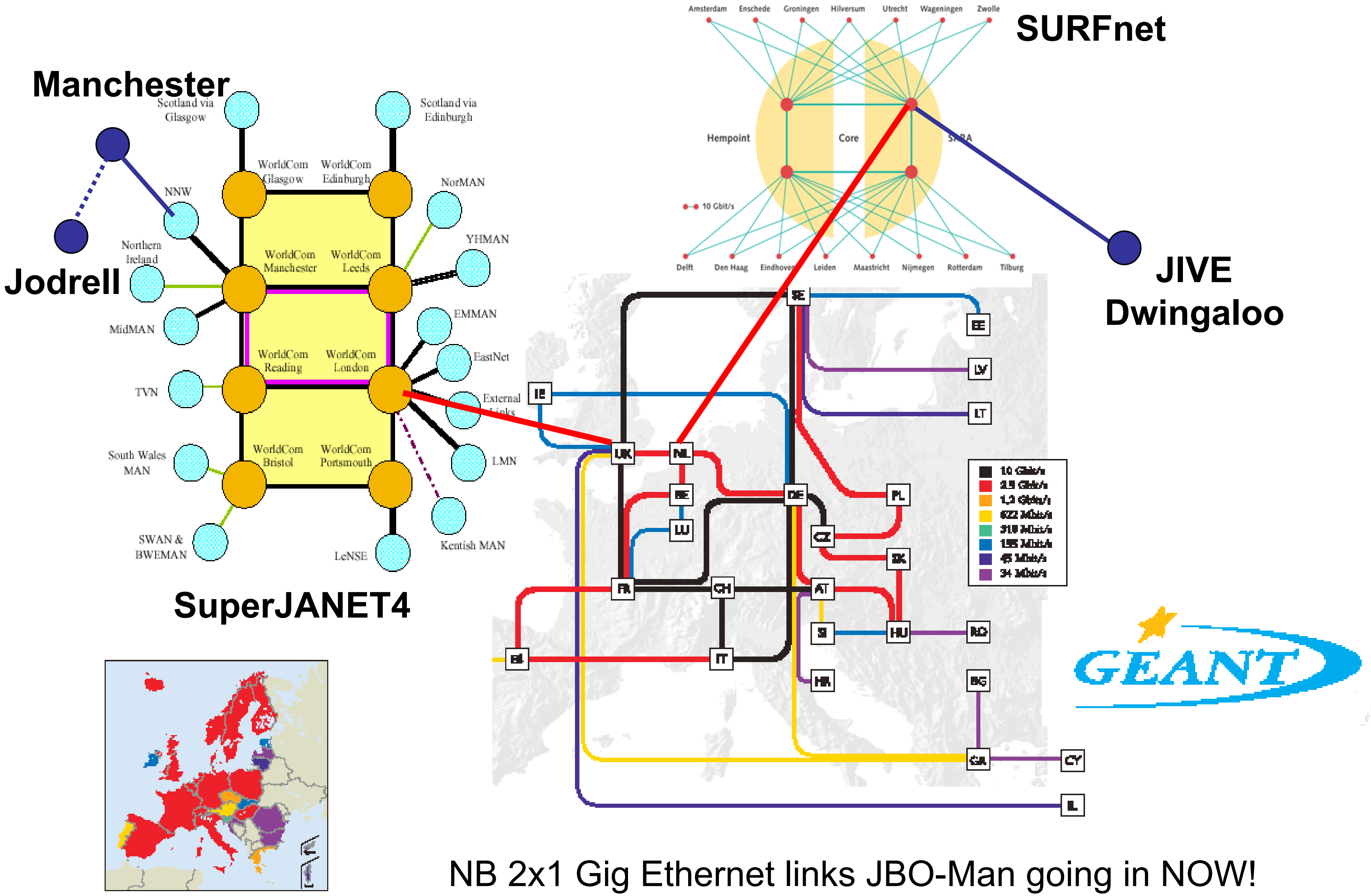
- The bandwidths possible, and hence the sensitivity that can be obtained, will eventually be significantly higher than that of disk-based systems.
- Real-time VLBI will be vastly more reliable than the current system in which data are often not correlated until several weeks after recording.
- Permanent connections via fibre and real-time VLBI will result in a major culture change in VLBI
  - Move away from three sessions a year to regular VLBI intervals
  - Enable EVN to perform monitoring observations of radio sources; a domain which to date has been the preserve of the VLBA in the USA.
  - EVN will be able to respond rapidly to targets of opportunity: supernovae, GRBs, microquasar bursts etc.

eVLBI and Internet Protocols

- IP handles the packet addresses, looked at by routers. TCP or UDP etc. run within IP, but ignored by routers.
- TCP Transmission Control Protocol
  - generally the default in most computer systems, used by ftp, email, www, ssh etc.
  - has in built congestion control, interprets packet loss as congestion
  - reliable data transfer, at expense of data rate
  - various implementations around
- UDP User Datagram Protocol
  - has no congestion control,
  - can lose packets, but runs at high data rates
  - Useful for diagnostic tests on networks (UDPmon)

Which is best for eVLBI?

VLBI Project: Test Topology

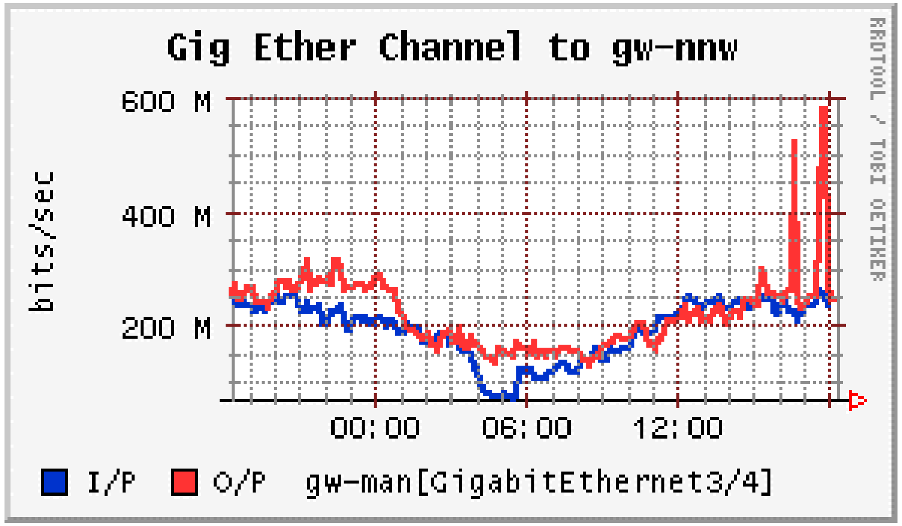


NB 2x1 Gig Ethernet links JBO-Man going in NOW!

Tests on the Network, Manchester-Dwingeloo: investigation of packet loss

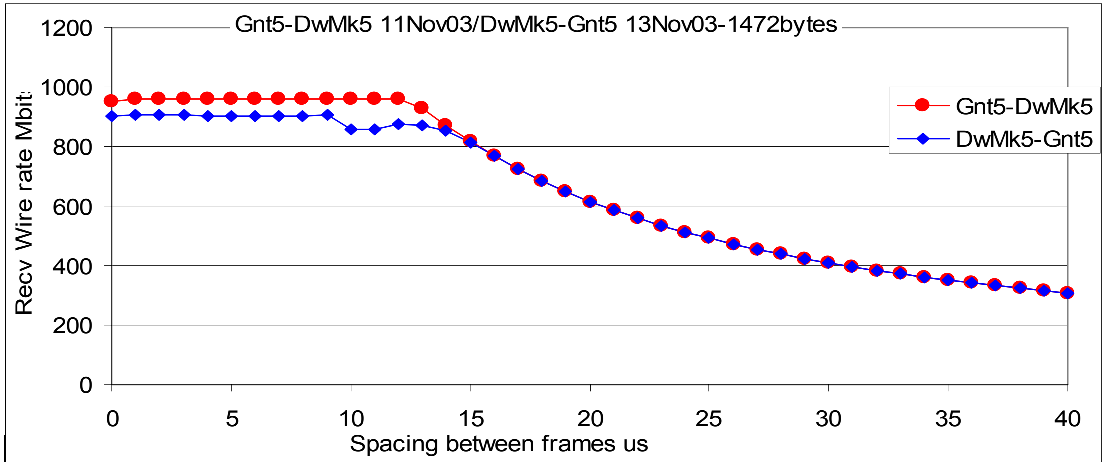
• 4th year MPhys Project at The Univertsity of Manchester Oct-Dec 2003, using campus network and SuperJANET 4 academic network in the UK.
• UDPmon used to test throughput and packet loss
• NB tune up the end machines – see **http://grid.ucl.ac.uk/NFNN.html**

Effect on the local traffic:

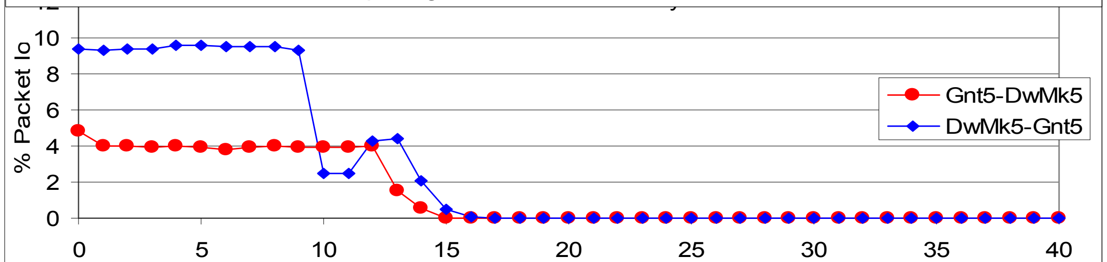


UDP Throughput Manchester-Dwingeloo

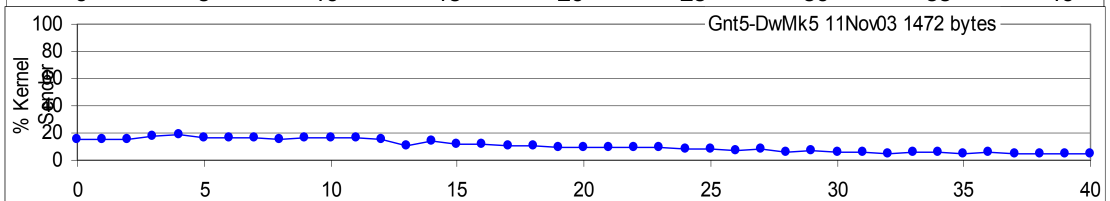
- **Throughput vs packet spacing**
- **Manchester: 2.0G Hz Xeon**
- **Dwingeloo: 1.2 GHz PIII**
- **Near wire rate, 950 Mbps**
- **NB record stands at 6.6 Gbps SLAC-CERN**



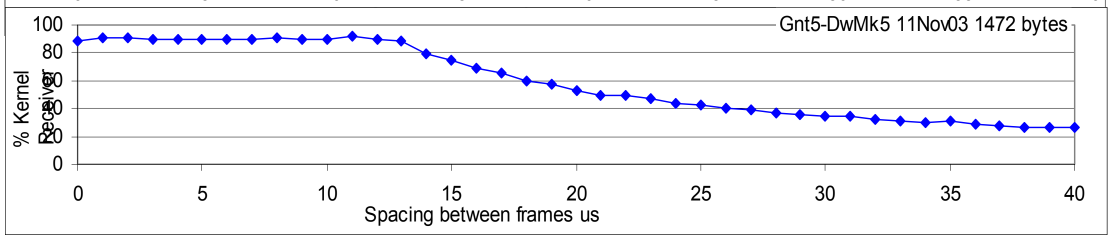
- **Packet loss**



- **CPU Kernel Load sender**

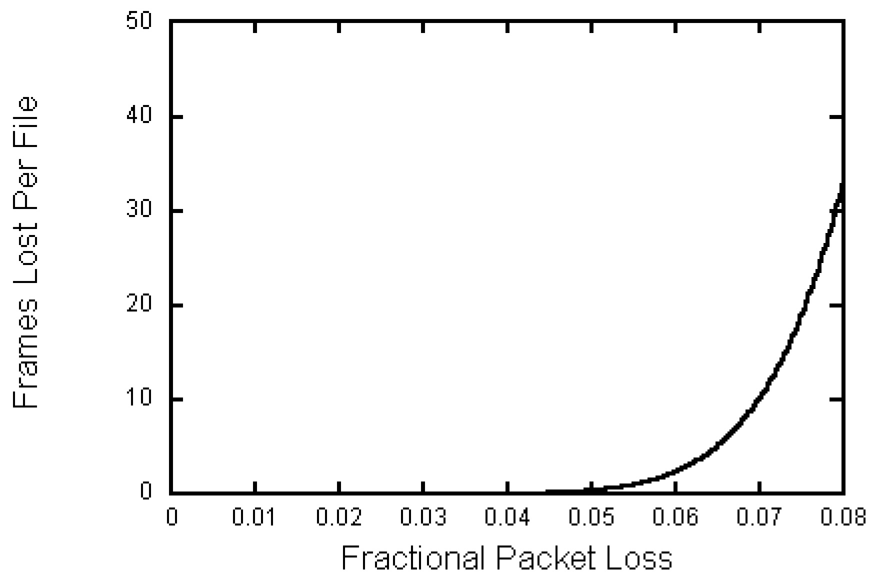


- **CPU Kernel Load receiver**



- **4th Year project**
  - **Adam Mathews**
  - **Steve O'Toole**

Effect of Packet loss on Correlator

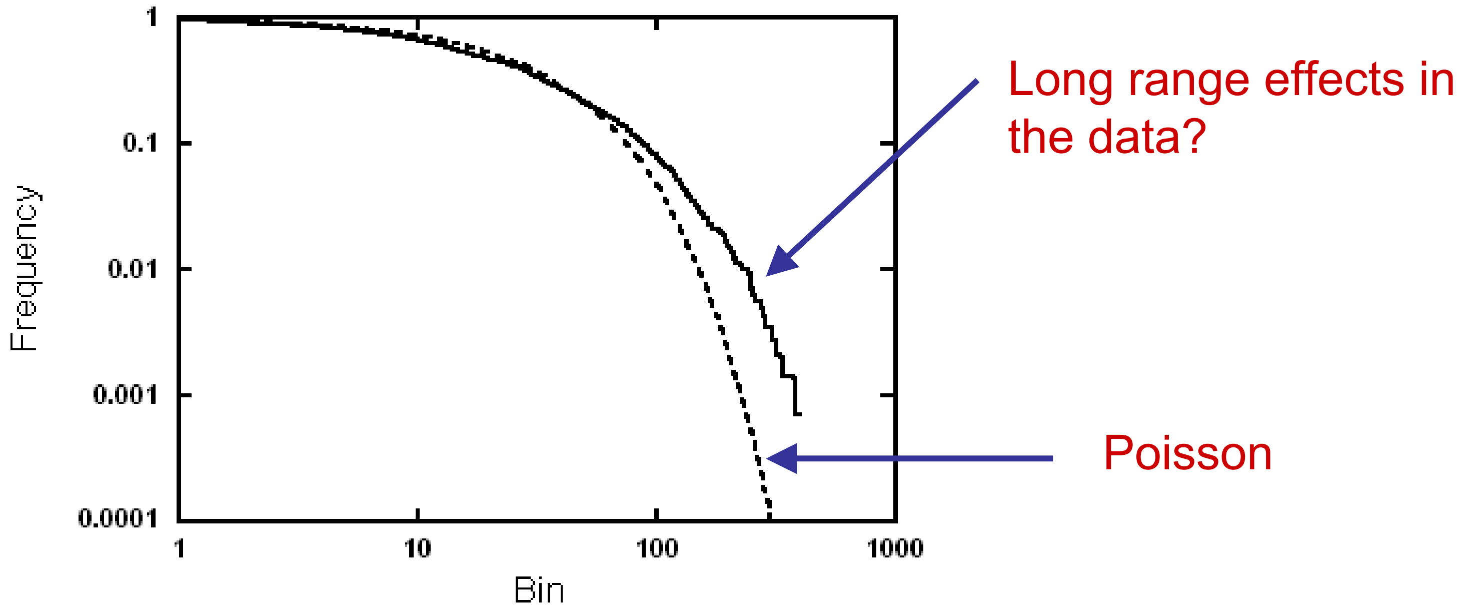


Number of VLBO MkIV Frames lost per 1.8 Gbyte file vs fractional packet loss

Useful rule of thumb – frame sync lost if packet loss > 2%
(cf ALMA/EVLA/eMERLIN spec. 10^{-6})

Packet loss distribution:

Cumulative distribution $\int_{t}^{\infty} p(\tau)d\tau$

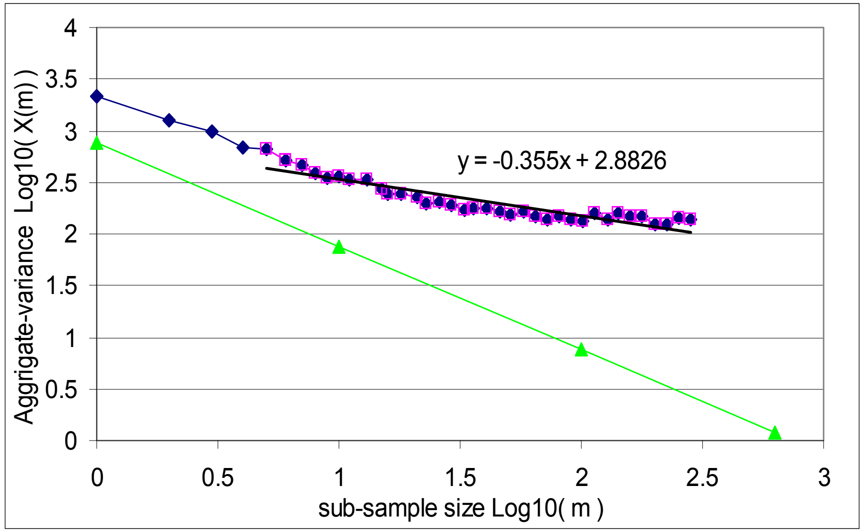


Cumulative distribution of packet loss, each bin is 12 μsec wide

Are there Long Range Effects?

- Aggregated Variance Method
- Divide time series length N into blocks of size m
- Calc mean of each section Xm(k) k= 1 … N/m
- Calc variance VXm of these Xm(k)
- Vary m size of the blocks

- Plot on log-log & fit slope β
- Hurst parameter H
  $\beta = 2H - 2$
- Measure β = -0.355, which gives H=0.822
- H =1 no long range dependence



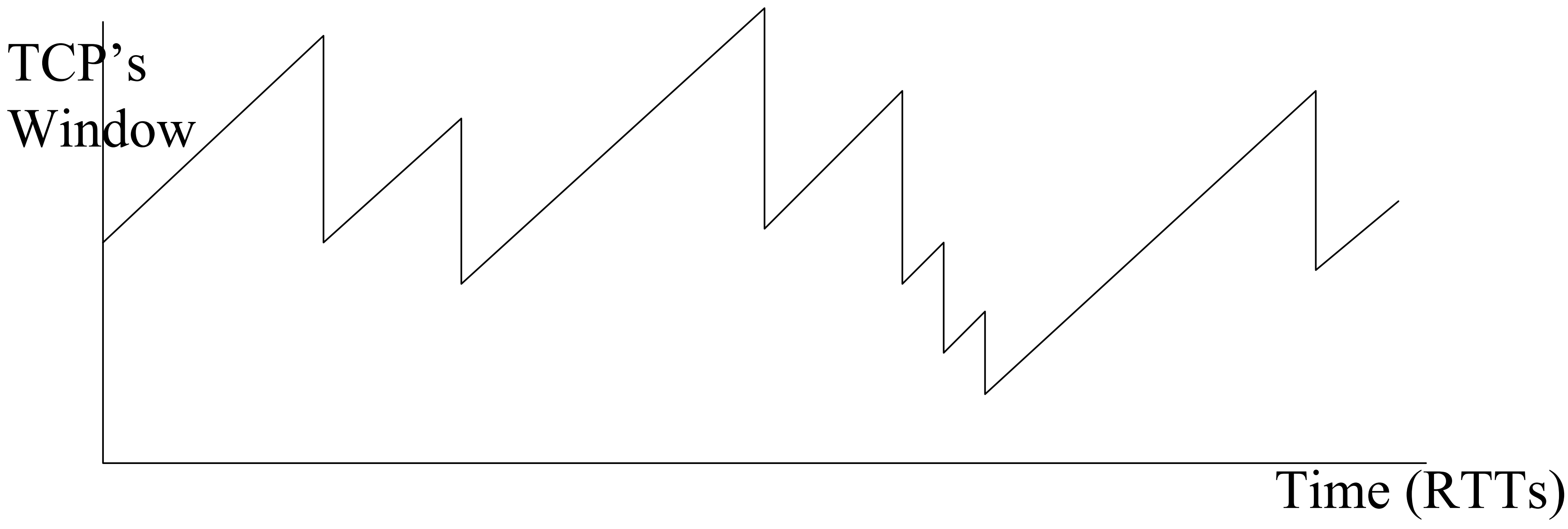
TCP/IP (Thanks to Mark Handley UCL)

- IP handles addressing (and some other stuff).
  - Routers look at the IP headers to move packets from sender A to receiver B.
  - Sometimes the routers will break, or get congested, or re-route your traffic over a piece of wet string, and then they'll drop packets.
- TCP packets are carried in IP packets.
  - The routers don't look at TCP.

TCP Adaptive Congestion Control

Basic behaviour: Additive Increase, Multiplicative Decrease.

- Maintain a *window* of the packets in flight:
  - Each round-trip time, increase that window by one packet.
  - If a packet is lost, halve the window.

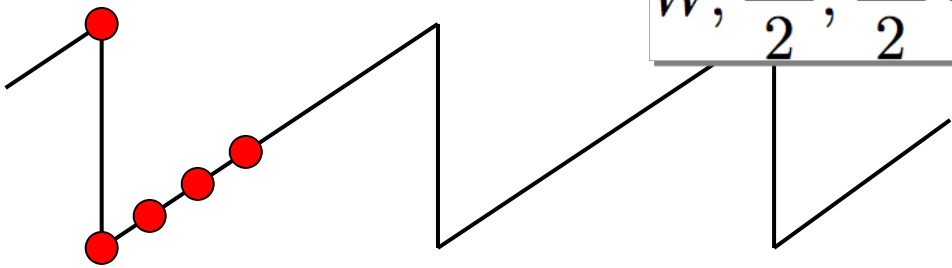


TCP Modelling: The "Steady State" Model

The model: Packet size *B* bytes, round-trip time *R* secs, no queue.

- A packet is dropped each time the window reaches W packets.
- TCP's congestion window:

$$W, \frac{W}{2}, \frac{W}{2}+1, \frac{W}{2}+2, ..., W, \frac{W}{2}, ...$$



- The maximum sending rate in packets per roundtrip time: *W*
- The maximum sending rate in bytes/sec: *W B / R*
- The average sending rate *T:* *T* = (3/4)*W B / R*
- The packet drop rate *p:*

$$p = \frac{1}{\frac{3}{8}W^2}$$

- **The result:**

$$T = \frac{\sqrt{6}B}{2R\sqrt{p}} = \frac{\sqrt{3/2}B}{R\sqrt{p}}$$

An Improved "Steady State" Model

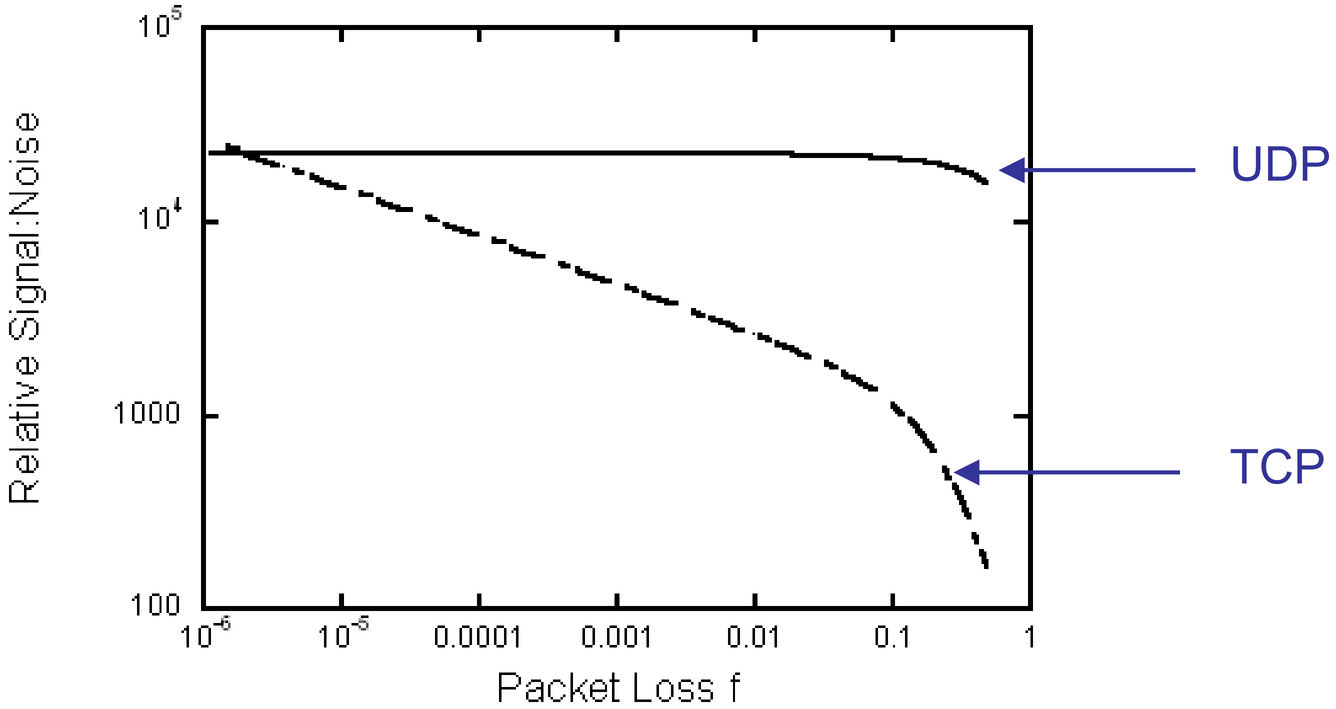
A pretty good improved model of TCP Reno, including timeouts, from Padhye et al, Sigcomm 1998, ACM Tr 2000

$$T = \frac{s}{R\sqrt{\frac{2p}{3}} + t_{RTO}(3\sqrt{\frac{3p}{8}})p(1+32p^2)}$$

T : sending rate in bytes/second]
R : round trip time
p : fraction of packets lost
t_{RTO} : TCP retransmission timeout

Effect on Signal:Noise

Throughput of TCP is fundamentally lower than that of UDP



Signal to noise for TDP and TCP

Conclusion

- UDP can give high data rates over networks in EVN, but could lead to denial of service for other users!
- TCP is reliable, but bandwidth more important for VLBI since some loss of data can be tolerated. Recovery times long for long links (minutes for EVN)
- Newer protocols to be investigated by ESLEA* project using UKLight network (2 postdocs on protocols, 1 postdoc in eVLBI**, 1 PhD stud)
- No fundamental reason why real time data rates >512 Mpbs can't be achieved *now* in EVN.

Get Connected!

* ESLEA - Exploitation of Switched Lightpaths for Escience Applications

** Talk to me if you need a job