Protocols and eVLBI

Jouko Ritakari
Metsähovi Radio Observatory

Background

Metsähovi developed a data acquisition system in year 2002

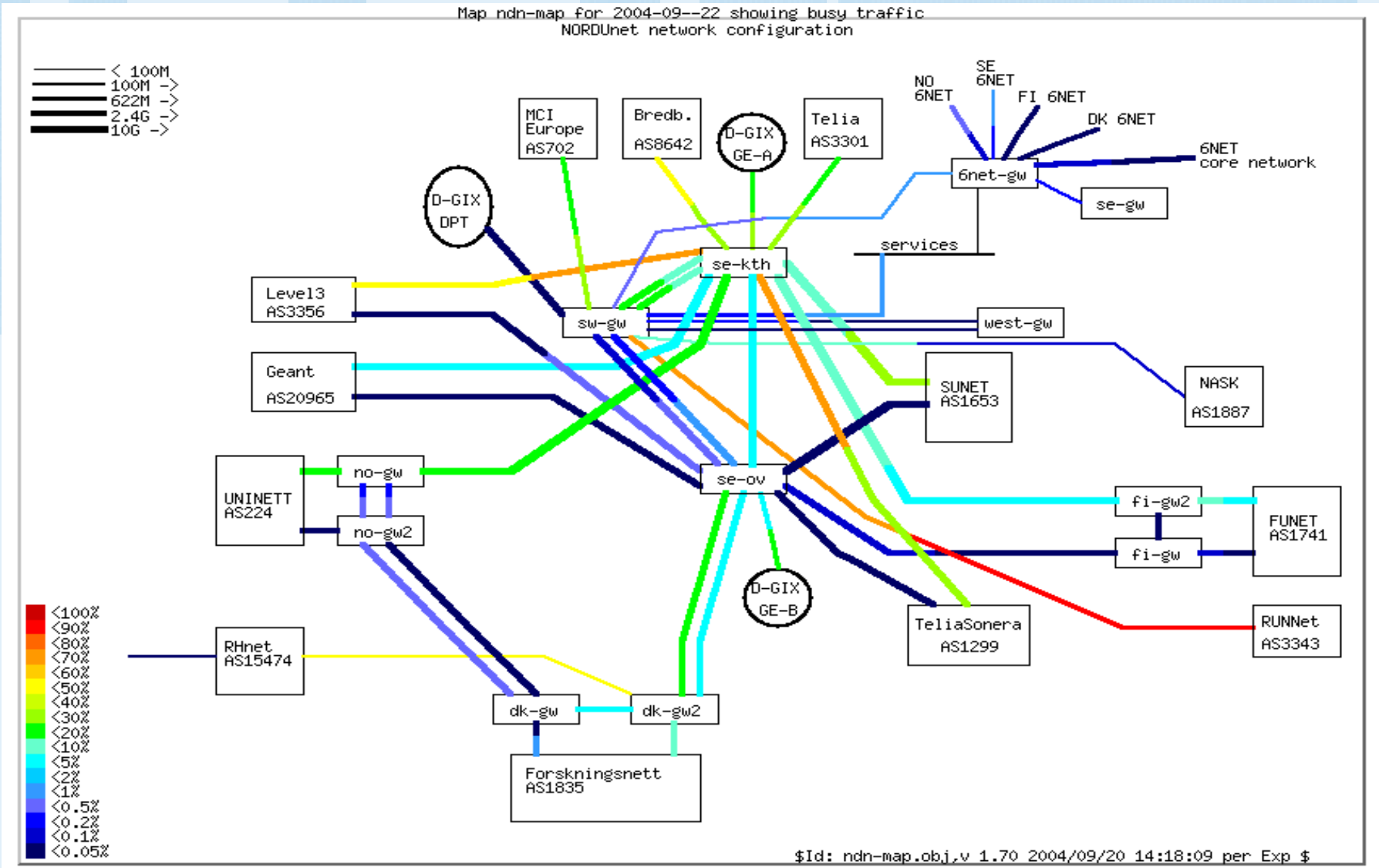
TCP, UDP and RTP all unsuitable

- TCP not suitable for long fat pipes
  - Parameter tuning helps, so do switched lightpaths
- UDP loses packets at receiving end
  - FPGA-based special 10GBE receiver would help
  - The SETI people already have it, called iBOB
- Third solution: New reliable UDP-based protocols
  - UDP with retransmission, why wasn't it invented before?

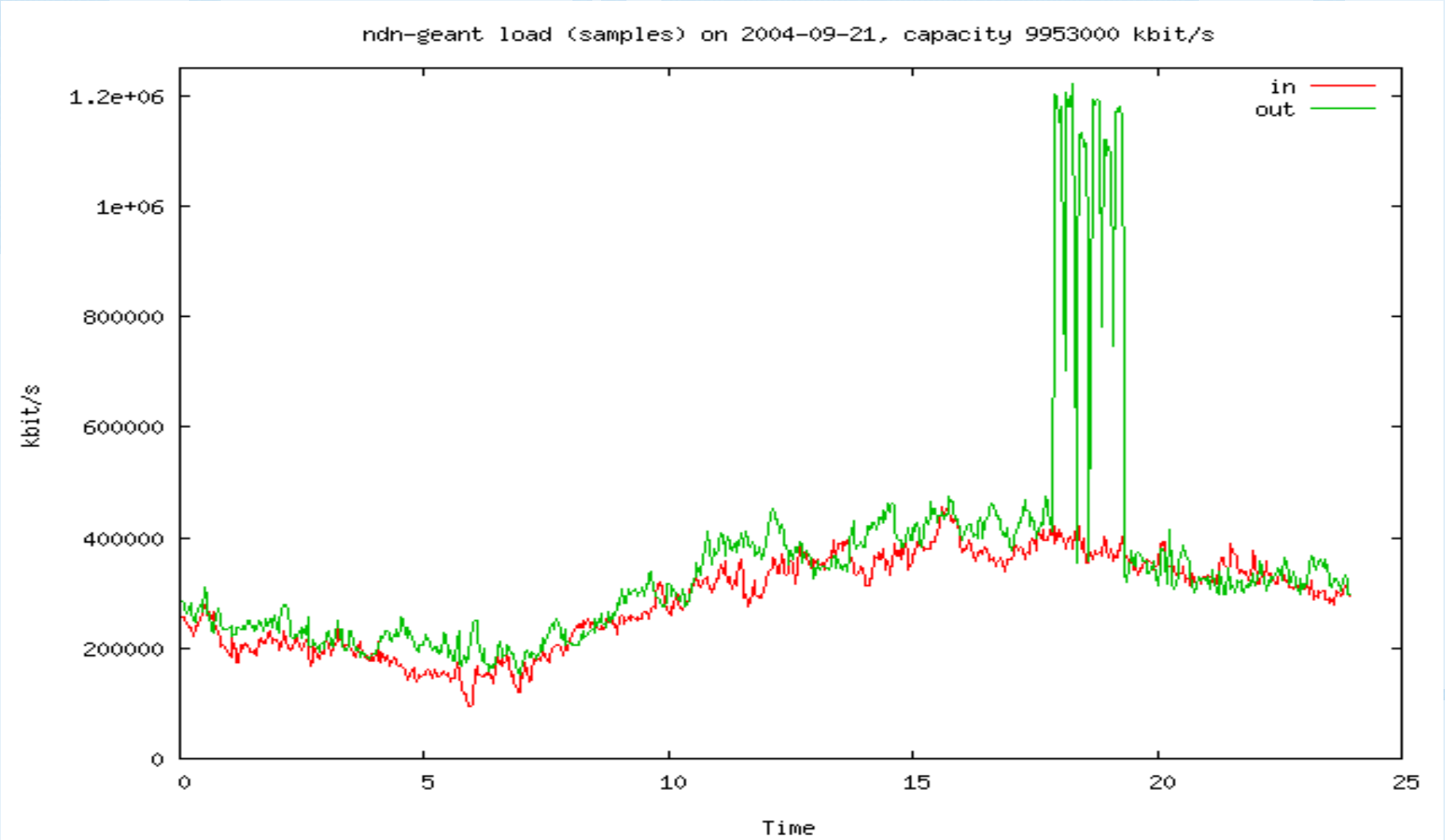
New reliable UDP-based protocols

- Several variants: RBUDP, fobs, UDT/Sabul, Tsunami
  - Concerns about TCP-friendliness
- Tsunami was chosen to eVLBI tests because of clear programming style
- Caused no packet-loss to other traffic
- Only routers know the congestion
  - Explicit Congestion Notification (ECN)

Nordunet load map at the time



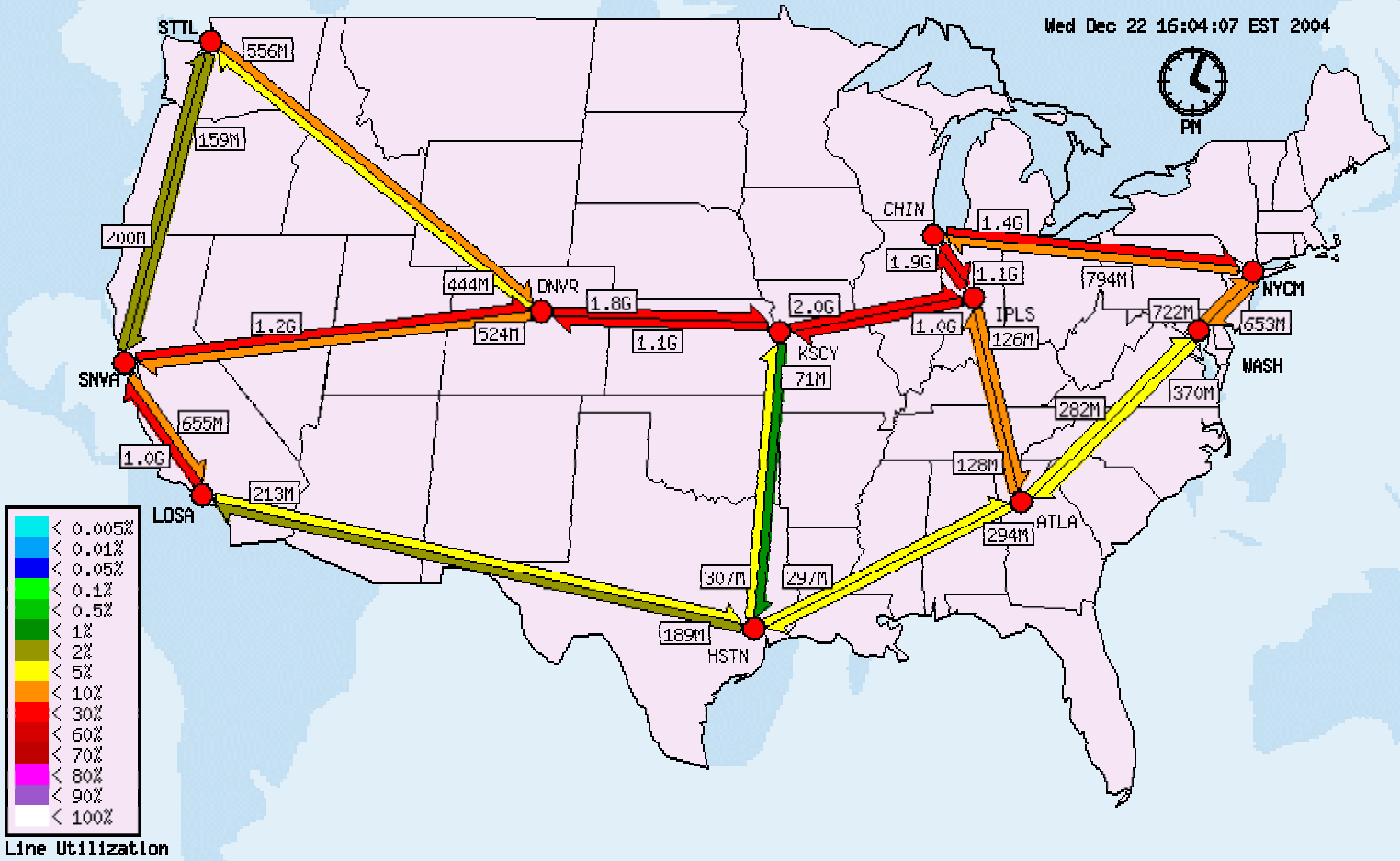
With improved 4-disk raid array we got 640 Mbit/s



Japan-JIVE-Japan tests

- 22th December 2004 at 23:00 Finnish time the JIVE gigabit connection started working
- During the first half an hour 355 Mbit/s speed from Nict, Japan to JIVE, Netherlands was achieved
- On the next day the speed was improved to 400 Mbit/s with simple parameter tuning

Nict-JIVE transfer 2004-12-22



New protocol developments

- New features for the Tsunami protocol
  - Csiro: several bug fixes, wildcard filenames
  - Metsähovi: Realtime 512 Mbit/s version of protocol
  - Metsähovi: eVLBI filename convention
  - Metsähovi: Distributed transfer for software correlation

Realtime version of Tsunami

- Realtime server & client by Jouko Ritakari
- Works with standard tsunami programs
- Works reliably in lab at 512 Mbit/s
- Both server and client can make backup copies on disk “on the fly” at 512 Mbit/s
- Server interpretes the “filename” as start time

Error rate calculations

- Assume 512 Mbit/s speed, 10% packet loss
- Assume 32kB packet => 2000 packets per second
- Assume 10 second = 640MB buffer
- Assume retransmission once per second
- Result: Residual packet loss 10E-1 to the power of 10, that means one packet in 10E10 packets will be lost
- Mean time between packet loss: 10E10/2000/60/60/24/365.25 years = 1.58 years
- For a 20-second buffer the same calculation gives one lost packet every 15.8 Gyears (Age of the Universe, 1/H0, estimated to be 13.7 +/- 0.2 Gyears)

Experiences with reliable udp protocols

- Networks are reliable, virtually no packet loss
- Jumbo frames or dedicated lightpaths not needed
- Modest CPU load, 70% at 640 Mbit/s
- Almost all packet loss caused by receiving computer
- Important that all parts of computer are fast: Network adapter, CPU, disks, PCI bus