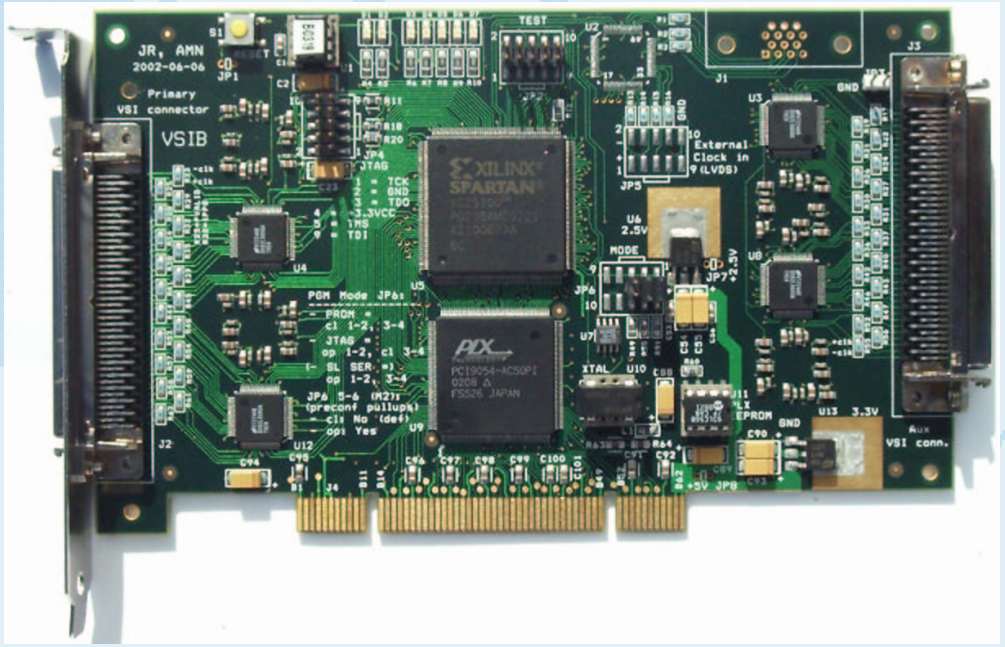
New reliable UDP protocols for eVLBI

Jouko Ritakari, Ari Mujunen
Metsähovi Radio Observatory
Jouko.Ritakari@hut.fi, Ari.Mujunen@hut.fi

Background

Metsähovi developed a data acquisition system in year 2002



Design decisions

- Off-the-shelf technology would be used
- Data stored in normal Unix files
- VSI-standard input port
- Adapters to various systems would be external modules
- Compatible with Mk4, VLBA, S2, K4, ADS1000 and A/D converters

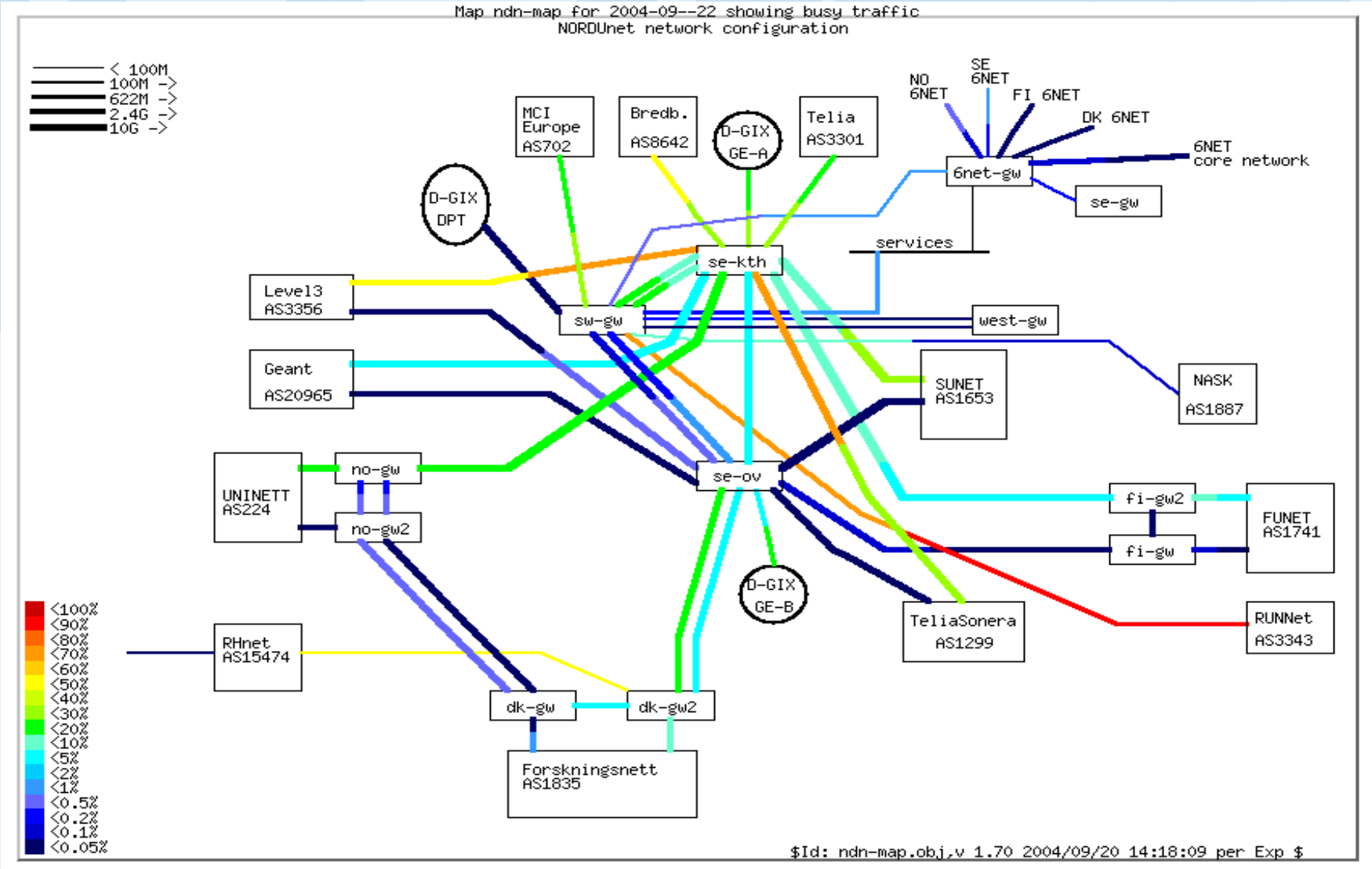
The system keeps on improving

- Third generation of motherboards now in use
- Especially e-VLBI capabilities have improved
- Until June 2004 the PCI bus limited throughput to roughly 450 Mbit/s
- The new nVidia nForce3 250 gb chipset has the first native 1Gbit/s Ethernet controller
- Nforce4 arrived, first tests very promising

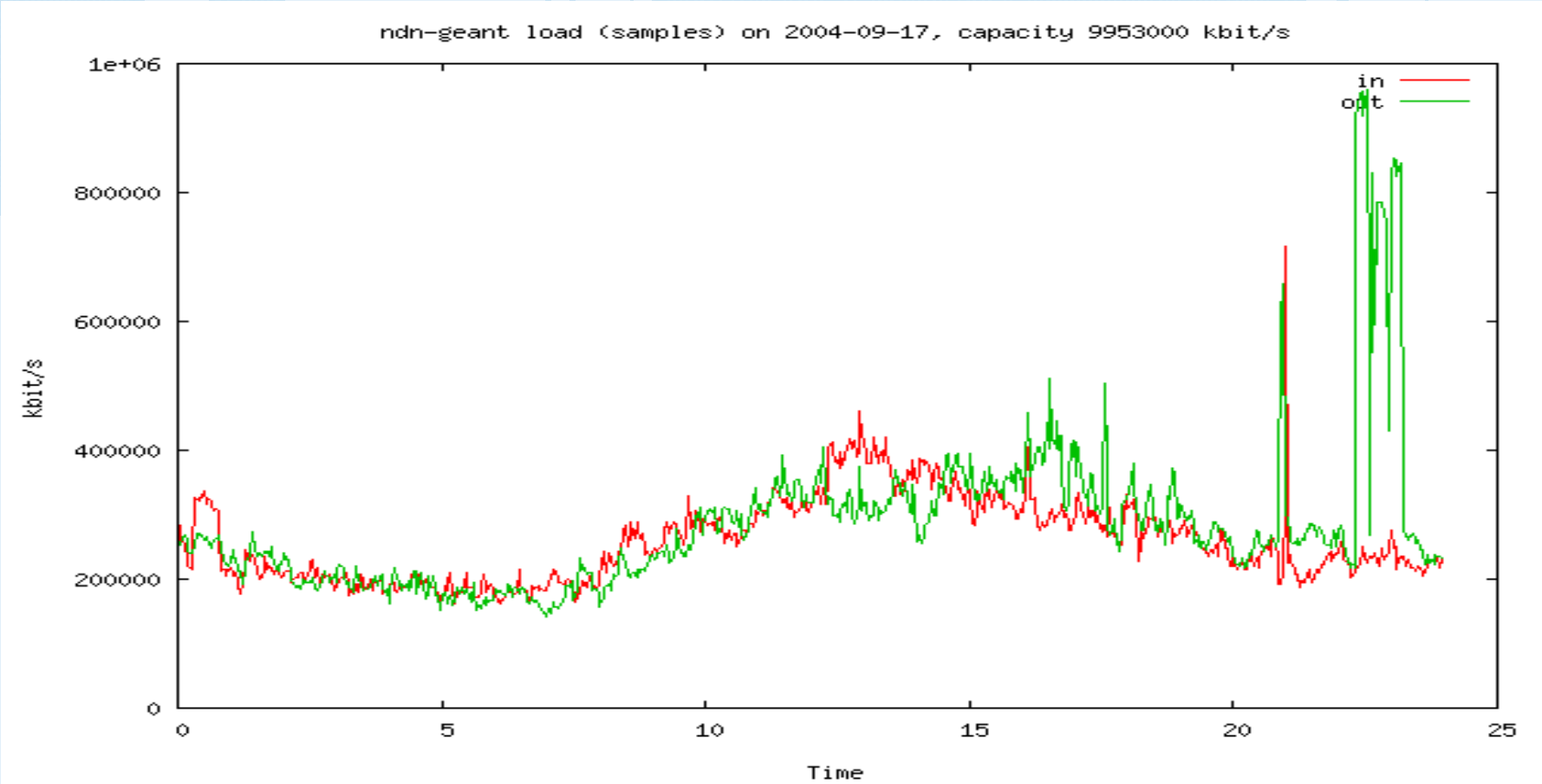
e-VLBI file transfer tests in September 2004

- MRO contacted Funet, the Finnish University Network and got permission to run eVLBI tests
- Ten 58GB VLBI data files were transferred
- Route: Funet – Nordunet – GEANT – SURFnet
- Tests were very easy to set up, we achieved 575 Mbit/s disk-to-disk transfer speed almost immediately
- Used mainly Tsunami, some tests with UDT
- No jumbo frames or parameter tuning needed

Current Nordunet load map

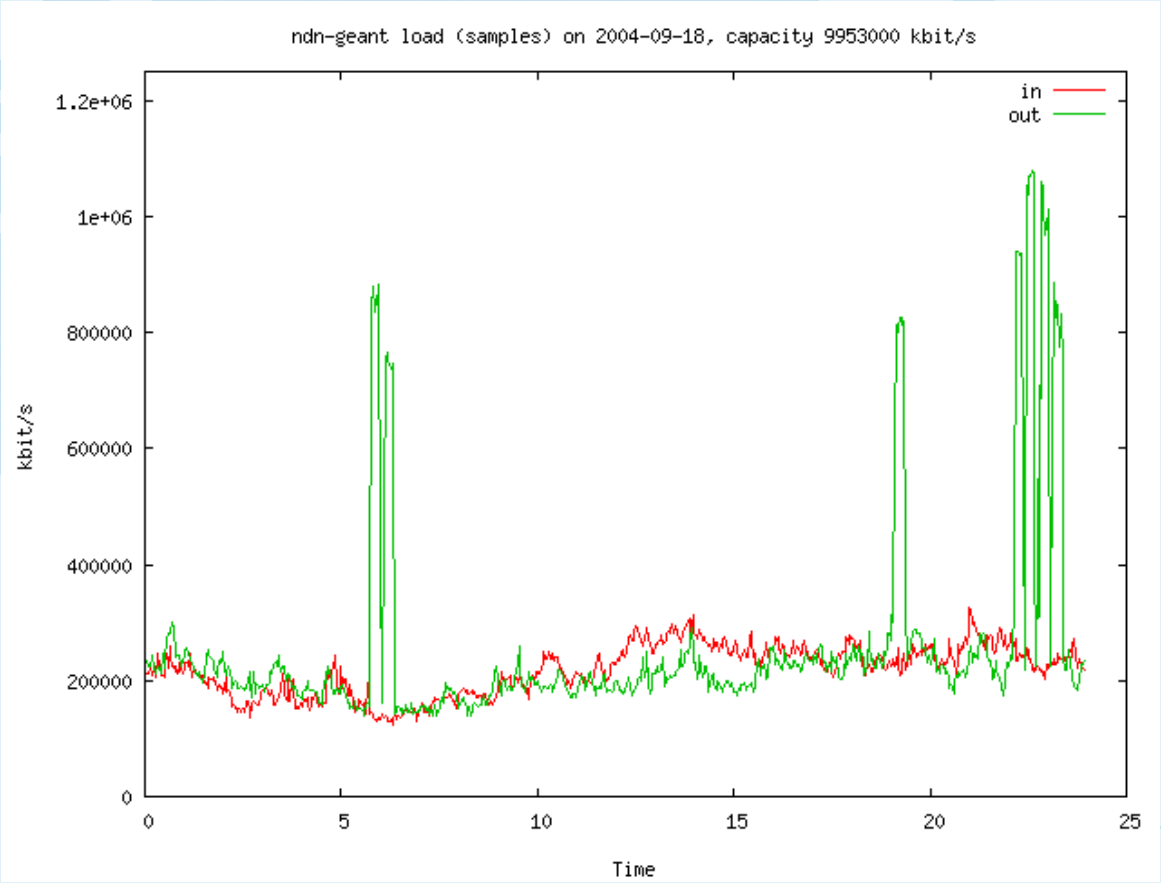


Our first transfer tests to JIVE

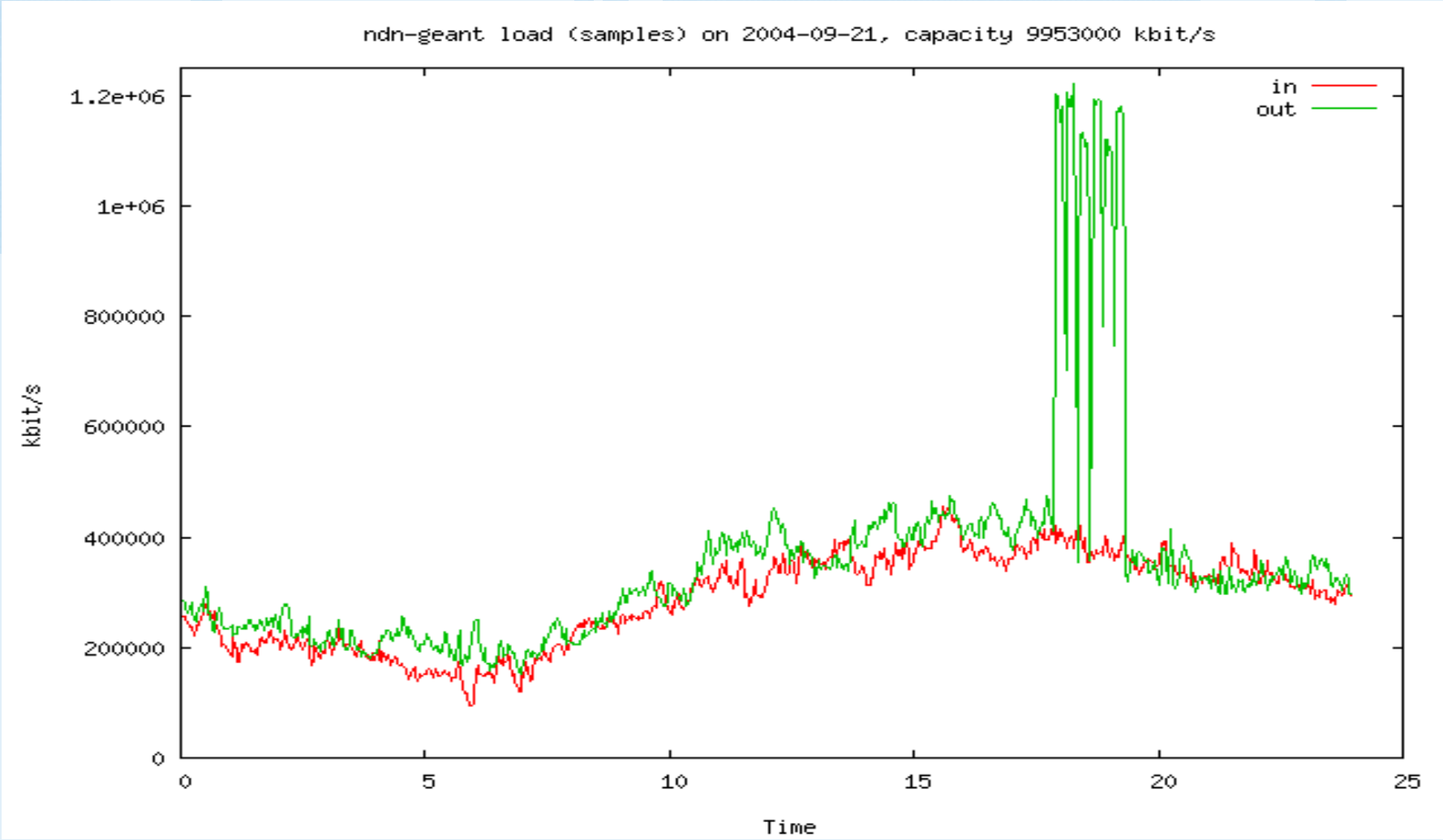


Tests continued next day...

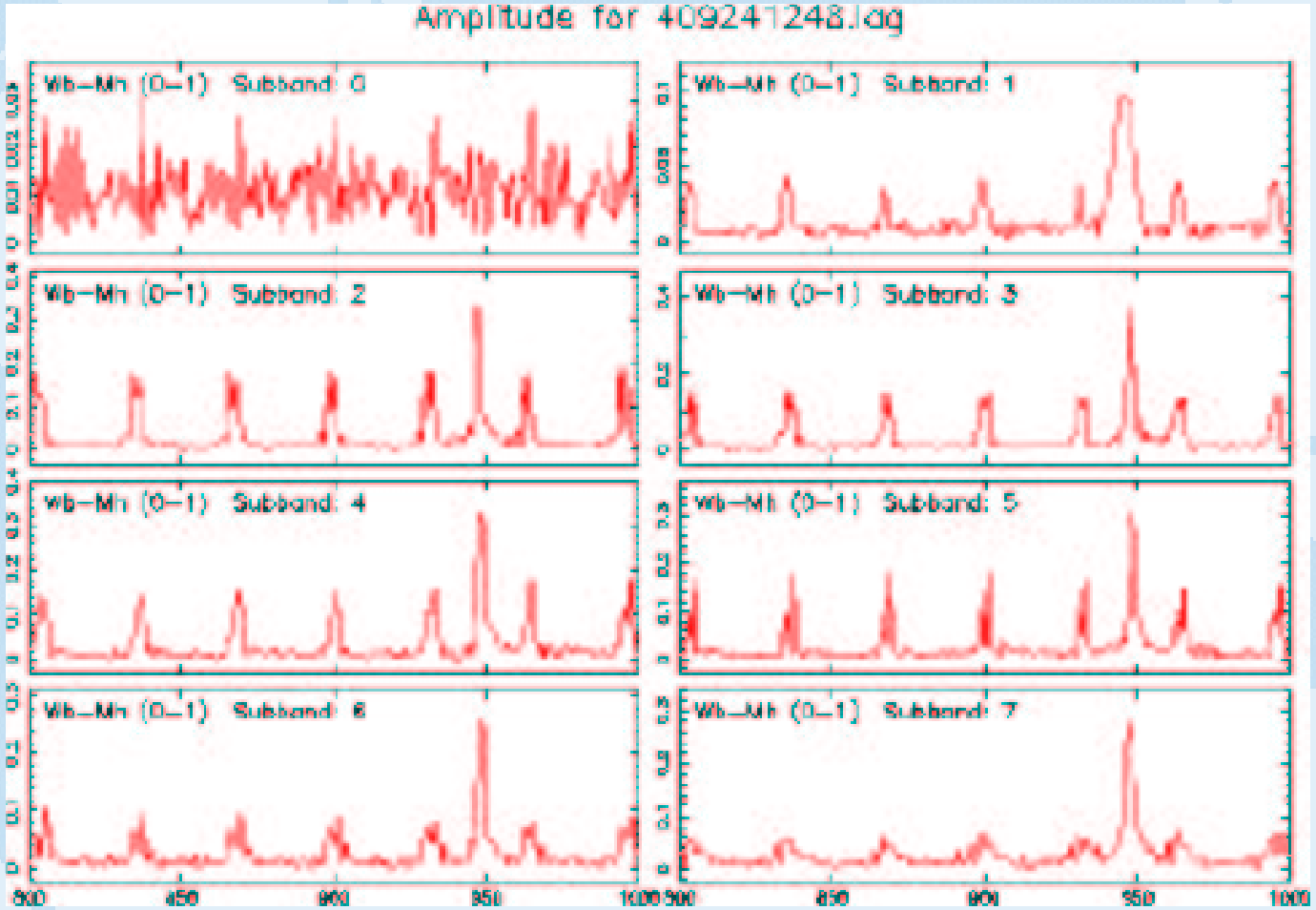
- Consistent >512 Mbit/s transfer rate
- Rate limited by 2-disk RAID array in JIVE



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



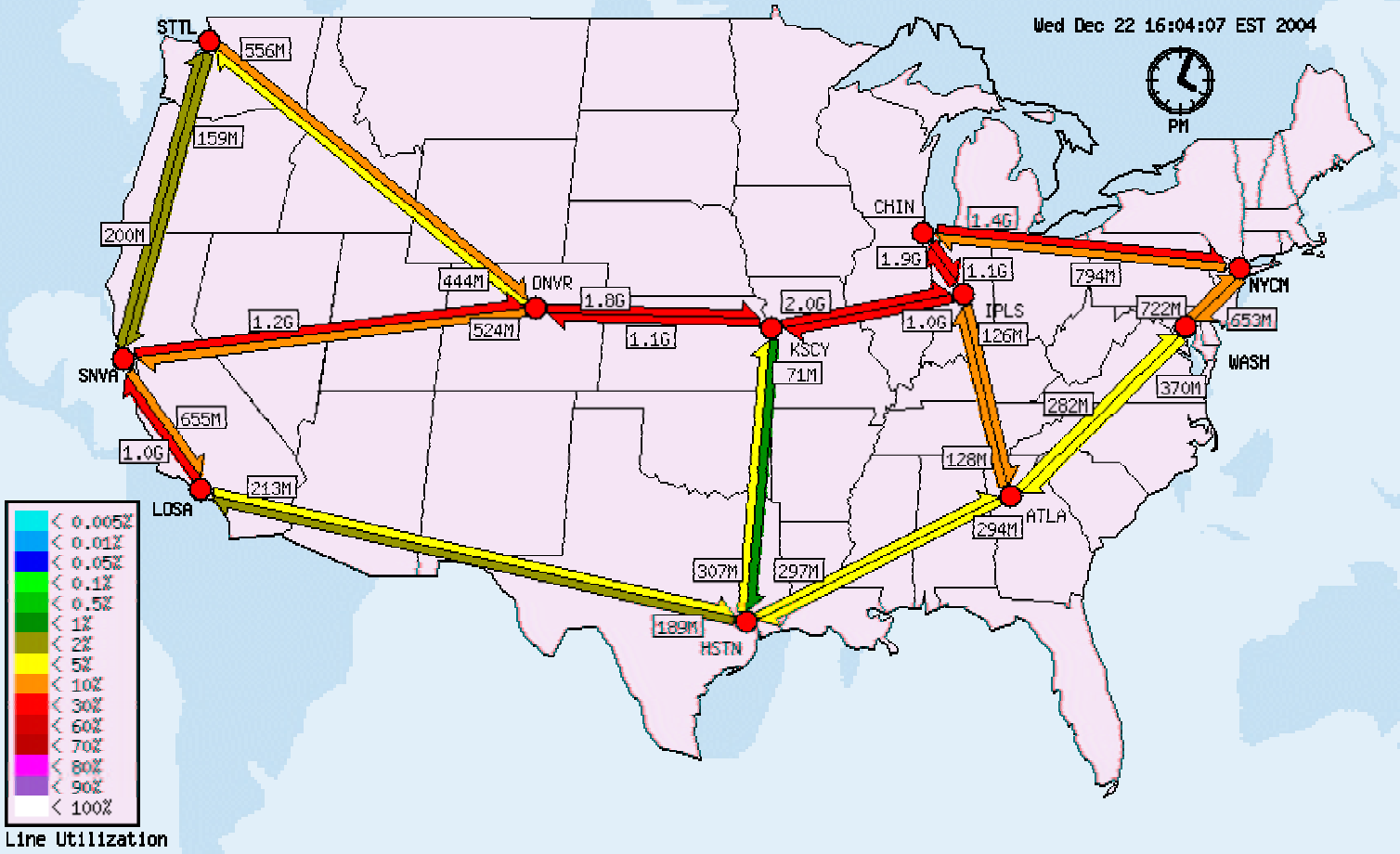
The data was correlated in JIVE on 24th September 2004



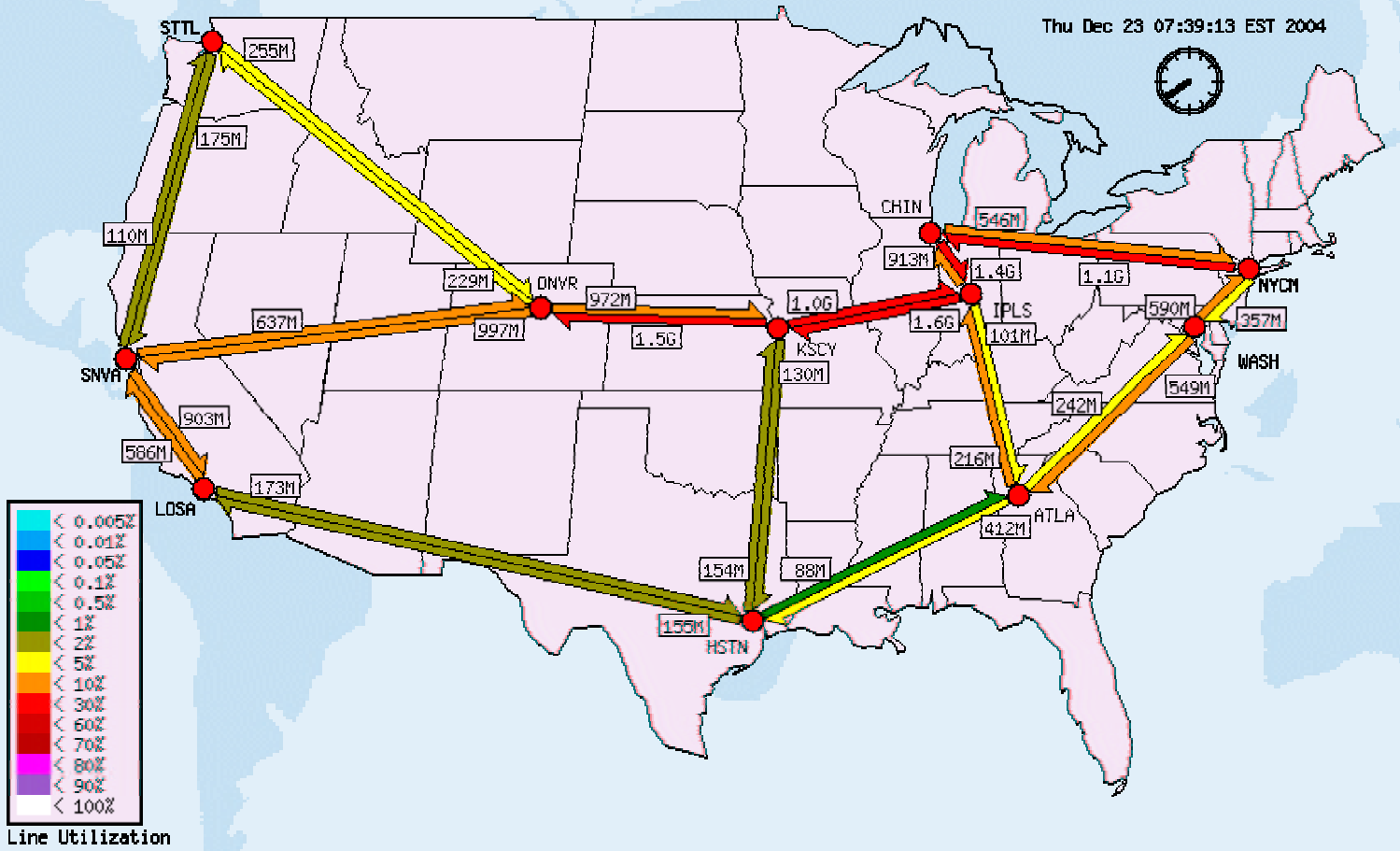
Japan-JIVE-Japan tests

- 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



JIVE-Nict transfer 2004-12-23



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)

Australian developments

- Total of 17 PCEVN recorders in Australia
- Chris Phillips, Tasso Tzioumis and Jamil Zaman of the CSIRO have experimented with Tsunami
- Jamil Zaman has improved the code
- October 2004: Finnish-Australian transfer test, 100 Mbit/s out of two Australian 155 Mbit/s trunk lines
- 14th of January 2005: Two 50GB files transferred to JIVE at 300 Mbit/s

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