A feasibility study of a 4 Gbps data recorder for VLBI

Prepared for EXPReS

17 June 2008

Jouko Ritakari , Jan Wagner, Guifré Molera, Ari Mujunen, Minttu Uunila

Metsähovi Radio Observatory, TKK

EXECUTIVE SUMMARY

This is a preliminary report about the feasibility of building a next-generation four-gigabit-per-second VLBI data recorder using commercial-off-the-shelf computer compo-nents and standard Linux operating system. The preliminary results are very promising, reliable 4Gbps recording from 10Gbps Ethernet to disks has been achieved. This speed will further improve in the future when new faster disks and computer hardware become available.

The 4G-EXPReS data recorder can have a Mark5C compatibility mode, so extra software effort in other parts of the VLBI station can be avoided. The work has been partially funded by the European Commission FP6 EXPReS project and this completes the requirements for its deliverables.

GENERAL

This study has been done mainly by Jan Wagner and Guifré Molera from November 2007 to February 2008. Full lab notes, description of hardware etc. are available at <http://www.metsahovi.fi/en/vlbi/10gbps/>. The hardware is based on an Asus dual-processor motherboard that has twelve native SATA-II ports. Dual-core AMD processors were used.

DISK WRITE TESTS

The first tests were performed using twelve one-year-old 320GB and 250GB disks. These limited the transfer speed to the range of 3.5 Gbps.

Purchasing a set of new Samsung 750 GB hard disks removed this problem, disk write speed increased 50%. One terabyte Samsung disks would have been an even better choice, but more expensive. In the best case a RAID-0 array with 750 GB disks works at 9 Gbps (outer tracks) to 6 Gbps (inner tracks). The single RAID-0 throughput ranged between 6 and 4 Gbps and is shown in Figure 1.

10 Gbps ETHERNET TESTS

The original idea was to use the programmable features of the Chelsio 10 Gbps Ethernet controller to place the incoming packets into a huge ring buffer. This would have produced a zero-processor-load 10 Gbps data reception. Unfortunately, the Chelsio chip was programmable for TCP streams only so this approach failed. However both Chelsio and Myrinet boards proved to be fast enough, reaching 4.5 Gbps using standard 1500-byte frames and more than 9 gigabits per second using jumbo frames.

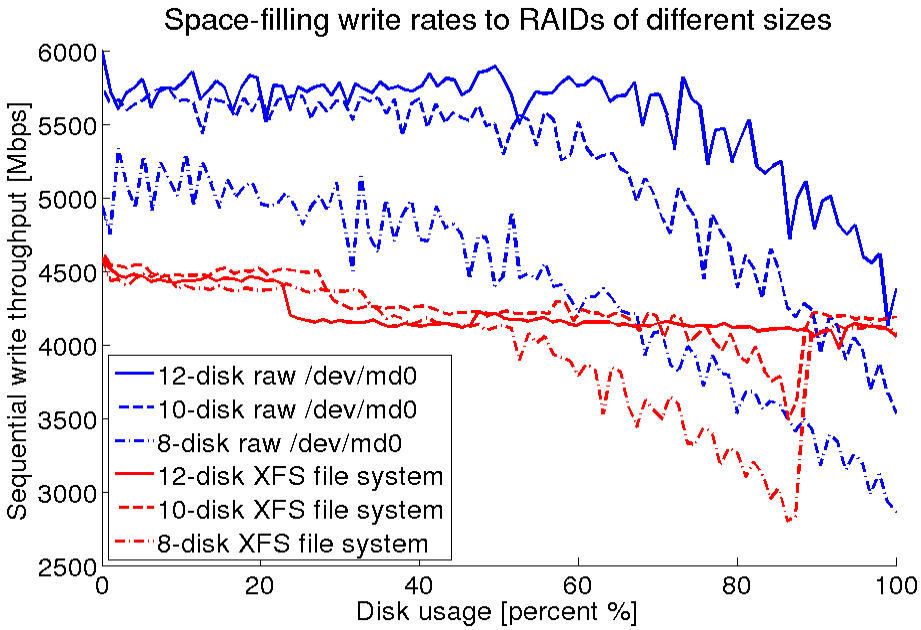

Figure 1: Sequential write rate for the RAID-0 for different numbers of 750GB disks and with or without an XFS file system

COMBINED NETWORK & DISK PERFORMANCE

Both UDP packet receiving and disk writing are CPU-intensive tasks. At 4 Gbps speed these tasks use fully the power of two processor cores. The test system had two dual-core processors so the CPU usage was kept on a comfortable level. Further improvement can be achieved by buying four-core processors.

It is also necessary to divide the program into several threads. The threads can then run on different cores. Fortunately Linux does this automatically, no programming effort was needed and the workload was evenly divided. Fastest transfer speeds were achieved with two six-disk RAID arrays and two different Tsunami programs. These already have two threads each. Even with a single Tsunami program a transfer rate of 4 Gbps can be sustained, see Figure 2. It must be noted that these numbers include the Tsunami protocol overhead. For Mark5C-like operation the speeds will be even better.

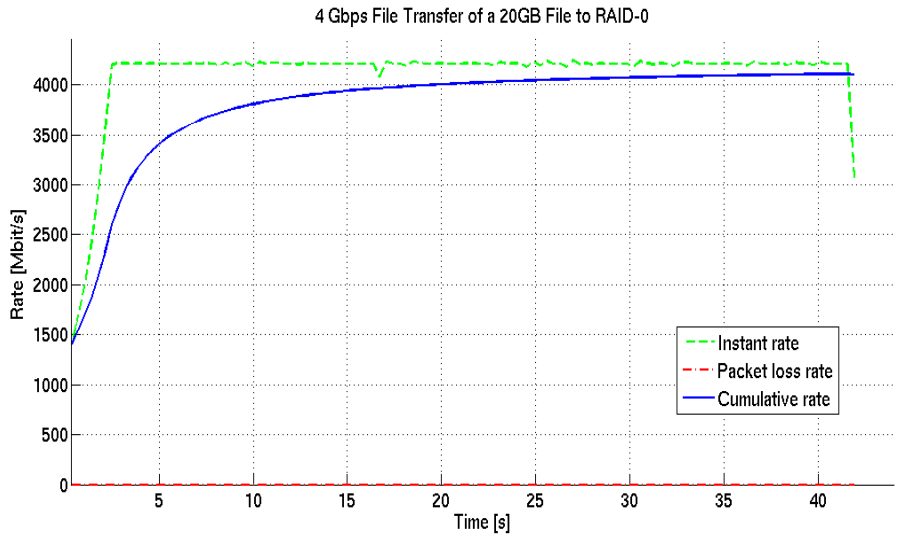

Figure 2: Transfer throughput for the Petabit Tsunami v1.2 transferring a 20GB fileover 10G LAN to a 12-disk RAID-0 at a 4.3 Gbps target rate