Metsähovi Four-Gbps Data Recorder for VLBI and eVLBI

Jouko Ritakari, Jan Wagner,
Guifré Molera
Metsähovi Radio Observatory

Starting point: The Metsähovi correlator engine

- Developed for the Cell processor (Sony Playstation 3) in 1H 2007

- About ten times more powerful than Intel-based solutions

- Less platform-dependent: Only a few pages of code need to be rewritten if platform changes

- Needs high-speed input data

Problem(s):

- Getting data in and out of specialised disk controller and old PATA disk packs is slow

- Data input is moving to 10 Gbps Ethernet when iBob or DBBC are used

- Testing four-Gbps streaming is easier with four-Gbps recorder

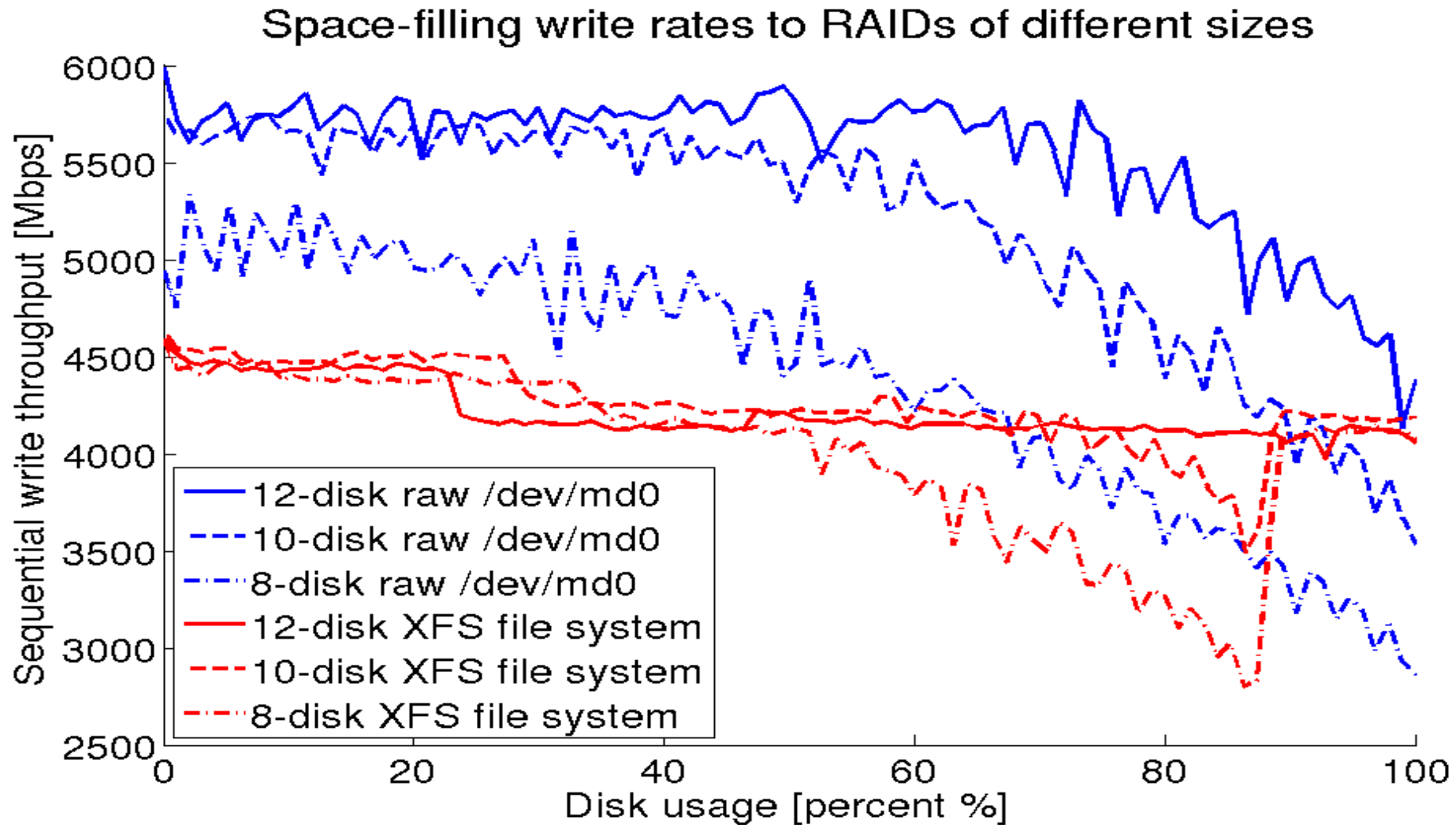- The seven-year-old VLBI data recorder designs would benefit from speedup

Solution:

➢ Feasibility study: a four-Gbps recorder made from only commercial off-the-shelf components

➢ 10 Gbps Ethernet for input/output

➢ Fast international data streaming using the Tsunami protocol

➢ Fallback plan: If it does not work at four Gbps, use two at two Gbps each

Disk subsystem tests

- Original tests with 12-disk raid array, one-year-old Seagate 250 GB disks

- Best write speed with ext2 fs 5.5Gbps, decreases to half of that on inner tracks

- Replaced the disks to newest and fastest Samsung Spinpoint F1 750 GB

- Consistent > 4Gbps performance

Samsung 750GB disk performance

Space-filling write rates to RAIDs of different sizes

Legend:
- 12-disk raw /dev/md0
- 10-disk raw /dev/md0
- 8-disk raw /dev/md0
- 12-disk XFS file system
- 10-disk XFS file system
- 8-disk XFS file system

Y-axis: Sequential write throughput [Mbps]
X-axis: Disk usage [percent %]

10 Gbps Ethernet tests

- Equipment: Myrinet and Chelsio NICs
- Initial iperf tests: 2-3 Gbps, 100% CPU
- Improved to 8 Gbps with jumbo frames
- Turned off udp checksum, Ethernet has already one.
- Final iperf UDP results:
  - MTU 1500: 4.88 Gbps
  - MTU 4470: 8.55 Gbps
 - MTU 9000: 9.04 Gbps
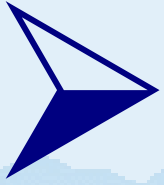
Combined 10GBE + disks

- Both tasks are CPU-limited
- Used two dual-core AMD 2212 processors and threads to divide the load
- Optimum performance with two six-disk raid arrays and two udp streams
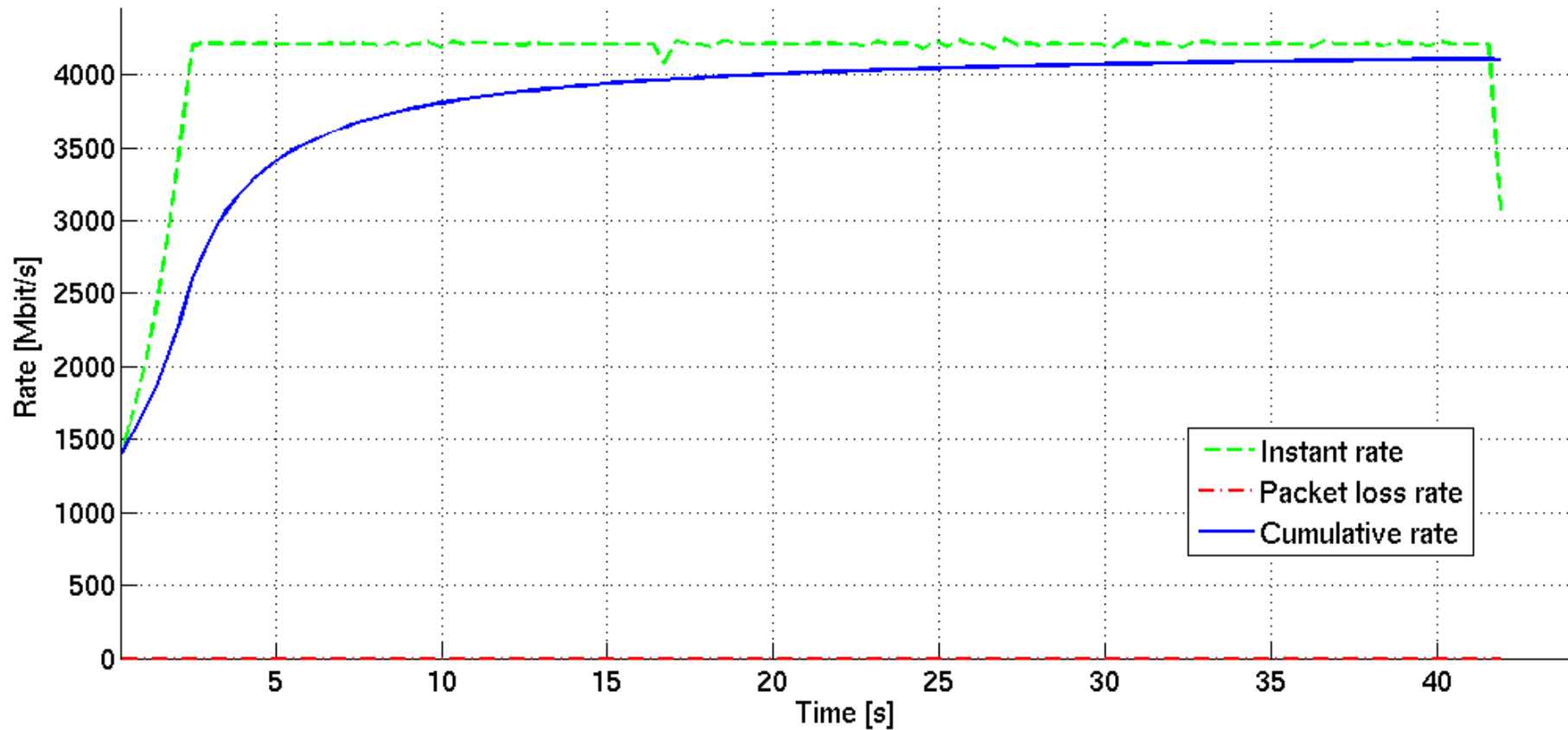- Sustained 4 Gbps transfer from 10GBE to disks

Tsunami tests

- Four-Gbps tsunami would be very useful

- ...so would be direct transfer from iBob udp/ip packetizer to correlator...

- A single tsunami transfer runs at 4 Gbps, two parallel transfers are a little faster

- And: Zero packet loss!!!
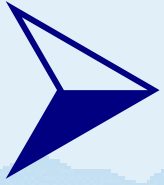
Lossless 4 Gbps Tsunami transfer

4 Gbps File Transfer of a 20GB File to RAID-0

Conclusions

➢ A 4 Gbps recorder is feasible now

➢ At this time the recorder is CPU-limited, situation will improve with time or with the use of four-core processors

➢ Only the newest and fastest SATA2 disks have the required performance

➢ Both Chelsio and Myrinet boards are fast enough

➢ Jumbo frames are needed at 4 Gbps

Recorder hardware

➢ Asus L1N64-SLI WS motherboard, nForce 680A chipset, 12 SATA2 ports

➢ Two dual-core AMD 2212 CPUs

➢ 4GB RAM

➢ Chelsio N320-CX 10Gbps Ethernet NIC

➢ Twelve Samsung Spinpoint F1 750 GB disks