Towards 10 Gbps e-VLBI

Ari Mujunen, Jouko Ritakari, Jan Wagner, Guifré Molera
Metsähovi Radio Observatory
Ari.Mujunen@hut.fi, Jouko.Ritakari@hut.fi
Outline

- Case: establishing the 10Gbps connection at Metsähovi
- Computer, disk, networking progress and trends in the market
- Tapping the power of mixing 1G/10G networking for a continually upgradable modular system
Metsähovi 10Gbps Connection

- Leased dark fiber to Funet, ~50.2km
- 1U Extreme X450 switch
  - 24 1G <---> 2 10G XENPAKs
  - 10GBase-ZR XENPAK
- Most of the cost is this “-ZR” optics to reach >40km
  - The switch is 3.7k€, optics 9.5+1.6k€
Test Setups

- Added a few VSIB PCs with 1Gbps copper Eth
Fierce Speed of Developments

- Already after 3 months, the price of -ZR optics dropped from 9500€ --> USD 5622
- The older, larger, more expensive XENPAK optics are going out of fashion
  - Replaced by smaller XFP modules
- Media 10GBase-CX4(copper) / (optics) -SR/-LR/-ER/-ZR
  - CAT6 copper dev group? CX4 needs expensive Infiniband cabling
The “1G” Computers

- **tetricus**: a 900€ “multimedia PC”
  - Intel i945G-based Asus P5LD2-VM
  - A comparable Dell OptiPlex GX620 tested, too
- **maximus**: a 300€ “budget PC”
  - nVidia C51G-based AcerPower M6
- **juliano**: a 1500€ “max gaming PC”
  - Intel i975X-based Asus P5WD2-E Premium
    - 2x4xSATA2 disks, 2x1GE via internal PCIe
The “1G” Computers Delivered

- All were capable of reliably streaming 800Mbps from the old PCI32 VSIB board
- PCI Express-based chipsets seem to dedicate one channel to legacy PCI32 alone
- The disk subsystem exceeds 1.1Gbps (w+r) already with only 3 SATA disks
- Used new 99€ Seagate 320GB SATA2 disks (in slower SATA1 mode) that use PMR (perpendicular magnetic recording)
Talking about Disk Storage...

- Celebrating the 50th anniversary of HDs :-)  
  - (IBM RAMAC 5MB system, 13-Sep-2006)
- SATA2 (3Gbps) dominates the market
- Perpendicular magnetic recording (PMR)
- Keeping industry on track with predictions
- E.g. Hitachi:
  - PMR demo of 345 gigabits per sq.in. in Sep-2006
  - Predicts 1TB drives before end-of-2007, 2TB before end-of-2009
Serial ATA Disks Everywhere

➢ Extra broad range of SATA products

www.copansys.com:
14 SATA drives/pack
112 /shelf
896 /cabinet

“MAID”, massive array of inexpensive disks
The “10G” Computers?

- E.g. successors of Dell PowerEdge 2900
- Intel 5000X-based multilane PCI Express
- FBD memories (bw up from 6-->21Gbytes/s)
- Have true potential of driving “10Gbps-class” PCIe x8 peripherals
- Are these coming up?
Manufacturers Believe in “10G”

- 10G Ethernet PCI Express boards appearing
  - www.myricom.com: Myri-10G PCIe x8
  - 10GBase-CX4 USD795, 10GBase-R USD 895+XFP
  - www.netxen.com: PCIe x8 for est. USD600
  - NXB-10GXxR for XFP or NXB-10GCX4 for CX4 copper

left arrow Myri

NetXen right arrow
Not only 10Gbps Ethernet...

- Disk controllers with “10G-class” bandwidth have appeared, too

- For example:
  - www.promise.com:
    - Promise SuperTrak EX16350
      - 16 SATA2 ports
      - PCIe x8

- The motherboards will quickly follow to allow these boards to operate at 10G rates
Why Will 10G Enter the Mainstream?

- It doesn't cost much more
- Faster memories and multi-core CPUs need “a way out of the box”
- Virtualization boom strikes data centers
  - Everyone wants to centralize disks into huge SAN boxes
  - Servers will be consolidated into fewer high-performance boxes
- 10G data center networking is needed between boxes!
10G LAN Switches

- 1U “pizza box” generation already in shops
- Such as:
  - www.smc.com: SMC8708L2 8-port 10G XFP
    - in stock www.pcnation.com, ships in 24h at USD 6220
  - www.hp.com: 6400CL series
    - ProCurve 6400cl, 6 CX4 copper ports + 2 XFP slots
      - in stock www.lacc.com, USD 3384
    - ProCurve 6410cl, 6 + 2 XFP slots
      - in stock www.lacc.com, USD 5648
      - at www.hp.com, USD 8099...
All the 10G parts are already available!

- Not only available, but in (or already past) the “rapid price decline” phase
- Mix-and-match:

![Diagram](image-url)
Use 1/10G to Bring Flexibility

- Use 10G samplers with a “farm” of 1G or 2x1G PCs
- Local disks for “store and forward”
- Start with 1—2 1G remote links if 10G n/a
- Upgrade just the switches when 10G available
- Use alternately direct eVLBI or a “storage farm”
- Use the net to distribute data to sw correlator processing nodes, local or remote
An Example of “Store and Forward Farm”

- Six econo-1G-PCs with 1TB swappable SA-TA2 disk space in each, Extreme X450, -SR XENPAK for local connection to 10G sampler, a slot for remote (or 10G switch) XENPAK 10G link
  - $6\times(300+64+3\times99)+(3700+1600+2000) = 11.3k\$€$
  - Supports at least 4.8Gbps from 10G sampler to disk buffer for over 2.5h recording
- 10G is not only available, it is affordable!