Local Area Networking at the CMC: Upcoming Changes

Peter Silva -- Head

Data Acq. & Dist. Syst.

“Transparent” Networking Changes.

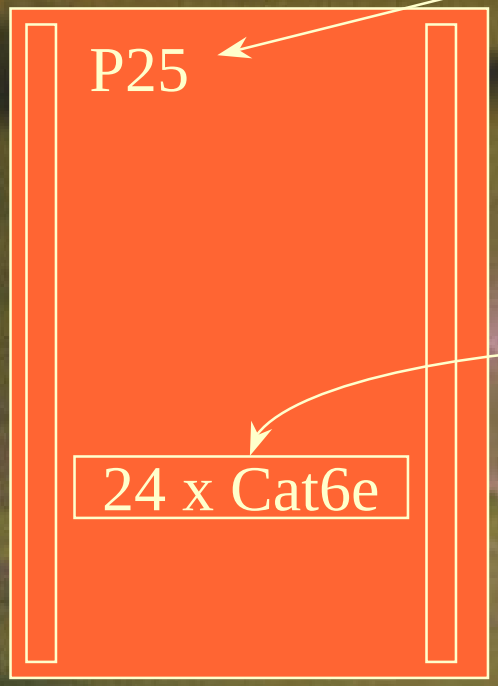
- Except when it breaks.
- Entire office LAN replaced last spring (provided fiber uplinks enabling move.)
- cabling between 1st floor and the new room will be run.
- stuff will move (no before/after difference)
- Informatics will be very busy.

The New Room: Better

- Power -- dual distribution systems
- Cabinets -- fewer ad-hoc installs
  - standardized (cable management)
  - Dual power fed
  - network via in cabinet patch panels.
  - named by grid location.
- Simpler, faster, thorough installations

The Standard Cabinet

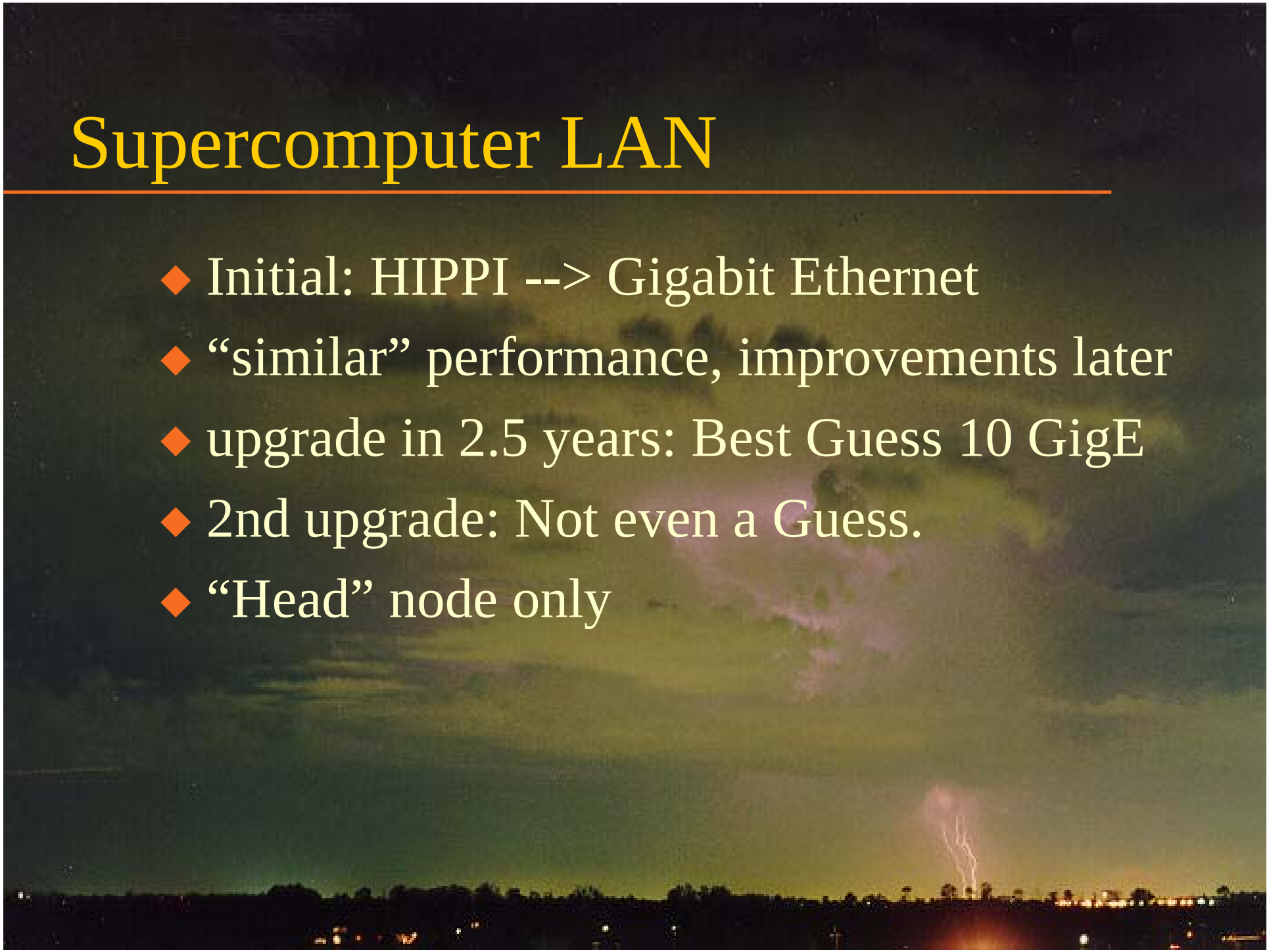
- extra-wide cabinets to allow space to run cables.
- Dual, separately fed power
- bulk cabling.



Cabinets are named by grid location

Integrated Patch Panel console & eth.

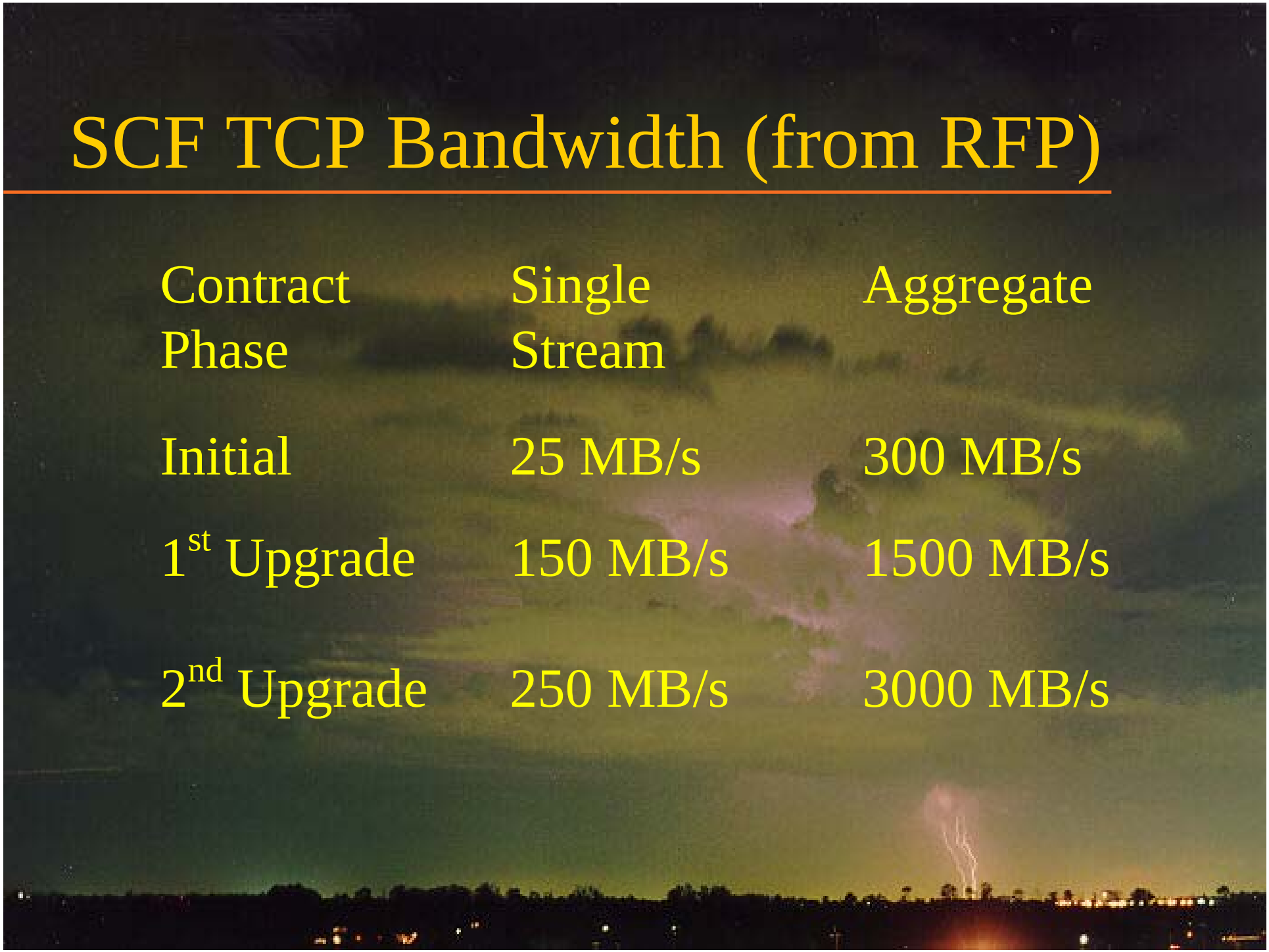
Supercomputer LAN



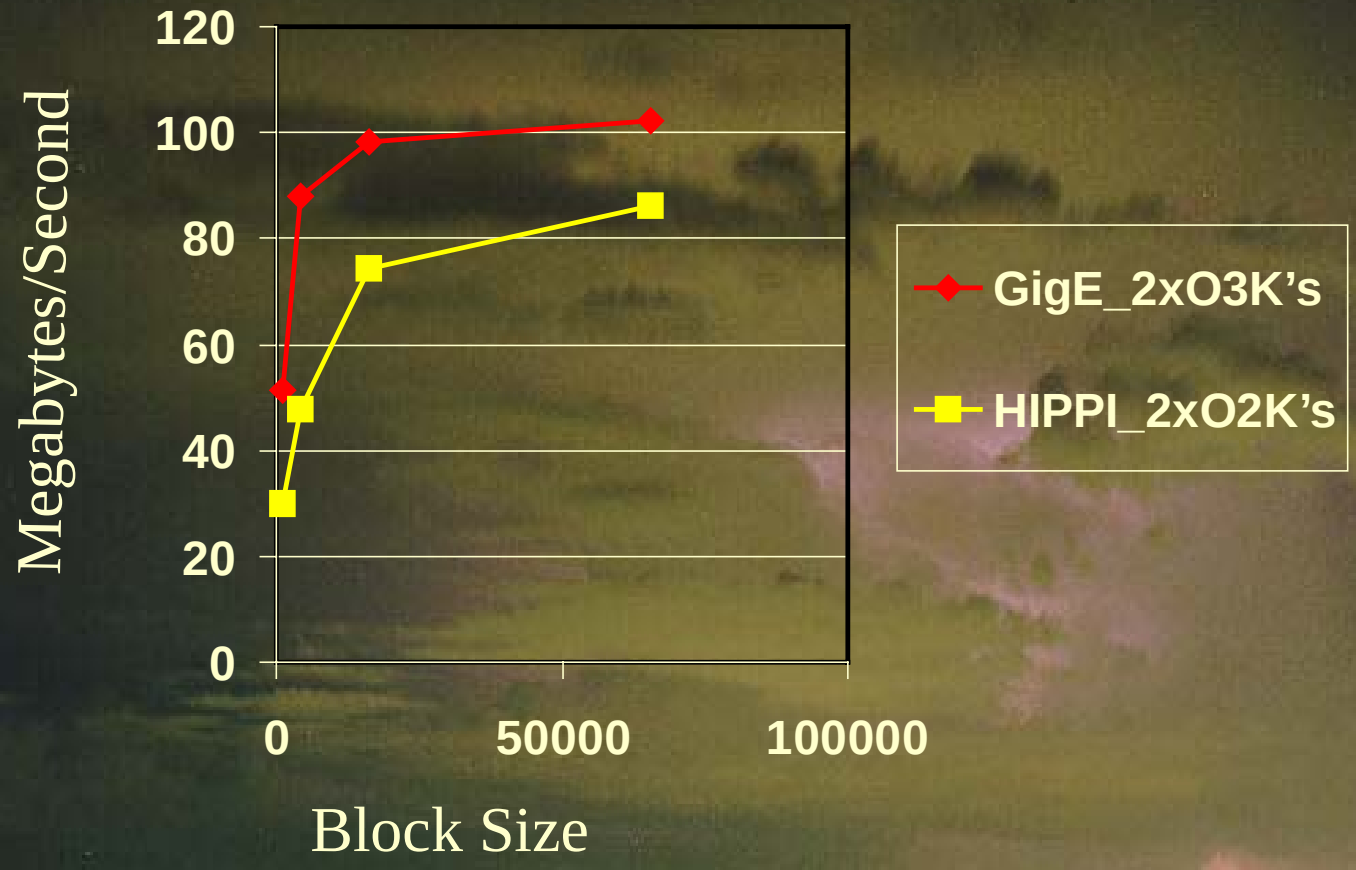
- Initial: HIPPI --> Gigabit Ethernet
- “similar” performance, improvements later
- upgrade in 2.5 years: Best Guess 10 GigE
- 2nd upgrade: Not even a Guess.
- “Head” node only

SCF TCP Bandwidth (from RFP)

Contract Phase	Single Stream	Aggregate
Initial	25 MB/s	300 MB/s
1st Upgrade	150 MB/s	1500 MB/s
2nd Upgrade	250 MB/s	3000 MB/s



HIPPI vs. GigE Performance

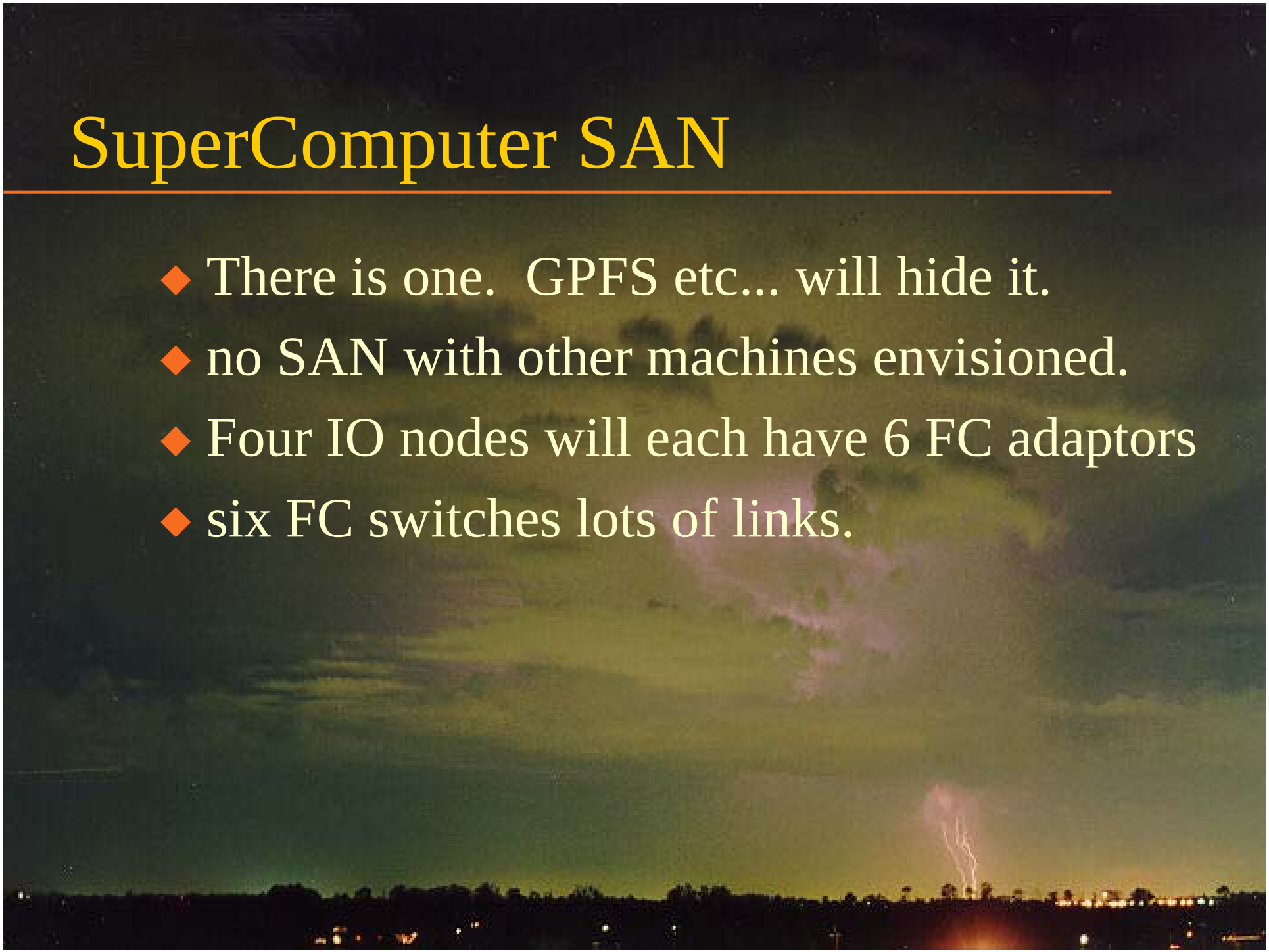


single channel: GigE-tcpspeed, HIPPI-smcp.

GigE Transition

- SGI o2000 gigE performance: Not!
- NEC SX-6: functional, 20% worse than HIPPI, use more interfaces.
- SGI o3000 GigE 20% better than HIPPI (SGI approaches wire speed in both cases.)
- IBM ? Que Sera, Sera
- FoundryNetworks.net FastIron 1500

SuperComputer SAN



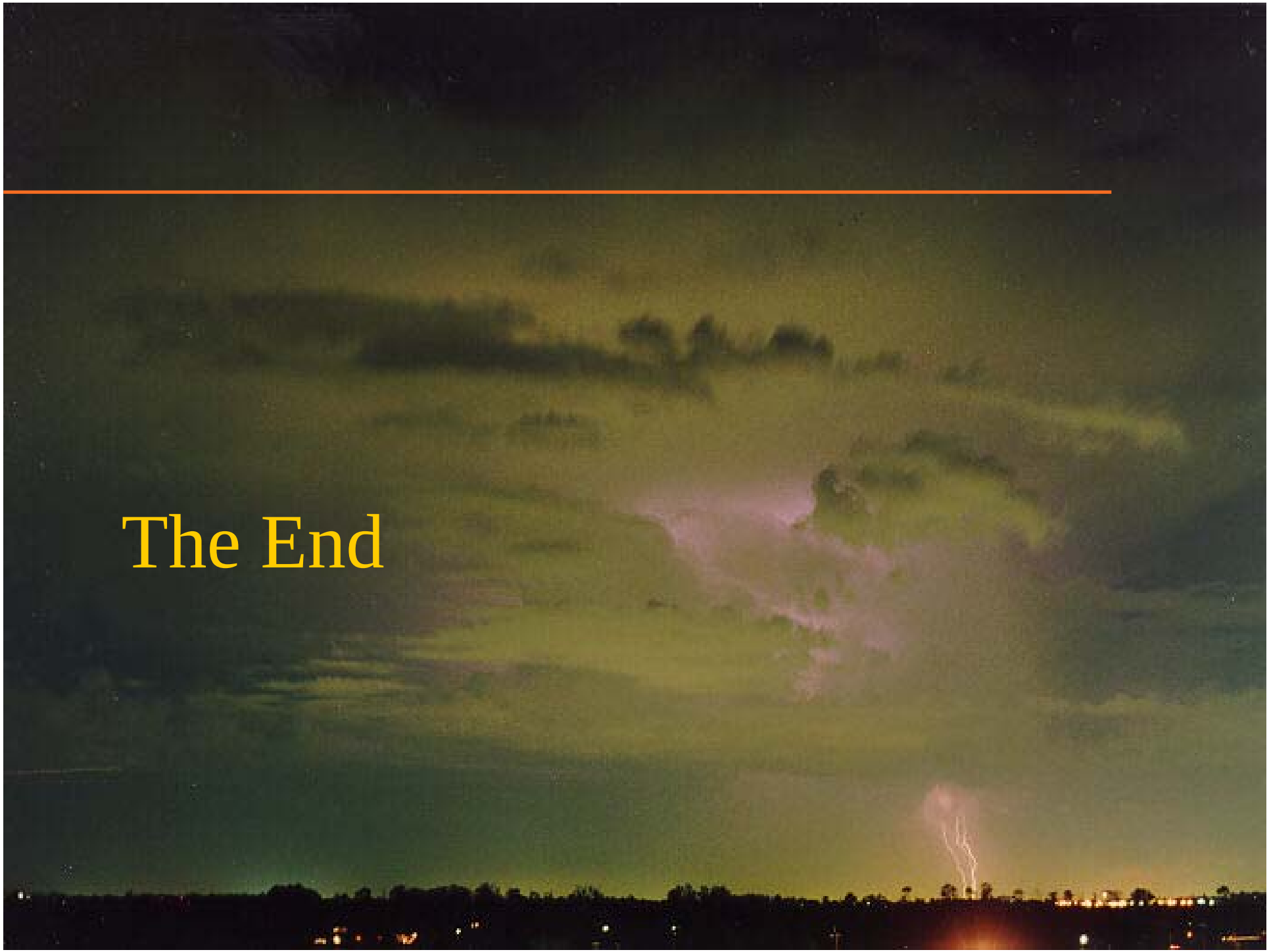
- There is one. GPFS etc... will hide it.
- no SAN with other machines envisioned.
- Four IO nodes will each have 6 FC adaptors
- six FC switches lots of links.

Stuff We Don’t Know

- 10 GigE is standardized. SCSI over IP using 10 GigE could be interesting. Replace FC?
- CA*Net: The Terabyte transfer.
- desktop networking:
  - Will people ever use it? NoW/Grid...
  - Will people ever have fewer machines ? (addressing on network 4)

Conclusions

- A lot of “invisible” activity, very busy informatics group, for at least one year, maybe two.
- Improved processes will accellerate and improve the quality of installations
- SCF LAN will be commodity tech, more reliable, head node centric, not much faster initially.


The End

LAN Future

- Trends affecting planing
- Transition to the new computer room
- Transition to the new supercomputer (LAN and SAN)
- focus:
  - impact
  - opportunities

Example

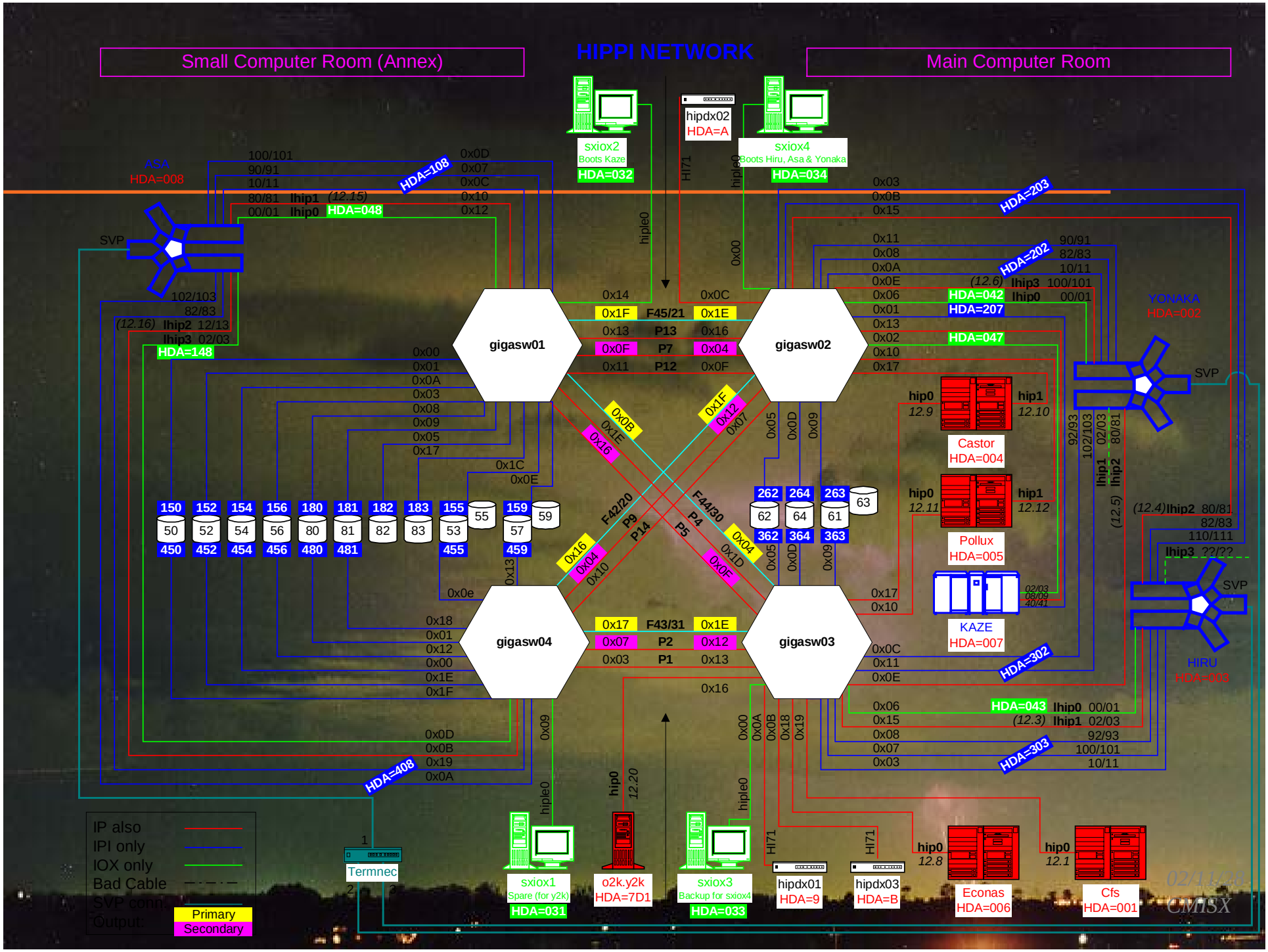
- Single Window: 4 LANS, 6 switches, 3 or 4 interfaces/host. And that is without host-based HA or SAN. (maybe insert a slide here.)
- AWWS -- 6 interfaces per machine (host based HA) + 4 for SAN.

Trends

- large numbers of standard, smaller components replace small numbers of larger specialized ones
- many installations, more management needed.
- More interfaces per system.
- expectations of High Availability
- lowering tolerance for maintenance
- need for (sw) maintenance remains.

Adieu HIPPI

- Was the greatest in 1992.
- First SAN in Environment Canada (1993)
- First combined storage and interconnect network. Lessons Learned still valuable.
- performance still competitive.
- minimal industry/market support


Small Computer Room (Annex)
HIPPI NETWORK
Main Computer Room
ASA
HDA=008
SVP
HDA=108
HDA=048
HDA=148
sxiox2
Boots Kaze
HDA=032
hipdx02
HDA=A
sxiox4
Boots Hiru, Asa & Yonaka
HDA=034
HDA=203
HDA=202
HDA=042
HDA=207
HDA=047
YONAKA
HDA=002
gigasw01
gigasw02
gigasw04
gigasw03
F45/21
P13
P7
P12
F44/30
P4
P5
F42/20
P9
P14
F43/31
P2
P1
Castor
HDA=004
Pollux
HDA=005
KAZE
HDA=007
HIRU
HDA=003
HDA=302
HDA=043
HDA=303
HDA=408
Termnec
sxiox1
Spare (for y2k)
HDA=031
o2k.y2k
HDA=7D1
sxiox3
Backup for sxiox4
HDA=033
hipdx01
HDA=9
hipdx03
HDA=B
Econas
HDA=006
Cfs
HDA=001
IP also
IPI only
IOX only
Bad Cable
SVP conn
Output:
Primary
Secondary
02/11/06
CMSX