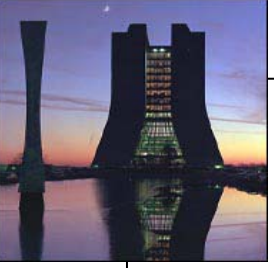
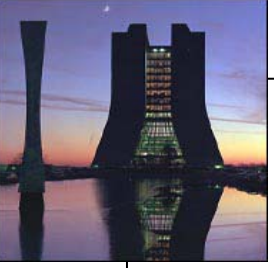


PASTA Review
Technology for the LHC Era

21 October 2002

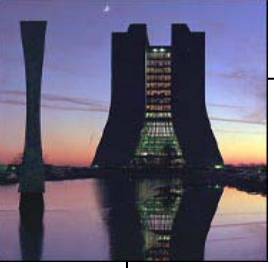
Michael Ernst, FNAL

Ernst@fnal.gov



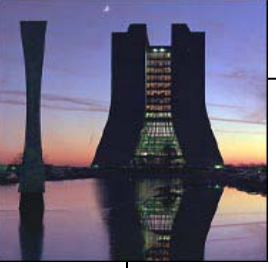
Approach to Pasta III

- Technology Review of what was expected from Pasta II and what might be expected in 2005 and beyond.
- Understand technology drivers which might be market and business driven. In particular the suppliers of basic technologies have undergone in many cases major business changes with divestment, mergers and acquisitions.
- Try to translate where possible into costs that will enable us to predict how things are evolving.
- Try to extract emerging best practices and use case studies wherever possible.
- Involve a wider number of people than CERN in major institutions in at least Europe and the US.



Participants

- A: Semiconductor Technology
  - Ian Fisk (UCSD) Alessandro Machioro (CERN) Don Petravik (Fermilab)
- B:Secondary Storage
  - Gordon Lee (CERN) Fabien Collin (CERN) Alberto Pace *(*CERN*)*
- C:Mass Storage
  - Charles Curran (CERN) Jean-Philippe Baud (CERN)
- D:Networking Technologies
  - Harvey Newman (Caltech) Olivier Martin (CERN) Simon Leinen *(*Switch*)*
- E:Data Management Technologies
  - Andrei Maslennikov (Caspur) David Foster (CERN)
- F:Storage Management Solutions
  - Michael Ernst (Fermilab) Nick Sinanis (CERN/CMS) Martin Gasthuber (DESY)
- G:High Performance Computing Solutions
  - Bernd Panzer (CERN) Ben Segal (CERN) Arie Van Praag *(*CERN*)*

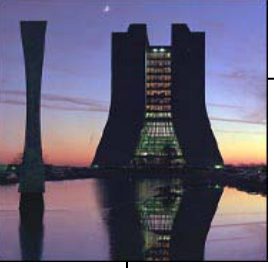


Current Status

- Most reports in the final stages.
- Networking is the last to complete.
- Some cosmetic treatments needed.

Draft Reports can be found at:

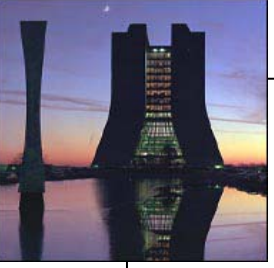
http://david.web.cern.ch/david/pasta/pasta2002.htm



SIA 1997 technology forecast

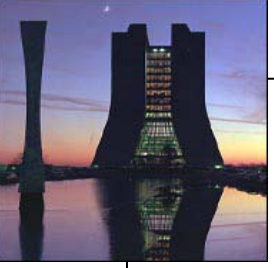
Year	1997	1999	2001	2003	2006	2009	2012
Technology requirements							
Dram ½ pitch (um)	.25	.18	.15	.13	.10	.07	.05
uP channel length	.20	.14	.12	.10	0.7	.05	.035
Tox equivalent (nm)	4-5	3-4	2-3	2-3	1.5-2	<1.5	<1.0
Gate Delay Metric CV/I (ps)	16-17	12-13	10-12	9-10	7	4-5	3-4
Overall Characteristics							
Transistor density (M/cm2)	3.7	6.2	10	18	39	84	180
Chip size (mm2)	300	340	385	430	520	620	750
Maximum Power (W)	70	90	110	130	160	170	175
Power supply voltage (V)	1.8-2.5	1.5-1.8	1.2-1.5	1.2-1.5	0.9-1.2	0.6-0.9	0.5-0.6
OCAC clock (high perf.)	750	1200	1400	1600	2000	2500	3000
OCAC clock (MHz) (cost perf.)	400	600	700	800	1100	1400	1800

(Known solution/Solution being pursued in 1999/No known solution in 1999)



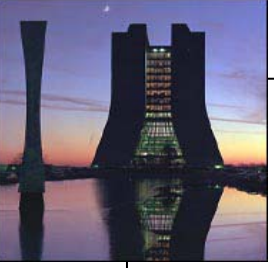
Year	1997	1999	2001	2003	2006	2009	2012
Basic cost of DRAM							
Dram capacity	256 Mb	1 Gbit		4 Gbit	16 Gbit	64 Gbit	256 Gbit
Cost/Mbit (USD/year1)	1.2	0.6		0.15	0.05	0.02	0.006
Processor cost							
Cost MTR (USD) year1	30	17.4	10	5.8	2.6	1.1	0.49
Processor cost (year1)	330	365	400	440	510	570	680

SIA 1997 pricing forecast



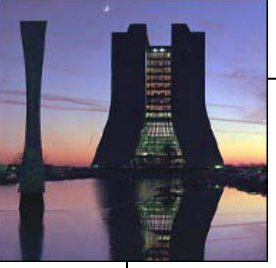
2002 SIA Technological Forecast

Year	2001	2002	2003	2004	2005	2006	2007
Technology Requirements							
DRAM ½ Pitch (nm)	130	115	100	90	80	70	65
Gate Length (nm)	90	75	65	53	45	40	35
Overall Characteristics							
Transistor Density (M/cm^2)	39	49	61	77	97	123	154
Chip Size (mm^2)	310	310	310	310	310	310	310
Maximum Power (W)	130	140	150	160	170	180	190
Power Supply Voltage (V)	1.1	1.0	1.0	1.0	0.9	0.9	0.7
OCAC Clock (MHz)	1,700	2,300	3,000	4,000	5,200	5,600	6,800



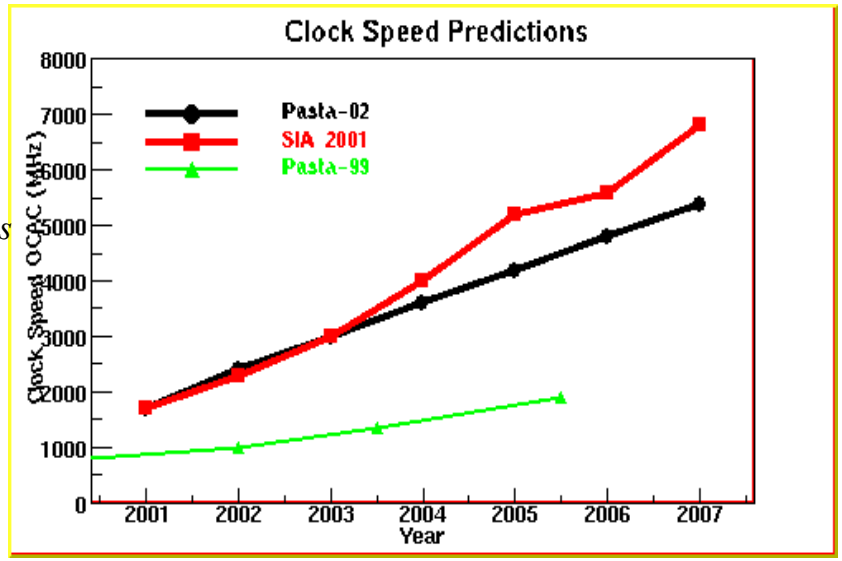
SIA long-term technology predictions

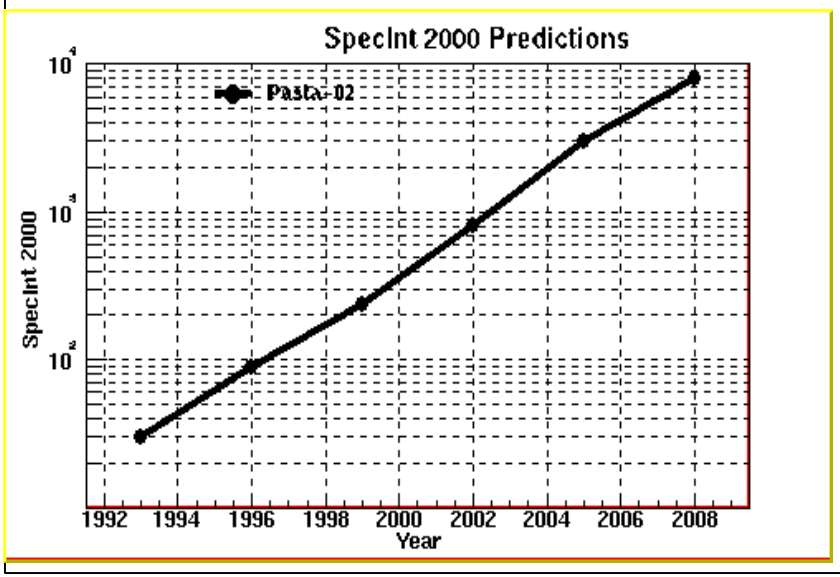
Year	2010	2013	2016
Technology Requirements			
DRAM ½ Pitch (nm)	45	32	22
Gate Length (nm)	18	13	9
Overall Characteristics			
Transistor Density (M/cm^2)	309	617	1235
Chip Size (mm^2)	310	310	310
Maximum Power (W)	215	250	290
Power Supply Voltage (V)	0.6	0.5	0.4
OCAC Clock (MHz)	11,500	19,300	28,800



Basic System Components - Processors

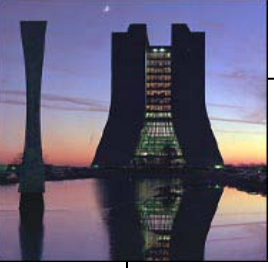
- *1999 Pasta report was conservative in terms of clock speed BUT, clock speed is not a good measure, with higher clock speed CPU's giving lower performance in some cases*
- *Predictions beyond 2007 hard to make, CMOS device structures will hit limits within next 10 years, change from optical litho to electron projection litho required => new infrastructure*





Specint 2000 numbers for high-end CPU.
Not a direct correlation with CERN Units.
P4 Xenon = 824 SI2000 but only 600 CERN units

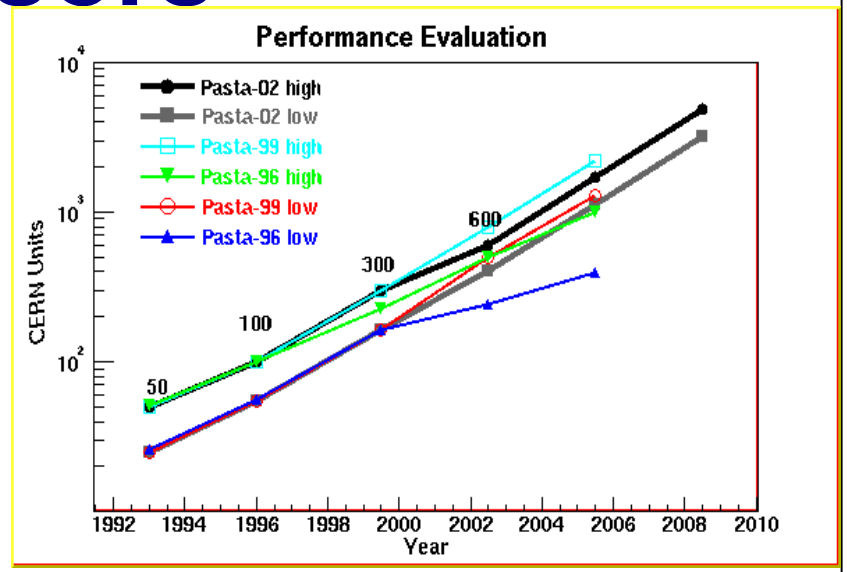
Compilers have not made great advances but Instruction Level Parallelism gives you now 70% usage (CERN Units) of quoted performance.

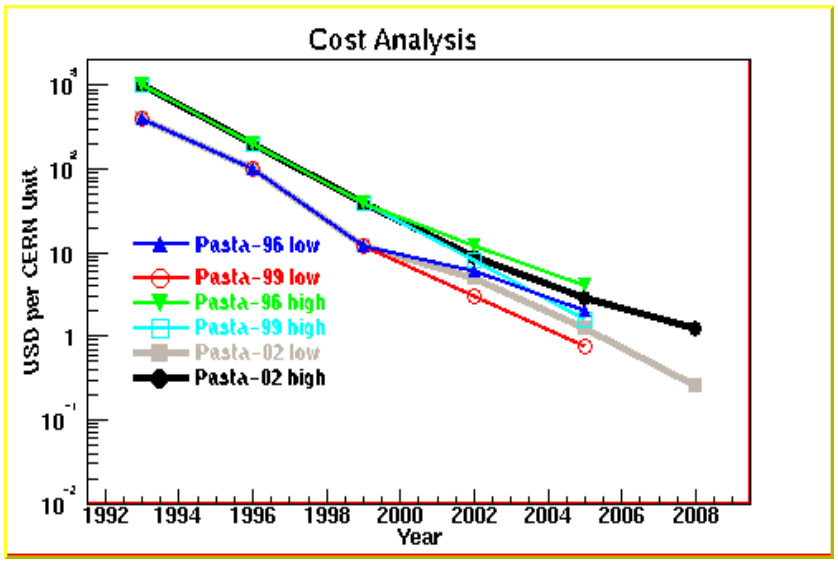


Basic System Components - Processors

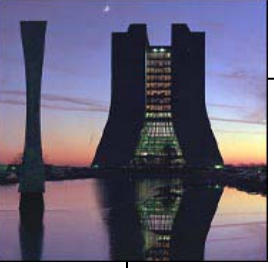
Performance evolution and associated cost evolution for both High-end machines (15K$ for quad processor) and Low-end Machines (2K$ for dual CPU)

Note 2002 predictions revised down slightly from the 1999 Predictions of actual system performance
- *'99 report: expect 50% of what Intel quotes, trend holds*
- *with hyperthreading (P4 XEON) agrees with '96 predictions reducing the gap from 50% to 30%*
- *ILP has not increased significantly*
- *IA-64 still not as good as recent P4*



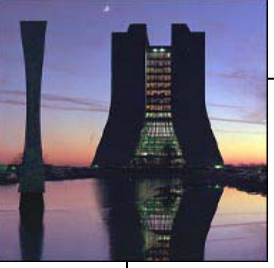


Fairly steep curve leading to LHC startup suggesting delayed purchases will save money (less CPU's for the same CU performance) as usual

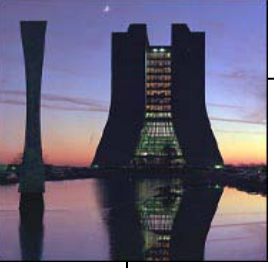


Basic System Components

- Predictions on physical properties in ’96, rev. in ‘99 too conservative
  - Much of clock improvements from changing pipeline structure
  - With 2000 CU systems in 2007 this is 1 year delay from ’99 prediction
- Memory capacity increased faster than predicted, costs around 0.15 $/Mbit in 2003 and 0.02 $/Mbit in 2005
- Many improvements in memory systems 300 MB/sec in 1999 now (2002) in excess of 1.5 GB/sec
  - Keeping pace with improvements in CPU performance
- Intel and AMD continue as competitors. Next generation AMD (Hammer) permits 32bit and 64bit code. And is expected to be 30% cheaper than equivalent Intel 64bit chips.
  - Comparison of Intel and AMD Processors (courtesy of Ian Fisk, UCSD)

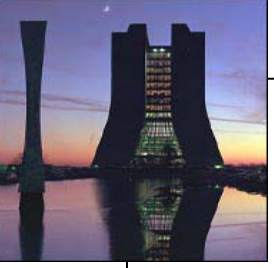


The little difference …



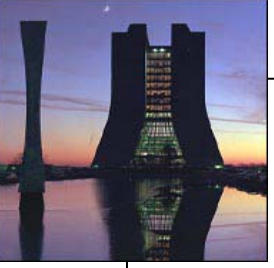
Basic System Components - Interconnects

- PCI Developments
  - PCI 66/64 mostly on servers
  - PCI-X introduction slow
  - PCI-X 2 standard with 266 MHz (2.13 GB/s) and 533 MHz (4.26 GB/s)
    - Supports DDR and QDR technology
  - PCI Express (alias 3GIO, project Arapahoe)
    - Internal Serial Bus, NOT an Interconnect
    - Primarily for high-end systems
- New Interconnects
  - 3GIO, Intels industrial proposal
  - HyperTransport, AMD (12.8 GB/s asym, bi-directional, 64 bit Bus)
  - Chipset includes routing crossbar switch
  - Connection to outside to connect peripherals
  - Superior to Intel, but will the market accept it ?

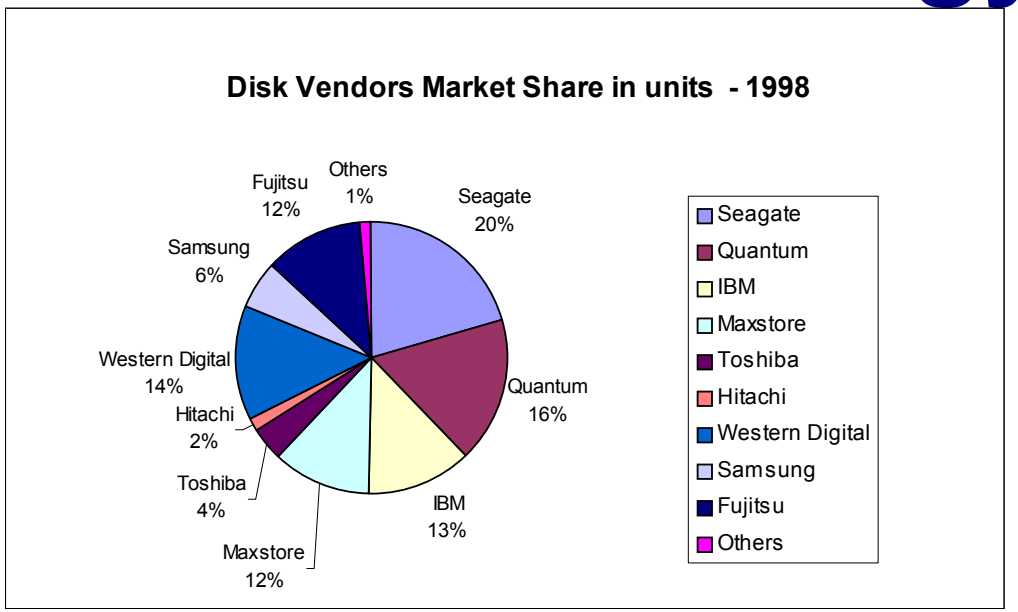


Basic System Components Some conclusions

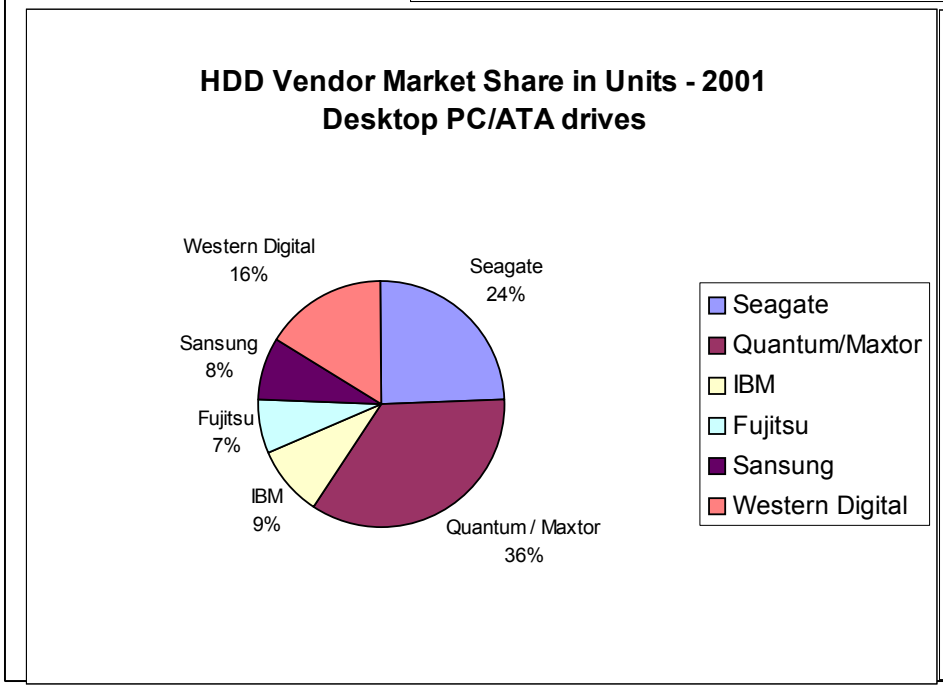
- No major surprises so far, but
  - New Semiconductor Fabs very expensive squeezing the semiconductor marketplace.
  - MOS technology is pushing again against physical limits – gate oxide thickness, junction volumes, lithography, power consumption.
  - Architectural designs are not able to efficiently use the increasing transistor density (20% performance improvement vs. 60% more transistors)
- Do we need a new HEP reference application ?
  - Using industry benchmarks still do not tell the whole story and we are interested in throughput.
  - Seems appropriate with new reconstruction/analysis models and code

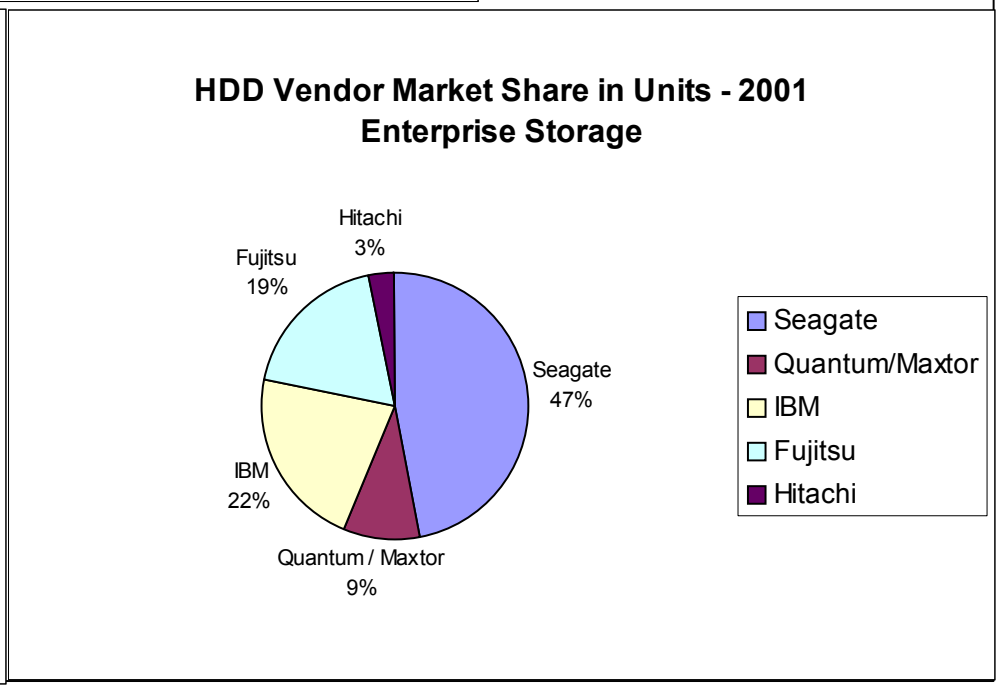


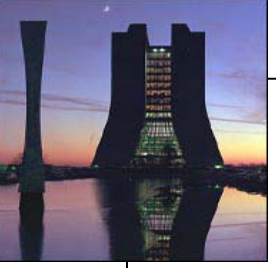
Disk Technology



Specialisation and consolidation of disk manufacturers

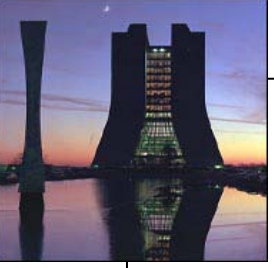




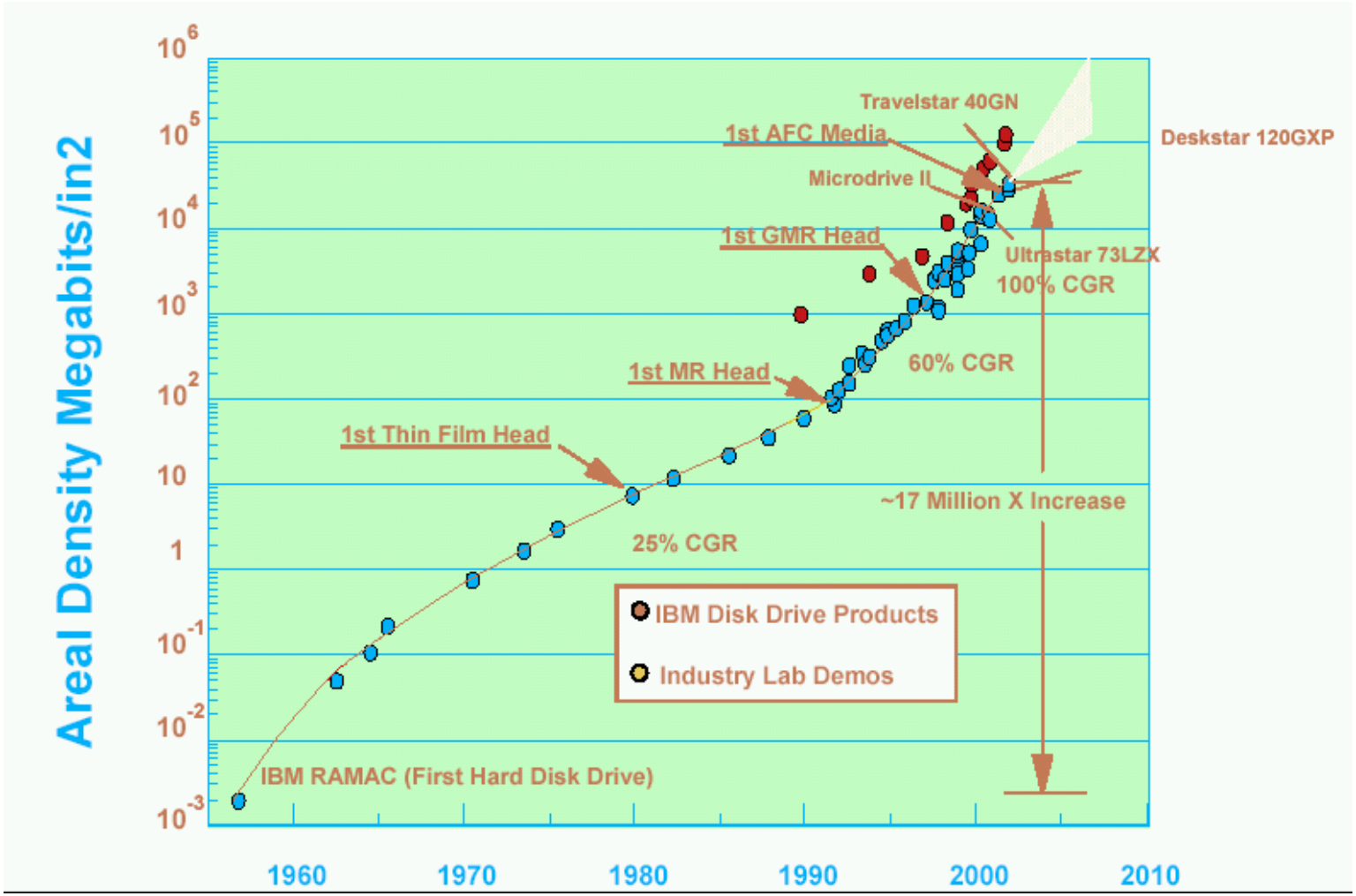


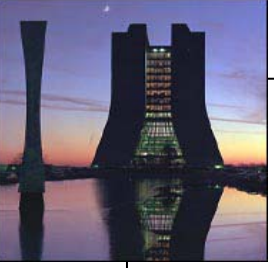
Disk Technology Trends

- Capacity is doubling every 18 months
- Super Paramagnetic Limit (estimated at 40GB/in^2) has not been reached. Seems that a platter capacity of 60 GB can be manufactured today, resulting in 500GB Drives.
- “Perpendicular recording” aims to extend the density to 500-1000GB/in^2. Disks of 10-100 times today’s capacity seem to be possible. The timing will be driven my market demand.
- Rotational speed and seek times are only improving slowly so to match disk size and transfer speed disks become smaller and faster. 2.5” with 23.500 RPM are foreseen for storage systems.



Historical Progress





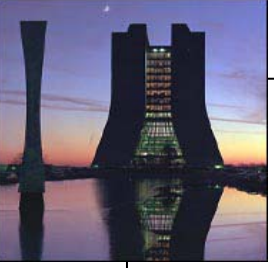
Disk Drive Projections

Performance Desktop

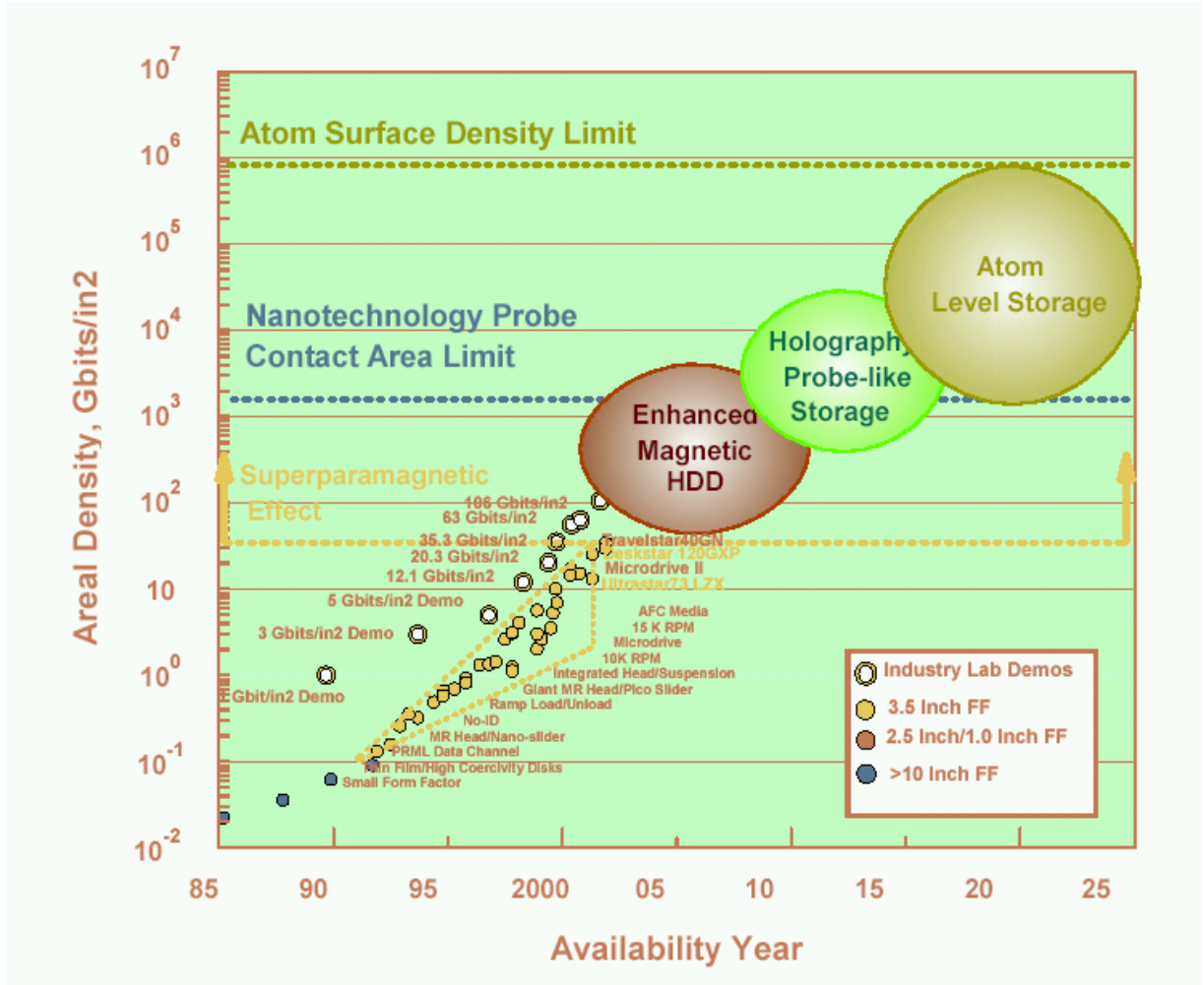
	RPM	Avg. Seek (ms)	1KB Random SIO/s (RPO)
2000 (75GB)	7200	8.5	137
2001	7200	7.8	146
2002	10K	7.0	171
2003	10K	6.3	187
2004	10K	5.6	205
2005 (1050GB)	15K	4.8	252

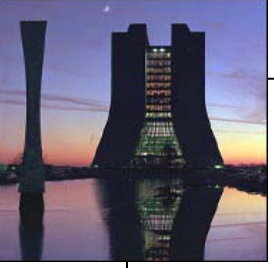
Mainstream Server

	RPM	Avg. Seek (ms)	1KB Random SIO/s (RPO)
2000 (36GB)	10K	4.9	226
2001	10K	4.5	244
2002	15K	4.1	283
2003	15K	3.6	317
2004	15K	3.2	345
2005 (500GB)	20K	2.8	408



Advanced Storage Roadmap

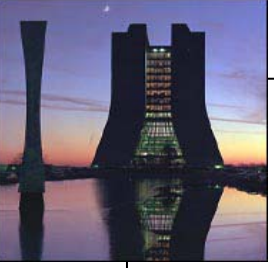




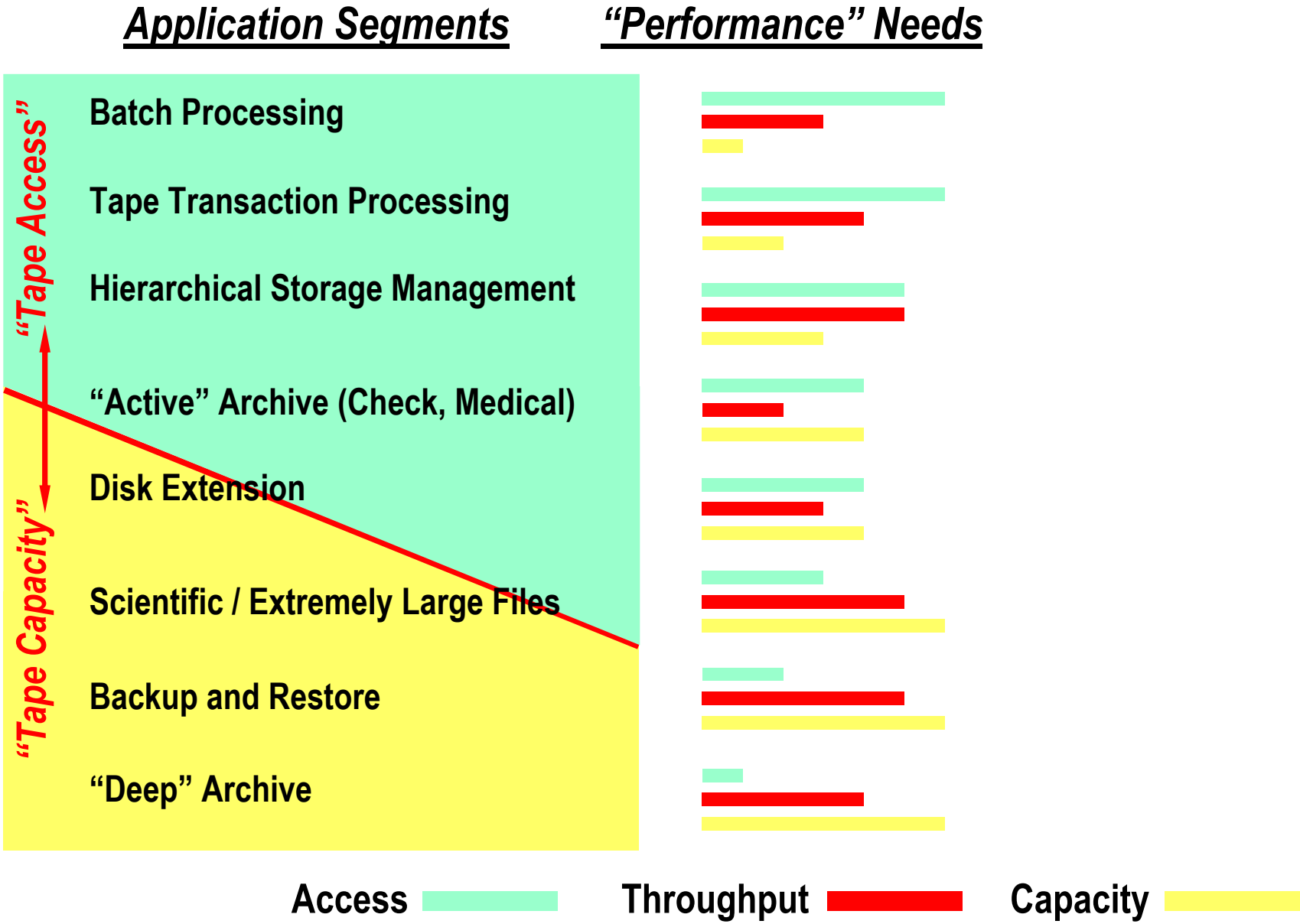
Disk Trends

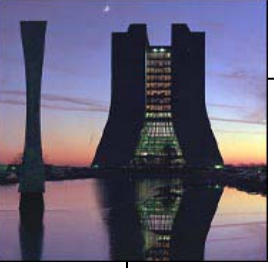
- SCSI still being developed, today at 160 MB/s with 320MB/s transfer speed announced.
- IDE developments
  - Disk Connections from parallel => Serial ATA (150MB/s – 600 MB/s)
  - Serial ATA is expected to dominate the commodity disk connectivity market by end 2003.
- Fiber channel products still expensive.
- DVD solutions still 2-3x as expensive as disks.
  - No industry experience managing large DVD libraries.


Tape Storage Technology

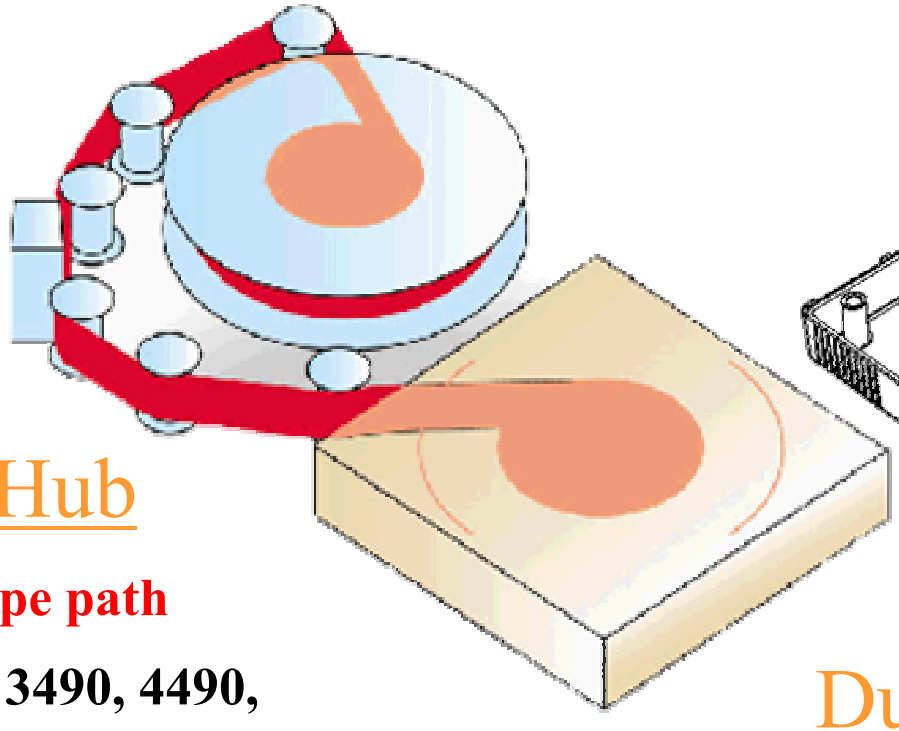


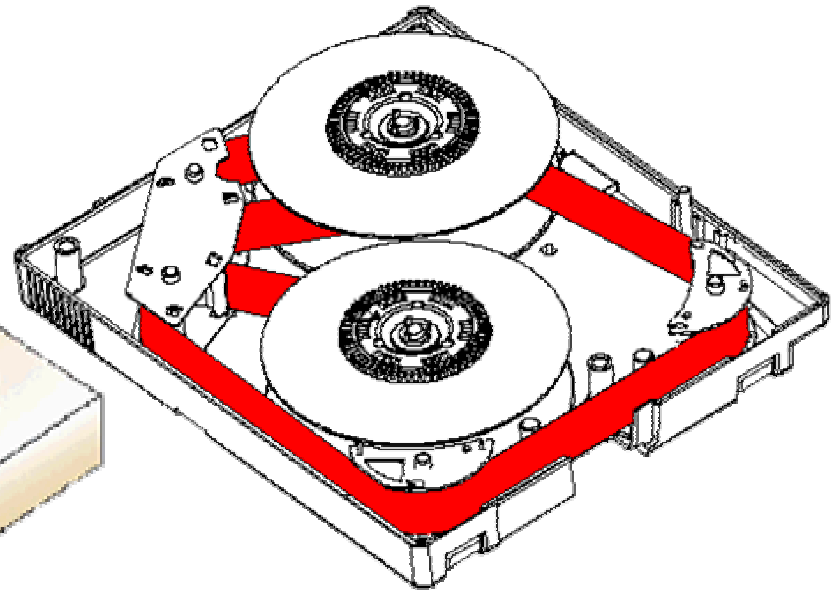
Tape Drive Target Applications





Tape Path and Cartridge Types

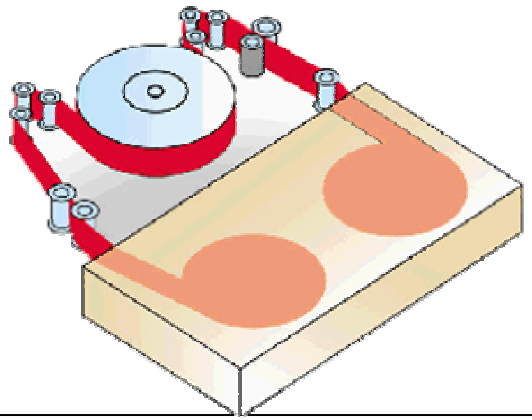




Single Hub

External tape path

DLT, LTO, 3490, 4490, 9490, 3590, SD-3, 9940



Dual Hub

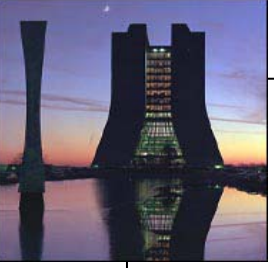
Internal tape path

STK 9840, IBM 3570, QIC

Dual Hub - Cassette

External tape path

AIT, Mammoth and other 4/8mm Helical scan



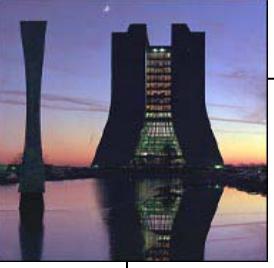
LTO Ultrium Roadmap

	Generation 1	Generation 2	Generation 3	Generation 4
	GA Start of Q3-00	18-24 Months after Gen1	18-24 Months after Gen2	
Capacity	100 GB	200 GB	400 GB	800 GB
Transfer Rate	10-20 MB/s	20-40 MB/s	40-80 MB/s	80-160 MB/s
Enabling Technology	?	?	?	?
Number of Channels	8	8	16	16
Recording Method	RLL 1,7	PRML	PRML	PRML
Media Type	MP2	MP	MP	Thin Film
Tape Length	580 m	580 m	800 m	800 m

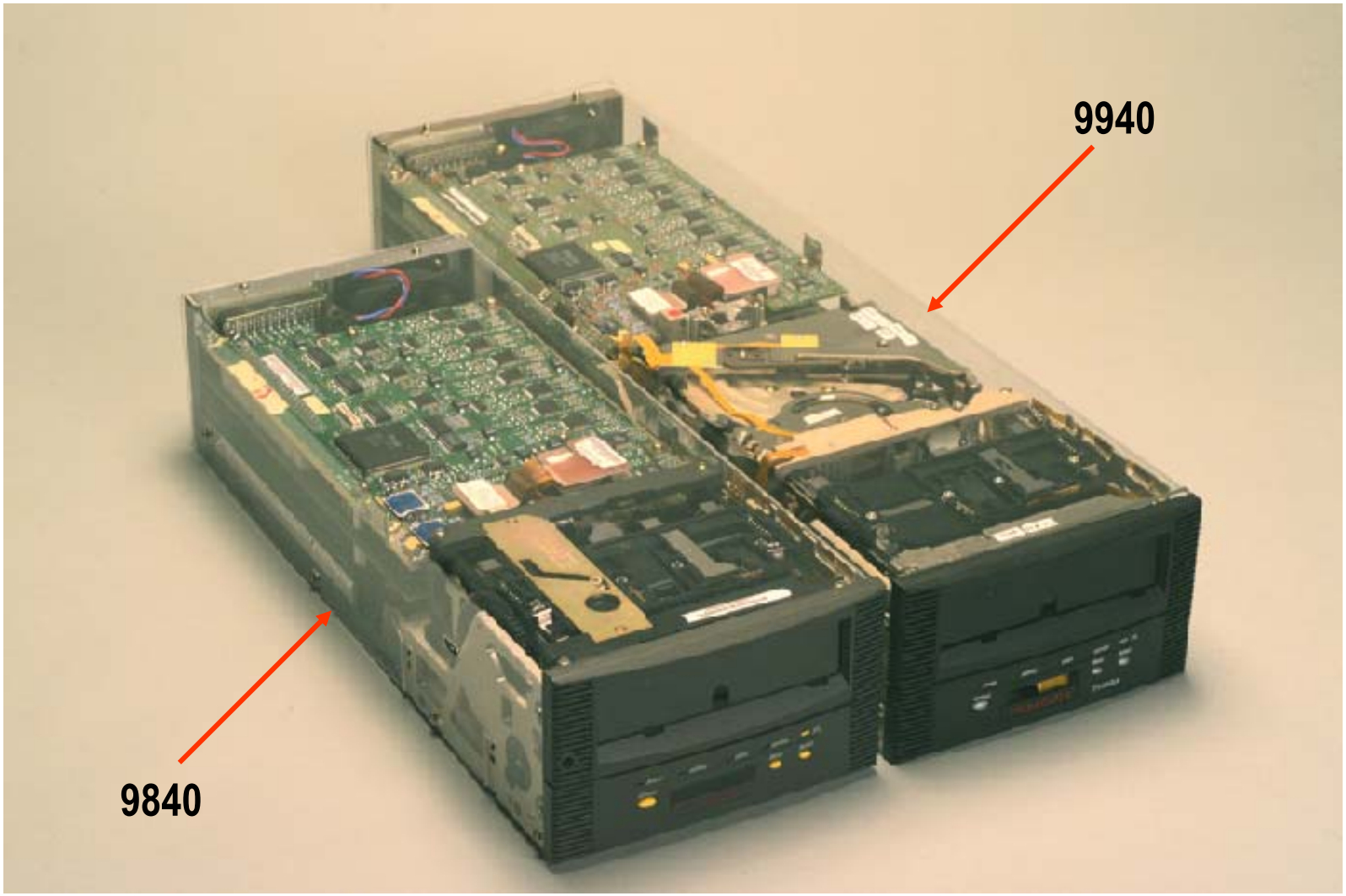
Media Swap

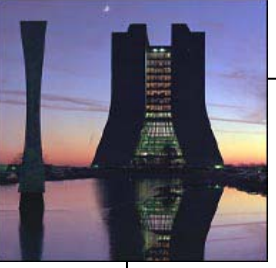
Media Swap

Media Swap

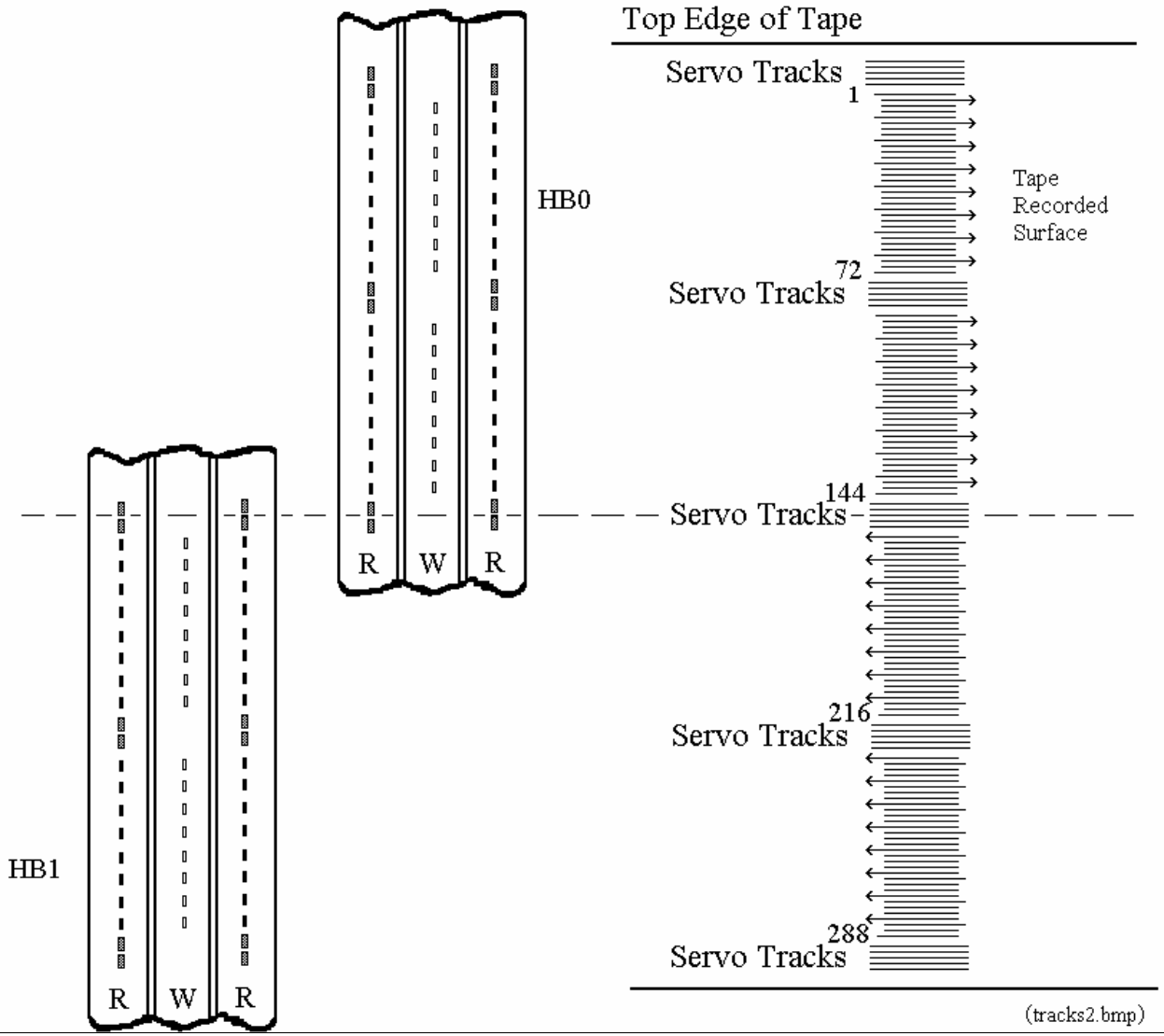


Recent Tape Drives (STK)

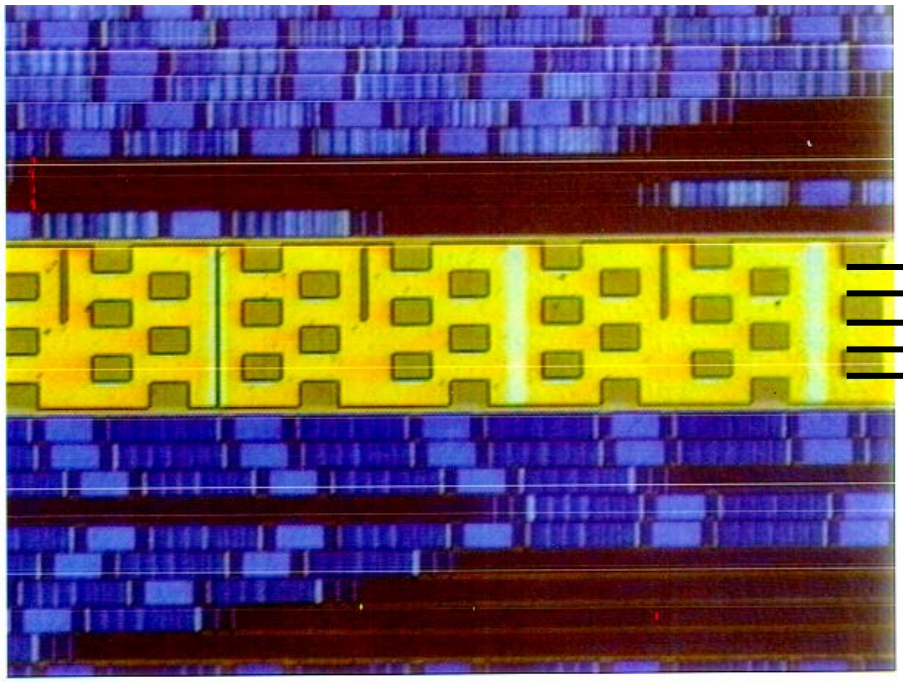




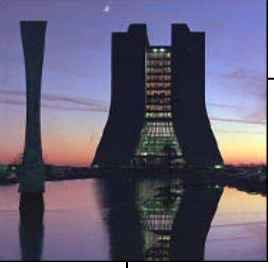
Track Layout (STK 9840/9940)



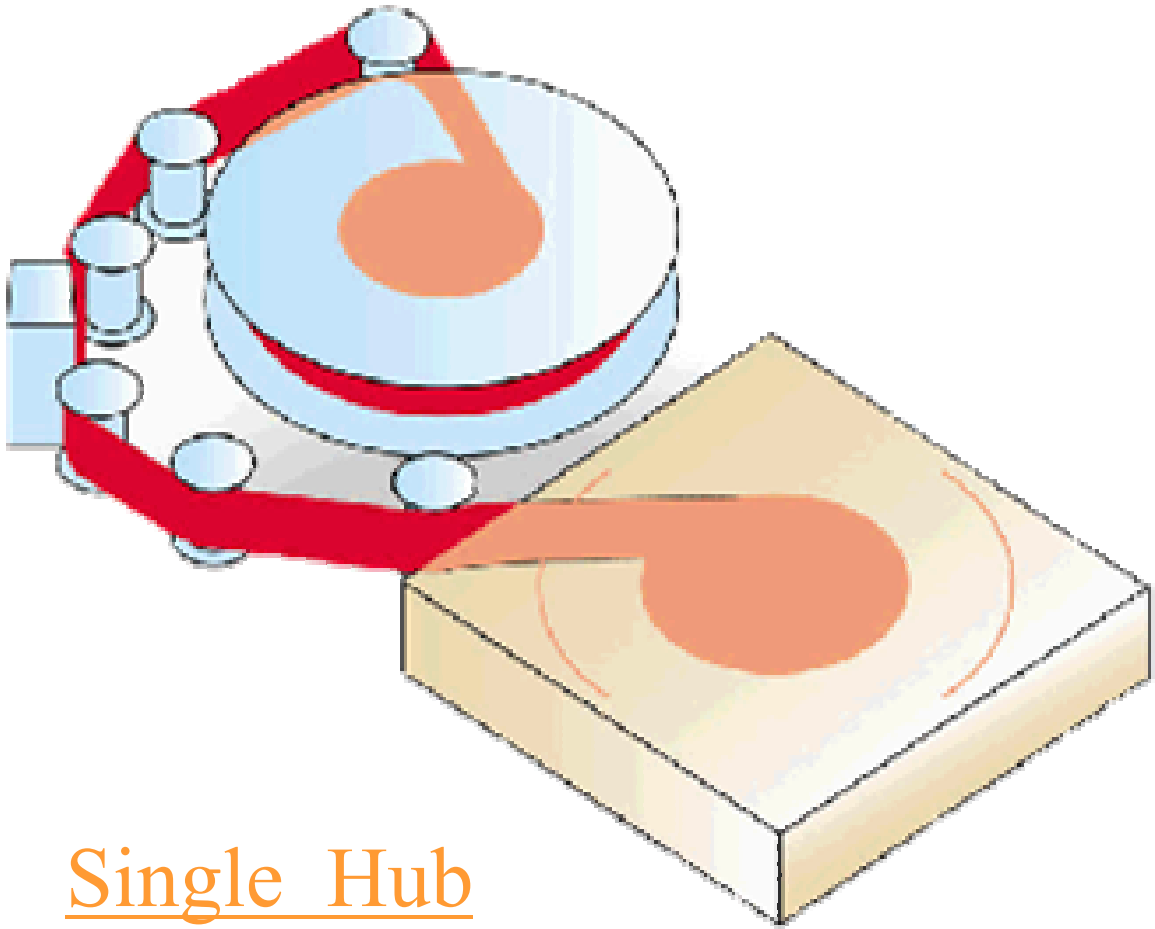
Ferrofluid/MFM Images of Data Tracks and Amplitude Based Servo Tracks (STK 9840/9940) Drives



6 servo tracks with bursts of servo signal. Servo head positions between tracks. 5 positions

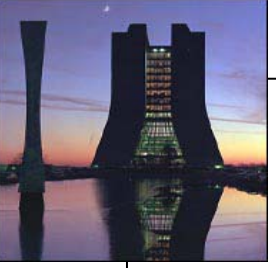


9940 Operational Principle



Single Hub

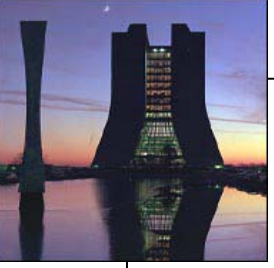
External tape path



Data Access Comparison

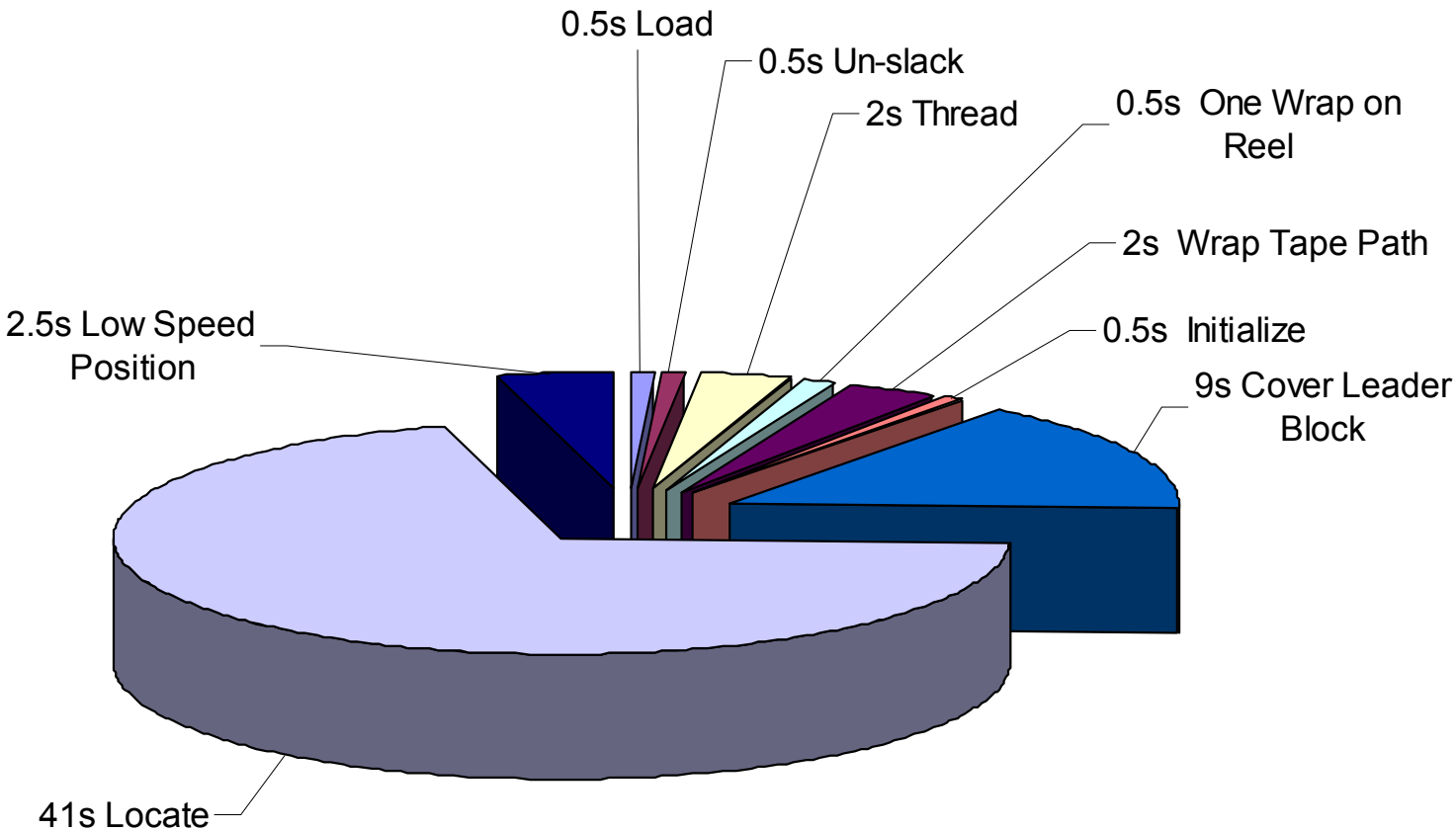
Data Access: Time from cartridge insertion in drive to first byte of data - timed to mid-point and does not include "*robotic time*"

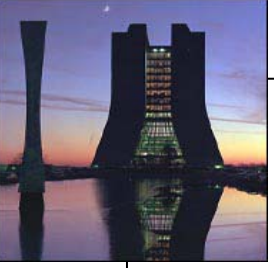
Access	*Seconds*
9840	12
Timberline	25
Capacity	*Seconds*
9940	59
RedWood	66
Magstar 3590	64
Magstar 3590E	90+
IBM LTO 3580	110



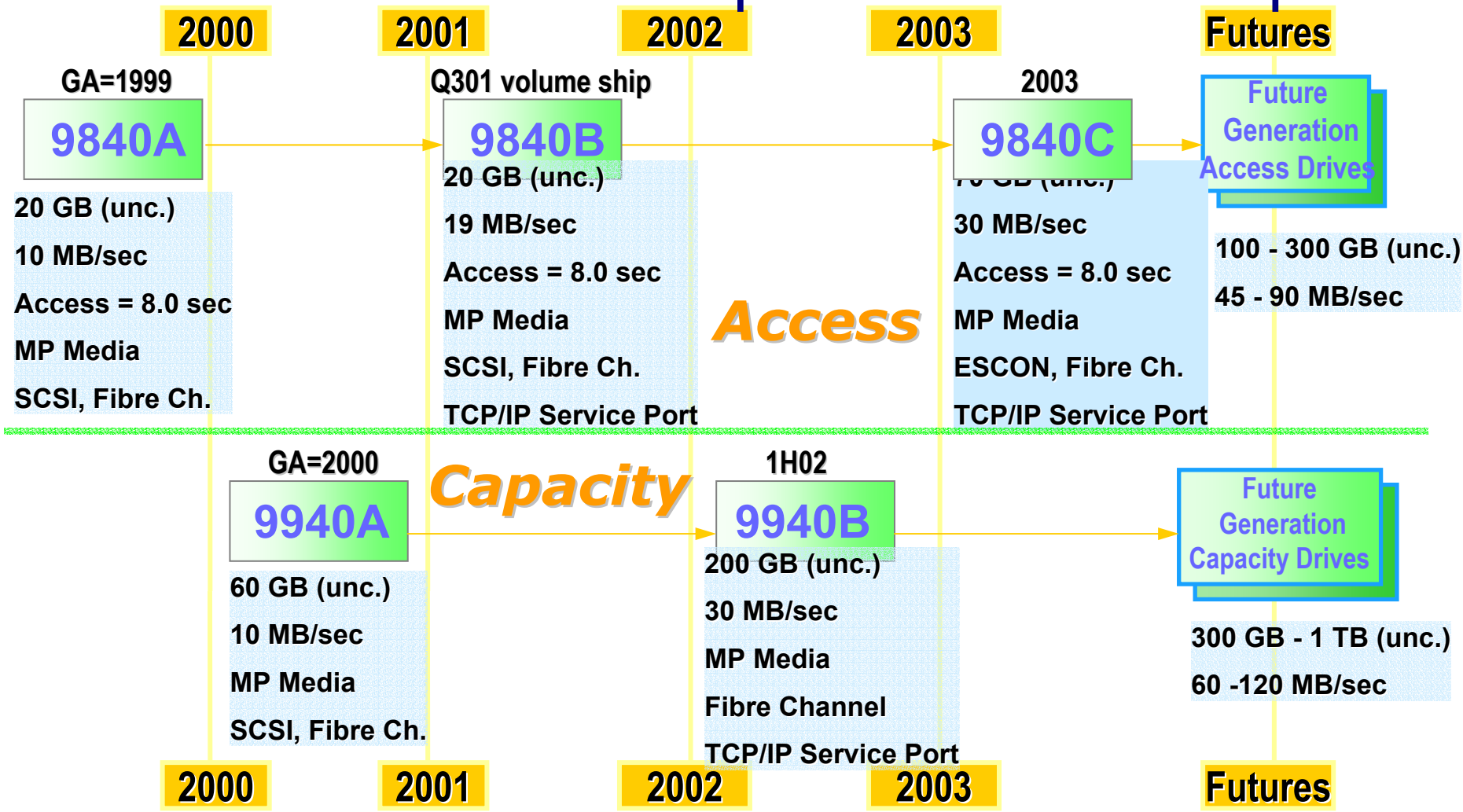
9940 Load Time

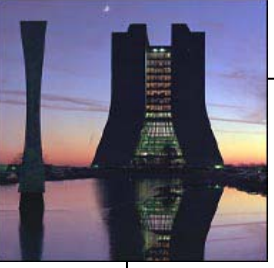
9940 Load Time To First Data (Total: 58.5 Seconds)





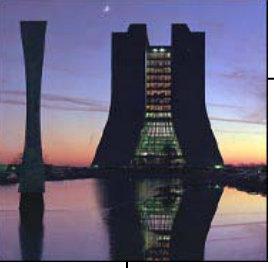
9840 *and* 9940 Tape Drive Roadmap

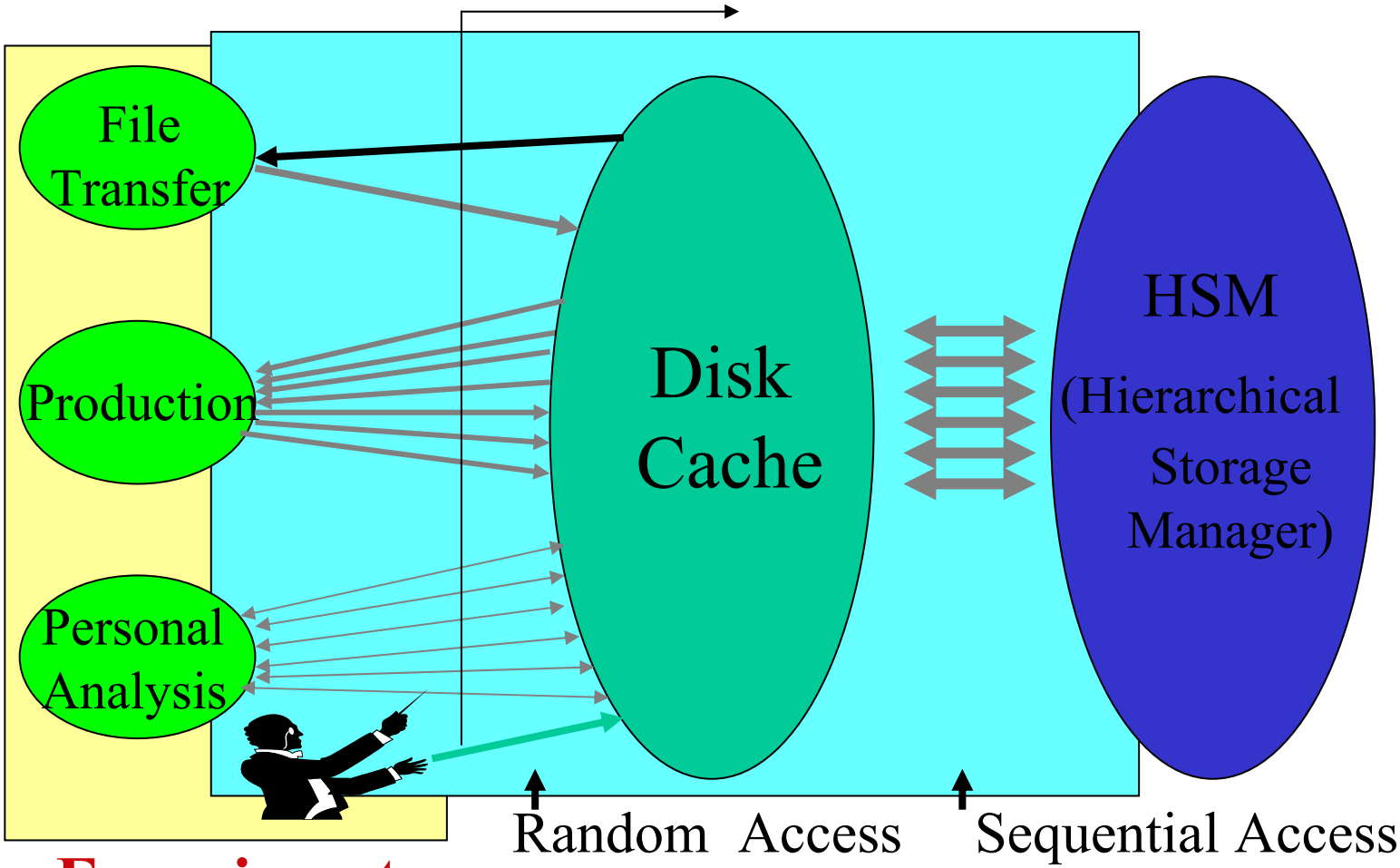




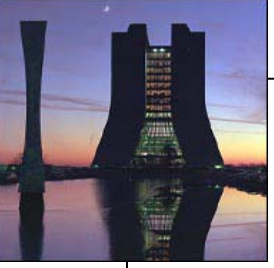
Tapes - 1

- Recent tape drive technology (9840, 9940A/B, LTO) is installed at CERN.
- Current Installation are 10 STK silo's capable of taking 800 new format tape drives. Today tape performance is 15MB/sec so theoretical aggregate is 12GB/sec (no way in reality!)
- Cartridge capacities expected to increase to 1TB before LHC startup but it's market demand and not technical limitations driving it.
- Using tapes as a random access device is no longer a viable option
  - Need to consider a much larger, persistent disk cache for LHC reducing tape activity for analysis.



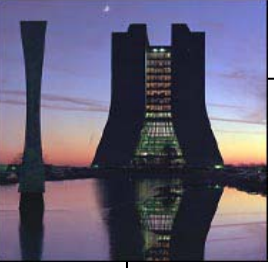


Experiment-specific



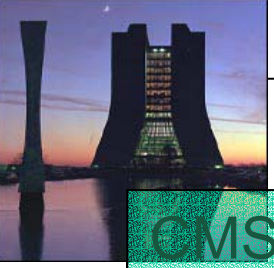
Tapes - 2

- Current costs are about $33/slot for a tape in the Powderhorn robot.
- Current tape cartridge (9940A/B, 60GB) costs $86 with a slow decrease over time.
- Media dominates the overall cost and a move to higher capacity cartridges and tape units sometimes require a complete media change.
  - Current storage costs 0.4-0.7 USD/GB in 2000 could drop to 0.2 USD/GB in 2005 but probably would require a complete media change.
- Conclusions: No major challenges for tapes for LHC startup but the architecture has to be such that “random access” is avoided



Networking

Interregional Connectivity is the key ….



CMS as an example …

CMS has adopted a distributed computing model to perform data analysis, event simulation, and event reconstruction in which two-thirds of the total computing resources are located at regional centers.

Experiment

Online System

~PBytes/sec

~100 MBytes/sec

Tier 0 +1

Offline Farm, CERN Computer Center

France Center

UK Center

Italy Center

US Center @ FNAL

Tier 1

Tier 2

Tier2 Center

Center

nter

Center

Center

~2.4 Gbits/sec

Tier 3

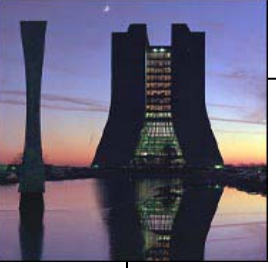
Institute

itute

stitute

nstitute

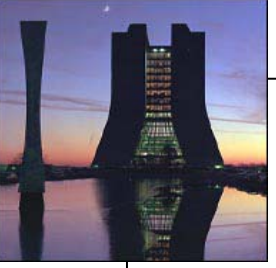
The unprecedented size of the LHC collaborations and complexity of the computing task requires that new approaches be developed to allow physicists spread globally to efficiently participate.



Transatlantic Net WG (HN, L. Price) Bandwidth Requirements [*]

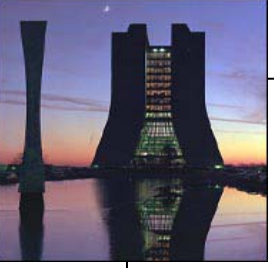
	2001	2002	2003	2004	2005	2006
CMS	100	200	300	600	800	2500
ATLAS	50	100	300	600	800	2500
BaBar	300	600	1100	1600	2300	3000
CDF	100	300	400	2000	3000	6000
D0	400	1600	2400	3200	6400	8000
BTeV	20	40	100	200	300	500
DESY	100	180	210	240	270	300
CERN BW	155-310	622	2500	5000	10000	20000

[*] Installed BW. Maximum Link Occupancy 50% Assumed



Network Progress and Issues for Major Experiments

- Network backbones are advancing rapidly to the 10 Gbps range
  - "Gbps" end-to-end throughput data flows will be in production soon (in 1-2 years)
- Network advances are changing the view of the net's roles
  - This is likely to have a profound impact on the experiments' Computing Models, and bandwidth requirements
- Advanced integrated applications, such as Data Grids, rely on seamless "transparent" operation of our LANs and WANs
  - With reliable, quantifiable (monitored), high performance
  - Networks need to be integral parts of the Grid(s) design
- Need new paradigms of real network and system monitoring, and of new of "managed global systems" for HENP analysis
  - These are starting to be developed for LHC

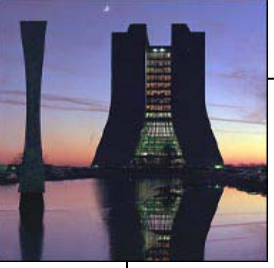


Signs of the Times: Uncertainty But No Change in Outlook

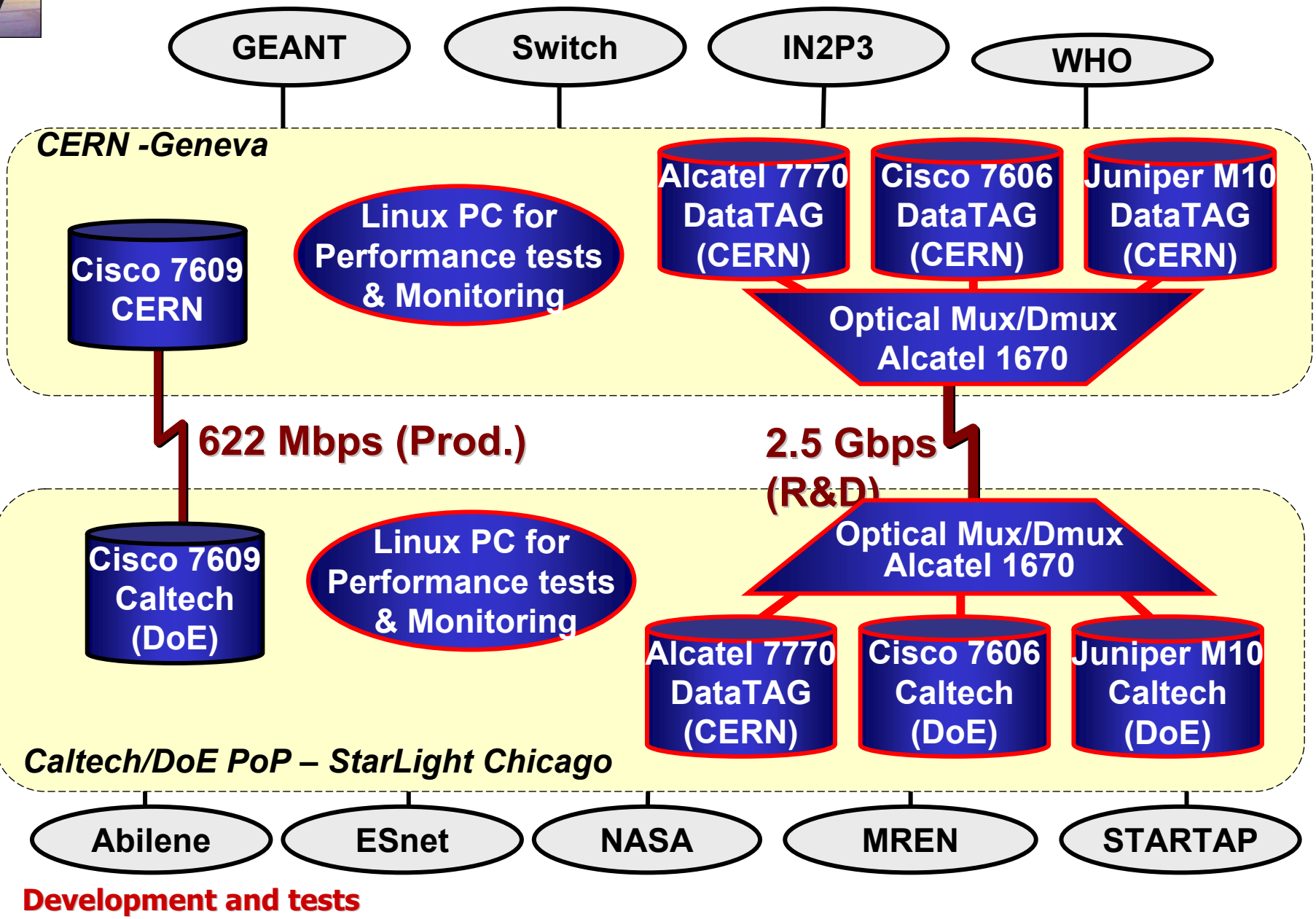
- Key Providers in Bankruptcy

  KPNQwest, Teleglobe, Global Crossing, FLAG; *Worldcom*

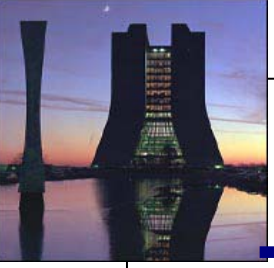
- Switching to Others, Where Needed and Possible
  - E.g. T-Systems (Deutsche Telecom) for US-CERN
- Strong Telecom Market Outlook
  - Good pricing from DT
  - MCI/Worldcom network will continue to operate (?): 20 M customers in US; UK academic & research network
  - Aggressive plans by major and startup network equipment providers
- Strong Outlook in R&E Nets for Rapid Progress
  - Abilene (US) Upgrade On Schedule; GEANT (Europe), and SuperSINET(Japan) Plans Continuing
  - ESNet Backbone Upgrade: 2.5 Gbps "Now"; 10 Gbps in 2 Yrs.
  - Regional Progress, and Visions; E.g. CALREN: *"1 Gbps to Every Californian by 2010"*



LHCnet Network : Late 2002



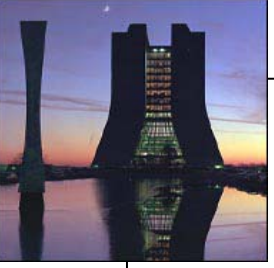
Development and tests



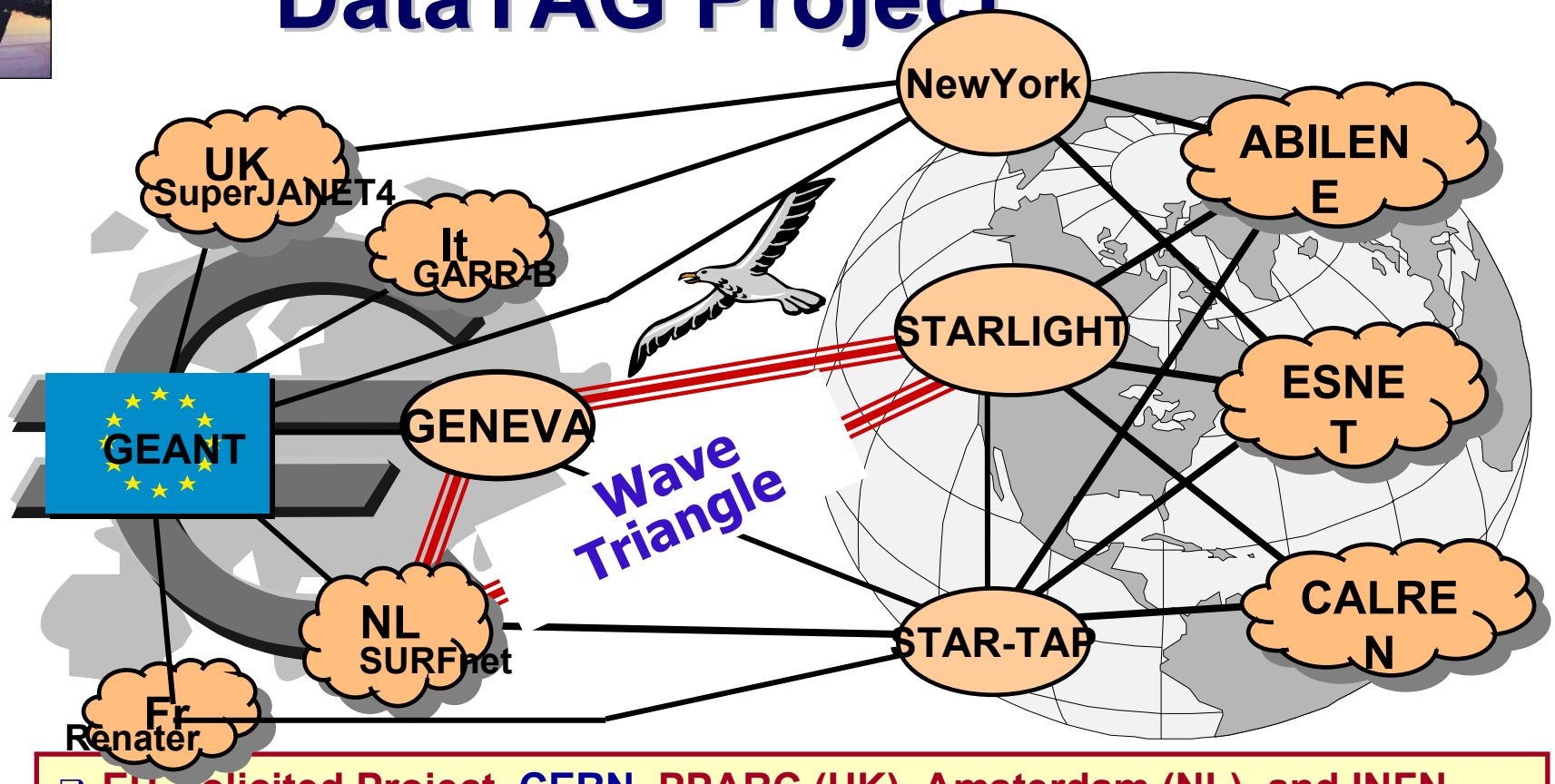
The Rapid Pace of Network Technology Advances Continues

Within the Next One to Two Years

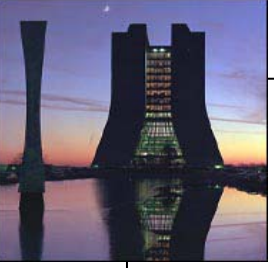
- 10 Gbps Ethernet on Switches and Servers; LAN/WAN integration at 10 Gbps
- 40 Gbps Wavelengths Being Shown
- HFR: 100 Mpps forwarding engines, 4 and more 10 Gbps ports per Slot; Terabit/sec backplanes etc.
- Broadband Wireless [Multiple 3G/4G alternatives]: the drive to defeat the last mile problem
  - 802.11 ab, UWB, etc.



DataTAG Project

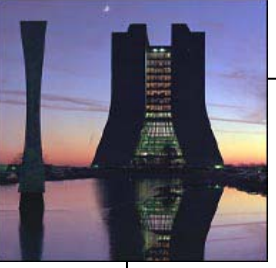


- EU-Solicited Project. CERN, PPARC (UK), Amsterdam (NL), and INFN (IT);
  and US (DOE/NSF: UIC, NWU and Caltech) partners
- Main Aims:
  - Ensure maximum interoperability between US and EU Grid Projects
  - Transatlantic Testbed for advanced network research
- 2.5 Gbps Wavelength Triangle 7/02 (10 Gbps Triangle in 2003)



HENP Major Links: Bandwidth Roadmap (Scenario) in Gbps

Year	Production	Experimental	Remarks
2001	0.155	0.622-2.5	SONET/SDH
2002	0.622	2.5	SONET/SDH DWDM; GigE Integ.
2003	2.5	10	DWDM; 1 + 10 GigE Integration
2005	10	2-4 X 10	λ Switch; λ Provisioning
2007	2-4 X 10	~10 X 10; 40 Gbps	1st Gen. λ Grids
2009	~10 X 10 or 1-2 X 40	~5 X 40 or ~20-50 X 10	40 Gbps λ Switching
2011	~5 X 40 or ~20 X 10	~25 X 40 or ~100 X 10	2nd Gen λ Grids Terabit Networks
2013	~Terabit	~MultiTerabit	~Fill One Fiber

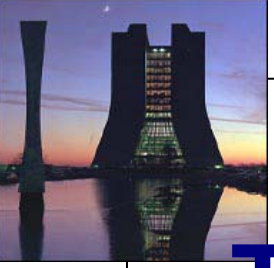


HENP Lambda Grids: Fibers for Physics

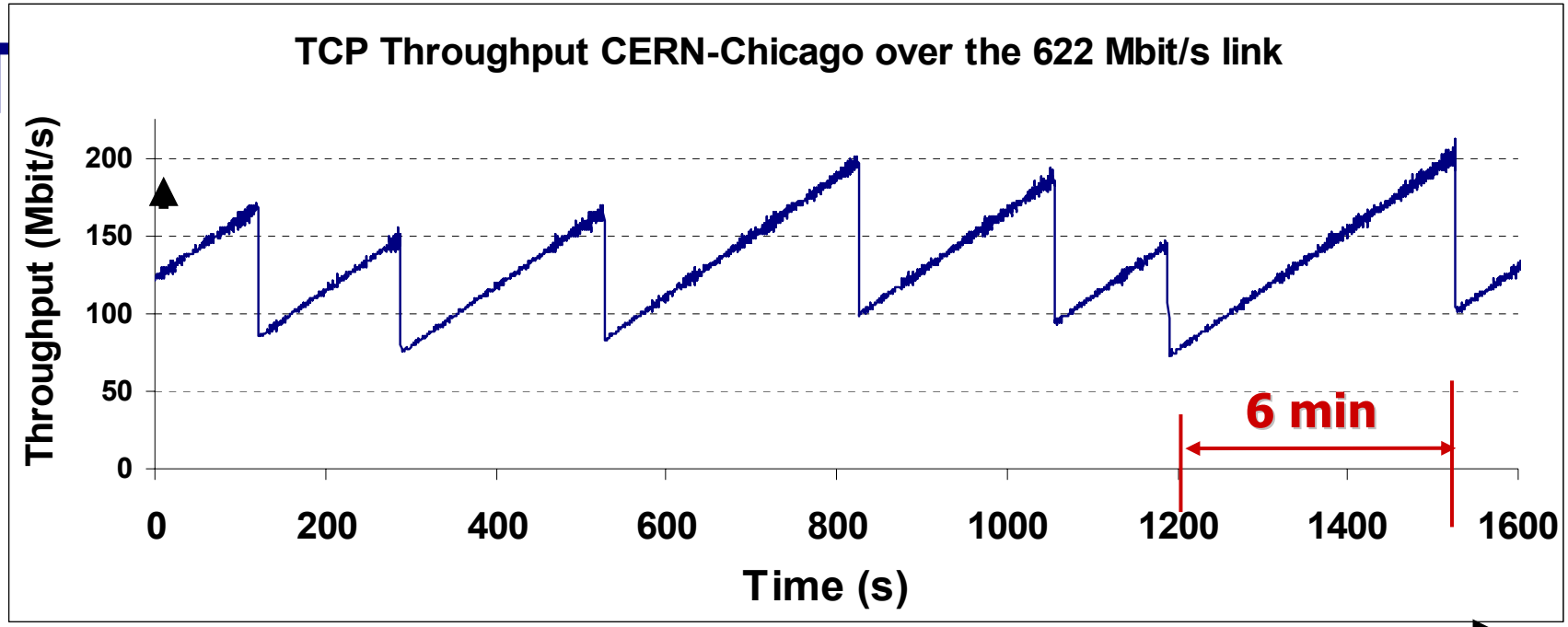
- Problem: Extract "Small" Data Subsets of 1 to 100 Terabytes from 1 to 1000 Petabyte Data Stores
- Survivability of the HENP Global Grid System, with hundreds of such transactions per day (circa 2007) requires that each transaction be completed in a relatively short time.
- Example: Take 800 secs to complete the transaction. Then
-

Transaction Size (TB)	Net Throughput (Gbps)
1	10
10	100
100	1000 (Capacity of Fiber Today)

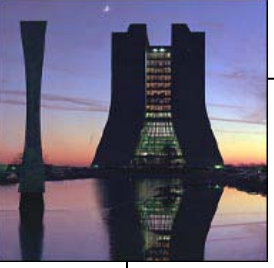
- Summary: Providing Switching of 10 Gbps wavelengths within ~3 years; and Terabit Switching within ~6-10 years would enable "Petascale Grids with Terabyte transactions" within this decade, as required to fully realize the discovery potential of major HENP programs, as well as other data-intensive fields.



Network Protocol Issues

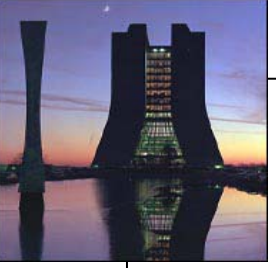


- TCP reactivity: Due to the Basic Multiplicative-Decrease Additive-Increase Algorithm to Handle Packet Loss
  - Time to increase the throughput by 120 Mbit/s is larger than 6 min for a connection between Chicago and CERN.
- A single loss is disastrous
  - A TCP connection reduces bandwidth use by half after a loss is detected (Multiplicative decrease)
  - A TCP connection increases slowly its bandwidth use (Additive increase)
  - TCP is much more sensitive to packet loss in WANs than in LANs

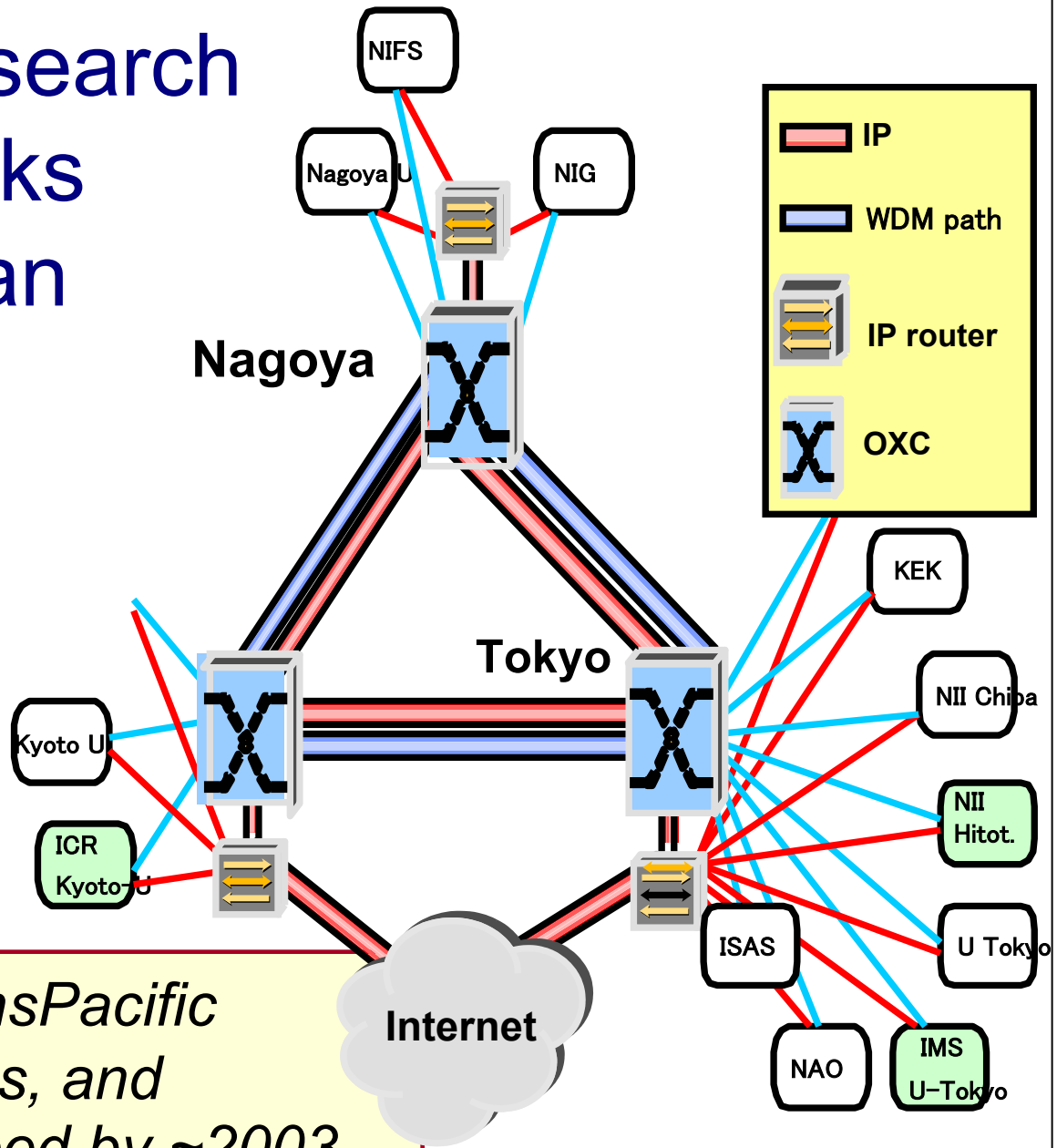


TCP Responsiveness

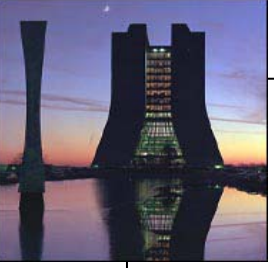
Case	Capacity	RTT (ms)	MSS (Byte)	Responsiveness
Typical LAN in 1988	10 Mbps	[2 ; 20]	1460	[1.5 ms ; 154 ms]
Typical WAN in 1988	9.6 Kbps	40	1460	0.006 sec
Typical LAN today	100 Mbps	5 (worst case)	1460	0.096 sec
Current WAN link CERN – Starlight	622 Mbps	120	1460	6 minutes
Future WAN link CERN – Starlight	10 Gbit/s	120	1460	92 minutes
Future WAN link CERN – Starlight	10 Gbit/s	120	8960 (Jumbo Frame)	15 minutes



National Research Networks in Japan

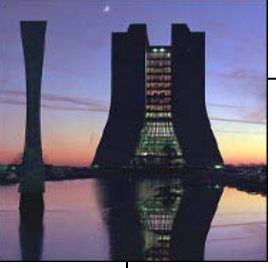


Proposal: Two TransPacific 2.5 Gbps Wavelengths, and Japan-CERN Grid Testbed by ~2003



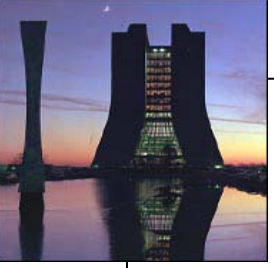
Networking

- Major cost reductions have taken place in wide-area bandwidth costs.
  - 2.5 Gbps common for providers but not academic in 1999. Now, 10Gbps common for providers and 2.5Gbps common for academic.
- Wide area data migration/replication now feasible and affordable.
  - Tests of multiple streams to the US running at the full capacity of 2Gbps were successful.
- Local Area Networking moving to 10 Gbps and this is expected to increase. First10Gbps NIC's available for end systems.



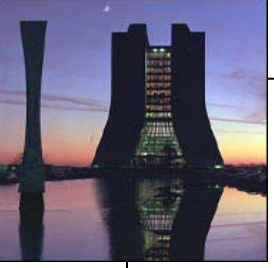
Networking Trends

- Transitioning from 10Gbit to 20-30 Gbit seems likely.
- MPLS (Multiprotocol Label Switching) has gained momentum. It provides secure VPN capability over public networks. A possibility for tier-1 center connectivity.
- Lambda networks based on dark fiber are also becoming very popular. It is a “build-yourself” network and may also be relevant for the grid and center connectivity.

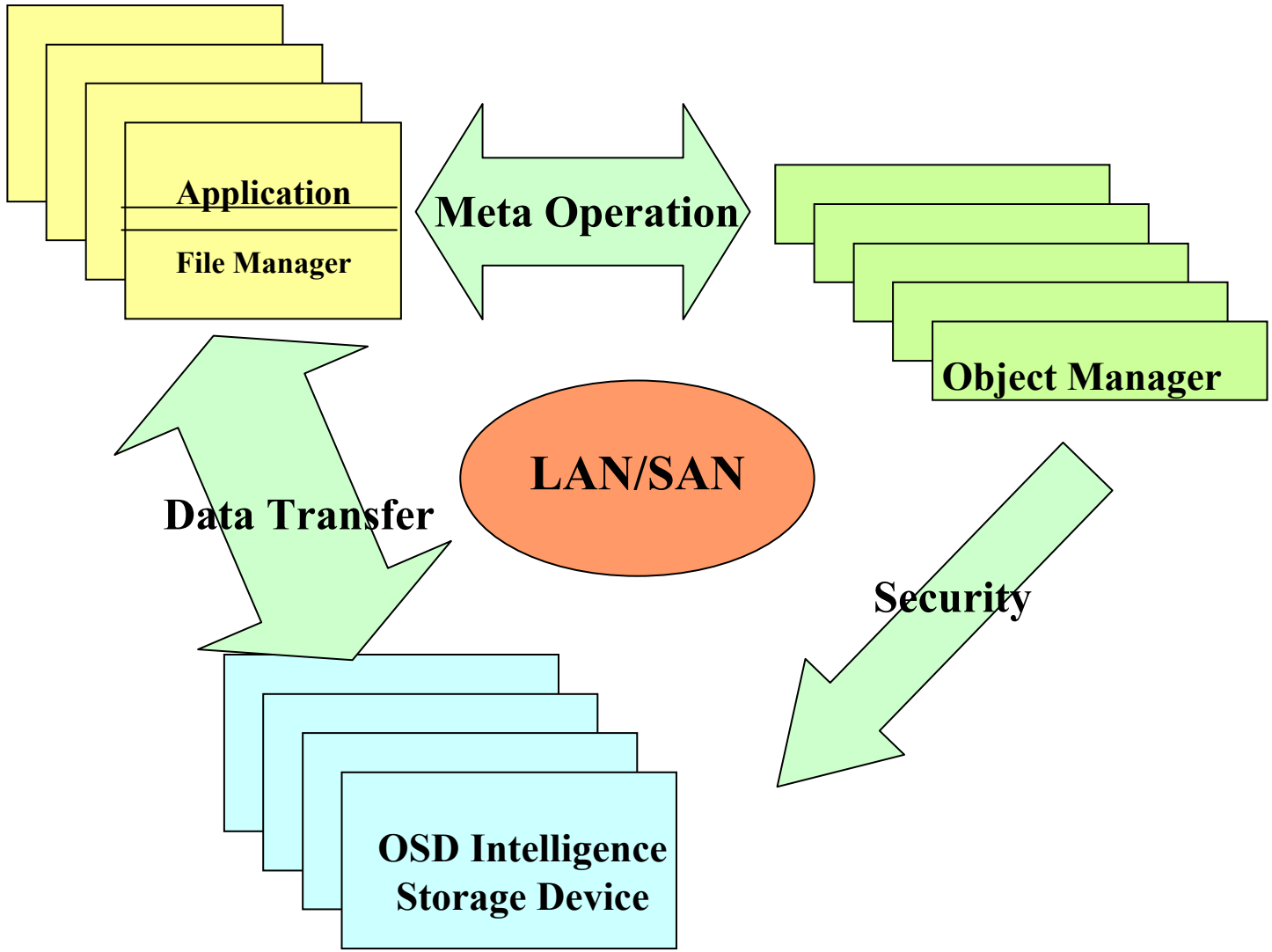


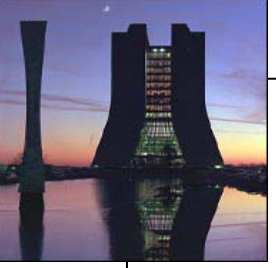
Storage - Architecture

- Possibly the biggest challenge for LHC
  - Storage architecture design (seamless integration from CPU caches to deep archive required)
  - Data management. Currently very poor tools and facilities for managing data and storage systems.
- SAN vs. NAS debate still alive
  - SAN, scalable and high availability, but costly
  - NAS, cheaper and easier to manage
- Object storage technologies appearing
  - Intelligent storage system able to manage the objects it is storing
  - Allowing “light-weight” Filesystems



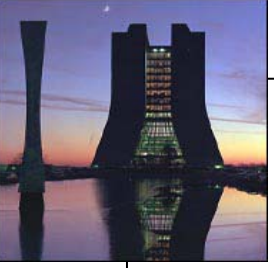
Object Storage Device Architecture





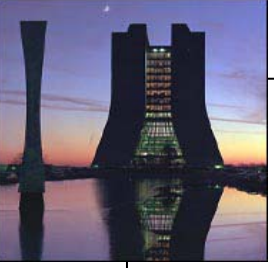
Storage Management

- Very little movement in the HSM space since the last PASTA report.
  - HPSS still for large scale systems
  - A number of mid-range products (make tape look like a big disk) but limited scaling possible
- HEP still a leader in tape and data management
  - CASTOR, Enstore, JASMine
  - Will remain crucial technologies for LHC.
- Cluster file systems appearing (StorageTank, Lustre)
  - Provide “unlimited” (PB) file system (e.g. through LAN, SAN)
  - Scale to many 1000’s of clients (CPU servers).
  - Need to be interfaced to tertiary storage systems (e.g. Enstore)

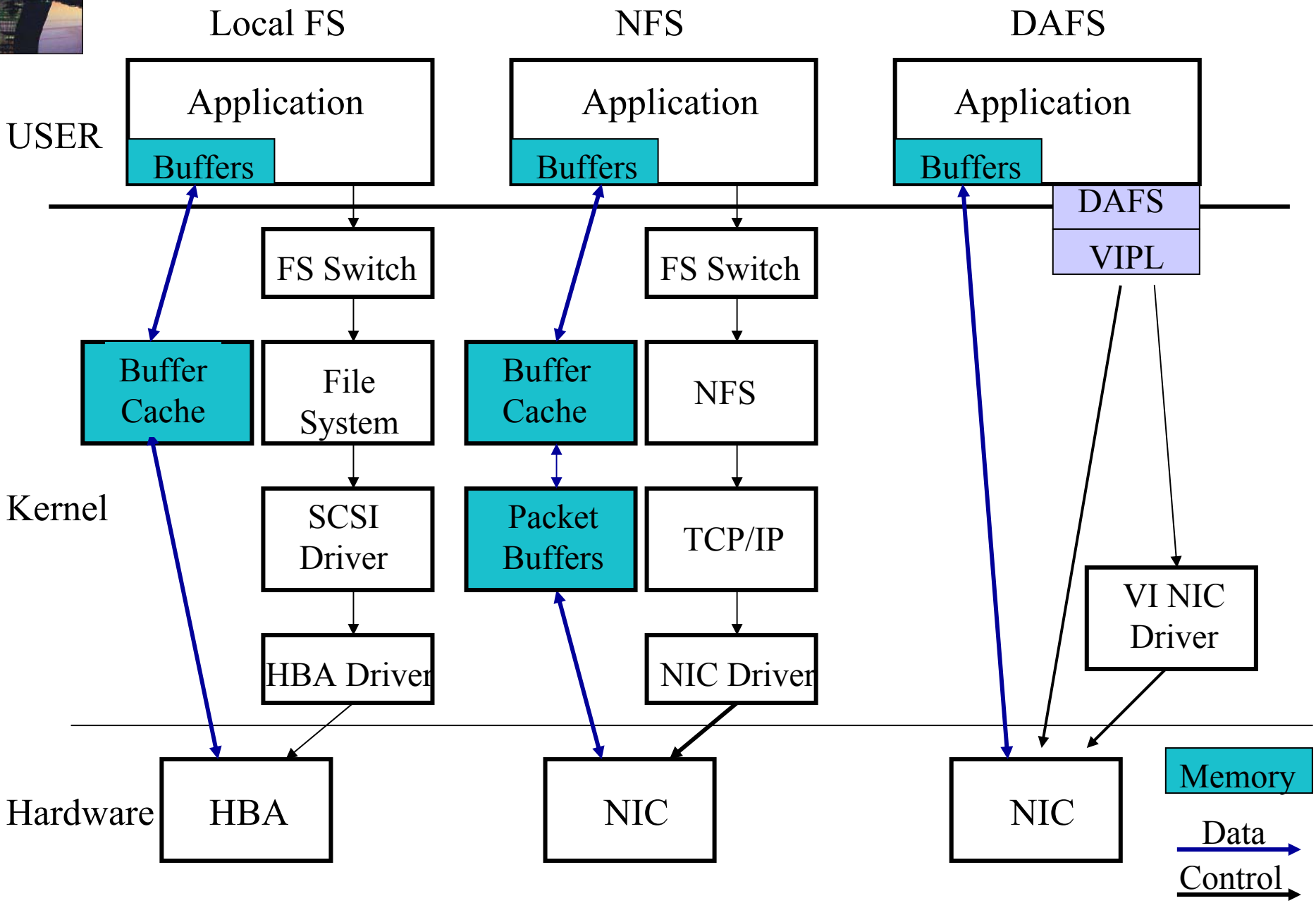


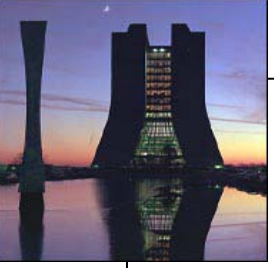
Storage - Connectivity

- FiberChannel market growing at 36%/year from now to 2006 (Gartner). This is the current technology for SAN implementation.
- iSCSI or equivalent over Gigabit Ethernet is an alternative (and cheaper) but less performant implementation of SAN gaining in popularity.
  - It is expected that GigE will become a popular transport for storage networks.
- InfiniBand (up to 30 Gbps) is a full-fledged network technology that could change the landscape of cluster architectures and has much, but varying, industry support.
  - Broad adoption could drive costs down significantly
  - FIO (Compaq, IBM, HP) and NGIO (Intel, MS, Sun) merged to IB
  - Expect bridges between IB and legacy Ethernet and FC nets
  - Uses IPv6
  - Supports RDMA and multicast



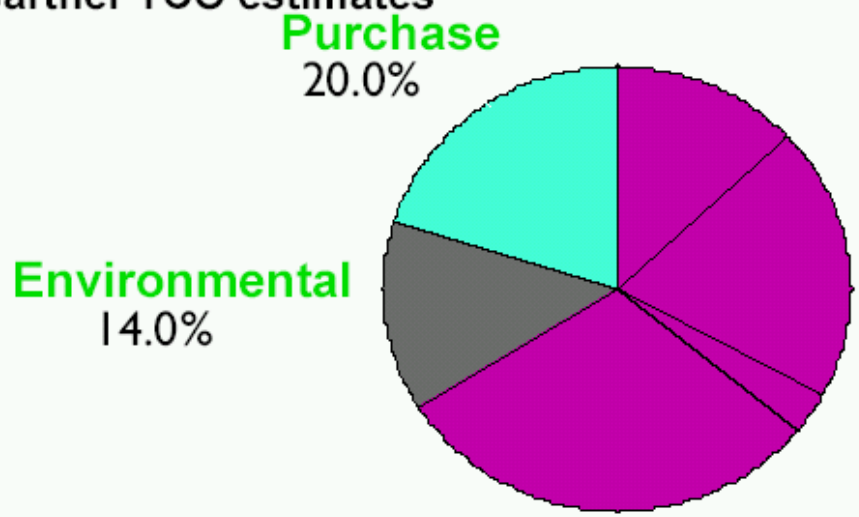
Data Transfer Overhead





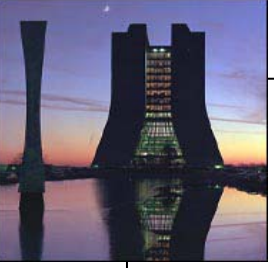
Storage Cost

- Gartner TCO estimates

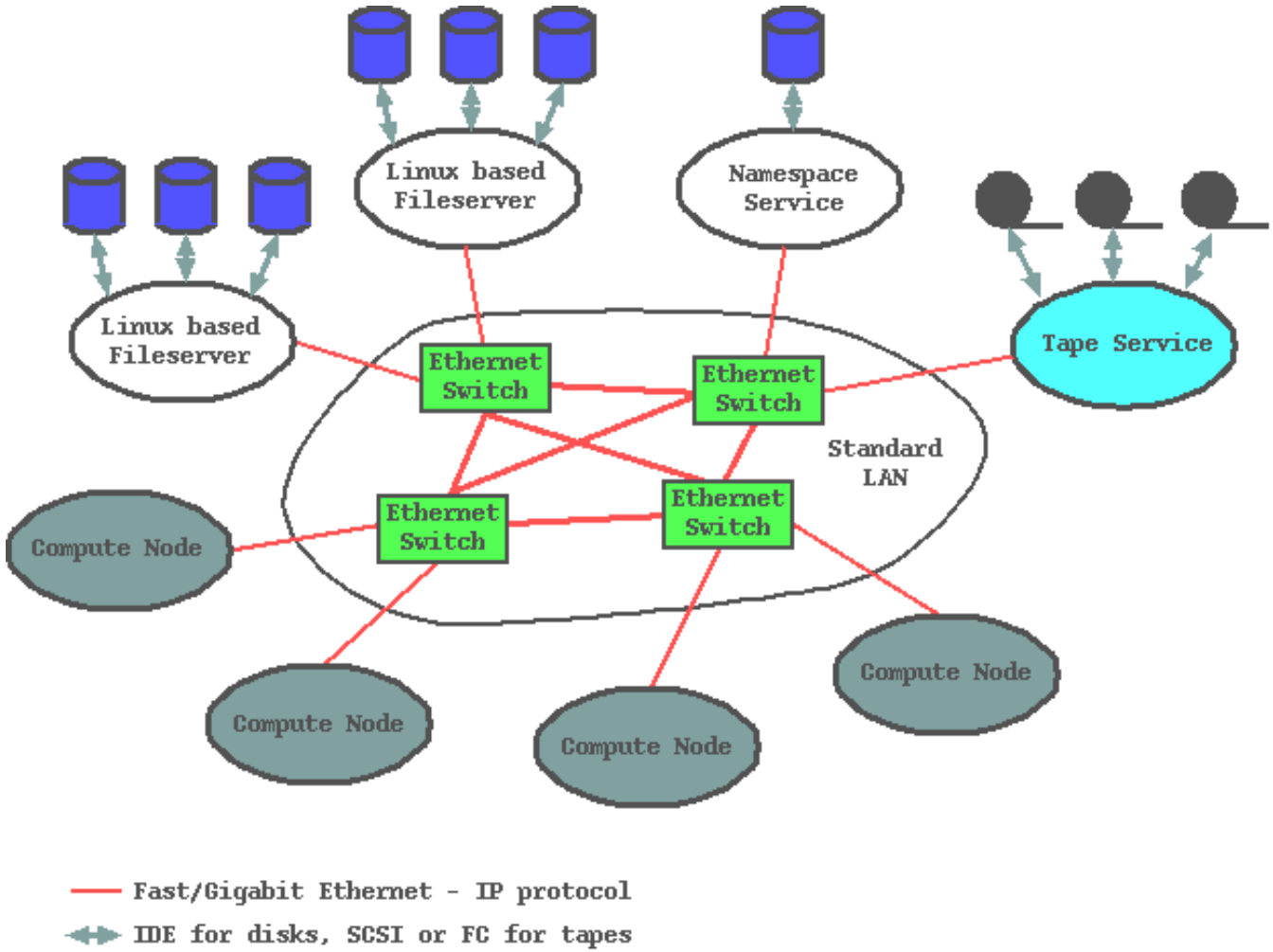


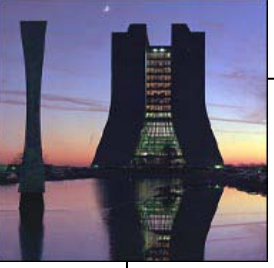
Source: Gartner Research, 2000

Cost of managing storage and data are the predominate costs

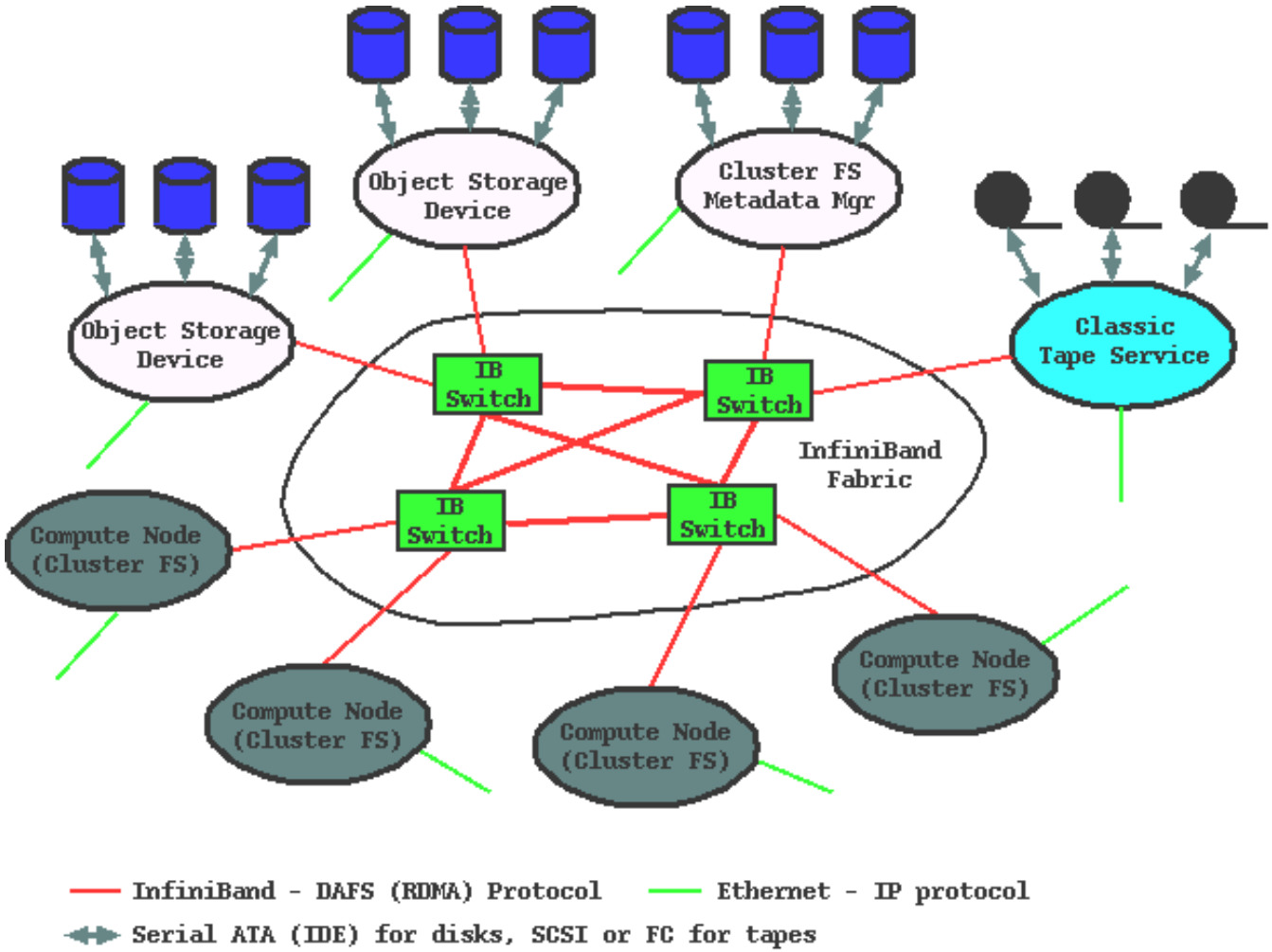


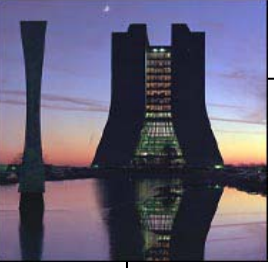
Storage Scenario - Today





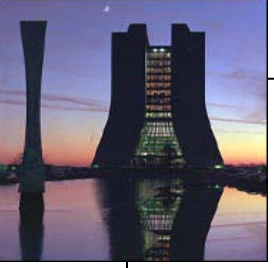
Storage Scenario - Future



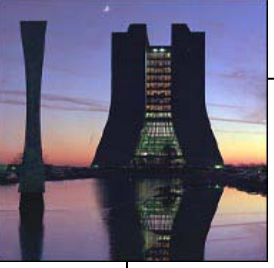


Some Overall Conclusions

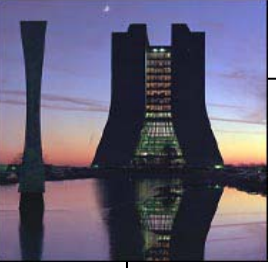
- Tape and Network trends match or exceed our initial needs.
  - Need to continue to leverage economies of scale to drive down long term costs.
- CPU trends need to be carefully interpreted
  - The need for new performance measures are indicated.
  - Change in the desktop market might effect the server strategy.
  - Cost of manageability is an issue.
- Disk trends continue to make a large (multi PB) disk cache technically feasible, but ….
  - The true cost of such an object a bit unclear, given the issues of reliability, manageability and the disk fabric chosen (NAS/SAN, iSCSI/FC etc etc)
  - File system access for a large disk cache (RFIO, DAFS, …) under investigation (urgent !)
- More architectural work is needed in the next 2 years for the processing and handling of LHC data.
  - NAS/SAN models are converging, many options for system interconnects, new High Performance NAS products are (about to be) rolled out (Zambeel, Panasas, Maximum Throughput, Exanet etc)



… Sounds like we are in pretty good shape …..



… but let's be careful …



PASTA has addressed issues exclusively on the Fabric level

- It is likely that we will get the required technology (Processors, Memory, Secondary and Tertiary Storage Devices, Networking, Basic Storage Management)
- Missing: Solutions allowing truly distributed Computing on a Global Scale

  Will the Grid Projects meet our Expectations (in time) ?