The Premier Advanced Recording Technology Forum

THIC Inc.

HITACHI DATA SYSTEMS

Disk Technologies in High-End Storage Arrays

Roberto Basilio
Hitachi Data Systems
750 Central Expressway, Santa Clara, CA, 95050-2627
Phone: +1-408-970-5990 FAX: +1-408-562-5477
E-mail: Roberto.Basilio@hds.com

Presented at the THIC Meeting at the Sony Auditorium, 3300 Zanker Rd, San Jose CA 95134-1940

April 19-20,2005

Topics

- Introduction
- Hard Disk Drives
  - Evolution and directions
- Storage Systems
  - Architectures and Deployment
- What’s next
  - New technologies and deployment in large enterprise

Hitachi, Ltd. Worldwide Storage Group

Two Powerful Storage Businesses:

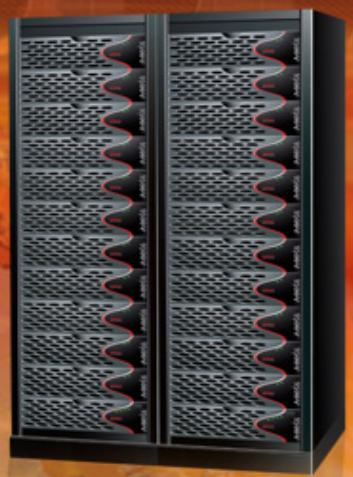
- HDD Business/ Hitachi Global Storage Technologies
- Storage Systems & Solutions Business Hitachi,Ltd. & Hitachi Data Systems

Leveraging Storage Technology Leadership:

- Enterprise Storage Platforms, Midrange Storage Systems, HDDs, Storage Management Software, Consulting Services

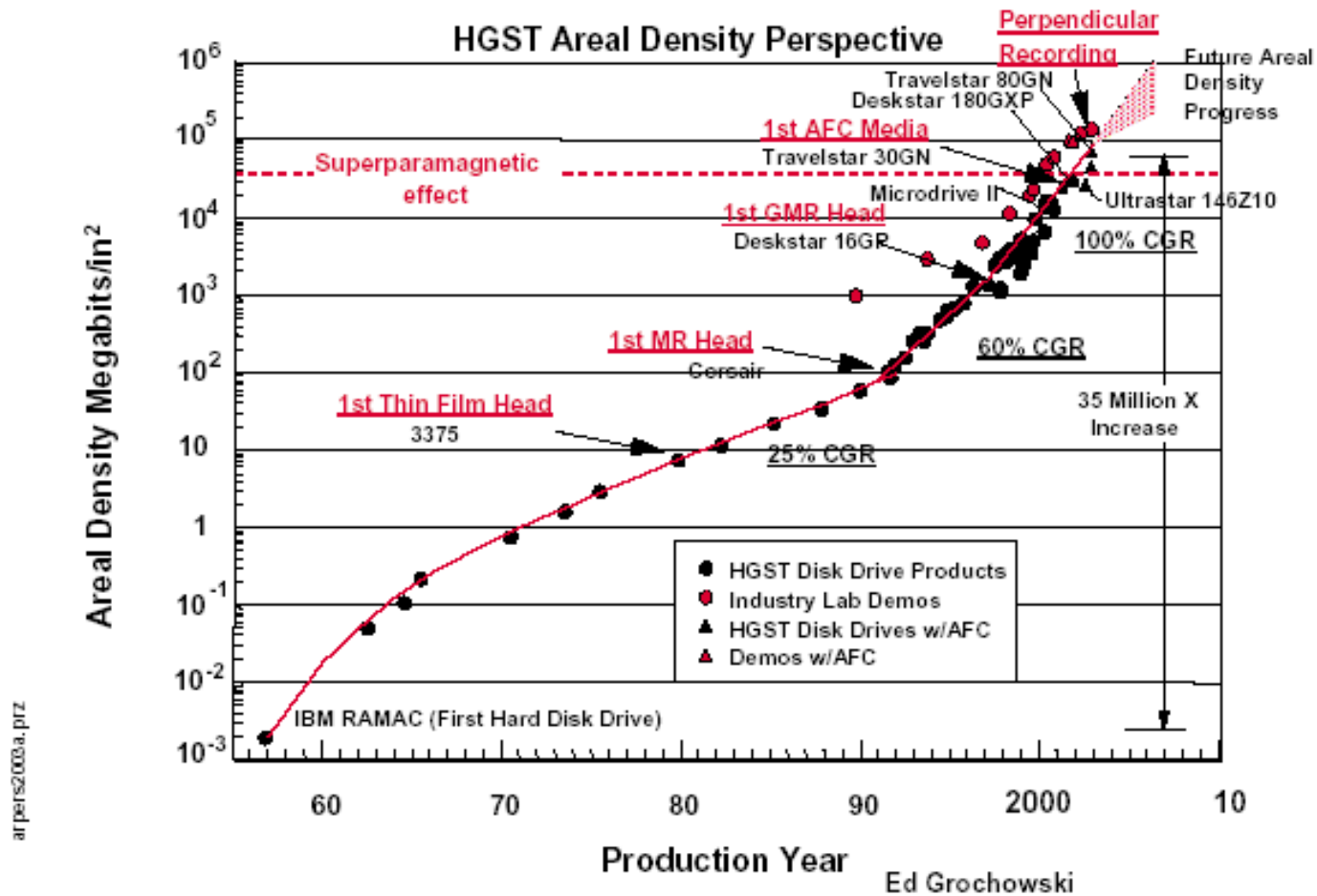
Expanding and Investing in Emerging Markets:

- Investing over $500M annually
- Compliance Solutions, Rich Media, Mobile Music, Game Players, “Terabyte Home”



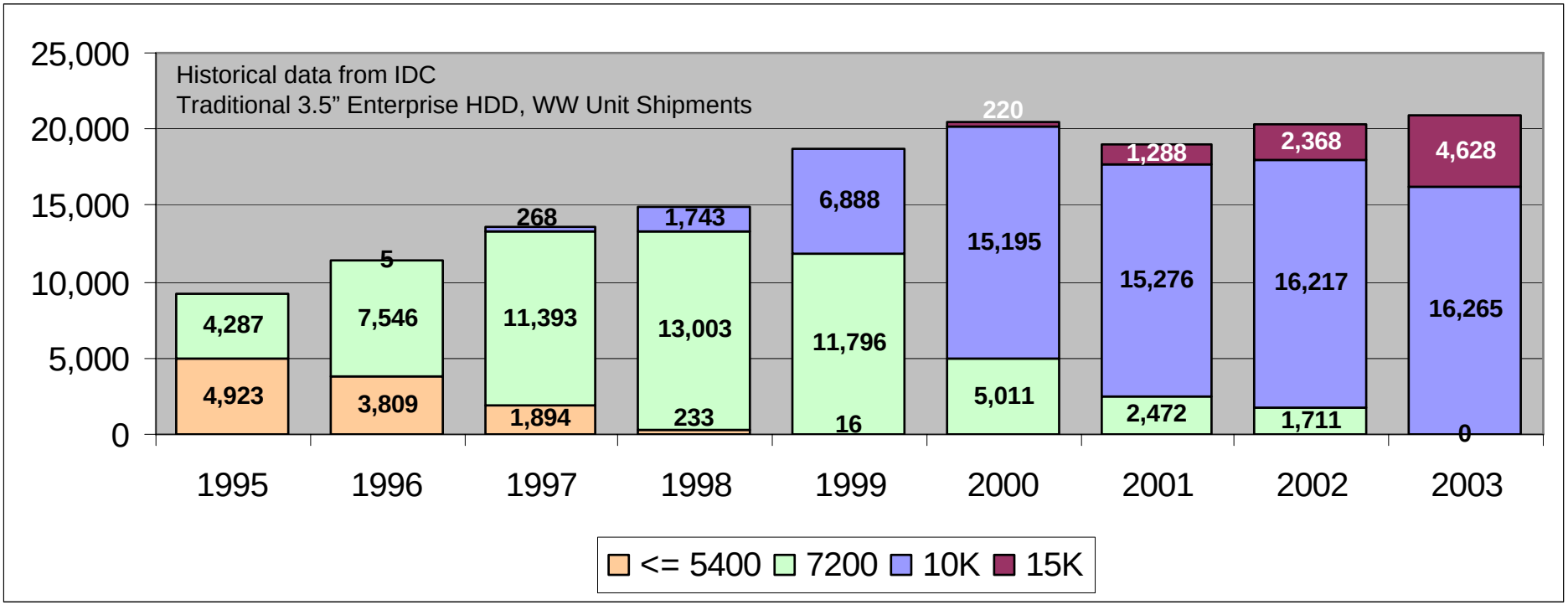
Disk Evolution

Disk Recording Densities Continue to Grow



Ed Grochowski

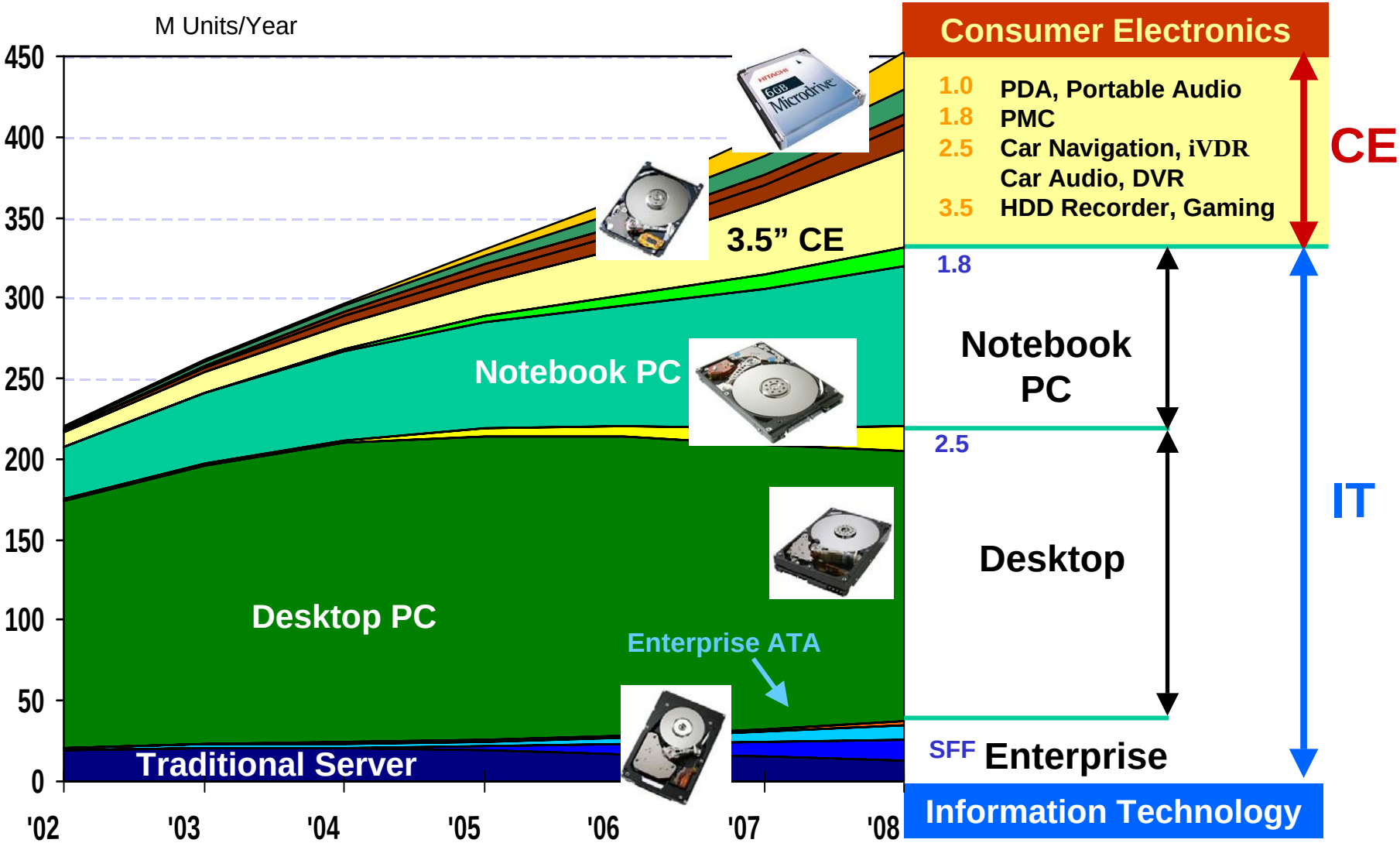
Server HDD Transition – RPM History



RPM Transitions typically quick after new Dominator

- 7200 RPM eliminated 5400 RPM 2 years after it reached >50%
  - 7200 RPM was very quick to dominate the market after only 2+ years
- 10K RPM domination in 2000 led to 7200 RPM extinction 2 years later
  - 10K RPM took approx. 3+ years to become dominate
- 15K RPM has now been in the market 5 years.
  - Adoption slowed during the 2001 economic downturn
  - Many customers beginning to position 15K to become dominate
  - If history holds: 10K also destined to fall off relatively quickly by '07 or '08

Strong Growth Ahead for Consumer Electronics

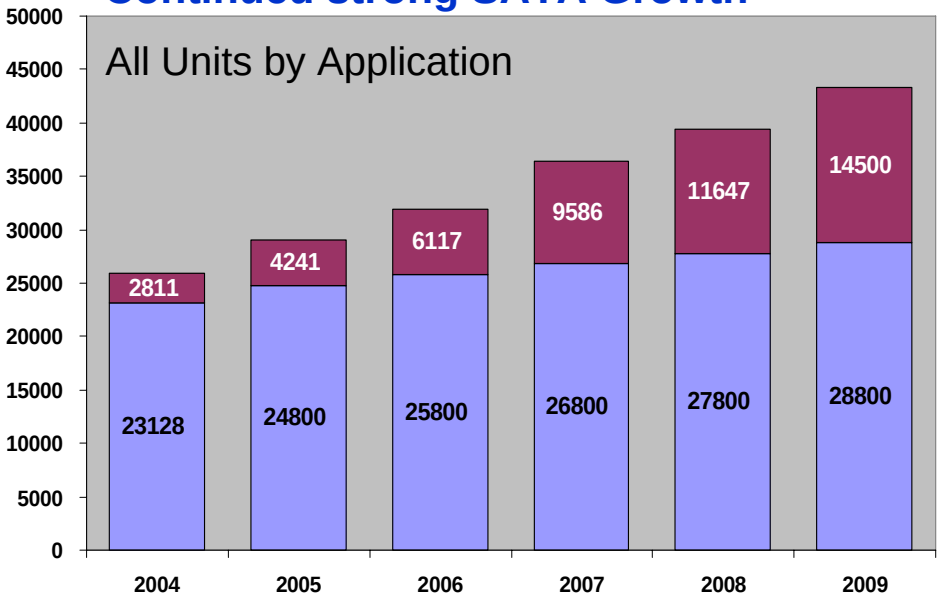


Source : Hitachi GST estimates

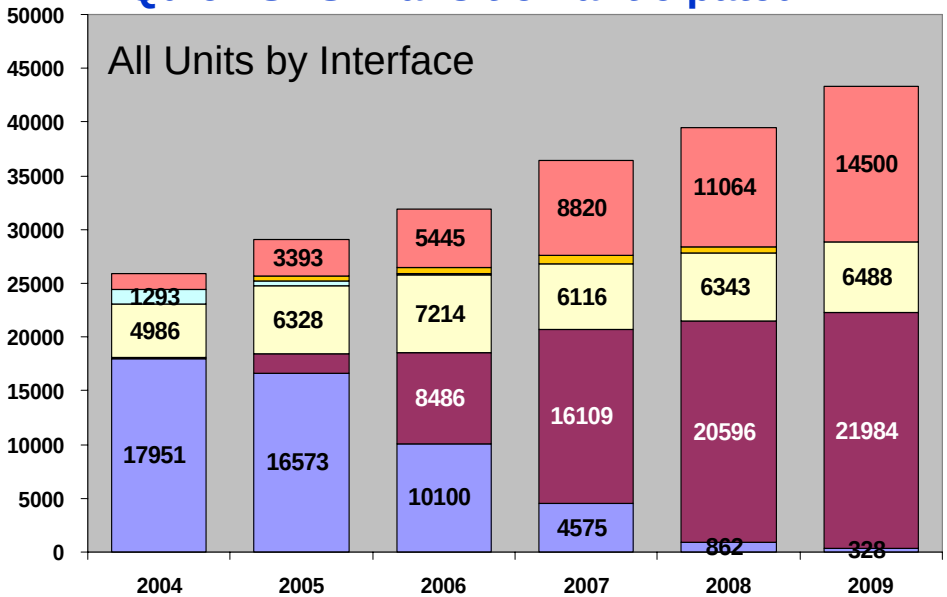
Enterprise by Application and Interface

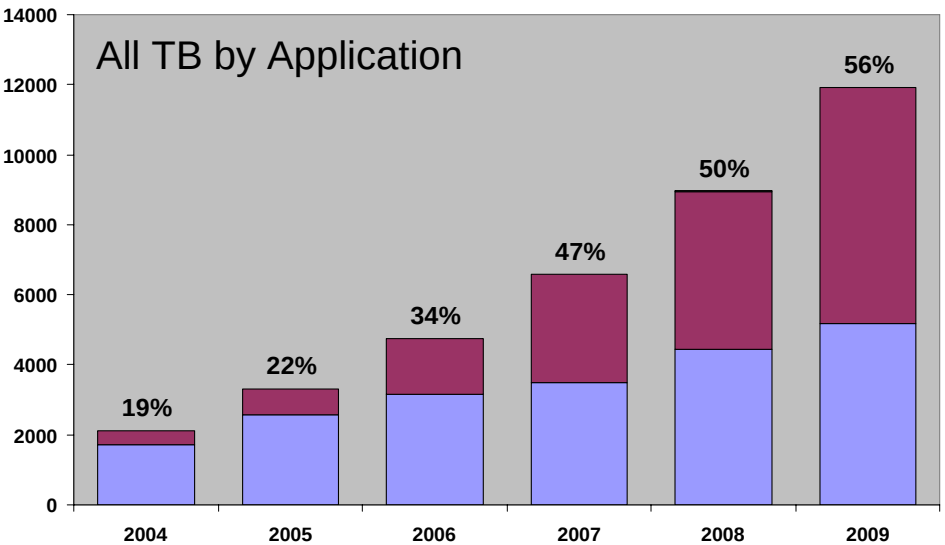
(Hitachi GST – Feb ’05)

Continued strong SATA Growth



Quick SAS Transition anticipated





CAGRs: 2004-2009

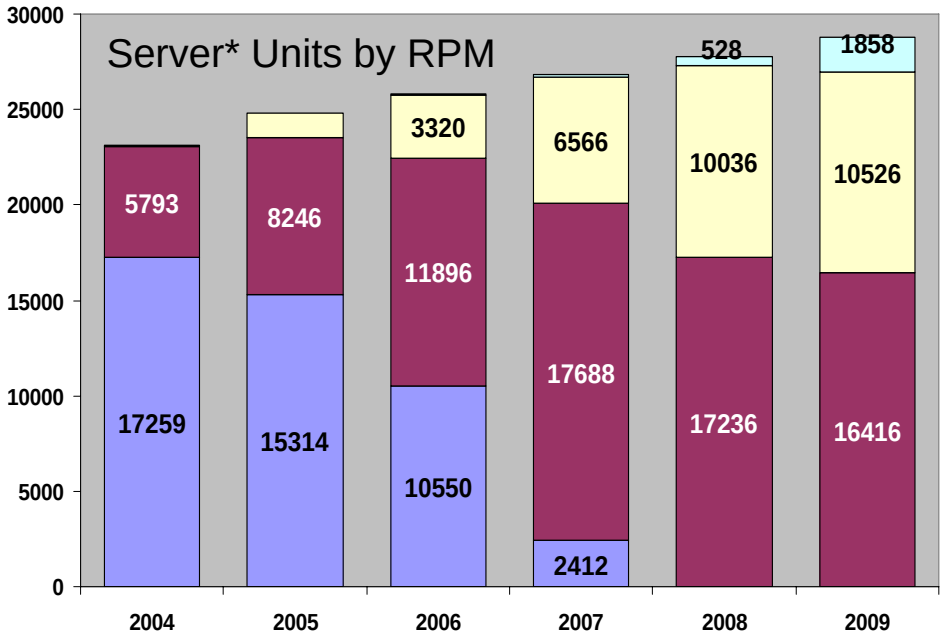
IDE / SATA:	39%
FC / LCFC:	5%
SAS:	84%
SCSI:	(55%)

CAGRs: 2004-2009	**UNIT**	**TB’s**
All Enterprise:	11%	41%
Enterprise ATA:	39%	76%
Traditional Serve	5%	25%

All Server* by RPM and Form Factor

(Hitachi GST – Feb' 2005)

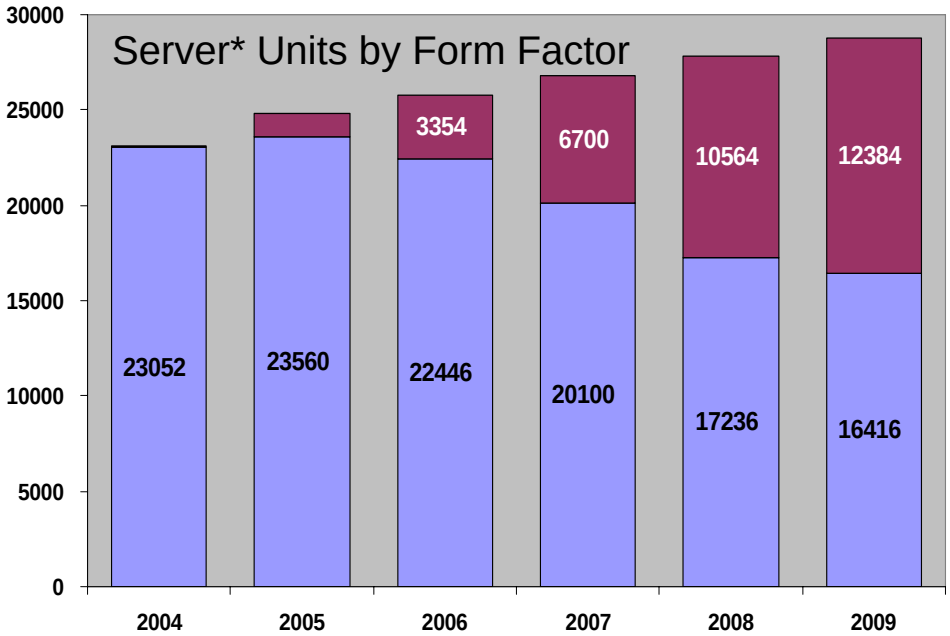
3.5" 15K transition accelerating



CAGRs: 2004-2009

SFF, 15K RPM	
SFF, 10K RPM:	104%
3.5", 15K RPM:	23%
3.5", 10K RPM:	(33%)

SFF Ramp slower than original projections



CAGRs: 2004-2009

SFF:	115%
3.5":	(7)%

*Server: consists of SCSI, FCAL, and SAS

Storage Growth Drivers

Technology Trends and Drivers

- **Traditional large SMP architectures to diminish in numbers - Server technology is moving from enterprise to clusters to blades**
  - What does this mean to storage?
    - Higher bandwidth requirements
    - High data accessibility (equivalent to high port counts today)
    - Clustered file system technology
    - Cluster partitioning
- **New computing model requires many more server images to access a common storage pool**
  - Increasing need for connectivity and safe multi-tenancy on storage
  - Increasing need for scalable performance and QoS
  - Data Management workload (snapshots, replication, backup, HSM) need to be offloaded from the servers
- **Fibre Channel, S-ATA, SAS, ?**
  - Common electrical interface will influence the decision
- **ATA disks will become the new high-end HDD**
  - Serial ATA is the growth tier drivers include Virtual Tape & Disk to Disk
  - Biggest “Pain Point” manual movement of data between Tiers
  - something else will fulfil the role of the low-end high volume disk
  - A shift in form factor - smaller

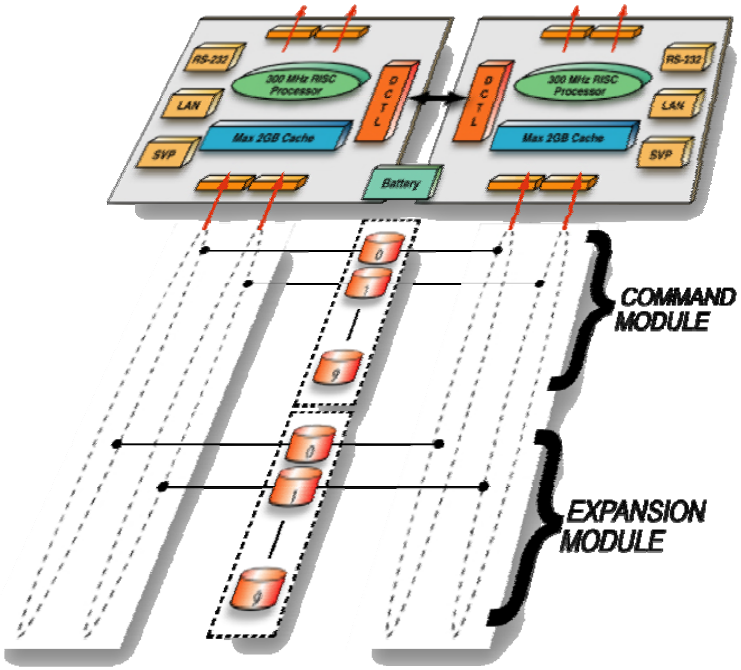
Storage Growth Drivers

- Disaster Recovery
  - Creating Distance Needs
  - High Availability
- Regulations
  - Sarbanes-Oxley Act compliance alone expected to reach US$6.1 billion in 2005, according to AMR Research
  - Basel II requires increase in capital reserves for operational risk
- Consolidation
  - Especially with Servers
- Email & Desktop
  - Video & audio email (multi- media), Desktop Back up
- Internal Requirements
  - What goes to SAN/NAS/DAS
- Getting closer to the data
  - Data Warehousing “Get close to the customer”
- Application & Database Growth
  - Oracle, SAP – the usual suspects
- iSCSI & FC over IP –
  - gets SAN ’s out of the data center and opens up more ways for the use to increase
    - But still low in the current users priorities

Storage Subsystems Technology

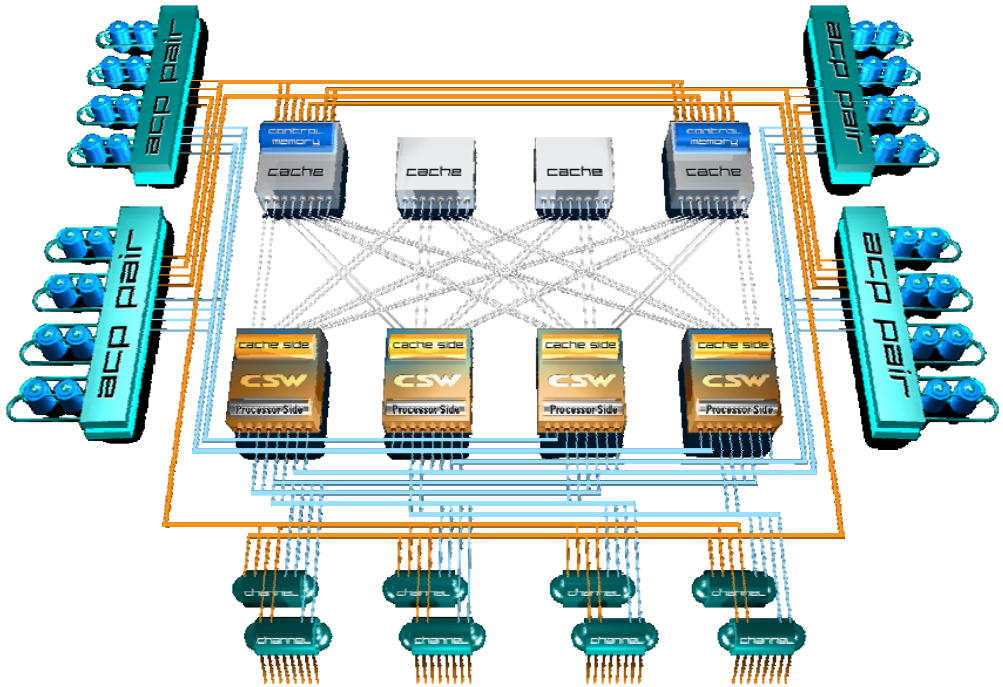
Storage Architectures – Distributed vs. Centralized

Distributed Architecture



- Lower cost, granular design
- Service a specific application, few users
- Dual processor controllers, dual caches
- Bus-based implementation

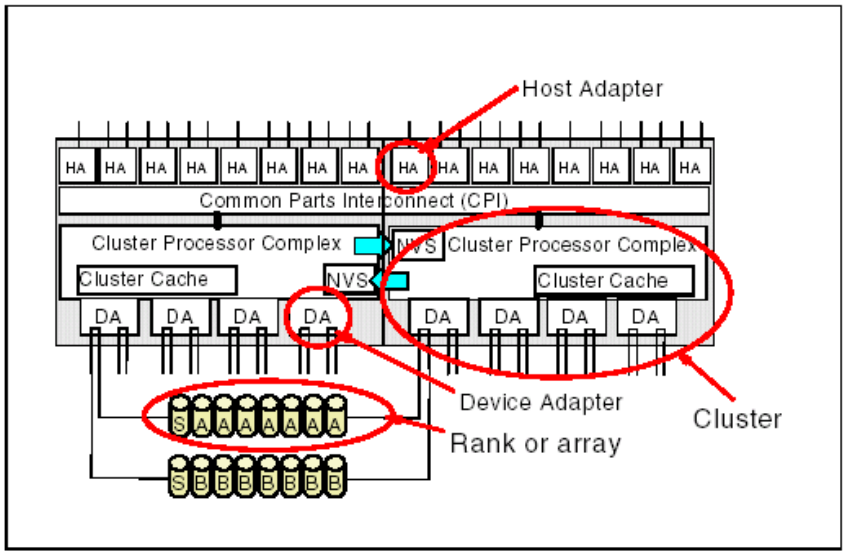
Centralized Architecture

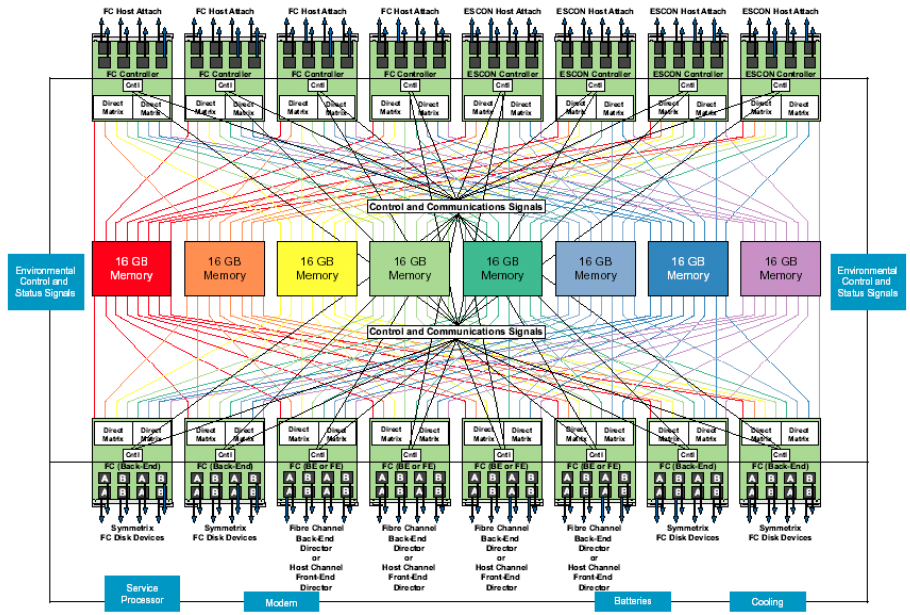


- Higher functionality, centralized design
- Service multiple applications, many users
- Multiple processor controllers, shared global cache

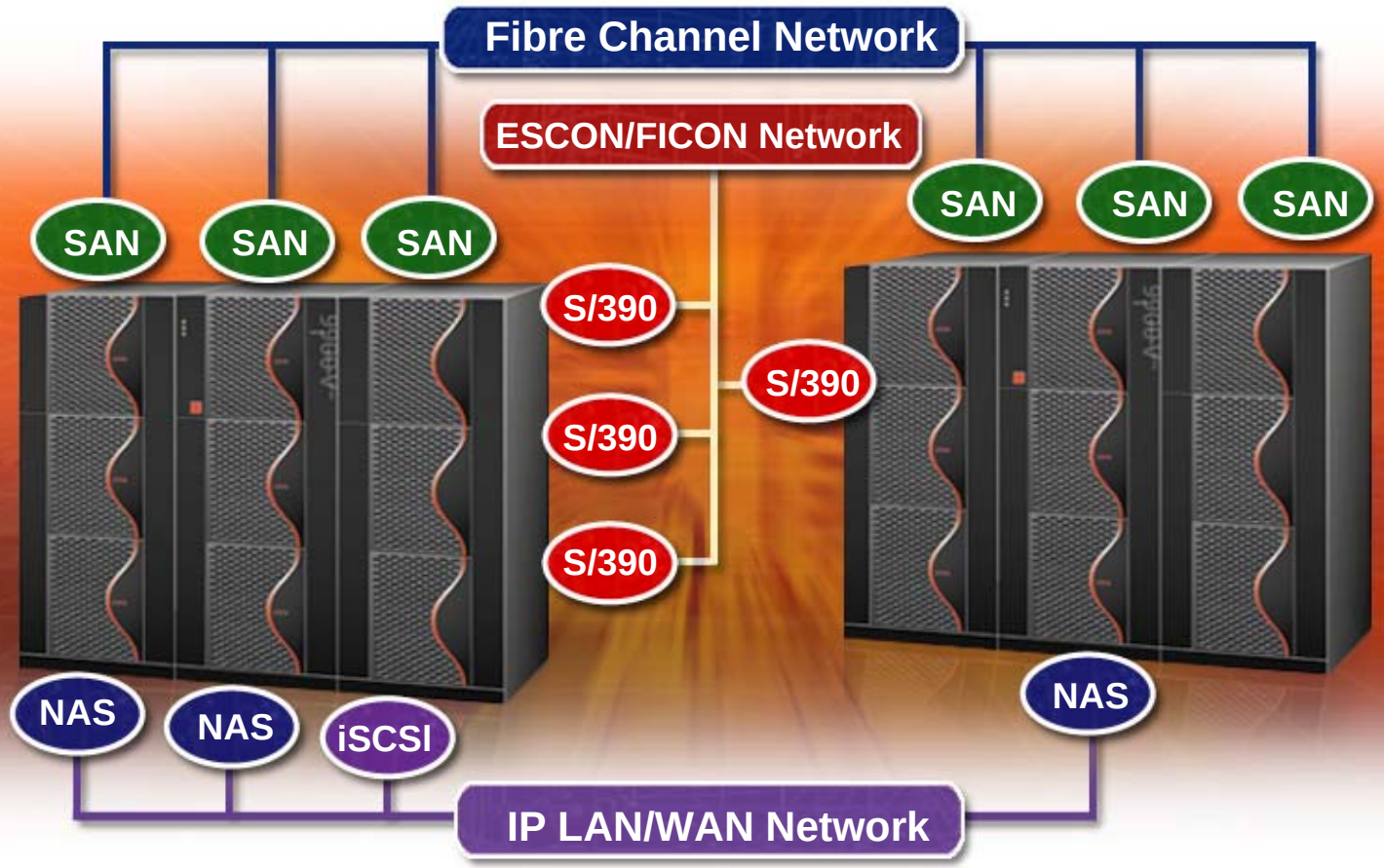
- Shared Bus, Massively Parallel Crossbar Switch or Matrix contraption-based implementation

Centralized Architecture





Extensive, Multi-Protocol Connectivity



Future in High-end

- What comes next and what is going to be used in the next five years?
  - Virtualization and management
  - New connectivity

A look at what is hot today

- **Tiers, Tiers, Tiers….for both SAN & NAS**
  - Serial ATA is the growth tier drivers include Virtual Tape & Disk to Disk
  - Biggest "Pain Point" manual movement of data between Tiers
  - Missing Tier: IP SANs > sluggish adoption within the Fortune 1000

- **ILM Still "Interesting" ; Implementations are Sparse**
  - Potential to automate data movement of unstructured data is viewed as entry

- **Virtualization: A "Momentum Play"**
  - **HDS, IBM & smaller firms**

- **NAS is a starting point for Virtualization, ILM**

- **Storage Management Software "Challenged"**
  - Virtualization adding value for core Management needs
  - Unbundling of the purchase causes some surprises in budget allocations
  - Very "Tactical" functionality being implemented ie Reporting

- **Too Early: WAFS, Storage Grids, Thin Provisioning, Fabric Intelligence**

- **OS Functionality more robust? – Volume Management, Multi-pathing**

THEINFOPRO

Enter Storage Virtualization

The SNIA Dictionary

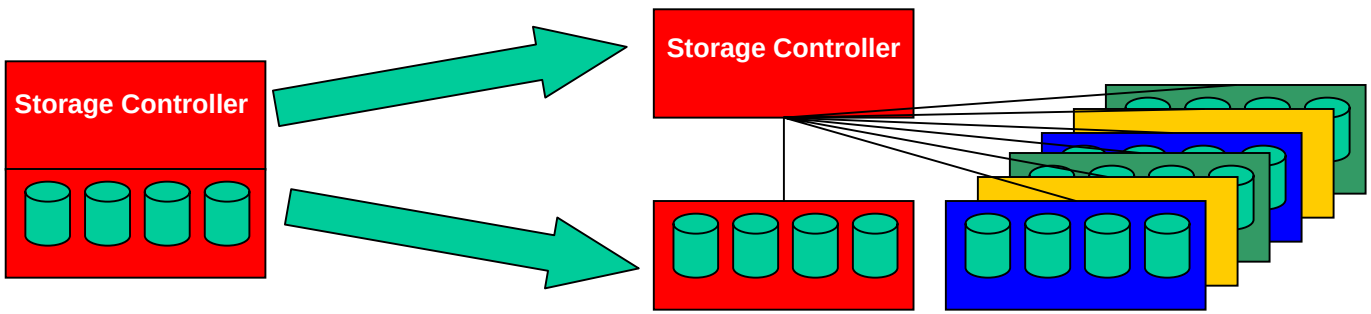
http://www.snia.org/education/dictionary/v/
defines virtualization as

"An act of integrating one or more (back end) services or functions with additional (front end) functionality for the purpose of providing useful abstractions. Typically virtualization hides some of the back end complexity, or adds or integrates new functionality with existing back end services."

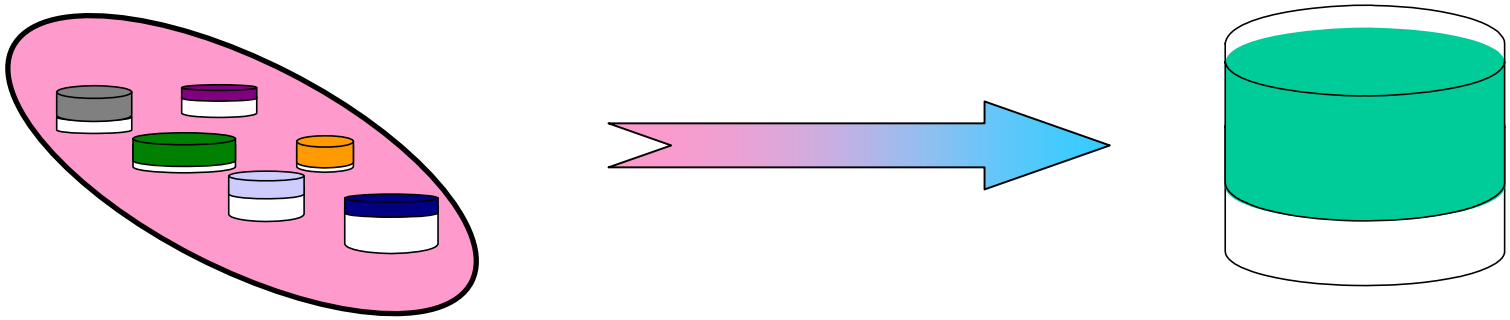
The Innovator's Dilemma

Clayton M. Christensen

- Innovators who continually enhance their products are ambushed by disruptive technologies that are "good enough"
  - SATA storage is a disruptive technology
- Chasing disruptive technologies will compromise innovation
  - Requires a move to another dimension
- Virtualization changes the dimensions of the storage market by separating the storage controller from the disks
  - The real value is in the controller
  - Disks are becoming a commodity
  - Unbundling the storage controller enables more choices



What is the promise of Storage Virtualization?

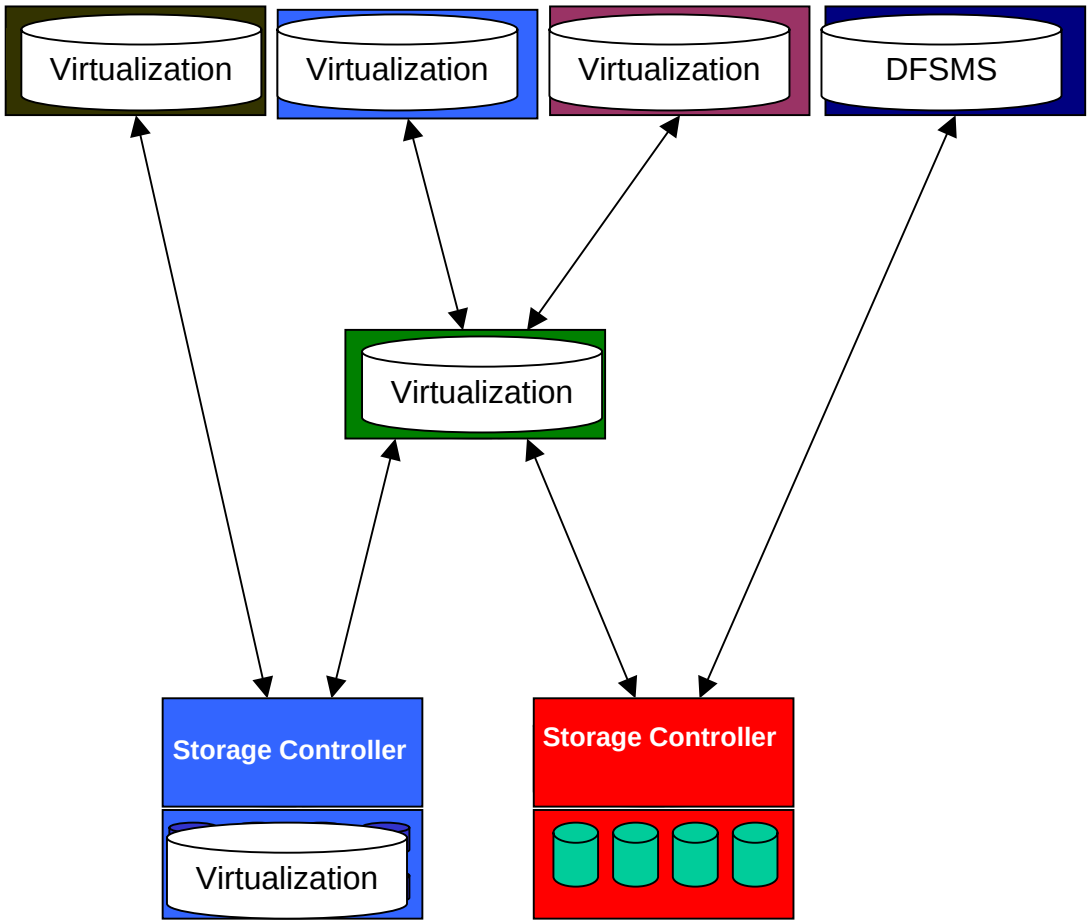


- View the entire SAN as one pool of storage independent of physical location
- Mask the differences between heterogeneous devices
- Simplify control and management under a single interface.
- Centralize storage volume management on the SAN
- Dynamically allocate capacity to the applications that need it
- Facilitate matching storage resources to business requirements

Requirements For A Comprehensive Storage Virtualization Solution

- **Volume Mapping and data mobility across heterogeneous vendor platforms**
- **Support multiple platforms with standard protocols**
  - Direct attach, Fabric, FC, ESCON, FICON, NAS
- **Does not introduce another layer of complexity**
- **High performance and low latency**
- **Scalable to thousands of host connections and petabytes of storage**
- **Secure multi-tenancy for thousands of users sharing the same physical assets**
- **Logical partitioning for Quality of Service**
- **Common end to end management for heterogeneous vendor platforms**

Where Should Virtualization Reside?

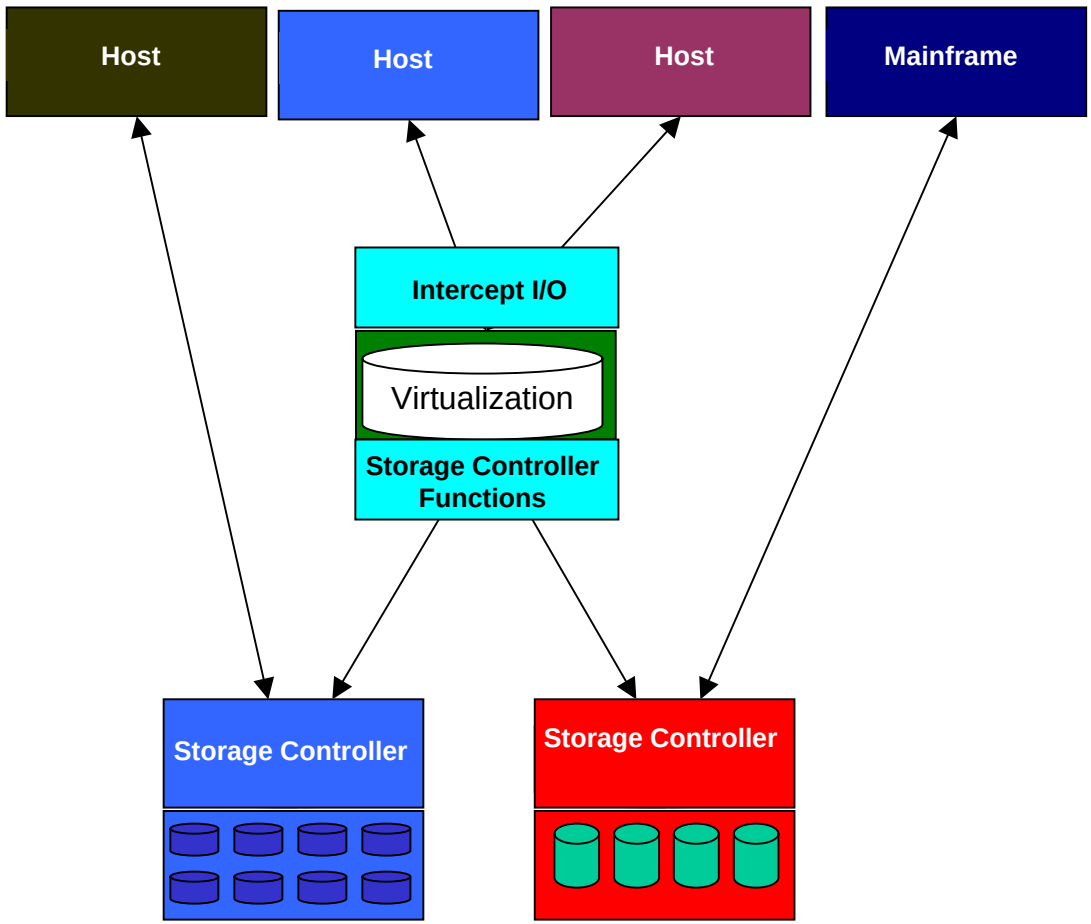


Virtualization in the host?

In the SAN?

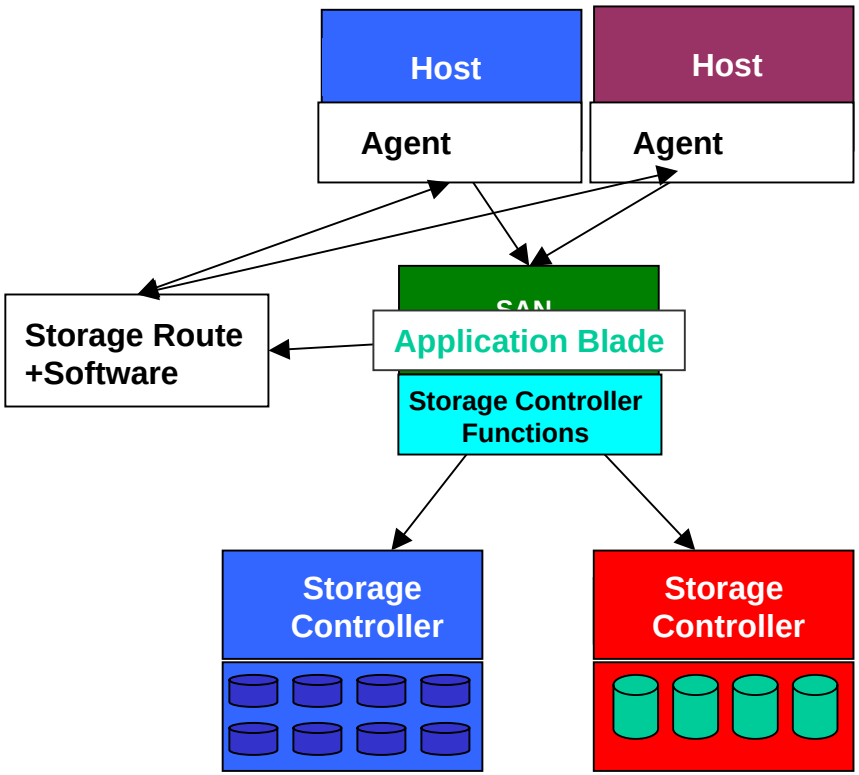
In the Device?

Challenges of Inband SAN Virtualization



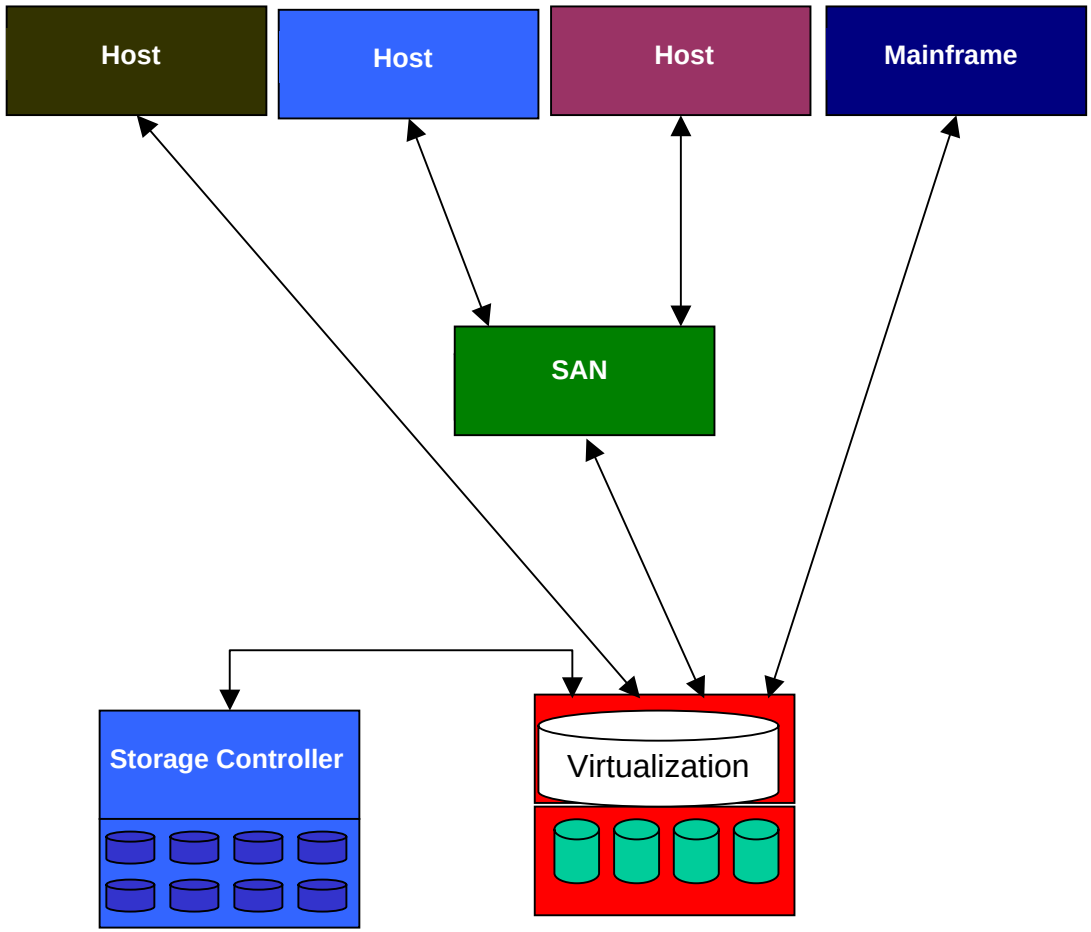
- Additional complexity
- SAN only attachment
- Increased latency
- Limited data mobility functions
- Lacks scalable cache and access ports
- Requires APIs
- Violates end-to-end security

Challenges of Storage Router Virtualization In the SAN



- Agents in the host intercept I/O
- Storage Router redirects the I/O
- State is kept in the Switch and creates a vendor lock in
- Re-invent storage controller in the switch
- Lacks scalable cache and access ports
- Requires APIs and/or standards
- Violates end to end security

Virtualization in the Storage Controller



- No added complexity or latency between the host and storage
- Virtualization is done behind the Controller
- External storage attaches through standard Fibre Channel protocols

Connectivity – What is the future

Technology Trends & Drivers

Fibre Channel – Roadmap (www.fibrechannel.org)

Base 2

Product Naming	Throughput (MBps)	Line Rate (Gbaud)	T11 Spec Completed (Year)	Market Avaiability (Year)
1GFC	200	1.065	**1996**	1997
2GFC	400	2.125	**2000**	2001
4GFC	800	4.25	**2003**	2005
8GFC	1,600	8.5	**2006**	2008
16GFC	3200	17	**2009**	2011
32GFC	6400	34	2012	Market Demand
64GFC	12800	68	2016	Market Demand
128GFC	25600	136	2020	Market Demand

Base 10

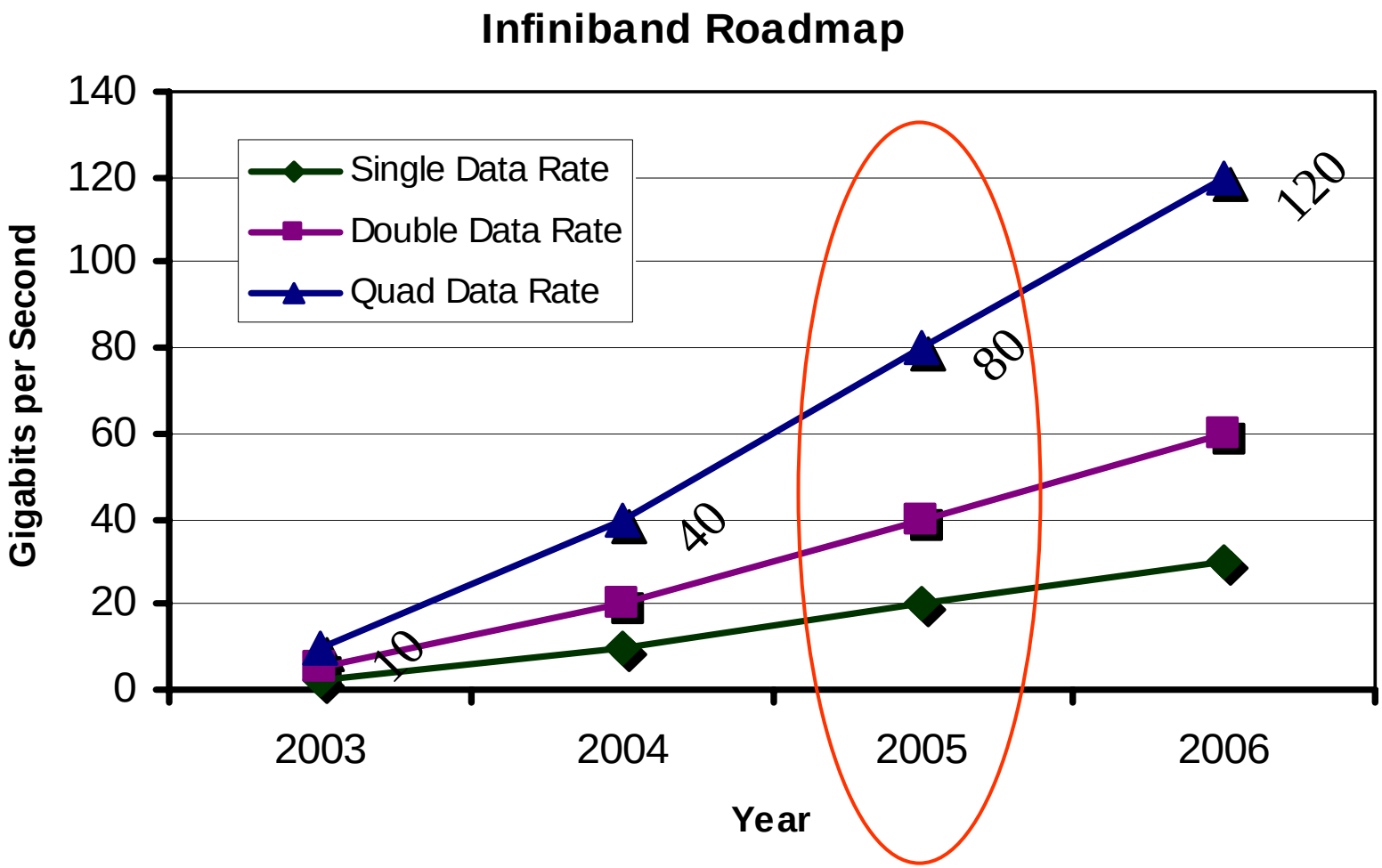
10GFC	2400	10.52	**2003**	2004

*Base2 used throughout all applications for Fibre Channel Infrastructure and devices. Each speed maintains backward compatibility at least two previous generations (I.e., 4GFC backward compatible to 2GFC and 1GFC)

**Base10 commonly used for ISL, core connections, and other high speed applications demanding maximum bandwidth

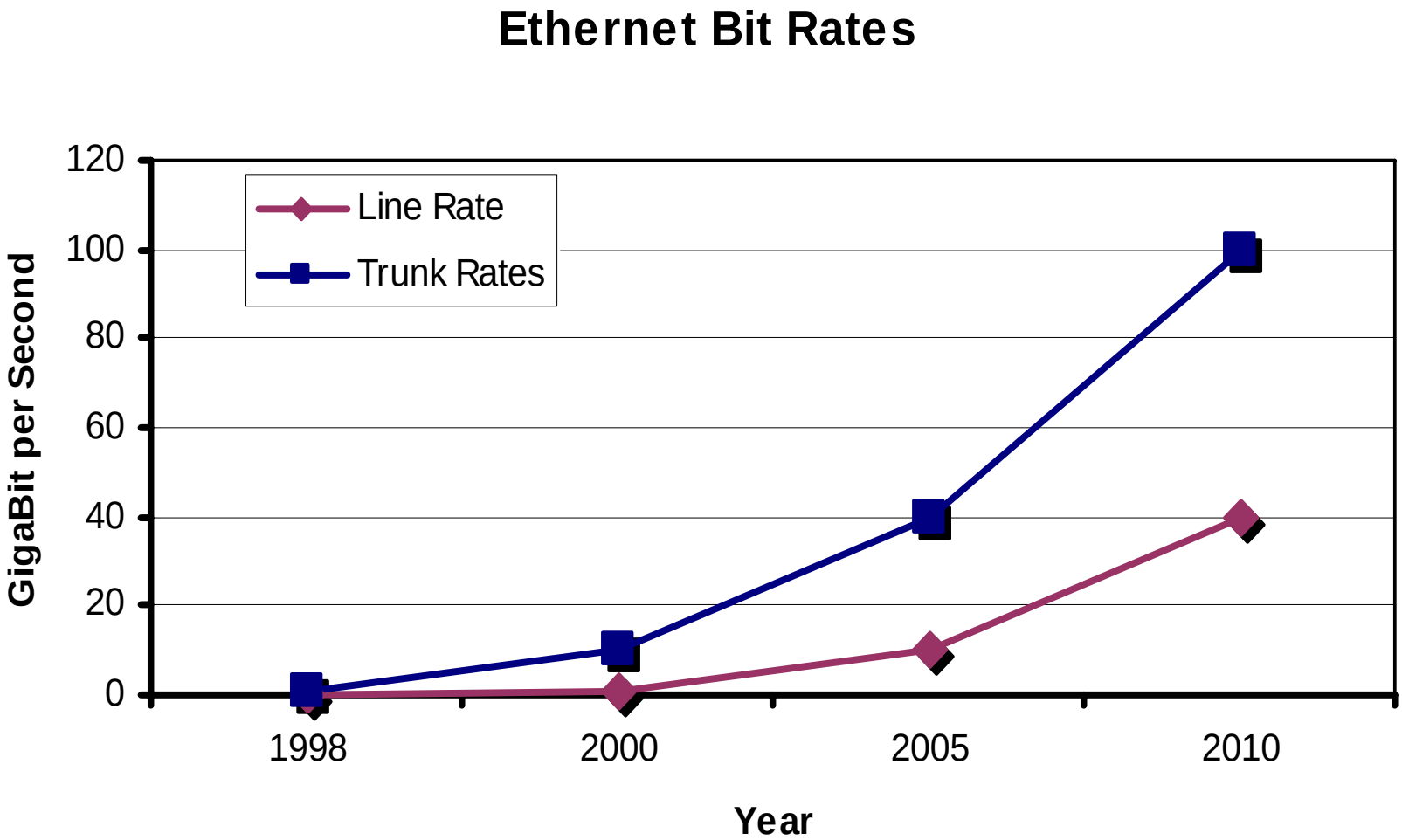
Technology Trends & Drivers – *cont.*

InfiniBand (www.infinibandta.org)



Technology Trends & Drivers – *cont.*

Ethernet



Conclusions

- Larger and faster units will greatly leverage forms of virtualization
- Multi-connectivity within the same system is a requirement
  - Use the best connection for the application to serve
- Hard Disk technology less important as long as it gets the job done
  - FC, SAS, SATA will have their common space for a while
- More Virtualization
  - Any form, multiple vendors
  - Ease the storage management

Thank you!

Roberto.Basilio@hds.com