



Storage Scaling in Support of Changing Data Center Trends

Claus Mikkelsen

Chief Scientist

Hitachi Data Systems



The Way We Were

SNIA



SNW
COMPUTERWORLD

October 11-14, 2010
The Gaylord Texan
Dallas, Texas

10 Megabyte Hard Disk
\$3,495*



5440-12 Top Load Drive
* Factory rebuilt 10MB cartridge disk drive only.
A new Cameo Data Systems controller is available for \$1,495
\$4,495 for a brand new Ampex 10MB drive only



We are the CP/M** and MP/M** specialist of Southern California. We can supply you with the latest CP/M (\$150) or MP/M (\$300) and with Standard BIOS (\$150) or Custom BIOS (\$300). Immediate delivery worldwide. Domestic and foreign inquiries invited... dealers too.



We are a full service computer retailer. We totally integrate hardware and software into high quality, high reliability systems. Systems for use in development, process control and general business. Word processing naturally, multi tasking and multi processing too.

COMPUTER COMPONENTS

Circle 279 on inquiry card. 5848 Sepulveda Boulevard Van Nuys, California 91411 213•786-7411

BYTE July 1980 291

$$\$3500 + \$1500 = \$500/\text{MB}$$

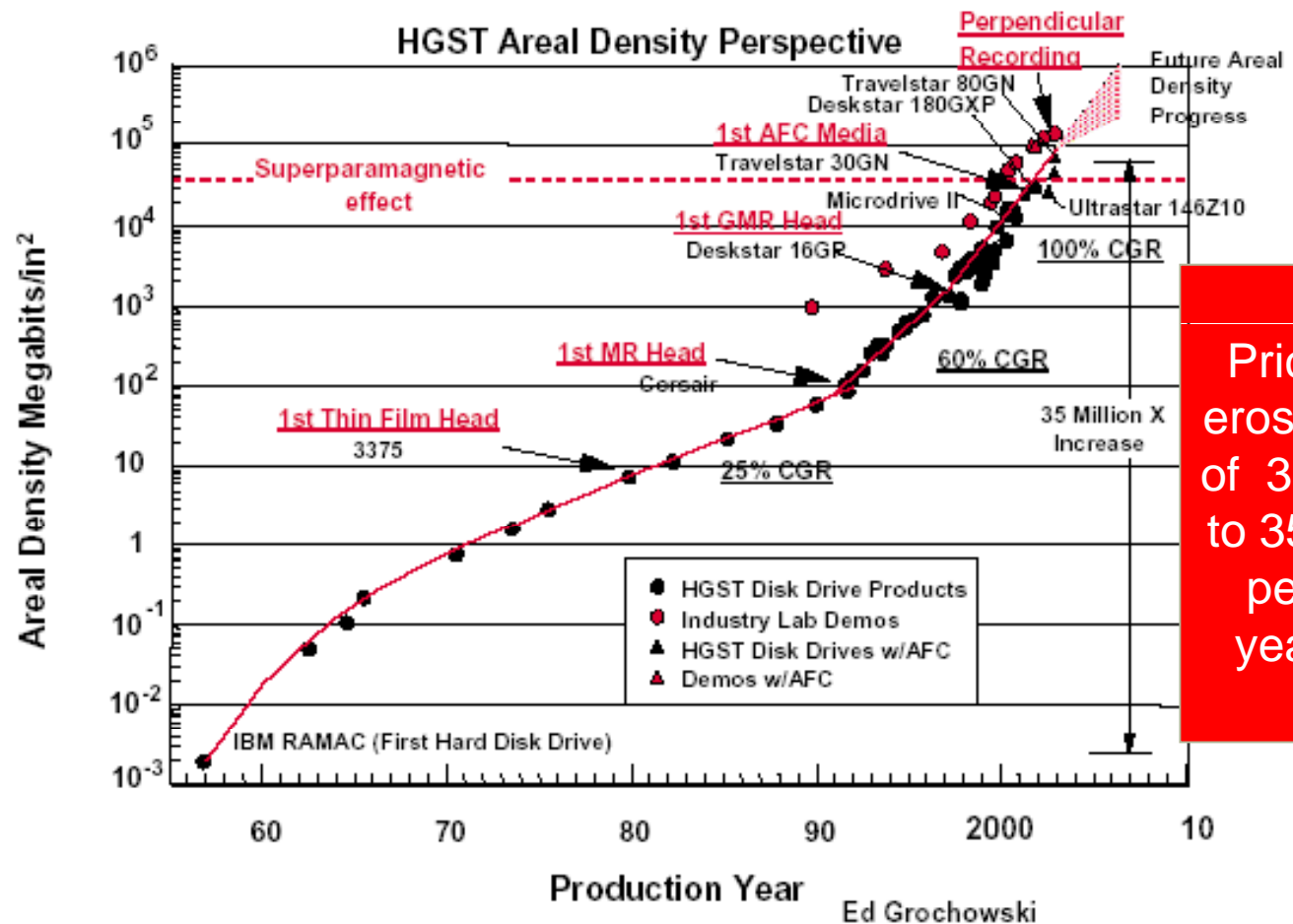
Or

$$\$500,000/\text{GB}$$

Or

$$\text{\$ } \frac{1}{2} \text{ Billion / TB}$$

Storage Capacity Will Continue to Increase



San Jose Research Center

Hitachi Global Storage Technologies

SNIA

SNW

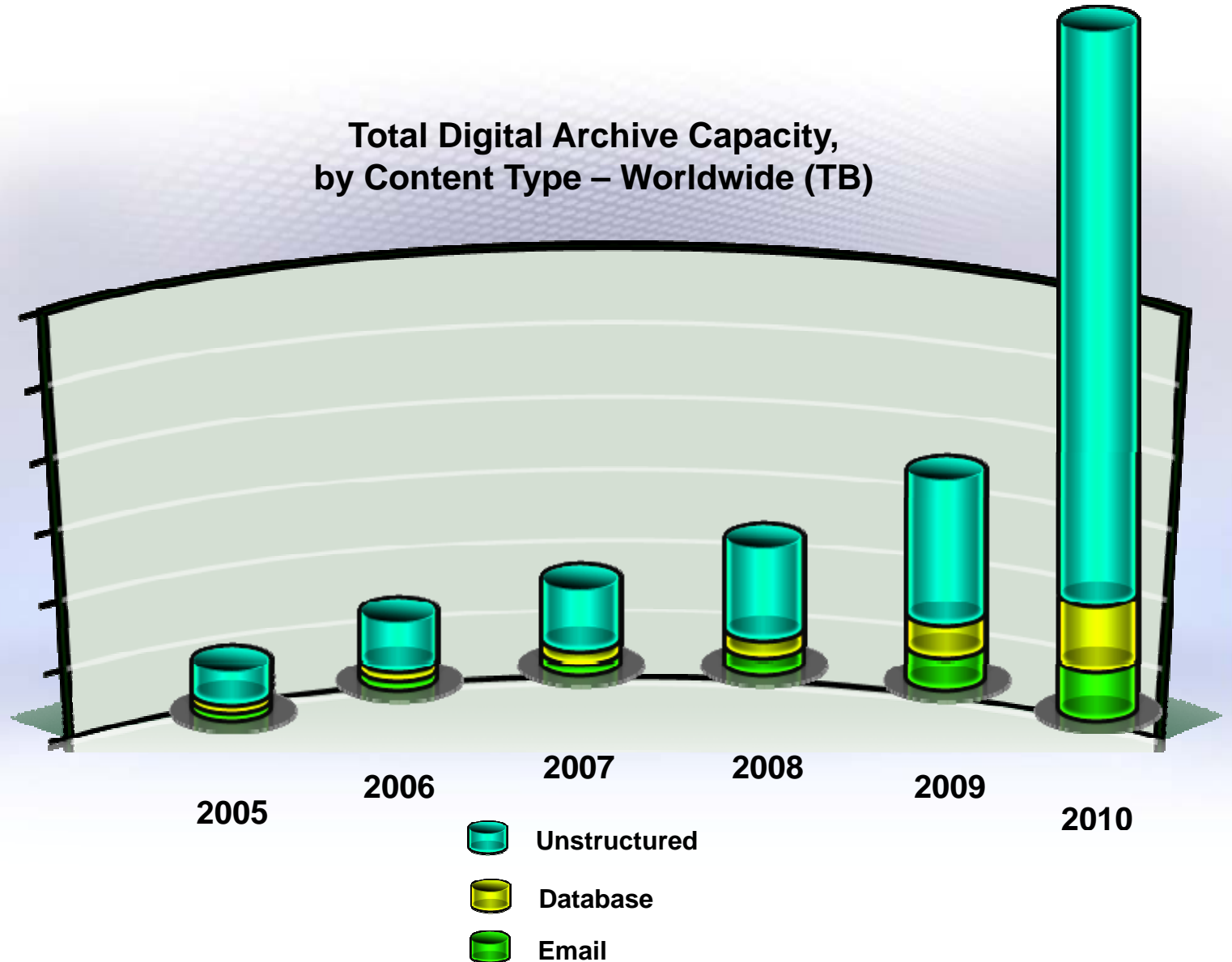
COMPUTERWORLD

October 11-14, 2010
The Gaylord Texan
Dallas, Texas

arpers2003a.prz

The Changing Forms of Data

Total Digital Archive Capacity,
by Content Type – Worldwide (TB)





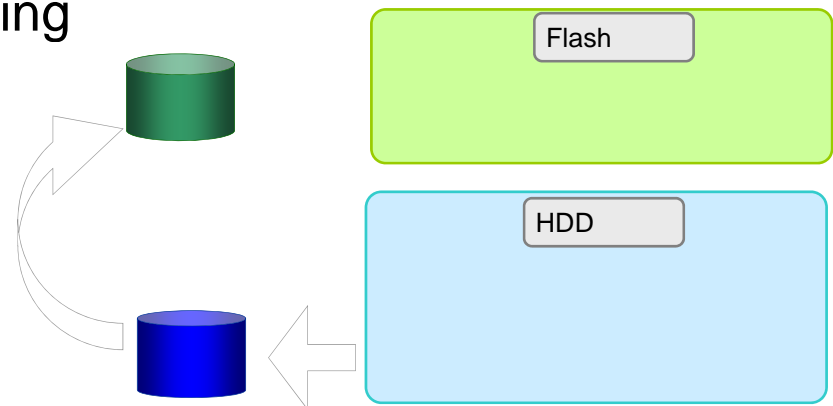
But a Bit on Structured Data

- Generally, most critical applications
- Access density to “short stroking”
- Little relief from the disk drive guys



Flash Drives

- Very high performance for critical applications
 - 70x random reads, 14x random writes
 - 50% reduction in power and cooling
 - Enterprise SLC flash has 100,000 write/format cycles
 - Wear leveling, error recovery, spares
 - Currently 10 times higher cost than HDD
-
- Optimize use of flash with tiering, but where?
 - Eliminate waste of allocated unused flash capacity with dynamic provisioning





The Industry Still Needs to Improve

Utilization

- 51% of open system data is unnecessary
- 22% of data is duplicate
- 68% of data has not been accessed for 90 days or more
- 47% of total storage management cost is spent on file management activities, 12% on backup and restore, remainder is hardware servicing

Availability and Capacity Planning

- 47% of open systems capacity is available, but in the wrong place
- 55% of unplanned server outages occur from out-of-control disk space consumption

How does your company compare?

Sources: Strategic Research Corporation, and SNIA



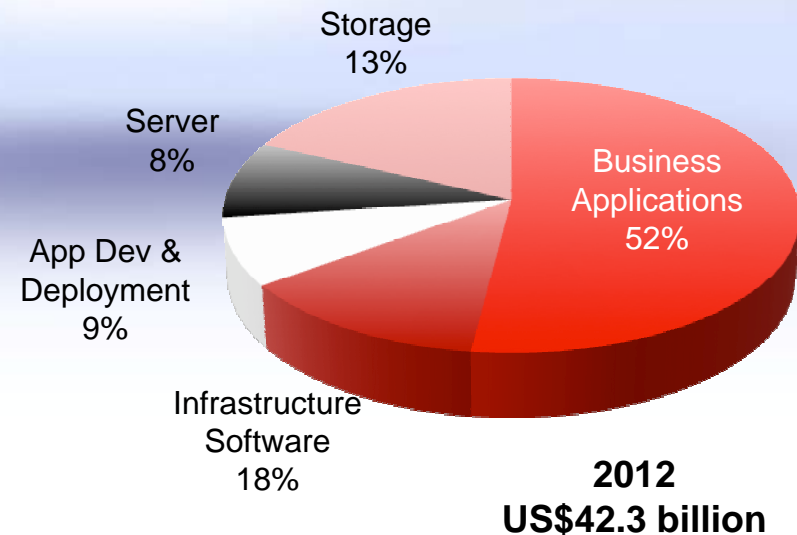
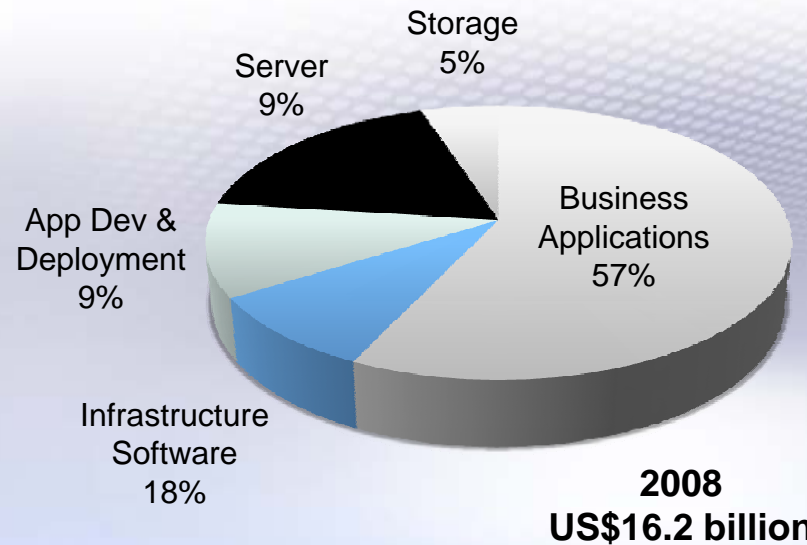
From Half a \$Billion per Terabyte to...

- We consider storage **capacity** to be “free”
- We consider storage **capacity** to be “unlimited”
- We buy more **capacity** when we know our utilization rate is ~30%
- We buy more **capacity** to solve access density
- We buy more **capacity**



Storage Quality: Key to Controlling Cost Growth

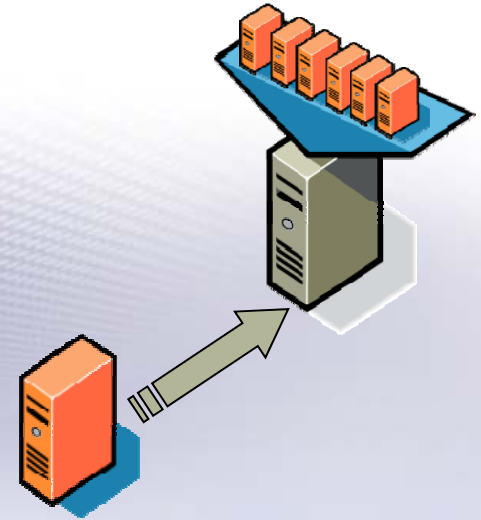
IDC: Worldwide IT Cloud Services Spending* by Product/Service Type 2008, 2012





Servers and Networks Are Scaling Up

- Nahelum/Xeon
 - Intel Architectures
 - Multi-core – Dual, quad, octo
 - SMT – Simultaneous multi-thread
 - L1, L2, L3 cache
- Virtual servers and hypervisors
 - Server clusters with 200 to 240 OS instances
- Networks
 - FC 8Gbs going to 16Gbs
 - FCoE 10Gbs to 40Gbs to 100Gbs



Scale-up storage is required to support virtual servers and networks

A Perfect Storm



The number of virtual servers in an organization is expected to grow by 50% in 2010

1st
TIME

In 2009, more virtual servers than physical servers were deployed



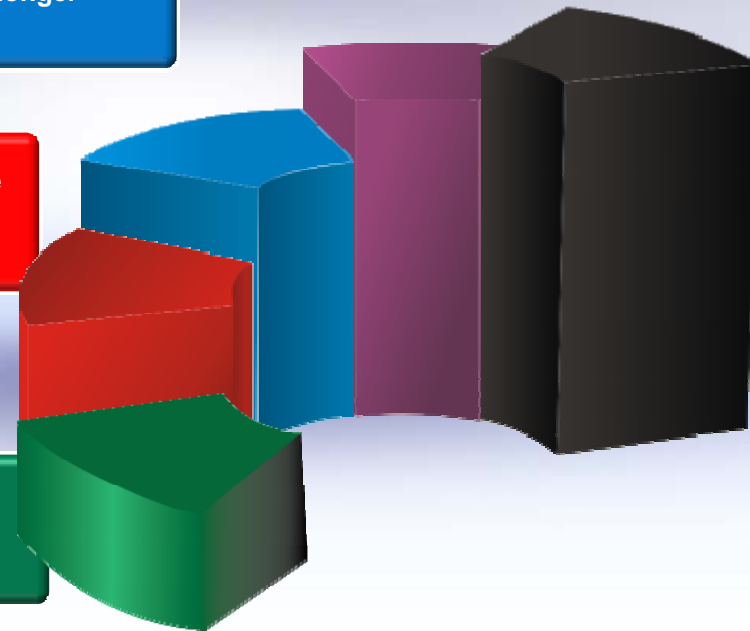
Regulations and laws now require more data to be protected longer



The volume of new healthcare data quadruples every two years



Data grows an average of 50%-60% each year



Today's Customer Requirements

- **Storage growth has become irrational** – It does not track to business performance
- **Keeping data longer** – Compliance driving data discovery and retention.
- **File and content data is not well managed today**
- **Global economy** – increasing pressure on IT budgets and investments
- **Green solutions** – Power and cooling optimization is a key consideration
- **“Help me save costs”** – while providing “right” performance and access to my data.





Data Center Transformation Today

Data Center Trends

Customer Requirements

Aging Infrastructures

Ease of migration to new technologies

Servers Scaling Up with Multiple Cores

Scaling up and scaling out of applications and storage

Virtual Servers Are Mainstream

Enterprise class scalability, availability and performance

Cloud Adoption

More agile infrastructure with safe multitenancy

Application Growth

Classify data to optimize cost and performance



But is There Hope?

Single Platform – All Data

Thin Provisioning

Capacity Efficiency



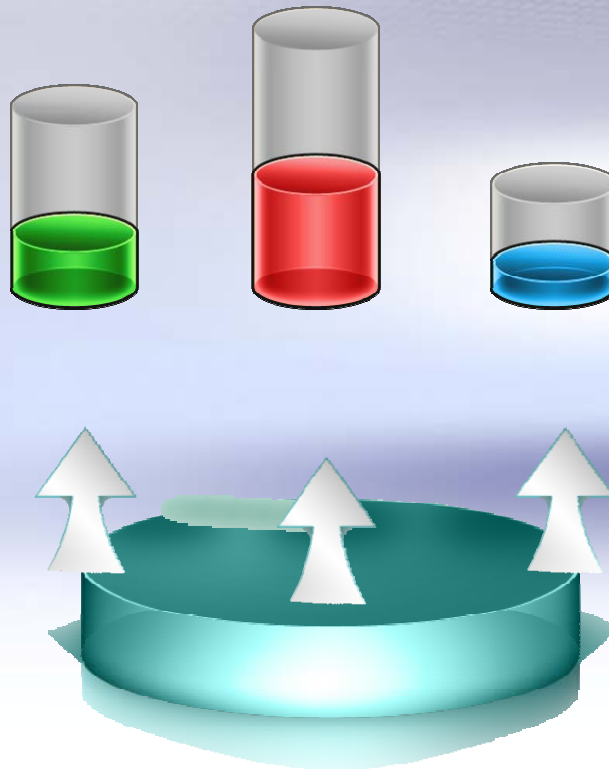
Sub-LUN Tiering

Storage Virtualization

Wide Striping



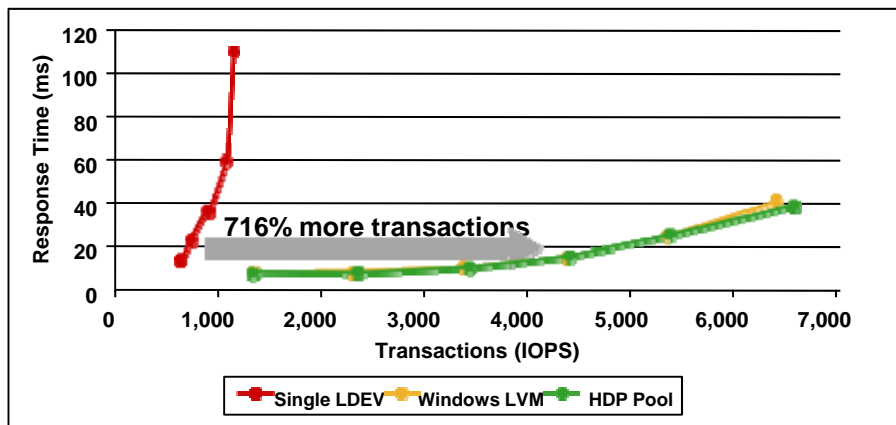
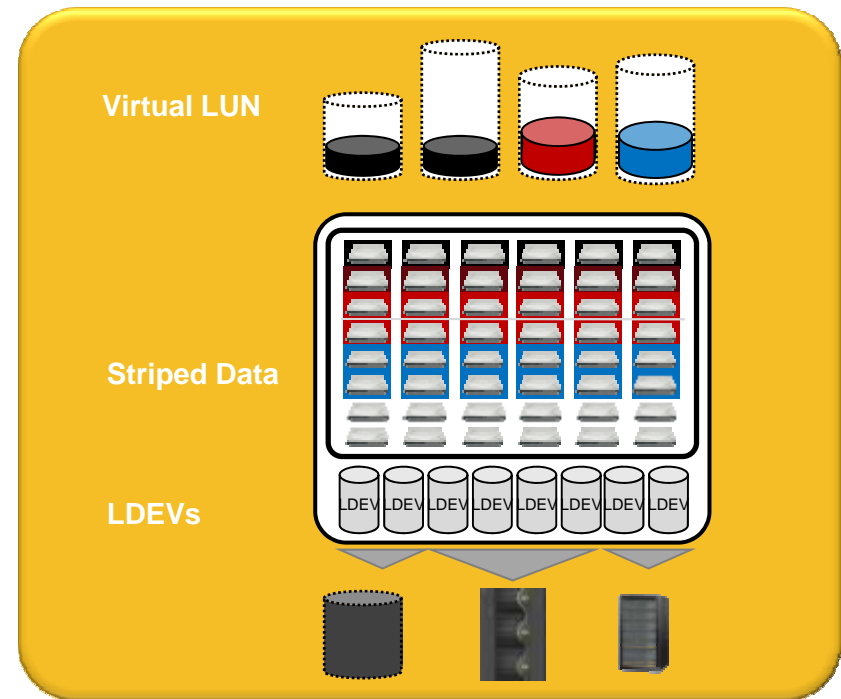
Thin Provisioning (or whatever your vendor calls it)



- New physical disk resources are allocated to virtual volumes as applications grow naturally over time
- Proactively monitors and alerts storage administrators before more physical storage is required
- New physical capacity can be added nondisruptively to storage systems when needed

Wide Striping for Performance

- Virtualize devices into a pool of capacity and stripe data across many drives
- Optimize storage performance by spreading the I/O across more arms

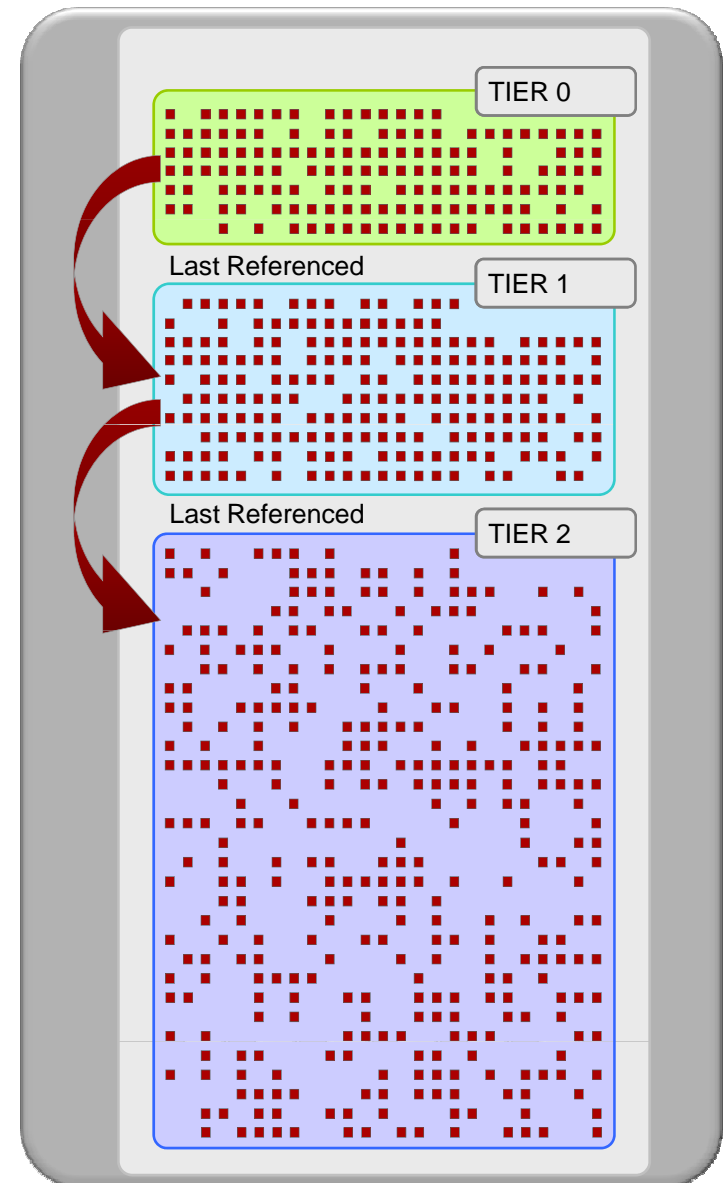


October 11-14, 2010
The Gaylord Texan
Dallas, Texas



Automatic Sub-LUN Tiering

- Saves costs – Right storage at the right use; self-optimizing automates data placement
- The volume's most active data moves to the highest performance tier
- Less active data moves to the lowest cost tier
- Focused performance – Automatically puts the right data in the right place at the right time
- Operational efficiency – Automation reduces human error and takes out the guess work of using flash drives



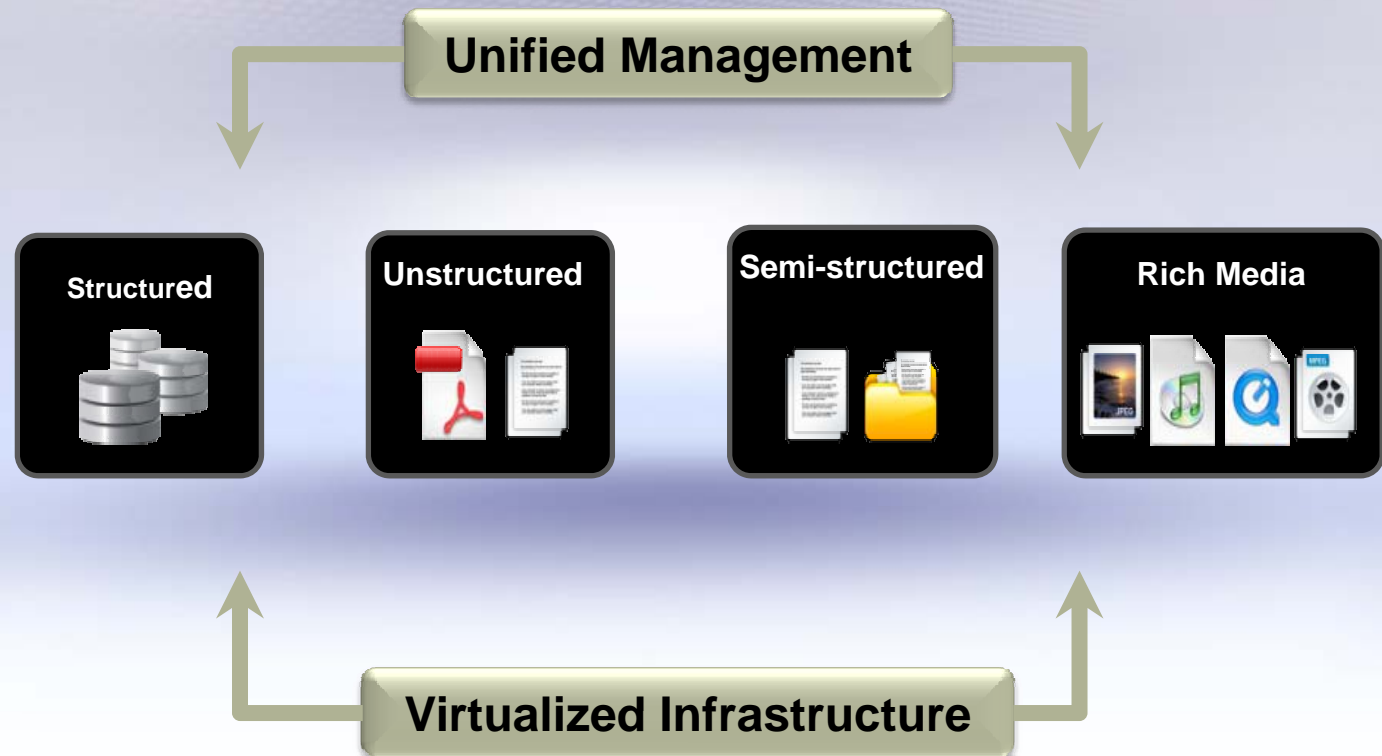


Capacity Efficiencies

- Thin provisioning
 - Zero page reclaim (or whatever your vendor calls it)
 - Write-Same
- “Spaceless” CoW snapshot
- “Thick to thin” and “thin to thin” replication
- De-dupe
- Archive stale data
- Single instance store
- RAID 5/6 rather than RAID 10

Single Platform for All Data

A single, virtualized platform for all data, content and information





Predictions



- Low utilization will become a key business issue
- Storage virtualization will provide the data mobility to consolidate and migrate data
- New tools for replication, copy on write, de-duplication, thin provisioning, and single instance store will reduce the amount of data that needs to be stored
- Archiving will reduce the toxic waste of stale data and ensure compliance for records retention
- Storage will be viewed in the context of a deliverable service as IT moves toward a dynamic data center
- Green engineering capabilities will become synonymous with power and cooling requirements



Thank you!



COMPUTERWORLD

SNIA

SNW

OCTOBER
11-14
2010

The Gaylord Texan, Dallas, Texas