Virtualization and Scale up Storage Architecture

April 13, 2010

Hubert Yoshida

Chief Technology Officer

Hitachi Data Systems

HITACHI
Inspire the Next
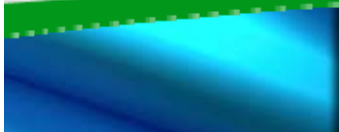
IDC World Wide IT Spend on CAPEX vs OPEX

CAPEX spending has remained flat while overall costs continue to Increase due to OPEX
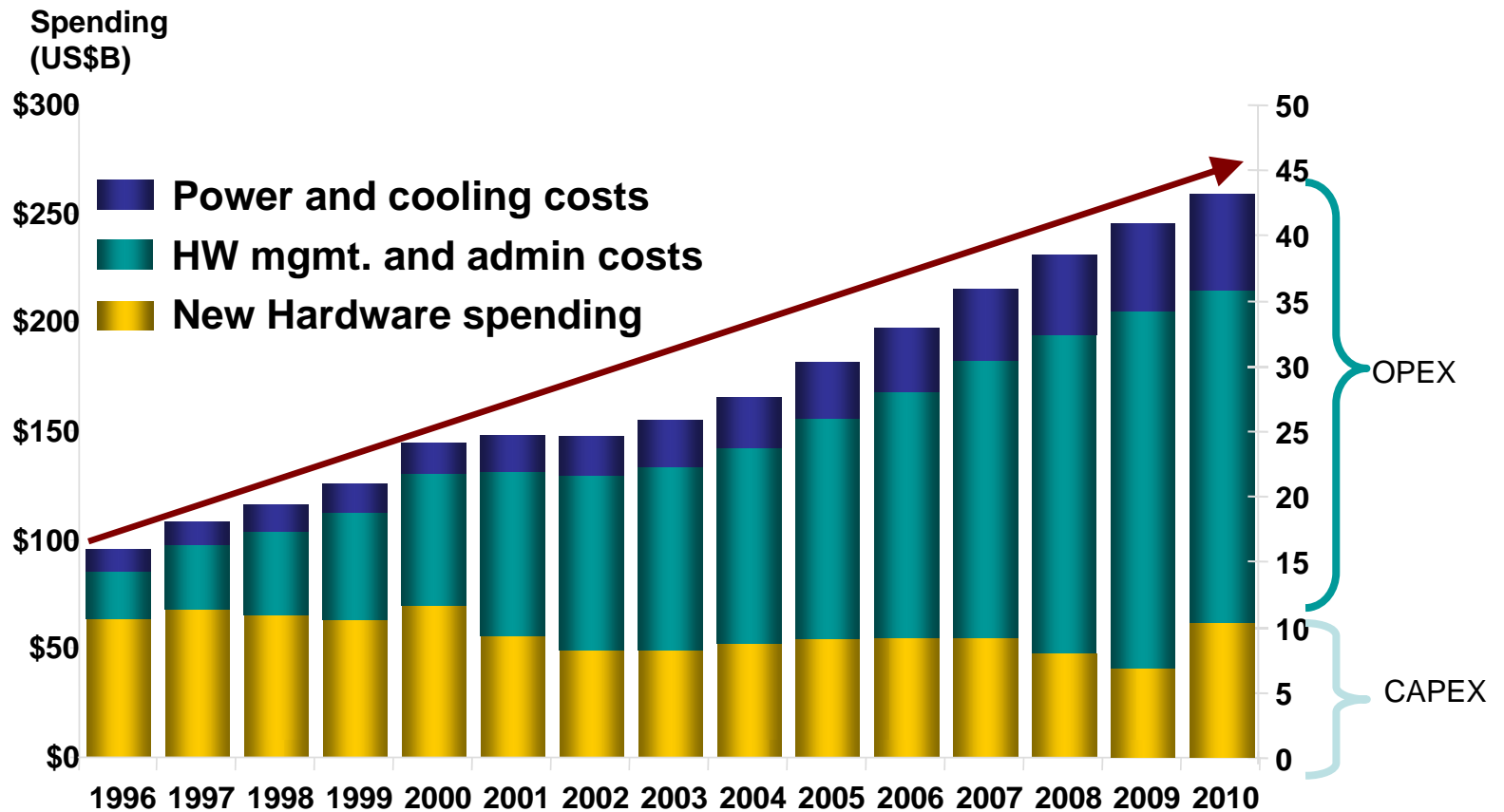
A Sustainable Future

Apply Technologies to Reduce Operational costs
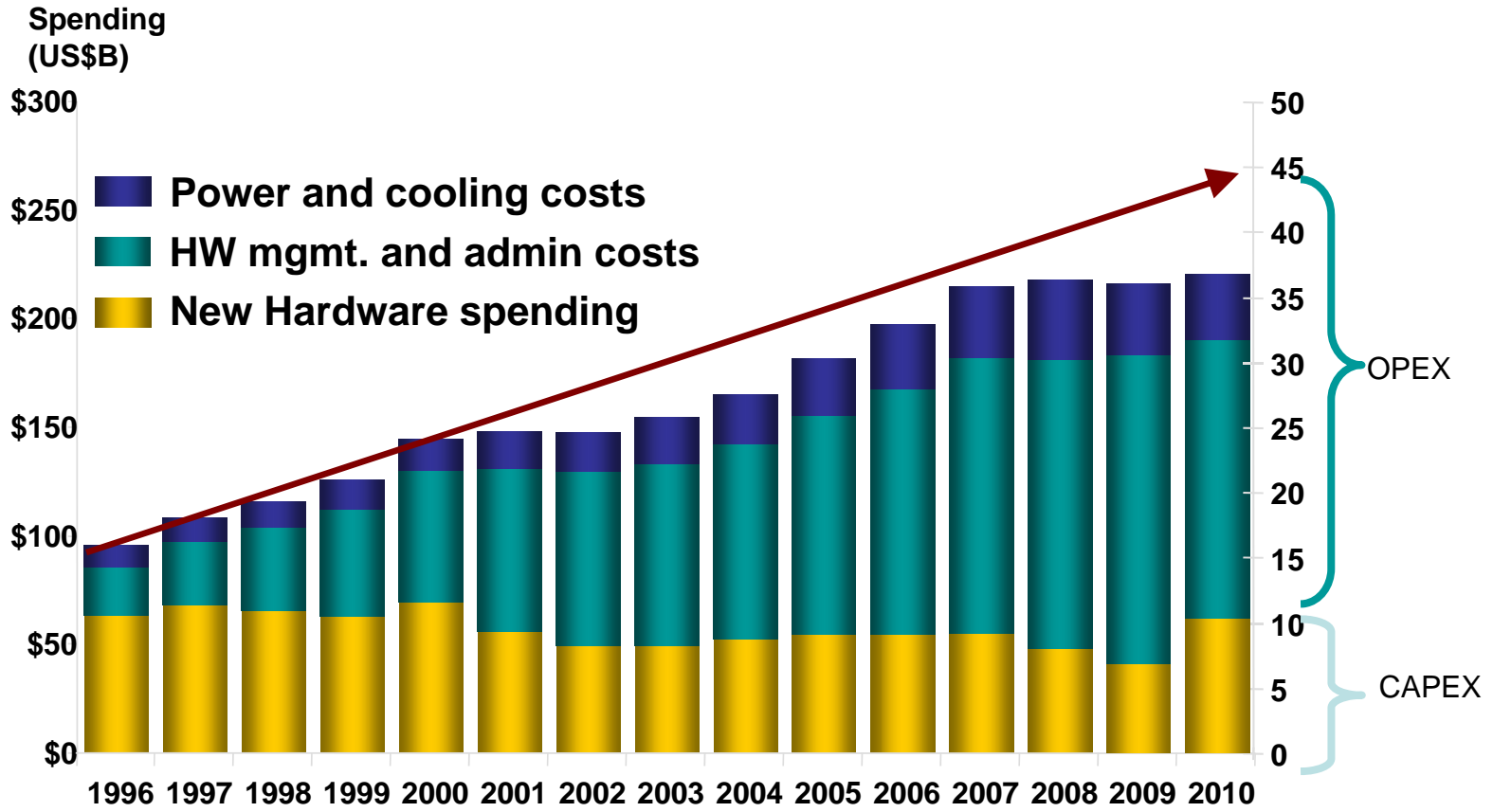
Virtualization, dynamic provisioning, archive address operational costs

Exploding Data Growth

Compounded Annual Growth Rate of 66%

(Exabytes)
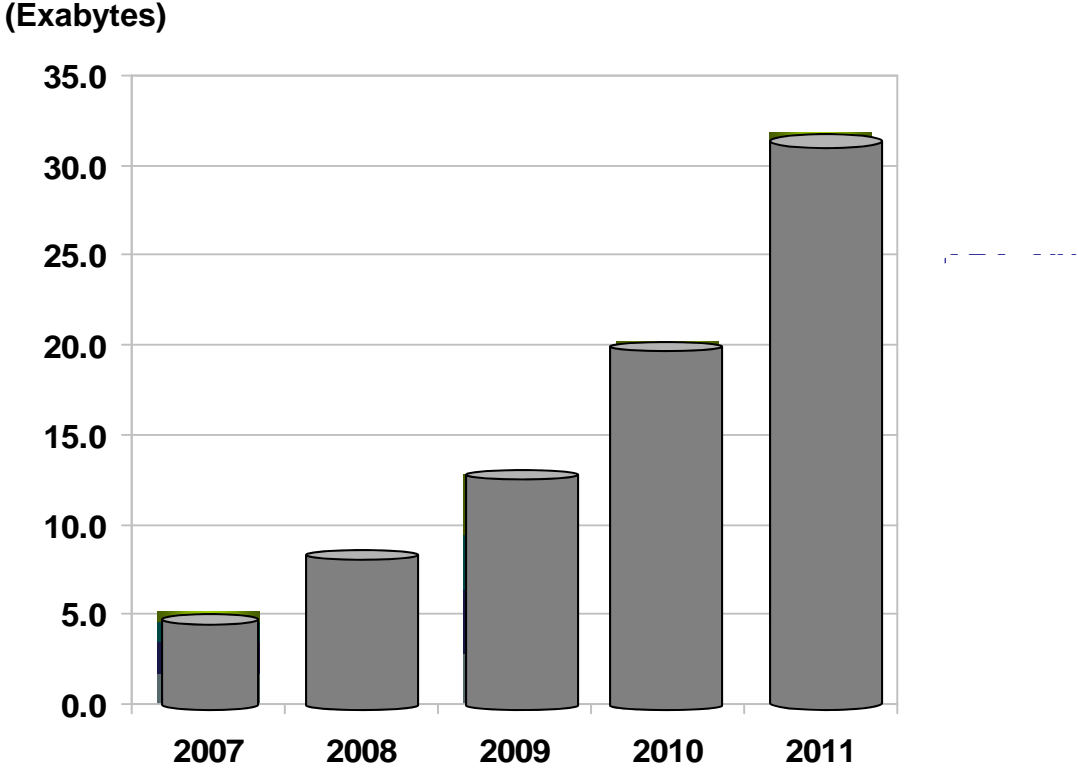
Understanding the types of data will help to address costs

Source: IDC 2008

What We Are Keeping: Changing Enterprise Data Profile

Consumption of Enterprise Disk Capacity by Type

(EB)

CAGR

■ Content Depots

■ Unstructured data

■ Replicated data

■ Traditional Structured data

121.1%

63.7%

43.9%

32.3%

2007 2008 2009 2010 2011
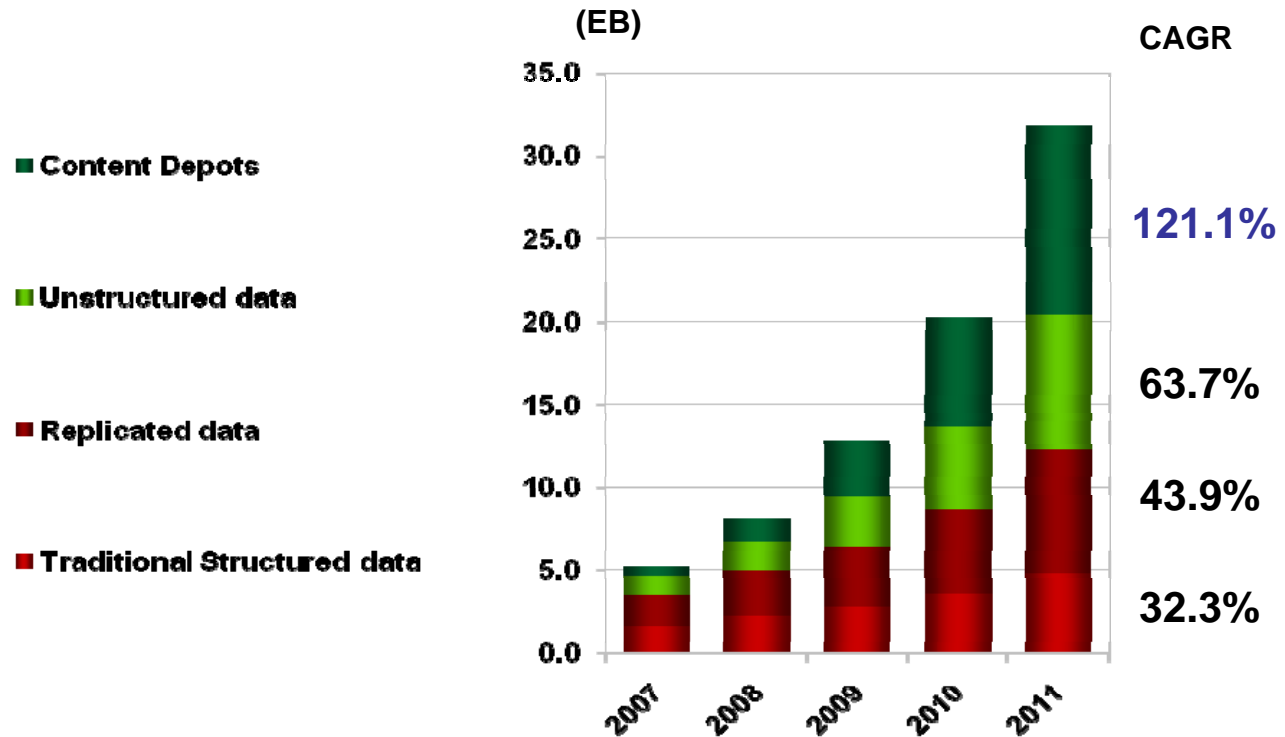
- Content data and Unstructured data are growing the fastest
- Replicas of data do not have to be on the same tier of storage as primary data
- Majority of the data does not need tier 1 storage
- Tiering of data and archive will play a major role in controlling cost

Source: IDC 2008

Servers and Networks Are Scaling Up

- Processor technology is scaling up
  - Multi-core – Dual, Quad, 8 way
  - SMT – Simultaneous Multi Thread
  - L1, L2, L3 cache

- Virtual Servers and hypervisors
  - Multiple Virtual Machines per physical server
  - Multiple VM I/O streams per physical server
    - Increasing I/O load
    - Increasing randomness of I/O

- Network bandwidth is increasing
  - FC 4Gbs going to 8Gbs
  - FCoE 10 Gbs to 40 Gbs to 100 Gbs
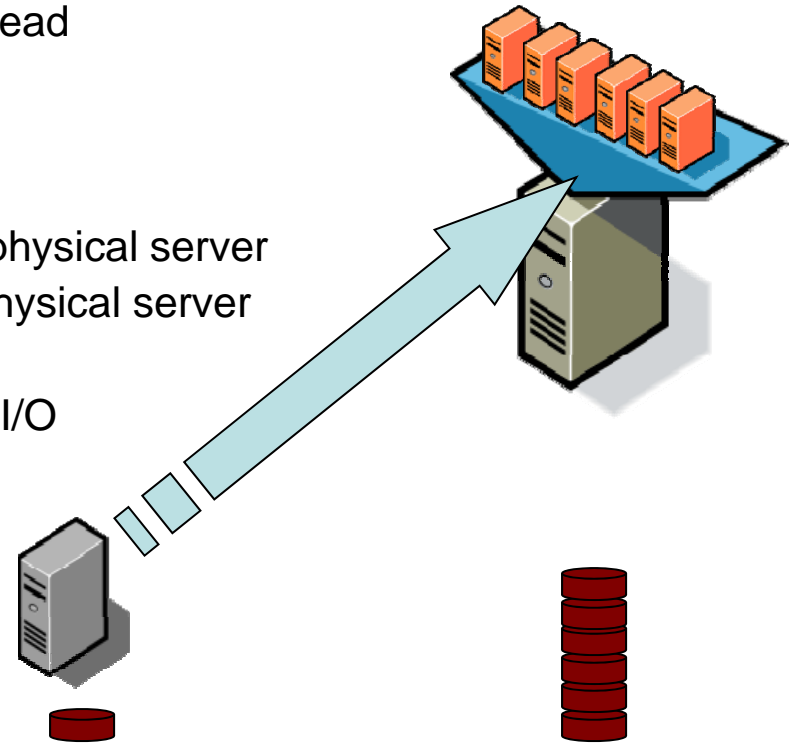
__Scale-up Storage is required to support virtual servers and networks__

Scale Out Storage
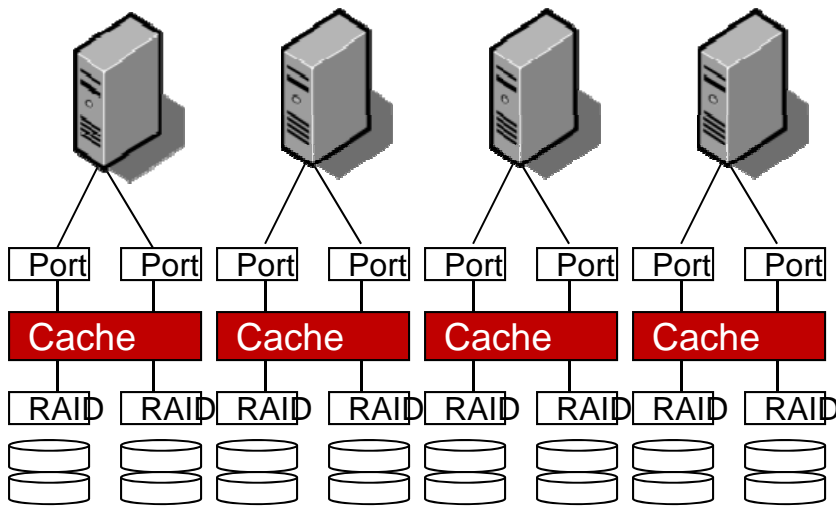
- **Scale out is loose coupling with external Switches**
- **Resources in one node can not be used to increase the resources in another node**
  - **Additional management required to distribute workload**
  - **Availability comes from active/passive redundancy**

Scale Up Storage

Port	Port	Port	Port	Port	Port	Port	Port

Cache

RAID	RAID	RAID	RAID	RAID	RAID	RAID	RAID
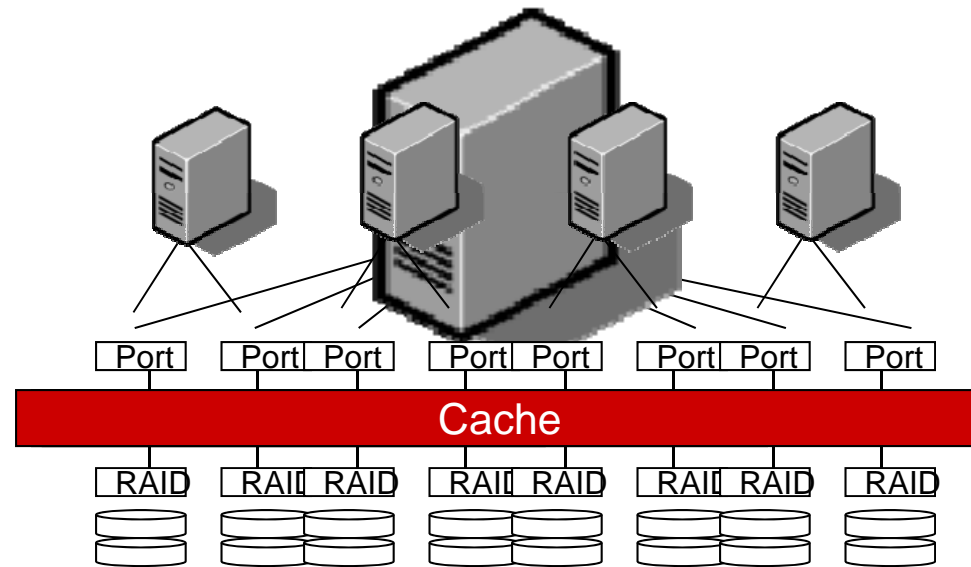
- **Scale Up is tight coupling through a Global Cache**
  - **Resources can be pooled together to give maximum performance and capacity**
  - **Scale up can be partitioned to work like scale out**

Cartesian Scaling

- **Two Dimensional scaling**
  - Increase the number of port processors to scale out
  - Increase Global cache to tightly couple port processors
  - Return on Investment

Amount of Global Cache
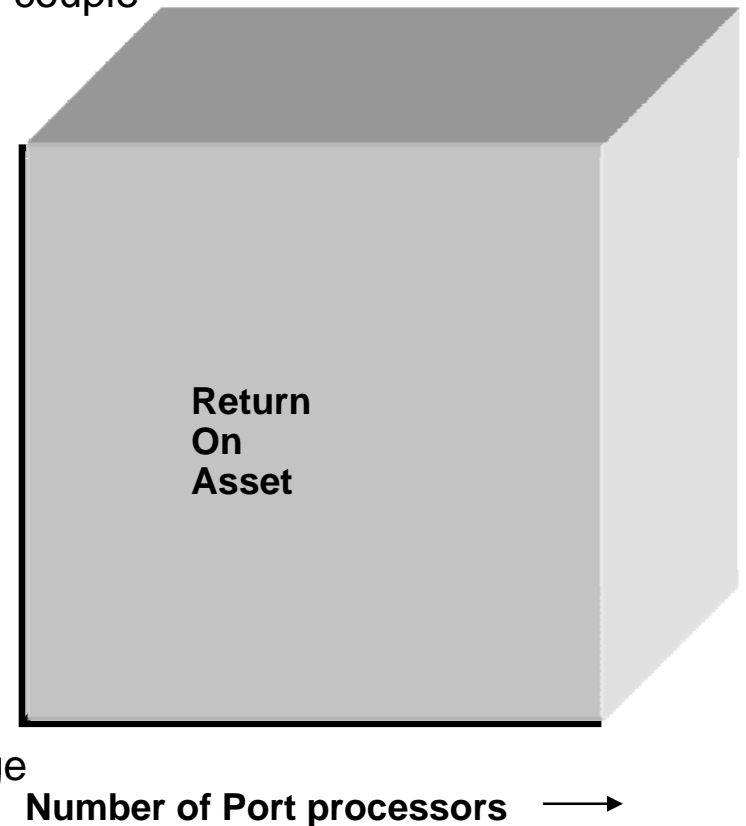
Return On Asset

- **Three Dimensional Scaling**
  - Virtualization of External Storage
  - Return on Asset

Number of Port processors

Scale Up Existing Storage With USP V

Consolidation with Virtual servers
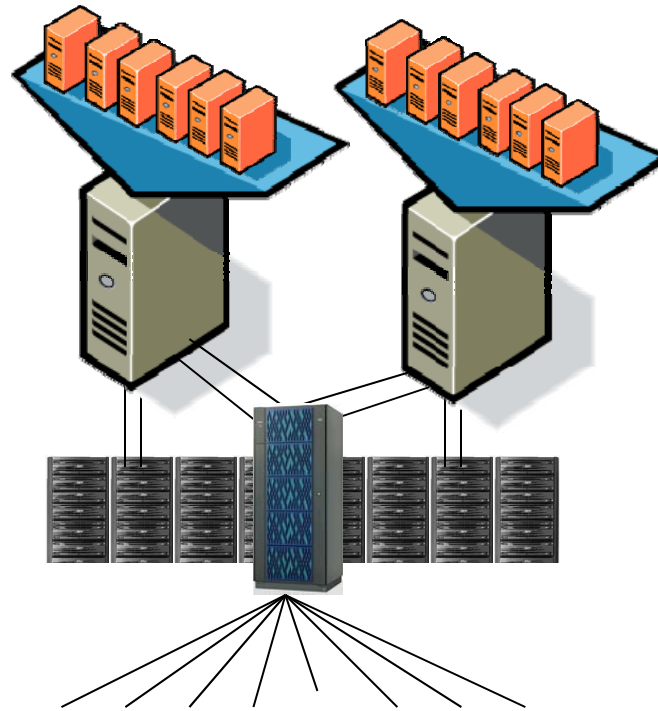
**Virtual Servers become the new Mainframe
Scale up Storage is required**

Scale Out For Data Migration

No more messy data migrations and painful "downtime"

Service disruption today

USP V — Small service disruption

USP V — NO service disruption

USP V | USP Vx — NO service disruption
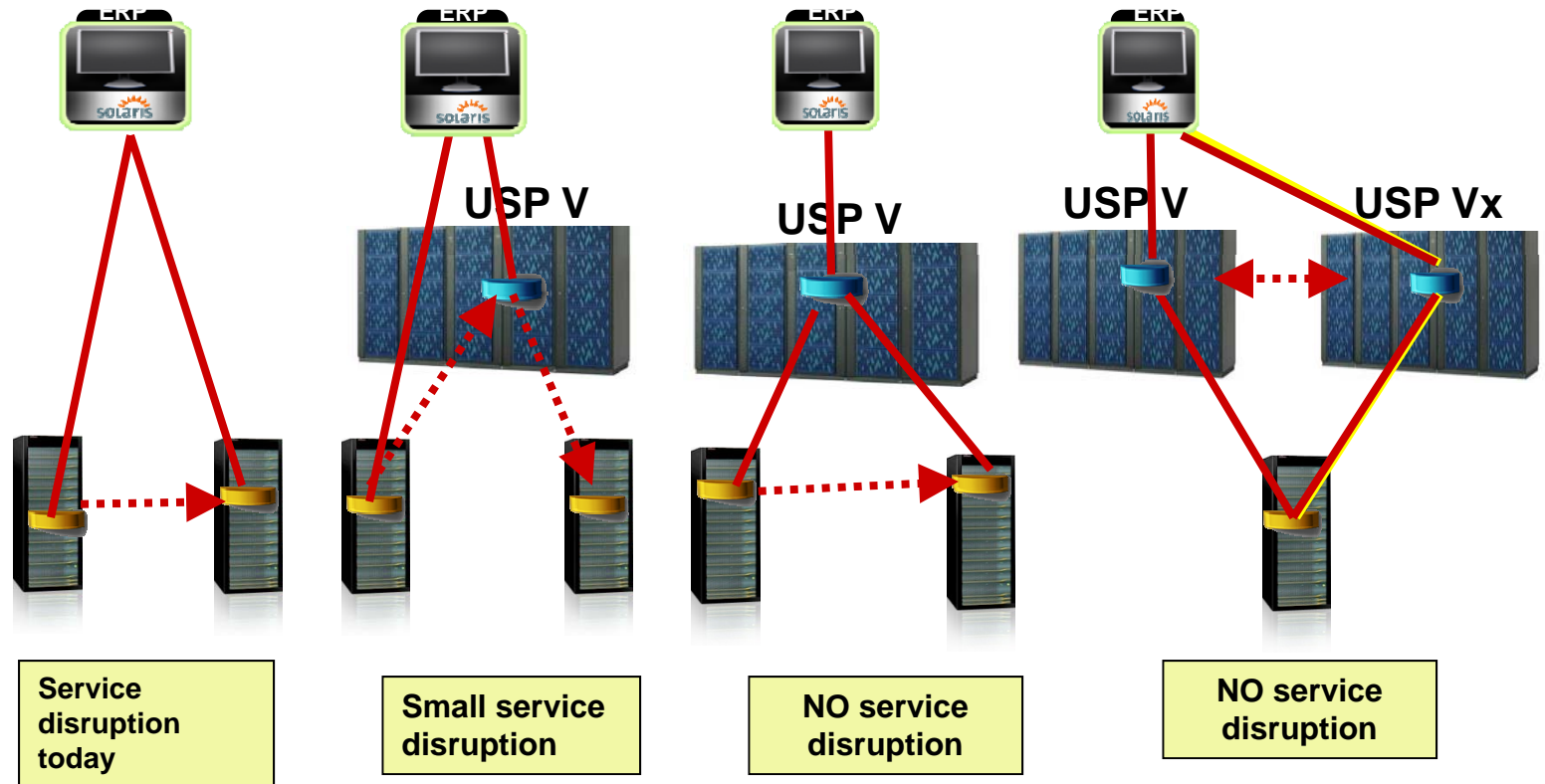
The USP V Can Scale out with Virtualization and Loose Coupling

SNIA Definition of Storage Virtualization

http://www.snia.org/education/dictionary/v/

The act of abstracting, hiding, or isolating the internal functions of a storage system or service from the application, host computer, or general network resources, for the purpose of **enabling application and network** <u>**independent**</u> **management of storage or data.**

The application of virtualization to storage services or devices for the purposes of aggregating functions or devices, hiding complexity, or **adding new capabilities to lower level storage resources**.

Virtualized vs Non-Virtualized Storage Services

Virtualized Storage Services
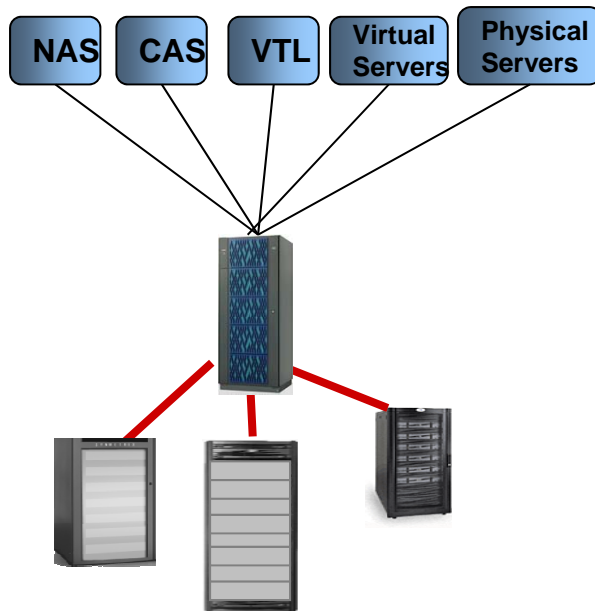
NAS　CAS　VTL　Virtual Servers　Physical Servers
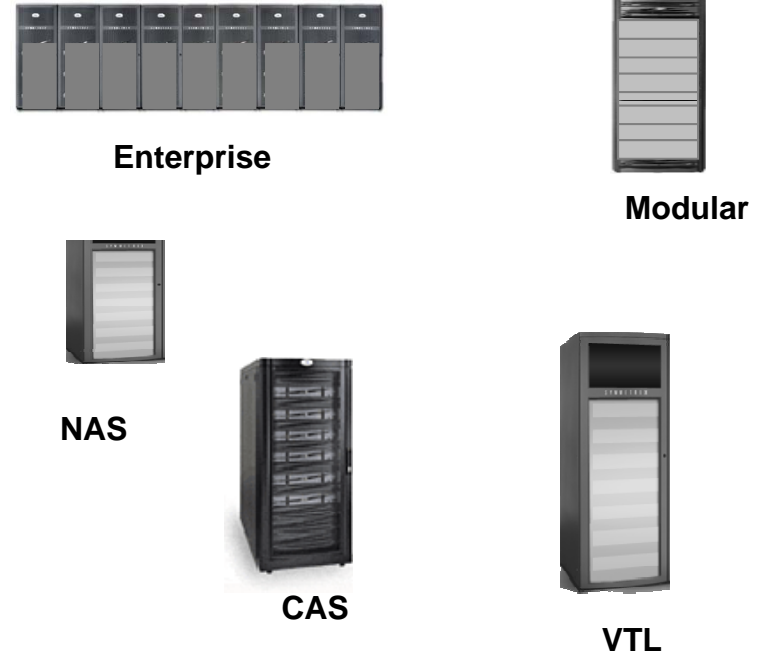
Non-Virtualized Storage services

Enterprise

Modular

NAS

CAS

VTL

Dynamic Provisioning

- **Virtualize Devices into a pool of capacity and allocate by pages**
- **Eliminate allocated but unused waste by allocating only the pages that are used**
- **Dynamically provision new servers in seconds**
- **Extend Dynamic Provisioning to external virtualized storage**
- **Convert Fat volumes into thin volumes by moving then into the pool**
- **Optimize Storage performance by spreading the I/O across more arms**
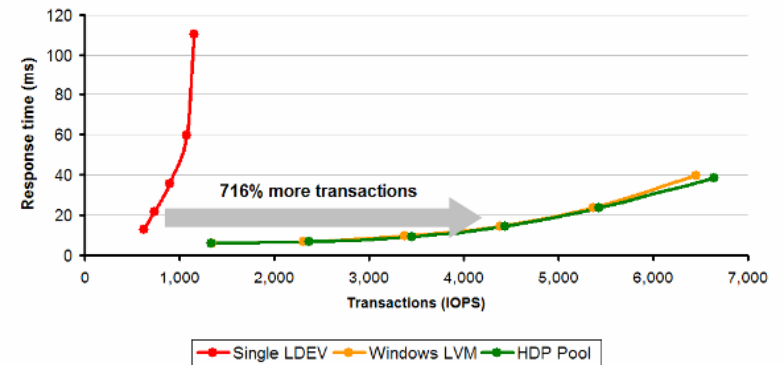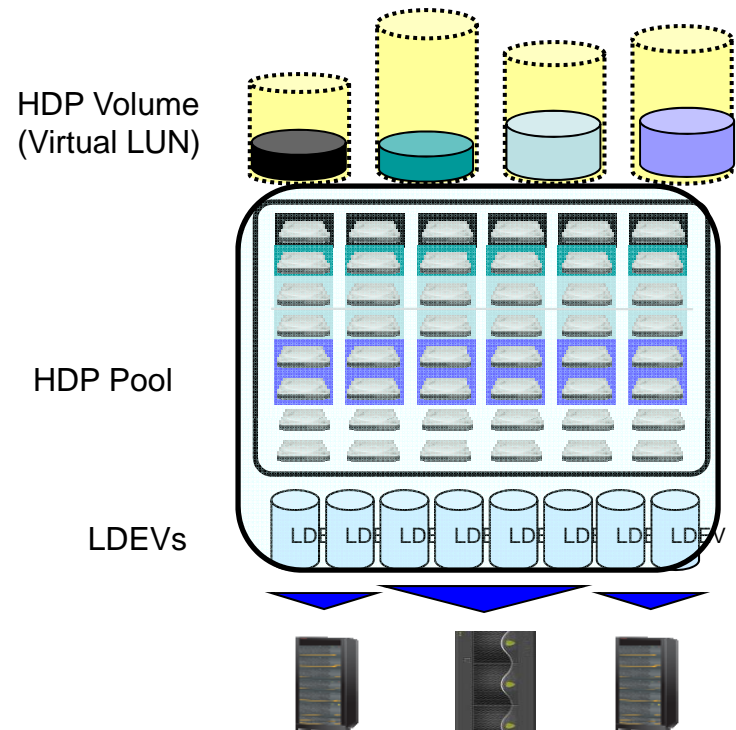
Flash Disks

- 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 dynamic tiering

- Eliminate waste of allocated unused Flash capacity with Dynamic Provisioning
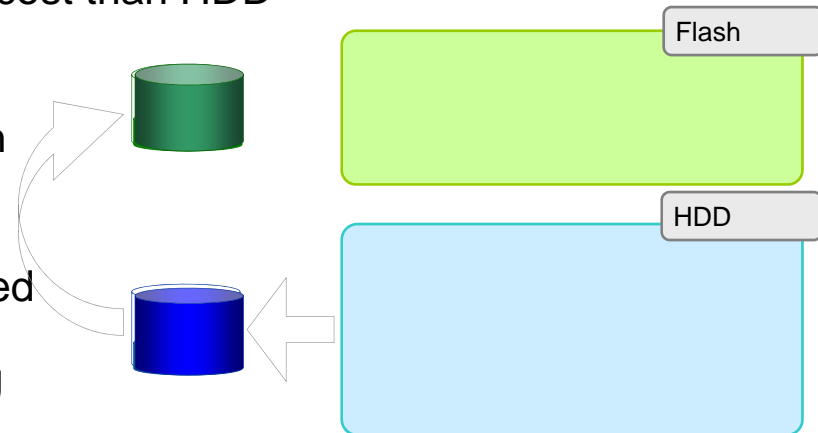
Flash

HDD

Convergence of File, Content and Block Virtualization

File/content Services Leverage Advanced Block Services Architecture

- **Global Name Space**

- **Intelligent Tiering Management**

- **Push files down to a content archive**

- **Tiering Through Virtualization**

NAS

Archive   Enterprise   Modular

Heterogeneous Vendor disks
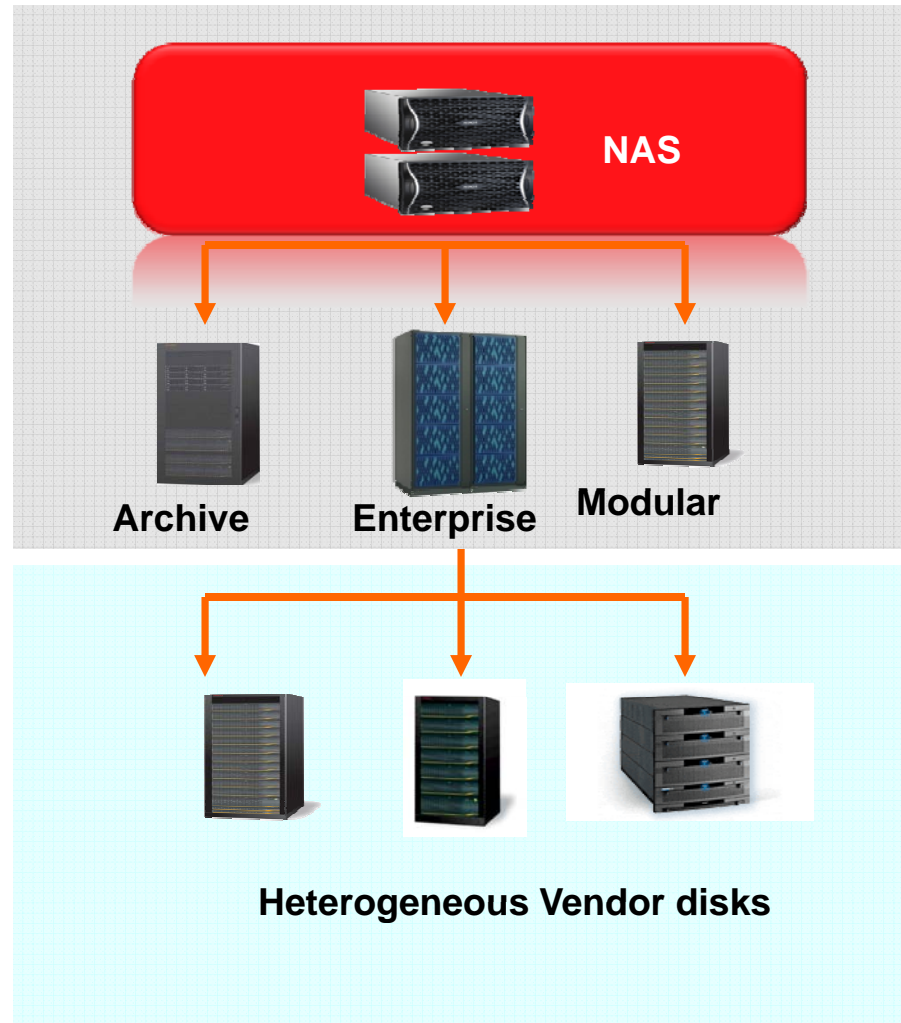
Automated Data Mobility

- **Two layers of data mobility**
  - **Infrastructure layer / block level**
  - **Object layer / file level**
- **Use virtualization and data mobility tools to move (or re-provision) the volume without disruption to any other pool or tier based on**
  - Promotion or demotion
  - Consolidation or migration
  - SLO, performance or cost change
- **Automated with policy based management**
  - Based upon pre-set SLAs

Tier 0

Tier 1

Tier 2

Infrastructure /block level

Object /file level

Data Management General Use Case:
File HSM (ILM)

/share
- file001
- file002
- file003
- file004
- file005

file001
file002
file003
file004
file005

ObjectStore
- file002
- file004
- file005
- file006
- file008
- file010
- file011

Active

NAS

Active
- file005
- file006
- file008

- file002
- file004
- file010
- file011

Tier 1   Tier 2   .mp3

Inactive

Archive

Tier 3

File Aware Data Movement

Move all MP3 to SATA tier

Content Aware Data Movement

Move all PDF older than 90 days and containing word "Budget" to Archive

Files from multiple NAS filers can be stubbed out to Archive

Virtualization is a Key Enabler for the Cloud

"Cloud is a way of using technology, not a technology in itself – it's a self-service, on-demand pay-per-use model. Consolidation, virtualization and automation strategies will be the catalysts behind Cloud adoption."

– The 451 Group

Key characteristics of the Cloud are:

• The ability to scale and provision dynamically in a cost efficient way

• The ability to make the most of that power without having to manage the complexity of the underlying technology

"The Cloud architecture can be private (hosted within an organization's firewall) or public (hosted on the internet)."

– Open Cloud Manifesto

Cloud architectures have public, private and hybrid use cases

Hitachi Data Systems

Summary

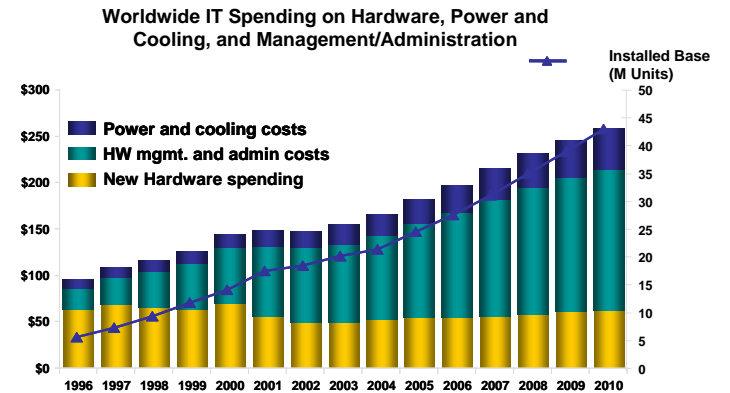- **While hardware costs remain flat, IT costs are escalating due to operational costs**

Worldwide IT Spending on Hardware, Power and Cooling, and Management/Administration

Installed Base (M Units)

- Power and cooling costs
- HW mgmt. and admin costs
- New Hardware spending

1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

IDC 2008
Vernon Turner

- **Virtualization can address operational costs through virtualization and scale up architectures**

COMPUTERWORLD

SNIA

SNW

April 12-15, 2010 | Rosen Shingle Creek Resort | Orlando, Florida