STORAGE GROWTH AND ETHERNET

Scott Kipp
September 12, 2011
What is an Exabyte? – 1 Million Terabyte Drives

- Earth created or replicated over 1,000 Exabytes of data in 2010 – that’s 143GB for each of 7 Billion people

<table>
<thead>
<tr>
<th>SI decimal prefixes – short scale</th>
<th>Binary usage</th>
<th>IEC binary prefixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Name (Symbol)</td>
<td>Value</td>
<td>Name (Symbol)</td>
</tr>
<tr>
<td>Thousand kilobyte (kB)</td>
<td>$10^3$</td>
<td>kibibyte (KiB)</td>
</tr>
<tr>
<td>Million megabyte (MB)</td>
<td>$10^6$</td>
<td>mebibyte (MiB)</td>
</tr>
<tr>
<td>Billion gigabyte (GB)</td>
<td>$10^9$</td>
<td>gibibyte (GiB)</td>
</tr>
<tr>
<td>Trillion terabyte (TB)</td>
<td>$10^{12}$</td>
<td>tebibyte (TiB)</td>
</tr>
<tr>
<td>Quadrillion petabyte (PB)</td>
<td>$10^{15}$</td>
<td>pebibyte (PiB)</td>
</tr>
<tr>
<td>Quintillion exabyte (EB)</td>
<td>$10^{18}$</td>
<td>exbibyte (EiB)</td>
</tr>
<tr>
<td>Sextillion zettabyte (ZB)</td>
<td>$10^{21}$</td>
<td>zebibyte (ZiB)</td>
</tr>
<tr>
<td>Septillion yottabyte (YB)</td>
<td>$10^{24}$</td>
<td>yobibyte (YiB)</td>
</tr>
<tr>
<td>Googol GoogolByte?</td>
<td>$10^{100}$</td>
<td></td>
</tr>
</tbody>
</table>

The world created over a ZB last year!
An Exabyte is not Infinite

1 TB Hard Disk Drive (HDD)
10 HDD – Just A Bunch of Disks (JBOD)
100 HDD – 1 Storage Subsystem
  - Controller + JBODs + Cache + IO Cards
1,000 HDD – 1 Row of Storage Subsystems
  At 1TB/HDD = 1PB per row
10,000 HDD – 1 Large Data Center
  With 10PB of Data
100,000 HDD – 100PB - Storage capacity of European Grid Infrastructure
1,000,000 TeraByte HDDs
  – 1 Exabyte

1 Storage Subsystem
with 72 Disk Drives
The Source of the Data

The Digital Universe Study

• Data is growing 40-50% per year –doubling every two years – compared to IP traffic growth of 30-40%

• 75% of the data is created by individuals, but enterprises have some liability for 80% of it
  • For data creation, think of computer files, music files, Digital Video Recorders, DVDs, backup drives, digital pictures...
  • They don’t explain the 80% number well, but I bet an example is that the cable company has liability for the shows on your DVR

• 25% of data is generated by machines and that is growing fast with sensors and remote monitoring

• Over the next decade, the number of servers (physical and virtual) will grow by a factor of 10, storage will grow by a factor of 50 and files will grow by a factor of 75

How much will it grow? Into the Zettabytes

1,000 Exabytes is a Zettabyte

• We create more digital data every couple of years than was created in history

• 500,000 Trillion files in 2011

Replication is the Great Multiplier
Look back at Andy Bach’s NYSE presentation

This little blue sliver of data is replicated about a hundred times to different organizations.

How many people make their own copy of the data within each organization?

CERN Case Study

CERN’s LHC generates 15 PB of data every year that is distributed over their core network with a 10Tbps capacity

Tier 0 storage at CERN – 15 Petabytes

Tier 1

150+ Petabytes
At 11 Tier 1 storage sites around the world

Tier 2

150+ Petabytes at 160 Tier 2 sites with partial sets of data

LHC = Large Hadron Collider
Transferring Large Data Sets – Big Data

- To transfer 15PB would take about:
  - 3.8 Years at 1GbE
  - 137.5 days at 10GbE
  - 13.75 days at 100GbE
  - 33 Hours at TbE

<table>
<thead>
<tr>
<th>Size of Data to Exchange</th>
<th>Latency of 1GbE</th>
<th>Latency of 10GbE</th>
<th>Latency of 40GbE</th>
<th>Latency of 100GbE</th>
<th>Latency of 400GbE</th>
<th>Latency of 1TbE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Gigabyte</td>
<td>8</td>
<td>0.8</td>
<td>0.2</td>
<td>0.08</td>
<td>0.02</td>
<td>0.008</td>
</tr>
<tr>
<td>10GB</td>
<td>80</td>
<td>8</td>
<td>2</td>
<td>0.8</td>
<td>0.2</td>
<td>0.08</td>
</tr>
<tr>
<td>100GB</td>
<td>800</td>
<td>80</td>
<td>20</td>
<td>8</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>1 Terabyte</td>
<td>8,000</td>
<td>800</td>
<td>200</td>
<td>80</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>1 Petabyte</td>
<td>8M</td>
<td>800000</td>
<td>200,000</td>
<td>80,000</td>
<td>20,000</td>
<td>8,000</td>
</tr>
<tr>
<td>10 PB</td>
<td>80M</td>
<td>8M</td>
<td>2M</td>
<td>800,000</td>
<td>200,000</td>
<td>80,000</td>
</tr>
<tr>
<td>100 PB</td>
<td>800M</td>
<td>80M</td>
<td>20M</td>
<td>8M</td>
<td>2M</td>
<td>800,000</td>
</tr>
</tbody>
</table>
Video Content – The Growth Component

• Cisco’s Visual Network Index (VNI)* predicts ~1 ZB of content will be distributed over Global IP networks in 2015 while there will be almost 8ZB of data produced and replicated that year

• Consumer video streaming is the main bandwidth driver in the future according to VNI*

• 1GB of content can produce 1PB** of data transfers so the storage component is one millionth compared to the networking component in some applications

Breaking IT Down into Information Technology (IT)

Users

Network

Servers

SAN – Storage Area Network

Storage

SAN – Storage Area Network

Tape Library – Racks and racks of tape drives with 1,000s of tapes

Storage Array – Think of a rack full of disks with memory cache

Email, Databases Applications

IT Server Team

Ethernet

FC

LAN

IT Network Team

IT Help Desk

IT Storage Team

9/26/2011
Explanation of Storage Access

DAS = Direct Attached Storage
NAS = Network Attached Storage
iSCSI – Internet Small Computer Systems Interface
LOM = LAN on Motherboard
NIC = Network Interface Card
HBA = Host Bus Adapter
CNA = Converged Network Adapter

SAN
Ethernet Switch

iSCSI Storage Arrays
Fibre Channel Over Ethernet Storage Arrays
Fibre Channel Storage Arrays
Fibre Channel Tape Library

Ethernet Switch

8G, 16G, 32G

Fibre Channel Switch

10G, 16G, 32G

Tape Library

9/26/2011
External Storage Sales in Exabytes

- 17EB in 2011 growing to 90 EB in 2015 – About 1% of the digital universe
- Ethernet-based storage expected to grow to over 50% of storage capacity in 2015
- NAS is usually unstructured data, but its supporting more applications

**NAS – Network Attached Storage**

An application server that serves files

- **High End NAS**
  - NAS – Think of a rack full of disks from with an Ethernet network interface

- **Low End NAS**
  - Can be as small as an individual disk drive or tens of disk in a 1U or 2U chassis

- Network connections:
  - Mostly 1GbE, moving to 10GbE and soon 40GbE

Users

- Ethernet
Fibre Channel Storage Area Network (SAN)

SCSI over Optical Fiber

Percent of servers connecting to FC Storage

15-25% of standard servers

40-50% of blade servers

80-90% of mainframes

Source: Brocade
iSCSI – SCSI over Ethernet

Still Two Networks

Users

Network

LAN

Server

iSCSI connections mostly 1GbE, moving to 10GbE and then 40GbE

Ethernet Switch

iSCSI SAN

iSCSI Storage Array – Think of a rack full of disks with an iSCSI network interface

iSCSI Tape Drives

Ethernet
FCoE SAN – Networked Storage Access
Mainly for Top of Rack deployments

FCoE Switch with Data Center Bridging

10GbE and eventually 40GbE

FCoE Storage

Fibre Channel SAN

SAN A
8G, 16G, 32G

Rack of Servers

SAN B

LAN

10GbE and eventually 40GbE

Ethernet

FC
Trends in Storage

• Application migration benefits from networked storage compared to DAS
• Cloud computing requires major data moves
• Virtual Desktop Infrastructure (VDI) leads to centralized storage and increased network traffic
• Solid State Drives (SSDs) or Flash Storage leads to higher bandwidth demands on the network
Server Virtualization and Migration

Data needs to move with the application

For application A to move to a new server with DAS, the data has to be moved over the LAN.

In a SAN, the data stays still, but access to the data changes via SAN configuration.
Data in Cloud Computing in 2015
Over 10% of storage could be in the clouds!

THE DIGITAL UNIVERSE AND THE PUBLIC CLOUD, 2015

0.8ZB
Cloud Services

1.4ZB
“Touched” by the Cloud

7.9ZB
Total Digital Universe

9/26/2011
Cloud Computing

• Cloud Computing offers the grand vision of hosting and scaling applications from your data center to the cloud provider or another data center on demand.

• To enable this transition, the data needs to be exchanged or mirrored first.
Data Mirroring Between Storage Arrays

Application A data needs to be mirrored before the application can move.

Primary Data Center

Cloud Provider or Secondary Data Center

Router  WDM  Router

Server  WDM  SAN Switch

A  A  A  A

Ethernet  FC

9/26/2011
3 Main Ways to Mirror Data over the WAN

1. Native Fibre Channel, Fibre Channel Internet Protocol (FCIP) or Ethernet Over WDM
2. SAN Switch to SAN Switch via IP and FCIP
3. Server to Server that backs the data up to storage
Virtual Desktop Infrastructure (VDI) Architecture

- VDI enables centralized management and simple upgrades to software and applications and increases LAN traffic.

Users on Virtual Desktops

Centralized Virtual Desktops

Linked Clones

SAN

LAN
# Comparing Server Technologies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU</strong></td>
<td>1 x Pentium 4 1.5 GHz</td>
<td>5 x Pentium D 2.6 GHz</td>
<td>15 x Nehalem Quad 2.6 GHz</td>
<td>45 x? Haswell 2.6 GHz?</td>
</tr>
<tr>
<td><strong>DRAM</strong></td>
<td>1 x DDR1</td>
<td>4 x DDR2</td>
<td>8 x DDR3</td>
<td>32 x? DDR4?</td>
</tr>
<tr>
<td><strong>Network</strong></td>
<td>1 x 100Mb Ethernet</td>
<td>10 x Gigabit Ethernet</td>
<td>100 x 10 Gigabit Ethernet</td>
<td>400 x 40 Gigabit Ethernet</td>
</tr>
<tr>
<td><strong>Bus</strong></td>
<td>1 x PCI 32-bit/33 MHz</td>
<td>15 x PCIe Gen1 x8</td>
<td>30 x PCIe Gen2 x8</td>
<td>60 x PCIe Gen3 x8</td>
</tr>
<tr>
<td><strong>Fibre Channel</strong></td>
<td>1 x 1GFC</td>
<td>4 x 4GFC</td>
<td>8 x 8GFC</td>
<td>32 x 32GFC</td>
</tr>
<tr>
<td><strong>Disk</strong></td>
<td>1 x 15K rpm hard drive</td>
<td>1 x 15K rpm hard drive</td>
<td>1 x 15K rpm hard drive</td>
<td>1 x 15K rpm hard drive</td>
</tr>
</tbody>
</table>

Source: Nimbus and Brocade
SSDs – Solid State Drives

- Application performance is limited by multiple factors with disk drive latency being one factor.
- Order of magnitude improvements in performance:
  - While traditional spinning disk drive seek times are in the millisecond range, SSD seek times are in the microsecond range.
  - SSDs often referred to as Tier-0 storage while disk drives are Tier-1.
  - Capacities in the hundreds of GBs per drive.
  - Very energy efficient compared to spinning disks.
  - Most SSDs provide over 50,000 IOPs per drive.
- One flash storage system supports 500,000 IOPS and 8 GBps (64 Gbps) of throughput.

<table>
<thead>
<tr>
<th></th>
<th>Latency</th>
<th>Drive IOPs</th>
<th>Array IOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDD</td>
<td>2-10 mS</td>
<td>100-300</td>
<td>400-40,000</td>
</tr>
<tr>
<td>SSD</td>
<td>50-250 uS*</td>
<td>40k-150k</td>
<td>50k-500k</td>
</tr>
</tbody>
</table>

* This is based on Flash memory and multiple parallel processing.
Conclusion

• We entered the Zettabyte era last year – 1M TB/year of new data

• More data is created every two years than all previous years combined

• Virtualization causes the need for networked storage of all varieties (SAN and NAS)

• All storage technologies are improving except disk drive access times and disk rotational speeds

• New applications and devices are driving more data access and higher bandwidths
Thank You