NearStore® Storage Efficiency

Examining Space and Power Efficiency in Disk-to-Disk Data Protection

Larry Freeman, NetApp

TR-3539

Overview

Data center resources have become a precious commodity. IT managers are carefully analyzing each piece of equipment entering the data center, scrutinizing the power, heat, cooling, and space requirements vs. the operational benefits of each device.

One factor behind this movement to economize data center resources is today's high-density servers. The widespread need for a variety of applications has necessitated the implementation of larger quantities of servers to provide database, e-mail, Internet, and file services and many other functions within and outside the enterprise organization. Server vendors have responded to this demand by producing products with more processing power and smaller footprints. However, these servers, while providing far greater computational ability than their predecessors, are unfortunately also consuming alarming amounts of energy as the available rack space in data centers becomes more and more dense.

In conjunction with the growth of application servers, the thirst for larger and faster data storage volumes in the data center is equally demanding. Ever-expanding data must be available for quick access by these servers, and still more resources are consumed protecting this data. This combination of the growth of servers and associated primary and secondary data storage is compounding the resource provisioning currently being juggled by data center managers.

The "megawatt" enterprise data center is today's reality. What does a megawatt data center look like? On average, the typical enterprise data center houses approximately 300 UNIX® and Windows® servers within a 3,000- to 5,000-square-foot area. According to a recent paper from APC, *"A typical value for the cost of electrical power is $0.12 per kW hr. Given this cost, the annual electrical cost per kW of IT load is approximately $1,000."* If we apply APC's estimated costs to our megawatt data center, the energy costs alone are $1 million per year! Clearly, conserving space and power in the data center has data center managers thinking more about floor tiles and watts—and less about bits and bytes!

Typical Enterprise "Megawatt" Datacenter

Equipment Type	Floor Space (sq ft)	% of Total	Power (watts)	% of Total
300 High Density Rack Servers [1]	400	40.0%	240,000	24.0%
5 Unix Enterprise Servers[2]	100	10.0%	150,000	15.0%
Primary and Secondary Disk Storage [3]	100	10.0%	75,000	7.5%
Networking / SAN Switches	60	6.0%	20,000	2.0%
Tape Archival	40	4.0%	15,000	1.5%
HVAC System / UPS System	300	30.0%	500,000	50.0%
Total Requirements	**1000**	**100%**	**1,000,000**	**100%**

Source: Manufacturers' Data

[1] i.e. Dell PowerEdge 2950

[2] i.e Sun Fire E25K

[3] Generic 100 TB FC primary storage array + 200TB SATA secondary storage array

Data Storage Space and Power Consumption

While the servers will continue to be the power hogs of the data center, data storage equipment is the next likely offender, consuming approximately **15%** of data center resources when their power requirements are combined with the HVAC and UPS energy required to maintain these storage arrays.

Storage vendors can optimize data center resources through efficient use of power and space in the design of their products. The real benefits are:

- Reduction in the amount of physical space needed for data storage
- Reduction in the amount of power/heat and cooling requirements for data storage

Implementation of ”Storage Efficient” products translate into:

- Additional resources that can be reprovisioned by servers
- Reduced operational costs (i.e., energy savings) within the data center

Recently, data center managers have taken advantage of a new classification of low-cost disk arrays (led by the NetApp™ NearStore platform) to improve the speed and reliability of data backups when compared to traditional tape backups. The trend towards disk-to-disk backups is a double-edged sword. While it provides a valuable function, it also further accelerates power consumption and decreased server space within the data center.

Growth of Secondary Data Storage

Traditionally, low-cost disk storage arrays were implemented to store small, disk-based “snapshot” copies of primary file systems. Snapshot™ technology, pioneered by NetApp in the 1990s, was valued for its ability to quickly and easily recover lost files. Another key benefit of early Snapshot technology was that the snapshot only contained the changes made to an active file system, and thus required a relatively small amount of secondary storage space.

As the use of secondary disk storage for data protection expanded, system administrators began to maintain an entire “baseline” copy of their file system data on secondary disks along with regular updates —similar to the method of storing full and incremental data backups to tape. This led to maintaining several copies of primary data on secondary disks and the desire by data center managers to use disks for all data backups on a regular basis, with periodic archival on tape. The advent of low-cost disk storage arrays made this possible, but further added to the consumption of space and energy resources in the data center.

Recognizing this, NetApp has played a vital role in reducing the burden typically associated with adding more disks into the data center for secondary storage. The remainder of this paper will focus on how the NetApp NearStore platform can help make the enterprise data center a more “storage efficient” environment.

“Storage Efficiency” Criteria

When integrating disk-to-disk secondary storage products into the data center, there are two key elements that together constitute **storage efficiency**:

1. Is the secondary storage unit designed with **storage capacity optimization**, reducing the dependency on high power (and low storage density) primary storage? Does the secondary storage unit reduce the required number of disk drives and thus provide the greatest space-to-capacity utilization?

2. Does the secondary storage unit have provisions for **optimal power efficiency** without sacrificing reliability?

Appendix B – Typical NearStore Use Case Examples

Use Case #1 - File Archival
File Archival To Secondary Storage With A-SIS Deduplication

Data Environment	
Amount of User Data Replicated	50 TB
NearStore on FAS Platform	FAS3020
Disk Interface	SATA
Capacity Per Disk	500GB
RAID Level	RAID-DP
Disk-Based Backups via SnapVault?	No
Remote Office Replication via SnapMirror?	No
Automatic Data Migration via IS1200?	No
Cloned Data Sets vis FlexClone?	No
A-SIS Deduplication Enabled?	Yes

Qualifying Information	
Amount of Data Replicated	N/A
% of Primary Data Migrated	N/A
Capacity of Cloned Data	N/A
Deduplication Reduction Percentage	50%

NearStore Storage Efficiency	
NearStore Storage Capacity	25 TB
Effective Storage Capacity	50 TB
# of Disk Drives Required	58
Watts Required	1,750 watts
Floorspace Required	6.13 sq ft
NearStore Power Efficiency	35 watts per TB
NearStore Cooling Efficiency	119 BTH/h per TB
NearStore Space Efficiency	8 TB's per sq ft

Use Case #2 - D2D Backup
Disk-to-Disk Backup Using SV-NBU and Deduplication

Data Environment	
Amount of User Data	100 TB
NearStore on FAS Platform	FAS3050
Disk Interface	SATA
Capacity Per Disk	500GB
RAID Level	RAID-DP
Disk-Based Backups via SnapVault?	Yes
Remote Office Replication via SnapMirror?	No
Automatic Data Migration via IS1200?	No
Cloned Data Sets vis FlexClone?	No
A-SIS Deduplication Enabled?	Yes

Qualifying Information	
Amount of Data Replicated	N/A
% of Primary Data Migrated	N/A
Capacity of Cloned Data	N/A
Deduplication Reduction Percentage	30%

NearStore Storage Efficiency	
NearStore Storage Capacity	77 TB
Effective Storage Capacity	110 TB
# of Disk Drives Required	176
Watts Required	4,150 watts
Floorspace Required	12.26 sq ft
NearStore Power Efficiency	38 watts per TB
NearStore Cooling Efficiency	129 BTH/h per TB
NearStore Space Efficiency	9 TB's per sq ft

Use Case #3 - D2D Backup, Replication, and File Migration
Disk-to-Disk Backup Using SV-NBU and Deduplication
Remote Office Replication Using SnapMirror
Automatic Migration of Inactive Files Using IS1200

Data Environment	
Amount of User Data	100 TB
NearStore on FAS Platform	FAS3070
Disk Interface	SATA
Capacity Per Disk	500GB
RAID Level	RAID-DP
Disk-Based Backups via SnapVault?	Yes
Remote Office Replication via SnapMirror?	Yes
Automatic Data Migration via IS1200?	Yes
Cloned Data Sets vis FlexClone?	No
A-SIS Deduplication Enabled?	Yes

Qualifying Information	
Amount of Data Replicated	10 TB
% of Primary Data Migrated	20%
Capacity Reserve for Cloned Data	N/A
Deduplication Reduction Percentage	30%

NearStore Storage Efficiency	
NearStore Storage Capacity	107 TB
Effective Storage Capacity	140 TB
# of Disk Drives Required	214
Watts Required	5,760 watts
Floorspace Required	12.26 sq ft
NearStore Power Efficiency	41 watts per TB
NearStore Cooling Efficiency	140 BTH/h per TB
NearStore Space Efficiency	11 TB's per sq ft

Use Case #4 - D2D Backup, Replication, File Migration, and File Cloning
Disk-to-Disk Backup Using SV-NBU and Deduplication
Remote Office Replication Using SnapMirror
Automatic Migration of Inactive Files Using IS1200
Cloned Storage Space For Test Environment With Flex Clone

Data Environment	
Amount of User Data	100 TB
NearStore on FAS Platform	FAS6030
Disk Interface	SATA
Capacity Per Disk	500GB
RAID Level	RAID-DP
Disk-Based Backups via SnapVault?	Yes
Remote Office Replication via SnapMirror?	Yes
Automatic Data Migration via IS1200?	Yes
Cloned Data Sets vis FlexClone?	Yes
A-SIS Deduplication Enabled?	Yes

Qualifying Information	
Amount of Data Replicated	10 TB
% of Primary Data Migrated	20%
Capacity of Cloned Data	5 TB
Deduplication Reduction Percentage	30%

NearStore Storage Efficiency	
NearStore Storage Capacity	112 Tb
Effective Storage Capacity	145 TB
# of Disk Drives Required	224
Watts Required	6,600 watts
Floorspace Required	12.26 sq ft
NearStore Power Efficiency	46 watts per TB
NearStore Cooling Efficiency	155 BTH/h per TB
NearStore Space Efficiency	12 TB's per sq ft