


DIY: Build Your Own Private or Hybrid Storage Cloud

Howard Marks
Chief Scientist

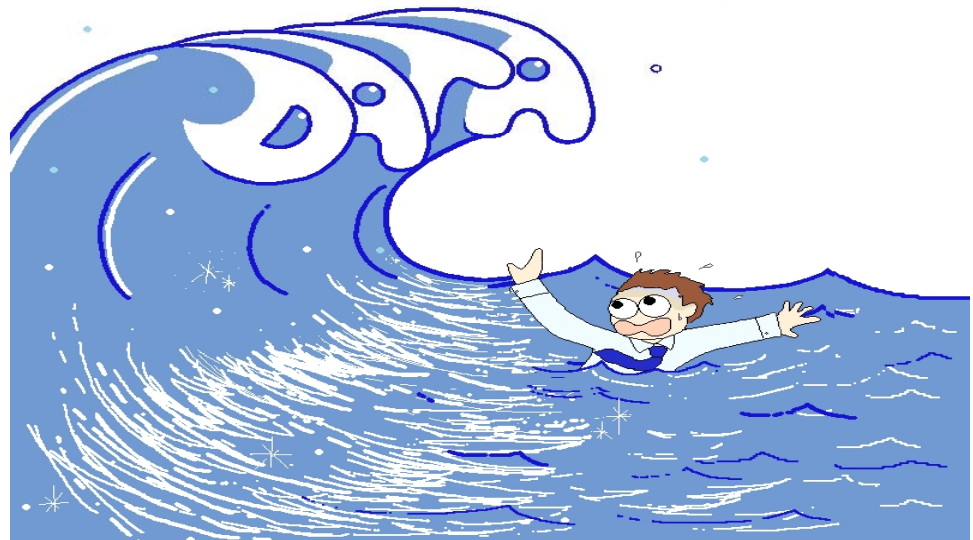


Where I Come From

- 25 years as a consultant in the midmarket
- From the school of been there, done that, broke that too
- Reviewing products for magazines since '87
- Concentrating on storage/servers this century
- Now running independent test lab/analyst firm  DeepStorage.net
- Twitter @DeepStoragenet Hmarks@deepstorage.net

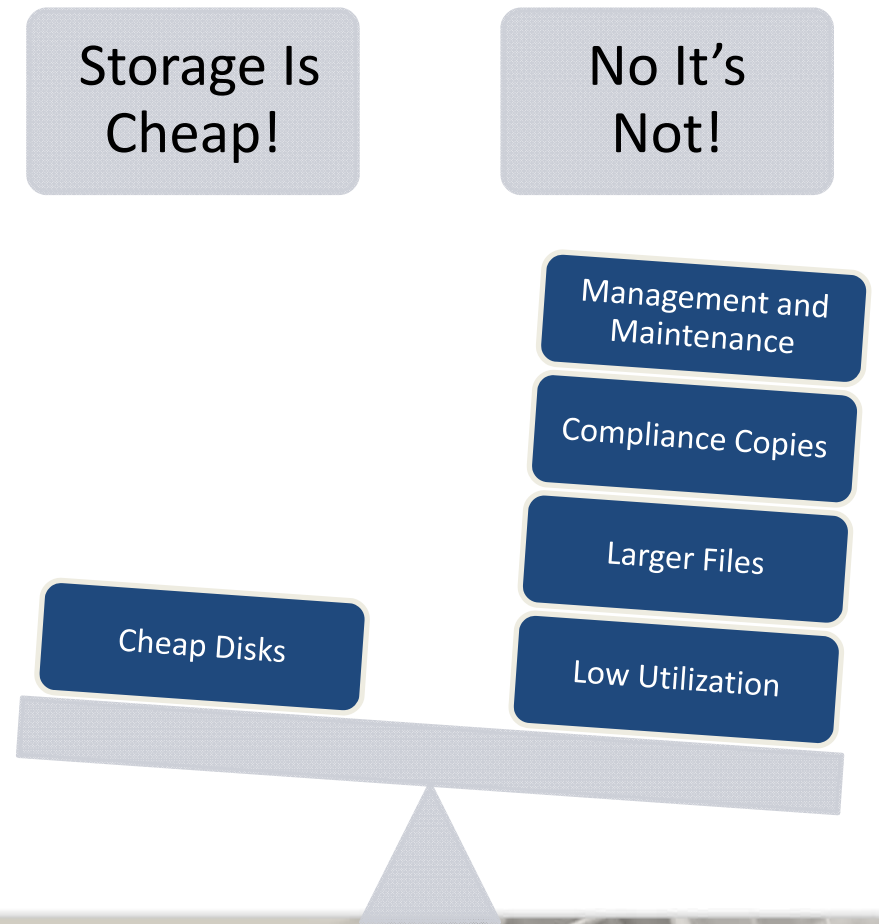
As If You Didn't Know

- **We're drowning in data**
 - Standard file systems breaking down
 - Too Small
 - 16TB-100TB
 - A few million files
 - 255 byte path



Storage is Supposed to Be Getting Cheaper!

- Disk cost is dropping rapidly
- \$250 buys:
 - 1994: 2 GB
 - 1999: 20 GB
 - 2004: 200 GB
 - 2009: 2000 GB
- **But enterprise storage costs keep rising!**



What is Cloud Storage

- Well it's what public cloud providers sell
 - Infinitely Elastic
 - Relatively low cost
 - S3 ~15¢/GB/Mo
 - Typically object interface
 - Internet accessible
 - Multi-tenant
 - You don't have to manage it!

So What's a Private Storage Cloud

- Massively scalable
 - Not really elastic
 - Step function for growth
 - Can't shrink
- Data protection by policy
 - Fault tolerance, copies, retention, etc.
 - Includes location
 - Store 2 copies in location 1 plus 1 in each of 3 other locations

What's Cloud Storage Good For?

- Reduced TCO through reduced management
 - Includes reduced backup
- Large data stores
- Low change rates
 - Especially of individual objects
- Not latency sensitive
- Archives, rich data stores, etc.

An Advertising Agency

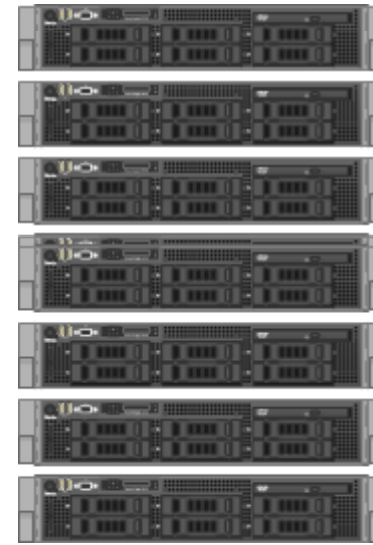
- Digital Asset Management System
 - Holds ads (print and video)
- Accessed by creative types in NY, Chicago, LA, SF
- Metadata and index in each office
- Media on cloud storage
 - Public cloud
 - Private cloud replicating across offices

Public or Private

- Public cloud more elastic and pay as you go
- Simplifies multi-site access
- But:
 - Network adds latency
 - Security concerns
 - But you could encrypt data before it leaves
 - Control concerns
 - International issues like privacy/Patriot Act

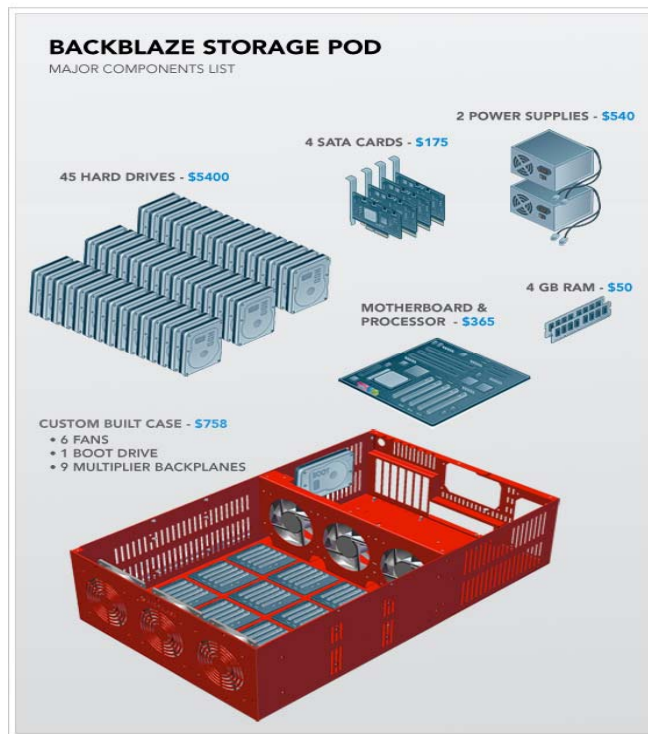
Typical Cloud Storage Architecture

- Redundant Array of Independent Nodes (RAIN)
 - All peers or
 - Access nodes and storage nodes
- Low cost x86 servers
- Low cost SATA DAS
- Object storage
 - Could include NAS access



Which is Your Idea of Low Cost Storage?

Backblaze Storage Pod



CLARiiON Full of 1TB Drives



The Difference is Philosophy

- Enterprise systems have high subsystem reliability
 - Use redundant components to make sub-systems fault tolerant
 - Subsystem failure creates a crisis
- Cloud systems accept node failures
 - Reliability comes from software and redundant data

File Systems and Object Stores

File System

- Limits:
 - Disk Capacity (16-100TB)
 - Path (255 char)
 - Number of files
 - Metadata
- Syntax:
 - Open(file)
 - Lock(2343,100)
 - Write(2343,"hello")
 - Close(file)

Object Store

- Store/retrieve file by URI/URL
- Usually has extended metadata:
 - Retention
 - Protection policy
- No limits
 - Path depth
 - Files/folder
 - Total files

Object Access

- Object stores use HTTP derived syntax
 - Put object to save
 - Get object to read
- New object replaces old one
- API usually based on SOAP or REST
- Content Addressable Storage
 - Special case of object store where URI=data hash
 - Inherent single instance storage

Data Protection Methods

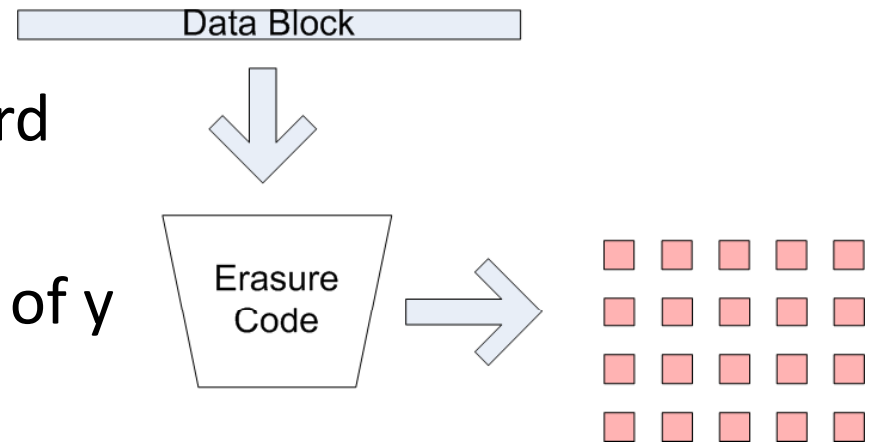
- Conventional RAID
- Object replication/dispersal
- Object replication with RAID
- Erasure and dispersal codes

Erasure Codes

- Beyond RAID for protection and integrity
- Usually based on Reed-Solomon math
- Can provide higher protection at lower overhead
 - EG: Survive 4 drive failures w/25% overhead
- Dispersal codes add location

How Erasure Codes Work

- Breaks data into blocks
- Blocks contain data and forward correction codes
- System can return data from x of y blocks
- If 10 of 15 *any* 10 blocks can reconstruct data
- Store each block on different disk/node
- Store 5 blocks each in 3 locations



Private Cloud Storage Products

- EMC Atmos
- Caringo Castor
 - SW only powers Dell Dx6000
- Hitachi Content Platform
- NetApp Storage Grid
 - Was Bycast
- DDN Web Object Store

Erasure Code Based Systems

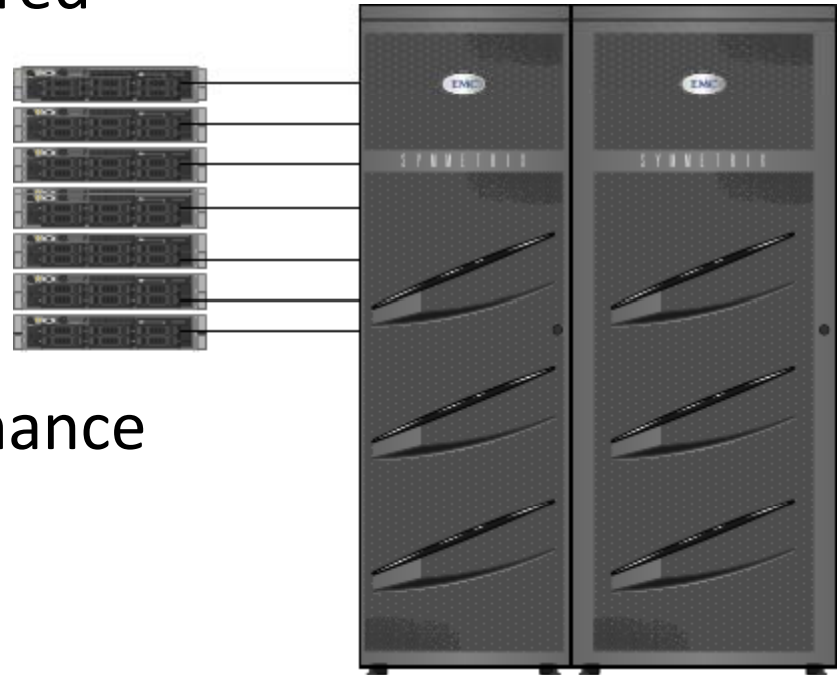
- OceanStore
 - Academic project at Berkeley
- Cleversafe
 - RAIN dispersal system
- Amplidata
 - RAIN dispersal system w/Atom storage nodes
- NEC HYDRAsstor
 - Deduplicating grid for backup/archive
 - No dispersal

Hybrid Cloud Options

- Any combination of on-premise and internet storage could be called hybrid cloud
- Models:
 - Cluster on-premise replicates to public provider
 - Atmos to Atmos (AT&T Synaptic Cloud)
 - On-premise replicates to colocation cluster
 - Gateway/archiving system writes to both
- Also used for dedicated infrastructure by public provider

Clustered NAS

- Cluster of NAS heads w/shared storage
- Can have file/policy based replication Etc.
- Can scale to 10+TB
- Low latency/higher performance
- Examples:
 - IBM SONAS
 - Symantec FileStore
 - Gluster



Scale-Out NAS

- NAS cluster with distributed storage
- Adding nodes adds performance and capacity
- Scales, performs like clustered NAS
- Simpler management
- Examples:
 - Panasas
 - EMC/Isilon
 - HP IBRIX



And Now It's Time to Play...



Stump the chump