Cloud Storage’s “Organic” or Living Evolution

Marc Staimer
Dragon Slayer Consulting
Marc Staimer - President & CDS

- 13+ years consulting
  - Storage, SANS, SW, Networks, Server, Data Centers, MSPs, Cloud Storage, Cloud Computing, Cloud BURR, & more
  - Consults vendors (> 200)
  - Consults end users (> 600)
  - Market and Technology Analysis
  - Publishes consistently with Tech Target
  - Periodically published for trade magazines

- 31+ years industry experience
Cloud Storage is changing the storage game

1st storage system that emulates living organisms
Lessons of Socrates
Agenda

- Defining Cloud Storage
- Types of Cloud Storage
- Cloud Storage vs Cloud Computing w/Storage
- Problems Solved by Cloud Storage
- “Cloud Washing”
- Who Does What
- Why Cloud Storage is “Life-Like”
Defining Cloud Storage – What Is It?

- Massively scalable
- Not tied to geographic location
- Based on commodity components
- Secure multi-tenant
- Exceptional Self-healing
- Enduring data resiliency/permanence
- Allocation on-demand
- Billed or licensed per usage
- Application agnostic
- Primary access via REST or SOAP
Massively Scalable

- **Capacity Scalability**
  - From petabytes to exabytes even zettabytes

- **Performance Scalability**
  - Increases linearly as capacity scales
    - Throughput and IOPS per TB improves or stays the same

- **Object or File Scalability**
  - Multiple billions
Loose Geographic Coupling

- Geographically aware
  - Geographically distributable
  - On-demand data movement based on policy
    - Round-trip latency/response time
Based on Commodity Components

- Off the shelf components
  - x86 servers
  - HDDs – SATA, Nearline SAS, SAS
  - SDDs – SLC, EMLC, MLC
  - Ethernet – 1/10/40 Gbps
Secure Multi-Tenant

- No unauthorized user or employee
  - Can ever read/write someone else’s data
  - Including some or all following levels of security:
    - Encryption – AES, Erasure Codes, FIPS 140-2
    - ACL
    - Automated rotating passwords
Exceptional Self Healing

- Extremely low probabilities of data ever being
  - Lost
  - Corrupted
  - Silently corrupted
Enduring Data Resilience/Permanence

- Accessible online for years or decades
  - Searchable
  - Capable of locking down data (immutability)
    - WORM
Allocation On-Demand

- Capacity, Performance, and Location
  - Policy based
  - User transparent
Billed or Licensed per Usage

- True “Pay-as-you-go”
  - Charged on usable storage
  - Versus traditionally raw storage
  - A.K.A. “Pay-by-the-drink”
Application Agnostic

- Transparent to all applications
  - Structured or unstructured
  - Only difference is REST and/or SOAP interface
    - Apps are easily modifiable
    - Hardware & software gateways are also available
Primary Access REST and/or SOAP

- Same interface as World Wide Web & Web 2.0
  - REST – Representational State Transfer
    - HTTP puts and gets
  - SOAP – Simple Object Access Protocol
    - XML, RPC, and HTTP
Types of Cloud Storage

- Public
- Private
- Hybrid
Public Cloud Storage

- Storage as a service (STaaS)
  - Over the Internet or VPNs
  - Pay-by-the-drink
    - Only for what’s actually being used
  - Accessed via REST and/or SOAP
    - Also accessed via hardware & software gateways
  - Typically has multiple data centers
    - Geographically separated by regions
- 3 classes
  - Consumer, Industrial, Enterprise
Private Cloud Storage

- IT owns/operates/manages their cloud storage
  - On customer’s premises or Co-lo
  - Leveraging cloud storage technology
  - On VLAN and/or VPN
  - Can charge-back to departments if desired
  - Still accessed primarily via REST and/or SOAP
    - Also accessed via hardware & software gateways
- Provides many public cloud advantages
  - Just requires own or co-lo data centers
Hybrid Cloud Storage

- Best of both worlds
  - Combination of public and private cloud storage
  - Infinite variations
    - Mostly private with public for DR or extra copy
    - Mostly public with private for local caching
    - Everything possible in-between
Cloud Computing w/Storage
A.K.A. Storage in the Cloud

- Not the same as Cloud Storage
- Cloud computing can use any type of local storage
  - DAS, NAS, SAN, or Object
  - Examples include:
    - Google docs, Office 365, Shutterfly, SalesForce.com, iCloud, BURR, etc.
- This is storage in the cloud, not cloud storage
Problems Solved by Cloud Storage

- Minimal storage containers as data scales
- Greater automation w/reduced management
- Multi-tenancy
- Long-term data permanence with data reliability
- Ending tech refresh disruptive data migrations
- Cost effective DR
- Reduced TCO
Minimal Storage Containers as Data Scales

- Traditional storage not designed for massive data
  - Used to be 1PB was a lot of storage…not any more
  - Dozens of PBs, 10EBs, 100EBs, even a ZB is a lot of storage
  - Leads to storage system sprawl
  - And lots of data migrations
Manually Labor Intensive Data Migration

- Rip-Out-And-Replace Architectures require DM
  - Required as storage systems approaches limitations
  - And larger more capable versions are implemented
  - Also required for equalizing storage sprawl
  - And expensive
Cloud Storage Minimizes Storage Containers Because it’s Object Storage

- No known object storage limits
  - Each add’l node increases capacity & performance
    - Always positively
  - Objects measured in the billions vs. millions
    - Capacity measured from PBs to EBs to ZBs
  - Capacity and/or performance scales
    - In small or large increments
    - without limitations
How Object Storage Works

- Loosely federated data
  - Vs. consistent storage system
    - Across all resident data
  - E.g. No requirements for
    - Cache consistency
    - Nodal awareness of objects owned by other nodes
    - Single aggregated namespace
- Data scales based on rules
  - Rules About the data itself
  - Rather than about the system
  - Can scale nearly indefinitely
Mgmt Overload – Complicated

- Traditional Storage
  - Manually labor intensive
    - Load balancing
    - Operations
- Infrastructure
  - Ports
  - Switches
  - Cables
  - Connectors
  - Multi-pathing software
Object Storage Has More Metadata

- More customized control over the data
  - Vs. file system w/ fixed amount of metadata
    - File type, creation date, & last-accessed date
  - Vs. SAN storage which typically has virtually none
- Object Storage increases # of possible metadata fields
  - Customizable for specific business and system functions
    - Allows system to manipulate data based on policy triggers
    - Data scales based on rules
- Rules that also automate many traditional manual tasks
  - Tiering, security, migration, redundancy, and deletion
Greater Automation = < Mgmt

- Data dynamically laid out & optimized automatically
  - Vs Traditional storage systems
    - Require physical data location assignments
    - Data manually migrated & optimized as system changes over time
  - Object storage data moves nondisruptively based on policies
    - w/o admin decisions about placement of each individual object
Power & Cooling Becoming a Crisis

- In the dot com era\(^1\)
  - Major data centers were built w/2 megawatts of power
- Today\(^1\)
  - Same data centers are built with > 50 megawatts of power
- The key culprit…traditional storage

\(^1\)Sources: Gartner, IDC, 451 Group, Forester
HDDs – Electro Mechanical Devices

- HDDs
  - Capacity storage device of choice
  - Spinning platters
  - Consume power
  - Generate heat requiring cooling

- Storage growth rates
  - ~ 62% CAGR\(^1\) = capacity doubling every 18 mos
  - HDD capacity improvements slowing
    - .5TB to 1TB to 2TB to 3TB to 4TB
    - 100% growth to 50% to 25% means more HDDs over time

\(^1\)Sources: Gartner, IDC, 451 Group, Forester – CAGR varies & all agree with this rate for unstructured data
Object Storage Inherently More Efficient

- Provides more usable storage
  - No RAID overhead
  - More usable efficiency
    - = less physical storage
    - Means a lot less of everything else too
      - Switches
      - Ports
      - Cables
      - Connectors
      - Power/Cooling/Battery BU
      - Etc.
Built-in Multi-Tenancy

- Objects have own custom metadata
  - Functioning as autonomous data instances
  - Carries a broad swath of access policies
  - Controlled & restricted access for distinct parties

- Unparalleled Security
  - Built-in encryption in-flight and at-rest
  - Erasure codes

- Embedded billing and/or chargeback
  - Only charges what’s actually used

Traditional Storage
- Bolted on
- Not built-in
Unparalleled Self Healing Data

- Erasure Code Data
  - Minimally 10,000 times more resilient than RAID\(^1\)
  - In some configurations much greater

- Multiple Copy Data
  - Minimally 1,000 times more resilient than RAID\(^2\)

\(^1\)Calculations based on 16/4 number of slices or data chunks
\(^2\)Calculations based on 4 total copies of the data
Preserving data for long time periods is challenging

- Especially periods longer than storage systems’ refresh cycle
- Regulations require lengthy retention periods
  - A big issue for more organizations in more industries
  - Traditional storage makes it complicated & Time consuming
    - Requires manual labor intensive data migrations from old to new
Long-term Data Permanence w/Data Reliability

- Object data is location independent
  - Object storage systems are based on grid technology
  - Natively preserves data
  - Data is inherently WORM (write once read many)
  - Once data is saved it’s tagged with a unique identifier
    - Guarantees immutability of that object – mods/edits create new object
    - Old object can be deleted or kept as P.I.T. version
- Traditional storage require additional software
  - And a lot more effort
Traditional Storage
3 yr Tech Refresh Cycle

- Must refresh every 3 to 4 years
  - Because costs skyrocket thereafter
    - Maintenance costs go up dramatically from vendors
    - Parts become impossible to find
    - Storage systems break down more frequently
    - Greater unscheduled as well as scheduled downtime
Storage Refresh is Incredibly Time Consuming

- Currently averaging 9 to 12 months
  - Because of the manually intensive data migration
  - Each refresh takes longer because 62%* stored data CAGR
    - Data volume grows 4x every 3 years
    - Making the transfer alone incredibly time consuming
  - Data gets corrupted or is lost in each migration
  - Server remediation takes longer with each migration

*Per IDC, Gartner, 451 & DSC, stored data grows ~62% per year quadrupling in 3 years
Storage Technology Locked Into Refresh Cycle

- New innovations or savings must wait
  - Projected cycle’s capacity all bought upfront
  - Newer higher capacity HDDs or SSDs often not compatible
    - Newer software as well
  - Technology lock-in means pre-paying for unused assets
    - Often, never used
How Cloud Storage Makes Tech Refresh Painless

- Each node is added online, no scheduled downtime
- As newer more advanced nodes are added
  - New nodes are auto discovered & integrated into system
  - Older nodes can be removed from system at leisure - online
    - Each object system node can be a mix of old and new nodes
- Data is copied seamlessly
- Eliminates
  - Data migration
  - Rip-Out-And-Replace
  - Scheduled downtime
Costly DR

- Traditional Storage DR requires 2 or more sites
  - 2nd site can be own, 3rd party, or outsourced (SUNGARD)
  - Duplicate storage: power, cooling, etc.
  - Complicated
  - Expensive
Cloud Storage DR is a Different Story

- Cloud Storage is the 2nd site
  - Data is secure in one or multiple Cloud data centers
  - Only paying for consumed storage
  - In-expensive
Cloud Storage is an entirely new costing paradigm

<table>
<thead>
<tr>
<th>Costing Paradigm</th>
<th>Traditional Storage</th>
<th>Cloud Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost basis</td>
<td>Raw Storage</td>
<td>Actual used storage</td>
</tr>
<tr>
<td>Pay timing</td>
<td>Pre-usage</td>
<td>Post-usage</td>
</tr>
<tr>
<td>OpEx costs</td>
<td>Yes</td>
<td>Public = no</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Private = yes (less)</td>
</tr>
<tr>
<td>OpEx over 3 yrs = 5 x CapEx</td>
<td></td>
<td>Hybrid = yes (lot less)</td>
</tr>
<tr>
<td>Data migration costs</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Hidden costs</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
What is Cloud Washing?

- Calling a storage system “Cloud Storage”
  - When in reality it has only some cloud characteristics
  - It’s an attempt to leverage the market hype
  - Just because they say it’s cloud storage
    - Does not make it true
General Object Storage Issues

- High latency
- Good throughput, generally not so good IOPS
  - Frequently changing data is not a good fit…usually
    - Scality Organic RING is one exception
- Each vendor has own uniqueness & advantages
  - Erasure Codes resilience
  - Lower cost
  - Public, private, and hybrid offerings
  - OEM relationships
  - Packaged w/hardware or as software only
- Open source “Open Stack”
  - Not ready for prime time
## Who Does What in Cloud Storage

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Product</th>
<th>Public-Private-Hybrid</th>
<th>Interfaces</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>S3</td>
<td>Public</td>
<td>REST/SOAP</td>
<td>Public leader</td>
</tr>
<tr>
<td>Amplidata</td>
<td>AmpliStor</td>
<td>Private</td>
<td>REST/SOAP</td>
<td>Erasure codes</td>
</tr>
<tr>
<td>Basho</td>
<td>Riak</td>
<td>Private</td>
<td>REST/SOAP</td>
<td></td>
</tr>
<tr>
<td>Caringo</td>
<td>CAStor</td>
<td>Private</td>
<td>REST/SOAP/NFS/CIFS</td>
<td></td>
</tr>
<tr>
<td>Citrix</td>
<td>CloudStack</td>
<td>Private</td>
<td>REST/SOAP</td>
<td></td>
</tr>
<tr>
<td>Cleversafe</td>
<td>Slicestor</td>
<td>Private</td>
<td>REST/SOAP</td>
<td>Open source. Just acquired CloudStack</td>
</tr>
<tr>
<td>DELL</td>
<td>DX Object Store</td>
<td>Private</td>
<td>REST/SOAP/NFS/CIFS</td>
<td>OEM from Caringo.</td>
</tr>
<tr>
<td>DDN</td>
<td>Web Object Scaler</td>
<td>Private</td>
<td>REST/SOAP</td>
<td></td>
</tr>
<tr>
<td>EMC</td>
<td>ATMOS</td>
<td>Private</td>
<td>REST/SOAP</td>
<td></td>
</tr>
<tr>
<td>Gluster</td>
<td>GlusterFS</td>
<td>Private</td>
<td>REST/SOAP/NFS/CIFS</td>
<td>Largest storage containers in production. Some &gt; several EB. Erasure codes.</td>
</tr>
<tr>
<td>Mezeo</td>
<td>Cloud Storage</td>
<td>Private</td>
<td>REST/SOAP</td>
<td></td>
</tr>
<tr>
<td>Microsoft</td>
<td>Azure</td>
<td>Public</td>
<td>REST/SOAP</td>
<td></td>
</tr>
<tr>
<td>NetApp</td>
<td>StorageGRID</td>
<td>Private</td>
<td>REST/SOAP/NFS/CIFS</td>
<td></td>
</tr>
<tr>
<td>Nirvanix</td>
<td>CloudComplete</td>
<td>Public-Private-Hybrid</td>
<td>REST/SOAP/NFS/CIFS</td>
<td></td>
</tr>
<tr>
<td>Rackspace</td>
<td>Open Stack</td>
<td>Public</td>
<td>REST/SOAP</td>
<td>Open source.</td>
</tr>
</tbody>
</table>
What About Other Scale-out Storage Tech?

- Scale-out storage often perceived as Cloud Storage
  - Scale-out NAS
  - Scale-out SAN
Scale-out NAS

- Similar to Object Storage
  - Utilizes x86 servers with internal HDDs and/or SSDs
  - Or shared external DAS, SAN Storage, or NAS systems
  - Sometimes leverages GNS or File Virtualization
  - But most require a metadata database = object bottleneck
    - GlusterFS is an exception – acts very similar to Object Storage
General Scale-out NAS Cloud Issues

- Each add’l node has diminishing marginal returns
  - Eventually next node has a negative impact
- Tends to be higher cost than NAS
  - Too high for Cloud Storage requirements
- Limited objects/files
  - GlusterFS is exception – no metadata database similar to object
- No geographical awareness
- No automated intersystem storage tiering
  - Limited automated intrasystem storage tiering
- Resilience based primarily on RAID
  - Some do multi-copy mirroring
- Primarily still a “Rip-Out-And-Replace” architecture
  - Just a bigger bucket and especially with tech refresh
Scale-out Vendors & Products

- Avere – FXT
- BlueArc – Mercury & Titan
- DDN
  - NAS Scaler/GRIDScaler/EXAScaler/xSTREAM Scaler
- EMC – Isilon
- GlusterFS – Open Source
- GRIDStore (Low End)
- HP – IBRIX
- IBM – SONAS
- Lustre – Open Source
  - Wham Cloud
- MogileFS – Open Source
- NetApp – FAS Ontap 8.1
- Panasas
Scale-out SAN

- Utilizes x86 servers with internal HDDs and/or SSDs
- Clustered Controllers
  - With proprietary ASICs
  - IBA, Ethernet, FC, PCIe, Crossbar, pt-to-pt
- Lower end is switched iSCSI on 1G or 10G Ethernet
- High end tends to be within the rack
  - Or Limited # of racks
  - Mostly layer 2 networking – deterministic
    - FC, FCoE (CEE) IBA, AoE
  - Switched internal Controller
    - Or stacked controller mesh interconnect
General Scale-out SAN Issues

- Nodes have diminishing marginal returns
  - Eventually a node has a negative impact
- Tends to be far & away the most costly
  - CORAID is the exception w/AoE (ATA or IDE over Ethernet)
    - Most not priced for Cloud Storage requirements
  - All of SAN’s complexity & cost, just at a bigger storage scale
    - iSCSI is less complex as is AoE
  - Limited to metropolitan areas
- No automated intersystem storage tiering
- Limited automated intrasystem storage tiering
- Resilience based primarily on RAID
- Still a “Rip-Out-And-Replace” Architecture
  - Especially on tech refresh
## Scale-out SAN Vendors & Products

- **CORAID** – EtherDrive (AoE)
- **DDN** – S2A & SFA series (FC, FCoE, iSCSI)
- **DELL** – EqualLogic (Low end iSCSI)
- **EMC** – VMAX/VMAXe (FC or FCoE)
- **GRIDStore** (Low end iSCSI or FC)
- **HDS** – VSP (FC or iSCSI)
- **HP**
  - Low end – LeftHand P4000 series (iSCSI)
  - High end – 3PAR series (FC)
- **IBM**
  - Low end – SVC (FC)
  - High end – XIV (iSCSI)
- **Pure Storage** – iSCSI SSDs
- **SolidFire** – iSCSI SSDs
Why Cloud Storage is “Life-Like”

- Organics
  - Self replicate
  - Self heal
  - Regenerate – renew automatically
  - Are adaptive
  - Are resilient
  - Are geographically aware
  - Have autonomic nervous system
    - Crucial function automation
    - Without conscious intervention
- As does Cloud Storage
Conclusions

- Cloud Storage’s ability to mimic life
  - Changes the storage game
  - Makes storage more resilient than ever
  - Stronger
  - More permanent
  - Better
Paying Attention
Marc Staimer, President & CDS
Dragon Slayer Consulting
marcstaimer@comcast.net
503-579-3763