Deploying Public, Private, and Hybrid Storage Clouds

Marty Stogsdill
Oracle
The material contained in this tutorial is copyrighted by the SNIA unless otherwise noted.

Member companies and individual members may use this material in presentations and literature under the following conditions:

- Any slide or slides used must be reproduced in their entirety without modification.
- The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.

This presentation is a project of the SNIA Education Committee.

Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.

The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.
Abstract

Deploying Public, Private, and Hybrid Storage Cloud Environments

Everyone has heard talk of storage clouds but do you know how to deploy one? This session will be a technical dive into implementations of popular storage cloud use cases and how cloud fits into your existing enterprise IT operations.

- Cloud Storage Overview
- Deploying Cloud Storage
- Popular Use Cases Deployment Food for Thought
  - Backup Cloud
  - Long Term Archive Cloud
  - Application Data Cloud
- SNIA CSI’s CDMI
Cloud Computing Definitions

▷ NIST Definition:

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

5 Essential Characteristics
• On-demand self-service
• Resource pooling
• Rapid elasticity
• Measured service
• Broad network access

3 Service Models
• SaaS
• PaaS
• IaaS

4 Deployment Models
• Public Cloud
• Private Cloud
• Community Cloud
• Hybrid Cloud

Source: DRAFT "NIST Cloud Computing Definition", NIST SP 800-145
What Is Cloud Storage Used For?

- Elastic demand for web based media (video, eBooks, audio)
- Backup to the cloud
  - Restore, Recovery, “Seed” the backup with hard drive
- Sync of files to the cloud and multiple devices
  - Internet “Drive” secondary storage
- Archive/Preservation in the cloud
  - Including Compliance, Retention and eDiscovery
- Storage for Cloud Computing
  - Support for legacy storage interfaces key
  - Self Service Storage Administration
Data Storage Interface Evolution

- Born Digital data is being generated by an ever diversifying set of devices
- 75% of data being generated by individuals (IDC)
  - Amount of Data doubling every two years
- 5 Billion Mobile Phones
  - Tablets on the rise
- Storage space on devices will always be limited (although continues to grow exponentially)
- Device Apps are accessing content/storage via the Internet
  - Cloud Storage largely driven by device based use cases
- Traditional data storage interfaces (filesystems) are not up to this task!
Where are your users putting their data?

With corporate users bringing their own devices (BYOD), where are they putting files?

Devices have limited storage capacity

They are saving them in storage clouds!

Corporate data is already in the public cloud and unmanaged/protected
Distilled Storage Cloud Technical Requirements

- Storage – and lots of it
- Need a Global Namespace for data location
  - FILE handle -> URL + Unique Object ID globally
- Need support for Rich Metadata associated with the data
  - Plus Metadata Query and Index based Search
- Need HTTP access – browser support
- Data Services need to be transparent to the user, but extensible and powerful
  - Driven by Metadata
- Support for large objects, many small objects
- Support for active preservation operations
  - Driven by preservation Metadata
The Cloud Storage Big Picture

- Cloud Storage Clients:
  - Connect to
  - Manage
  - Interact with Cloud storage
Corporations already offer their own email (instant messaging, etc.) services to their employees.

Cloud Storage is the next type of service offering for employee devices.
- Enterprise “Dropbox” implementations

Essentially a “private” cloud with access over the Internet from corporate and other “Apps”.

Data is retained “in house”, protected, available, secured and compliant.

Check out SNIA Tutorial: Building The Business Case for the Cloud
Storage Cloud Use Cases

✦ Determine your storage cloud use case:
  ✦ Backup Cloud
    › Capacity and cost larger concerns than latency
    › Enterprise “dropbox”
  ✦ Long Term Archive Cloud
    › Low cost and high capacity larger concerns than latency
  ✦ Application Data Cloud
    › Low latency larger concern then high capacity or cost

✦ Separations across use cases are normally prioritization of: capacity, latency, and cost
Backup Clouds

- Traditional backup software on local machines pointing to a Backup Cloud

- Central backup server pointing to a Backup Cloud
Long Term Archive Cloud

- Archive Cloud as Last Tier in ILM deployment

- Direct Archive Cloud
Application Data Cloud

Storage Cloud as Backend Storage

Storage Cloud as a Compliment Environment
Diagnose Potential Bottlenecks

- Diagnose potential bottlenecks for your deployment scenario
- Does your LAN have available capacity?
  - Do you have 1gb to desktop? Is 1gb enough?

SANs and LAN-free backups/storage access was deployed for a reason
Should your cloud deployment look more like:

- Evaluate cost/performance trade-offs with:
  - Dedicated circuits / Internet QoS
  - Caching/Accelerators
After defining your cloud storage use case evaluate your organizations needs for:

- Cloud Storage Capacity
  - High Importance
  - Low Importance
  - Not Important

- Cloud Storage Latency
  - High Importance
  - Low Importance
  - Not Important

- Cloud Storage Cost
  - High Importance
  - Low Importance
  - Not Important

Most organizations will have trade-offs between capacity, latency, and cost when designing their storage clouds.
Define If Shared Tenancy Is a Fit

What is multi-tenancy:
- “the terms multi-tenant and multi-tenancy are not new; both have been used to describe application architectures designed to support multiple users, resource owners or “tenants” for many years. With the advent of cloud computing, this terminology has simply been extended to include any cloud architecture”
  - Source: “Storage Multi-Tenancy for Cloud Computing” Whitepaper, Paul Feresten, SNIA CSI Member

Secure multi-tenancy:
- Application layer
- Server layer
- Network Layer
- Storage Layer
Secure Multi-Tenancy in Cloud Storage

- Isolation/Security/Defense In-Depth:
  - Firewall/Edge network security
  - Tunneling across network/internet connections
  - HTTPS SSL/TLS
  - Private volumes, partitions, LUNs, or datastores
  - Encryption of data
Document Concerns - Security

▶ Security Concerns
  ▶ Legal Jurisdiction
  ▶ Regulations
  ▶ Who ensures compliance/privacy?

▶ Data Governance
  ▶ Integrity
  ▶ Classification
  ▶ Regulation / legal

▶ Audit
  ▶ Forensics
  ▶ Data
SNIA Cloud Storage Initiative

- Gaining Momentum for Cloud Storage
  - Supporting the development and adoption of CDMI, Cloud Storage
  - Marketing, Outreach, Education on Cloud Storage
  - Requirements gathering
  - Premier Organization promoting Cloud Storage and associated Standards

- 23 Member companies
  - Multiple events including Cloud Burst event focused exclusively on cloud storage
Cloud Data Management Interface

Applicable to three deployment models of Cloud Storage:

- Cloud Storage for Cloud Computing
  - Whitepaper at snia.org/cloud – the management interface for the lifecycle of storage in a compute cloud

- Public Storage Cloud
  - Whitepaper at snia.org/cloud – both a Data Path for the Cloud and a Management Path for the Cloud Data

- Private Cloud Storage
  - As well as hybrid clouds
  - An API for Storage Vendors selling into Cloud based solutions

Semantics

- Simple Containers and Data Objects with tagged Metadata
- Data System Metadata expresses the data requirements

Protocol

- RESTful HTTP as “core” interface style
- JSON (JavaScript Object Notation)– format of the representations are extensible
Why not just adopt one of the existing interfaces?

- Despite the “open” licensing of several existing cloud storage interfaces, they all remain under the change control of a single vendor.

- No cloud vendor wants to have a competitor have change control over their interface.
  - Thus they release their own interface which they do have change control over.

- This leads to the propagation of multiple interfaces, each essentially locking developers/customers into that service.

- CDMI is under change control of a standards body, accommodates requirements from multiple vendors and can be extended for proprietary functions.
Where is CDMI Being Used?

- Commercial Cloud Storage vendors started announcing implementations in April 2011
- Increasing Adoption by eScience and Academic community
  - Especially in Europe (OpenNebula open source cloud, others)
- Quarterly Plugfests with growing participation
  - > September plugfest in Europe and North America
- CDMI 1.0.2 is now an International Standard
  - ISO/IEC 17826:2012
- CDMI has a reference implementation (Java™ under BSD)
  - Downloaded by hundreds of developers worldwide
A single cloud computing infrastructure can implement both the OCCI and CDMI interfaces.

The infrastructure abstracts the configuration of the networking and virtual machine details and uses the standard interface merely to define connectivity.

A cloud computing client can then utilize the interfaces to both specify the data requirements and then use that data for guests.

Get whitepaper at snia.org/cloud
Build vs. Buy

Weigh start-up and operating costs of leveraging a service provider / vendor vs. building your own

PUT /MyContainer HTTP/1.1
6-Common_Operations Page 1 of 4
Host: cloud.example.com
Accept: application/cdmi-container
Content-Type: application/cdmi-container
X-CDMI-Specification-Version: 1.0
{
  "metadata": { }
}

HTTP/1.1 201 Created
Content-Type: application/cdmi-container
X-CDMI-Specification-Version: 1.0
{
  "metadata": {},
  "objectURI": "/MyContainer/",
  "objectID": "0000706D0010D538DEEE8E38399E2815",
  "objectName": "MyContainer/",
  "parentURI": "/",
  "domainURI": "/cdmi_domains/MyDomain/",
  "capabilitiesURI": "/cdmi_capabilities/Container/",
  "completionStatus": "Complete",
  "metadata": {
    "cdmi_size": "0"
  },
  "childrenrange": "",
  "children": []
}

Source: CDMI Reference Implementation Developers Guide
http://cdmi.sniacloud.com/CDMI_Spec/6-Common_Operations/6-Common_Operations.htm
Deployment Considerations

Think through control, access, administration

- SLA

Diagram:

- Public Cloud Service Provider
- Public Cloud Subscriber
- No Access or Control
- Administers
- Physical Control
- Application Data
  - Server OS
- Storage Container Data
  - Hypervisor
  - LUN
  - Server Hardware
  - Storage Hardware
- Control
- Requests / Self Service
- No Control
CDMI Reference Deployment

Front End and Middle

CDMI Server
Restful JAX-RS Front-End

Java App Server (Glassfish)

CDMI Security filter

TLS Admin Security
Apache CXF
Spring Framework

CDMI Mid Layer

Objects
Containers
Capabilities
Notifications

Serialize / Deserialize
OCCI Export

Logging / Audit
Queues
Encryption
Retention
Hashing
Accounts
Query

Back-End

CDMI Reference Implementation Architectural Diagram
Green: SNIA Developed Code
Blue: 3rd Party Code

Source: CDMI Reference Implementation Developers Guide

Learn more at snia.org/cloud

Date: July 20, 2010
Back-End Reference Deployment

Source: CDMI Reference Implementation Developers Guide

* File System Naming:
  Container Objects = Folders named with the container name
  Data Objects = Files named with the object name, if one was given, else the Object ID
  Metadata = Files named with the same name as the corresponding object with an additional "." in front

Examples:
  Container: /mnt/cdmi server/MyContainer
  Container Metadata: /mnt/cdmi server/MyContainer/

  Data Object's Data: /mnt/cdmi server/MyContainer/MyDataObject.txt
  Data Object's Metadata: /mnt/cdmi server/MyContainer/.MyDataObject.txt

  Data Object's Data: /mnt/cdmi server/MyContainer/0000706D0010B84FAD185C425D8B537E
  Data Object's Metadata: /mnt/cdmi server/MyContainer/. 0000706D0010B84FAD185C425D8B537E
Begin with bare-bones secured Cloud and phase in functionality:
Expand into a mature Cloud leveraging common framework/standards:
The SNIA Education Committee thanks the following individuals for their contributions to this Tutorial.

**Authorship History**

Marty Stogsdill, Oracle

Updates:
- Alex McDonald
- Mark Carlson

**Additional Contributors**

- David Slik
- Alex McDonald
- Mark Carlson

Please send any questions or comments regarding this SNIA Tutorial to tracktutorials@snia.org