Education

Cloud Archiving

Paul Field
Consultant

SNIA Legal Notice

- The material contained in this tutorial is copyrighted by the SNIA.
- 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

Cloud Archiving

- Many organizations are moving to cloud computing and cloud storage. This session will explore how organizations can meet their information retention needs using cloud technologies and resources. We look at how data is moved to the cloud and how retention policies are managed in the cloud. We review various architectures on how a cloud archive can be built, from private to public and hybrid models using an organization's existing information management resources.

Learning Objectives

- Define cloud archiving
- Understand how cloud resources can be leveraged to provide information retention
- Show examples of how a cloud archive could be built using public, private and hybrid cloud resources

Where's your data?

Agenda

- Defining cloud archiving
- Public and private cloud services
- Compliance in the cloud
- Leveraging cloud resources for archiving
- Summary

Archive – A Definition

- A collection of data objects, perhaps with associated metadata, in a storage system whose primary purpose is the long-term preservation and retention of that data.

Source: The 2009 SNIA Dictionary

Archiving Objectives

- Supports business information retention and regulatory compliance needs
- Information typically protected and secured
- Supports operational efficiencies
- Moves data off of production systems
- For historical information
- Long-term in nature
- Easily searched

Cloud Storage versus Cloud Archiving

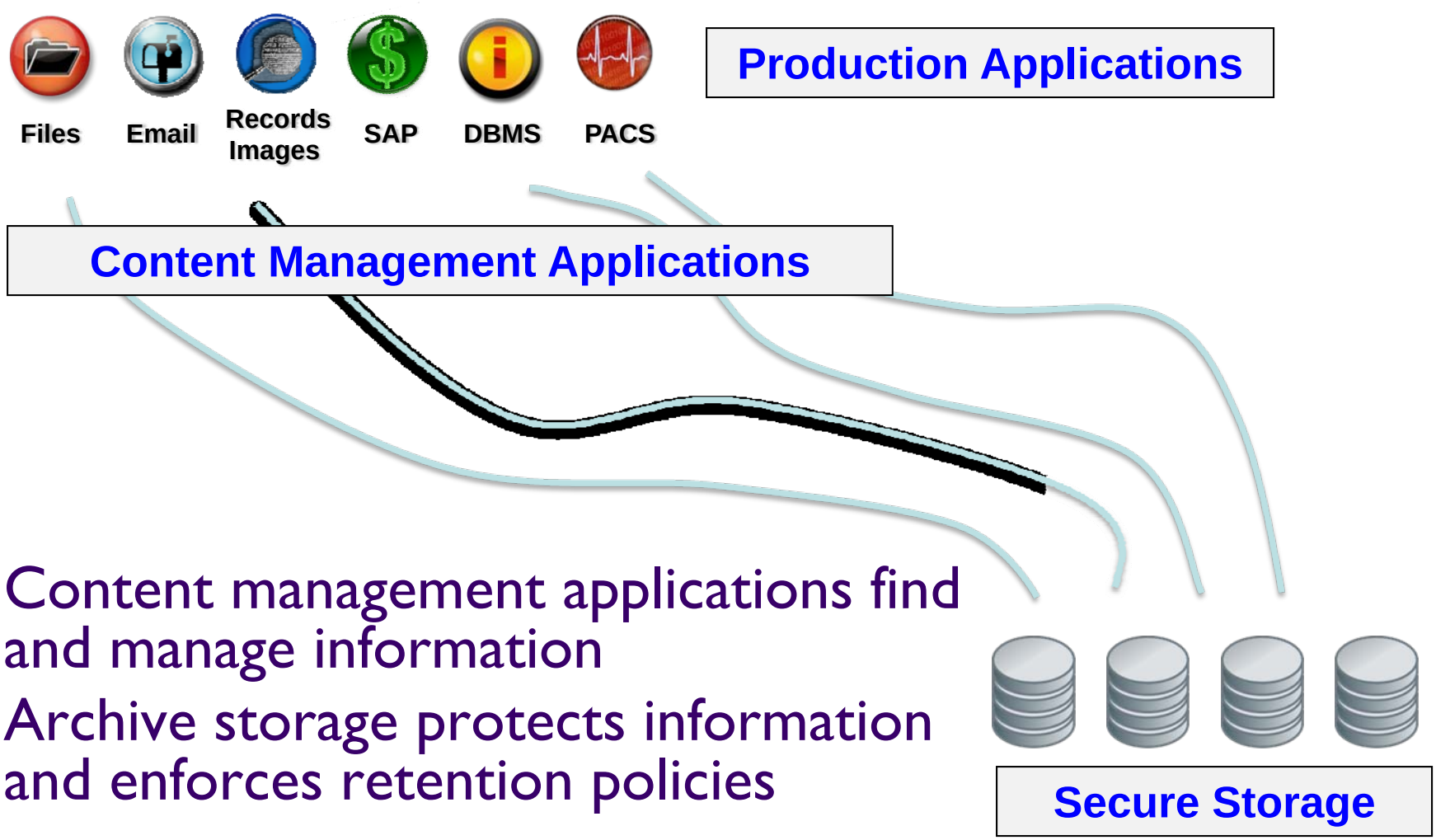
- Key Differentiators

Cloud Storage	Cloud Archiving
No protection from overwrites	Overwrite protection
No retention policy	Time or event based retention policy
No automatic deletion	Deletion per retention policy
No extended metadata	Archiving metadata
Limited search capabilities	Full search capabilities

Why Use the Cloud for Archiving?

- Cost
  - Leverage economies of scale of cloud storage providers
- Flexibility
  - Pay as you go cloud pricing model
  - Increase and decrease usage easily
- Low investment costs
  - Limited infrastructure purchase required
  - Pay by the capacity used
- Fewer IT skills and staff needed
  - Focus on business efficiencies and innovation
- Access from anywhere

Archiving Architecture

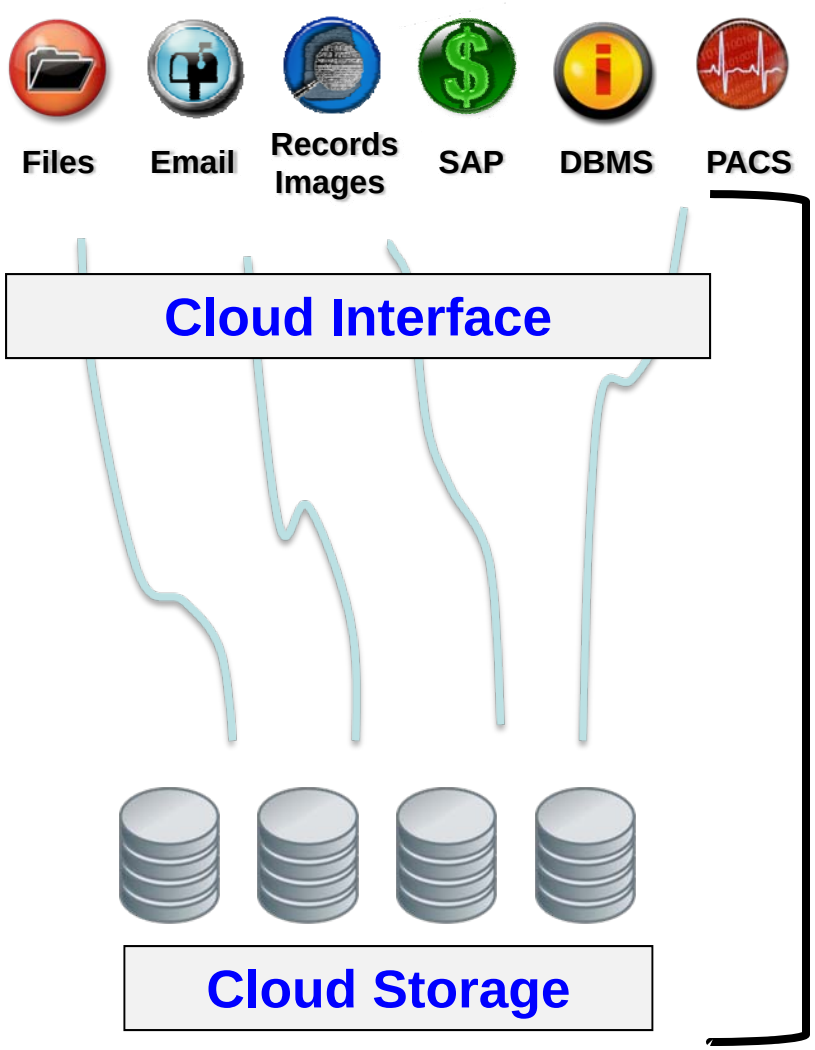


- Content management applications find and manage information
- Archive storage protects information and enforces retention policies

Cloud Storage - Examples

Public Clouds	Private Clouds
amazon web services™	EMC Atmos
the rackspace cloud cloudfiles	BYCAST StorageGRID®
nirvanix™	Azure™ Services Platform: Live Services, .NET Services, SQL Services, Microsoft SharePoint Services, Microsoft Dynamics CRM Services; Windows Azure

Public versus Private



- Public
  - Ease of setup, access and use
- Private
  - Greater control and privacy
- Hybrid
  - Customize based on needs and skills

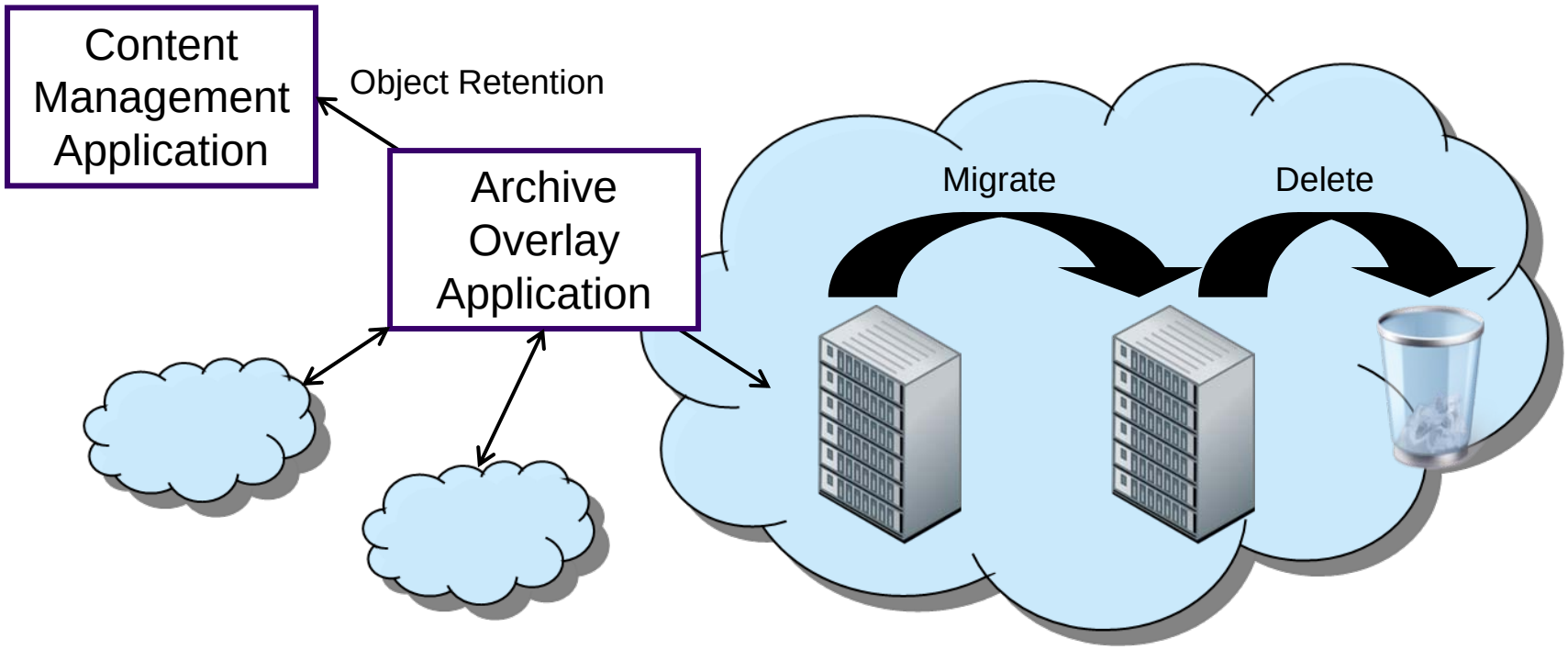
Retention and Compliance in the Cloud

- Amorphous nature of cloud storage makes retention and regulatory compliance in the cloud problematical
  - Overlay software could verify accuracy and completeness of data retrieved
  - Control access to data
  - Validate hash marks of objects
  - Protect data privacy with encryption

Cloud Logistics

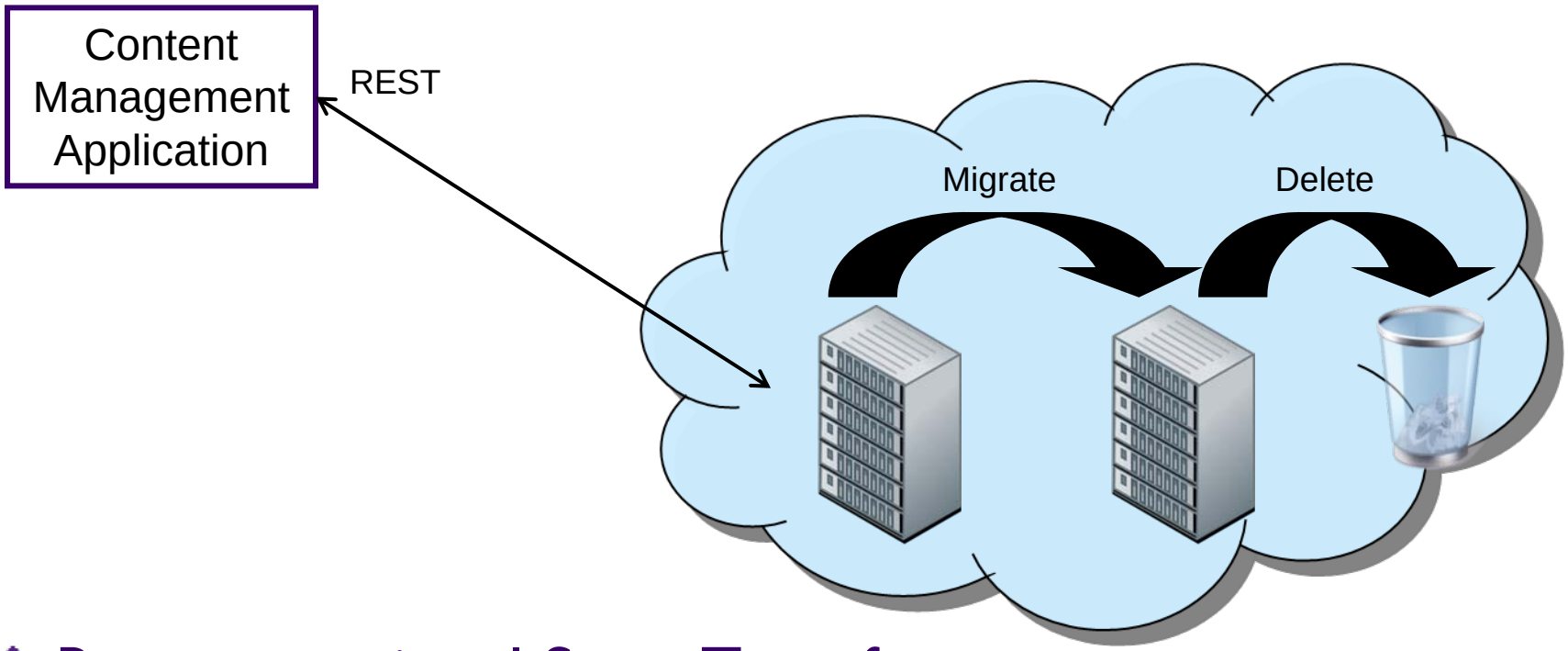
- Cloud access
  - Archive enhancements needed
    - New standards like the Cloud Data Management Interface proposal
    - Overlay applications
- Cloud network speeds
  - Wide area network costs limit ability to store and retrieve large amounts of data quickly

Archive Overlay on Cloud Storage



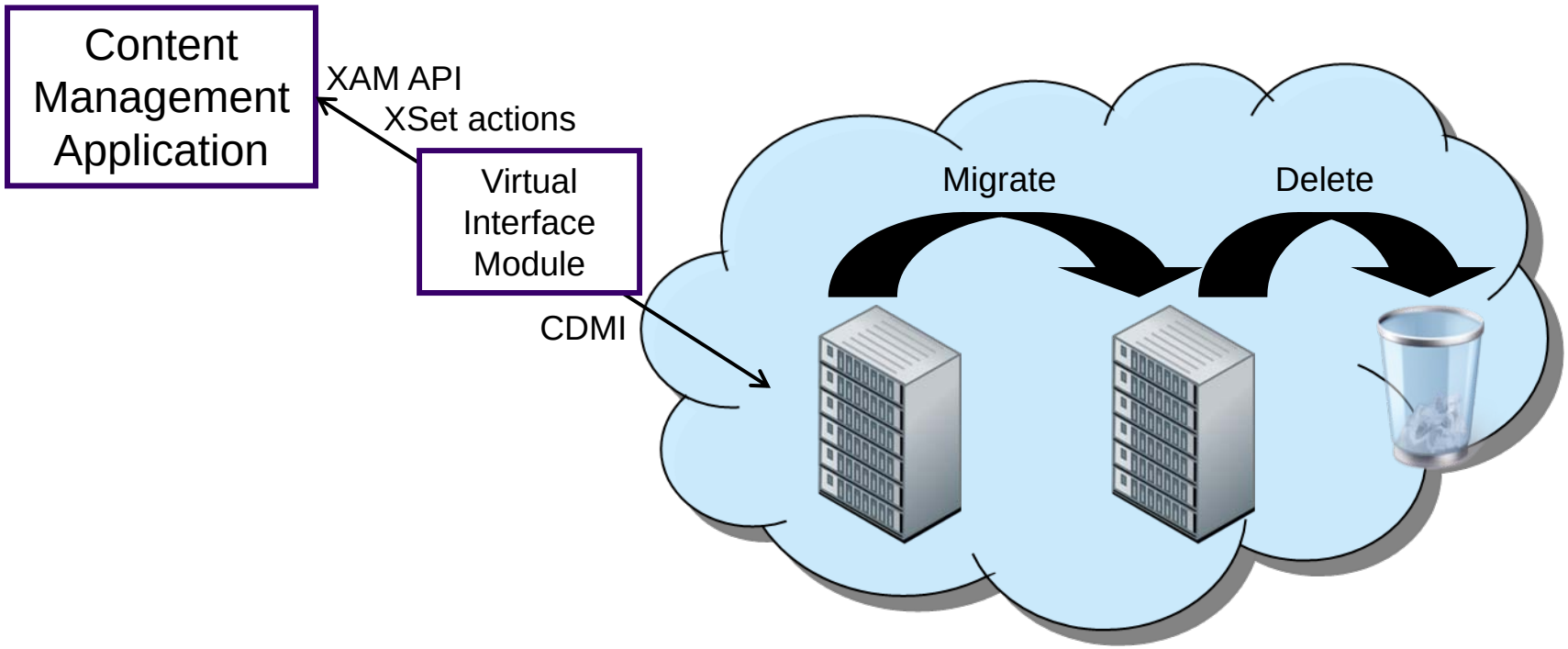
- Overlay application enforces retention as directed by the content management application
  - Cloud storage unaware of retention policies

REST and Cloud Archiving



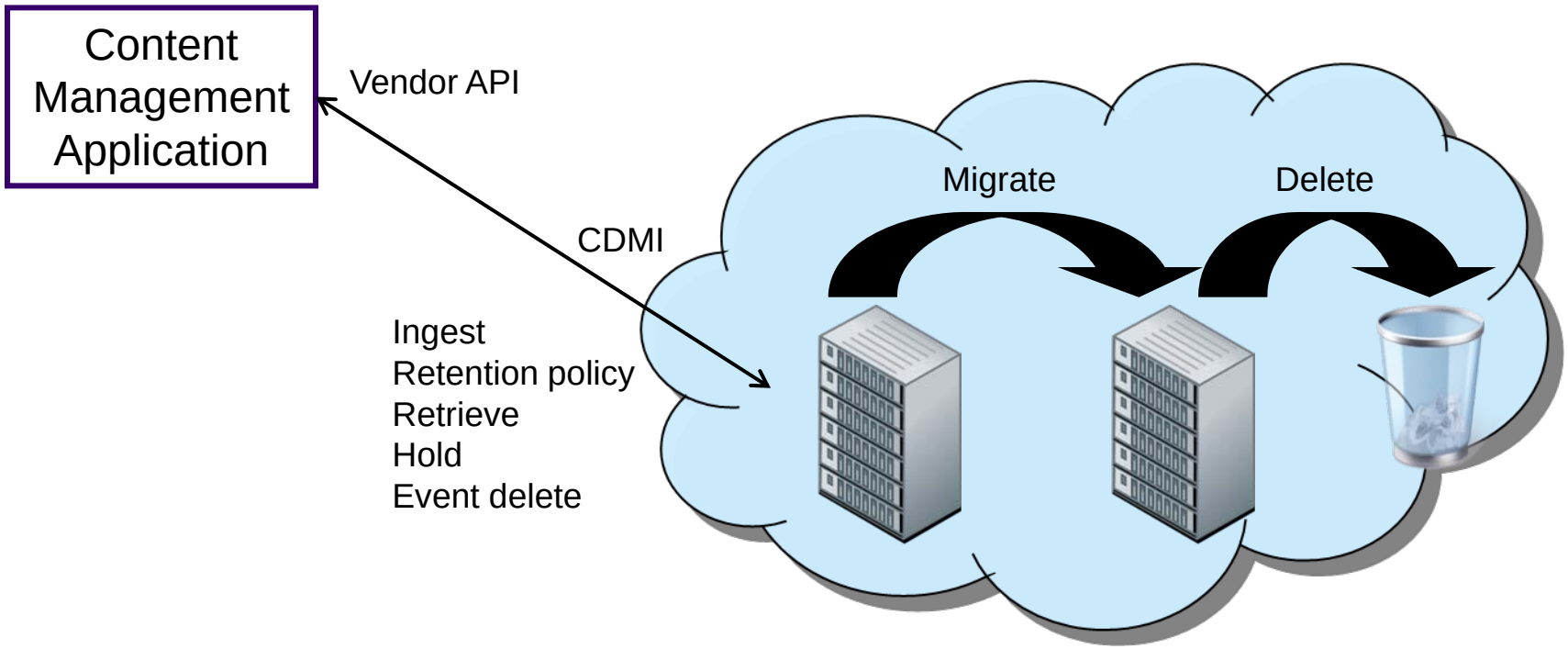
- Representational State Transfer
  - Operations for creating, reading, updating and deleting the individual data objects via HTTP
  - Standardized by CDMI

XAM and Cloud Archiving



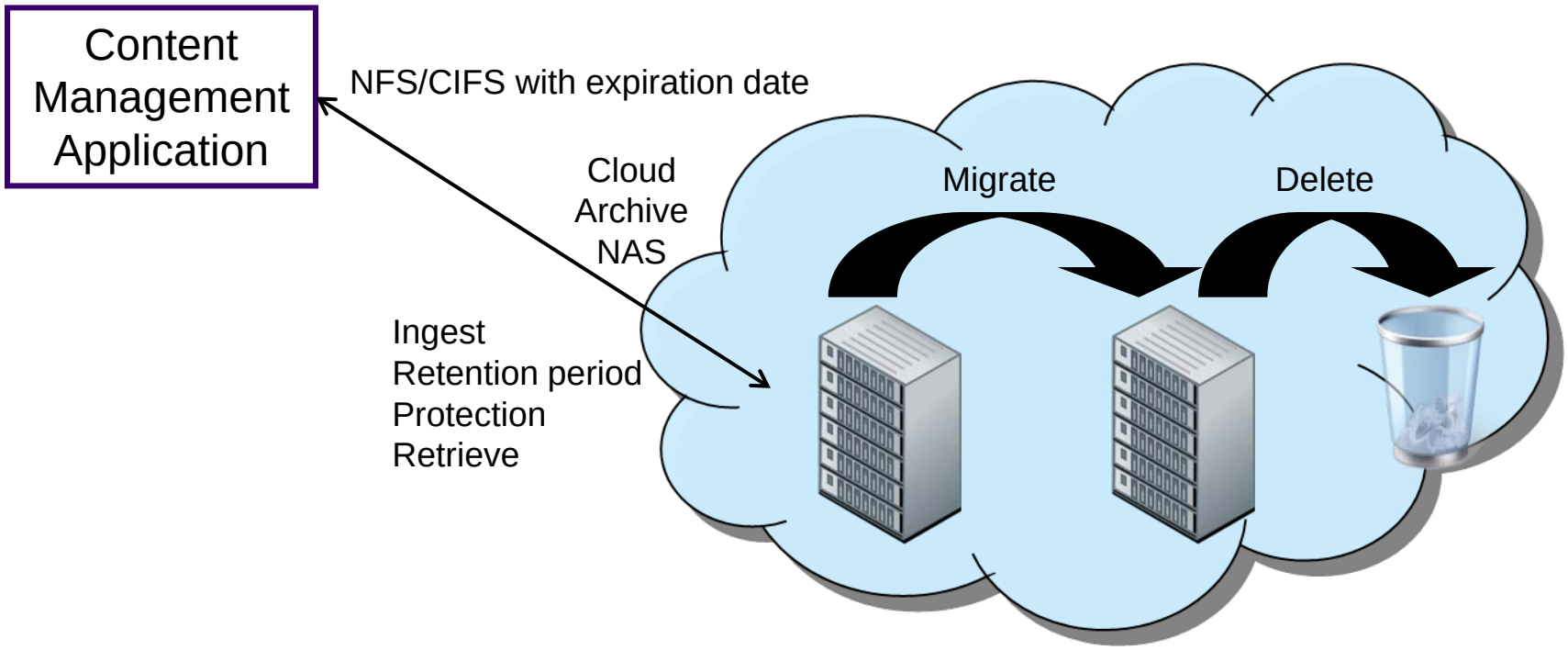
- XAM API passes retention policies to cloud services via CDMI
- Cloud manages object lifecycle

Vendor API and Cloud Archiving



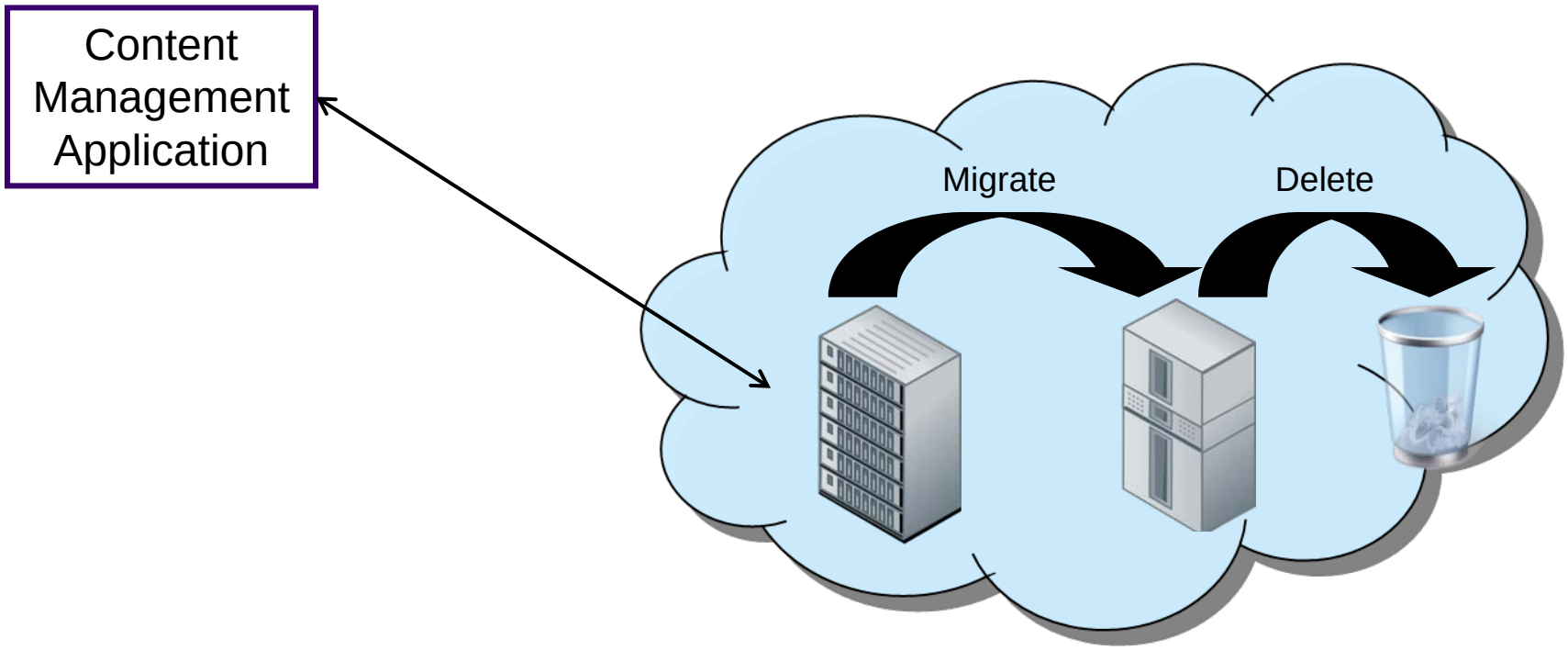
- Vendor API passes retention policies to cloud services via CDMI
- Cloud manages object lifecycle

NFS/CIFS and Cloud Archiving



- NAS protocol passes retention period to cloud services
- Cloud manages object lifecycle with protection

Cloud Archiving with Storage Tiers



- Class of service defined to allow tiering to tape
- Cloud moves data between disk and tape based on policies

CDMI – The Cloud Storage Standard

- Cloud Data Management Interface
  - Enables interoperable cloud storage and data management
  - Facilitates data portability, compliance and security
  - Provides a data path to the cloud service and a management path for the cloud data
  - Supports multiple types of cloud data storage interfaces
  - Create, Retrieve, Update and Delete
  - Discover cloud storage capabilities
  - Manage cloud containers and data
  - Can work with the Open Cloud Computing Interface (OCCI) to use compute resources
- See the CDMI tutorial by Mark Carlson

Summary

- Cloud storage can be used to implement cloud archiving with appropriate overlay software and interfaces
- Organizations may be able to leverage cloud resources to provide some or all of their archiving needs

Check out SNIA Tutorials:

How eDiscovery will affect storage professionals

Cloud storage security

About the SNIA CSI

- The mission of the SNIA Cloud Storage Initiative is to foster the growth and success of the market for what is generally referred as cloud storage and more generally the use of data storage resources and services in the cloud.
- The CSI is promoting the adoption of cloud storage as a new delivery model, that provides elastic, on-demand storage billed only for what is used.

Q&A / Feedback

- Please send any questions or comments on this presentation to SNIA: trackcloudtechnologies@snia.org

Many thanks to the following individuals for their contributions to this tutorial.

- SNIA Education Committee

Mark Carlson
Rob Peglar
Joshua Tseng
Terry Yoshii
Wendy Betts