Module 3.5
Content Addressed Storage (CAS)

Upon completion of this module, you will be able to:

• Describe the features and benefits of a CAS based storage strategy.

• List the physical and logical elements of CAS.

• Describe the storage and retrieval process for CAS data objects.

• Describe the best suited operational environments for CAS solutions.

So far we have looked at DAS, SAN and NAS. All of these technologies have advantages in solving specific business challenges. One unique business need is to store vast amounts of static, unchanging information. The best solution for that challenge in Content Addressed Storage (CAS). This lesson looks at CAS and how it can be implemented.
Lesson: CAS Description and Benefits

Upon completion of this lesson, you be able to:

• Define CAS.
• Describe the key attributes of CAS.
• List the features and benefits of CAS.

CAS is a newer technology and combines both physical and logical components to make a complete solution. In this lesson, we will look at the definition of CAS and benefits of using a CAS storage strategy.
What is Content Addressed Storage (CAS)?

- Object-oriented, location-independent approach to data storage.
- Repository for the “Objects”.
- Access mechanism to interface with repository.
- Globally unique identifiers provide access to objects.
- Extensible metadata that enables automated data management practices and applications.

CAS is:

- An object-oriented, location independent approach to data storage.
- A repository for objects, or bundles, of data and metadata.
- An access mechanism to interface with that repository.
- Large, flat address space where globally unique identifiers (Content Addresses) provide access to objects.
- Extensible metadata that enables data management applications and practices.
Data that does not change is referred to as fixed content. It is any informational object retained for future reference and/or business value including electronic documents and many types of newly digitized information. Unlike transactions or files, it is typically unchanged once created.
Challenges of Storing Fixed Content

- A significant amount of newly created information falls into the category of fixed content.
- Fixed content is growing at more than 90% annually.
- Often, long-term preservation is required (years-decades).
- Simultaneous multi-user online access is preferable to offline, or near-line storage.
- New requirements and service level agreements have created the need for faster access to records.
- Need for location independent data, enabling technology refresh and migration.
- New regulations require retention and data protection.
- Traditional storage methods are inadequate.

Currently, fixed content data is the fastest growing sector of the data storage market. Assets such as X-rays, MRIs, broadcast content, CAD/CAM designs, surveillance video, MP3s and financial documents are just a few examples of an important class of data that is growing at over 90% annually.

- User access to fixed content is changing also. Simultaneous, rapid access is key and thus, storage cannot be offline.
- The increase in fixed content is driven by regulatory needs and by advances in technology.
- Because of this explosive growth and changes to user requirements, traditional storage technologies are inadequate.
Shortcomings of Traditional Archiving Solutions

- Tape is slow, and standards are always changing.
- Optical is expensive, and requires vast amounts of media in order to store data of any size.
- Both solutions require 3rd party media management.
- Many times companies retire tape products without warning.
- Many times recovering files from tape and optical is time consuming.
- Data on tape and optical is subject to media degradation.

Traditional archive solutions store fixed content offline where you can not readily access it; for this reason CAS emerged. It simplifies fixed content storage and retrieval. With CAS, you get fast, affordable online access to your fixed content assets and benefit from lessons learned from the shortcomings of traditional solutions such as tape and optical.
Benefits of CAS

- Immutability and authentication
- Location independence
- Single instance storage
- Faster record retrieval
- Record-level retention, protection, and disposition
- Technology independence
- Online (like Disk)
- Optimized TCO
- Scalability

The benefits of CAS include:

- Immutability and authentication – A unique Content Address is derived from the content. It is used to enable Write Once Read Many (WORM) media and to authenticate content. It protects against malicious modifications.
- Location independence – CAS completely insulates the client application from the physical location of the content. CAS provides content mobility across physical placement and geographic locations. It uses a unique identifier that applications can use for retrieval.
- Single instance storage – CAS uses multiple metadata tags, each tailored to specific user’s requirements, can point at the same piece of unique content
- Faster record retrieval – CAS maintains all content on-line. Thus, there is immediate, rapid access to information.
- Record-level retention, protection, and disposition
- Technology independence – Object based systems are neutral to the storage media in use, making it much easier to migrate to new storage without disturbing the integrity of achieved materials.
- Online (like Disk) – immediate availability of all content.
- Optimized TCO
- Scalability
Lesson: Summary

Key points covered in this lesson:

• CAS Definition
• CAS Description
• Benefits
Lesson: Elements of CAS

Upon completion of this lesson, you will be able to:

• Describe the Physical Elements of CAS.
• Describe the Logical Elements of CAS.

We have looked at what CAS is and the benefits of a CAS strategy, now let us look at the details of the elements of a CAS solution.
Physical Elements of CAS

- Storage devices (CAS Based)
- Servers (to which storage devices get connected)
- Client

The physical elements of CAS include the storage devices and the servers to which the storage devices get connected. The storage devices generally are disk-based storage. Clients get connected to the server to access the CAS data.
Logical Elements of CAS

- The Logical Elements of CAS include the Object-Level Access Protocols.

Logical attributes of CAS include the use of Object-Level Access Protocols. As has been previously mentioned, CAS uses content addressing to store and retrieve data. To follow the sequence of data from a client to the CAS based storage, new terminology must be defined.

- Application Programming Interface (API) - A set of function calls that enables communication between applications, or between an application and an operating system.

- Content Address (CA) - An identifier that uniquely addresses the content of a file and not its location. Unlike location-based addresses, Content Addresses are inherently stable and, once calculated, they never change and always refer to the same content. If the content changes, then a new CA is calculated for the new content.

- Metadata - Metadata, or data about data, describes the content, quality, condition, and other characteristics of data.
Lesson Summary

Key points covered in this lesson:

- Physical Elements of CAS
- Logical Elements of CAS

CAS Environments contain both hardware (physical) components as well as APIs, Content Addressing and Metadata (logical) components.
Lesson: Data Object Storage and Retrieval

Upon completion of this lesson, you will be able to:

• Describe how data gets stored in a CAS environment.
• Describe how data is retrieved from a CAS environment.

In this lesson, we will look at how CAS stores a data object and how it retrieves an object.
To store data in a CAS environment:

1. End users input their data to content management applications that interface with the CAS system via the Applications Program Interface (API). The API separates the actual data from the metadata.

2. The Content Address (CA) is calculated from the object's binary representation.

3. The content address and metadata about the object, such as its file name and creation date, are then inserted into an XML file known as the Content Descriptor File (CDF) and transferred to the CAS. Note: XML refers to Extensible Markup Language. It allows the flexible development of user defined document types.

4. CAS recalculates the object’s Content Address as a validation step, and stores the object. This is to ensure that the content of the object has not changed. If the data has been modified, then a new CA will be generated, and the object will be stored separately.

5. An acknowledgment is only returned to the API once a mirrored copy of the Content Descriptor File (CDF), and safely stored in the CAS repository. Once a data object is stored in the CAS repository, the API is given a Content ID (also called a content handle).

6. The Content ID is a content address of the CDF, which contains the Content Address of the actual data on the CAS. It is also referred to as a content handle and content reference. Using the content handle, the application can read the data back from the CAS at any time. There is no centralized directory in the CAS and no path names or URLs are used. Where the data is stored on the CAS is transparent to the application.
The process for how CAS retrieves a data object is described below.

1. An object is required by a user or an application.
2. The application queries the local table of content addresses for the needed object.
3. Using the CAS API, a retrieval request is sent along with the content address to the CAS.
4. CAS delivers the requested information to the application which, in turn, delivers it to the client.
Lesson: Summary

Key points covered in this lesson:

• How data gets stored in a CAS environment.
• How data is retrieved from a CAS environment.

Content is stored and retrieved in CAS systems using the unique content address.
Healthcare represents the fastest growing market CAS market segment. A typical installation includes a PACS (Picture Archive Communication System) front end that captures and acquires the images. From there the specific environment design will determine how images are stored. The diagram above shows Images initially stored to both storage platforms. Storing images on tier 1 storage satisfies the performance requirements of the implementation while storing data on CAS satisfies the regulatory requirements. As patient data ages, information can be migrated to CAS for long term archive. Any data previously stored will be “single instanced” and will also serve to free up space on the primary storage platform where the cost per megabyte is much higher.
Financial Example: CAS Solution

Check images maintained in tier 1 storage for 60 days then migrated via HSM to “active archive”

For the next 60 days, check images may be requested by regional banks or individual consumers for verification purposes, at a rate of about \( \frac{1}{2} \) percent of the total check pool. Beyond 60 days, access requirements drop dramatically, to as few as 1 for every 10,000 checks. In this case, the check images would be stored on a CAS system starting at day 60 and held there indefinitely. A typical check image archive can approach 100TB.

Check imaging is one of many financial service applications requiring the content storage facilities of CAS. Customer transactions initiated by e-mail, contracts, and security transaction records also need to be kept online for as long as 30 years.
Module Summary

Key points covered in this module:

- Benefits of CAS based storage strategy.
- Overview of physical and logical elements of CAS.
- Storing and retrieving data from CAS.
- CAS application examples.
✔ Check Your Knowledge

- What are the key features of a CAS implementation?
- What are the benefits of a CAS Storage Strategy?
- What are 3 business applications that would benefit from CAS technology?
- What are the logical elements of a CAS system?
- How does data get stored in a CAS environment?
Apply Your Knowledge

• After completing this topic, you should be able to describe the features of a Centera CAS solution.
CAS represents a new paradigm in archive technology and leverages the lessons learned from the shortcomings of traditional archive solutions.

From a business perspective issues dealing with TCO, multi-user online access, availability, reliability and compliance all need to be addressed.

From a functional perspective, a new solution needs to archive any data type, from any application on any platform, that assures the authenticity of your content, for the life of that content.
Archive data is fastest growing data type in most organizations. All of the bullets in this slide are requirements for solutions in the archive space going forward. Users have to look past simple acquisition cost to the total cost of ownership, reporting requirements and new service level agreements. In most cases traditional solutions like tape and optical do not measure up.
**EMC Centera**

The World’s Most Simple, Affordable, and Secure Repository for Information Archiving

- Purpose-built for information archiving
- More than 1,200 customers
- 400+ partners—works with any application from virtually any platform
- More than 30 PB shipped

Though Centera’s acquisition cost is higher than tape and optical, acquisition cost only accounts for 20% of the total cost of ownership over time. Human resource and media management cost do not scale using tape or optical, as capacity of these solutions increase so do the costs of people and media management. Centera on the other hand is built on top of a self managing, self healing platform that requires very little if any scaling of management or operational personnel.
One of the big differentiators of Centera as compared to traditional archive solutions is that the Centera archive is online which increases the value of the information. Tape and optical solutions often store information offsite which diminishes the value of the information.

Additionally, the Centera archive is accessible using most any application or platform.
From a hardware perspective, all Centera Nodes are the same. The nodes are Pentium based servers with four disk drives and redundant input power supplies. When the cluster is installed, the nodes are configured with a “roll” defining what functionality it will provide for the cluster. A node can be configured as storage only, access only or dual roll providing both storage and access capabilities.

Storage nodes are used to store and protect data objects; they are sometimes referred to as “back end nodes”.

Access nodes are used to provide connectivity to the customer’s LAN. The number of access nodes is determined by the amount of throughput the customer requires in and out of the cluster. If a node is configured as a “pure access node” the disk space of the node can not be used to store data objects. This configuration is generally found in older installs.

Dual roll nodes are nodes that provide both storage and access node capabilities. This node configuration is more typical than a pure access node configuration.
Centera architecture commonly referred to as a “redundant array of independent nodes” (RAIN) provides hardware redundancy in every major subsystem resulting in no single point of failure in the cluster. As objects are stored, they are mirrored on two different nodes.

The content of one drive could be mirrored to any multiple number of nodes within Centera. This enables a very short rebuild time, because Centera does not wait for one drive to rebuild. Rather, it asks the nodes that hold a single instance of content to quickly make a second copy on another available node. Thus, a number of nodes could be mirroring the same content at the same time.

- Centera continuously checks content integrity by:
  - Constantly validating the integrity of its data objects and structure
  - Performing ongoing background data scrubbing: Are there two copies for every object? If not, it automatically makes a second copy.
  - Constantly checking authenticity to prevent data corruption

- Centera determines the physical placement of content objects—System Administrators need not specify and manage the underlying file/LUN structure.

- Centera also performs self-load balancing:
  - Uses multiple active network paths to each node
  - Hands off work to the least busy node
  - Does not send work to nodes that are offline
  - Handles wild load fluctuations
If any component in the node or if the entire node fails, data will be regenerated to a different part of the cluster maintaining two copies at all times. In addition to this “organic regeneration” process, there are other processes that run continuously in the background verifying objects with a scrubbing process ensuring objects have not been corrupted.
Here’s a summary of how Centera handles component failures:

- In the event of a disk failure, a new copy of the failed drive’s objects are made.
- If a node fails, Centera, after ensuring that the node is truly down (not just offline for such reasons as a short-term network issue, a software upgrade, etc.), will send out a request to all nodes that have the one instance of content, and request that a duplicate copy be created. No content will be replicated to a node on the same power rail.
- There are two switches in a rack configuration. If one switch or a switch port/connection fails, all traffic to affected nodes is automatically routed through the other switch/surviving LAN connections.

And here’s some additional information on self-healing:

- System-level regeneration is granular to the disk level.
- A watchdog layer detects software faults and restores any processes.
- Multiple schemes/techniques verify content integrity.
In closing, Centera provides a new concept for storing fixed content data. This simple but clever process protects customers from in-technology changes on a self managing self configuring platform that scales extremely well. All of the Centera features add up to a superior solution for storing and protecting archive data with varying requirements for weeks, months or years.