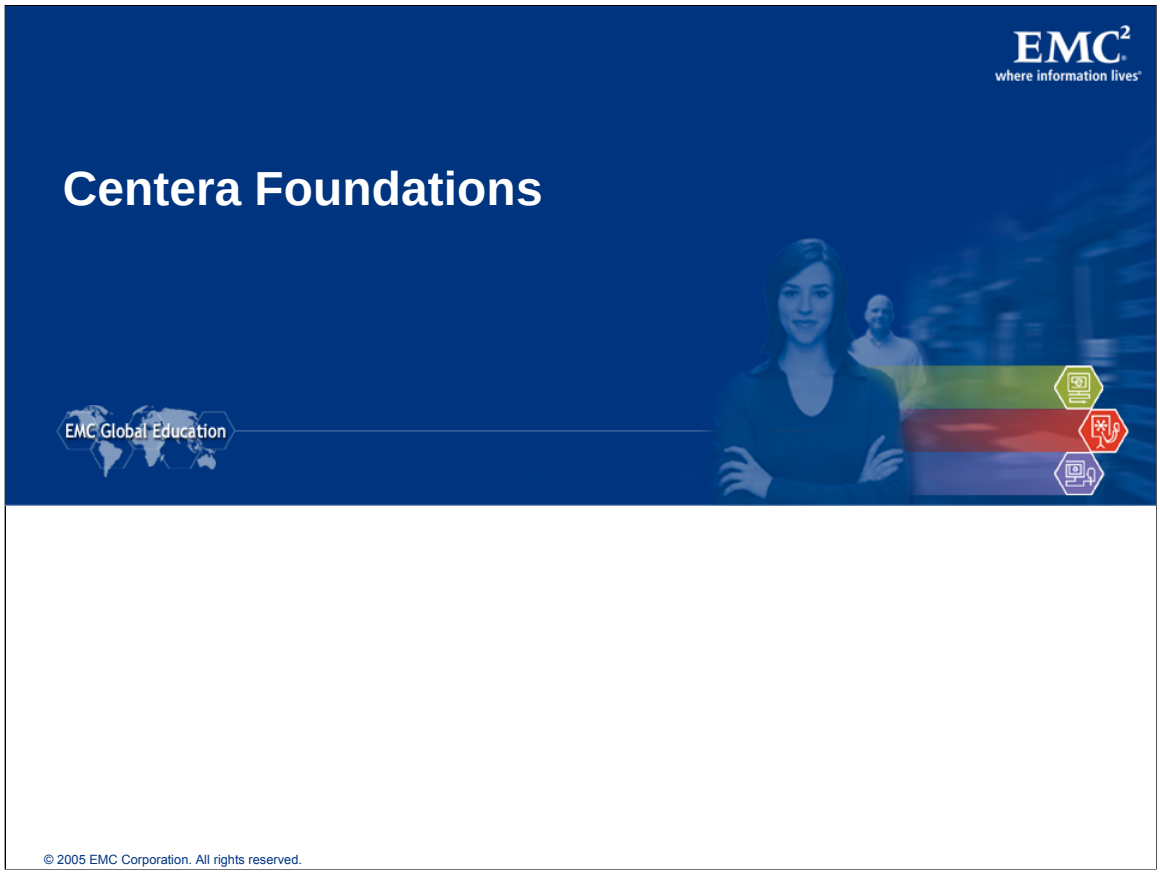
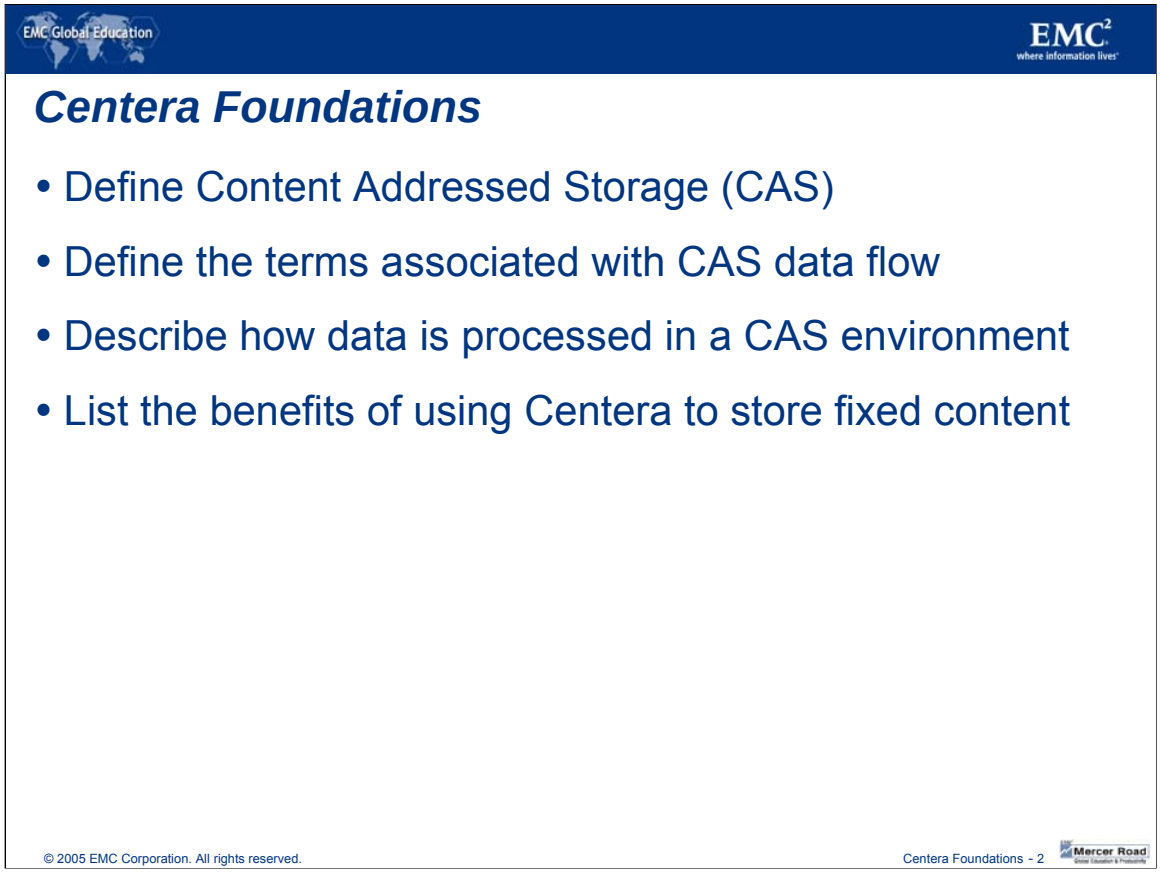


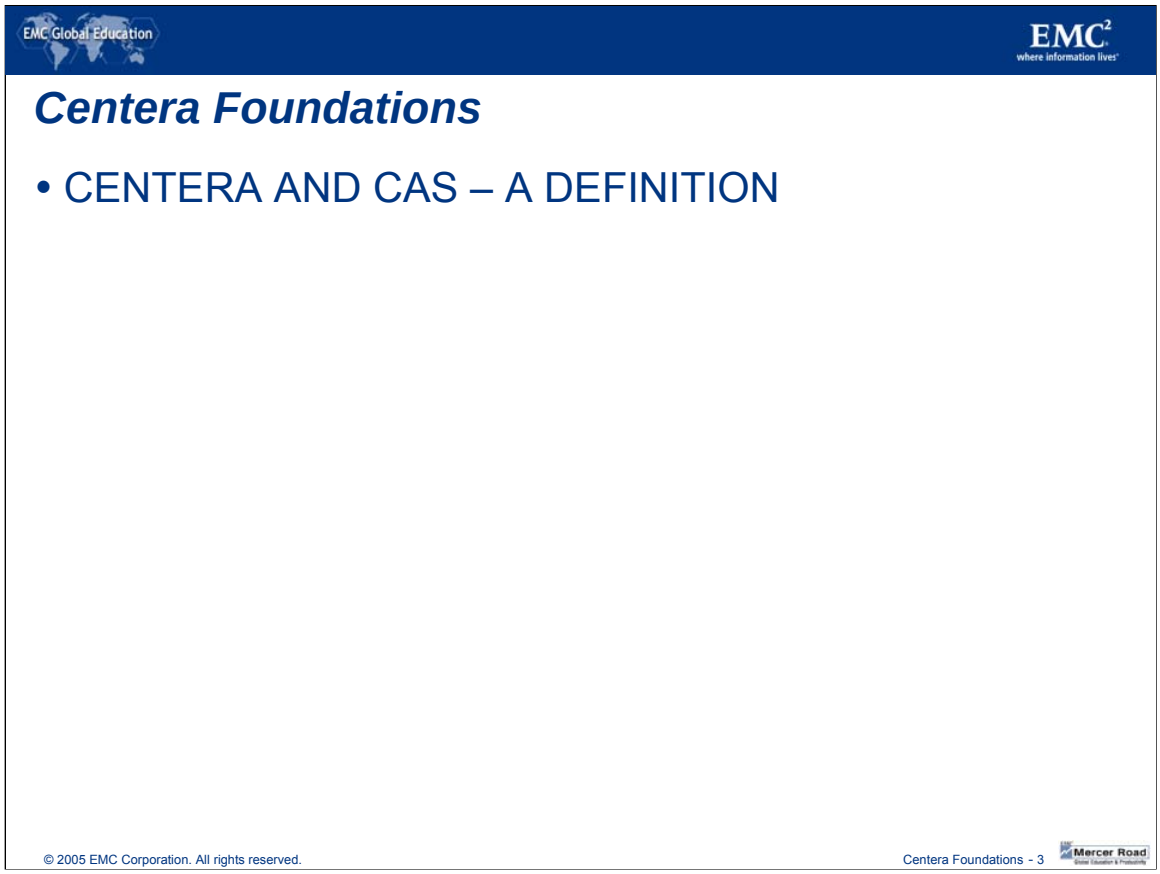
Welcome to Centera Foundations.

The AUDIO portion of this course is supplemental to the material and is not a replacement for the student notes accompanying this course. EMC recommends downloading the Student Resource Guide from the Supporting Materials tab, and reading the notes in their entirety.

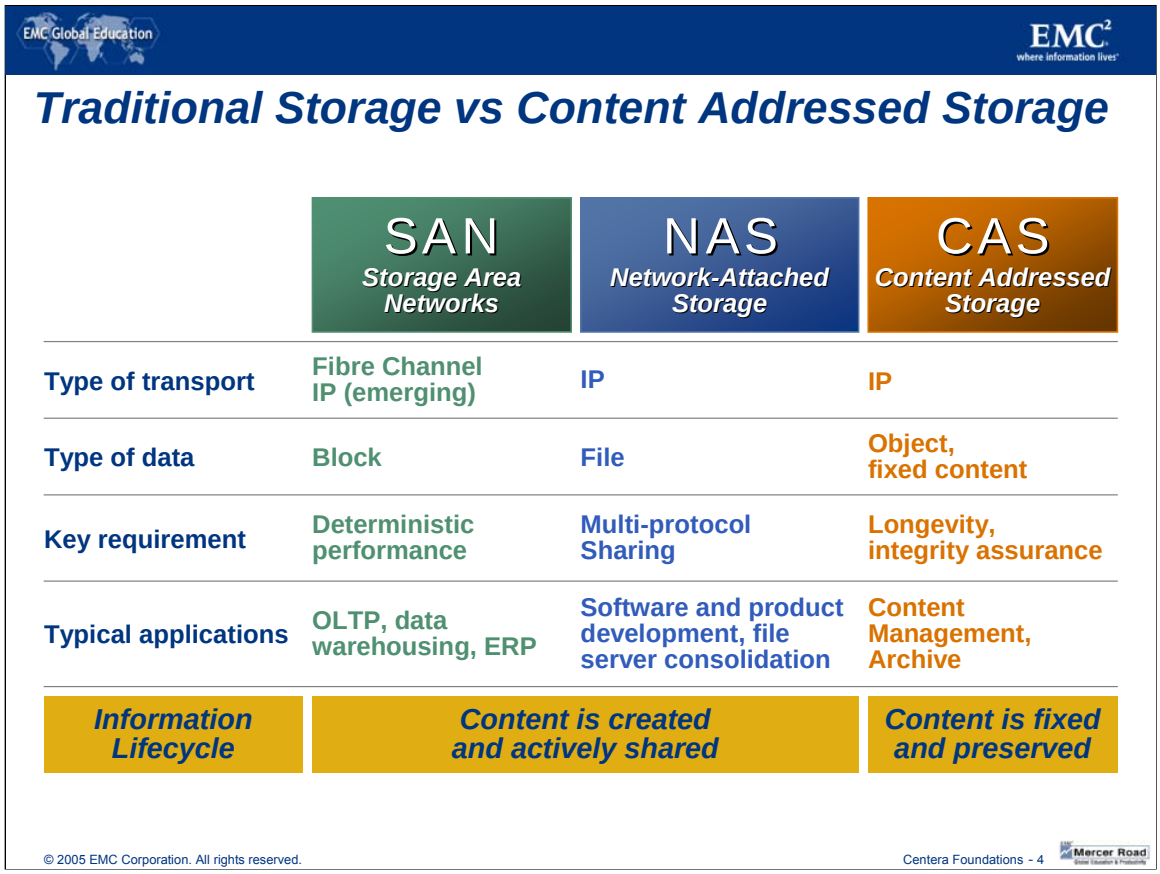


Economic changes in communications and storage have made it necessary to move fixed data content to a network-accessible format. EMC's Centera is the first of its kind platform, designed and optimized specifically to deal with the characteristics of fixed data content. Huge forces are driving customers to manage that content on-line. For example, new regulatory requirements compel healthcare providers and insurance companies to provide access to medical records. Centera's RAIN (Redundant Array of Independent Nodes) architecture and use of C-Clip technology for content addressing allow organizations to meet these needs.

The objectives for this course are shown here. Please take a moment to read them.



In this first section, we will briefly define what Content Addressed Storage (CAS) is and discuss the hardware platform that supports it, namely, Centera.

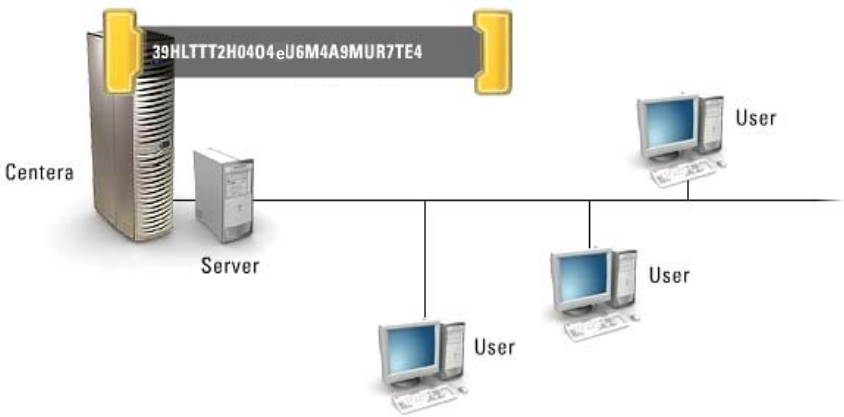


Traditional disk storage systems use block or file access schemes that are well suited to transaction oriented, update intensive, data storage solutions. In a fixed content environment, it becomes a challenge to manage the logistics of data placement and capacity scaling, while also assuring authenticity of the content over its lifetime.

EMC Global Education

EMC² where information lives

What is CAS?

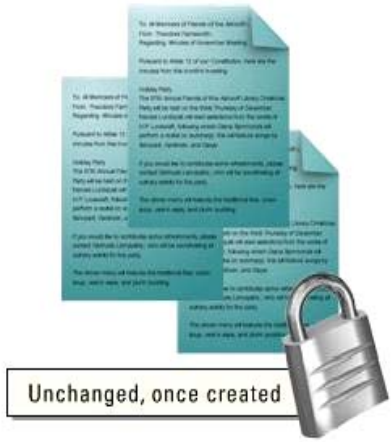


Mercer Road

CAS (Content Address Storage) is a new category of storage designed for the secure online storage and retrieval of fixed content. Rather than access a data object by its file name at a physical location, a CAS device uses a Content Address to store and retrieve the object, where the address of the object (e.g. a file) is created from the unique content of that file. The Content Address is a globally unique identifier, generated by a hashing algorithm. Content Address or "content addressing" will be discussed more thoroughly throughout this module.

The CAS market cuts across multiple vertical industry segments. In each of these market segments, content must be preserved intact for years, if not decades. This kind of content has often ended up on tape or optical disk where, if the data can still be accessed, retrieval may be so delayed that the usefulness of the information is often negated.

What is Fixed Content?



Data in its final form

Fixed content refers to any informational object retained for future reference and business value, including electronic documents and many types of newly digitized information. Unlike transactions or files, it is typically unchanged once created. If you think about the lifecycle of information, it ultimately all leads to fixed content. Content, like email, clinical trial data, CAD/CAM drawings, or electronic documents, may begin as transactional or collaborative work but ultimately becomes fixed content. It is at this point that its value comes from expanded use and not its ability to change.

Fixed content is often contained in large, long-lifetime objects. The quantity is constantly expanding. Regulatory, auditing, and consumer access needs prevent changes to the information. Frequent and fast retrieval is often required, and there are typically many users in many locations. Online availability significantly increases the business value of archived reference information.

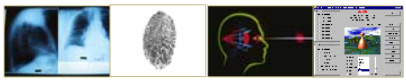
EMC Global Education

EMC² where information lives

Examples of Fixed Content

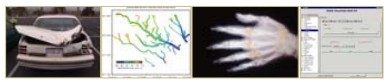
Unchanging Data Objects With Long-Term Value

X-rays
MRIs
CAT scans
Blueprints
Insurance photos
Contracts



Newspapers
Periodicals
Books
Marketing collateral
Check images
White papers
CAD/CAM originals
E-mail and attachments

MP3s
Movies
Recordings
Transcripts
Professional photos
Consumer photos
Educational videos
Surveillance videos
Seismic data
Astronomic data
Spreadsheets
Graphics
Source code



Training materials
Manuals

Genomic data
Proteomic data
Clinical trial results
Biometric data
Lab notebooks
Backups
Historical documents
Presentations
Monthly reports
Video conferences
Audio conferences
News clips
Sports videos
Government records

Mercer Road

Traditionally, the majority of fixed content has been stored on tape or optical technologies. While these technologies can store this content, none of them, nor traditional magnetic disk solutions, were built to handle the very unique requirements for storing final form content. CAS can:

- Provide online access with assured content authenticity
- Efficiently store the content by eliminating the storage of duplicate content
- Scale easily and seamlessly to hundreds of terabytes
- Provide low administration costs by having self-configuring, self-healing, and self-managing functionality

When looking for optimum solutions for fixed content, tape and optical solutions are inadequate. They are too slow, there have been too many in-technology changes that have resulted in lost or unusable content, and reliability is questionable (a tape concern), as is the industry's commitment to the technologies (a point specific to optical).

Common storage alternatives have not been designed with the storage management capabilities found in Centera. These typically do not scale beyond a few terabytes (and/or individual devices) before the operational complexities (and costs) become a significant barrier. For example, if an application requires more storage than fits within a single volume or physical storage device, management complexity increases significantly. Not only is the application challenged by the expanding filesystem hierarchy, but the storage manager is faced with time-consuming reallocation and data relocation, not to mention the complexities of replicating information to multiple sites for purposes of sharing or disaster recovery.

Mercer Road

Requirements to store and retrieve fixed content items are much different and, in many ways, more taxing, than those of transactional data (which is traditionally handled by SAN and NAS storage solutions). Organizations need new ways to manage these increasing amounts of reference information, which is typically unchanged once it is created, and may need to remain online for many years due to regulatory or consumer access requirements. Centera fulfills these requirements by providing faster record retrieval than traditional backups, single instance storage, guaranteed content authenticity, and self-healing, as well as numerous industry regulatory standards.

EMC offers networked storage solutions for every business need: SAN for business and technical applications requiring optimized transaction performance; NAS for high-availability file sharing and collaboration; and CAS for storage and retrieval of fixed content. Whether you need SAN, NAS, CAS, or a combination, only EMC can deliver and integrate all three to work together seamlessly in your environment.

Centera Models

- Centera Basic
  - Provides all functionality without enforcement of retention periods
- Centera Governance Edition
  - Process-centric on the lifecycle of electronic records and enabling policies and technologies
  - Restricts the retention and deletion of data but does not conform to SEC regulations
  - Suitable for most regulations
- Centera Compliance Edition Plus (CE+)
  - Designed for the strictest of regulation requirements, specifically SEC 17a-4
  - Restricts the retention and deletion of data according to SEC regulations

The Centera is offered in 3 different models, dependent on the needs of the individual customer. These needs are based on the data being stored and the stringency of the customers regulatory needs. The Centera Basic model is the least restricted of the models and does not provide any enforcement of retention periods or data shredding. The Centera Governance Edition is suitable for most regulatory needs and does restrict retention and deletion of data. Centera Compliance Edition Plus is the most secure model and follows the strictest regulation requirements demanded by the SEC.

We will give a few examples of the regulations, and how Centera features meet and fulfill their requirements, later in the presentation.

EMC Global Education

EMC² where information lives

Centera Advantages –

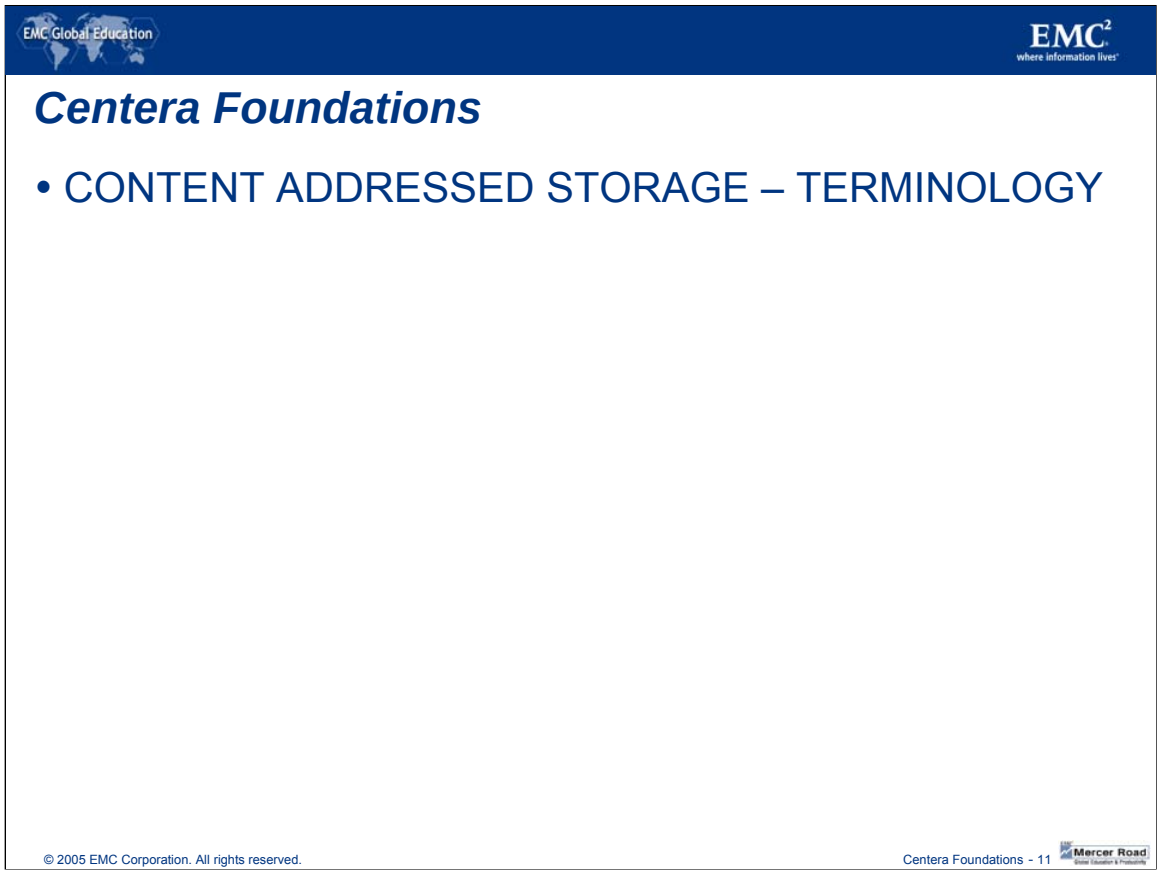
Faster record retrieval than other methods of storing fixed content

- Content Addressing
- WORM Technology
- GUID
- Single Instance Storage

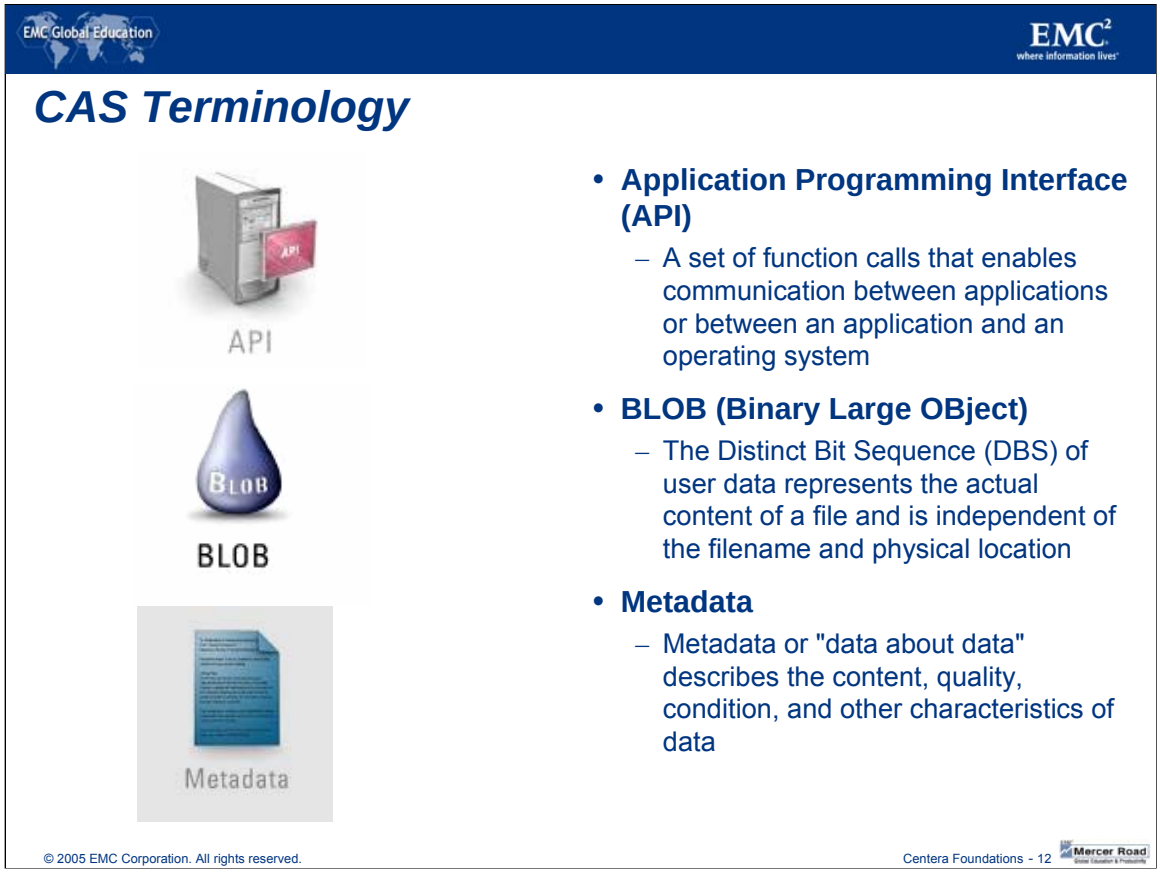
Mercer Road

Centera employs a unique storage/retrieval method called "content addressing", which is performed through the use of C-Clip technology. Centera offers a vast number of advantages such as WORM (Write Once Read Many) technology, GUID (Globally Unique Identifier), and one instance of data for multiple clients.

Centera enables shared, networked access to a single copy of fixed content at sub-second speeds, enhancing the value and usability of information previously stored in less accessible forms. It is critical in today's business environment to be able to quickly respond to new opportunities. Centera provides access to information that will help organizations be competitive in their marketplace.



As with any new technology innovation, there will always be a number of new concepts and names for those concepts. In this section, we will briefly describe the new objects within CAS.



As previously mentioned, Centera uses Content Addressing to store and retrieve data. To follow the sequence of data from a Client to the Centera, 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.

BLOB (Binary Large OBject)

The Distinct Bit Sequence (DBS) of user data. The DBS represents the actual content of a file and is independent of the filename and physical location.

Metadata

Metadata, or "data about data“, describe the content, quality, condition, and other characteristics of data.

Review of CAS Terminology (continued)

- **C-Clip**
  - A package containing the user's data and associated metadata
- **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
- **C-Clip Descriptor File (CDF)**
  - The additional XML file that the system creates when making a C-Clip. This file includes the content addresses for all referenced BLOBs and associated metadata
- **C-Clip ID**
  - The content address that the system returns to the client. It is also referred to as a C-Clip handle and C-Clip reference

C-Clip

A package containing the user's data and associated metadata.

C-Clip ID

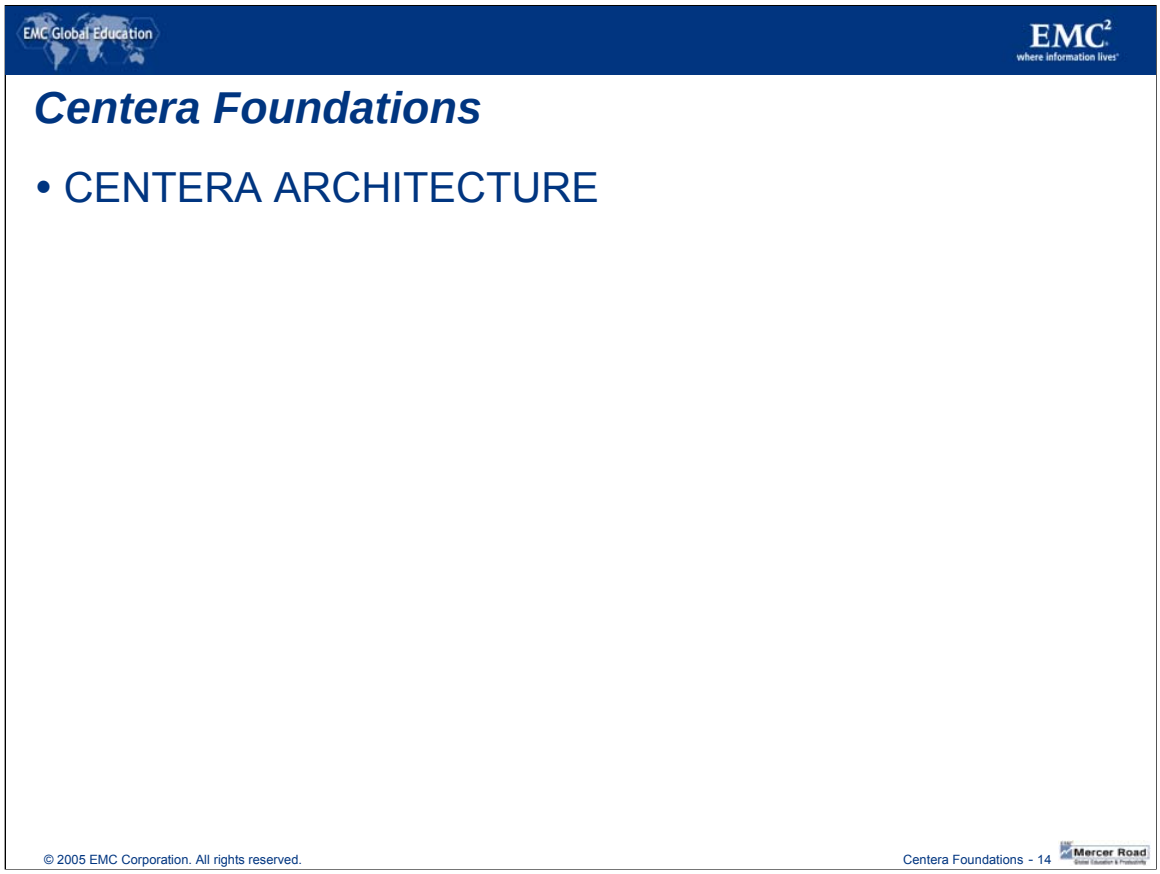
The content address that the system returns to the client. It is also referred to as a C-Clip handle and C-Clip reference.

C-Clip Descriptor File (CDF)

The additional XML file that the system creates when making a C-Clip. This file includes the content addresses for all referenced BLOBs and associated metadata.

Content Address (CA)

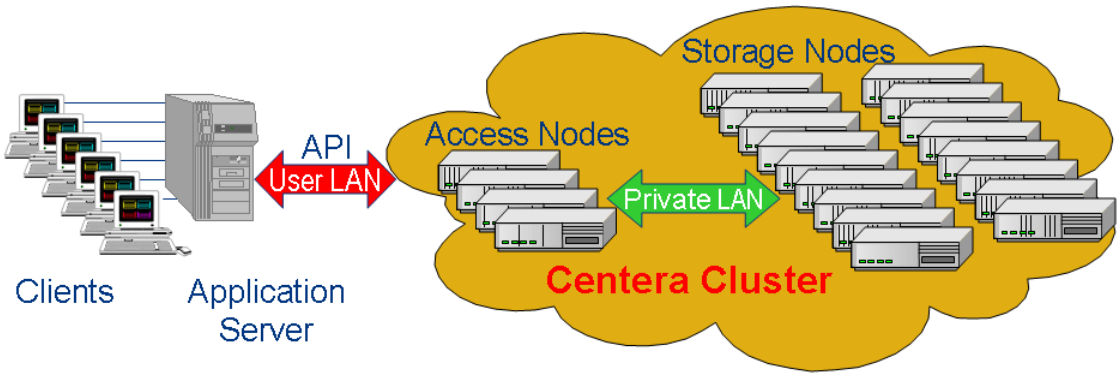
An identifier that uniquely addresses the content of a file and not its location. Unlike location-based addresses, content addresses are inherently stable and, once calculated, never change and always refer to the same content.



In this section, we will briefly discuss the hardware architecture of the solution.

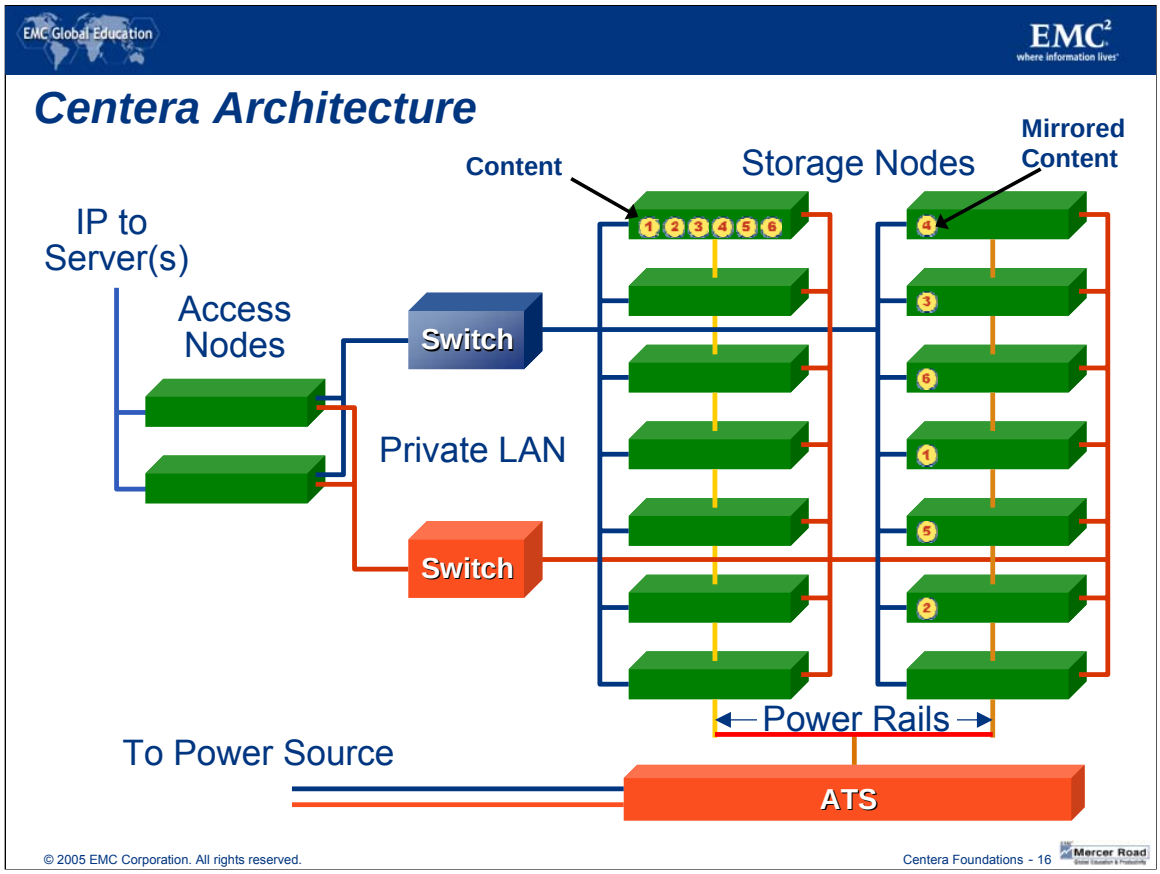
Centera Architecture

Redundant Array of Independent Nodes (RAIN)



The architecture of the Centera system, based on Redundant Array of Independent Nodes (RAIN), is designed to be highly scalable and hold petabytes of content. Each node has its own Linux OS and CentraStar microcode and utilizes a distributed workload. CentraStar is the application (microcode) that runs the Centera and all its features.

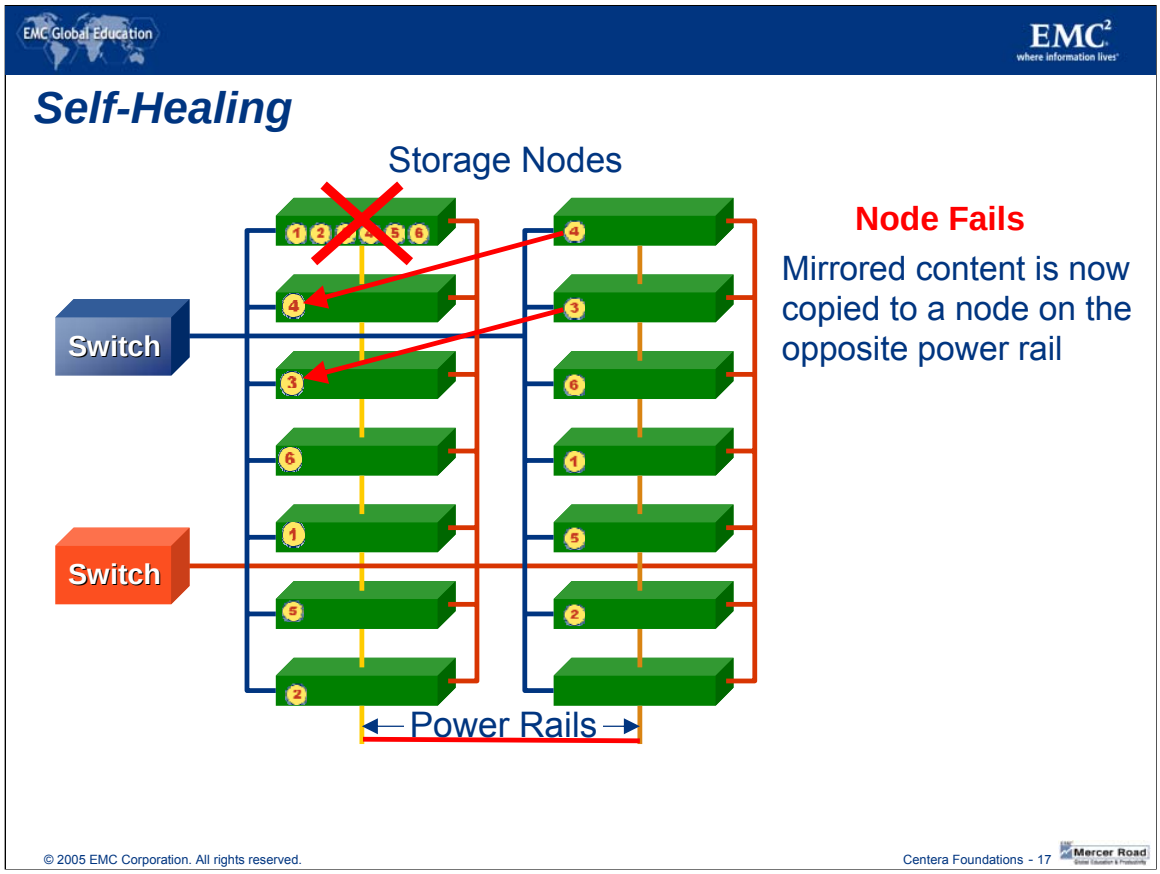
Client applications are written using the Centera API. Application store and retrieval requests are sent to the Access Node via “public” IP connections. The Access Node uses a unique Content Address to locate the requested information from the Storage Nodes over a private internal LAN, and supplies the necessary information back to the client through the API. The Storage Nodes are responsible for the long-term storage and protection of the BLOBs and CDFs.



The Centera cabinet may contain up to 32 nodes, with a minimum configuration containing as few as 8 nodes. Several cabinets can be connected to form a Centera cluster. Applications may support multiple clusters within a Centera domain. Each cabinet has from 4 TB to 23 TB or more of usable protected capacity, depending on whether they have Content Protection Parity (CPP) or Content Protection Mirrored (CPM) set.

There are a minimum of 2 access nodes that are connected to the customer's LAN and to the storage nodes via a private LAN. Each storage node contains more than 1 TB of usable capacity. The internal LAN has 2 48-port cube switches that provide communication between the nodes. Root switches are used for connecting 3 or more cabinets together into a single cluster.

Each cabinet is powered through an Automatic Transfer Switch (ATS), which ensures that power is supplied to the cabinet in the event that one of the two power sources fails.



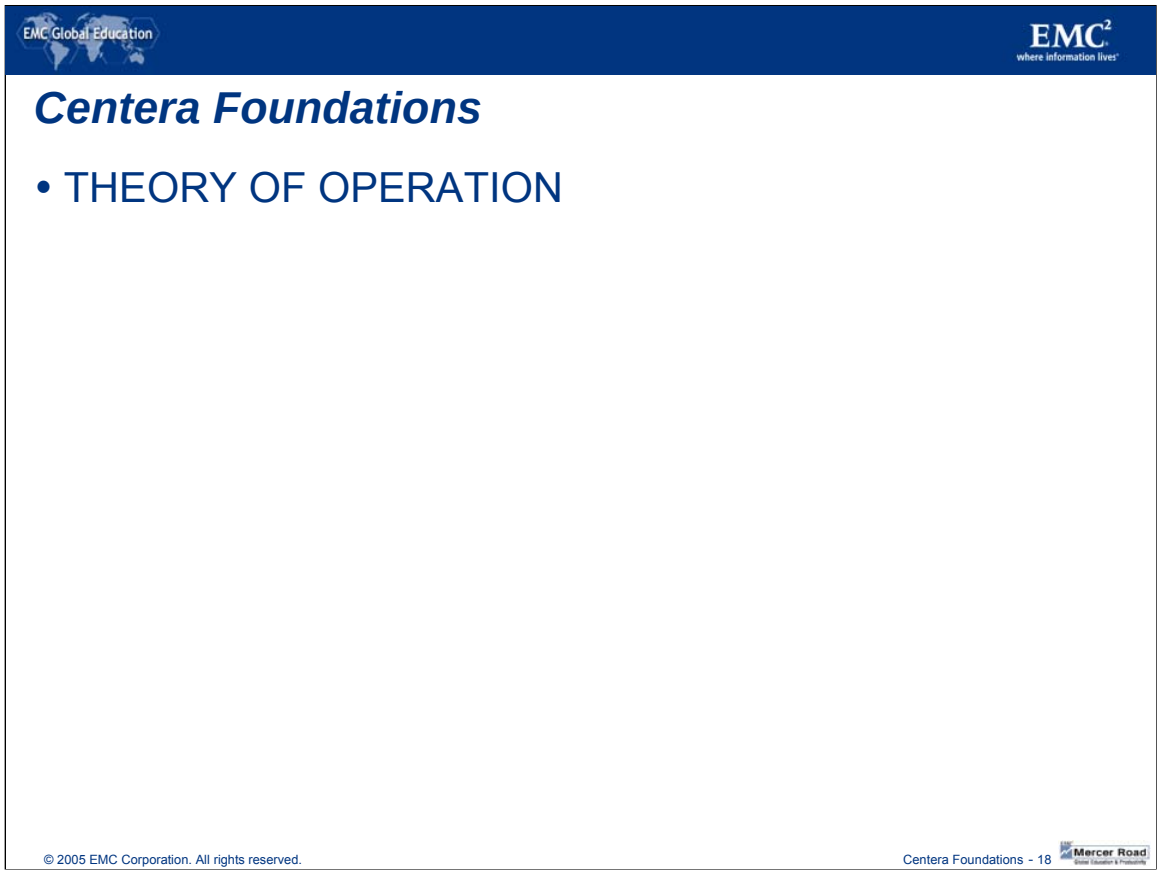
The Centera is a self-configuring, self-healing and self-managing solution. It offers different methods of content protection. The example in the above slide shows Content Protection Mirrored.

CPM (Content Protection Mirrored) is where every data object is mirrored. There are 2 copies of every piece of data sent to the Centera that will reside on different nodes. If a node or disk should fail, the Centera software would automatically broadcast to the node with the mirrored copy to regenerate another copy to a different node so that there will always be 2 copies available.

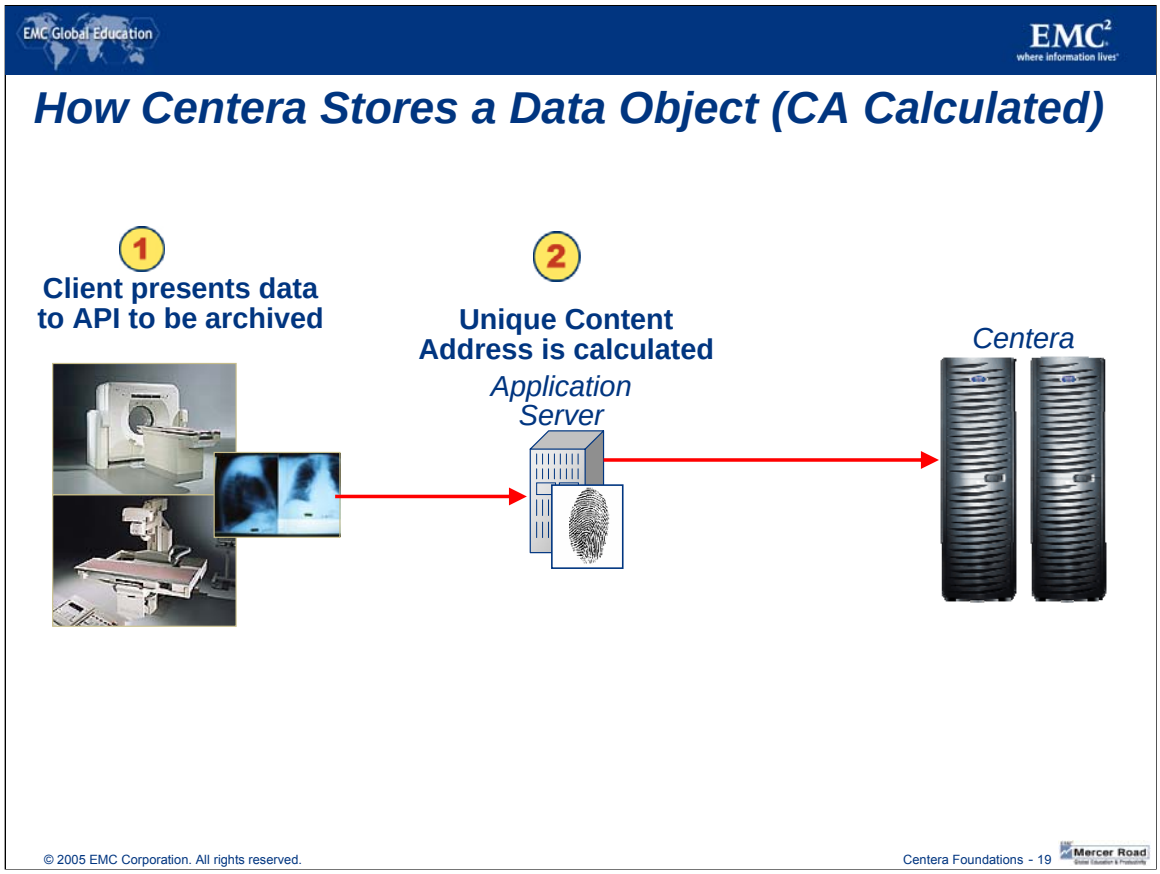
In CPP (Content Protection Parity), the data is fragmented into segments, with a parity segment. Each segment is on a different node, similar to a file type RAID. If a node or disk should fail, then the other nodes would recreate the missing segment and put it on a different node. Either protection scheme provides total protection against failure through the Centera's unique self-healing functions.

Centera performs continuous Content Integrity Checking:

- Centera constantly validates the integrity of its data objects and structure
- Centera does ongoing background data scrubbing
  - If the data is not protected, Centera will automatically ensure the content is protected
- Centera does constant authenticity checking to prevent data corruption
- Automated Garbage Collection



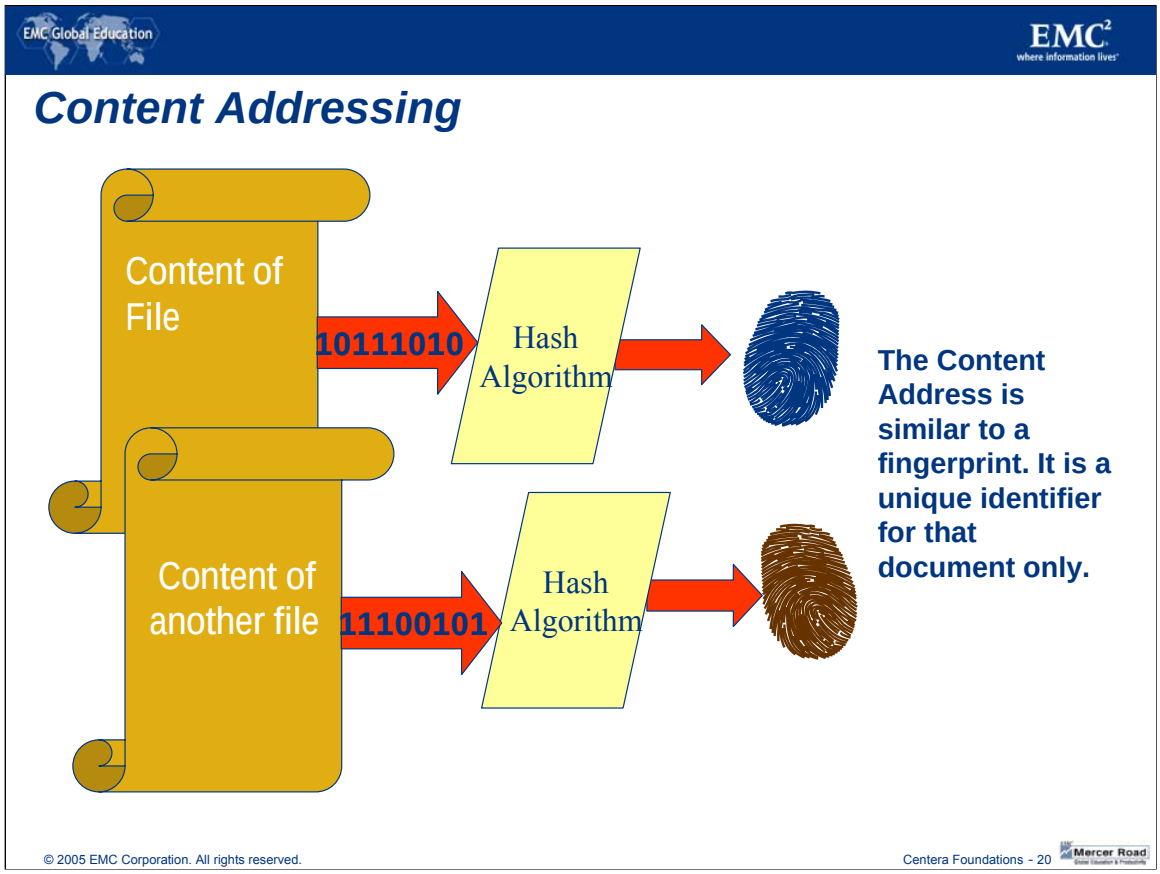
Having discussed the hardware aspect of the solution, we will now briefly discuss the theory of operation behind CAS.



The Centera API facilitates access from an application server to the Centera cluster. Applications that typically interface with Centera are applications that manage the storage and retrieval of fixed digital content such as X-rays, check images, scanned mortgage contracts, and more. This content is fixed in the sense that once it has been stored, it does not change. This type of content is stored according to the WORM principle: “written once and read many”.

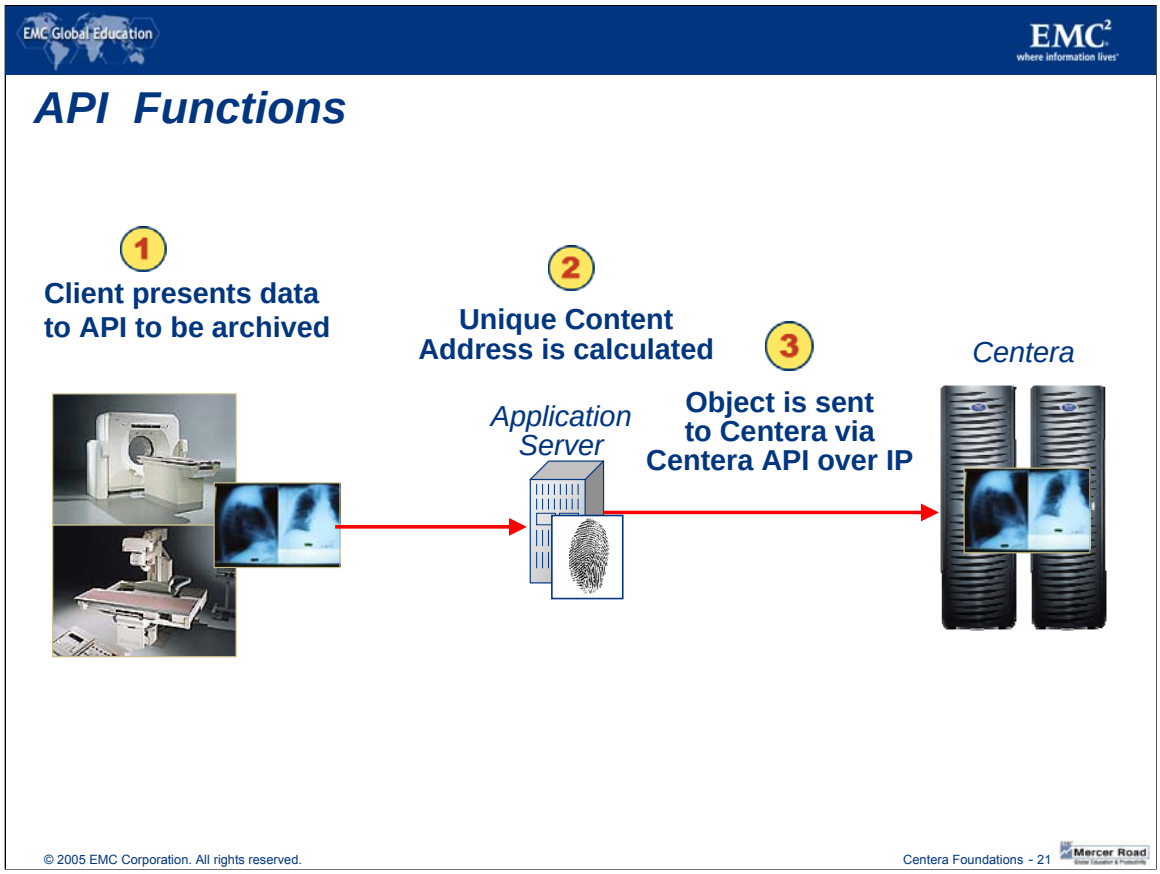
End users input their data to content management applications that interface with the Centera system via the Applications Program Interface (API). The API is part of the Centera SDK (Software Development Kit).

The API separates the actual data (BLOB) from the metadata and the Content Address (CA) is calculated from the object's binary representation.



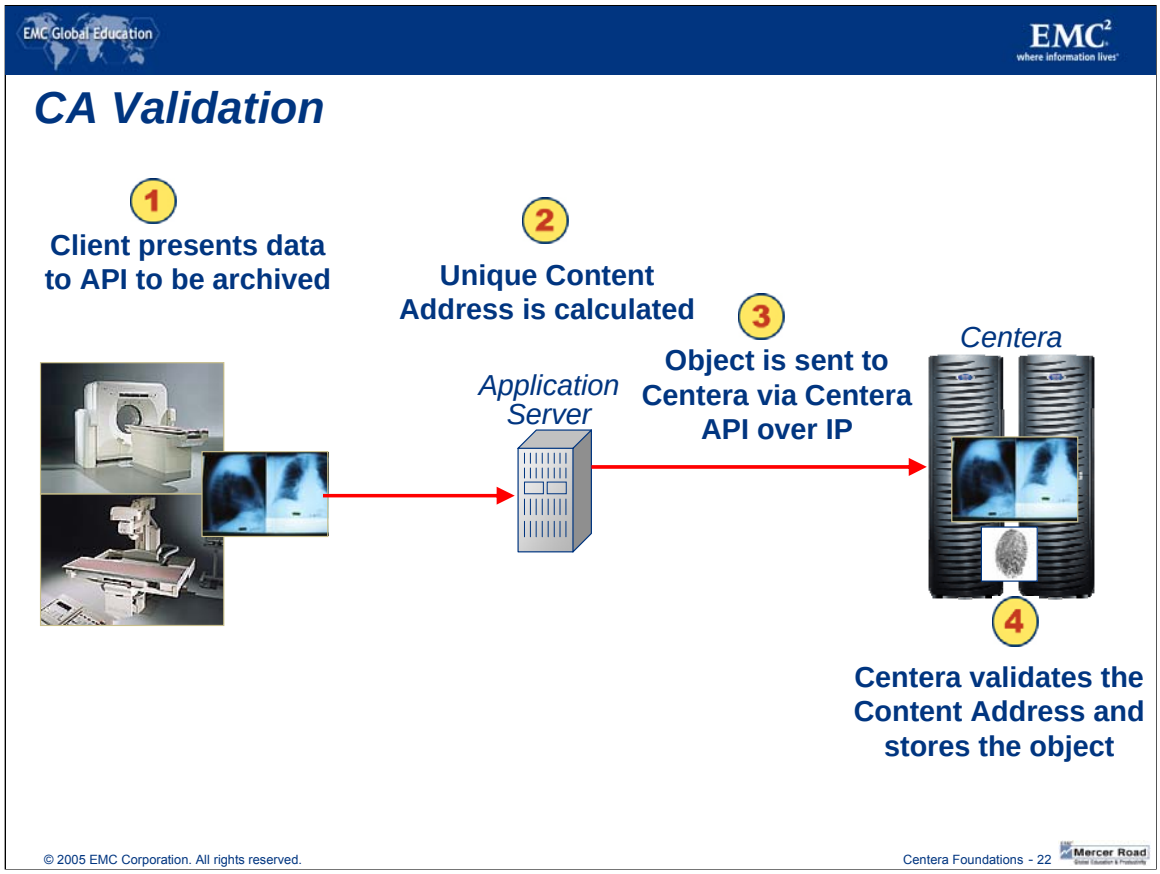
Centera's Content Address is a derived address. It is the result of a hashing algorithm run across the binary representation of the object. It takes into account all aspects of the content, even file type. And what it returns to a user's application is a content fingerprint unique to that content.

A unique number is calculated by the hash algorithm from the sequence of bits that constitutes the content of a file. If a single byte changes in the file, then any resulting calculation will be different. This fingerprint is now used as the Content Address for the data that is to be stored on the Centera. When viewed, this unique number will be displayed in a 27 or 53 character format, depending on the type of storage strategy chosen by the customer.

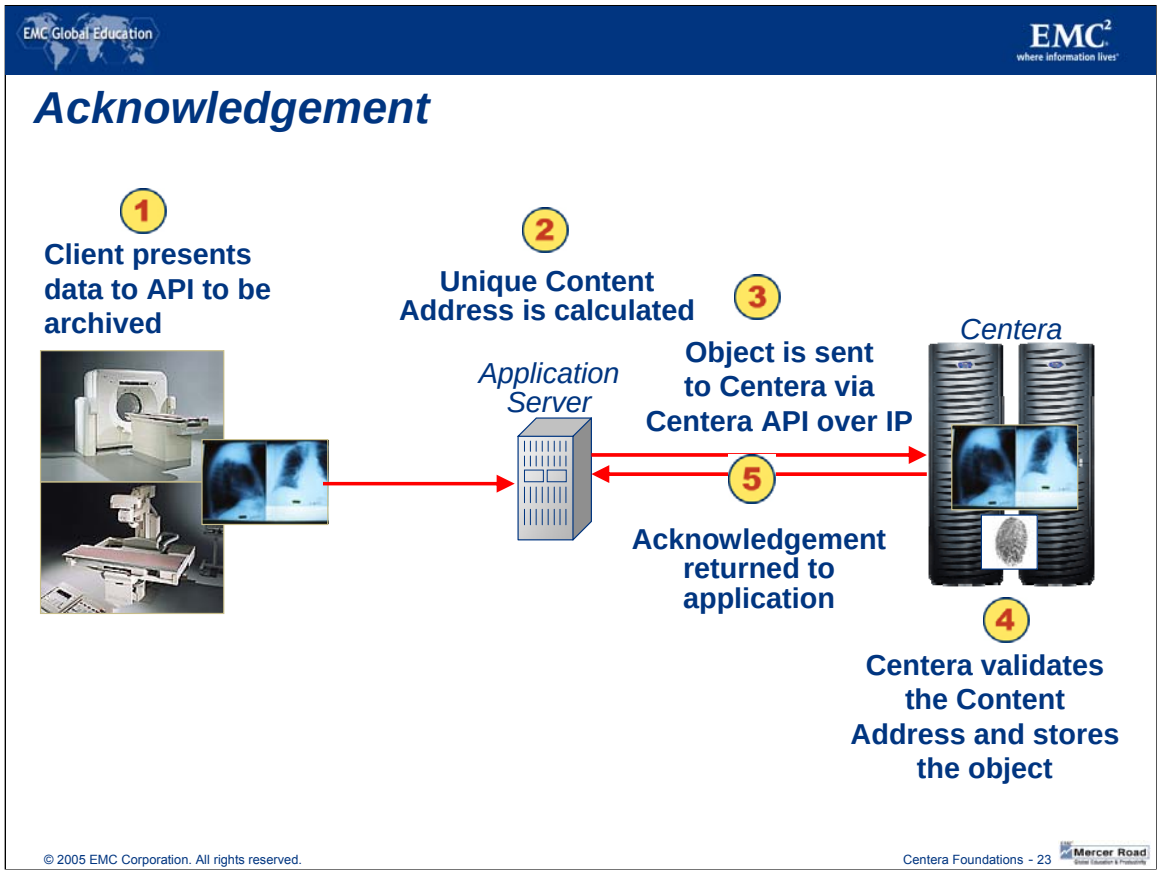


The content address and metadata about the object, such as its file name and creation date, are then inserted into an XML file known as the C-Clip Descriptor File (CDF) and transferred to the Centera. The combination of the BLOB and the CDF is referred to as the C-Clip, which is stored on the Centera.

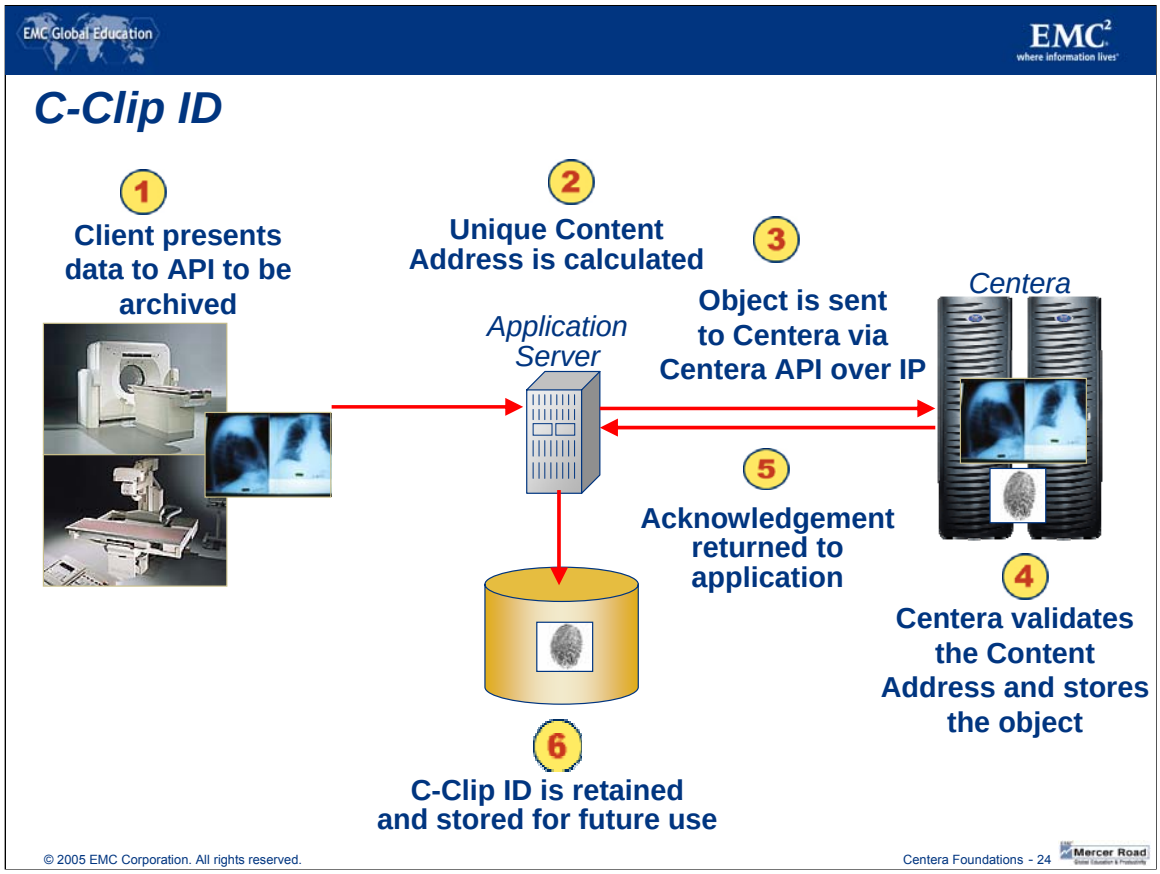
NOTE: XML refers to Extensible Markup Language. It allows the flexible development of user defined document types.



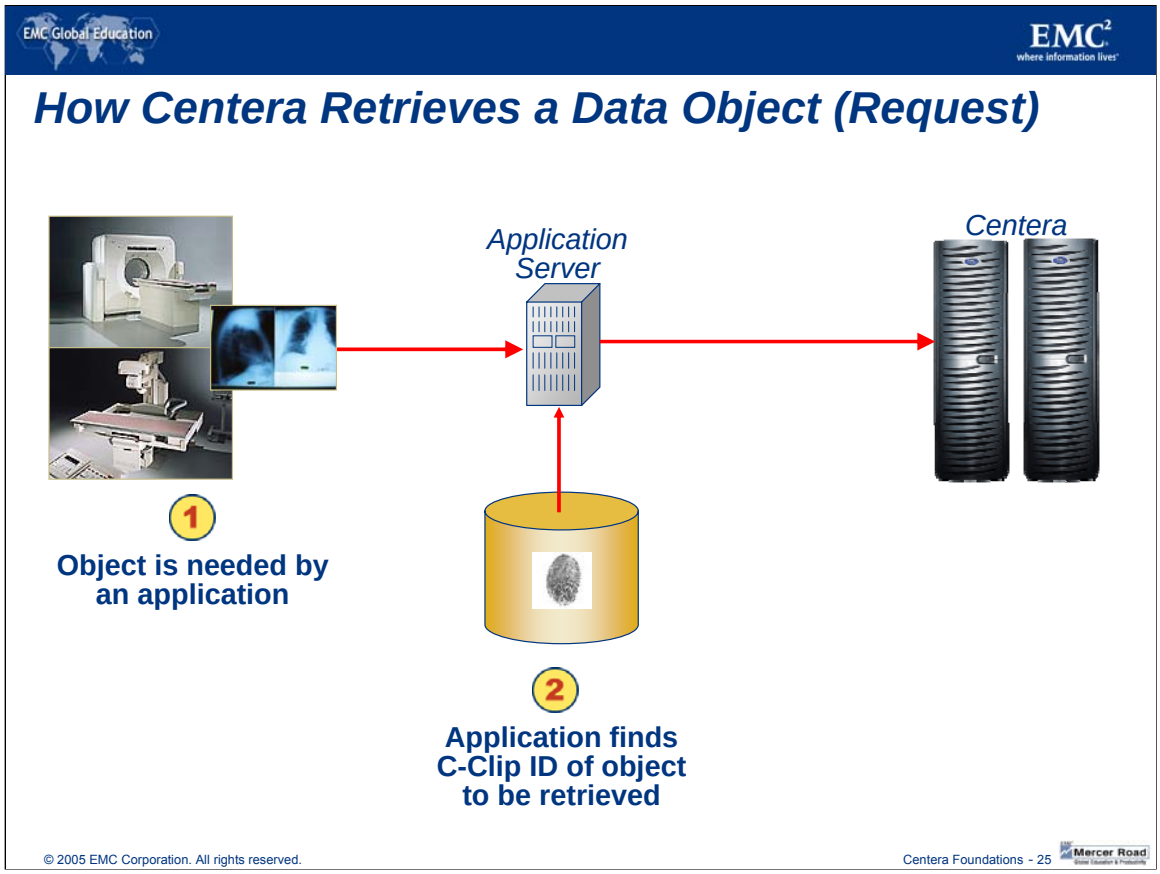
Centera recalculates the object's Content Address as a validation step and stores the object. This is to ensure that the content of the object has not changed. If the data has been modified, then a new CA will be generated, and the object will be stored separately as its own BLOB.



An acknowledgment is only returned to the API once a mirrored copy of the C-Clip Descriptor File (CDF), and a protected copy of the BLOB, have been safely stored in the Centera repository. Once a data object is stored in the Centera repository, the API is given a C-Clip ID (also called a C-Clip handle).



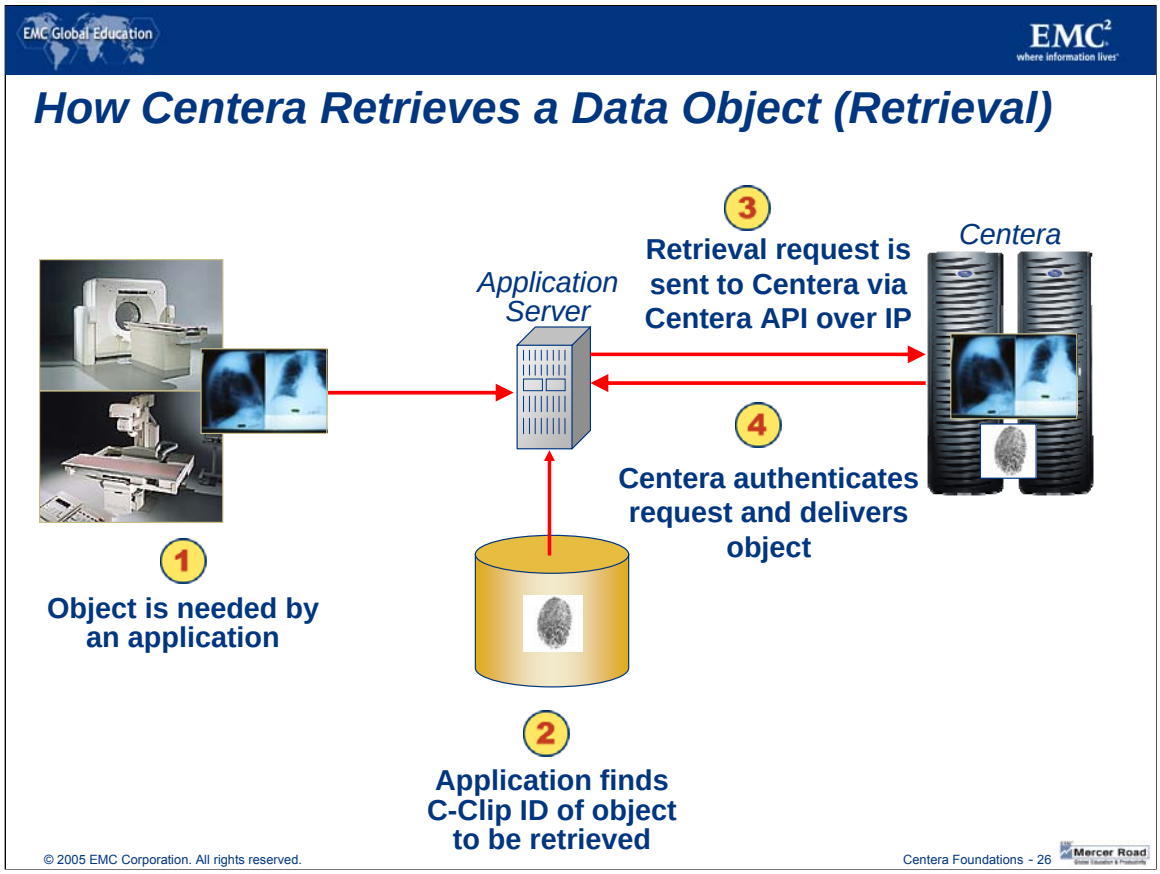
The C-Clip ID is a content address of the CDF, which contains the CA of the actual data on the Centera. It is also referred to as a C-Clip handle and C-Clip reference. Using the C-Clip Handle, the application can read the data back from the Centera at any time. There is no centralized directory in the Centera, and no pathnames or URLs are used. Where the data is stored on the Centera is transparent to the application.



Here is the process on how the Centera retrieves a data object:

Step 1. An object is required by a user or an application.

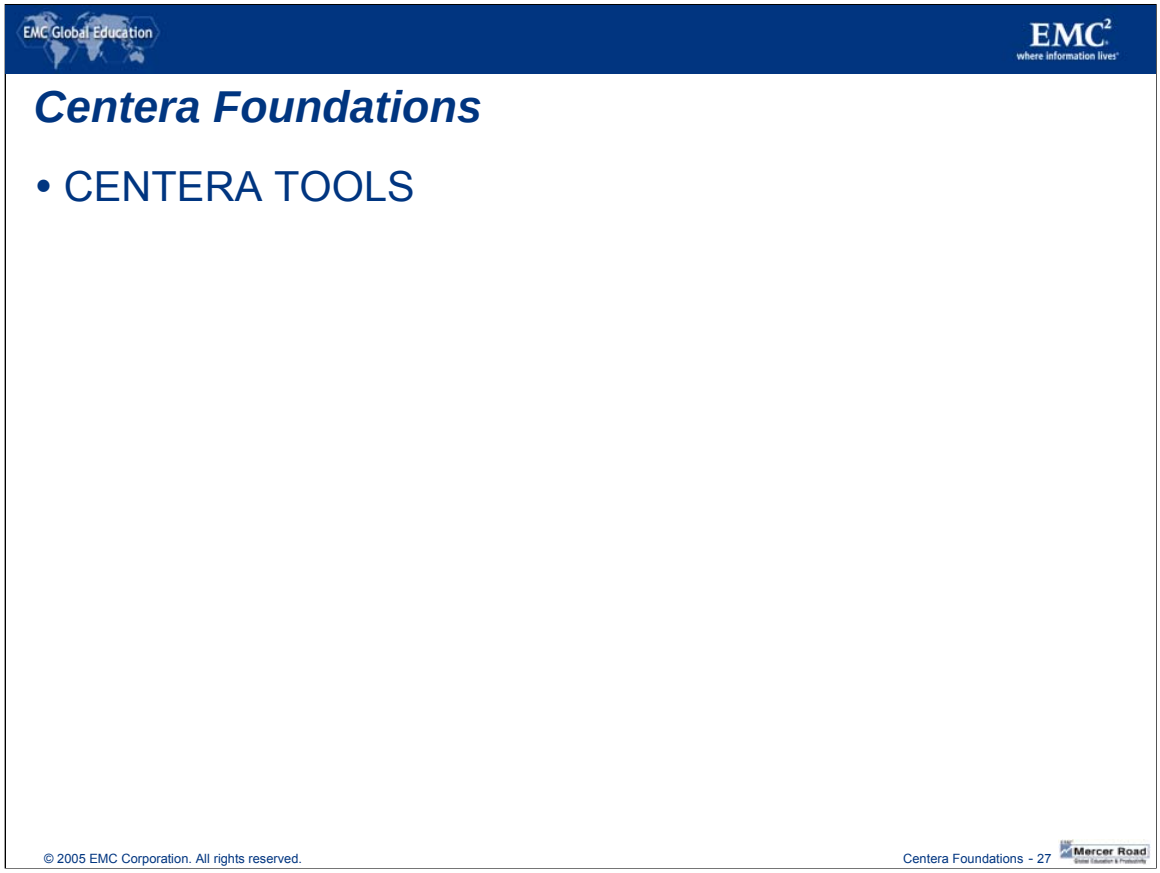
Step 2. The application queries the local table of C-Clip IDs and locates the C-Clip ID for the needed object.



Step 3. Using the Centera API, a Retrieval request is sent, along with the C-Clip ID, to the Centera.

Step 4. Centera delivers the requested information to the application which, in turn, delivers it to the client.

The Client does not need to know where the data is physically located, as it is the Centera that determines where the data is stored. It only requires the unique address that is used to identify the data so that it can be retrieved.



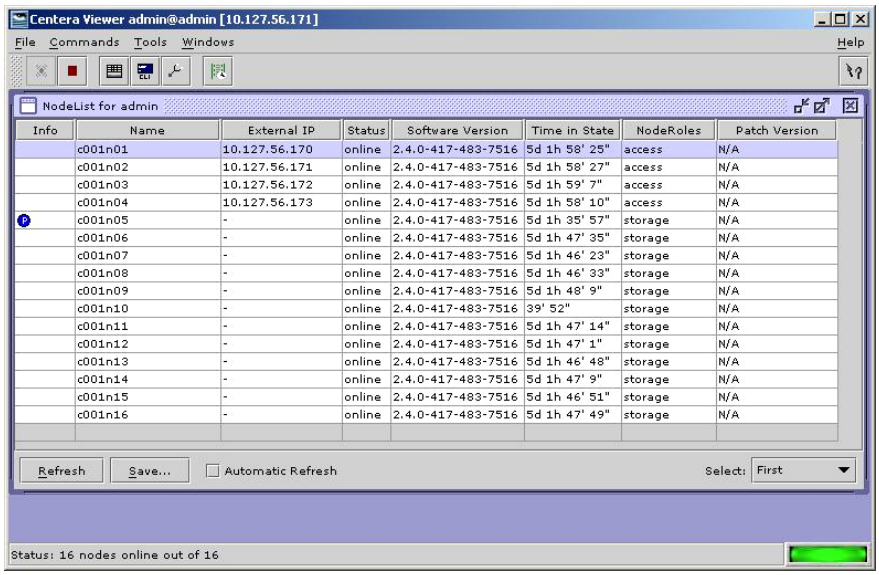
Some tools that are available for administration of the solution will be briefly discussed here.

EMC Global Education

EMC² where information lives

Centera Tools

Example of Centera Viewer



Info	Name	External IP	Status	Software Version	Time in State	NodeRoles	Patch Version
	c001n01	10.127.56.170	online	2.4.0-417-483-7516	5d 1h 58' 25"	access	N/A
	c001n02	10.127.56.171	online	2.4.0-417-483-7516	5d 1h 58' 27"	access	N/A
	c001n03	10.127.56.172	online	2.4.0-417-483-7516	5d 1h 59' 7"	access	N/A
	c001n04	10.127.56.173	online	2.4.0-417-483-7516	5d 1h 58' 10"	access	N/A
P	c001n05	-	online	2.4.0-417-483-7516	5d 1h 35' 57"	storage	N/A
	c001n06	-	online	2.4.0-417-483-7516	5d 1h 47' 35"	storage	N/A
	c001n07	-	online	2.4.0-417-483-7516	5d 1h 46' 23"	storage	N/A
	c001n08	-	online	2.4.0-417-483-7516	5d 1h 46' 33"	storage	N/A
	c001n09	-	online	2.4.0-417-483-7516	5d 1h 48' 9"	storage	N/A
	c001n10	-	online	2.4.0-417-483-7516	39' 52"	storage	N/A
	c001n11	-	online	2.4.0-417-483-7516	5d 1h 47' 14"	storage	N/A
	c001n12	-	online	2.4.0-417-483-7516	5d 1h 47' 1"	storage	N/A
	c001n13	-	online	2.4.0-417-483-7516	5d 1h 46' 48"	storage	N/A
	c001n14	-	online	2.4.0-417-483-7516	5d 1h 47' 9"	storage	N/A
	c001n15	-	online	2.4.0-417-483-7516	5d 1h 46' 51"	storage	N/A
	c001n16	-	online	2.4.0-417-483-7516	5d 1h 47' 49"	storage	N/A

Although System Administrators do not need to worry about volume creation/management or file system structure/maintenance or their data as the Centera controls where the data is stored (location independent). But customers do need to be able to monitor the Centera's capacity and performance, as well as EMC personnel, and partners need to maintain the Centera and to enable and disable its features as needed.

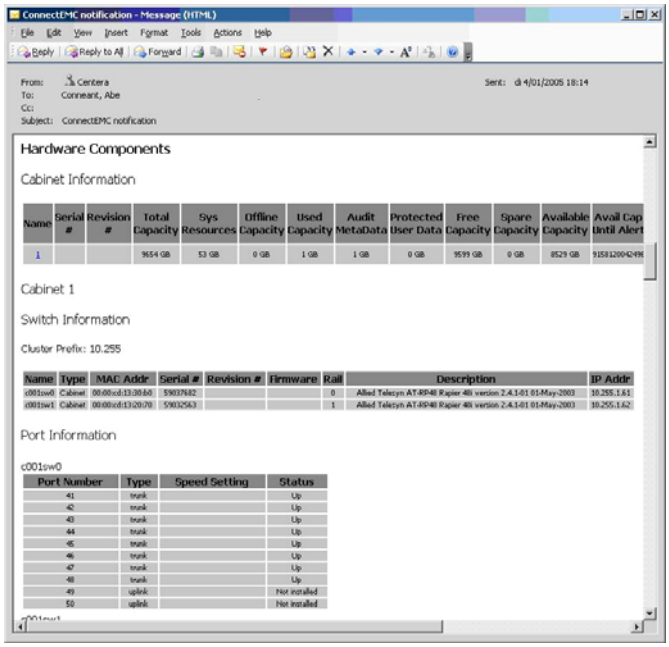
A group of tools is provided, both to the customers as well as the service personnel. The most commonly used tool is the Centera Viewer, a GUI (shown above) loaded onto a Windows PC with network access to the Centera that provides a simple means of displaying capacity utilization and operational performance of the Centera. It also enables the system administrator to change any site-specific information such as the public network information, as well as end-user contact information. It is also a tool most commonly used by EMC personnel and partners to troubleshoot failures and for upgrading the CentraStar code. There is also a Command Line Interface (CLI) that can be used with the GUI or as a standalone tool in a UNIX environment.

Other tools available are Centera Monitor, which allows the customer to monitor a CE+ Centera. Simple Network Management Protocol (SNMP) is used to alert an enterprise network management system to any faults that might occur within a Centera.

EMC Global Education

EMC² where information lives

Centera Health Reporting

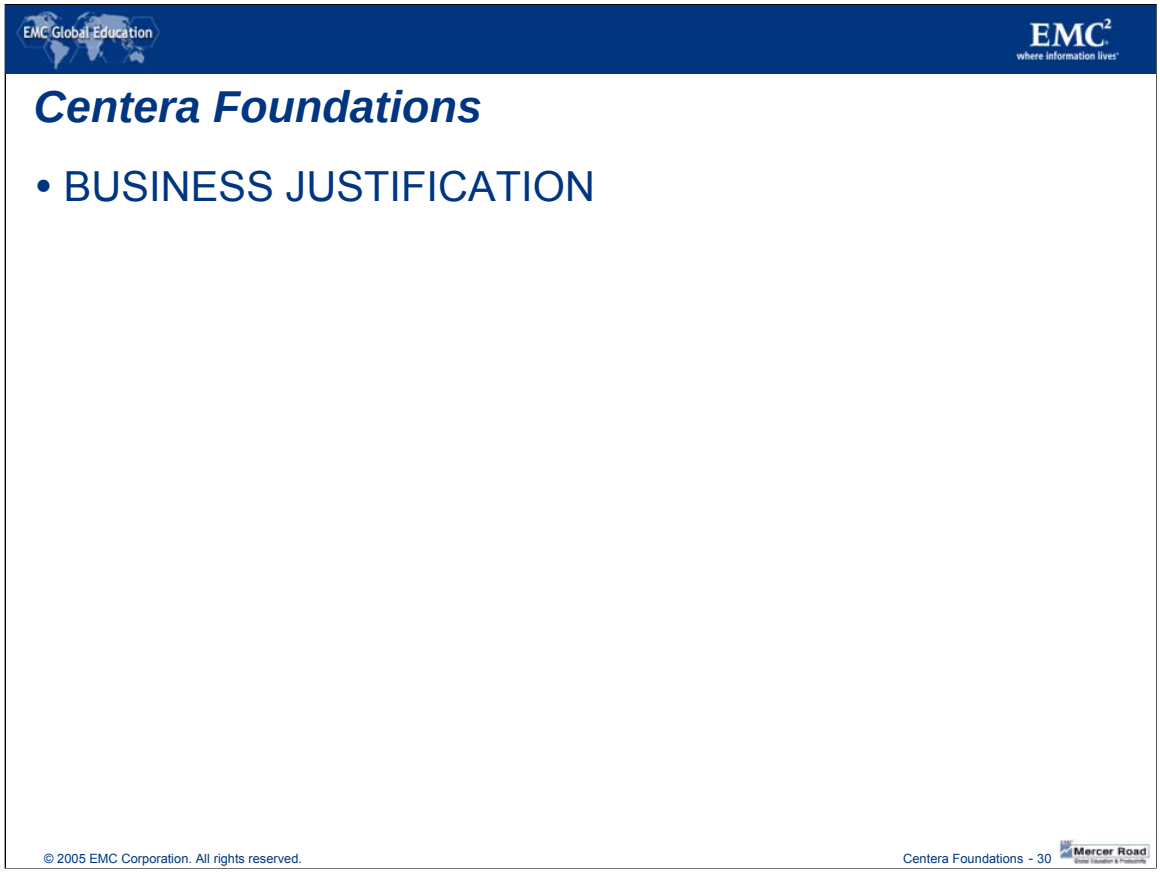


Example of a Health Report in HTML email format

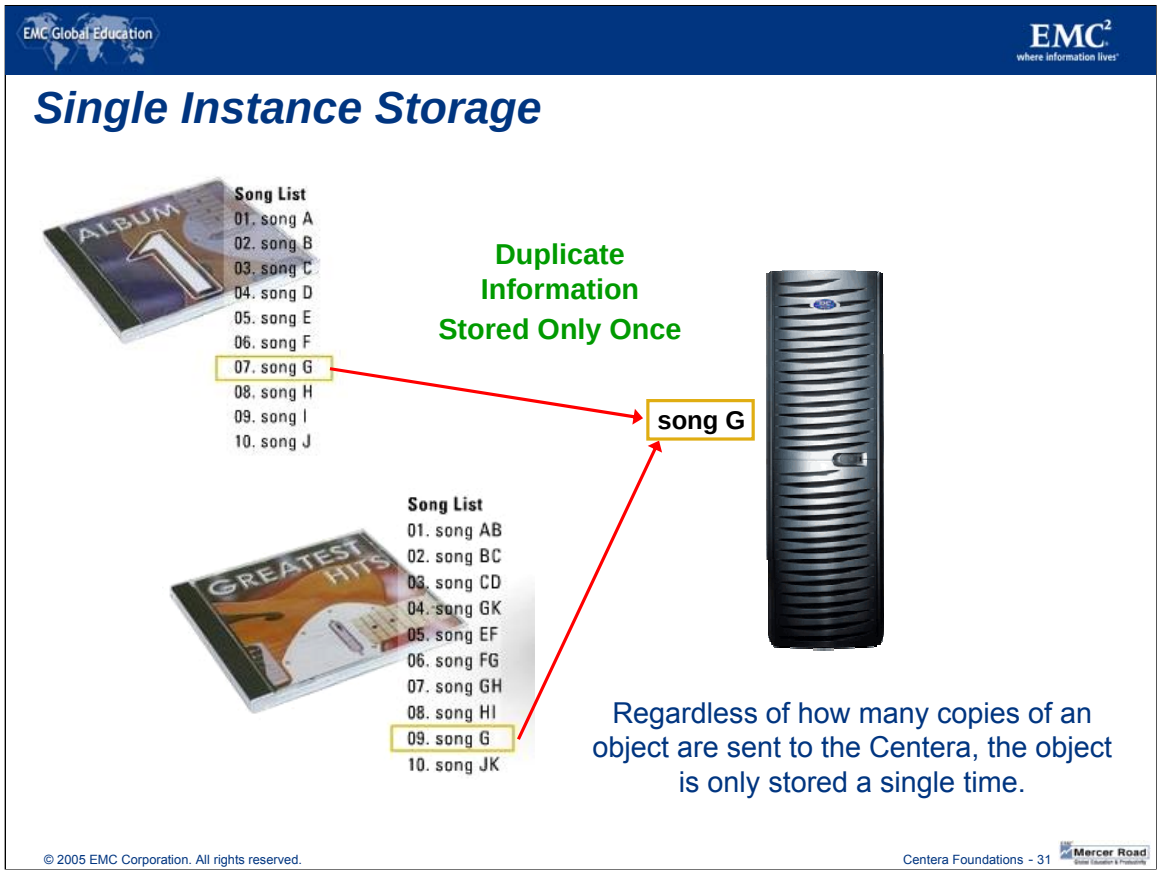
Mercer Road Global Education & Productivity

System Health Report is an automatic email message that a Centera cluster periodically sends to the EMC Customer Support Center with a list of predefined recipients. It reports on the current status of the Centera cluster. The report is sent to the EMC Customer Support Center in an XML format. When it is sent to other recipients, it is converted into an HTML format for easy reading of its data (as shown in the above slide). EMC is then able to monitor the Centera cluster remotely and detect any hardware or software problems.

Mercer Road Global Education & Productivity



How the Centera and the concept of Content Addressed Storage assist businesses will be discussed in this section.



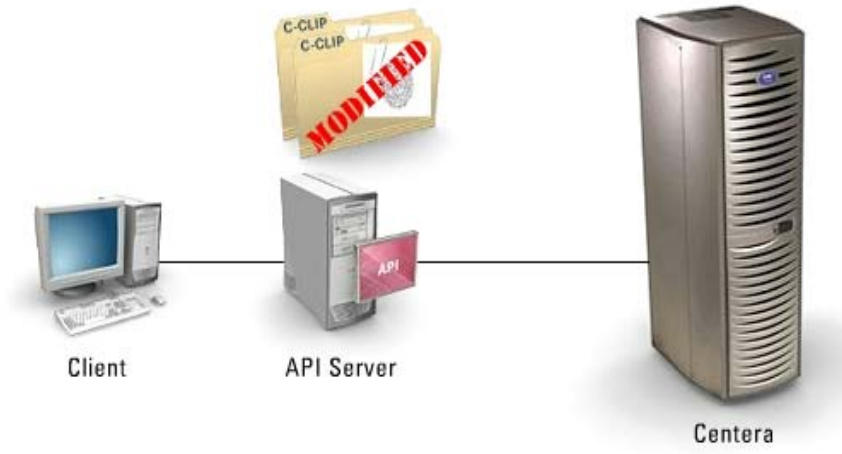
Rather than accessing a data object by its file name at a physical location, a CAS device uses a handle that is derived from each object's unique binary representation to store and retrieve the object. This is accomplished using breakthrough C-Clip technology, where subsequent access of the data object is made by simply giving the handle that uniquely identifies the object back to the repository. The data object is then returned. Content addressing greatly simplifies the storage resource management tasks, especially when handling hundreds of terabytes of static objects.

Also, this content-derived address is unique to ensure that only one protected (mirror or RAID) copy of the content is stored (single instance storage), no matter how many times applications store the same information. This significantly reduces the total number of copies of information stored, and is a key factor in lowering the cost of storing and managing content.

EMC Global Education

EMC² where information lives

Content Authenticity



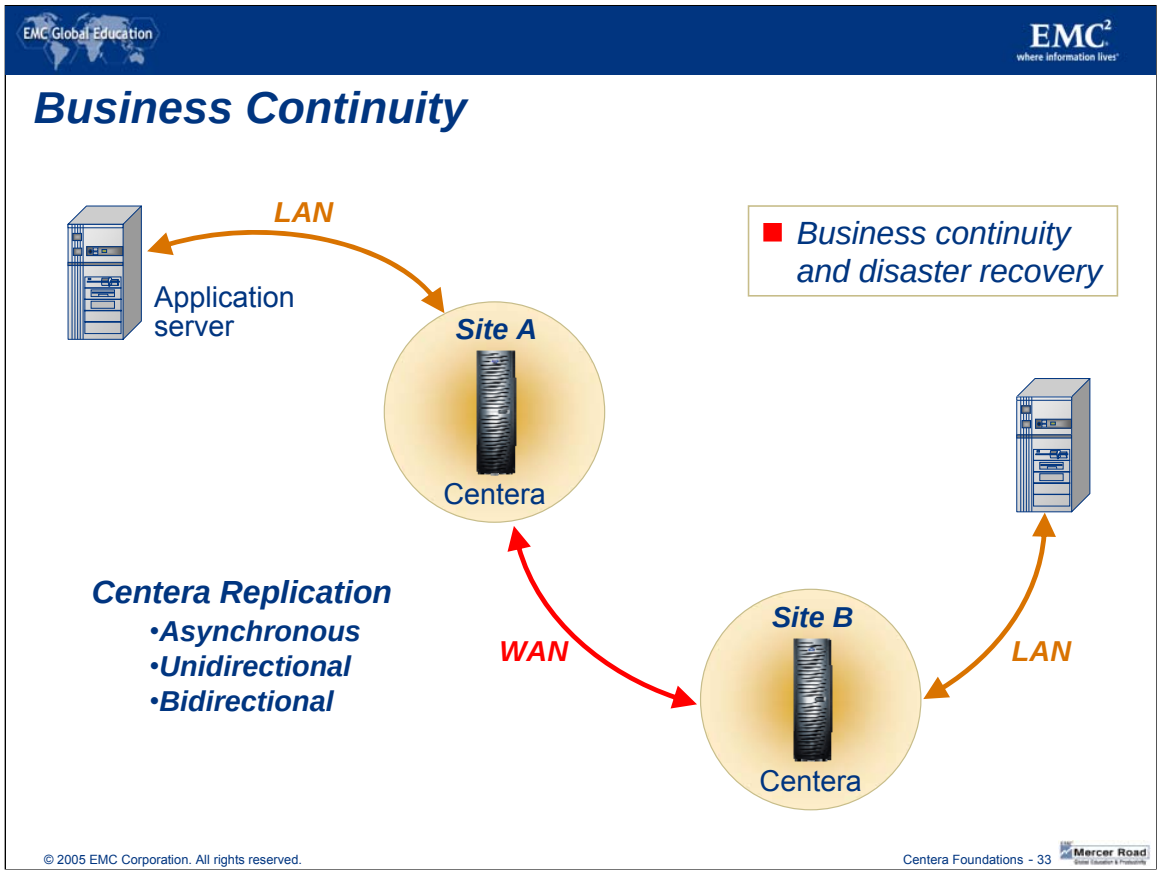
If a single bit in the content has been modified, a new unique content address is calculated. The modified content is then stored as a new object in the Centera.

The Content Address is a digital fingerprint for the content and it never changes.

Because a Content Address is globally unique, it ensures data mobility. Content can move as needed without concern on the part of the user. Applications present Centera with a Content Address and get the specific content in return.

In addition, the Content Address assures content authenticity. A change to content generates a new copy of that content with a new, unique Content Address.

Identical objects are only stored once, dramatically improving storage efficiency by eliminating redundant copies of content. An object's Content Address is derived from the content itself. This results in no more than one instance of identical content stored in a cluster.



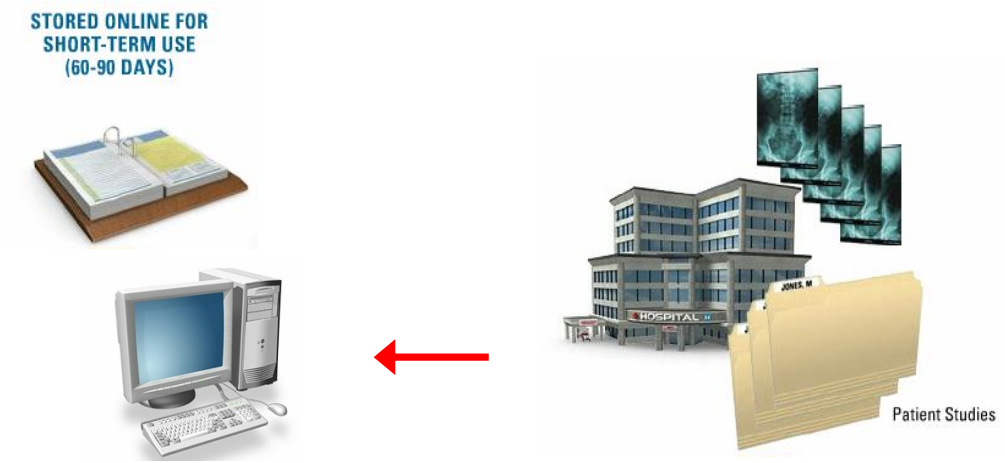
Centera Remote Replication replicates content from a local repository to a remote repository. When an object is initially stored in the local Centera, the object will be asynchronously and automatically replicated to the remote site over a wide area network (WAN), resulting in content being stored both locally and remotely. Centera replication currently can be implemented in 2 different ways for disaster recovery:

- Unidirectional
- Bidirectional

Unidirectional replication copies the data from the source Centera to the target Centera and is commonly used for disaster recovery and read only.

Bidirectional replication copies data in two directions, the local Centera to the remote Centera and from the remote Centera to the local Centera.

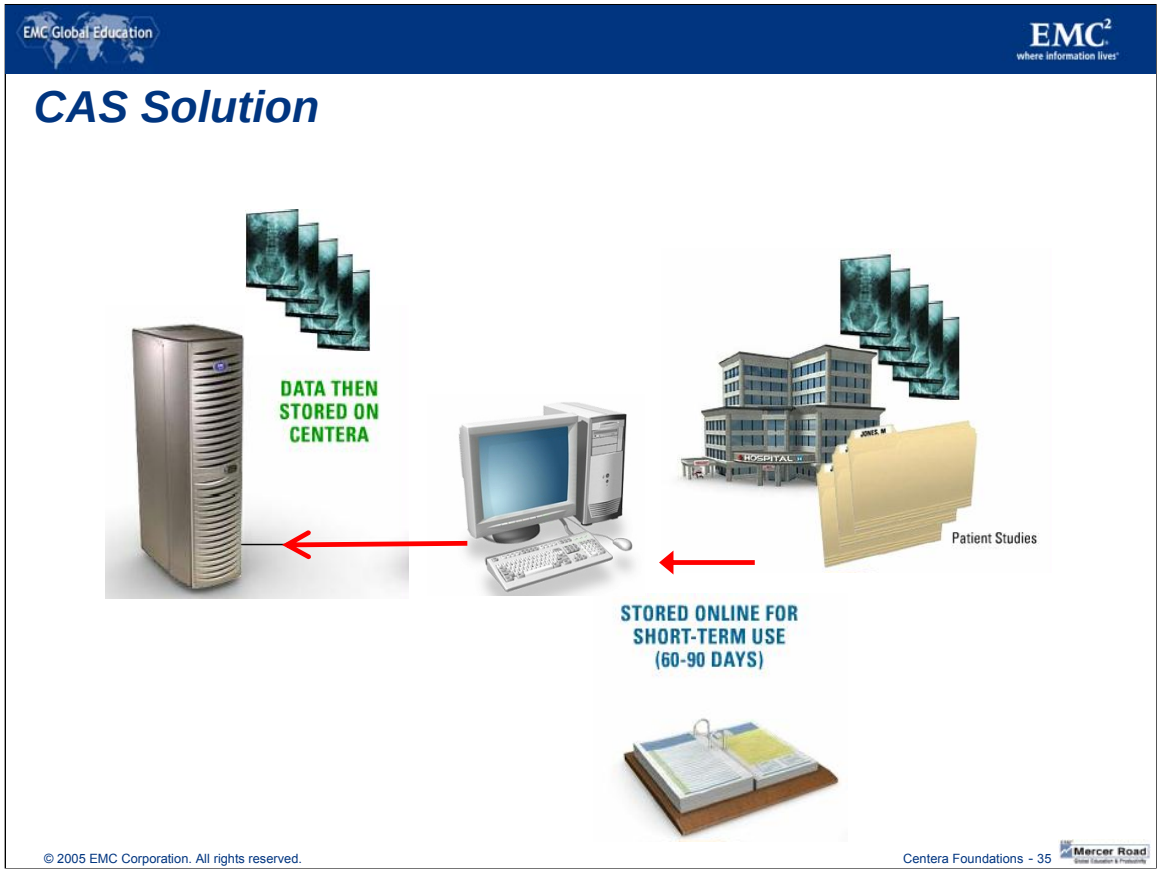
Healthcare Field Example



Over 400 million patient studies were completed in the U.S. in 2000. Each study was composed of one or a series of images, which range in size from about 15 MB for standard digital X-rays to over 1GB for oncology studies. As X-rays are created in the radiology department or hospital, they are stored online for immediate use by attending physicians for a period of 60-90 days.

At the point where the patients are cured or discharged, the access needs for their particular X-ray drop off dramatically. However, HIPAA* requirements stipulate that these studies must be kept in their full, glossy image formats for a minimum of 7 years.

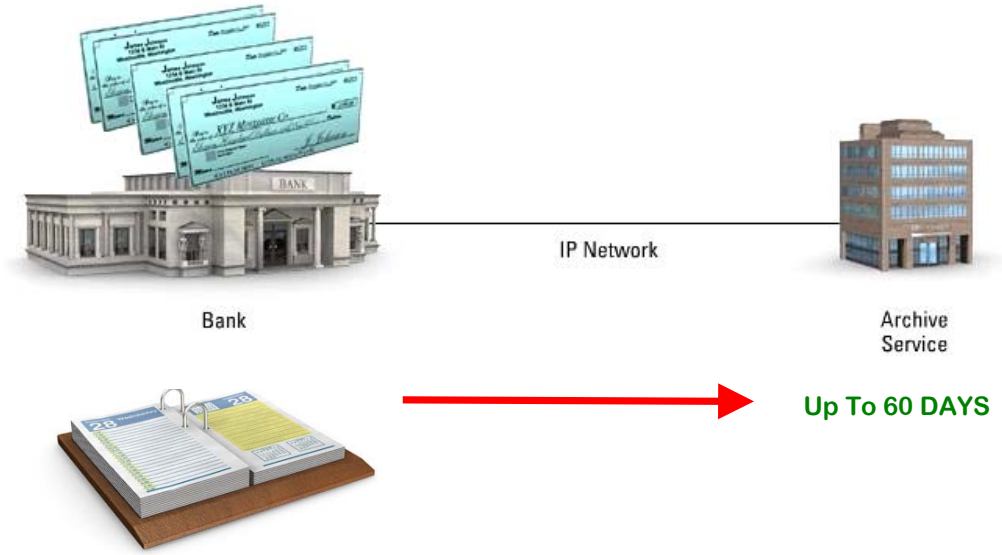
*HIPAA: Health Insurance Portability and Accountability Act of 1996



Beyond 90 days, hospitals may back up images to tape, or send them to an offsite archive service for long-term retention. The cost of restoring or retrieving an image when in long-term storage could be 5-10 times more expensive than leaving the image online. Long-term storage can also involve extended recovery times of hours or days.

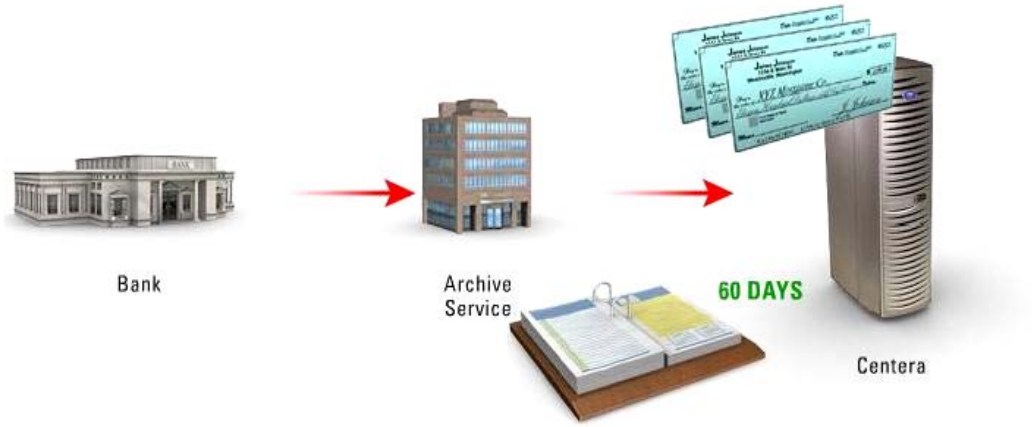
Medical image solution providers offer hospitals the capability of viewing medical studies, such as X-rays online, with sufficient response times and resolution to allow rapid assessment of patient situations. Centera is the optimum target storage device to facilitate long-term storage and immediate access of medical images online within a hospital or clinician's office.

Financial Field Example



Check images, each with a capacity of about 25 KB, are created at the bank and sent to archive services over a standard IP network. A check imaging service provider may process 50-90 million check images a month. Typically, check images are actively processed in transactional systems for about 5 days.

CAS Solution



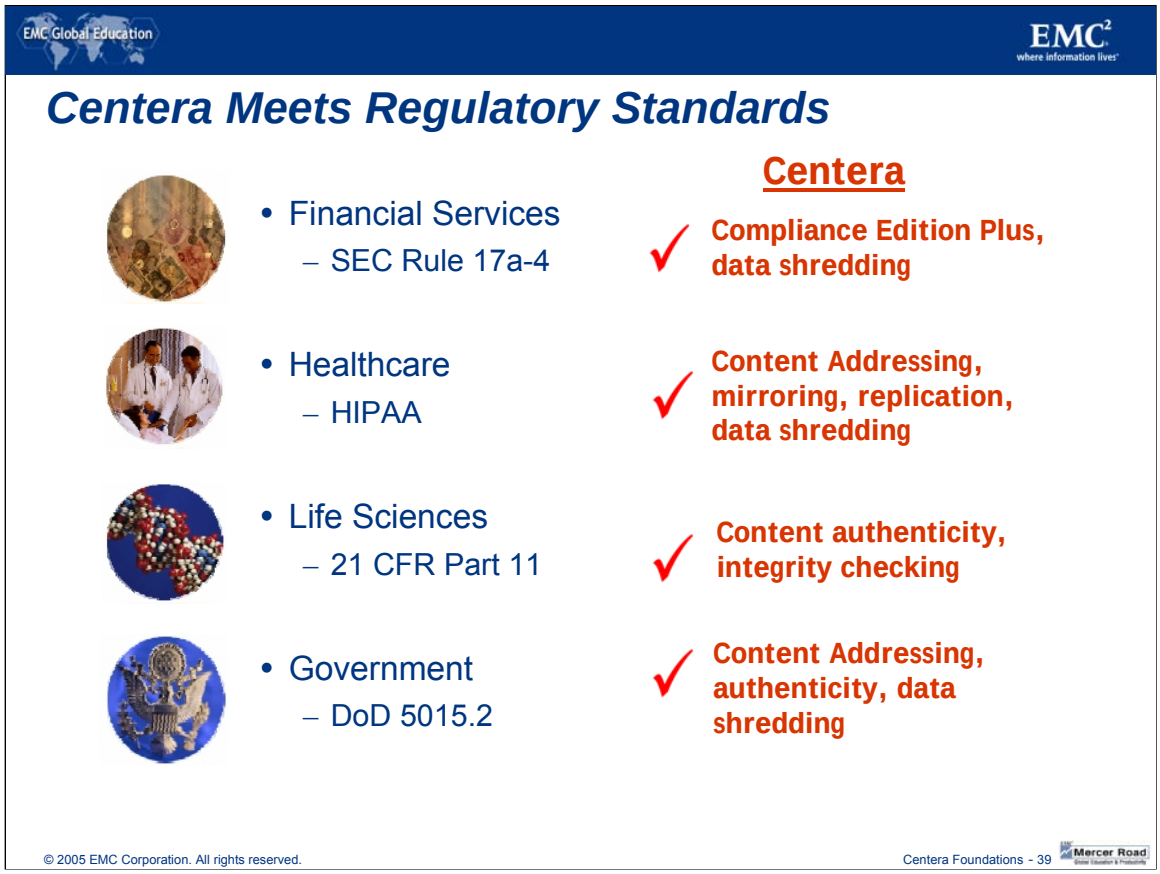
For the next 60 days, check images may be requested by regional banks or individual consumers for verification purposes, at a rate of about ½ percent of the total check pool (250,000-450,000 requests for check look-ups). Beyond 60 days, access requirements drop dramatically to as few as 1 for every 10,000 checks. In this case, the check images would be stored on Centera starting at day 60 and held there indefinitely. A typical check image archive can approach 100 TB.

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

Application/Technology Independent

CAS

Centera is easily installed and offers non-disruptive serviceability. The content addressed storage (CAS) technology allows for unique data representation and content authentication. In some applications, multiple customers concurrently accessing content may cause bottlenecks and access delays when using conventional solutions. Although these solutions may appear less expensive due to lower initial acquisition costs, in the long run, they cost more as they require increased focus on manual content management, such as movement to tapes and conversions to new formats. Centera alleviates the need to manage vast amounts of separated networked components and multiple file systems, caused by using low-end NAS or SAN alternatives to tape. Finally, Centera can be configured to maintain a replica of the fixed content at a remote site, eliminating the possibility of a site disaster destroying all copies of information.



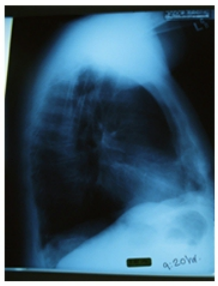
In all the regulated areas, the key elements of compliance are people, processes, and technologies. The mix and technology implications vary by regulation, and at our last count there were over 4,000 regulations across industries dealing with records authentication and retention. Here are only four of the many regulations that Centera addresses:

- SEC Rule 17a-4: The storage media requirements written into SEC (Security Exchange Commission) regulation 17a-4, viewed by many as one of the most stringent, if not the most stringent, regulated environments for information authenticity, retention, and protection. In Centera, data shredding and Compliance Edition Plus (CE+) handles these requirements with retention periods and data deletion restrictions as well as content authenticity.
- HIPAA: Centera complements the physical and administrative safeguards required by HIPPA, enhancing the protection and security of electronic patient information with content addressing, mirroring, replication and lifecycle integrity
- 21 CFR Part 11: Centera's digital fingerprinting and integrity checking of content throughout its retention provide functionality to detect altered, or compromised records beyond the grasp of any application, and is superior to that of conventional storage technology. With online access, Centera allows accurate and ready retrieval of all regulated records as required within Part 11.
- DoD 5015.2: Centera delivers additional levels of protection and security with Centera's Content Addressing, time/date stamping, and lifecycle integrity checking. Its Data Shredding feature exceeds the requirements of 5015.2 by ensuring privacy and eliminating liability.

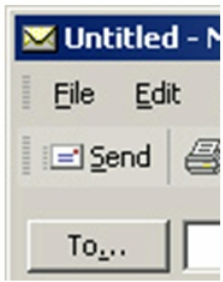
EMC Global Education

EMC² where information lives

A Few Integrated Applications and Partners



Medical Imaging Life Sciences
- GE Medical
- Fuji Medical
- AGFA/Mitra
- Scientific SW
- Amicas
- Mediface
- Merge



E-mail Archiving
- KVS
- Legato
- Educom
- IXOS
- Entelagent
- Assentor/Ilumin
- Tumbleweed
- NICE Systems

Enterprise Content Document Management
- Documentum
- FileNet
- Tower Tech
- daa Systemhaus
- System Ware
- Alliance
- Windream

Plus many more partners for the following areas:

Document/Check Imaging

HSM/FS Gateways

E-Learning

Oil and Gas

Audio Archiving

Backup/Archiving /Workflow

Media and Entertainment

Centera has hundreds of partners involved in its Centera Partner Program, taking advantage of Centera's free and open API. A small sampling of these partners is shown above.

Many of their integrated applications are already being used with the Centera by customers. Centera supports most OS platforms including Win2K/XP, Win2003, Linux, HP, AIX, IRIX, Solaris, and z/OS mainframe.

Centera Foundations

Addendum

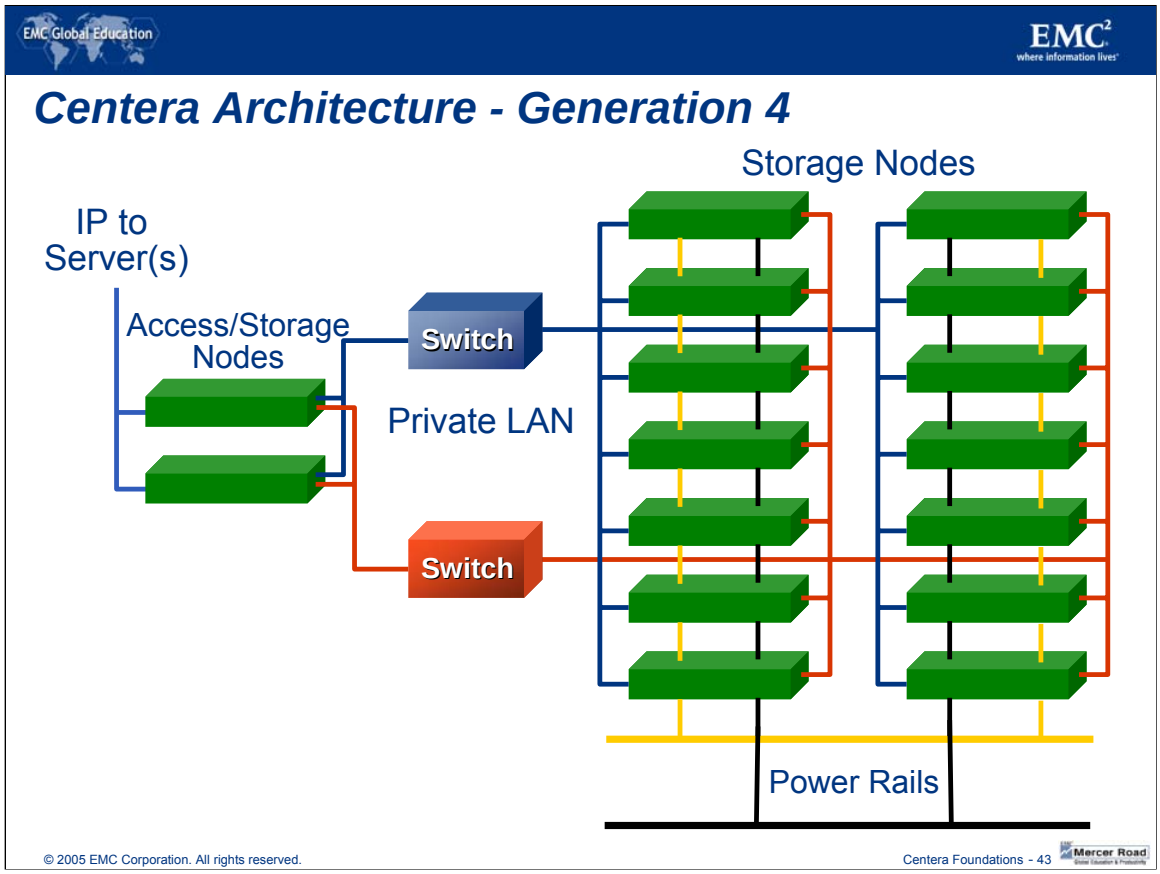
Welcome to Centera Foundations Addendum.

Module Objectives

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

- Identify changes made to Centera Hardware with Generation 4
- Describe the Virtual Pool feature and the new Replication Topologies

The objectives for this module are shown here. Please take a moment to read them.



As of May of 2005, Centera has introduced a new Generation of hardware and software. Several changes have been made to enhance performance and capabilities.

The new Generation 4 hardware (Gen4) includes changes to nodes, switches, power, and configurations. Each node now contains built in automatic transfer switches (ATS) to allow them to connect to two power sources. Each power outlet on a node will connect to a separate power supply. This has eliminated the need for the external ATS.

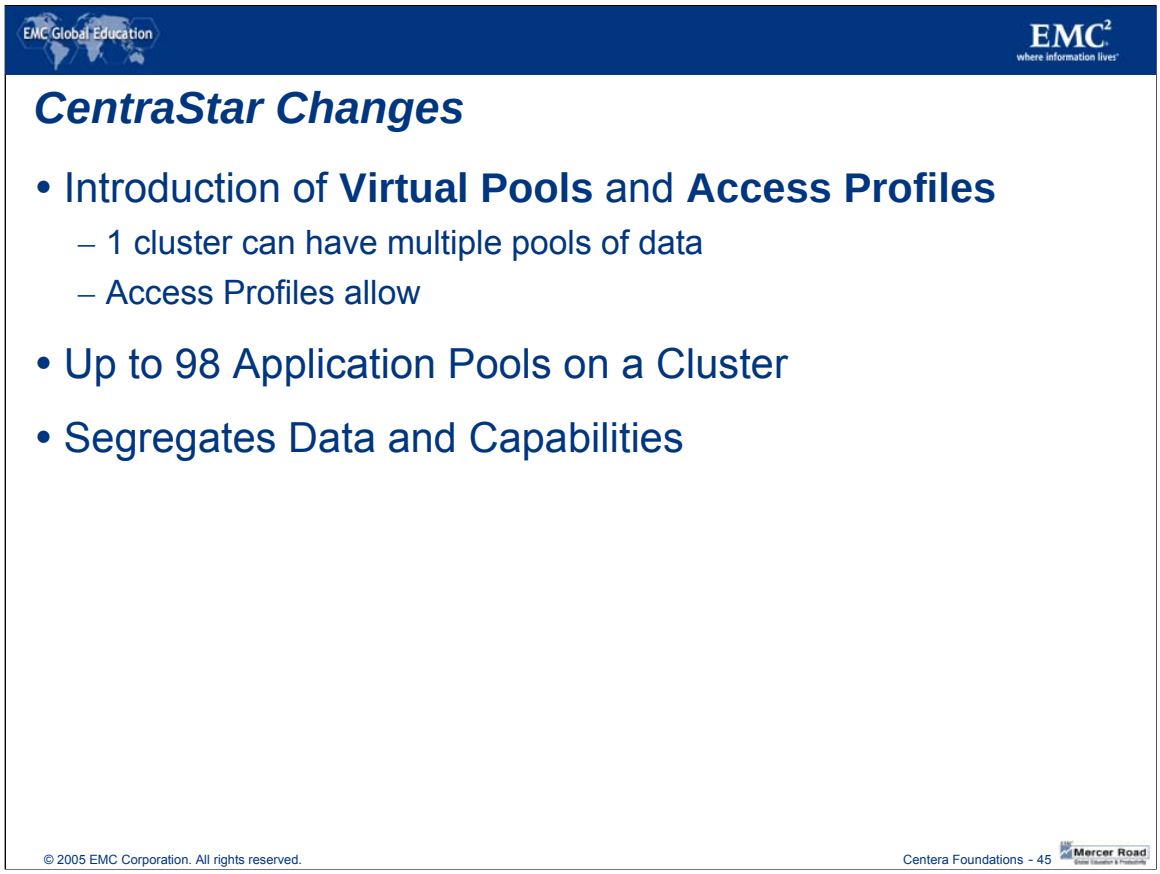
Each storage node still contains more than 1TB of usable capacity, but now both access and storage roles can reside on the same node. Also the nodes now have built in optical access capabilities where, before, a separate adapter was required.

The internal switches are now 24 port 2GB switches that provide communications for up to 16 nodes within the private LAN.

Centera Generation 4 Configuration

- New 4 Node entry system

While the Centera cabinet may still contain 32 nodes, the minimum configuration can now contain as few as 4 Gen4 nodes and 2 switches, and is expandable in increments of 4 nodes. A cube has also been redefined from a maximum of 32 nodes to 16 nodes. So a fully configured cabinet will be comprised of 2 cubes. The Gen4 hardware is also available as an unbundled solution. It can now be installed in an approved non-EMC cabinet.



As of May, 2005, CentraStar software, the application software that runs on the Centera nodes, now includes a new feature called Virtual Pools. Prior to this release, all the data was stored in one pool on the Centera that consisted of the entire cluster. All C-Clips were contained within this pool and any API client (with appropriate permissions) could access any clip, so there was no segregation of data.

With the introduction of Virtual Pools, a Cluster can now have multiple pools. Clips can be written to a specific pool where access can be granted to some applications but not others. Access to the Virtual Pool is granted to the Access Profile, which is used by the applications to gain access to the clips contained in a particular pool. This provides better security, flexibility and control of the data on the Centera.

Replication Topology

- Previous supported Replication Topologies
  - Uni-Directional
  - Bi-Directional
- Newly supported Replication Topologies
  - Star
  - Chain
- Replication of pools
  - Replicating individual pools instead of entire clusters

Another new addition to the Centera is the number of Replication Topologies supported. Previously, replication was only supported in pairs: two Centera clusters replicating either Uni-Directional or Bi-Directional. As of May 2005, Star and Chain topologies are supported as well.

Star topology consists of a Uni-Directional replication of three Satellite (Edge) clusters to one Central (Core) cluster (3:1). The Chain topology consists of three clusters in Uni-directional replication from one cluster to the next and then to the next (A→B→C). Replication is also now a pool level function, rather than a cluster function. Individual pools can be replicated to different targets (as long as they follow the supported topologies). This provides the System Administrator with the ability to selectively replicate individual pools within a cluster instead of replicating an entire cluster, or they can choose to replicate all the pools.

Module Summary

Key points covered in this module:

- Generation 4 hardware changes include nodes, switches, power, and configurations
  - A new 4 node configuration is allowed for entry system and expansion
- Virtual Pools along with Access Profiles allow for the segregation of Clips, providing better security and Replication of a subset of Clips
- Star and Chain Replication Topologies are now supported

These are the key points covered in this addendum. Please take a moment to review them.

EMC Global Education

EMC² where information lives

Course Summary

- Content Addressed Storage (CAS) is a new category of online disk storage designed specifically for fixed content, data that is in its final form
- Centera employs a unique storage/retrieval method called “content addressing” which is performed through the use of C-Clip technology
- Centera offers a number of benefits over that of other long-term backup solution such as:
  - Faster record retrieval
  - Single instance storage
  - Guaranteed content authenticity
  - Self-Healing
  - Meets regulatory standards
- Centera Addendum

Mercer Road

Key points covered in this course are shown here. Please take a moment to review them.

This concludes the training. In order to receive credit for this course, please proceed to the Course Completion slide to update your transcript and access the Assessment.

Mercer Road Global Education & Productivity