Archiving Calls to EMC Centera: Safe and Secure Storage for Compliance Purposes

Applied Technology

Abstract

This white paper is an overview of the integration between CyberTech Recording Solutions and EMC Centera® for the audio recording market segment.

July 2010
# Table of Contents

**Executive summary** .................................................................................................................. 4  
**Introduction** ................................................................................................................................. 4  
  Audience ........................................................................................................................................ 4  
**CyberTech reporting** ..................................................................................................................... 4  
  Recording Solutions ......................................................................................................................... 4  
  Short-term storage ........................................................................................................................... 5  
  Long-term storage .......................................................................................................................... 5  
    NAS/SAN ...................................................................................................................................... 5  
    EMC ......................................................................................................................................... 5  
**EMC storage** ................................................................................................................................ 6  
  EMC Centera ................................................................................................................................. 6  
  Archiving made simple .................................................................................................................... 6  
  The big picture .............................................................................................................................. 6  
  Setting up archiving to EMC Centera .............................................................................................. 6  
**Storage for compliance** ............................................................................................................... 7  
  Content flow .................................................................................................................................... 7  
**Retention management** .............................................................................................................. 8  
**Litigation hold** ............................................................................................................................. 9  
**Chargeback** .................................................................................................................................. 10  
**Implementation** .......................................................................................................................... 11  
  Data model ..................................................................................................................................... 12  
  EMC storage capacity ...................................................................................................................... 12  
  EMC Centera SDK ........................................................................................................................... 12  
**Conclusion** .................................................................................................................................... 13
Executive summary
Voice recording systems can produce terabytes of data that often need to be efficiently managed and stored for periods of up to 10 years. CyberTech Recording Solutions already offers a wide range of alternatives for archiving including DVD-RAM, storage area network (SAN), network-attached storage (NAS), and others. To meet the growing demand from customers, CyberTech has integrated with EMC Centera as an intelligent, purpose-built, disk-based storage archive.

EMC Centera is often used for centralized storage of recorded call archives. Centralized storage eliminates the manual processes associated with searching, loading and unloading, and management of individual media types (such as DVD, CD, external hard drive, digital tape), which in turn improves security and reduces staffing and operational costs.

The increasing number of compliance regulations within the financial industry requires secure and safe storage possibilities. CyberTech offers these possibilities with archiving to EMC Centera. Every recording stored on an EMC Centera system is allocated a unique MD5 fingerprint that guarantees integrity.

Extensive retention management is required for recorded calls. CyberTech offers full EMC Centera integration including complete retention management and litigation hold features. With "litigation hold" on EMC Centera, retention policies can be set for those calls that may be relevant to a lawsuit that has been actually filed, or even one that is "reasonably anticipated."

Many companies and departments require allocation of storage costs. CyberTech offers the EMC Chargeback application that generates reports to provide good insight into the amount of storage occupied by this archive. Chargeback reports can be generated for archived calls that are linked to specific users, user groups, channels, or channel groups.

This integration between CyberTech Recording Solutions and EMC Centera enables call recordings to be safely and securely stored in a clustered and resilient storage system. It unlocks the value in digital content, lowers costs, and makes it easier to manage and share information.

Introduction
This white paper begins by describing the long- and short-term options that CyberTech offers. It then discusses EMC Centera storage and topics such as storage for compliance, retention management, litigation hold, chargeback, and implementation.

Audience
This white paper is intended for an EMC and customer audience who wish to learn details on how CyberTech uses EMC Centera.

CyberTech reporting
Recording Solutions
CyberTech Recording Solutions is one of the first open and secure recording solutions, designed to use industry-standard hardware and software. This best-in-breed technology is being used by the world's leading financial institutions, governmental and public safety organizations, and call centers. By taking advantage of commercial off-the-shelf (COTS) hardware and customer-provided network storage devices, this future-proof solution for the capture, storage, retrieval, and playback of voice, radio, and data communications provides unsurpassed functionality while reducing a firm's total cost of ownership. This flexible, multi-media product delivers high-quality recordings of traditional or VoIP telephony, together with data and desktop screens for all applications, including verification and compliance, dispute resolution, training, and quality monitoring.
**Short-term storage**

Within CyberTech Recording Solutions all voice and call data is stored as WAV files on the system’s hard disk and as records in a central database. This MySQL database contains one record for every call. If the “encryption” option is enabled, all stored calls are encrypted with 256-bit Rijndael AES audio encryption. This means the WAV file can be replayed only if the encryption key is available. The number of hours of recording storage depends on the size of the hard disk and compression used. Table 1 is an example of storage capacities for three hard disk sizes.

**Table 1. Storage capacities for 80, 180, and 250 GB HDD**

<table>
<thead>
<tr>
<th>Compression type</th>
<th>80 GB HDD</th>
<th>180 GB HDD</th>
<th>250 GB HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncompressed 64 Kb/s</td>
<td>2.625</td>
<td>5.906</td>
<td>8.203</td>
</tr>
<tr>
<td>ADPCM 32 Kb/s</td>
<td>5.250</td>
<td>11.813</td>
<td>16.406</td>
</tr>
<tr>
<td>ADPCM 24 Kb/s</td>
<td>7.000</td>
<td>15.750</td>
<td>21.875</td>
</tr>
<tr>
<td>GSM 13 Kb/s</td>
<td>12.727</td>
<td>28.636</td>
<td>39.773</td>
</tr>
<tr>
<td>True Speech 8.5 Kb/s</td>
<td>19.765</td>
<td>44.471</td>
<td>61.765</td>
</tr>
<tr>
<td>Speex 8 Kb/s</td>
<td>21.000</td>
<td>47.250</td>
<td>65.625</td>
</tr>
<tr>
<td>Speex 5.95 Kb/s</td>
<td>28.235</td>
<td>63.529</td>
<td>88.235</td>
</tr>
<tr>
<td>Speex 3.95 Kb/s</td>
<td>42.532</td>
<td>95.696</td>
<td>132.911</td>
</tr>
<tr>
<td>Speex 2.15 Kb/s</td>
<td>78.140</td>
<td>175.814</td>
<td>244.186</td>
</tr>
</tbody>
</table>

The short-term storage to hard disk works with a selection of First In, First Out (FIFO) or Last In, First Out (LIFO) as a Disk Full error handling mode. To keep calls longer, a long-term archiving medium is required.

**Long-term storage**

For long-term storage of calls and data, the CyberTech Recording Solution archives these calls and data to offline storage media. Several types of storage media are supported, and are described next.

**NAS/SAN**

NAS or SAN hard disk drives are the most commonly used archive media as they are cheap and widely available.

**EMC**

CyberTech has added EMC storage integration to the existing archiving interfaces. This integration is approved and certified by EMC.

The CyberTech Remote Archive Service (RAS) is a Windows service that is by default not running. Once a license is loaded, the service can be activated and the user interface automatically enables all EMC Centera-specific features and functionality. Only one license per recording system is required. The RAS module will control archiving to EMC Centera.

![Figure 1. Data flow](image-url)
EMC storage

**EMC Centera**

EMC Centera is the world’s most simple, affordable, and secure repository for the archiving of information. This purpose-built, software-driven storage platform provides a myriad of new capabilities for unlocking the business value from unchanging or infrequently changing digital assets such as voice, radio, and data communications.

**Archiving made simple**

EMC Centera was specifically designed to store and provide fast, easy access to archived information. It is the first solution to offer online availability with long-term retention and assured integrity for this fastest-growing category of information. Whether integrated with an in-house-developed application or an application from a continually growing group of industry-focused EMC partners, EMC Centera is the optimal information archive for businesses and organizations that require a simplified solution to expanding amounts of information. EMC Centera greatly simplifies the task of managing, sharing, and protecting all sizes of content repositories. It cost-effectively puts information online in support of new sources of revenue generation, expanded business models, and increased service levels to users and customers. EMC Centera features the write once, read many (WORM) attributes of non-rewriteability and non-eraseability with disk performance and TCO superior to current archiving approaches.

EMC Centera was designed from the outset to provide simple, affordable, and secure access to archived information. The EMC Centera CentraStar® software operating environment employs an innovative content-addressing system to simplify management, ensure content uniqueness, and deliver the scalability needed for terabyte- to petabyte-level archive requirements. And the EMC Centera system accomplishes this while dramatically lowering overall management costs.

**The big picture**

EMC Centera has the following:

- Purpose-built magnetic-disk-based storage that overcomes the limitations of tape, optical, or traditional disk technologies
- The ability to address governance and compliance needs for information authenticity, retention, and preservation
- Self-healing and self-managing capabilities that allow you to easily manage 50-100 times more information than with other archive storage
- A tamper- and future-proof architecture that is designed to store millions and billions of pieces of information that will outlive the platform on which it resides

Please visit [www.emc.com/EMC Centera](http://www.emc.com/EMC Centera) for more information

**Setting up archiving to EMC Centera**

To use EMC Centera as an archiving medium within the Recording Solution, three steps are required:

1. Create a Marking and Archiving Schedule: Which channels need to be archived
2. Create Archiving Rules: To which media the call should be archived.
3. Select the EMC Centera as archiving media, as shown in Figure 2.
Storage for compliance

With the increasing number of compliance regulations, safe and secure storage becomes more and more essential for voice recording. Therefore CyberTech Recording Solutions offers standard options to comply with these regulations:

- Access to the voice recorder browser application can be protected by enhanced security features such as unique user IDs, alphanumeric passwords, domain authentication, and an account lockout mechanism, and complies with most password policies.
- Replay browser application passwords are all stored in an encrypted format.
- User account overviews can be downloaded for review.
- The level of access to the browser application and access to calls is controlled by the permissions given to each individual account.
- The system includes two audit trails that allow individual call and user access to be monitored, as well as more detailed reporting.
- All calls and data on the recorder server and archive media are encrypted using 256-bit Rijndael AES audio encryption and standard MD5 fingerprinting.
- Remote, restricted administration by CyberTech engineers is based on encrypted communication.
- Data and calls stored on the recorder can be deleted after a preset time interval.
- Standard integration with most IP and digital handsets allows either start/stop or the muting of parts of a call on handset function key depression.
- Customized development using CyberTech APIs may also provide for automated start/stop recording or muting of the conversation when completing certain fields within an application (subject to the appropriate telephony environment).
- The (optional) integration with EMC storage adheres to compliance regulations. Every recording stored on an EMC Centera system is allocated a unique MD5 fingerprint, which guarantees integrity.

Content flow

When archiving recording files from the CyberTech Recording Solution to an EMC Centera, the content flow is as follows:

1. On the recording system, the audio and call data are stored together in one file.
2. On the recording system, the “archive service” adds additional data to the file, such as retention dates, and sends it to the EMC Centera.
3. The file and retention dates are stored on the EMC Centera.
4. A unique MD5 fingerprint is returned from the EMC Centera and stored in the database of the recording system.
5. Using the fingerprint and retention dates, the recording system is able to find and play back calls that are stored on the recording system.

Figure 3. Content flow for archiving recording files

Retention management
As part of compliance regulation, retention management of recorded calls is very important. The CyberTech archive management process is fully automated with preset retention periods for each department or individual. CyberTech Recording Solutions maintains an EMC retention database that uses individual or group profiles than can be individually defined to meet specific needs. A “timed delete protection” feature ensures that recordings cannot be tampered with before the end of the predetermined retention period, at which point they are automatically deleted.

The CyberTech Recording Solution uses the EMC retention mechanism when it creates a call archive. Within this call archive, each call will be inserted with a retention period. When the retention period is expired, the CyberTech Recording Solution will automatically remove the call from the EMC Cluster and from the local hard disk (database record and call audio).

The CyberTech EMC retention periods are in most cases based on user profiles or user group profiles. Per user group a call retention period for users within that group is set. If needed, this retention period can be overruled for specific users within the group by defining a retention period in the user profile of these specific users. For systems where not all calls are linked to users or user groups, retention can be set per channel group. This allows defining a retention period per, for example, a trunk number. Channel group retention will be used only if no user or group is assigned to the channel.

The retention period will be linked to the call based on the following priority:

1. User profile (overruling user group retention)
2. User group profile
3. Channel group configuration

If a call needs to be archived not containing any of the above settings (no retention value specified) the default retention value set in the archive rule is used.

When creating a new user or editing an existing one, a retention period can be configured for the calls made by that user. The user administration page is located under User administration > User. In the section Details / properties for user account, a value can be configured for the retention period to be used for the
calls of this user. The retention values range is 1 to 6,000 days. The range is automatically adjusted to the minimum and maximum retention settings allowed for the EMC pool (defined under the EMC “location”).

![Image of configuring a retention period](image)

**Figure 4. Configuring a retention period**

When adding a new user, the default retention value will be **Inherit from group**. Once the user is stored, the value of this inheritance will be shown when the **Details / properties for user account** section is re-opened. When a user has their own retention value (the value is not inherited from their group), it will overrule the groups retention value.

Editing an already configured retention can result in two scenarios:

- If retention is increased, and no restrictions exist, historically archived calls will get the new retention value.
- If retention is decreased, only the calls to be archived will get this new retention value. Historically archived calls will not get this new retention value.

The CyberTech Recording Solution keeps track of the retention periods assigned to the calls that have been archived to the EMC Centera. The moment a retention period is exceeded, the CyberTech Recording Solution automatically deletes the corresponding call. This call is deleted from the EMC Centera as well as from the Recording Solution; both the database record and the call audio are deleted as well. This means that the local (that is, CyberTech core server) management of call audio and data is regulated by the configured retention periods (user profile, user group profile, and channel group configuration).

**Litigation hold**

A "litigation hold" is a suspension of EMC’s data (recorded call) retention policies for data that may be relevant to a lawsuit that has been actually filed, or even a suit that is "reasonably anticipated." By putting calls on hold, the system will temporarily prevent the deletion of the calls. While on hold, these calls cannot be removed from EMC Centera even if retention is expired. The CyberTech Recording Solution allows up to five “holds” per archived call.

The availability of the litigation hold functionality in the CyberTech Recording Solution depends on the following items:

- The EMC Centera software version must be version 3.1 or later (older versions don’t support the litigation hold functionality).
- The litigation hold functionality must be activated on the EMC Centera.
- The Recording Solution user must have the litigation hold right, which is defined in the user or user group management.

The litigation hold functionality can be configured in the user management section and user group management section. The functionality is split up in two rights:

- Place calls on hold
- Remove hold from calls

Per item, the set coverage determines for which calls the hold can be set or removed.
Figure 5. Call permissions

A call is put on hold using HOLD_ID. The call could be unlocked only using the same HOLD_ID as it was locked with.

Figure 6. Call details

Chargeback

The CyberTech chargeback application can generate reports that provide good insight into the amount of storage occupied by this archive. The application can create different reports. A report can be generated for archived calls that are:

- Linked to specific users
- Linked to specific user groups
- Recorded on specific channels
- Recorded on specific channel groups

The reports can all be combined within one single report. With this flexibility it is easy to create, for example, a report per organization or cost center. The report contains the used EMC Centera storage space of all the calls within the given criteria.

The generated report contains the data in a comma separated format. This makes it ideal for analyzing tools or to create graphs.
Implementation

The CyberTech Recording Solution creates different C-Clip™ layouts for the EMC storage, depending on the size of the file. These are shown in the following three figures.

Figure 7. Example of an imported report in Excel

Figure 8. For archived calls with a file size less than 100 KB

Figure 9. For archived calls with a file size greater than 100 KB
Figure 10. For the archived database

The retention of the database backup is set to 0 and is kept on the EMC cluster as a linked list, with a maximum of the five most recent database backups. The buffer size is set to 100 KB.

**Data model**

The type of content is recorded telecom conversations, including all metadata. The average content size is 300 KB. As a CDF Structure a simple one tag and one BLOB are used. The CyberTech Recording Solution supports EMC Centera as the archive.

The CyberTech Recording Solution supports content addresses of 65 bytes in length and stores all content addresses in their own database. The content address is not made visible to the end user of the solution. The user interface gives access to the stored content and shows the retention period.

The CyberTech Recording Solution makes use of multiple threads to write data to the EMC Centera. The default number of threads is 10, which is based on extensive testing and field experience. This seems to be a good balance between system load and archive speed. For fine-tuning, the number of threads is still configurable to accomplish the best balance.

**EMC storage capacity**

Storage capacity of the pool on the EMC Centera system is monitored by the Remote Archive Service of the CyberTech Recording Solution. Depending on the capacity status of the pool, the Remote Archive Service will generate the following warnings (if applicable):

- Error code 5005 – Warning: Archive location <location> is 80% full.
- Error code 5001 – Error: Archive location <Location> full.

Error code 5005 will be raised if the remaining available storage capacity of the EMC pool is 20% of the complete pool capacity; archiving to EMC will continue. This indicates that the amount of calls archived to the EMC pool exceeds the pool capacity. To meet the needs of the user, the only possible action to take on this warning is to expand the storage capacity of the pool by reconfiguring it (which might require extra EMC Centera storage hardware).

Error code 5001 will be raised if there is no more available storage capacity in the EMC pool. This indicates that the amount of calls archived to the EMC pool exceeds the pool capacity. To meet the needs of the user, the only possible action to take on this warning is to expand the storage capacity of the pool by reconfiguring it (which might require extra EMC Centera storage hardware).

**EMC Centera SDK**

The EMC Centera SDK used within the CyberTech Recording Solution is version 3.1, which is compatible with CentraStar 4.0. All EMC Centera SDK revisions are forward-compatible with future revisions of CentraStar.
Conclusion

Together CyberTech Recording Solutions and EMC Centera can meet the most demanding and challenging environments. The solution provides a framework for increasing operational efficiency and reducing storage requirements, while optimizing storage utilization and achieving the highest levels of performance. Given the explosive rate of data growth and longer retention requirements, the CyberTech and EMC Centera solution helps mitigate risk associated with the ever-increasing regulatory and corporate governance requirements.

About CyberTech

CyberTech International leads the voice logging and communications recording industry and is a recognized innovator of voice recording and quality monitoring applications. With products and services that drive efficiencies in the capture, storage, retrieval, and playback of voice, radio, and data communications, CyberTech offers a worldwide suite of innovative recording technologies that enable organizations to achieve the highest levels of performance, quality assurance, compliance, and liability protection while supporting existing business processes.

With solutions deployed widely throughout organizations in the financial services, public safety, and call center markets, CyberTech uniquely adds value through COTS products and applications that are scalable, turnkey, and easy to implement over a virtually unlimited number of channels. An unsurpassed commitment to customer satisfaction additionally assures maximized technology investments and quality of customer contacts without proprietary hardware, complex integrations, or expensive software upgrades.

CyberTech has global operations and offices throughout Europe, Asia, Middle East, and the Americas. In 2008, more than 100,000 recording channels were sold, and the company enjoyed a 175-percent increase in revenue for 2005-2008. Today, CyberTech is the third-largest global manufacturer of voice and data recording solutions and has its 25-year celebration in 2010. For more information, visit www.cybertech-int.com.