



# Repositories

## What are Repositories?

Repositories (AKA Vaults or Archives) are the self contained logical units, consisting of data, index and metadata that are created, populated and processed by Archive One.

Repositories can either be 'direct' (located on a local disk or on a network share, and the subject of this paper) or be controlled by a 'Storage Controller Module' (a pluggable interface layer to the backend storage device/media).

## Direct Repositories

### *Multiple Repository model*

Archive One provides support for Multiple Repositories to form the total archive store. Multiple Repositories provide greater flexibility to assist with ongoing operation, granular categorization by archive groups or data profile and the ability to simplify Disaster Recovery (DR) strategies. (It does not require the use of a database such as SQL Server therefore reducing costs and complexity – see separate document).

Each Repository is completely independent and self contained. Repositories can be located on any network share or direct attached storage, but are not fixed to that location. At any stage it is possible to relocate the Repository to a new location.

The process of relocation can be as simple as a native Windows file copy (Network security limits honored) and a few mouse clicks within the Archive One Administration console.

Multiple Repositories do not need to be located on the same drive or Storage Controller Module, they can be hosted where the access profile and drive IO performance best match.

### *A typical example of defining a Multiple Repository strategy*

Hosting email data based on an annual banding. E.g. if an Exchange server has 7 years of email data, it is possible to archive each year of the data to individual Repositories or to define Repositories for each business department within a time period, e.g. Finance 2007, Finance 2006, on so on.

Key advantages to this strategy include:

- Significant reduction on Exchange backup / restore requirement
- Elegant Repository backup strategy (it is only necessary to backup the latest Repository being complimented with data)
- Storage location profiling
- Consistent performance
- Maintenance free
- Data replication traffic to DR sites is kept to the absolute minimum
- Portable (each Repository can be moved without system operational impact)

## ***Reducing Exchange Backup / Restore Requirement***

One of the major benefits of email archiving is being able to reduce the size of the Information Stores of Exchange servers. Using Archive One makes it possible to archive data by many different criteria and policies.

The key to gaining maximum benefit from Archive One is to archive in a manner that rapidly reduces the size of the stores while providing an ongoing simplified backup strategy.

This is achieved through extracting the relevant subset of data to dedicated Repositories that are defined in a matching logical unit. For example: archive all Year 2000 emails to one Repository, Year 2001 to a second and continue this until in a position to archive data in the current year, where you may adopt a different archiving strategy, such as archive all emails greater than 30 days old.

This approach removes vast quantities of data from the Exchange stores, enabling the administrator to recoup the space in the stores and directly reduce the Exchange server backup commitment and restoration times.

### ***Elegant Repository backup strategy***

By providing a Multiple Repository technology within Archive One it is possible to take advantage of limited backup commitment and yet have full disaster recovery assurance.

By adopting a logical unit Repository model it is possible to enable a condition where changes to a Repository do not occur. For example, when all of the Year 2000 emails have been fully archived from Exchange to dedicated Repository. When this condition has been met the Repository structure becomes static or 'closed'.

As each Repository becomes closed it is no longer necessary to perform a daily backup of these Repositories, just maintain at least one disaster recovery copy.

Being able to reduce Exchange backup commitment significantly, plus a small daily archive requirement, enables a huge reduction in cost and complexity where disaster recovery is concerned. Furthermore, in a Disaster Recovery (DR) event, Repositories can be restored very quickly and in a sequence that provides access to the more recent data first and the older data later.

### ***Storage location***

Each Repository can be located on storage to best meet access requirements. Repositories that are being complemented with new archive data or will frequently serve archive requests should be positioned on higher performance drives to that of infrequently accessed Repositories.

This provides administrators the ability to position Repositories to best meet performance demands whilst keeping the cost of ownership to a minimum. It also allows for very simple relocation of Repositories / data onto storage with different access profiles as the Repository / data ages.

### ***Consistent Performance***

Unlike other archiving solutions, Archive One's closed Repository model does not suffer performance degradation over time. Other solutions may incur performance challenges due to the requirements to re-index, perform online maintenance or perform defragmentation.

It should be noted that as storage technology performance continues to improve and evolve it can be expected to see Repository performance enhance as Repositories are migrated to newer drive subsystems.

## ***Maintenance Free***

In line with providing consistent performance, administrators do not need to perform any regular maintenance to Repositories. This provides many advantages:

- Consistent performance
- Zero maintenance downtime for administrators
- No impact to user productivity
- No impact to operational costs

## ***Data replication traffic to DR sites is kept to the absolute minimum***

For many organizations it is essential to provide full disaster recovery site support. Archive One's Repositories provide administrators with the ability to provide full DR support without having to invest in complex solutions or technologies.

By adopting the Multiple Repository model in addition to the 'closed' Repository status, DR solutions are able to benefit by having to replicate the closed Repositories once. For the active Repository only the delta updates per archival run are required to replicate. This reduces the amount of traffic to be managed by the replication technology dramatically.

All premium replication technologies fully support Repositories without any special agents or complex configurations, indeed for a simple replication solution, applications such as Robocopy by Microsoft can be utilized free of charge.

## ***Portable***

An email archiving solution is a long term investment; the systems that host the solution initially are likely to be retired or experience some form of outage. An email archiving solution is another business critical application. As with email servers themselves, there is a business expectation of very high availability. Archive One's portable Repository design makes this possible.

To support planned or unplanned system downtime / upgrades each Repository is completely portable. Copy or restore the Repository to any network location that provides sufficient storage, reference this updated location within the Administrative console and the Repository content becomes available from its new location immediately.

It is possible to retire whole Repositories by deleting their definition within the Administrative Console and then removing from the storage. Any retired Repository can be re-established by simply defining the Repository values within the Administrative console and restoring the data to corresponding storage.

## **Summary**

Unlike other archiving systems Archive One couples the index with the archive data, creating one logical set of data - making backup and restore easy. Data re-acquisition is easy. Point and click at Repository and the data and index are restored, immediately.

Unlike other archiving solutions, each Repository is 'complete' and when a backup exists on media, it requires no further management, maintenance or to be a continued part of the daily DR regime.

This unique approach removes the significant cost of ownership and complexity to archived data whilst providing full functionality and scalability.

## Contact C2C:

C2C Systems Inc.  
1 Federal Street  
Bldg. 101-R 3W  
Springfield, MA 01105-1199  
T: 413-739-8575  
F: 413-739-4980

C2C Systems Ltd.  
6 Richfield Place, Richfield Ave  
Reading, Berkshire  
RG1 8EQ, UK  
T: +44 (0) 118 951 1211  
F: +44 (0) 118 951 1111