

# Archive360 Archive Migration

## Technical Guide

---

### Cloud-Based Approach

Your enterprise is moving to the cloud. Your on-premises legacy data is ready for migration. But that doesn't mean you must endure the time and expense of installing archive migration software on-premises. There is a better way: one that offers faster performance, less disruption and stronger security. Read on for a step-by-step walk through of how Archive360 combines patented, cloud-based technology with a fully managed service to extract and migrate legacy data from more than 20 archive systems.

Archive Migration	Cloud-based	On-premises-based
Reduces project start and finish timelines	✓	✗
Minimizes hardware and storage requirements	✓	✗
Eases network bandwidth needs	✓	✗
Eliminates the need for SQL server on-prem	✓	✗
Removes archive upgrade requirements	✓	✗

# Table of contents

## Technical Guide

---

1. Formulating a Strategy
2. Archive360 Migration Architecture
3. Step One: Accelerated Discovery
4. Step Two: High-Speed Transfer
5. Data Explosion, Bandwidth Optimization
6. Step Three: High-Speed Migration to Office 365
7. Step Four: Final Migration and Chain of Custody

## More than 50% of archive migration projects fail on the first try

Why? Even if the archive migration tool performs as advertised, problems with archive hardware, storage and databases, and network latency can derail a migration project. And because many different teams are needed to troubleshoot these issues, resolution can take a long time. This is why a cloud-based approach to archive migration makes sense.

This paper discusses Archive360's unique approach to extracting data safely from legacy systems, and common myths IT teams must de-bunk to ensure a smooth migration.

A legacy archive migration may include any combination of the following data types:

- Active users
- Inactive users
- Compliance journals
- Public folders
- Social media
- Files
- Stubs/shortcuts
- Documents on legal hold

## Formulating a Strategy to Tackle Legacy Migration Challenges

*Accurate data is the essential raw material that enables the value of enterprise applications*

Understanding the many factors that can contribute to migration failures is key in establishing a strategic and effective migration plan. One of the biggest issues that lead to migration failures is relying on the standard API approach used by most migration platforms.

This strategy uses the archive vendor's aging API to connect and move data to the cloud provider.

- Network overload causing enterprise issues
- Unplanned delays that can interfere with application sunsetting
- Potentially millions of dollars in additional spending on unwanted hardware and software renewals and support

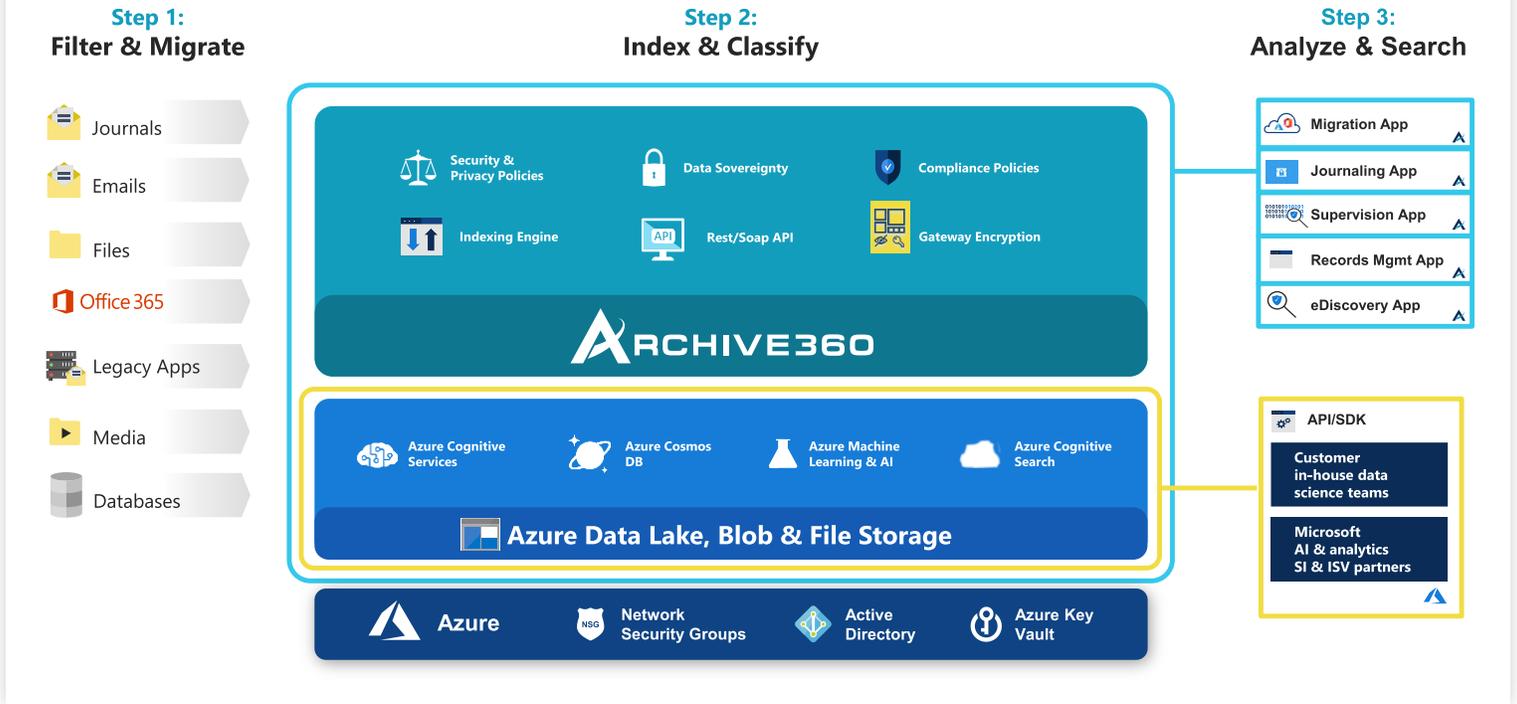
### Issues with this approach include:

- The migration project can take months or years to fully migrate a legacy archive containing terabytes or petabytes of data
- End-user productivity issues
- Error messages - resulting from issues in the legacy archive (corrupt index, etc.), with no clear insight into the cause of the problems - are ported and exacerbated in the cloud store
- Orphaned email stubs that trigger error messages to end-users

### Additional common challenges that arise in cloud migrations:

- Network and security issues
- Hidden database flaws
- High volumes of inactive mailboxes
- Discoverable metadata loss due to stub deletion or insufficient stub rehydration
- Limitations in Microsoft 365

## Archive360 Migration Architecture



### The value of the Microsoft Azure Cloud

The foundation of Archive360’s migration approach is our use of the customer’s Azure tenancy to transport on-premises legacy archives to the cloud of choice. Provisioning a migration in Azure (versus from on-premises) bypasses complexities and wait times and provides a range of efficiencies, including the ability to more easily right-size your on-premises infrastructure.

### At the outset of a migration, we help customers leverage Azure via two options:

1. As a certified Microsoft Cloud Solution Provider, Archive360 can provide the Azure infrastructure for customers that do not already have it in place.
2. For customers with an existing Azure tenant, we help configure the Azure storage account and obtain secure access to copy archive data to it.

### Four Steps to Accelerated Data Migration

Archive360 takes a different approach to overcome every issue that can occur during a legacy archive migration. Cloud-based and fully managed, our methodology provides data accountability and migrations at 10-20x the speed of conventional migration platforms, all without requiring any cloud experience from IT. The process includes the following four key steps.

## Step One: Accelerated Discovery

**De-Bunking Migration Myth #1:** IT often believes that moving data to the cloud requires the migration system to have pervasive, on-premises network access. But data discovery does not need to compromise the network. It is possible to collect and copy up to 60 terabytes of raw archive data in a matter of days, with nothing more than read-only access to the network.

Archive360 doesn't rely on archiving vendor APIs. Instead, it goes straight to the source of the data, providing a faster project start and finish. This low-touch approach ensures minimal disruption to the network, no API issues, no transfer speed restrictions, and full data accountability and autonomy from unhealthy or corrupt databases and indexes.

Beginning with the high-speed transfer of legacy archive data, Archive360's FastCollect supports the collection and migration of over 80 different data types across structured, semi-structured and unstructured data.

- Requires read-only access to the archive to execute a one-time sequential read of all raw storage data and metadata
- Can be installed, configured and migrating data in less than one hour
- Reads legacy systems with no disruption to existing hardware and software infrastructure
- Does not install software on the archive servers, ensuring the archiving software support and maintenance agreement is fully respected
- Reduces performance impact on the storage infrastructure - at a rate of up to 5-10 terabytes per day
- Data collections are fully audited and logged for documented chain of custody to support compliance and legal needs
- Data encryption in transit guaranteed
- Installed on a standalone virtual or physical computer

Once FastCollect is configured, an Archive Insight report is generated, showing the exact amount of data to be transferred for cloud staging.

### An Archive Insight Report is generated providing two important pieces of analysis:

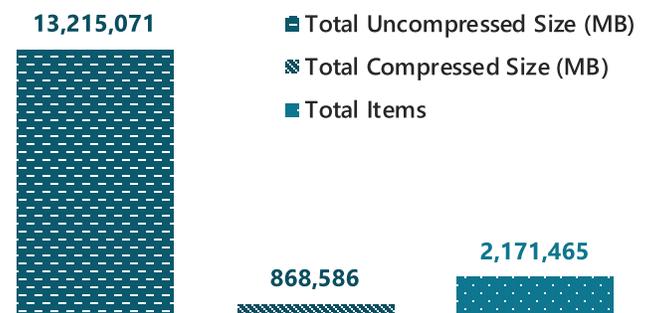
- The total number of messages and the size of the user archive
- The total amount and size of items segregated per archived mailbox, which may reflect users, shared mail, public folders and journal archive

One advantage of this report is the clarity around the total number of inactive users present in the archive. Most organizations do not have this insight, but generally, up to 60 percent of total users within the system are inactive. If not handled correctly, moving them into Microsoft Office 365 can add to issues that stall the migration.

### Decommission Faster

Faster migration means faster decommissioning of legacy systems. After accelerated discovery is complete, the process of shutting down on premises infrastructure can begin at once.

#### Totals



## Step Two: High-Speed Transfer

**De-Bunking Migration Myth #2:** Many migration providers will tell IT that the archive storage volume is the bottleneck that causes slow migrations. This is simply not true! The primary reason migrations move slowly is because of data duplication and the resulting explosion of transfer volumes. Our method of simply taking a one-to-one copy of raw archive data and moving it to the customer's Azure tenancy for staging eliminates this problem.

Archive360 does not conduct any advanced data processing. Instead, our solution is designed to copy data, in bulk, to the cloud, via HTTPS/TLS 1.2, with one click. Even better, while Archive360 is moving the data to the cloud, everything is moving to the customer's own Azure Cloud. Additionally, as the staged data is migrated from the customer's Azure cloud into Office 365, the customer retains total ownership over the data and all encryption keys used.

All data extracted from the archive storage subsystem, including specific SQL metadata in a compressed format and raw legacy archive data, is transferred at the exact same volume as it exists in the archive. There is no decompression or expansion of data. If the data is stored compressed by the archiving system, it is transferred as such.

### Centera Support & Compliance

*If the data is stored on EMC Centera, the raw C-Clips are read and transferred in that format.*

Transfer speed is dependent on the network bandwidth available. Where bandwidth is limited, customers can opt to use physical Drive Shipping. Currently, Archive360 supports Azure DataBox and Azure DataBox Heavy.

Because we circumvent data transfer throttling issues during migrations, Archive360 can transfer data at an average rate of 10 TB per day (on a 1 gigabyte/second line), with larger environments transferring at a rate of 50 TB per day (on a 5 gigabyte/second line). Though this approach reads data sequentially from storage devices, storage subsystem performance (e.g. EMC Centera devices) can impact transfer speed.

Once the data has reached the customer's Azure tenancy for staging, all SQL metadata will be processed, and raw data validated.

The screenshot shows the Archive360 Dashboard with two main sections: 'Databases' and 'Data sources - 100%'. The 'Databases' table lists seven databases with green progress bars for 'Script export' and 'Data export'. The 'Data sources' table lists four Centera data sources with progress bars for 'Discovery' and 'Copy'.

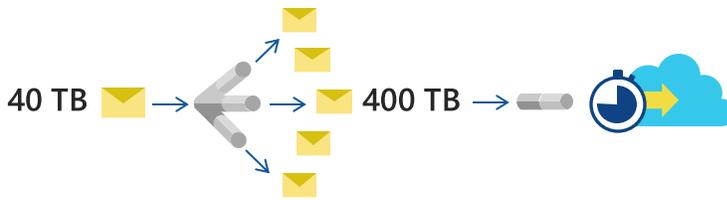
Databases			
Id	Name	Script export	Data export
1	EnterpriseVaultDirectory		
3	EVVSGBMSVAULTSTOREGROUP_1_1		
4	EVVSGBMSVSGIngestions_2_2		
5	EVVJournalVaultStoreEVSVR1A_8		
6	EVJournalVaultStoreEVSVR1		
7	EVVJournalVaultStoreEVSVR2_7		

Data sources - 100%							
Id	Type	Name	Discovery	Copy		%	
1	Centera	Test Vault Store Partition1	26	13.72 MB	26	13.72 MB	100%
2	Centera	Large Mailbox Test Partition1	1.3K	98.32 MB	1.3K	98.32 MB	100%
3	Centera	Journal VaultStore EVSVR1 Ptn1	0	0 B	0	0 B	0%
4	Centera	Journal VaultStore EVSVR1 Ptn1	0	0 B	0	0 B	0%

### On-premises model

*Increases costs, disrupts environment*



### Archive360's one-to-one copy approach

*Lean and low-touch*



## Data Explosion from Legacy Journals in a Conventional Archive Migration

While a legacy archive may only store one copy of an email, when the data is migrated using most migration platforms, the system will write out a duplicate for every recipient on an email. This means that a single email with five recipients would create five copies of the same email during migration. Consider this at scale: 1 megabyte of data instantly becomes 5 megabytes. It's easy to imagine how quickly 40 terabytes of data in an archive of 3,000 users could turn into 200 terabytes of email that will need to be migrated. Because Archive360 simply copies and moves the archive to Azure before any additional work takes place, we are moving the exact same amount of data as resides in the existing system. If it's 40 terabytes, we're only moving 40 terabytes - a huge cost savings. Once the data is in Azure, it is expanded into the full 200 terabytes that is then moved into Office 365. Archive360's unique approach makes a significant impact on time, efficiency, costs and project complexity.

### Making the most of your bandwidth

IT teams have options when it comes to how they schedule data transfers to limit network consumption. Archive360 offers a scheduled approach, so data can transfer only at night, on weekends or during certain low traffic hours. Or, data can move on an ongoing basis in a controlled fashion so that no more than a certain amount of data is transferred per second/hour. During this process, our solution can move data as fast as the network will allow.

#### Typical transfer times per network connection include:

- 1 GB/second line: 10 TB/day
- 5 GB/second line: 50 TB/day

## Step Three: High-Speed Migration to Office 365

**De-Bunking Migration Myth #3:** A common misconception is that archive migration always takes longer than mailbox migration or slows down mailbox transfer to Office 365. While this may be true in conventional migration methods, Archive360 offers a workaround. By staging the process - taking a copy of the data and moving it to Azure before active user migration begins - IT can use all of its network bandwidth to conduct a rapid mailbox migration.

With the Archive Insight Report in hand, Archive360 migration consultants will work with the customer to establish a migration plan, including defining user batches and the migration schedule, and will execute the migration plan on behalf of the customer. The plan includes selecting those archives that are to be migrated to:

- Office 365 End User and Shared Mailbox
- Active Mailbox and Personal Archive
- Archive2Azure

### As a best practice, we recommend:

- Active users are migrated to Office 365 online archive
- Inactive users are kept in Archive2Azure for eDiscovery and compliance purposes
- Journal data is kept in Archive2Azure for eDiscovery and compliance purposes

## Stub rehydration

Many migration tools recreate stubbed messages during migration. But a closer look reveals that metadata values, including message ID, flags and categories, have been changed. In litigation, this could be viewed by the Judge as destruction of evidence. Additionally, incorrect stub management can negatively affect end-users.

During Archive360's migration planning, the customer has the option of either rehydrating stubs (all or by date range) in place (i.e. in the same location/folder as before the migration, which is typically Active Mailbox) or deleting them.

## Archive360's Stub Rehydration Process

- Matches up each live mailbox message stub/shortcut with the original archived message in the legacy archive
- Performs in-place restoration (or rehydrating) of the stub/shortcut with the archived message, preserving all metadata and ensuring forensically complete messages
- Fully reconciles stubs with the original archived messages, and defensibly removes leftover stubs
- Is the only solution that provides this critical capability, breaking the enterprise's dependency on the legacy archive and allowing for a fully compliant and accurate cloud migration

## Migration and Reporting

At the direction of the customer, Archive360 Migration Consultants will begin the migration process. This includes validating each item to ensure its free from corruption. Each item that does not pass validation is flagged and reported for future investigation in a legally defensible exception report.

The following reports are made available to the customer after each batch run:

- User Migration Report that includes: Username | Display name | SMTP | archive msg count | Office 365 msg count | validation alert count
- Rehydration Report that includes: Username | Display name | SMTP | Rehydration count | Rehydration Size

## Step Four: Final Migration and Chain of Custody

**De-Bunking Migration Myth #4:** Common belief is that chain of custody and data fidelity remain when using Exchange Web Services (EWS) to push data into Office 365. That’s incorrect! When not addressed early, it can lead to additional processing headaches. This is true if the archive stores the message in EML format. Most archives use MSG format. Conventional migration tools convert the data from MSG to EML before migration, leading to processing headaches and fidelity loss.

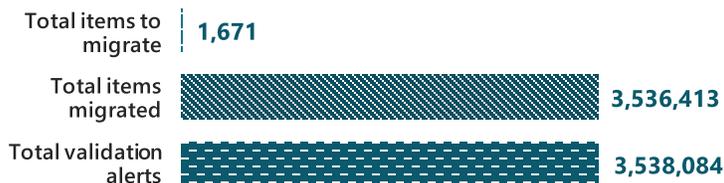
When moving data, most migration vendors first convert the data from its original MSG format to an EML file. Not only does data conversion create a processing overhead and extend the migration process, it also results in loss of data fidelity.

Archive360 doesn’t use EWS; all data is preserved in its native format. Rather than EWS, Archive360 uses the MAPI protocol to move data into Office 365. As a result, the data is not converted and there is no loss in data fidelity.

### Legal and Compliance

Upon successfully completing the migration, a final migration report is generated. Included in this report is the Legal Chain of Custody report, which is critical for legal defensibility so that the organization can account for every message that was in the archive and has been moved to the cloud and Office 365.

#### Total Items Processed



### Top Migration Considerations

Migrations are challenging and filled with misconceptions for even the most sophisticated IT departments. To fully prepare for a successful migration, IT should understand:

- The scope of data being migrated (active users, inactive users, social media, litigation holds, etc.)
- Files involved in the migration
- Legal, compliance and journaling requirements
- Whether stub rehydration is needed
- Security limitations and risks

### Conclusion

Cloud archiving has become a highly attractive and cost-effective alternative to siloed and costly on-premises archiving platforms. Most organizations are planning or are in the process of migrating the majority of their on-premises data to the cloud. However, organizations need to be aware of the issues related to the migration of large datasets. Regulatory compliance, data privacy, litigation/eDiscovery, and end-user productivity need to be taken into consideration. Choosing the right migration vendor with the best proven technology will save you time, costs, and risk in a large data migrations.

## About Archive360

Archive360 is the enterprise information archiving company that businesses and government agencies worldwide trust to securely migrate their digital data to the cloud, and responsibly manage it for today's regulatory, legal and business intelligence obligations. This is accomplished by applying context around the search, classification, and indexing of data including files, videos, and emails—all while allowing organizations to maintain full control over privacy, access, and compliance. Archive360 is a global organization that delivers its solutions both directly and through a worldwide network of partners. Archive360 is a Microsoft Cloud Solution Provider, and the Archive2Azure™ solution is Microsoft Azure Certified. To learn more, please visit [www.archive360.com](http://www.archive360.com).

Copyright © 2020 Archive360, Inc. Archive360 and Archive2Azure are trademarks or registered trademarks of Archive360, Inc. The list of trademarks is not exhaustive of other trademarks, registered trademarks, product names, company names, brands and service names mentioned herein are property of Archive360, Inc., or other respective owners. All rights reserved

A360TG0120009