

Executive Summary

Migrating huge amounts of data from on-premises systems to the cloud is time-consuming, costly, and error-prone if not done with the right planning and technology. Employee productivity can suffer, regulatory responsibilities can be put at risk, and the legal chain of custody can be lost - putting your legal defensibility at risk. The right technology can remove these issues by ensuring you use the correct technology and procedures instead of what a vendor wants to sell you.



Microsoft and Archive360 work together to ensure your valuable data - including email archives, retired application data, databases, Salesforce data, and medical data - arrives quickly and safely in your Azure tenancy for active management by the Archive2Azure platform.

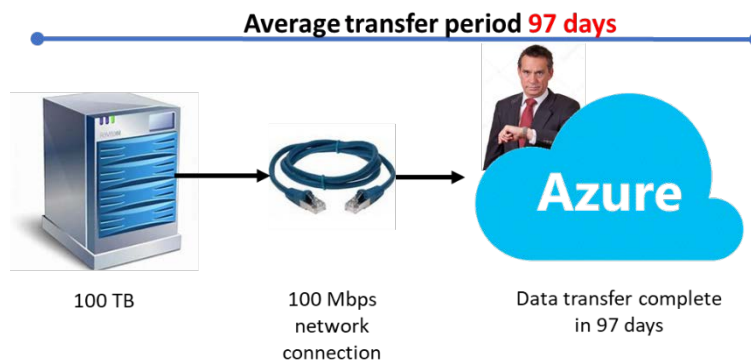
“...85 TB can be moved into Azure – start to finish – in less than ten days. That’s as fast as a 1 Gbps connection!”

Challenges in Transferring Large Data Sets to the Cloud

Microsoft Azure provides agility as well as built-in infrastructure, infinite scalability, security, and cost savings. However, moving existing workloads (applications and data sets) from your on-premises datacenter to Azure can be complex, costly, and time-consuming.

For small data transfers, corporate networks are the best first choice; however, when the volume of your data sets approach hundreds of terabytes or even petabytes, or your network is slow or overburdened, onboarding your data to Azure over the wire can be a challenge. Worse, some methods for moving data can raise risks associated with your regulatory compliance and litigation preparedness and response.

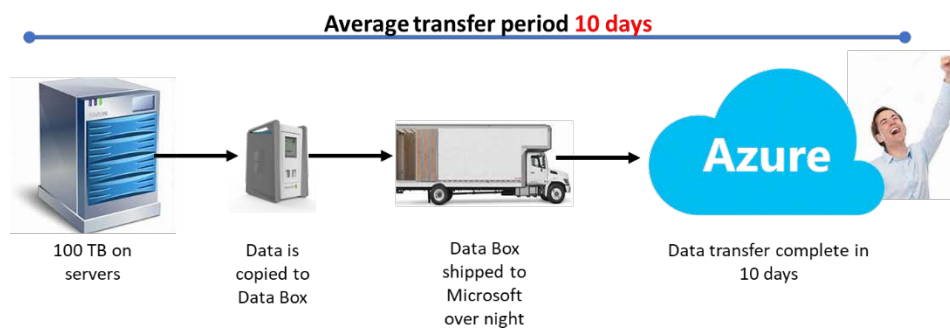
For example, 100 TB of data would take approximately 97 days to migrate into Azure using a network connection with 100 Mbps of bandwidth. As a result, the time-to-value for your data estate modernization goals is dramatically – and, as you will soon see – needlessly extended.



Data Migration with Azure Data Box

Azure Data Box is a service that enables the rapid transfer of large data sets from your on-premises repositories to your Azure tenancy, thereby reducing your costs, mitigating your risks, and accelerating your time-to-value. The Data Box migration process is simple and secure:

1. Request a Data Box from Microsoft or Archive360
2. Receive the Data Box via overnight shipping and copy your data sets to the Data Box
3. Ship the Data Box overnight to the Azure Data Center via standard logistics providers like UPS, FEDEX, or DHL
4. Microsoft inspects and uploads your data into your Azure tenant
5. Once your data has been successfully uploaded to Azure, the Data Box is wiped and sanitized in accordance with NIST 800-88r1 standards



The amount of data to be transferred and the capacity of your network will determine whether a Data Box is right for you. The table below outlines the approximate time it takes to transfer data to Azure at various bandwidth speeds. The Data Box transfer rate is equivalent to the 1Gbps column.

Data Qty	45 Mbps (T3)	100 Mbps	1 Gbps	10 Gbps
35 TB	76 days	34 days	3 days	8 hours
100 TB	216 days	97 days	10 days	1 day
500 TB	3 years	1 years	49 days	5 days
1 PB	6 years	3 years	97 days	10 days

The Archive2Azure Platform

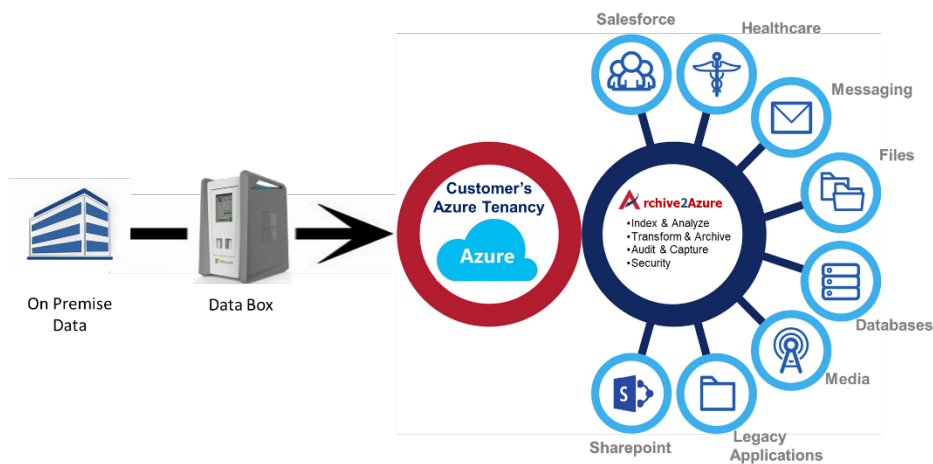
Archive2Azure is a native Azure solution for archiving and long-term information management. Archive2Azure provides a highly secure, low-cost, and compliant intelligent information management system for archiving and management of all types of data.

Archive2Azure works in conjunction with your Azure tenancy to capture, store, index, search, and manage large volumes of data. Archive2Azure creates containers which store and manage data in “cabinets” on specific Azure storage tiers – Hot, Cool, and Archive. Your data can then be moved across tiers, either via policies or manually, ensuring the most efficient performance level as well as greatest cost savings.

Archive2Azure further de-risks your data estate modernization initiatives because it frees you forever from vendor lock-in – meaning that onboarding your data to Azure will be the last migration you ever have to do:

- Data is stored in YOUR Azure tenant, so you never need to hand over sensitive data to a third-party, proprietary cloud archiving vendor who can hold your data at ransom if you ever want to get it out
- Your data is stored in its NATIVE format
- Your data is secured with YOUR own encryption keys to which no one else has access

Archive2Azure actively manages unstructured, semi-structured, and structured data, enabling you to gain deep insight into your data through powerful analytics to realize valuable business intelligence. And your time-to-value is accelerated by using Azure Data Box to quickly and securely transferring your data into Azure.



Archive2Azure – Functional & Technical Summary

Archive2Azure is an intelligent information management platform that can be configured in a Software-as-a-Service (SaaS) or Platform-as-a-Service (PaaS) infrastructure.

Archive2Azure runs in Microsoft's Azure datacenters. Data (along with the associated metadata) is stored in its native format in YOUR dedicated Azure subscription, encrypted with YOUR storage account encryption keys, in an open-standard format and full-text indexed by either Azure Search or Archive360 Search.

Archive2Azure orchestrates Azure's Cognitive Services for translation, transcription, and OCR functions. The platform also utilizes Azure AD integration for security, full platform audit and logging, data retention policies, matter creation, tagging, review, and native export. Immutable (WORM) storage, Microsoft's commitment to data privacy and regulatory certifications, is also available for those industries that have that specific regulatory requirements.

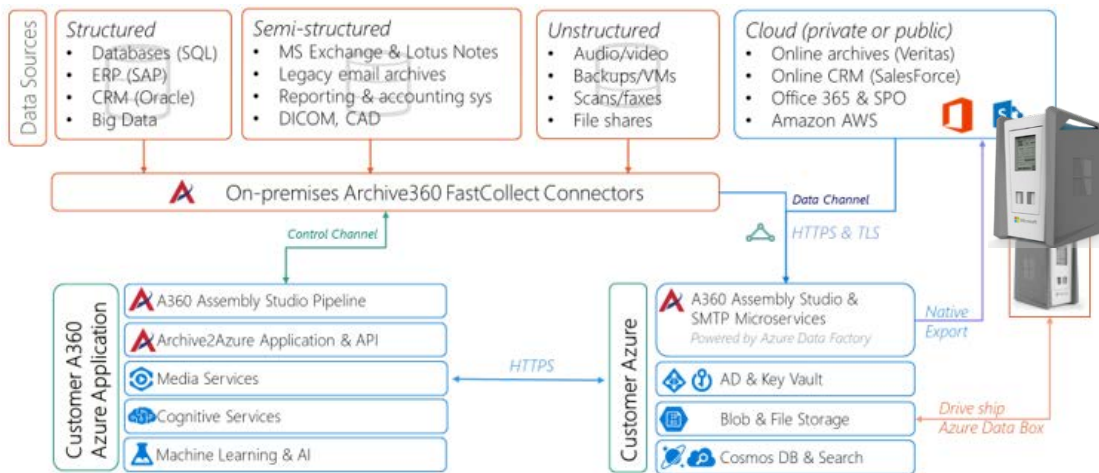
In summary, onboarding and validation, classification, analytics, search, review, and retention policies are Archive2Azure's strengths.

Archive2Azure Solution Design

Archive2Azure has more than 100 connectors for data ingestion which cover structured, semi-structured, and unstructured data - including custom connectors for customer-specific needs.

Onboarding is flexible: Archive2Azure can leverage source APIs, email journaling, polled-and-event-driven transfer, SMB protocols, FTP, as well as other methods. After data is onboarded, Archive2Azure orchestrates Azure's Machine Learning, AI, Cognitive Services (for translation and transcription), and Media Services for support of all searchable file types - including voice and video files.

Archive2Azure supports all Azure data storage tiers include WORM, Hot, Cool, and Archive. Metadata is stored in Cosmos DB as well as Blob Storage for compliance purposes.



Content Type/Data Flexibility

As you expand your data footprint within Archive2Azure, the ability to manage an expanding number of data formats will be a distinct advantage for information management as well as a key enabler for cost control. Archive2Azure can store and manage over 3,500 file formats of structured, semi-structured, and unstructured data types.

Archive2Azure tightly integrates with Microsoft Office 365, OneDrive, and SharePoint Online to help you meet your compliance requirements and budget constraints if you are migrating to and/or archiving data from these systems.

Validation

Data verification is at the core of both FastCollect and Archive2Azure. Archive2Azure uses MD5 hash (Digital Fingerprint) matches to ensure that records are accurately migrated and have not been edited once retained. Pre-migration of each item is hashed with a unique fingerprint (MD5). Once the item has landed in Archive2Azure, a new file fingerprint is generated. Only if the two hashes (digital fingerprints) match will the file be accepted into the system. This ensures file immutability. Only once the validation and source-item-ID to Archive2Azure-item-ID is established will the item be committed to Archive2Azure.

Data Storage Management

Archive2Azure supports all Azure storage regions, tiers, and redundancies. Hence, you have the option of hosting data in different geographic regions with full support for Hot, Cool, and Archive storage tiers. Intra-datacenter resiliency or region resiliency is supported as well. Azure storage provides object-level WORM capabilities as well as customer-provided object encryption keys to meet stringent security requirements. In effect, the Archive2Azure storage model meets many industry and government requirements. In addition, it enables a complete reconstruction of the entire application system in the event of a catastrophic failure. The use of Azure Blob Storage allows an unlimited storage capacity with Azure cloud storing Exabytes of data.

Advanced Search

Archive2Azure provides extremely powerful search capabilities. You have the option to create as many search fields as needed when ingesting data, and you can use any combination of those fields to build more elaborate search queries - which allows you to be very granular with the data that your searches retrieve.

Archive2Azure further allows you to build advanced search queries by combining all the above functionality with logical operators, wildcards, and special symbols in any desired combination. Fuzzy searches, logical operator searches, proximity searches, wildcard searches, and many more are all available and can be used against all the archived data.

In addition to all the above, Archive2Azure enables you to benefit from Azure Cognitive Services so you can do content searches against media files (audio and video) or scanned images.

Anonymization and Redaction

Many regulations specify sensitive data such as PII and PHI be protected and audited. However, blocking access to an entire file which contains sensitive information can cause productivity issues. Utilizing access controls and individual permissions, Archive2Azure can anonymize or redact specific content at the file level in real-time ensuring sensitive content within a file cannot be viewed by unauthorized employees. The advantage of Archive2Azure anonymization and redaction is that it still provides access to files (with sensitive data masked) for other work by employees without a specific level of authorization. Note: the redaction process replaces the sensitive data with a black block showing where in the document the PII was located.

On the other hand, the anonymization process replaces sensitive data with an encrypted character string making the PII unreadable. The encryption string is a series of random characters based on your encryption key (Archive2Azure never sees the encryption key). Using this anonymization process, PII can either be visually restored by using the correct key or can be made permanently unreadable by simply destroying the encryption key.

With either anonymization or redaction, both the original message and the rendition version can be indexed with different access authorizations. Additionally, Archive2Azure audits all actions in the system so any attempts at viewing PII by unauthorized employees (or outside hackers) is recorded.

eDiscovery

eDiscovery and litigation preparedness are unavoidable facts of life. Not taking it seriously - or worse, not being informed about the company's responsibilities - can cost you later in loss of cases and payment of huge fines and judgments. One of the main requirements of a compliant eDiscovery process is the need to ensure potentially relevant data be protected from loss, deletion, and change (editing).

Many corporate legal departments and law firms have begun to adopt the cloud as the best place to store their legal data, i.e., a secure legal repository for data on litigation hold, eDiscovery data sets, and attorney work product. Because of Azure's customizable services and world-class security, and the fact that corporate legal departments and law firm file systems have become a major target of hackers, the cloud has become a less costly, massively scalable, and more secure destination for legal data. Archive2Azure adds the key technology capabilities to enable the import, search, review, management, and export of data involved in litigation in a legally-defensible manner.

And if you are in the middle of litigation, the Azure Data Box adds the needed capability to move large eDiscovery datasets to your Azure tenancy in a legally-defensible manner quickly and with a reportable chain-of-custody history.

Litigation (Legal) Hold

Archive2Azure offers two types of litigation hold policies: Matter-based litigation hold and Custodian-based litigation hold:

- Matter-based litigation holds are managed via reviewer assigned tags. Each tag must be associated with a matter. A matter can have multiple litigation hold tags associated. Messages relating to a specific matter can be manually or automatically placed on legal hold at a saved search level, at the bulk message level, or at the individual message level. Within a matter, identified messages or search results can have one or multiple litigation hold tags applied without the need to save the message multiple times. A file can only be disposed of when all legal holds have been lifted from the individual file.
- Custodian-based litigation holds ensure all existing and new/incoming messages or files for a specific custodian are placed automatically on legal hold. Only once all legal hold tags are lifted can a message be disposed of.

Review and Production

Archive2Azure provides the ability to view individual files in returned search results. Archive2Azure allows you to preview individual items in the search results as well as to view all metadata related to specific items. You can also download all or a sub-selection of a results set. You can open and view selected records and, if authorized, export them or the entire results set. Export allows for the inclusion of metadata as well as manifest in either CSV, EDRM v2, XML, TIFF, PDF, and the file's original format.

Disposition

Archive2Azure manages disposition at the file level. In all cases, a legal hold overrides all associated disposition policies. Archive2Azure allows the file retention period to be extended or shrunk via either automatic or manual approval policies. Disposition policies are based on three individual components:

1. Record Date
2. Custodian
3. File classification tags and sub-tags

You have the capability to create sophisticated disposition policies so that data classified with a specific tag, within a specific matter, or for a specific custodian and date can be automatically destroyed. Ad-hoc data disposition also is supported. Disposition policies can be run against the original record which in turn will dispose of the associated renderings. Rendering specific disposition jobs can also be executed.

Data can be disposed of via disposition job schedule, by date, by custodian, and additional metadata. File disposition is accomplished in two ways: automatic or manual. Automatic file deletion occurs when the retention period for a file expires. If the file is not on litigation hold, the file is included in the deletion process which normally runs at midnight of that 24-hour period. Additionally, the automatic disposition process can be turned off so that records managers, etc. can access the system and look at those files designated as expired and manually dispose of them. This has a major benefit for records managers in that they have additional control of the disposition process. Customer attorneys can have greater/faster control of file deletion to ensure responsive files are not deleted due to automated disposition in case an unexpected eDiscovery situation arises.

Audit Logging

Archive2Azure includes audit logging which is designed to log every action or command taken by a user or administrator. All retention policy configurations, search results, permission changes, system accesses, read attempts, or dispositions (or attempts to dispose) of objects are fully audited. All audit logs are preserved and readily available for offline analysis and preservation. Audit data includes the following information:

- Who performed the operation (regardless of whether the operator is the customer's user, administrator, or third-party personnel working with you)
- Date and time when the operation was performed
- The actual commands used to perform the operation
- The details of the operation, such as the scope of the action
- The results of the operation (if applicable)

A Disposition Report is generated for each disposition task. The report lists the following:

- The exact retention policy being executed
- The objects affected by such policy
- The Record/Create Date, Object ID, Expiration Date, and other metadata for each object
- The result of the disposition job (i.e., success or failure due to retention date or legal hold)

Regulatory Compliance

Complying with regulatory retention requirements and responding to agency information requests is closely related to the eDiscovery process. Beyond the eDiscovery requirement to protect and turn over responsive content during the eDiscovery phase of a lawsuit, ensuring compliance also requires a records management regimen that ensures specific document classes be retained for specific periods of time and be available for inspection when requested by the regulatory authority.

Archive2Azure gives you a powerful information management capability which includes content retention based on customer-defined retention periods, disposition based on those same time periods, legally-defensible data export, automatic auditing, and powerful reporting tools.

About Archive360

Archive360 is the world's leader in intelligent information management for the Microsoft Cloud. Its proven platform enables organizations of any size to drive down the cost, risk and uncertainty of digital transformation to and in the cloud. Archive360 is a trusted partner for onboarding, validating, and managing all enterprise digital assets, while delivering predictability, defensibility and data insights. Archive360 provides non-proprietary information management that ensures security-focused infrastructure independence. Archive360 is a global organization with a worldwide network of partners. The Archive2Azure Platform is Microsoft Azure Certified. To learn more, please visit: www.archive360.com.

+1 (212) 731-2435 | info@archive360.com | www.archive360.com

Copyright ©2018. Archive360 is a registered trademark and FastCollect and Archive2Azure trademarks of Archive360, Inc.



Gold Cloud Platform
Gold ISV
Gold Messaging
Silver Application Development

