Tape Ark & Microsoft Azure
Financial Sector Cloud Migrations

A partnership to help liberate the world's largest collections of data.

Why does Tape Ark exist?

Tape Ark's mission is to liberate the largest collections of data on the planet into the Microsoft Azure cloud so that the data can be used to fuel big data, machine learning, AI and analytics.

Tape Ark's ground-breaking technology expedites the ingest of tape and disk-based data sitting underutilized and inaccessible to modern cloud technologies, where as most enterprise customers around the world use offsite storage warehouses to store their offline tapes.

Once the data is liberated, it can enable organizations to efficiently store and gain access to their data, scale their IT operations and keep their data and backup infrastructure up to date.

As the only dedicated specialists capable of mass migration of tape media in a wide array of data formats, Tape Ark has worked with customers across a wide range of industries including media and entertainment, oil and gas, and state and federal governments to help them liberate their valuable tape bound data to the cloud.

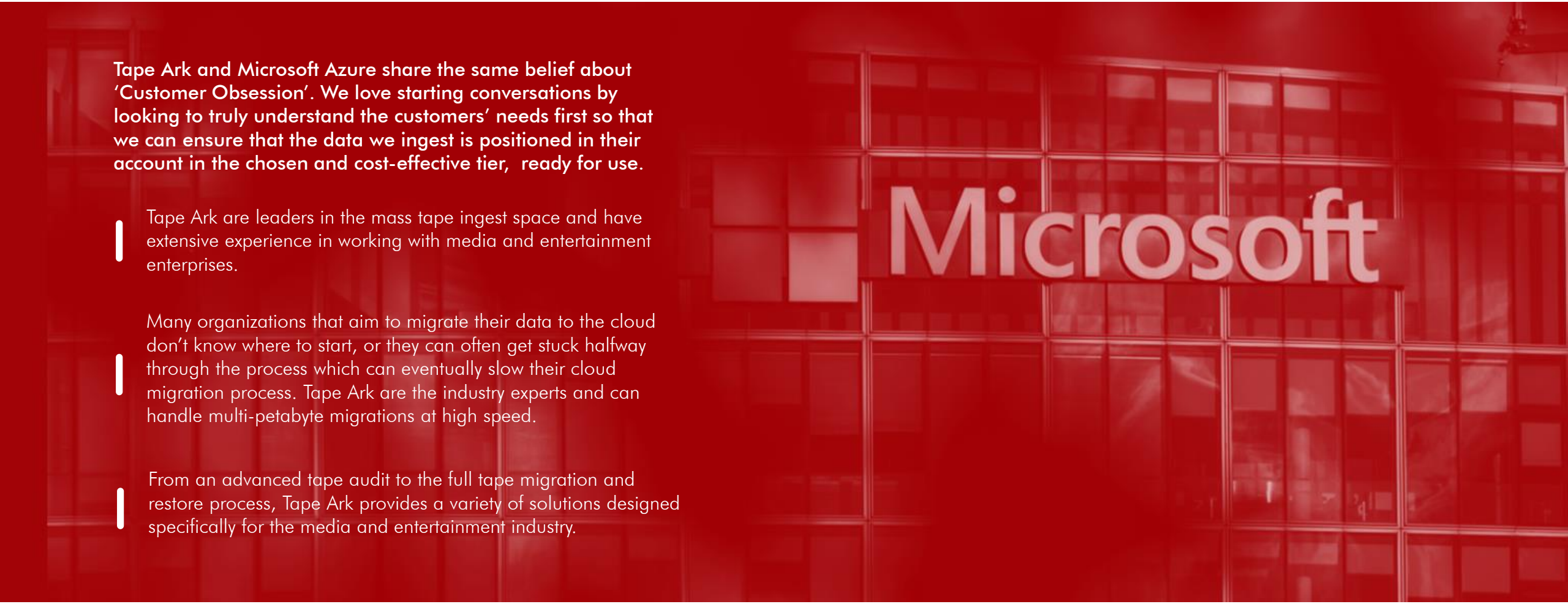
Tape Ark has the knowledge and expertise to make legacy datasets accessible from almost any format and our innovative solutions are powerful, cost-effective, and forward-thinking, plus we love solving customer problems.

No matter the size or complexity of the solution required, Tape Ark will have its 'A' team working on it.

Benefits of using Tape Ark

Tape Ark and Microsoft Azure share the same belief about 'Customer Obsession'. We love starting conversations by looking to truly understand the customers' needs first so that we can ensure that the data we ingest is positioned in their account in the chosen and cost-effective tier, ready for use.

- Tape Ark are leaders in the mass tape ingest space and have extensive experience in working with media and entertainment enterprises.
- Many organizations that aim to migrate their data to the cloud don't know where to start, or they can often get stuck halfway through the process which can eventually slow their cloud migration process. Tape Ark are the industry experts and can handle multi-petabyte migrations at high speed.
- From an advanced tape audit to the full tape migration and restore process, Tape Ark provides a variety of solutions designed specifically for the media and entertainment industry.



Mass Tape to Cloud Ingest

Tape Ark's scalable approach to tape migration rapidly liberates client data to their chosen cloud storage tier enabling organizations to leap into the world of opportunities on offer in the Microsoft Azure ecosystem. Then, with their data assets restored and highly accessible, they are positioned to get maximum value from their most precious assets. With this new access to their data, customers are able to harness new insights and monetize their previously inaccessible data.

- Mass tape data ingest irrespective of recording format
- Operate at a massive scale
- Provide an Evergreen solution
- Access to advanced analytics, AI and ML

The Tape Ark data restore process is done without the use of the customers original software or hardware. Preservation customers can elect to retain all the original tape formatting or have the data restored back to its original file state.

Tape Ark handles all of the major backup formats including TSM, Netbackup, Arcserve, Commvault, BackupExec, etc. Because the customer no longer has tape as part of their internal infrastructure requirements, they can then dispose of large expensive IT assets that in many cases can take up massive amounts of data centre space.

Mass Tape to Cloud Ingest

Tape data restore without the original software

For many organizations, their tape-based data is encapsulated in a format linked to the software that wrote it to the tape. Tape Ark's proprietary systems offer tape data liberation and restore without the need for the original software, and after restoring the data, it is presented to the customer in its the original native or raw format ready for ingest into their modern cloud-based systems.

Fast access

The Tape Ark tape to cloud migration service enables customers to get fast access to legacy content – anytime and anywhere, without the need to maintain expensive legacy hardware or software. Once the data is in the cloud, it can be connected to new applications and used in AI & ML workloads to make new discoveries from their historical content. The new use cases enabled by the Tape Ark process means greater utility and increased value from the legacy data which could never have been achieved from tape bound data. With demand increasing for fast remote access, there is no better time to migrate your data assets to the cloud.

Evergreen solution

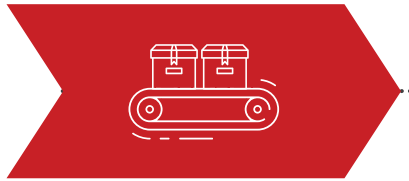
Tapes have a limited lifespan and come with the burden of tape refresh projects that need to be carried out on a frequent basis. Updating from old to new media tapes is time consuming, resource heavy, expensive and guarantees you will be in the same situation every few years. The Tape Ark solution we provide on Microsoft Azure is an evergreen solution that means customers never have to worry about updating tapes or infrastructure allowing them to focus on their business.

Restore on demand

For customers who have no immediate use case for their data but want to preserve their content and retire internal legacy hardware and software systems, the Tape Ark restore on demand is a perfect solution. Tape Ark provides a fully managed 'restore on demand' service that means reduced ingest and reformatting costs for customers, but with the confidence that the data can be restored if or when required. Tape Ark provide tape restoration services that can restore entire datasets or partial restores for selected files only. In addition, Tape Ark supports customers no matter which backup applications they prefer.

Tape Ark Audit and Ingest Process

Processing steps for every tape



01

Collect Media

02

QR Code

03

Photograph Media

04

RFID Chip Capture

05

Read Tapes

06

Generate Metadata

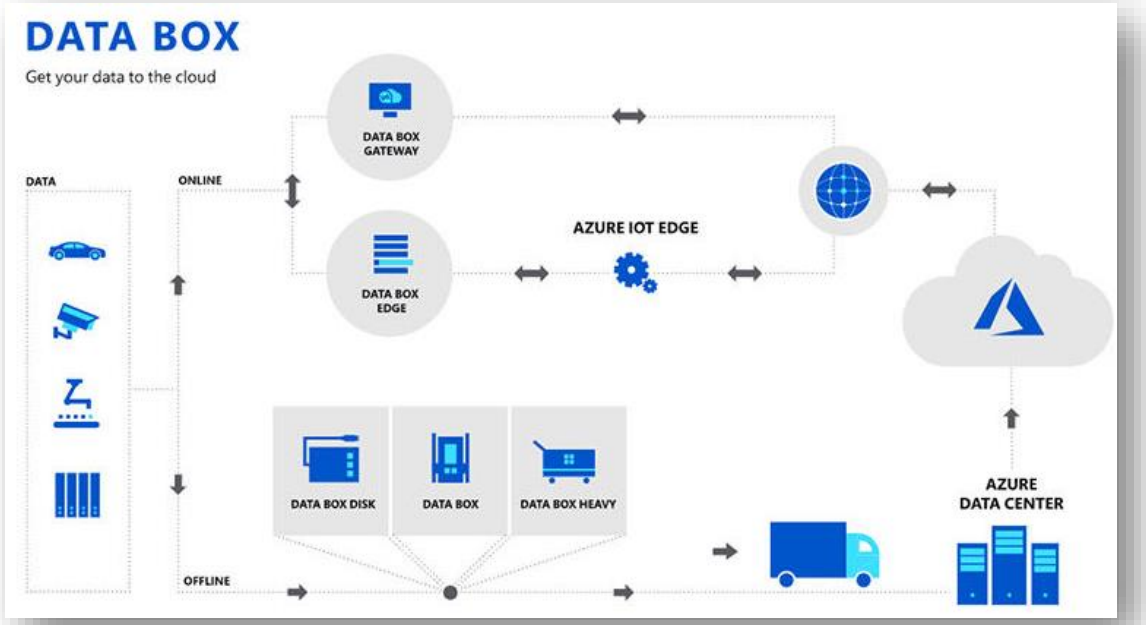
07

Mass Ingest

08

Dispose of Tapes

Microsoft Azure Architecture



Sample Only: Each Client Ingest Implementation is different

Tape Ark and Microsoft Azure have developed a solution to liberate NetBackup data for the financial sector to the cloud, making it available to analytics, machine learning, and collaborative workflows.

The high-level workflow solution starts by receiving media and performing a detailed tape media audit. This allows financial and insurance companies to predict the cloud footprint they will create from their data at a granular level, seek out duplicates, and remove data for ingest that may be licensed or not part of their core archive.

After audit, all data is ingested to the nominated cloud account using Tape Ark's highly scalable technology stack. As data is ingested into a client's account on Azure, automated checksum and name validation checks are carried out. These can be pre-prepared by the client so that real-time ingest quality control can be performed in an automated way.
After ingest and automated quality control, data tiering policies automate the movement of data to the nominated tier as requested by the client. As part of the transfer, JSON metadata manifest files can be created and placed into customers' Microsoft AZURE accounts to update their internal databases.

The workflow solution starts from Tape Ark's internal rapid mass tape ingestion platform Arkbridge, which was developed to utilize the strength and security of the Microsoft Azure Cloud. This internal system has resulted in a highly automated Internet of Things (IoT) approach that eliminates the need for any manual process to ingest media into Microsoft Azure client accounts, providing high accuracy with efficiency. It uses secure API's to execute Microsoft Azure functions that store the ingested tape metadata in Microsoft Relational Databases.

The solution uses a unique combination of photograph processing, OCR, and IoT for managing tape reads, which has helped Tape Ark to simplify and streamline the process of tape ingestion at scale.

The backup data is converted into objects and stored in Azure Hot Storage, either via direct tape to cloud ingest or via Microsoft Databox. Object data gets stored in Hot Storage for faster accessibility and or Azure Warm or Cold Storage for deep archiving. Finally, the object data in Hot storage becomes the source for end customers' machine learning applications.

Rock and Roll Hall of Fame

When the Library of Archives was commenced by the Rock and Roll Hall of Fame, an LTO based tape storage and backup system were liberated the content to the system was designed for preservation level content and allowed the Rock Hall team to manage large files such as three hour-long videos in 4K resolution.

As the years passed the technical infrastructure required to store the files became expensive to maintain and was fraught with technical challenges. Microsoft Azure introduced the Rock Hall to Tape Ark, and together we liberated the content from its TSM Backup format on LTO tapes to S3 Glacier Deep Archive. Tape Ark's tool set also generated matching file hash values to compare against the original file preservation tags to ensure the files were copied as a perfect matching copy.

Tape Ark managed the end-to-end migration including auditing, restoring, and ingesting the files onto Snowballs so it could be ingested into Amazon S3 Glacier Deep Archive.

"Not only did Tape Ark recover the files we needed, they also found an extra 109 files that we did not know where on the tapes. I would say that is an over 100% success rate"

Heidi Quicksilver, Senior Director of Systems and Strategy, Rock and Roll Hall of Fame.

Some of our customers

About Tape Ark

Tape Ark's mission is to liberate the world's largest collections of data from tape. These collections of data are often idle and inaccessible and are typically stored in offsite storage warehouses where it provides little value to its owners.

Customers want to both access their data and have the opportunity to apply today's analytics tools to gain greater value and insight from their legacy data, and Tape Ark has innovated every step of the tape processing stream to create a unique technology stack that scales and efficiently migrates legacy tape data to the cloud at an unparalleled scale. The application of today's data analytics, AI and ML tools has the potential to make profound discoveries and insights that will transform organisations and humanity. Tape Ark can help to capture, manage, secure and analyze legacy data so that customers can get maximum value from their data assets.

To find out how we can help liberate your data or for more information about Tape Ark, please visit tapeark.com.

For more information on Tape Ark & Microsoft Azure partnership, please visit our website at www.tapeark.com or contact Sales at sales@tapeark.com.

Tape Ark

Azure

Tape Migration Scoping

Many of our clients have used multiple backup software applications over their history. If you know all of your legacy systems, please provide as much info as you can for each backup system that you know and the volumes and type of media that you currently maintain. If you do not know the software formats used, where possible, please supply media volumes and types. This can usually be found on your offsite vault storage bill.

Please tell us about your legacy tape data requirements: Please specify the type of media? (Circle where more than one option is listed)

Qty		Qty	
_____	9 or 21 Reel Track Tape	_____	Sony AIT1/AIT2/AIT3/AIT4 & SAIT
_____	4mm DAT,DDS1-5,8mm Exabyte/Mammoth	_____	T10000A,B,C or D
_____	3480/3490/3490E	_____	Floppy, optical or CD & DVD
_____	3590/3590E/3590H or 3592 Jaguar	_____	QIC Tape SLR1-7, QIC DDS1-5,
_____	9840/9840B/9840C or 9490EE	_____	QIC TR1/TR2/TR3/TR4/TR5/TR6
_____	DLT2000/DLT4000/DLT7000/DLT8000	_____	LTO1/ 2/ 3/ 4/ 5/ 6/ 7/ 8/ 9 (circle one)
_____	DLTVS80 & DLT VS160	_____	Solid State Memory (HDDs)
_____	SDLT220/SDLT320/SDLT600	_____	Other – please specify
		_____	

System and Software

Please specify the format and version of the software that wrote the media? _____

What operating system does the data you need recovering originate from? _____

What file system/software were the media written under? _____

What Media Asset Management System was your data stored in? _____

Applied Hardware and Software Regimes (Circle)

Media compression applied?	Yes	No	Unknown
Media encryption implemented?	Yes	No	Unknown
Encryption key available?	Yes	No	Unknown
Was deduplication applied?	Yes	No	Unknown

Copy this page as many time as you need for each collection you have.