SOLUTION ARCHITECTURE

Dataverse Data Backup & Archival Solution - Architecture

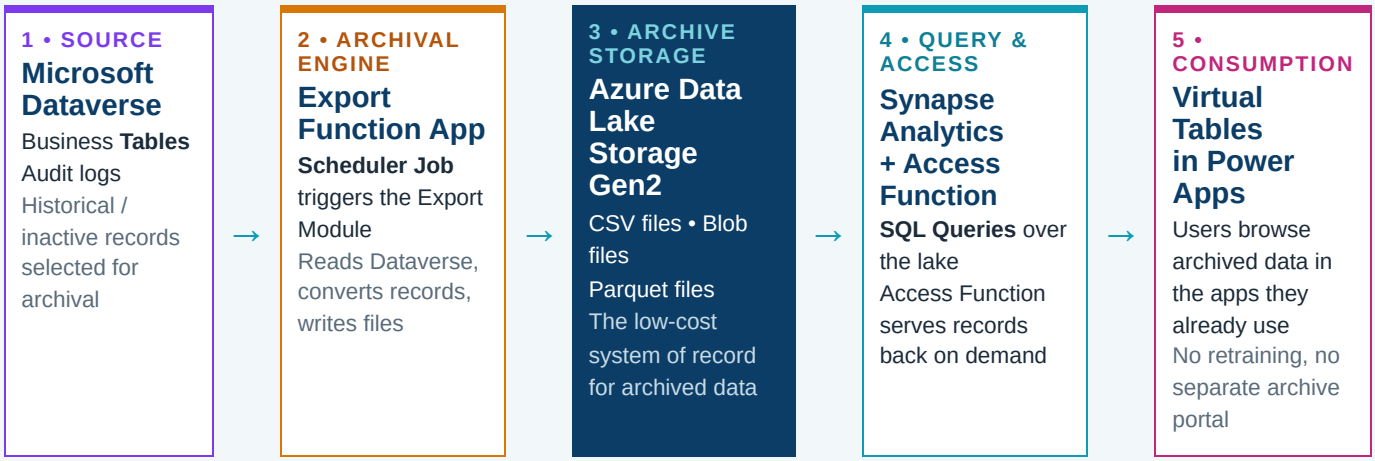
Historical Dataverse data is moved out of premium Dataverse storage into low-cost Azure storage - without disappearing. Users keep finding and opening archived records in Power Apps exactly as before, and analysts can query the same archive with SQL.

Lower storage cost
Ageing tables and audit history leave expensive Dataverse capacity for Azure Data Lake Storage Gen2.

Nothing is hidden
Archived records surface again through Virtual Tables in Power Apps and through Synapse SQL.

Built for audit
Business tables and audit logs are archived together, in open CSV, blob and Parquet formats.

ARCHITECTURE AT A GLANCE - READ LEFT TO RIGHT



Export path (top flow): Scheduler → Export Function → Data Lake. Runs on schedule, converts records to CSV / blob / Parquet.

Read-back path: Data Lake → Synapse SQL + Access Function → Virtual Table → Power Apps. Archived data behaves like live Dataverse data.

What moves
Historical table records, audit logs and their attachments - the data that grows fastest and is opened least often.

What stays live
Current, active Dataverse records stay in Dataverse. Archival targets history, not the working set users edit every day.

For reviewers
Page 2 details every component and the 5-step flow. Page 3 covers formats, access paths and design principles.

EVERY COMPONENT HAS ONE JOB

Two Azure Function Apps appear in the design, and they are deliberately separate. The Export Function only writes the archive; the Access Function only reads it back. Keeping those responsibilities apart makes scheduling, scaling and troubleshooting straightforward.

COMPONENT	ROLE IN THE SOLUTION	KEY DETAILS
Microsoft Dataverse	Source system	Business tables and audit logs. Historical and inactive records are the archival candidates.
Scheduler Job	Trigger	Starts the Export Module on a defined schedule, so archival runs predictably and off-peak.
Export Function App	Archival engine	Reads Dataverse records and converts them to CSV files, blob files and Parquet files in the lake.
Azure Data Lake Storage Gen2	Archive storage	Low-cost, scalable home for archived data in open formats - not a proprietary backup silo.
Azure Synapse Analytics	Query layer	Runs SQL queries directly over the lake files, for reporting, analysis and feeding the access layer.
Access Function App	Read-back service	Retrieves archived records from the lake / Synapse and supplies them to Virtual Tables on demand.
Dataverse Virtual Tables	Transparent presentation	Present archived data inside Dataverse / Power Apps as if it had never left.
Power Apps	User experience	Business users search, open and work with archived records in their existing apps.

THE FLOW, STEP BY STEP

1. **Schedule fires.** The Scheduler Job starts the Export Module in the Export Function App.
2. **Records are read.** The export reads the selected Dataverse tables and audit data.
3. **Records are converted.** Data is written to Data Lake Storage Gen2 as CSV, blob and Parquet files.
4. **Data becomes queryable.** Synapse Analytics exposes SQL queries over the lake files.
5. **Users see it again.** The Access Function feeds Virtual Tables, and Power Apps shows the archived records transparently.

OPEN FORMATS KEEP THE ARCHIVE USEFUL

The archive is not a black box. Each record type lands in the lake in a format chosen for how it will be used, so the data stays readable by standard Azure tools long after it leaves Dataverse.

Parquet files
Columnar, compressed format for table data. Efficient for Synapse SQL scans, reporting and analytics over large historical volumes.

CSV files
Plain, universally readable export of records. Simple to inspect, exchange and reload without specialised tooling.

Blob files
File and attachment content preserved as blobs alongside the record data, so documents stay attached to their history.

THREE WAYS BACK TO THE DATA

1. In the app - Virtual Tables in Power Apps. The everyday path. A user opens an old record from the same app and the Access Function retrieves it from the archive behind the scenes.

2. With SQL - Synapse Analytics. The analytical path. SQL queries run directly over the Parquet and CSV files in the lake, for reporting, compliance extracts and ad-hoc investigation.

3. In the lake - Data Lake Storage Gen2. The direct path. Archived files remain in open formats in Azure storage, under the organisation's own control.

DESIGN PRINCIPLES

- **Cost follows age.** Data that is rarely touched should not sit on premium Dataverse storage; the lake is the right cost tier for history.
- **Archival is not deletion.** Tables and audit logs move together, so a record keeps its history and its attachments.
- **No new user behaviour.** Access comes back through Virtual Tables and Power Apps - the tools people already know.
- **Serverless where it counts.** Function Apps and Synapse SQL scale with the archival and query workload, rather than running idle infrastructure.
- **Standard Azure building blocks.** Dataverse, Functions, Data Lake Storage Gen2 and Synapse - no exotic dependencies to support.

Confirm schedules, retention rules and table selection with the implementation team before deployment. The showcase that follows (pages 4-5) shows the solution running for a real customer.

SHOWCASE - PROOF, NOT PROMISES

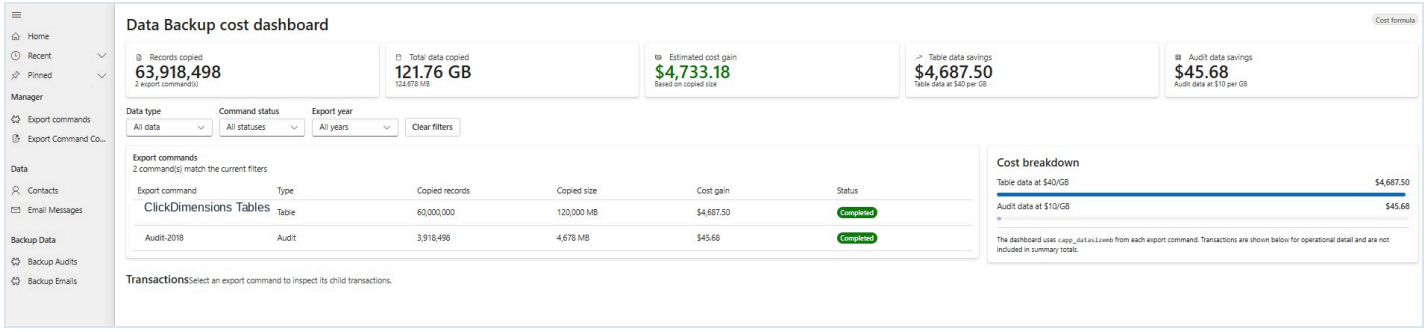
One customer saves $4,733.18 every month

A real production deployment, measured on the solution's own cost dashboard. Historical data left premium Dataverse storage - and stayed fully accessible.

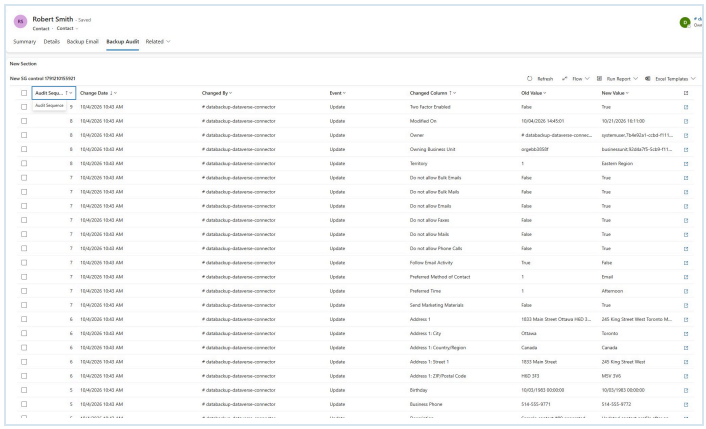
$4,733.18 /month
Estimated cost gain - about $56,798 per year

121.76 GB
Total data archived out of Dataverse (124,678 MB)

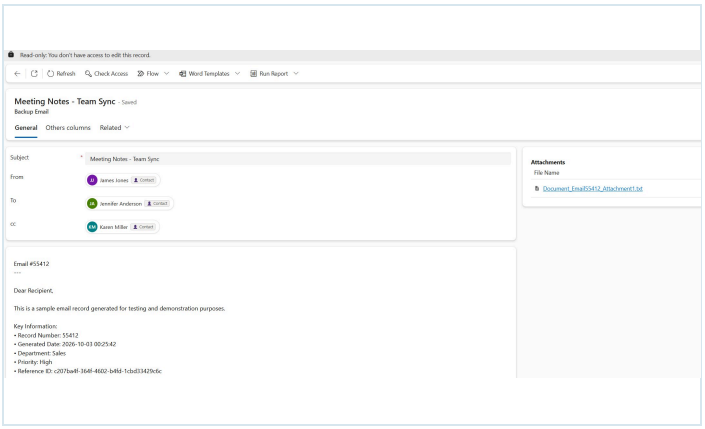
63,918,498
Records copied in 2 completed export commands



The cost dashboard. Table data is valued at $40/GB and audit data at $10/GB. Both export commands show Completed status.



Audit history, back in context. The Backup Audit tab on a contact record shows the full archived change history - changed column, old value, new value, who and when - inside the model-driven app.



Archived email, opening natively. A Backup Email record opens in the same app with sender, recipients, full body and its attachment. Read-only, as archived data should be.

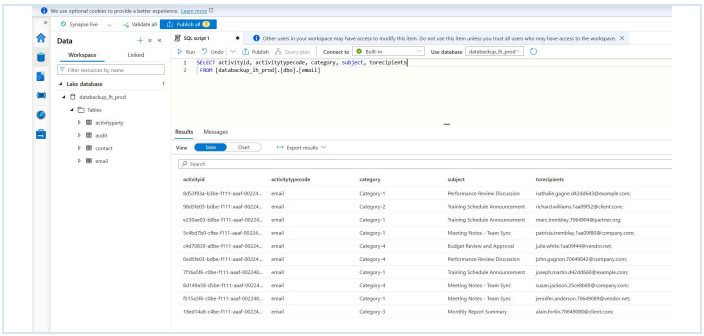
ClickDimensions Tables - 60,000,000 records • 120,000 MB • **$4,687.50/month** at $40/GB

Audit-2018 - 3,918,498 records • 4,678 MB • **$45.68/month** at $10/GB

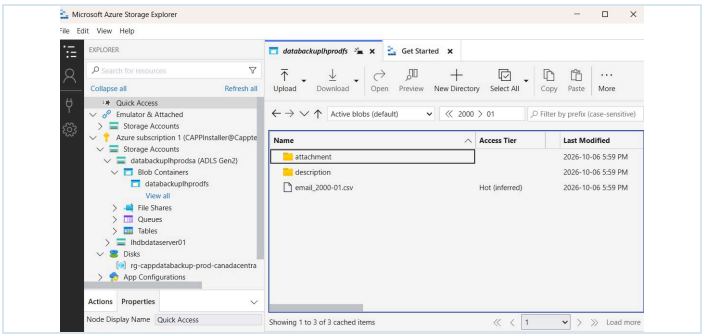
SHOWCASE - EVERY ACCESS PATH, WORKING

Query it, browse it, call it

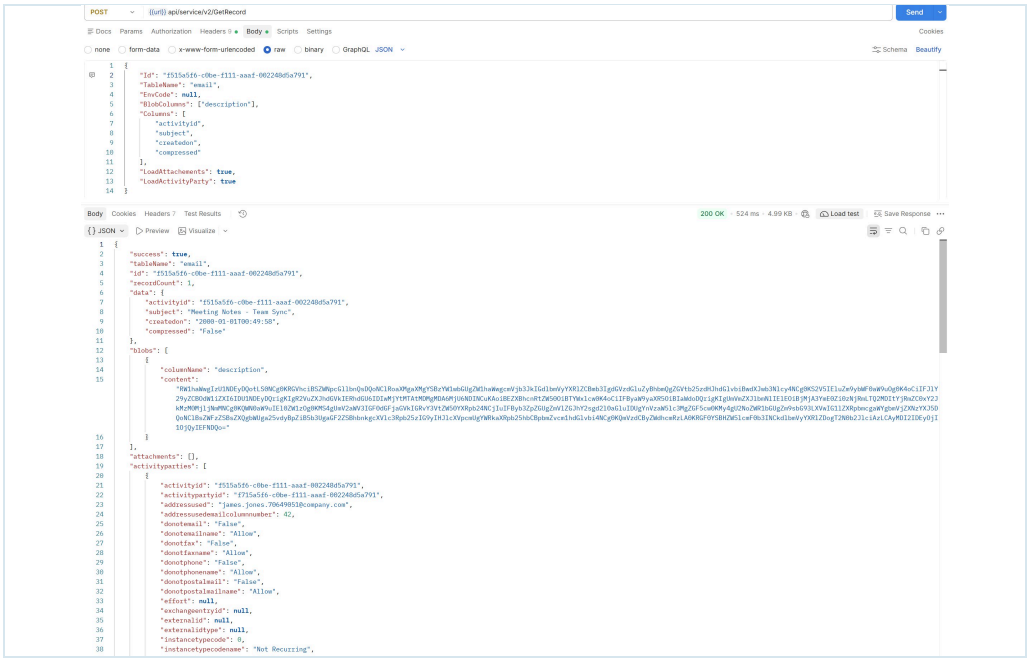
The same archive serves analysts, storage teams and developers - no restore step, no separate portal.



SQL over the archive. Synapse Studio queries the archived email table directly in the lake database, alongside audit and contact - results in seconds, in standard SQL.



Open formats, your storage. In Azure Storage Explorer, the archive sits in Data Lake Storage Gen2 as monthly files with attachment and description folders - no proprietary lock-in.



REST API, for developers. A POST to api/service/v2/GetRecord returns the archived email record as JSON - data, blob content, attachments and activity parties - with a 200 OK response.

THREE WAYS BACK TO THE DATA

1. Power Apps - Virtual Tables for everyday users, in the apps they already know.

2. Synapse SQL - Direct SQL queries over the lake for reporting and analysis.

3. REST API - Programmatic record retrieval, including attachments, for integrations.

Reduce Dataverse storage costs - without losing access to your history.

CAPP Techs Inc. • info@capptechs.com • https://capptechs.com

Full story: https://www.linkedin.com/pulse/dataverse-data-backup-archival-solution-capptechs-inc-g6m4e