

Clase 5. Bootcamp Bases de Datos en la Nube

# Almacenamiento de datos no relacionales con Azure Storage

Luis Beltrán  
Microsoft MVP



# Hola soy...

¡Hola a todos! Antes de sumergirnos en esta presentación, quiero tomarme un momento para hablar un poco sobre mí. [Soy Luis Beltrán](#), y he trabajado en algunos proyectos tales como desarrollo de aplicaciones móviles, soluciones informáticas de Inteligencia Artificial alojadas en la nube, bases de datos, consultoría de TI y capacitaciones oficiales de certificación de Microsoft.

Mi pasión por las bases de datos es contagiosa, así que prepárense para unirnos en un emocionante bootcamp juntos.



# Objetivos de la clase/Intro

Hoy nos adentraremos de forma teórica y práctica en el servicio Azure Storage para el almacenamiento de información de tipo no relacional. Describiremos las características y las funcionalidades de Azure Blob Storage, Azure Data Lake Gen2, Azure File Storage y Azure Table Storage. De forma práctica, aprovisionaremos y usaremos una cuenta de Azure Storage.





# Tabla de contenido

Clase #5

/01

## Azure Storage

Breve introducción

/02

## Azure Blob Storage

Almacenamiento y acceso de datos no estructurados a escala masiva en blobs en bloques.

/03

## Azure Data Lake Storage Gen 2

Soluciones de análisis de macrodatos empresariales.

/04

## Microsoft OneLake en Fabric

OneDrive para los datos



# Tabla de contenido

Clase #5

/05

## Azure Files

Recursos compartidos de archivos en la nube totalmente administrados.

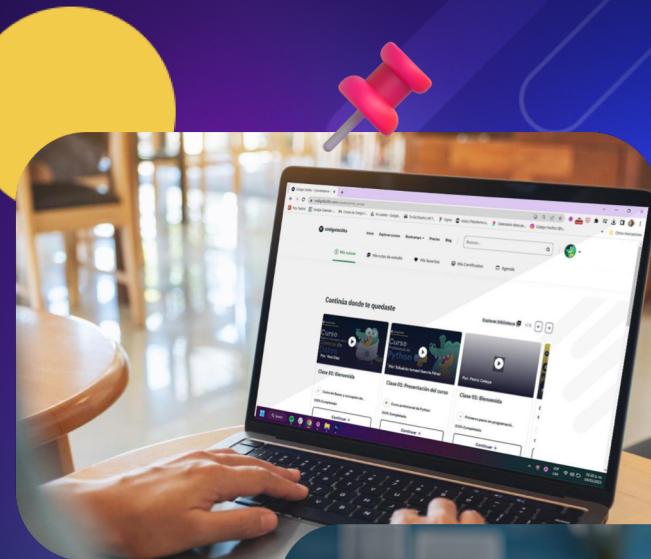
/06

## Azure Tables

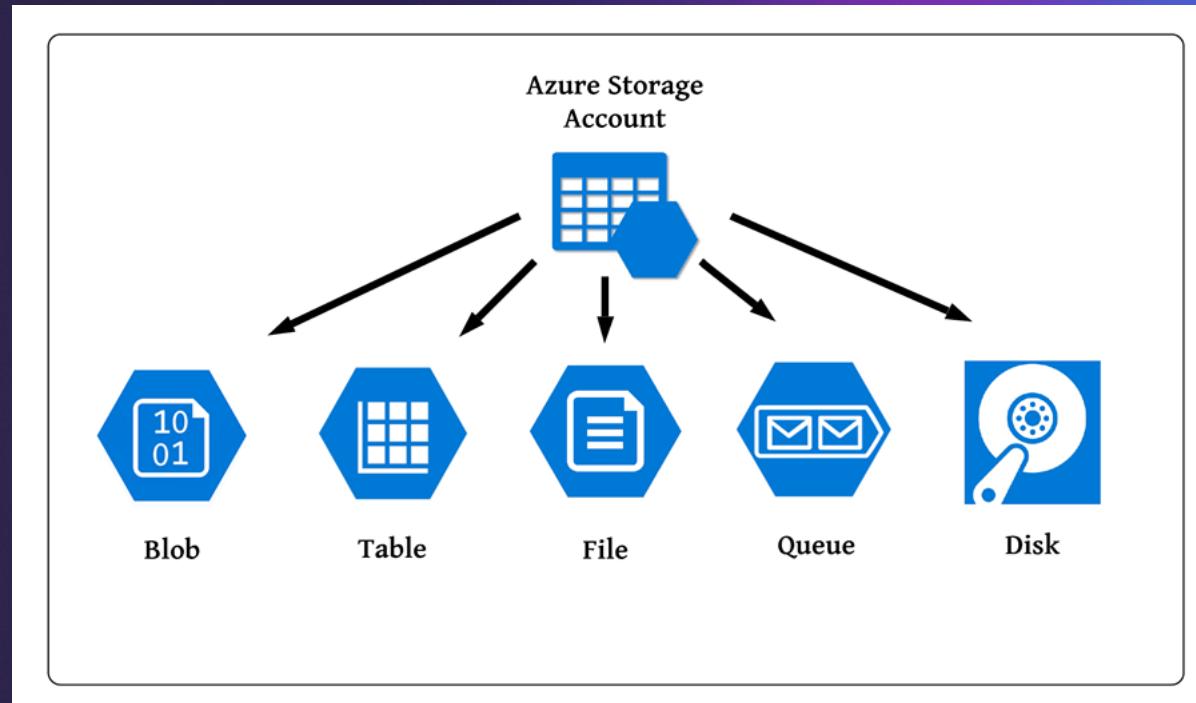
Almacenamiento de datos NoSQL estructurados en la nube.

# /01 Azure Storage

Introducción



# Introducción



# Azure Storage

- Una plataforma solución de almacenamiento en la nube de Microsoft para los escenarios modernos de almacenamiento de datos: almacenamiento de alta disponibilidad, escalable y seguro para diversos objetos de datos en la nube.
- Se puede acceder a los objetos de datos de Azure Storage mediante:
  - HTTP y HTTPS
  - API REST
  - Bibliotecas cliente (.NET, Java, Python, JavaScript, C++, Go)
  - Azure PowerShell y CLI de Azure
  - Azure Portal y Explorador de Azure Storage

# Servicios de datos en Azure Storage

- Azure Blobs
- Azure Files
- Azure Elastic SAN
- Azure Queues
- Azure Tables
- Azure Managed Disks
- Azure Container Storage (versión preliminar)

Se accede a cada servicio a través de una cuenta de almacenamiento con una dirección única.



códigofacilito

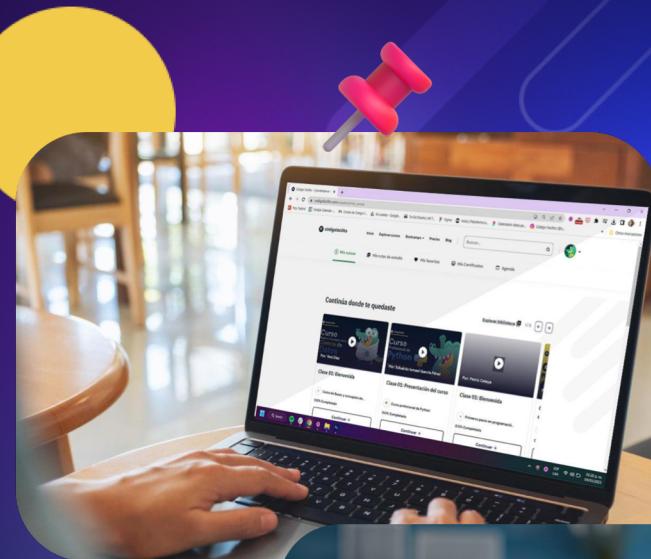
## Creación de una Cuenta de Azure Storage



/02

# Azure Blob Storage

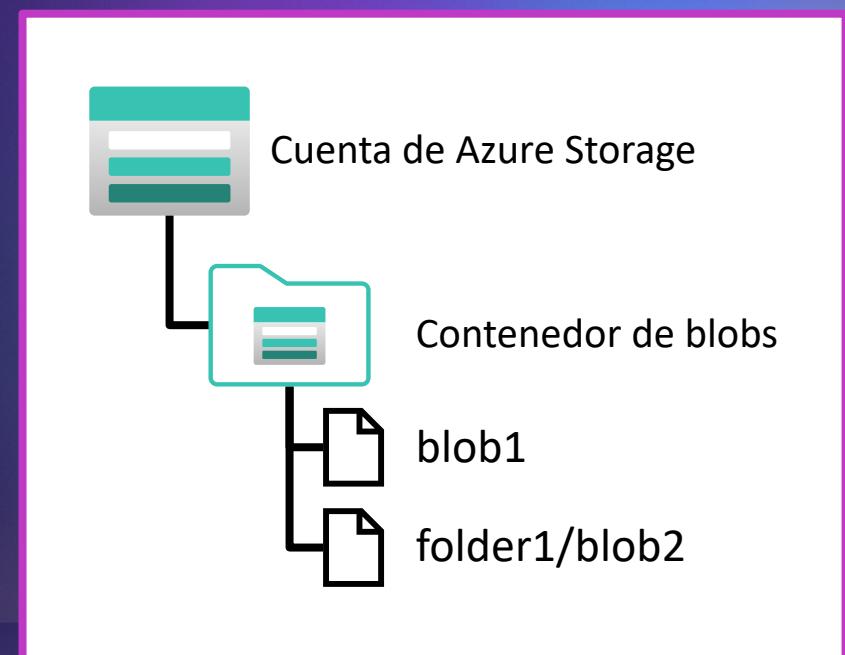
Almacenamiento para datos como  
BLOB



# Azure Blob Storage

Solución de almacenamiento para datos como objetos binarios grandes (BLOB)

Los blobs se pueden organizar en directorios virtuales, pero cada ruta de acceso se considera un único blob en un espacio de nombres plano: no se admiten operaciones de nivel de carpeta.



# Azure Blob Storage

Solución de almacenamiento para datos como objetos binarios grandes (BLOB)

## Blobs en bloques (Block blob)

- Objetos binarios grandes y discretos que cambian con poca frecuencia.
- Los blobs pueden ser de hasta 4,7 TB, compuestos por bloques de hasta 100 MB.
- Un blob puede contener hasta 50 000 bloques.

## Blobs en páginas (Page blob)

- Se usa como almacenamiento en disco virtual para máquinas virtuales.
- Los blobs pueden ser de hasta 8 TB, compuestos por páginas de 512 bytes de tamaño fijo.

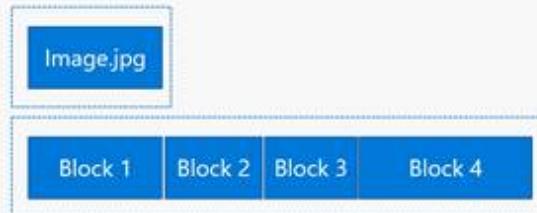
## Blobs en anexos (append blob)

- Blobs en bloques que se usan para optimizar las operaciones de anexión.
- Tamaño máximo de poco más de 195 GB, cada bloque puede ser de hasta 4 MB.

# Types of Blobs

## Block Blobs

Most object storage scenarios  
Documents, images, video, etc.



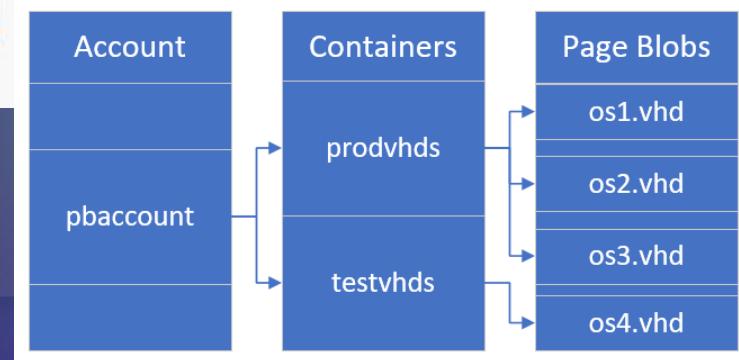
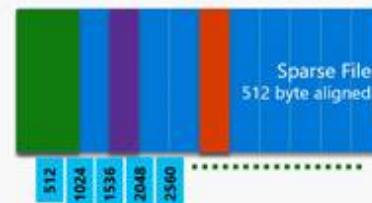
## Append Blobs

Multi-writer append only scenarios  
Logging, Big Data Analytics output



## Page Blobs

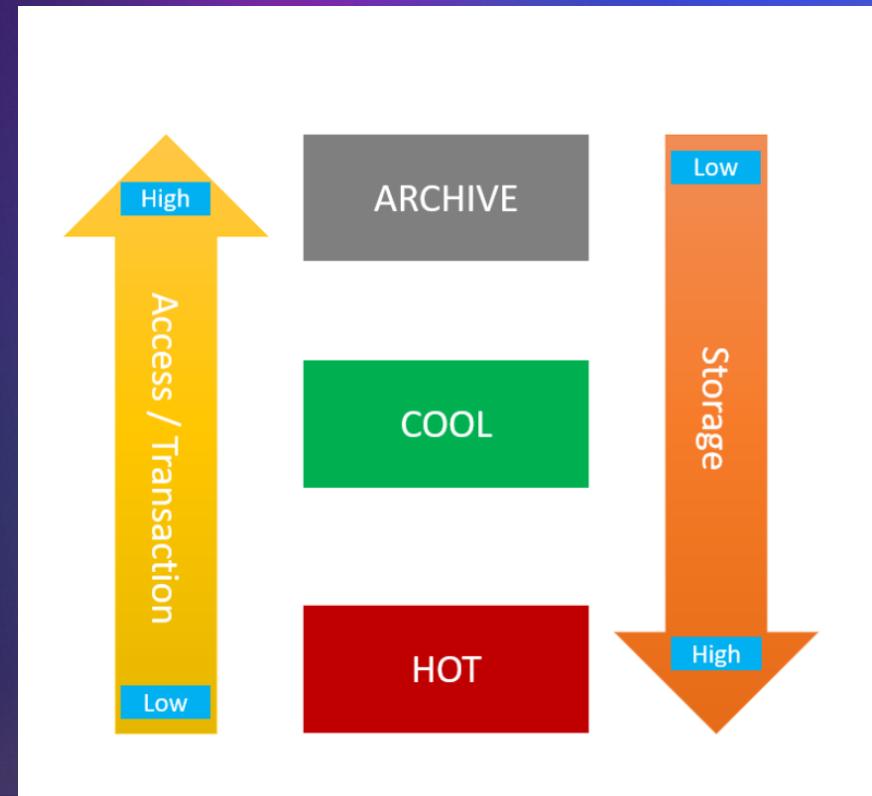
Page aligned random reads and writes  
IaaS Disks, Event Hub, Block level backup



# Azure Blob Storage

## Niveles de acceso por blobs

- Frecuente (Hot): el mayor costo de almacenamiento, el menor costo de acceso, la latencia más baja.
- Esporádico (Cool): menos costo de almacenamiento, más costo de acceso, más latencia.
- Archivo (Archive): el menor costo de almacenamiento, el mayor costo de acceso, la latencia más alta.





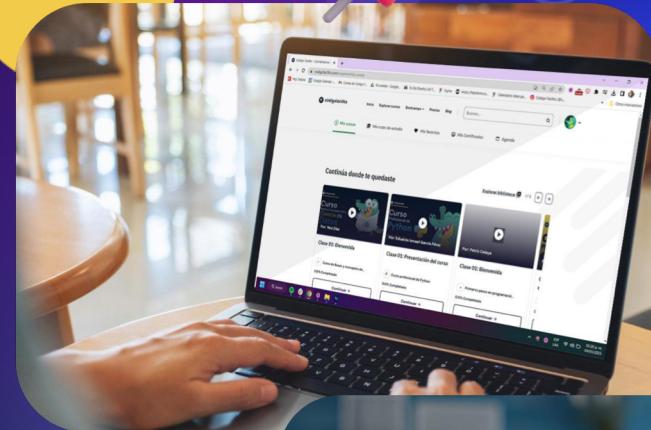
## Demo Azure Blob Storage



/03

# Azure Data Lake Storage Gen 2

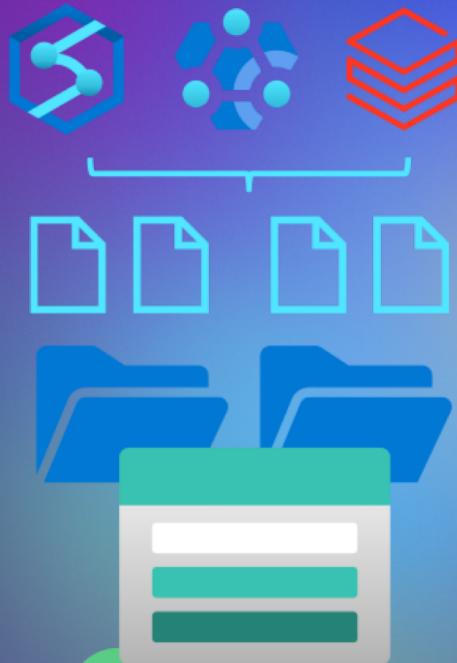
Sistema de archivos distribuido  
basado en Blob Storage



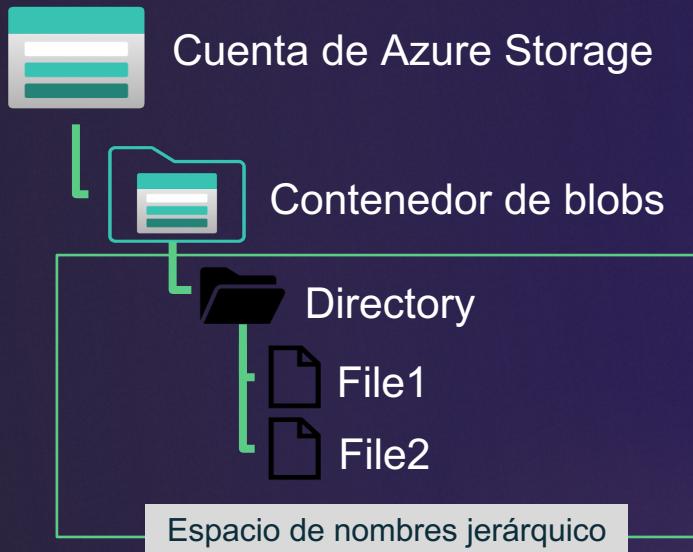
# Azure Data Lake Store Gen 2

Sistema de archivos distribuido basado en Blob Storage.

- Combina Azure Data Lake Store Gen1 con Azure Blob Storage para el almacenamiento y el análisis de archivos a gran escala.
- Habilita el control y la administración de acceso de nivel de archivo y directorio.
- Compatible con sistemas analíticos a gran escala comunes.



# Azure Data Lake Store Gen 2



**El sistema de archivos incluye directorios y archivos, y es compatible con sistemas de análisis de datos a gran escala, como Hadoop, Databricks y Azure Synapse Analytics.**

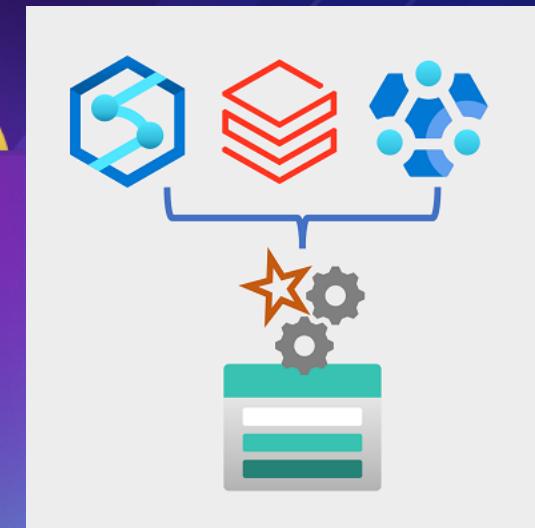
# Azure Data Lake Store Gen 2

Habilitado en una cuenta de Azure Storage mediante la opción:  
Espacio de nombres jerárquico

- Establecido durante la creación de la cuenta de almacenamiento.
- Se puede actualizar la cuenta de almacenamiento existente.
- Proceso de actualización unidireccional.

# Cargas de trabajo

- Procesamiento y análisis de macrodatos
- Almacenamiento de datos
- Análisis de datos en tiempo real
- Ciencia de datos y machine learning





## Demo Azure Data Lake Store Gen2



# Vamos a tomar un pequeño descanso

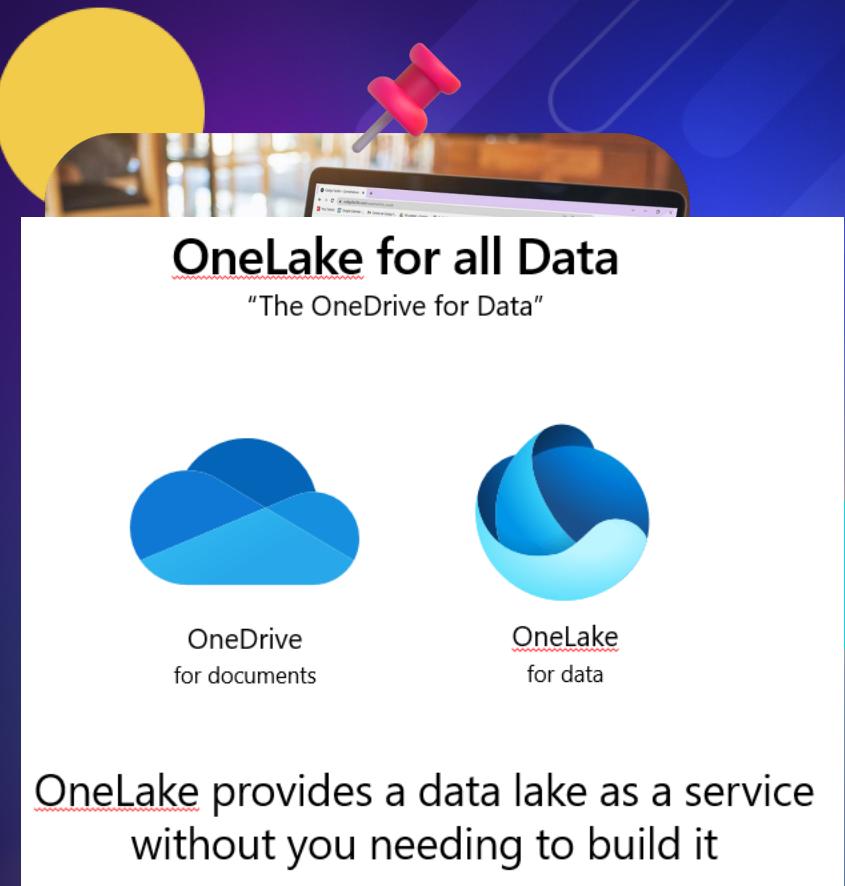


Break – 5 minutos

/04

# Microsoft OneLake en Fabric

OneDrive para los Datos



OneLake for all Data  
"The OneDrive for Data"

OneDrive for documents

OneLake for data

OneLake provides a data lake as a service without you needing to build it



# Microsoft Fabric

## The data platform for the era of AI



Data  
Factory



Synapse Data  
Engineering



Synapse Data  
Science



Synapse Data  
Warehousing



Synapse Real  
Time Analytics



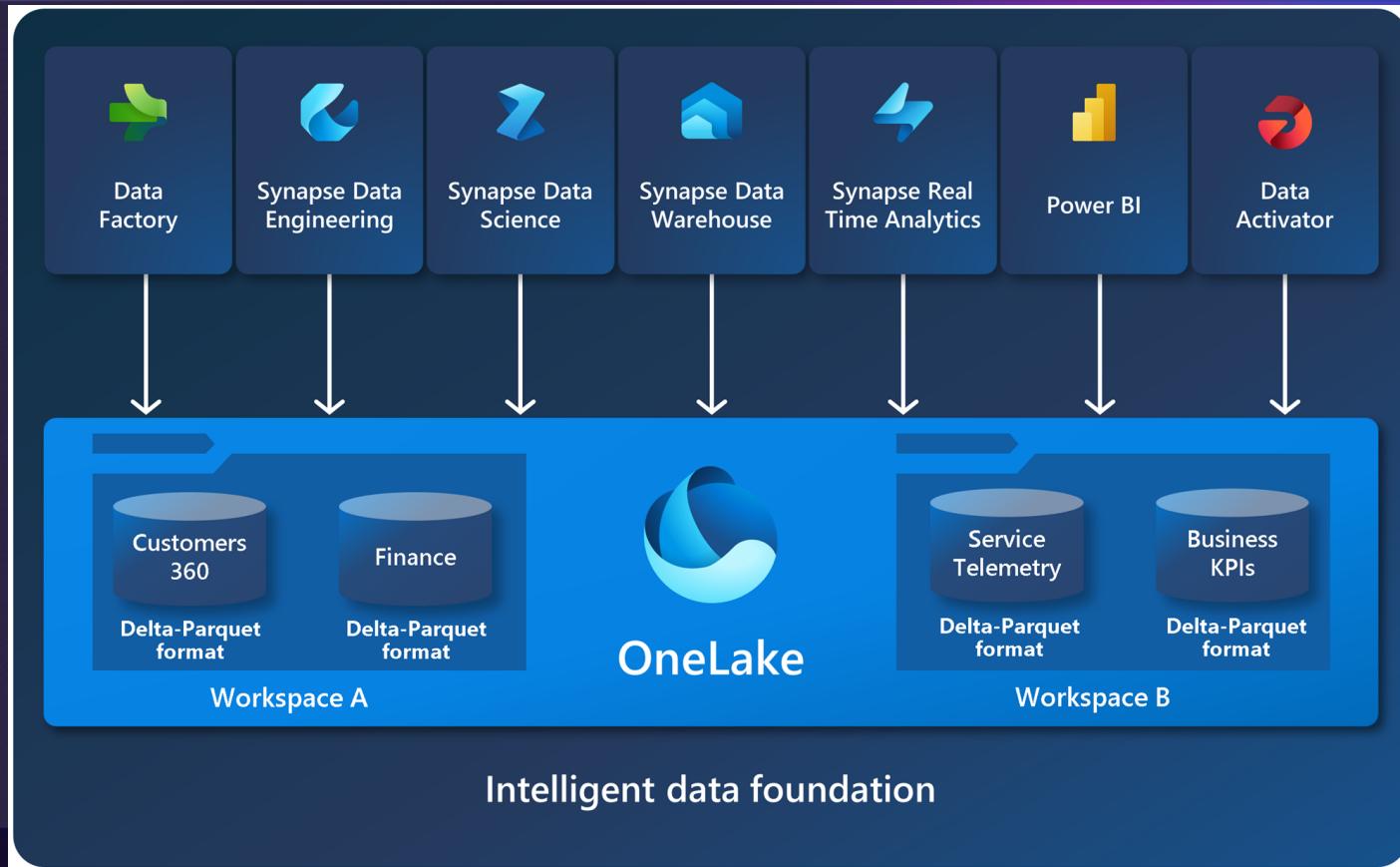
Power BI



Data  
Activator

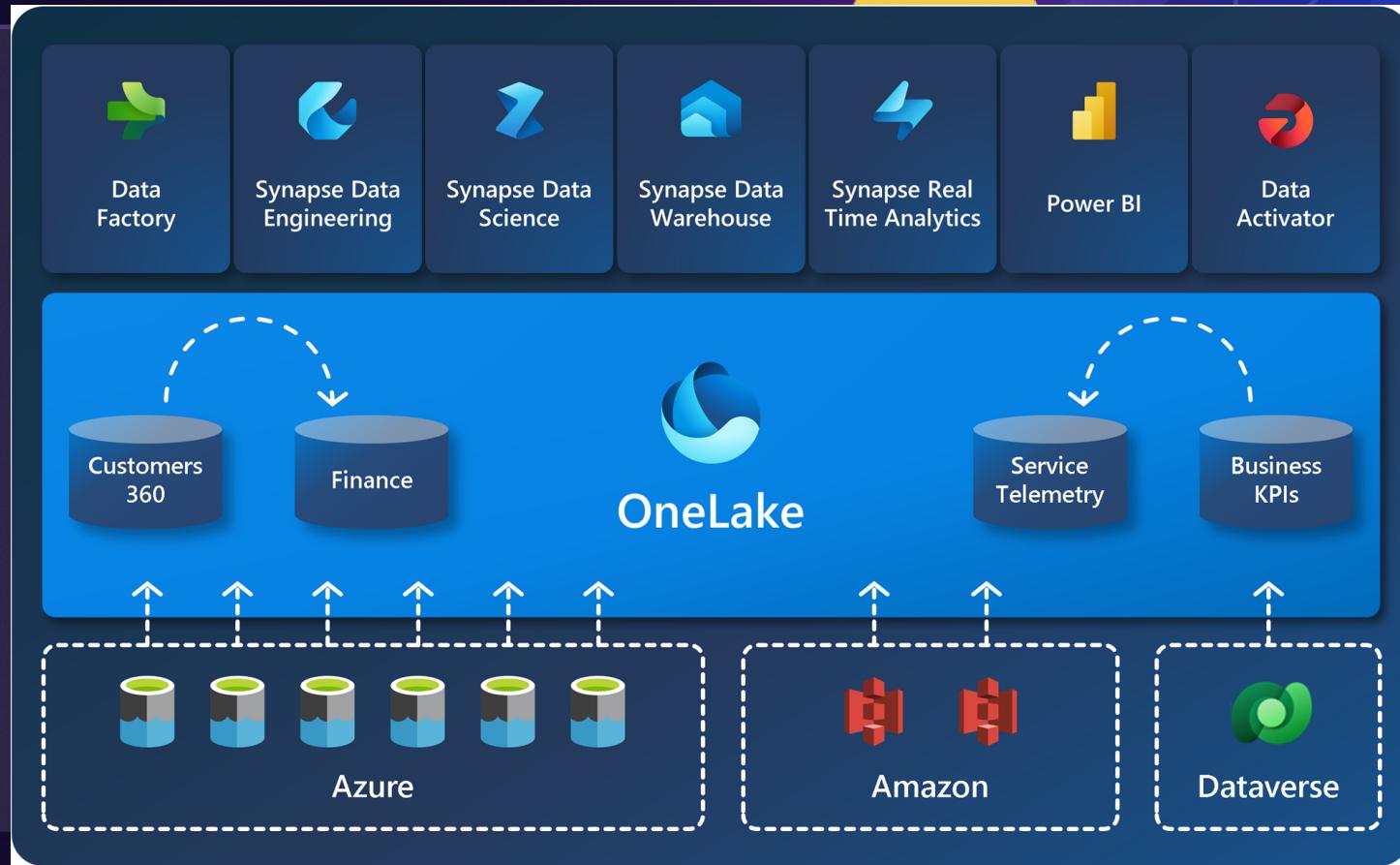


OneLake



# Microsoft Fabric – OneLake

- OneLake es un lago de datos único, unificado y lógico diseñado para toda su organización.
- OneLake viene automáticamente con todos los tenants de Microsoft Fabric y sirve como repositorio central para todos los datos de análisis.
- Ya sea estructurado o no estructurado, OneLake admite cualquier tipo de archivo y permite usar los mismos datos en varios motores analíticos sin movimiento de datos ni duplicación.



# Ventajas

- El lago de datos de toda la organización.
- Propiedad distribuida y colaboración.
- Abierto y compatible.
- Fácil de navegar.

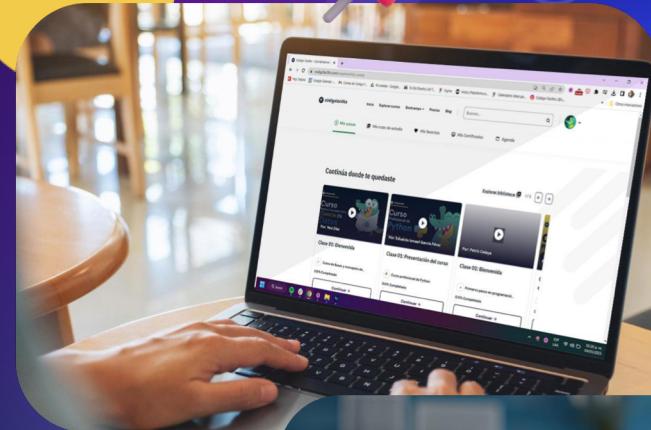


## Exploración de OneLake en Fabric



# /05 Azure Files

Recursos compartidos de archivos en la nube



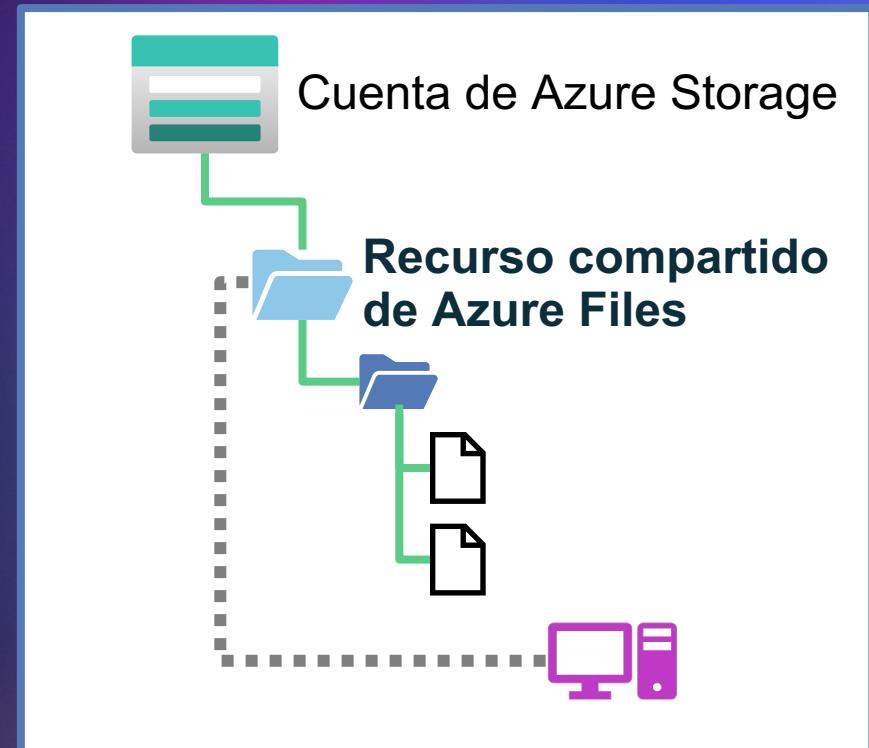
# Azure Files

**Recursos compartidos de archivos en la nube a los que se puede acceder desde cualquier lugar con una conexión a Internet.**

**Compatibilidad con protocolos comunes de uso compartido de archivos:**

- Bloque de mensajes del servidor (SMB)
- Network File System (NFS): requiere el nivel Premium

**Los datos se replican para la redundancia y el cifrado en reposo.**



# Azure Files

Azure File Storage ofrece dos niveles de rendimiento.

- Nivel Estándar → Usa hardware basado en disco duro
- Nivel Premium → Usa discos de estado sólido.

Azure Files admite dos protocolos comunes de uso compartido de archivos de red:

- SMB → Server Message Blocks (Windows, Linux, macOS)
- NFS → Network File System (algunas versiones de Linux y macOS)

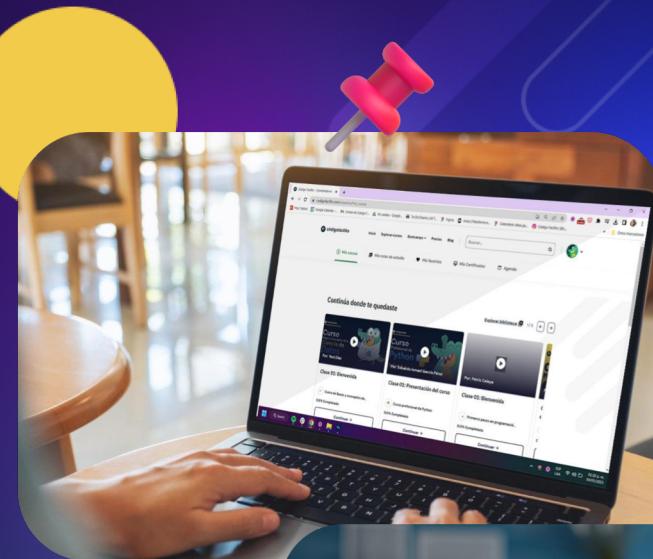


## Demo Azure Files



# /06 Azure Tables

Solución de almacenamiento  
NoSQL

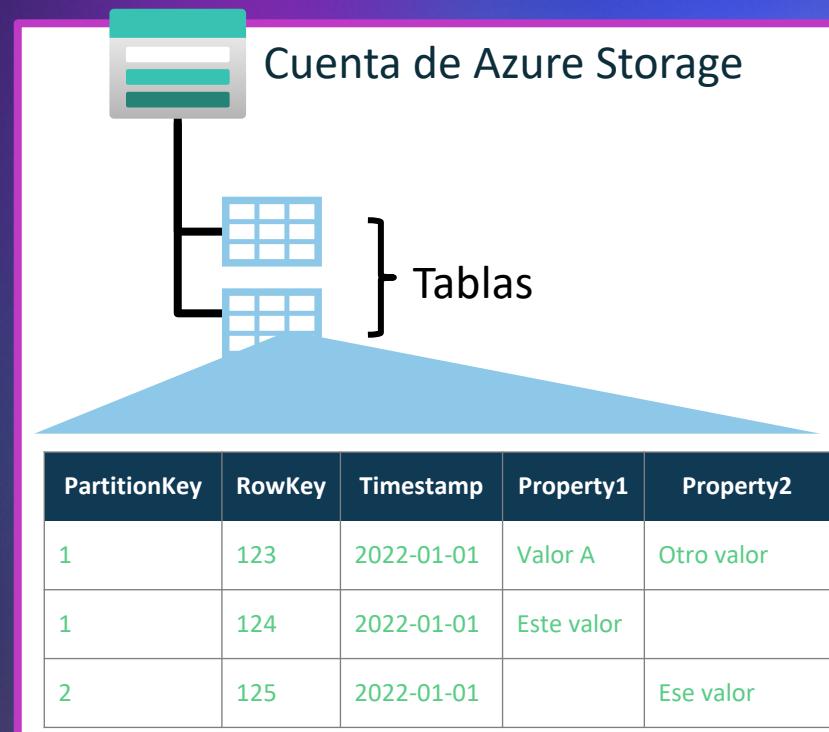


# Azure Tables

Almacenamiento de clave-valor para los datos de la aplicación

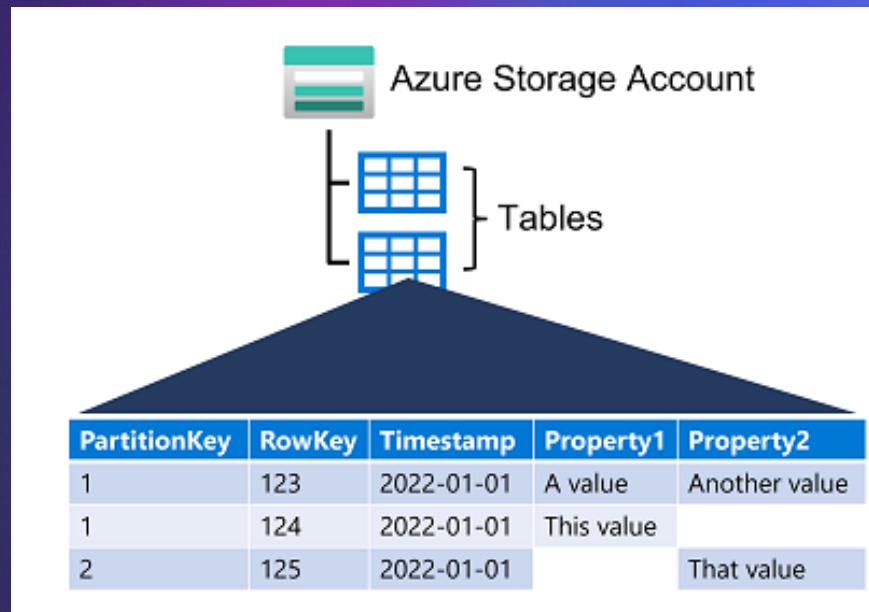
Las tablas constan de columnas de clave y valor

- Claves de partición y fila
- Columnas de propiedades personalizadas para valores de datos
- Se ha agregado automáticamente una columna Timestamp para registrar los cambios en los dato



# Azure Tables

- Las filas se agrupan en particiones para mejorar el rendimiento
- Se asignan columnas de propiedad a un tipo de datos y pueden contener cualquier valor de ese tipo.
- No es necesario que las filas incluyan las mismas columnas de propiedad





## Demo Azure Tables



1

¿Cuáles son los elementos de una clave de Azure Table Storage?

- A) Nombre de tabla y nombre de columna.
- B) Clave de partición y clave de fila.
- C) Número de fila.

2

¿Qué debe hacer con un inquilino de Microsoft Fabric existente para ser compatible con un lago de datos?

- A) Incorporación de un recurso compartido de Azure Files
- B) Creación de tablas de Azure Storage para los datos que quiere analizar
- C) No se requiere ninguna acción.

3

### ¿Para qué podría usar Azure File Storage?

- A) Para compartir archivos que se almacenan de forma local con usuarios ubicados en otros sitios.
- B) Para permitir que los usuarios de distintos sitios compartan archivos de red.
- C) Para almacenar archivos de datos binarios grandes que contienen imágenes u otros datos no estructurados.

4

### ¿Qué debe hacer en una cuenta de Azure Storage existente para admitir un lago de datos para Azure Synapse Analytics?

- A) Incorporación de un recurso compartido de Azure Files
- B) Creación de tablas de Azure Storage para los datos que quiere analizar
- C) Actualización de la cuenta para habilitar el espacio de nombres jerárquico y crear un contenedor de blobs

# ¿Qué has aprendido hoy?

Azure Storage es un servicio clave de Microsoft Azure y permite una amplia gama de soluciones y escenarios de almacenamiento de datos.



Describir las características y las funcionalidades de Azure Blob Storage, Azure Data Lake Gen2, Microsoft OneLake, Azure File Storage y Azure Table Storage

Aprovisionar y usar una cuenta de Azure Storage

# ¿Cuándo utilizar cada servicio?

## Azure Blob Storage

- Copia de seguridad o archivado de datos relacionados con aplicaciones y VMs
- Alojamiento y transmisión de imágenes, archivos de audio y vídeos
- Análisis de big data
- Almacenamiento en caché y distribución de contenido para sitios web
- Almacenamiento de objetos de archivos de registro y datos de telemetría para fines de supervisión, depuración o análisis

## Azure File Storage

- Almacenamiento de archivos compartidos para la colaboración en un área dispersa geográficamente
- Transferencia a la nube de aplicaciones que utilizan API nativas y SMB
- Reemplazo o complemento de servidores de archivos locales
- Agregación de registros y datos para procesamiento y análisis centralizados
- Gestión central de archivos de configuración para múltiples instancias de aplicaciones
- Almacenamiento de archivos para entornos de desarrollo y prueba

# ¿Cuándo utilizar cada servicio?

## Azure Table Storage

- Almacenamiento centralizado de registros, datos de telemetría y datos de monitoreo
- Almacenamiento de datos de catálogos y carritos de compras para aplicaciones de comercio electrónico
- Programación de tareas escalable y almacenamiento de metadatos
- Almacenamiento de datos sensoriales y datos de telemetría de IoT
- Almacenamiento de datos de configuración para aplicaciones en la nube



# Directorio de links

Exploración de Azure Storage para datos no relacionales

<https://learn.microsoft.com/es-mx/training/modules/explore-provision-deploy-non-relational-data-services-azure/>

Exploración de Azure Storage

<https://aka.ms/dp900-storage-lab>



# ¡Gracias por tu atención!

Luis Beltrán  
Microsoft MVP  
[about.me/luis-beltran](http://about.me/luis-beltran)

