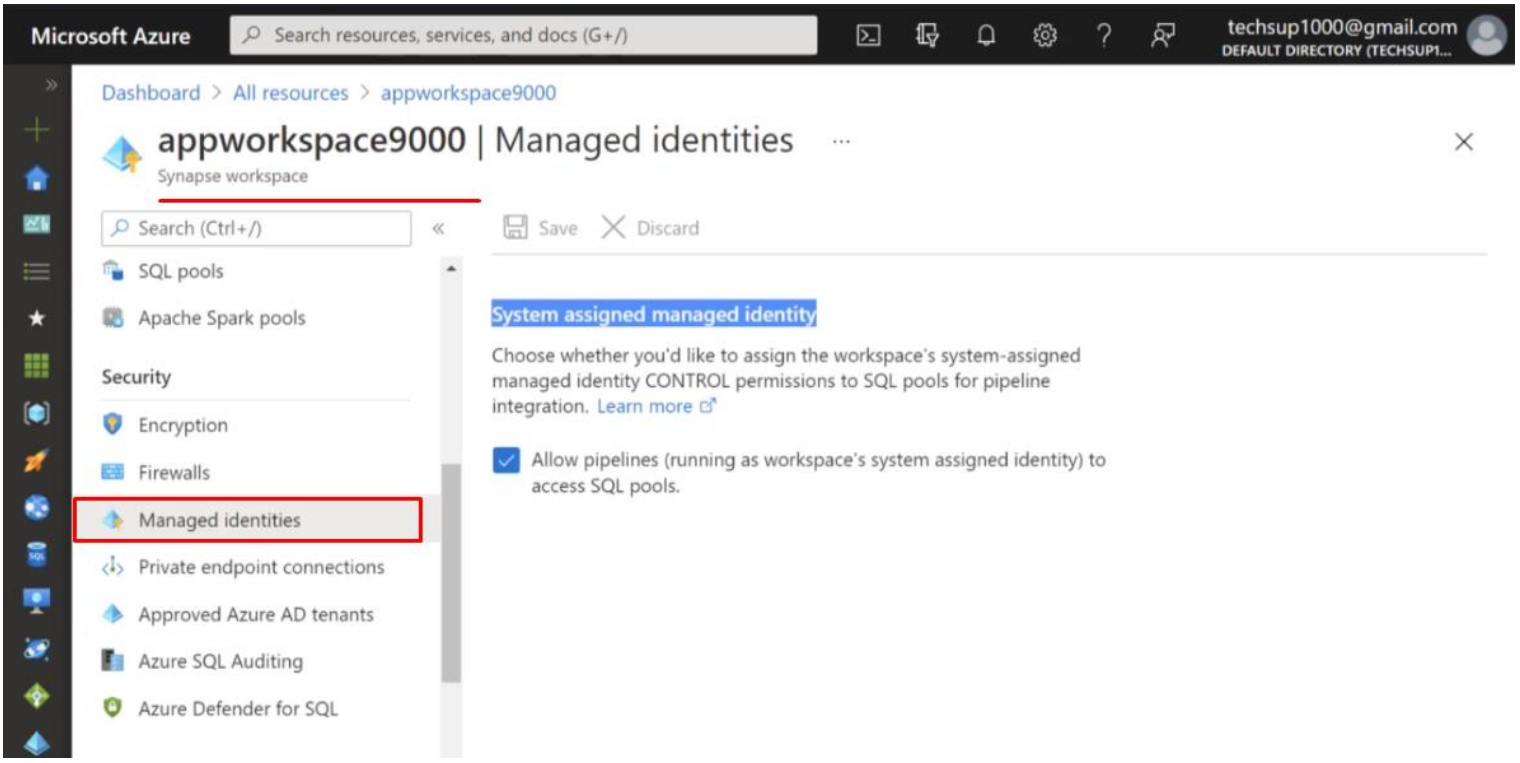
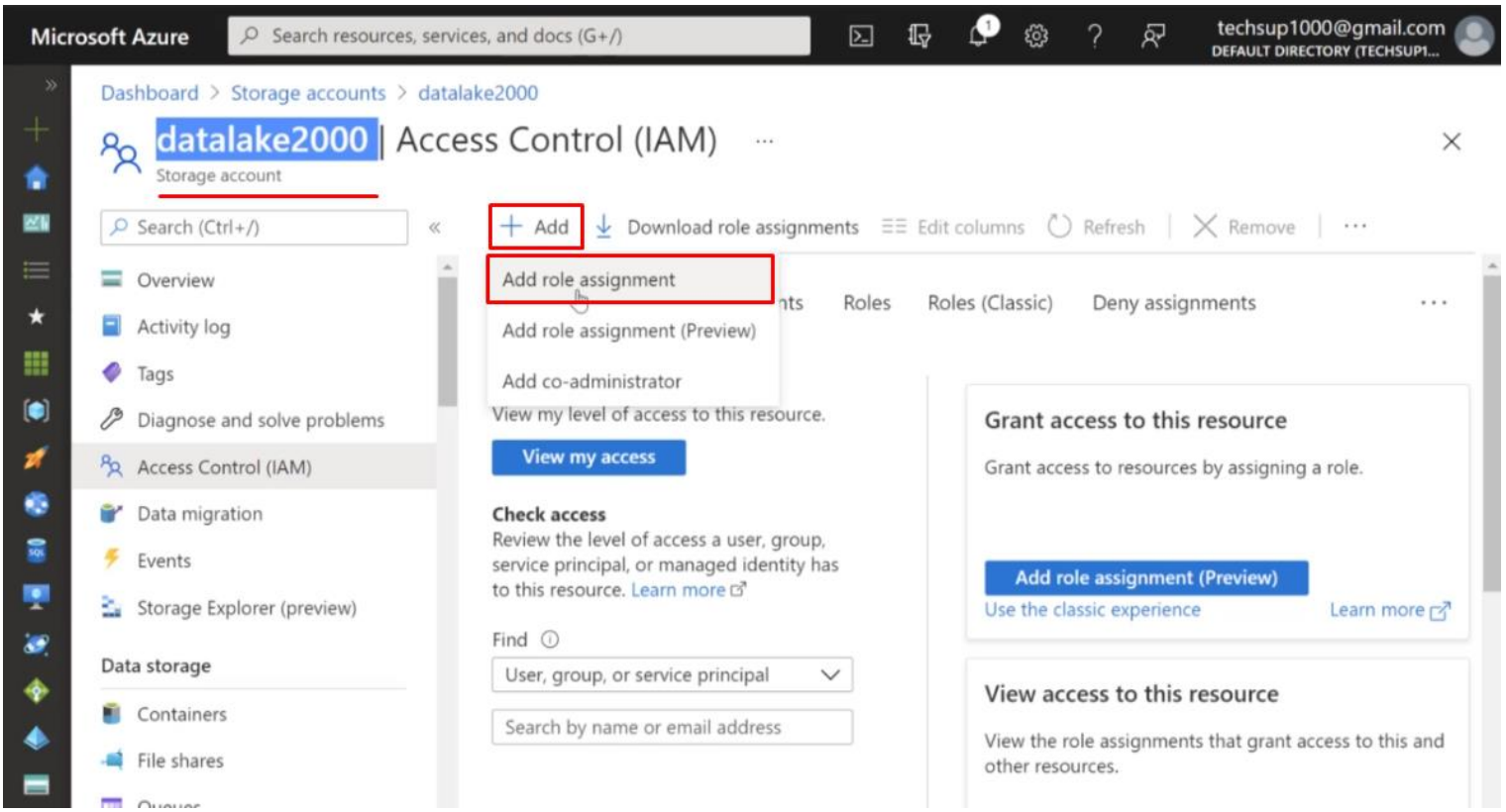
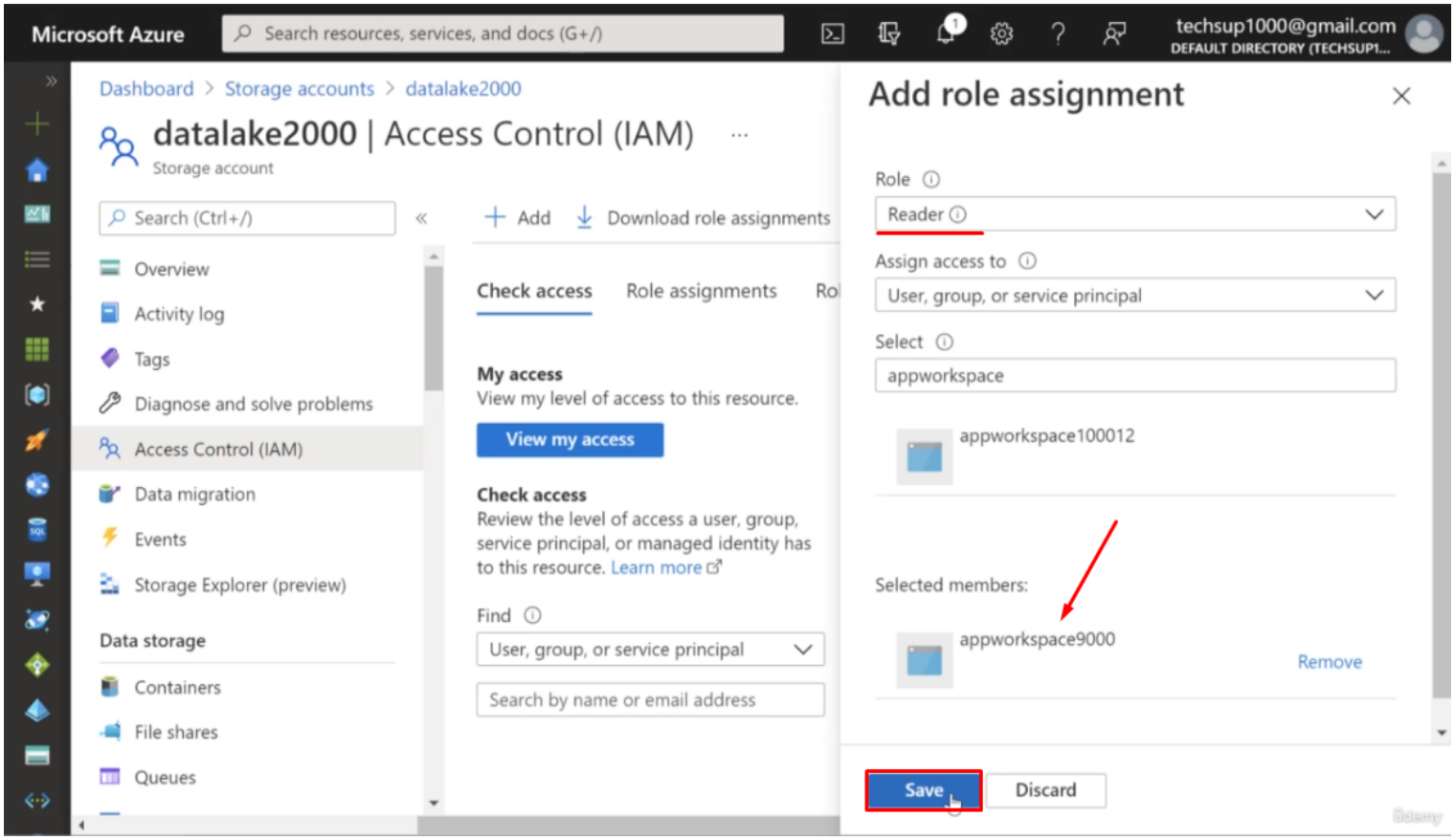
Azure Synapse External Tables Autorización via Managed Identity

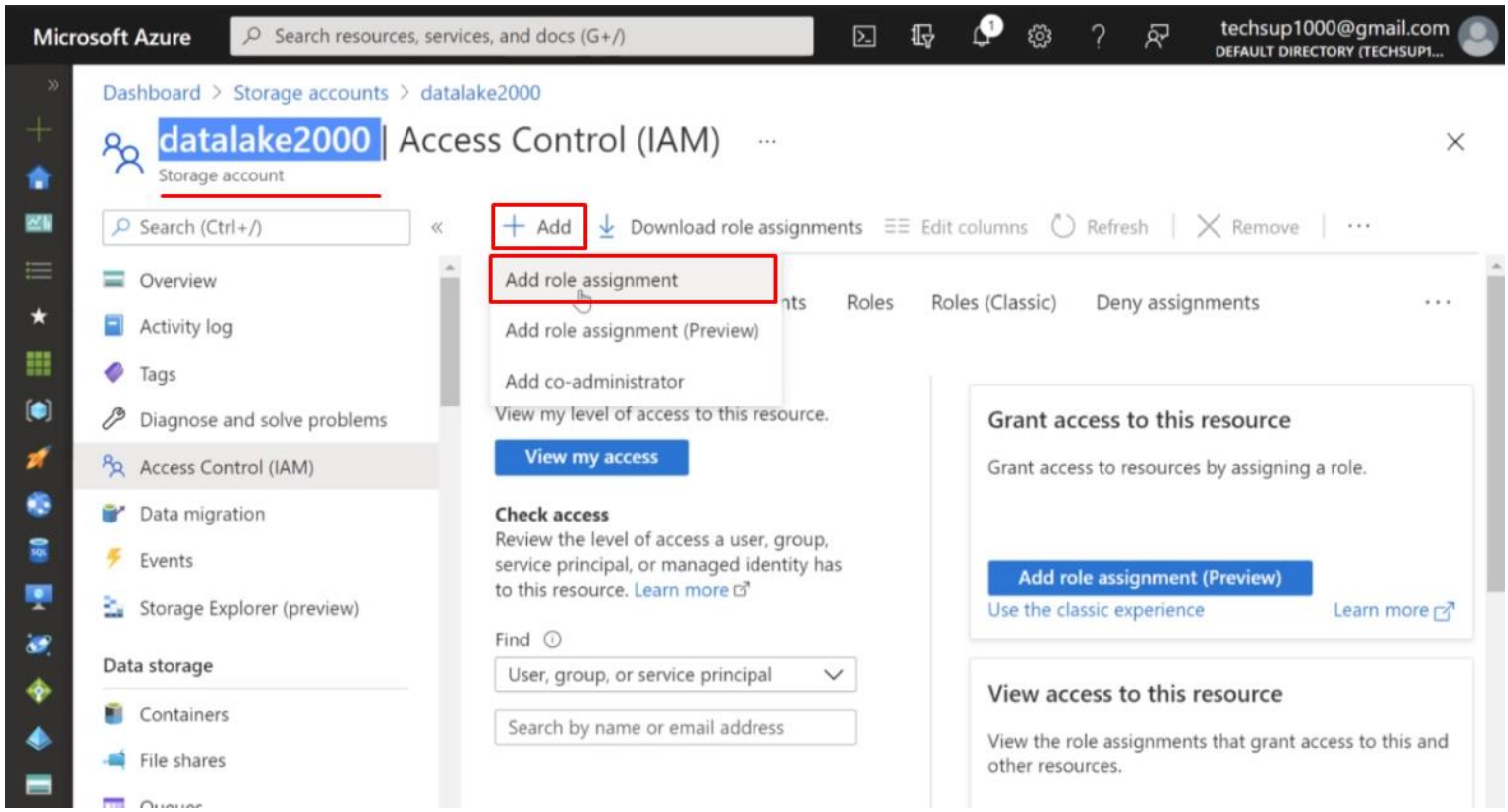
Ahora, en este capítulo, vamos a ver cómo crear una External table usando algo conocido como "Managed identity". Solo hemos visto como crear tablas externas cuando llegamos a la sección de Azure Synapse. Allí, creamos una tabla externa utilizando la "Access key" cuando se trata de autorización. Ahora, la "Access key" en si misma es como tener una contraseña y tener eso en el script no es un enfoque ideal. Así que hay algunas otras maneras en que podemos gestionar la autorización y aquí vamos a ver cómo lograr esto con la ayuda de la "Managed identity". Así que normalmente en un Azure, para un cierto número de recursos, en realidad se puede habilitar algo conocido como una "Managed identity ". Cuando se habilita una "Managed Identity", la identidad se crea en Azure Active Directory. Por ejemplo, si tienes una VM, una máquina virtual, si el nombre de esta VM es, digamos, 'demovm', y si habilitas la característica de "Managed identity" para esta máquina virtual, se creará una identidad con el nombre de 'demovm' en Azure Active Directory. Entonces podrás dar acceso a recursos como la cuenta de almacenamiento Azure Data Lake Gen2 a esa identidad. Así que esto ayuda a los recursos para acceder de forma segura a otro recurso basado en las identidades que están disponibles en Azure Active Directory. En nuestro script, vamos a crear una tabla externa basada en los datos del archivo "Log.csv", que tenemos en la cuenta de almacenamiento Azure Data Lake Gen2. Así que como dije antes, cuando se trataba de la autorización, estábamos buscando el uso de 'Access Keys' de nuevo. Esto es como tener un enfoque basado en contraseñas, porque al final, esto no es más que un secreto. Sí, hay maneras en que usted puede acceder a este secreto de una manera un poco más segura, entonces vamos a ver el enfoque de "Managed identity". Ahora, cuando se trata de Azure Synapse, la "Managed identity" está disponible para Azure Synapse, entonces vamos a dar acceso a nuestra cuenta de almacenamiento Azure Data Lake Gen2 como lo haríamos normalmente para cualquier otra identidad. Asi que recordemos que antes habíamos creado un usuario y para ese usuario habíamos dado acceso via "Role based access control" y tambien "Access control list", diciendo que ahora daremos acceso a Azure Synapse.

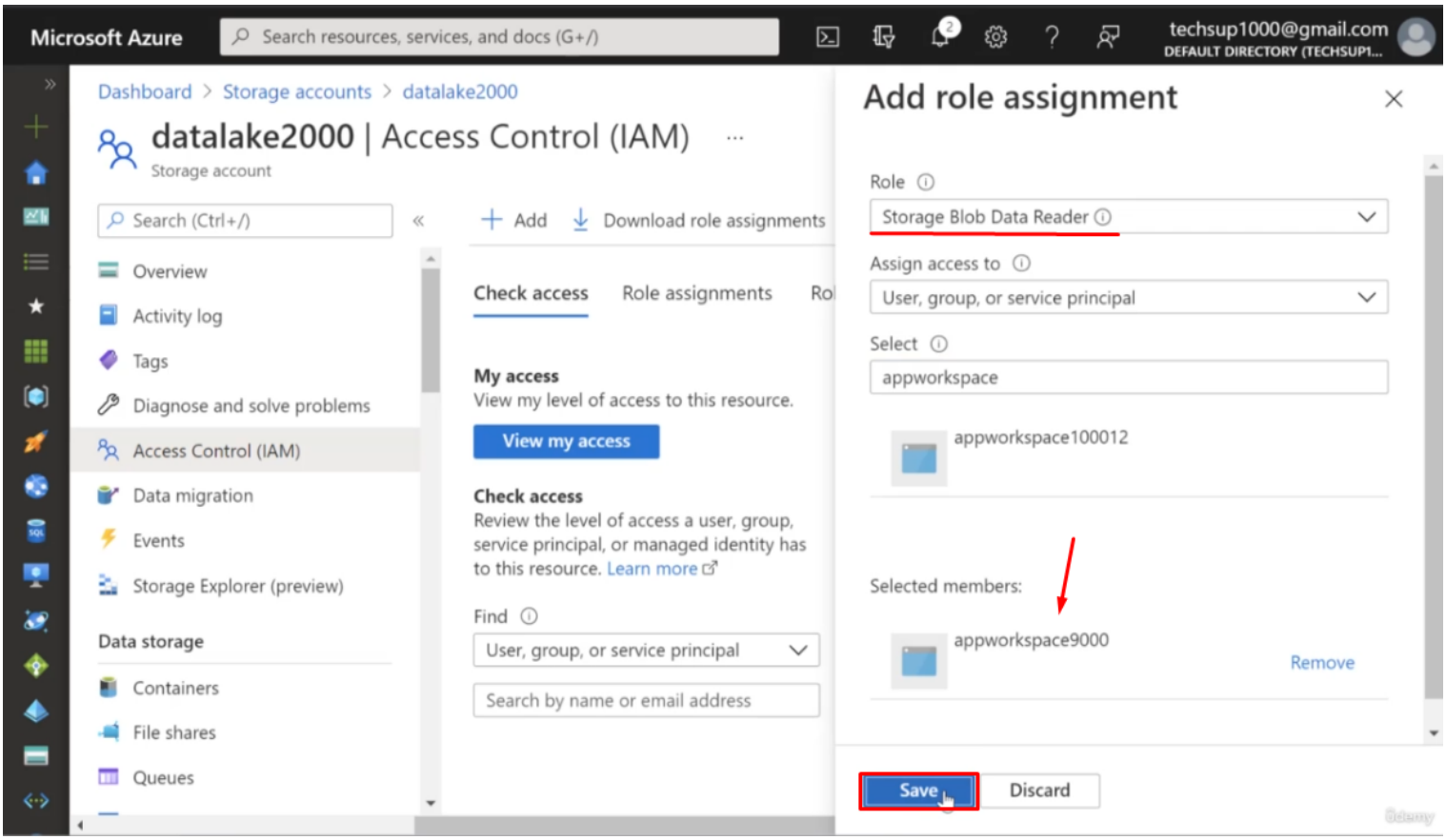

Microsoft Azure
Search resources, services, and docs (G+/)
techsup1000@gmail.com
DEFAULT DIRECTORY (TECHSUP1...
Dashboard > All resources > appworkspace9000
appworkspace9000 | Managed identities
Synapse workspace
Search (Ctrl+/)
Save
Discard
SQL pools
Apache Spark pools
Security
Encryption
Firewalls
Managed identities
Private endpoint connections
Approved Azure AD tenants
Azure SQL Auditing
Azure Defender for SQL
System assigned managed identity
Choose whether you'd like to assign the workspace's system-assigned managed identity CONTROL permissions to SQL pools for pipeline integration. Learn more
Allow pipelines (running as workspace's system assigned identity) to access SQL pools.


Microsoft Azure
Search resources, services, and docs (G+/)
techsup1000@gmail.com
DEFAULT DIRECTORY (TECHSUP1...
Dashboard > Storage accounts > datalake2000
datalake2000 | Access Control (IAM)
Storage account
Search (Ctrl+/)
Add
Download role assignments
Edit columns
Refresh
Remove
Overview
Activity log
Tags
Diagnose and solve problems
Access Control (IAM)
Data migration
Events
Storage Explorer (preview)
Data storage
Containers
File shares
Add role assignment
Add role assignment (Preview)
Add co-administrator
Roles
Roles (Classic)
Deny assignments
View my level of access to this resource.
View my access
Check access
Review the level of access a user, group, service principal, or managed identity has to this resource. Learn more
Find
User, group, or service principal
Search by name or email address
Grant access to this resource
Grant access to resources by assigning a role.
Add role assignment (Preview)
Use the classic experience
Learn more
View access to this resource
View the role assignments that grant access to this and other resources.

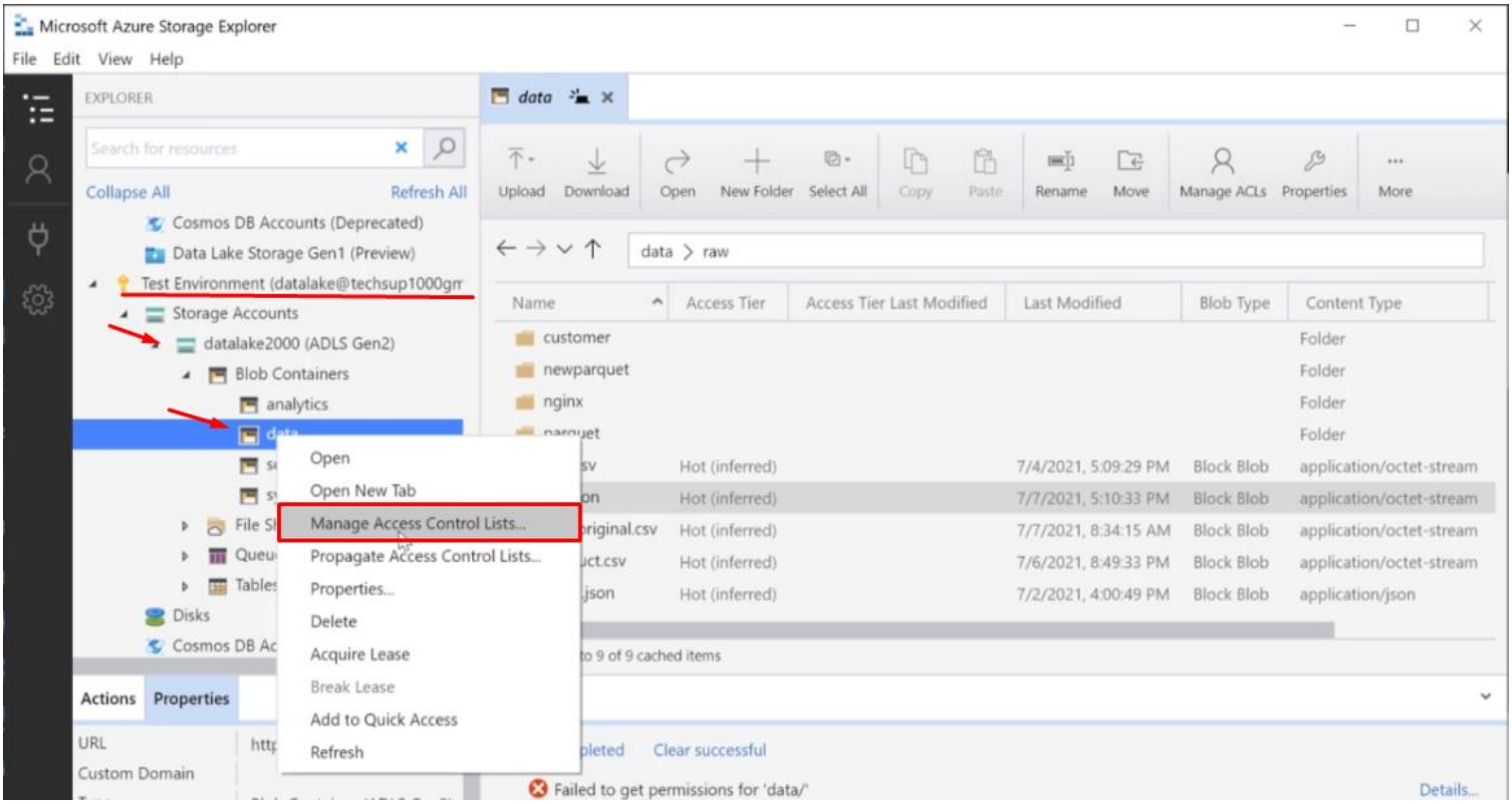
Damos acceso rol “Reader” al workspace de Azure Synapse. Así que podemos ver que también podemos elegir una **identidad** que se asocia a Azure Synapse workspace

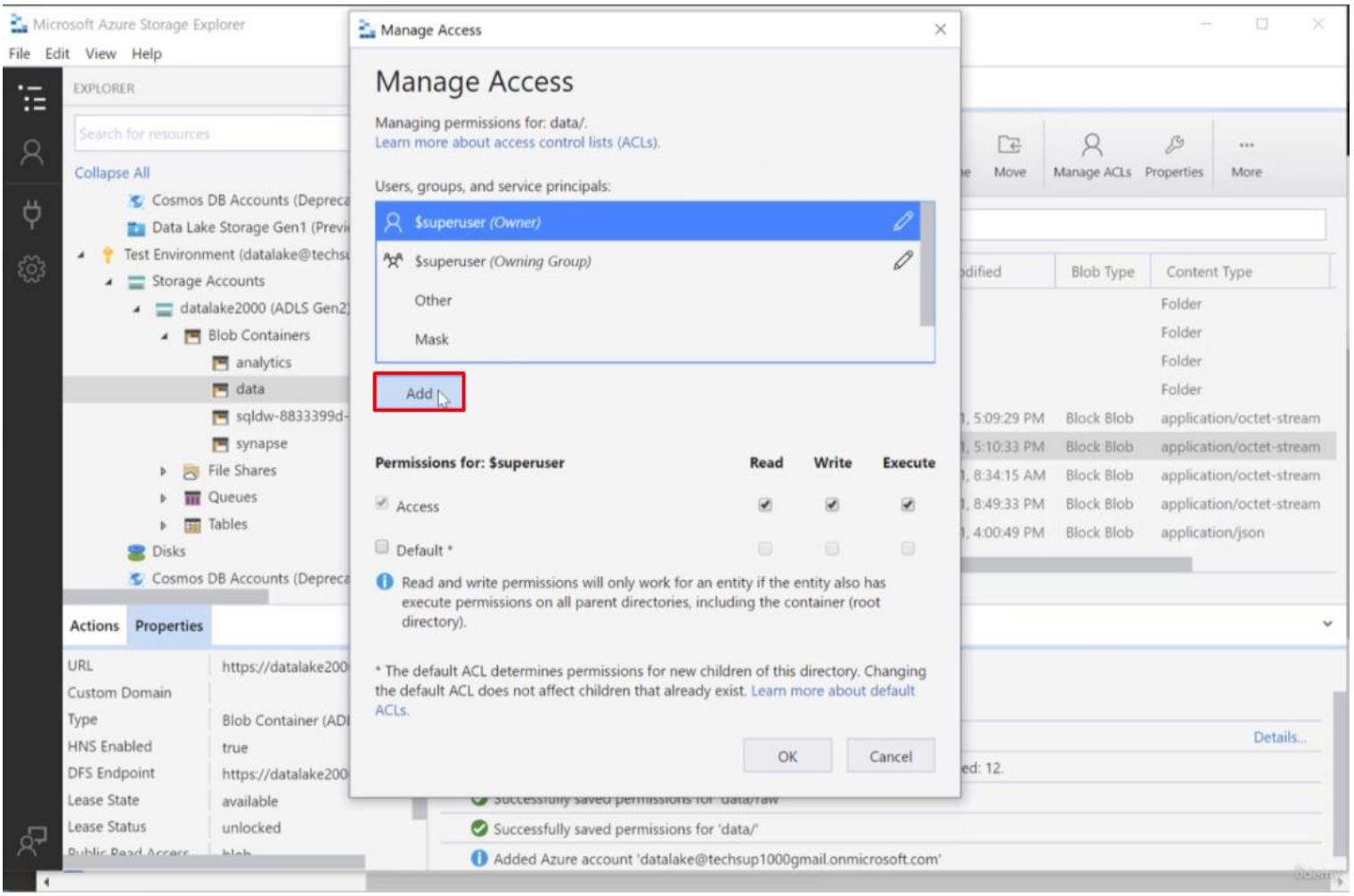


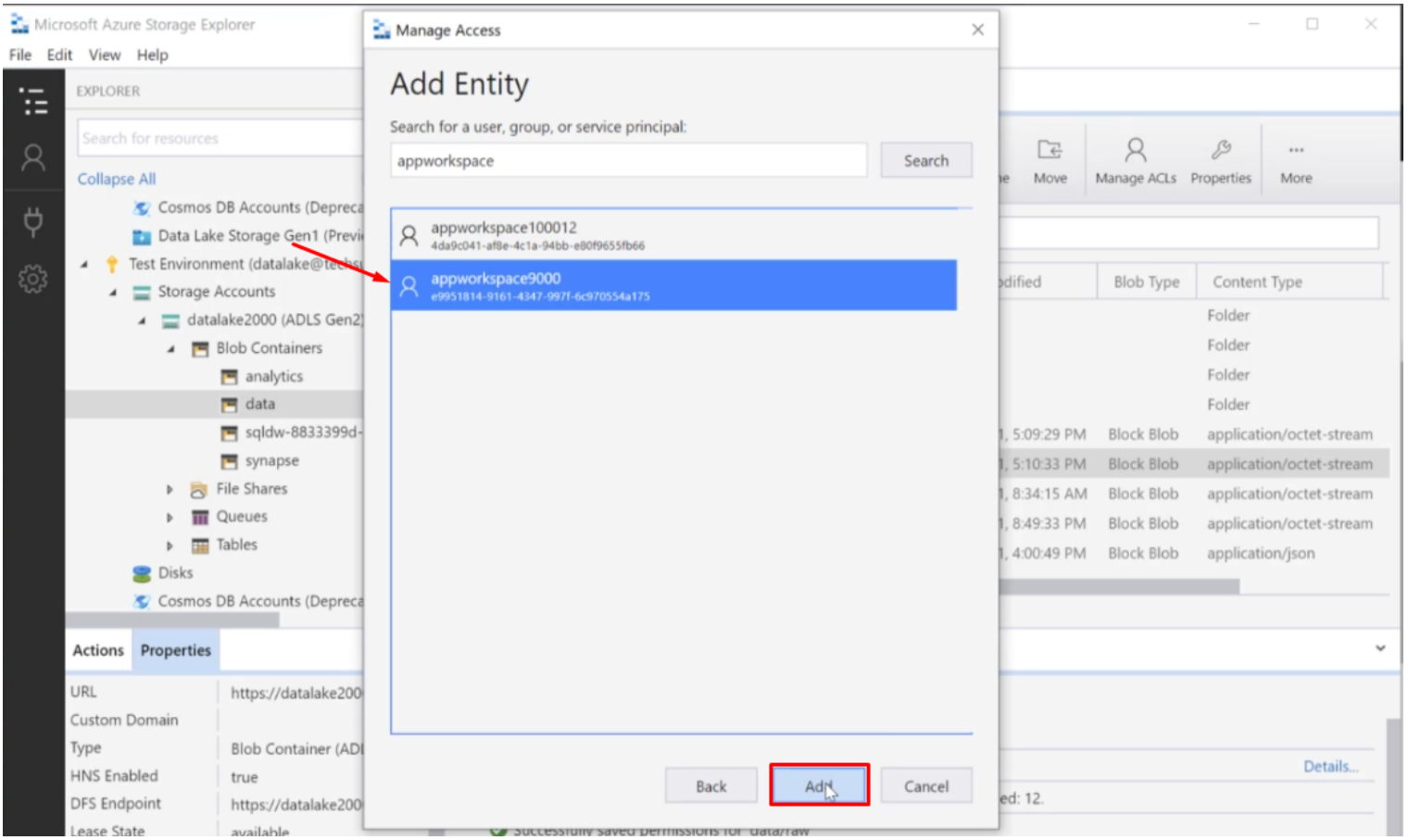


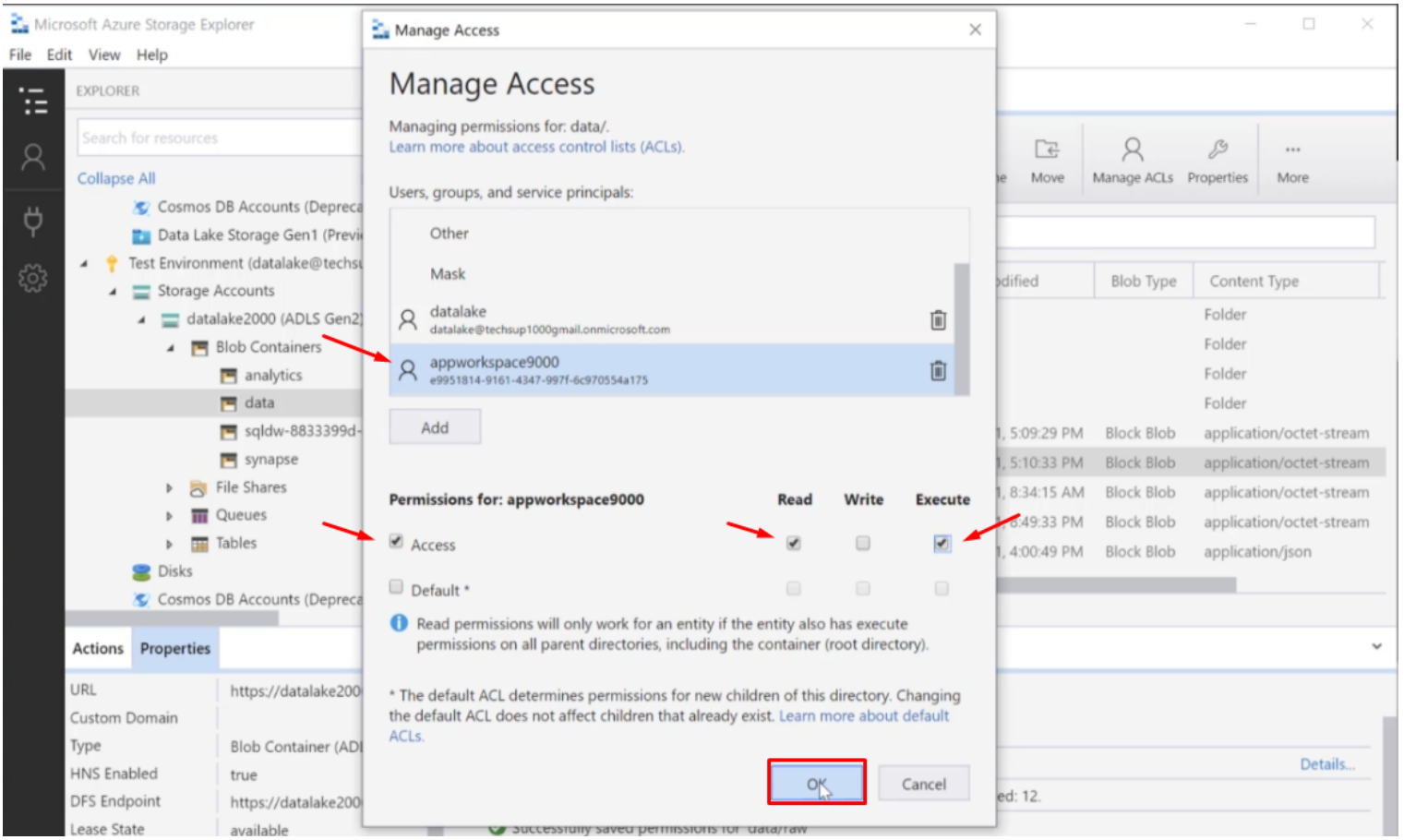


Ahora desde Azure Storage Explorer

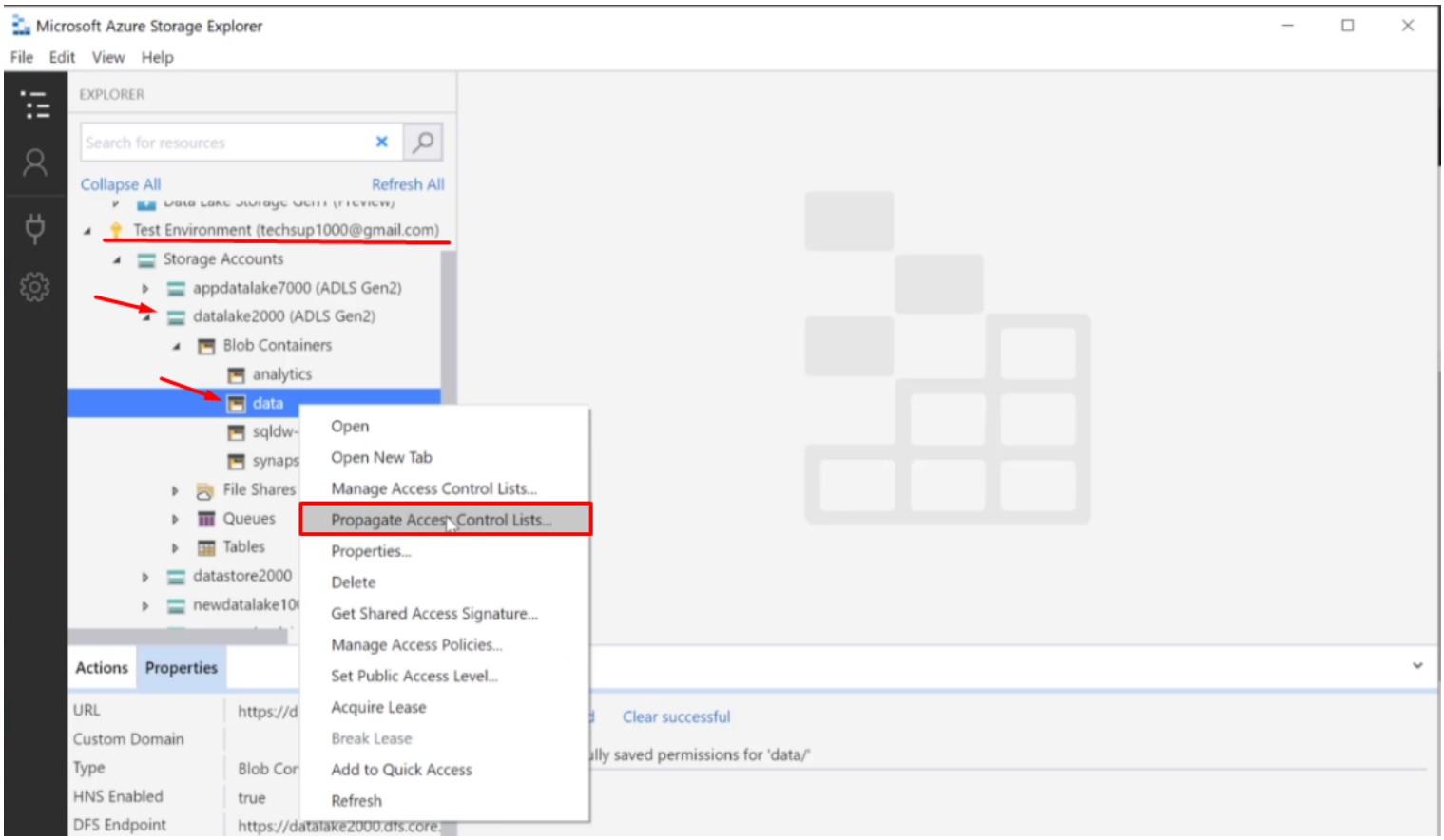


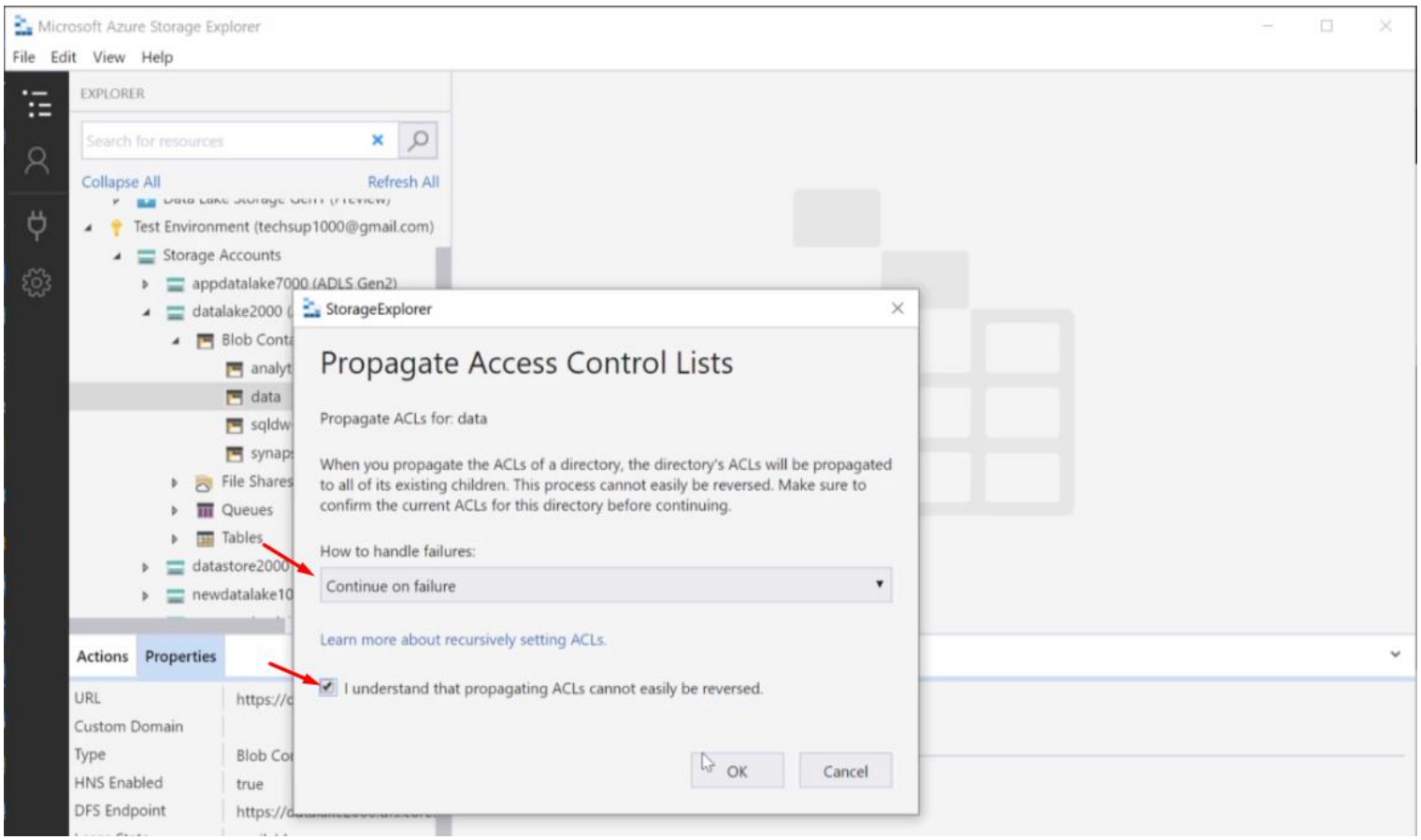

Microsoft Azure Storage Explorer
File Edit View Help
EXPLORER
Search for resources
Collapse All
Cosmos DB Accounts (Deprec
Data Lake Storage Gen1 (Previ
Test Environment (datalake@techsu
Storage Accounts
datalake2000 (ADLS Gen2)
Blob Containers
analytics
data
sqldw-8833399d-
synapse
File Shares
Queues
Tables
Disks
Cosmos DB Accounts (Deprec
Actions
Properties
URL
https://datalake200
Custom Domain
Type
Blob Container (AD
HNS Enabled
true
DFS Endpoint
https://datalake200
Lease State
available
Lease Status
unlocked
Manage Access
Manage Access
Managing permissions for: data/.
Learn more about access control lists (ACLs).
Users, groups, and service principals:
$superuser (Owner)
$superuser (Owning Group)
Other
Mask
Add
Permissions for: $superuser
Read
Write
Execute
Access
Default *
Read and write permissions will only work for an entity if the entity also has execute permissions on all parent directories, including the container (root directory).
* The default ACL determines permissions for new children of this directory. Changing the default ACL does not affect children that already exist. Learn more about default ACLs.
OK
Cancel
Move
Manage ACLs
Properties
More
Blob Type
Content Type
Folder
Block Blob
application/octet-stream
application/json
Details...
Successfully saved permissions for 'data/'
Added Azure account 'datalake@techsup1000gmail.onmicrosoft.com'

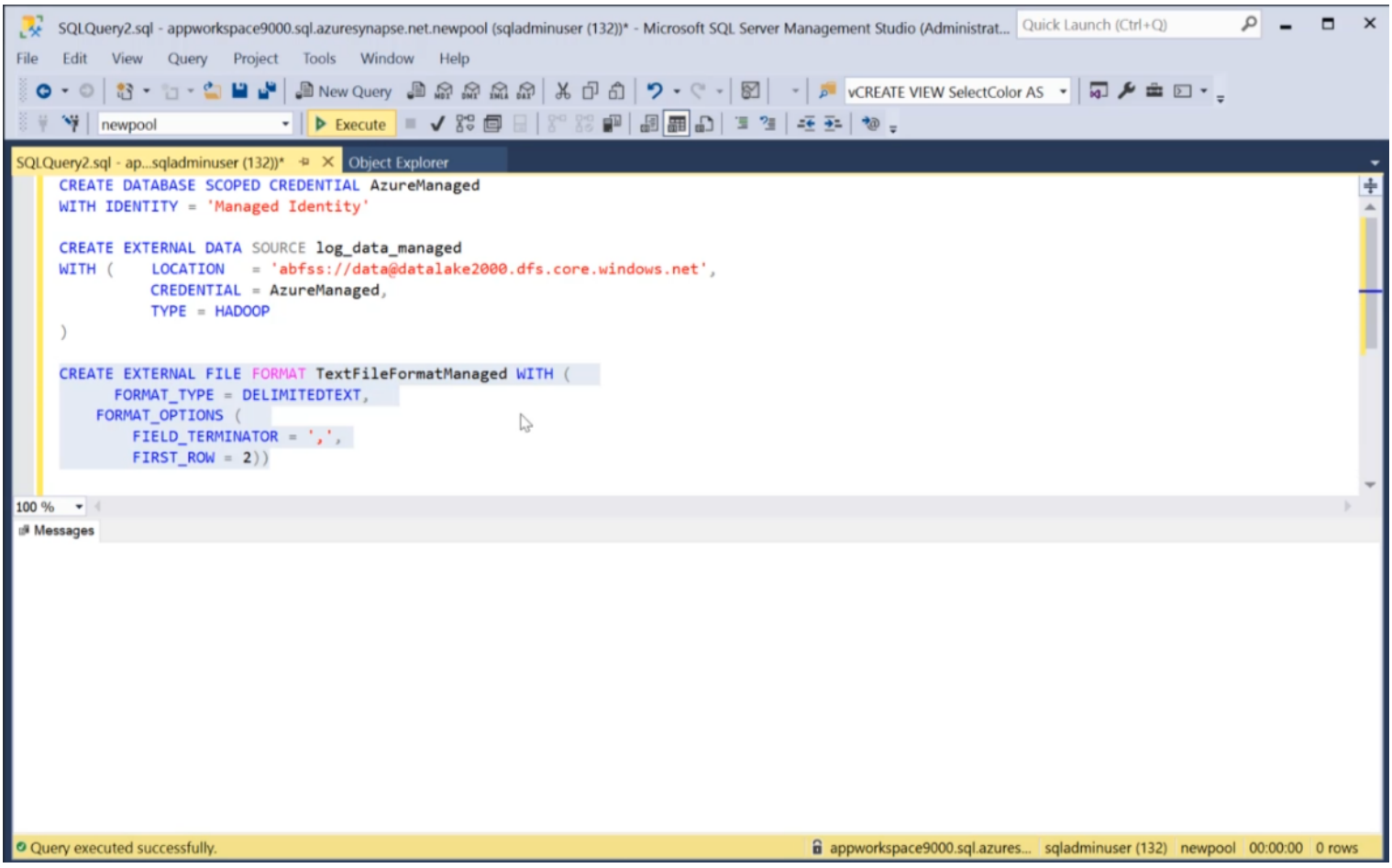

Microsoft Azure Storage Explorer
File Edit View Help
EXPLORER
Search for resources
Collapse All
Cosmos DB Accounts (Deprec
Data Lake Storage Gen1 (Previ
Storage Accounts
datalake2000 (ADLS Gen2)
Blob Containers
analytics
data
sqldw-8833399d-
synapse
File Shares
Queues
Tables
Disks
Cosmos DB Accounts (Deprec
Actions
Properties
URL
https://datalake200
Custom Domain
Type
Blob Container (AD
HNS Enabled
true
DFS Endpoint
https://datalake200
Manage Access
Add Entity
Search for a user, group, or service principal:
appworkspace
Search
appworkspace100012
4da9c041-af8e-4c1a-94bb-e80f9655fb66
appworkspace9000
e9951814-9161-4347-997f-6c970554a175
Back
Add
Cancel
Move
Manage ACLs
Properties
More
Blob Type
Content Type
Folder
Block Blob
application/octet-stream
application/json
Details...



Propagamos los permisos a todos los objetos del contenedor




Microsoft Azure Storage Explorer
File Edit View Help
EXPLORER
Search for resources
Collapse All
Refresh All
Test Environment (techsup1000@gmail.com)
Storage Accounts
appdatalake7000 (ADLS Gen2)
datalake2000
Blob Conta
analyt
data
sqldw
synaps
File Shares
Queues
Tables
datastore2000
newdatalake10
Actions
Properties
URL
Custom Domain
Type
HNS Enabled
true
DFS Endpoint
StorageExplorer
Propagate Access Control Lists
Propagate ACLs for: data
When you propagate the ACLs of a directory, the directory's ACLs will be propagated to all of its existing children. This process cannot easily be reversed. Make sure to confirm the current ACLs for this directory before continuing.
How to handle failures:
Continue on failure
Learn more about recursively setting ACLs.
I understand that propagating ACLs cannot easily be reversed.
OK
Cancel


SQLQuery2.sql - appworkspace9000.sql.azuresynapse.net.newpool (sqladminuser (132))* - Microsoft SQL Server Management Studio (Administrat...
Quick Launch (Ctrl+Q)
File Edit View Query Project Tools Window Help
New Query
vCREATE VIEW SelectColor AS
newpool
Execute
SQLQuery2.sql - ap...sqladminuser (132))*
Object Explorer
CREATE DATABASE SCOPED CREDENTIAL AzureManaged
WITH IDENTITY = 'Managed Identity'
CREATE EXTERNAL DATA SOURCE log_data_managed
WITH (LOCATION = 'abfss://data@datalake2000.dfs.core.windows.net',
CREDENTIAL = AzureManaged,
TYPE = HADOOP
)
CREATE EXTERNAL FILE FORMAT TextFileFormatManaged WITH (
FORMAT_TYPE = DELIMITEDTEXT,
FORMAT_OPTIONS (
FIELD_TERMINATOR = ',',
FIRST_ROW = 2))
100 %
Messages
Query executed successfully.
appworkspace9000.sql.azures...
sqladminuser (132)
newpool
00:00:00
0 rows

```sql

CREATE EXTERNAL TABLE logdatamanaged
(
    [Id] [int] NULL,
    [Correlationid] [varchar](200) NULL,
    [Operationname] [varchar](200) NULL,
    [Status] [varchar](100) NULL,
    [Eventcategory] [varchar](100) NULL,
    [Level] [varchar](100) NULL,
    [Time] [datetime] NULL,
    [Subscription] [varchar](200) NULL,
    [Eventinitiatedby] [varchar](1000) NULL,
    [Resourcetype] [varchar](1000) NULL,
    [Resourcegroup] [varchar](1000) NULL
)
WITH (
 LOCATION = 'cleaned/Log.csv',
    DATA_SOURCE = log_data_managed,
    FILE_FORMAT = TextFileFormatManaged
)



SELECT * FROM logdatamanaged

-- If you want to clean up your resources

DROP EXTERNAL TABLE logdatamanaged
DROP EXTERNAL FILE FORMAT TextFileFormatManaged
DROP EXTERNAL DATA SOURCE log_data_managed
DROP DATABASE SCOPED CREDENTIAL AzureManaged
```