ANOTHER BRICK IN THE FIREWALL

How to Secure your Azure Data Platform

Grace O'Halloran

INTRODUCTION

- Grace O'Halloran
- Principal Data Engineering Consultant @ Advancing Analytics
- Microsoft Data Platform MVP
- 7+yrs working with Azure Data Platforms
- Microsoft Certified Azure Developer & Administrator

Grace O'Halloran (grace-o-halloran)
@graceaohalloran
grace@advancinganalytics.co.uk
www.thinkingacloud.co.uk
https://github.com/gracedev94/GraceOH-CommunityContent

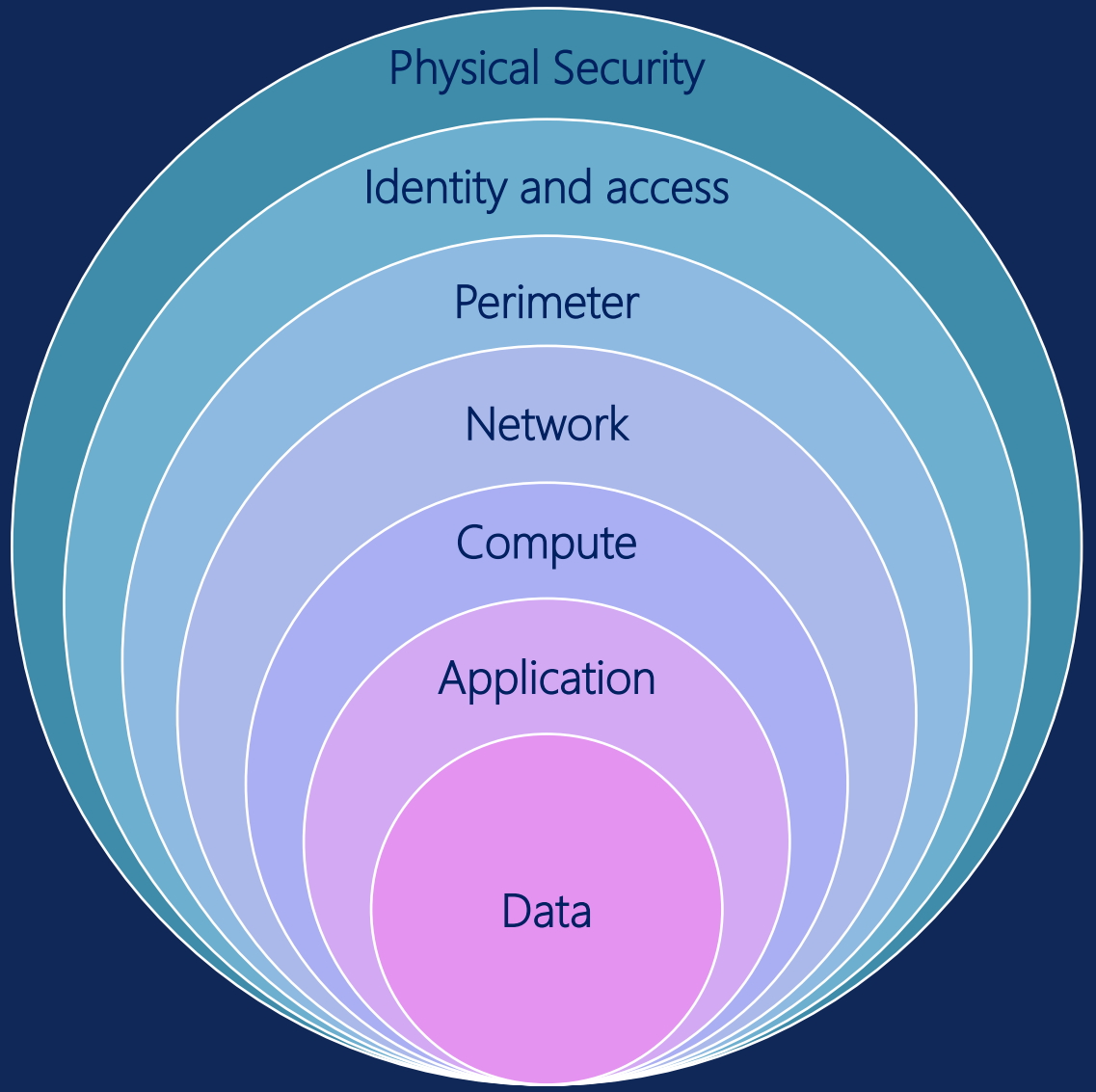
ANOTHER BRICK IN THE FIREWALL

Why care about Network Security	Networks	Ingress & Egress	Azure Private Link	Data Platform Components
• I do data – why is this relevant to me?	• Hub-and-spoke Topology • Address Space Considerations	• Firewalls & UDRs • Network Security Groups • Secure Development Access	• Private Endpoints • Azure Private DNS	• Azure Data Factory • Fabric • Azure DevOps • Databricks

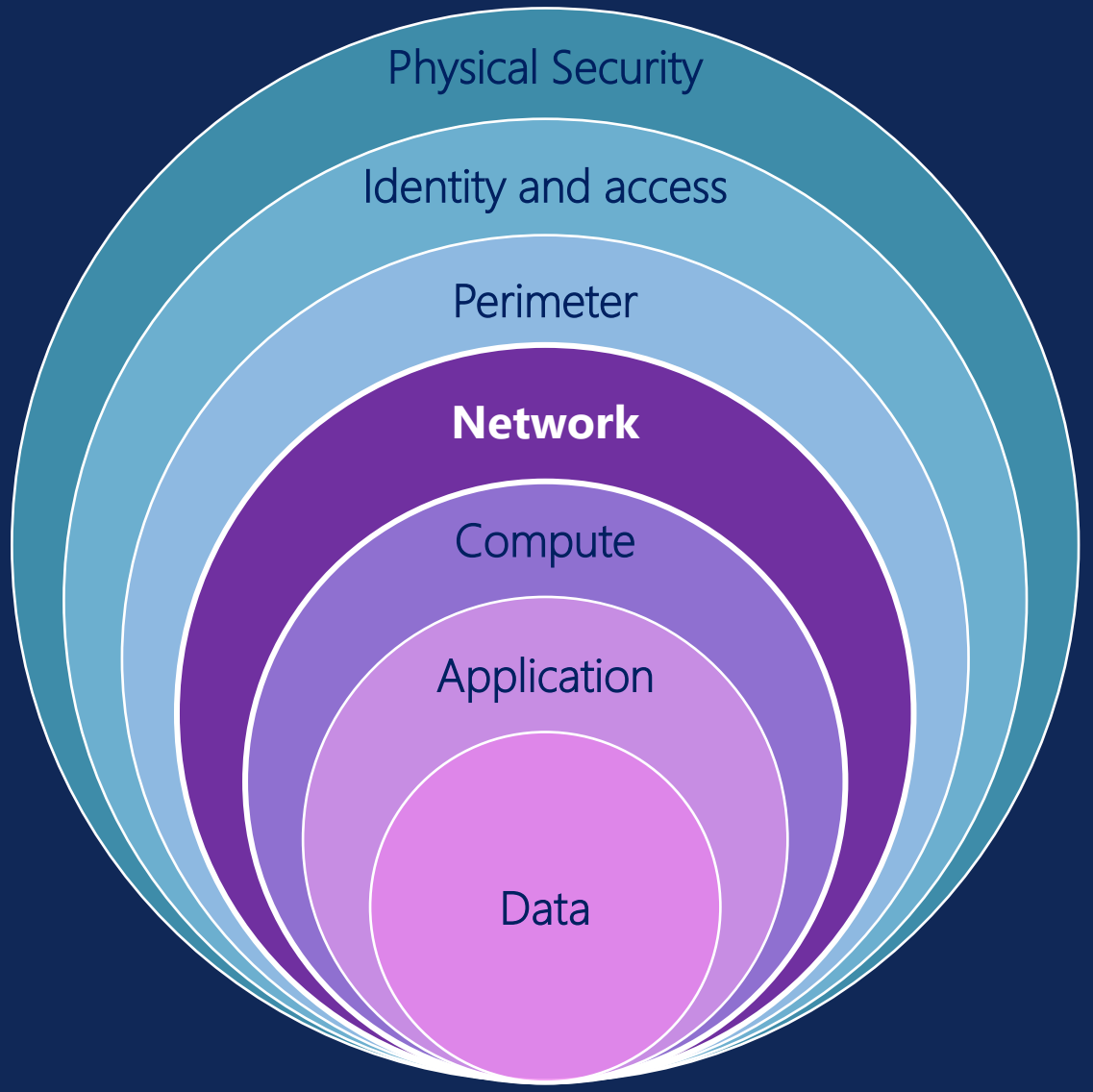
WHY CARE ABOUT NETWORK SECURITY

I do data – why is this relevant to me?

DEFENSE IN DEPTH



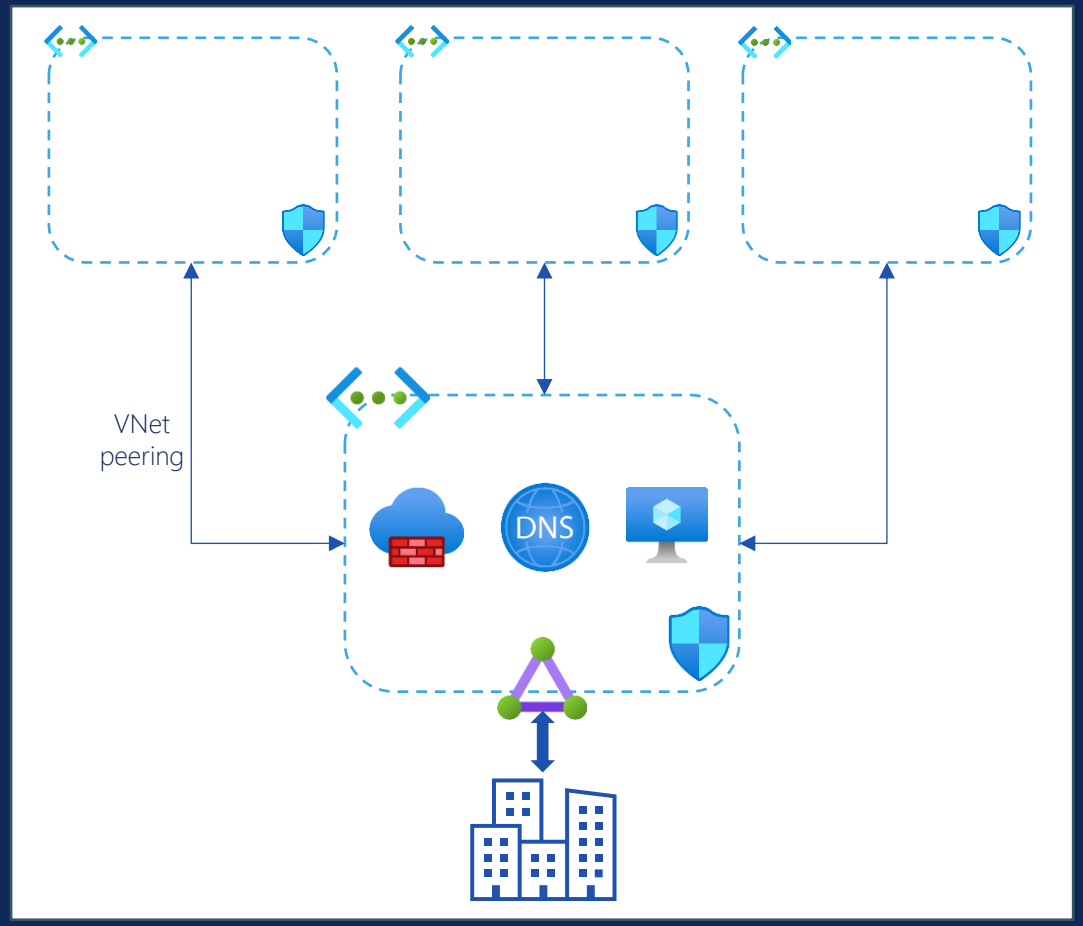
DEFENSE IN DEPTH



NETWORKS

Hub-and-Spoke Topology

HUB-AND-SPOKE TOPOLOGY



NETWORKS

Address Space Considerations

ADDRESS SPACE CONSIDERATIONS

Avoid Conflicts

Use IPAM

Size Requirements

Allow for Growth

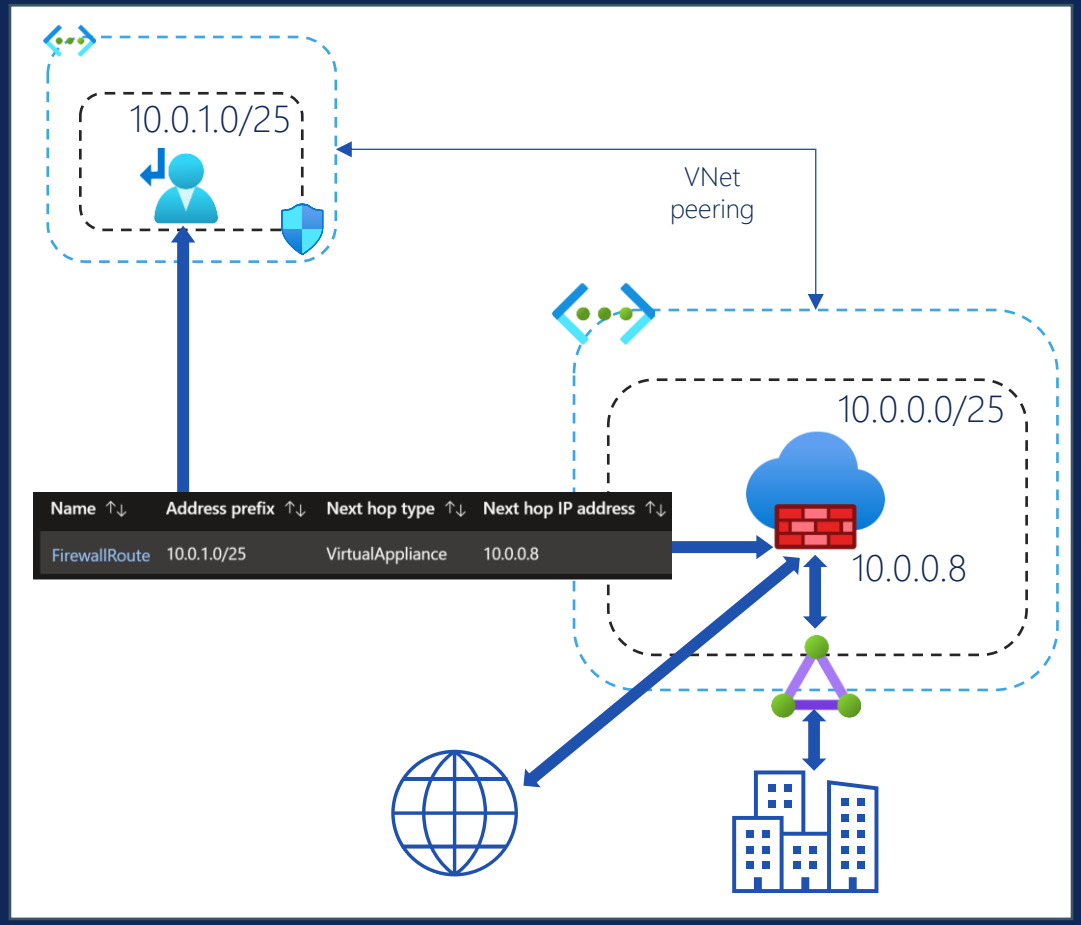
INGRESS & EGRESS

Firewalls & UDRs

“Ingress and Egress are fancy words for Inbound and Outbound.”

- Grace O’Halloran, now.

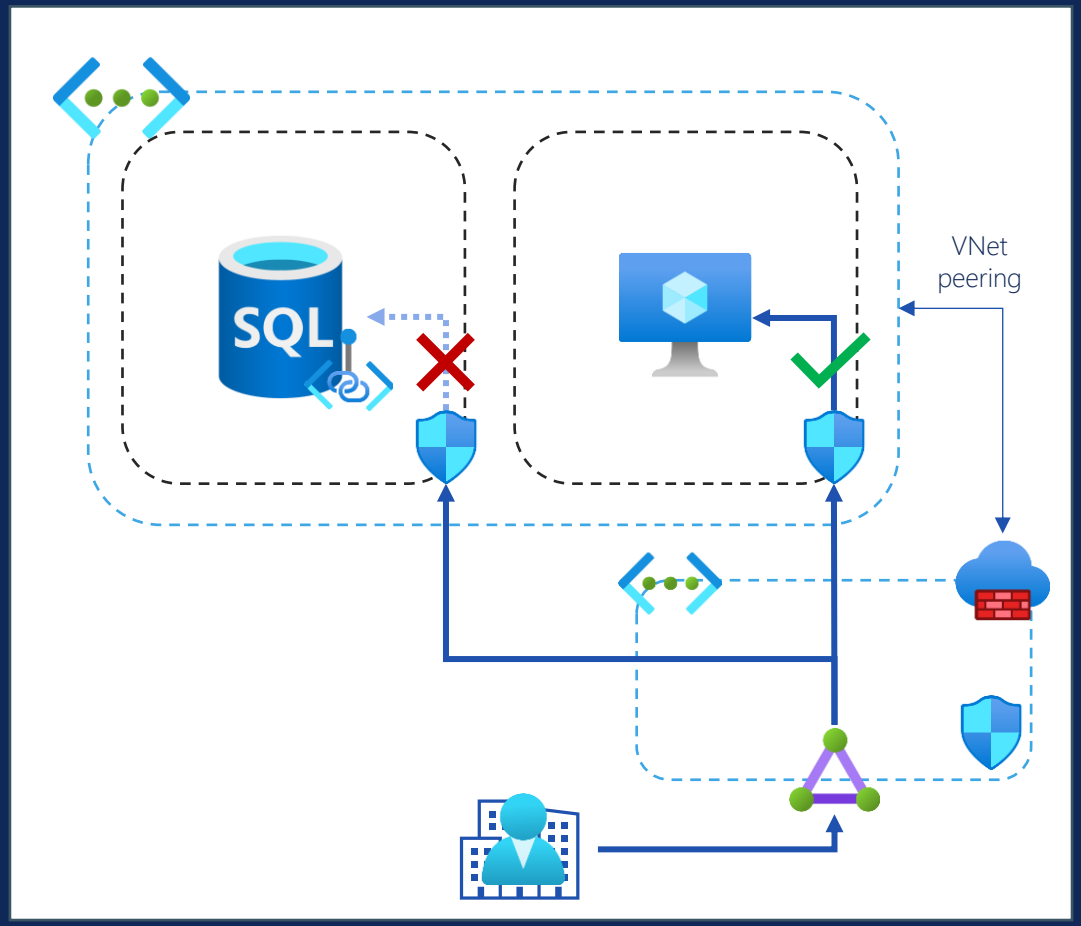
FIREWALLS & USER DEFINED ROUTES



INGRESS & EGRESS

Network Security Groups

NETWORK SECURITY GROUPS



INGRESS & EGRESS

Secure Development Access

SECURE DEVELOPMENT ACCESS

Virtualisation Tool

- Azure Virtual Desktop
- Windows Cloud PC
- Citrix
- VMWare

Microsoft DevBox

- DevBox provides preconfigured cloud-based developer workstations

Azure Bastion

- Azure Bastion provides a host for users to securely connect to Azure VMs.

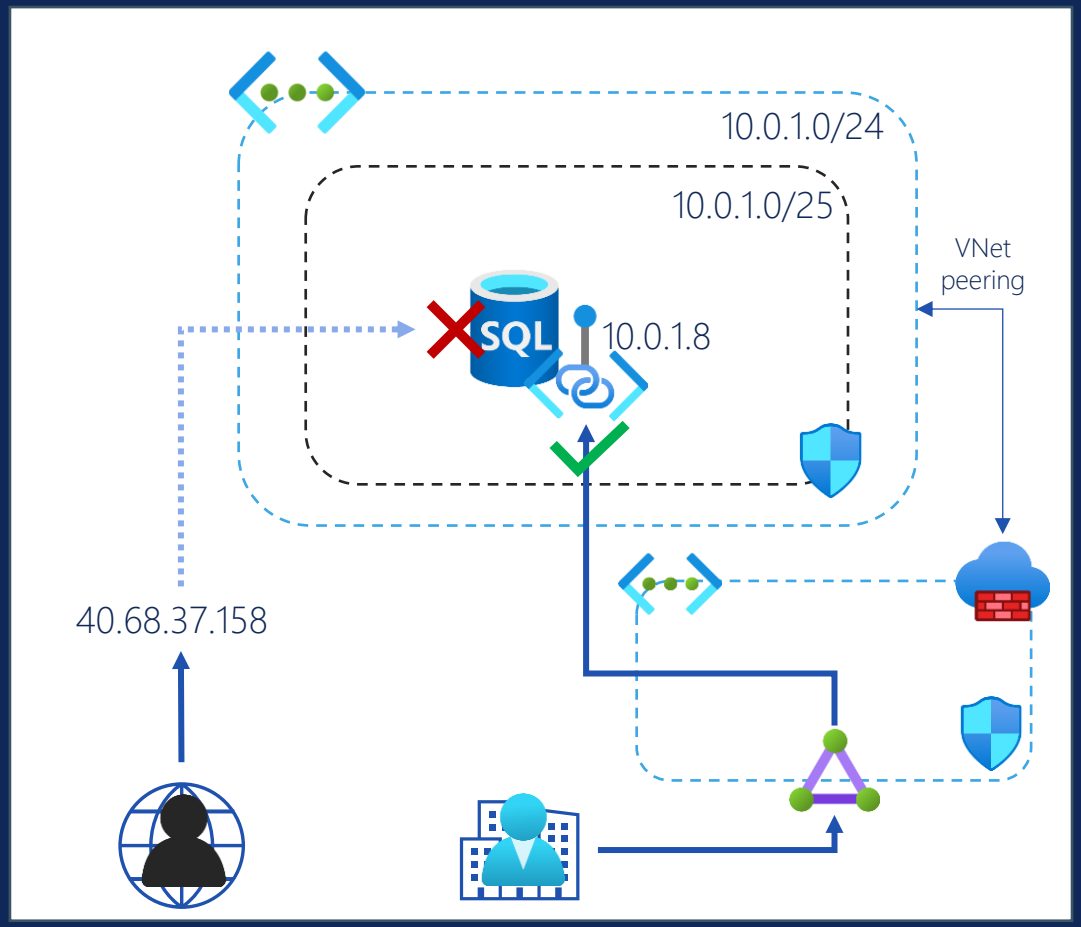
Jump Box

- Infra is responsible for maintaining the security of the jump box.

AZURE PRIVATE LINK

Private Endpoints

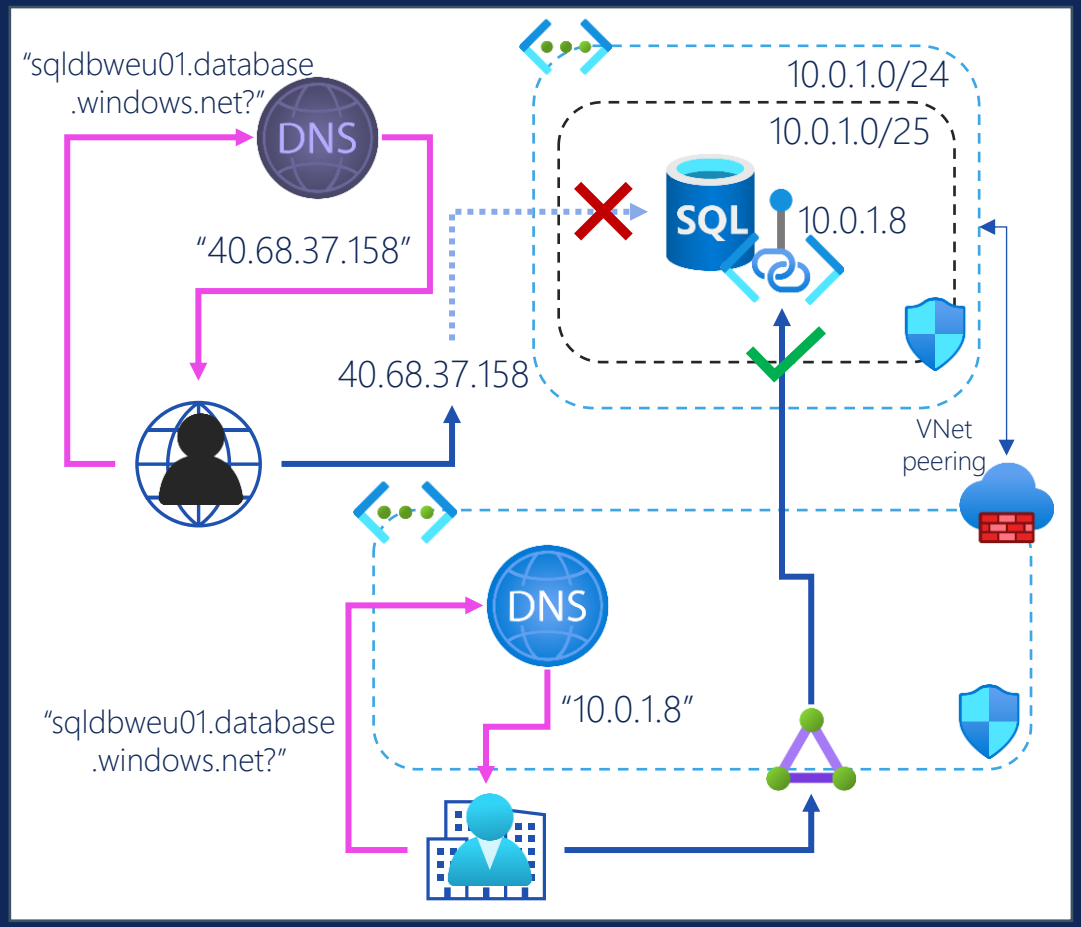
AZURE PRIVATE LINK



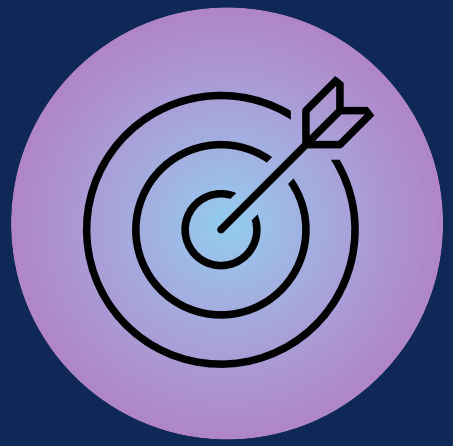
AZURE PRIVATE LINK

Azure Private DNS

AZURE PRIVATE DNS



PRIVATE DNS ZONES



Central

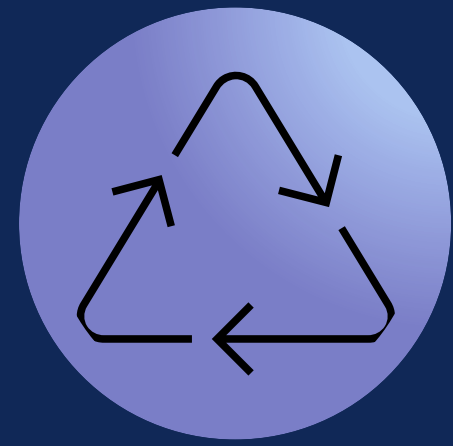
Private DNS Zones should be part of a central DNS solution

One per Domain

One Private DNS Zone required per Azure domain

Naming

Private DNS Zones used for Private Link must have specific names



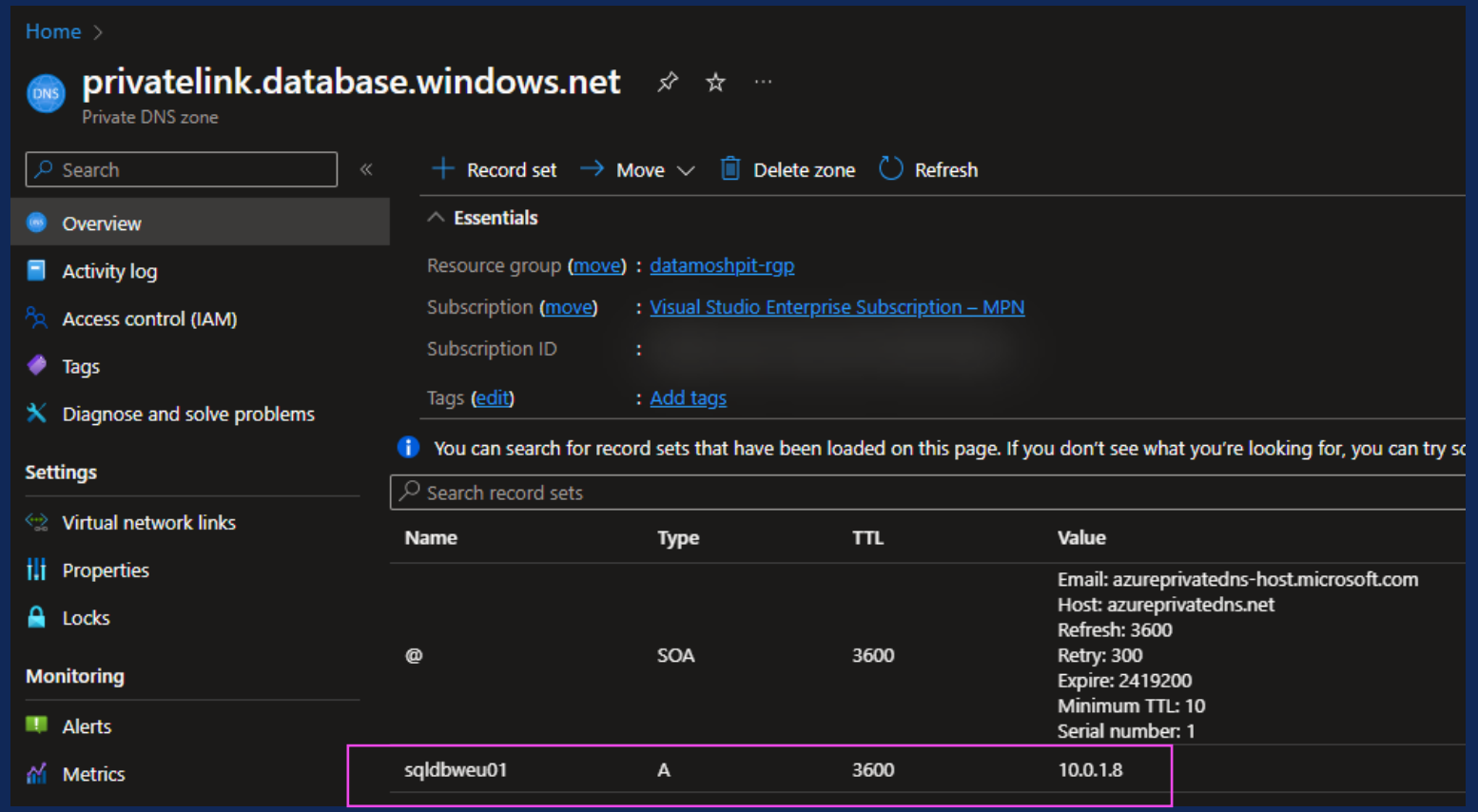
Reusable

Resources part of the same domain can reuse the same Private DNS Zone

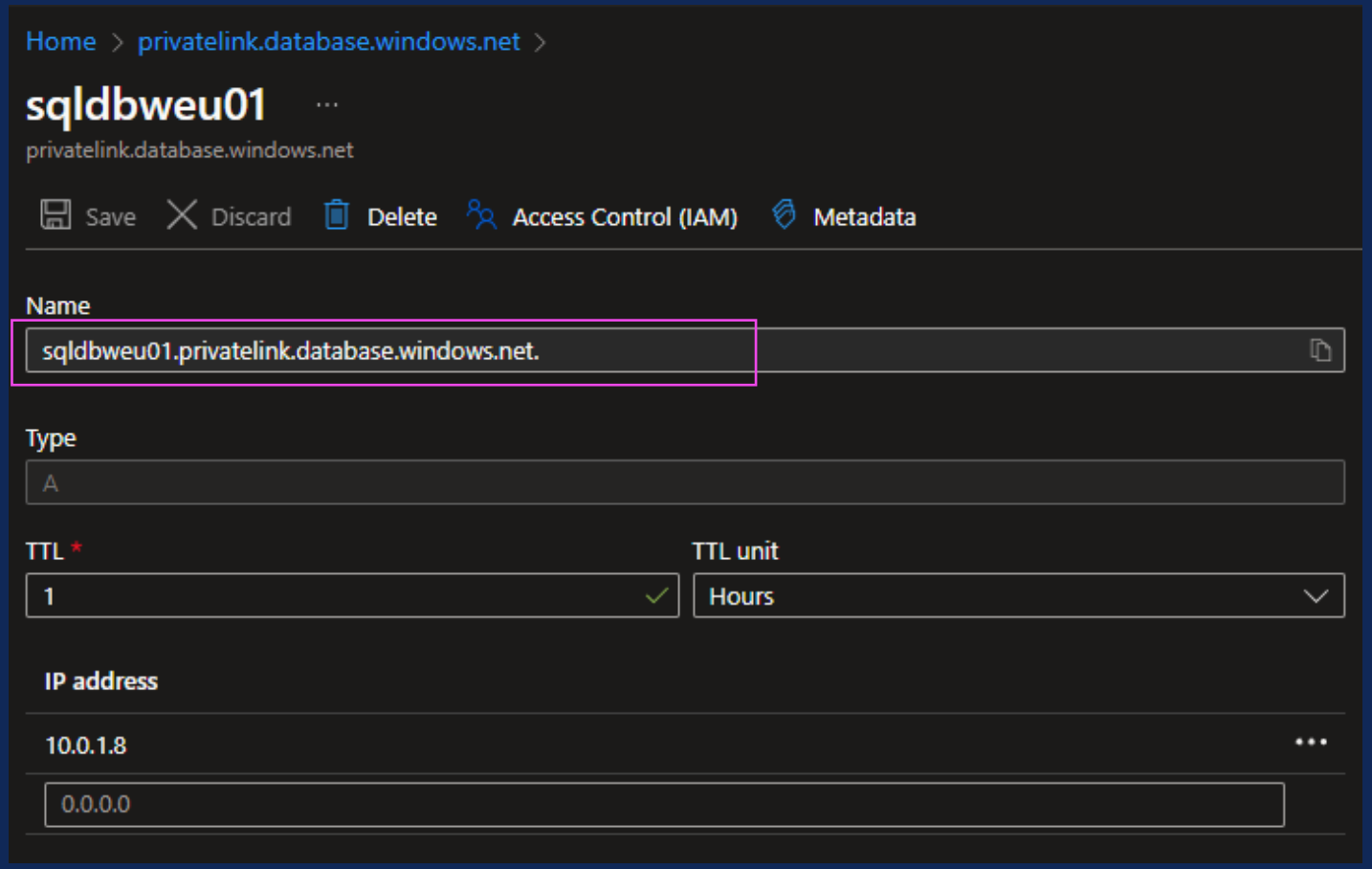
PRIVATE DNS ZONES

Azure SQL Server resource name:	sqldbweu01
Azure SQL Server public endpoint:	sqldbweu01.database.windows.net
Domain:	database.windows.net
Azure Private DNS Zone required:	privatelink.database.windows.net

A RECORDS


Home
privatelink.database.windows.net
Private DNS zone
Search
Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Settings
Virtual network links
Properties
Locks
Monitoring
Alerts
Metrics
Record set
Move
Delete zone
Refresh
Essentials
Resource group (move) : datamoshpit-rgp
Subscription (move) : Visual Studio Enterprise Subscription – MPN
Subscription ID :
Tags (edit) : Add tags
You can search for record sets that have been loaded on this page. If you don't see what you're looking for, you can try sc
Search record sets
Name
Type
TTL
Value
@
SOA
3600
Email: azureprivatedns-host.microsoft.com
Host: azureprivatedns.net
Refresh: 3600
Retry: 300
Expire: 2419200
Minimum TTL: 10
Serial number: 1
sqldbweu01
A
3600
10.0.1.8

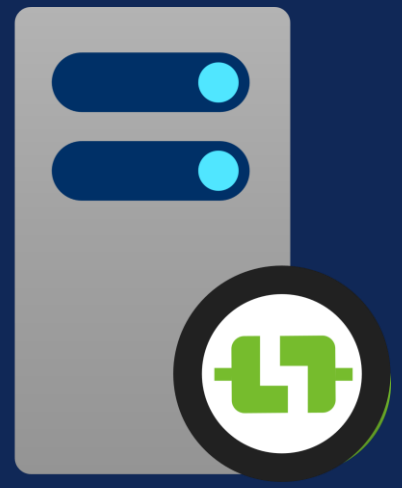
A RECORDS


Home > privatelink.database.windows.net >
sqldbweu01
privatelink.database.windows.net
Save
Discard
Delete
Access Control (IAM)
Metadata
Name
sqldbweu01.privatelink.database.windows.net.
Type
A
TTL *
1
TTL unit
Hours
IP address
10.0.1.8
0.0.0.0

DATA PLATFORM COMPONENTS

Azure Data Factory: Integration Runtimes

ADF: INTEGRATION RUNTIMES



ADF: INTEGRATION RUNTIMES

Azure IR with Managed VNet

Use the Azure-hosted IR with the Managed VNet enabled in order to secure the compute inside a private network.

You must use Managed Private Endpoints to allow your IR access to your protected resources.

Self-hosted IR (SHIR)

SHIRs are created by installing an IR application on your own machine, this can be an on-prem server or an Azure VM.

The SHIR server will utilise existing Private Endpoints to securely connect to your protected resources.

ADF: INTEGRATION RUNTIMES

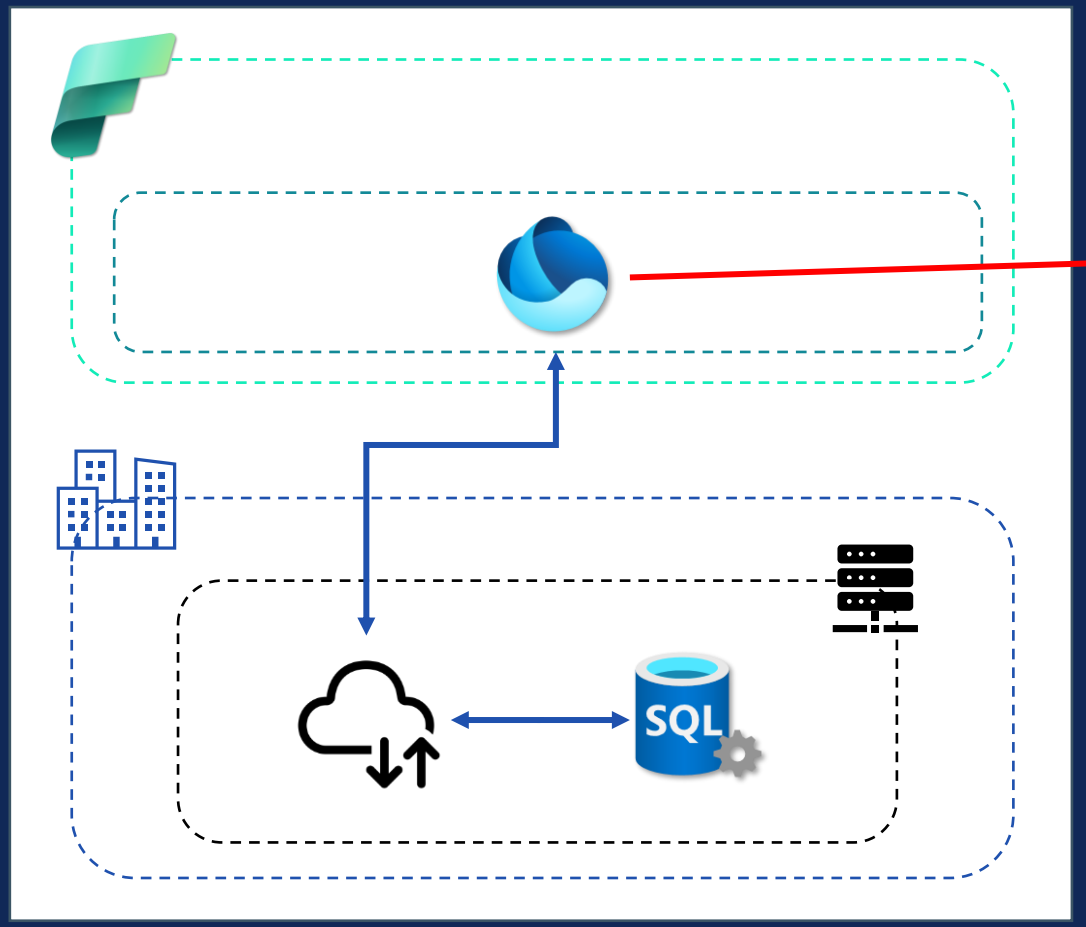
	Azure IR with Managed VNet	Self-hosted IR (SHIR)
Pros	Fully managed and serverless Elastic scaling No maintaining of firewall rules	High Availability options Runtime costs are cheaper Allows for easy connectivity to on-prem data sources
Cons	No control over address space Requires additional private endpoints Can increase cost Doesn't work easily with on-prem connectivity	Requires pre-existing infrastructure Responsible for providing and maintaining the server Maintenance of firewall rules Pay for compute resource

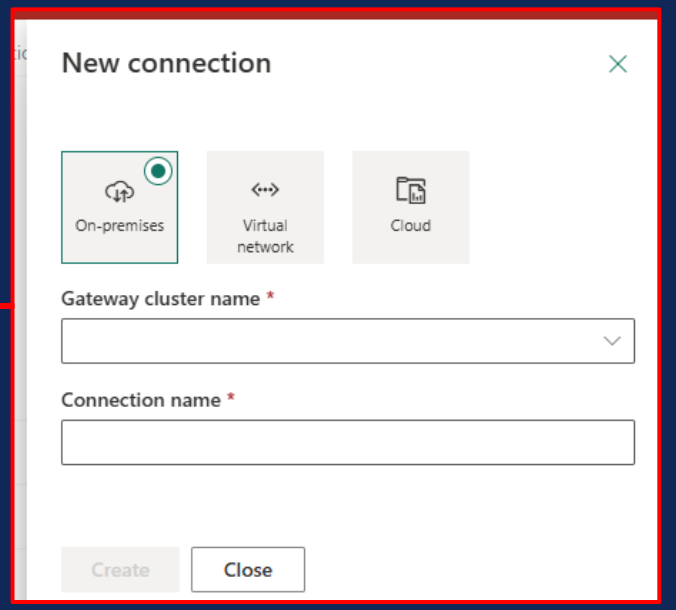
DATA PLATFORM COMPONENTS

Fabric: Data Gateways

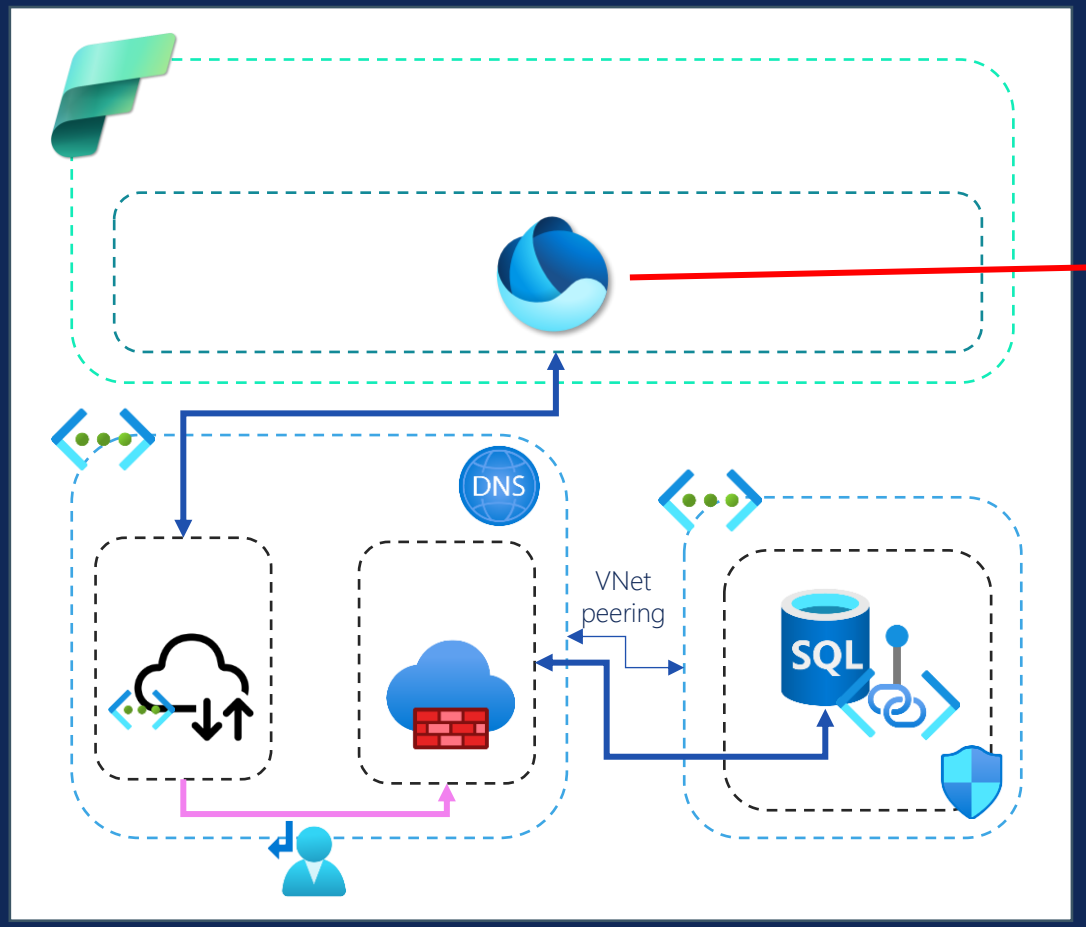
FABRIC: DATA GATEWAYS

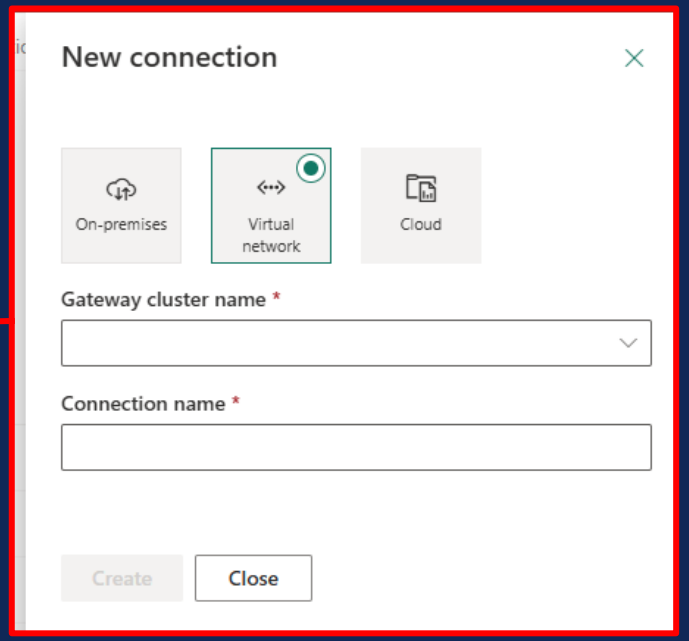
FABRIC: ON-PREM DATA GATEWAY





FABRIC: VNET DATA GATEWAY

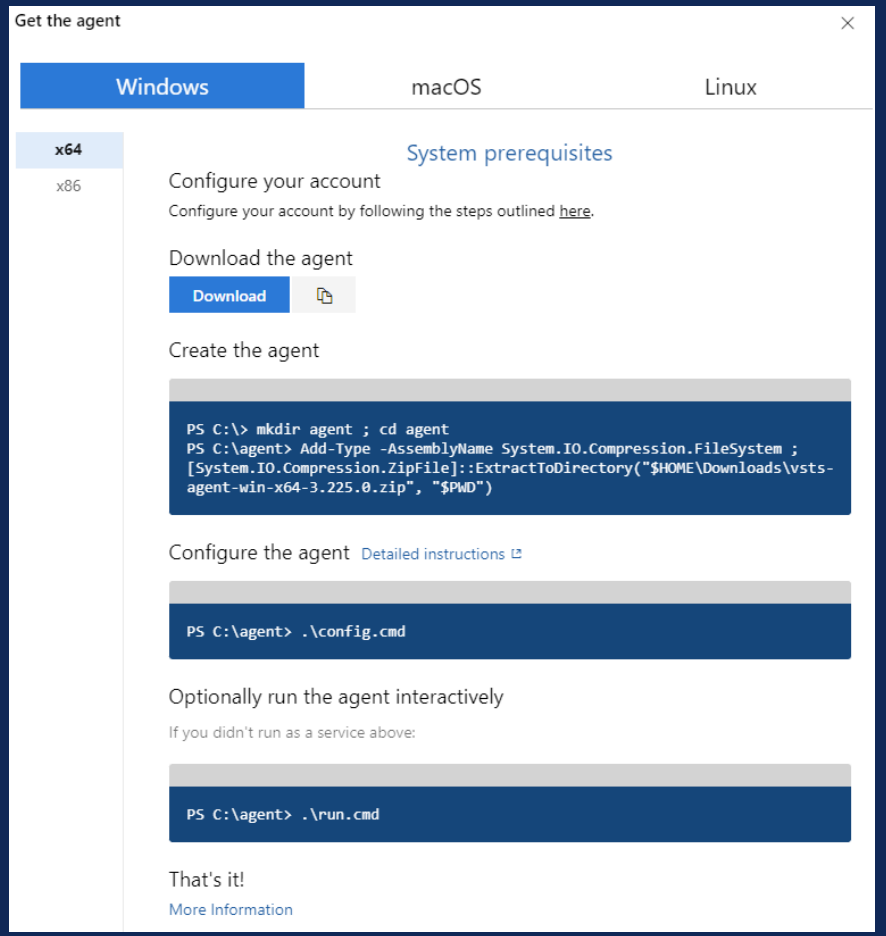




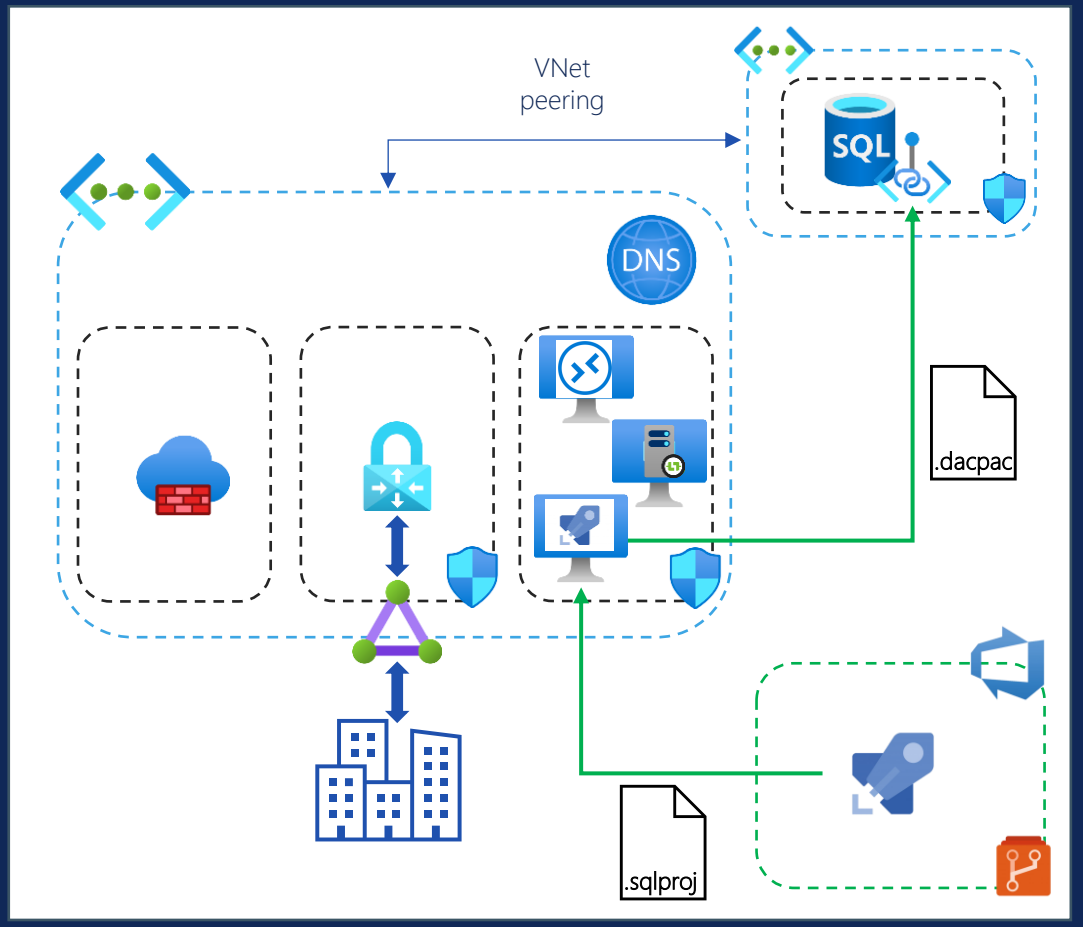
DATA PLATFORM COMPONENTS

Azure DevOps: Self-hosted Build Agents

AZURE DEVOPS: BUILD AGENTS



AZURE DEVOPS: BUILD AGENTS



DATA PLATFORM COMPONENTS

Databricks: VNet Injection, Secure Cluster Connectivity, Private Link, Network Connectivity, Configurations.

DATABRICKS

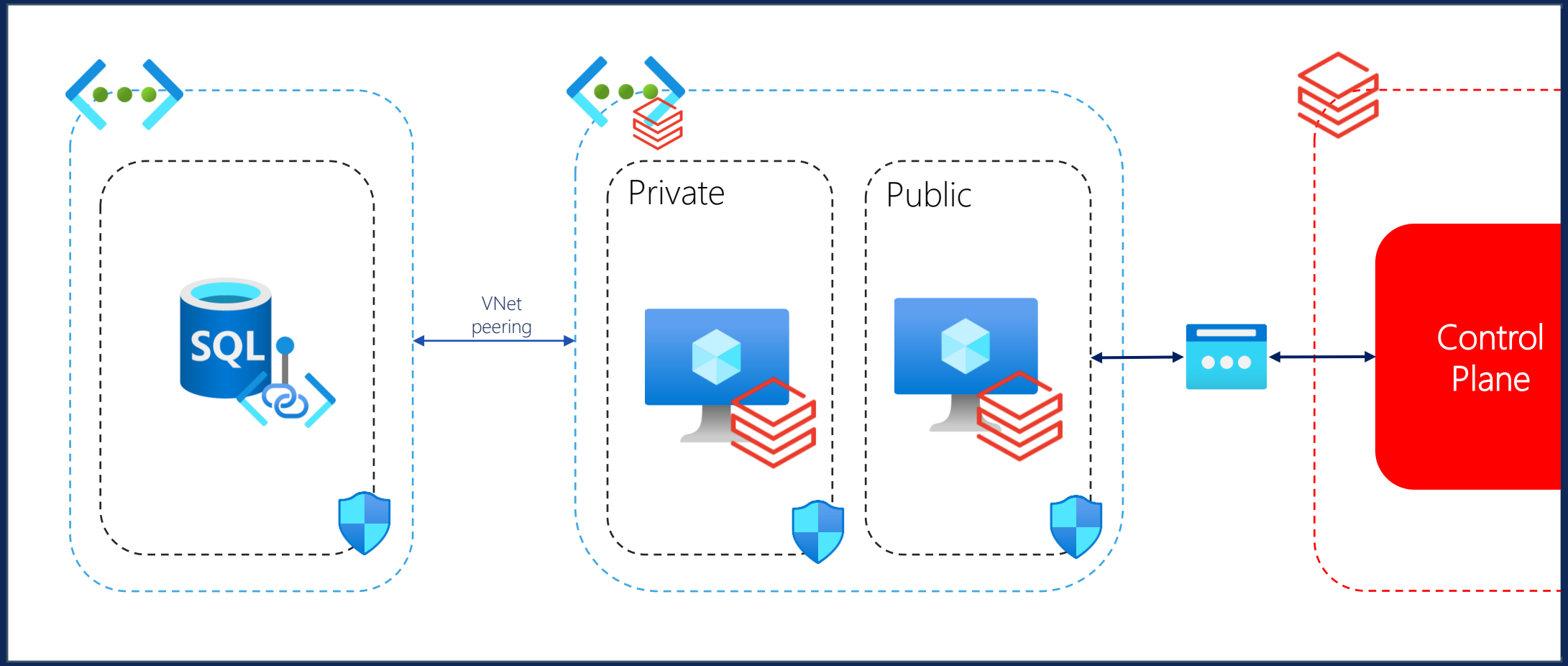
Managed VNet

VNet Injection

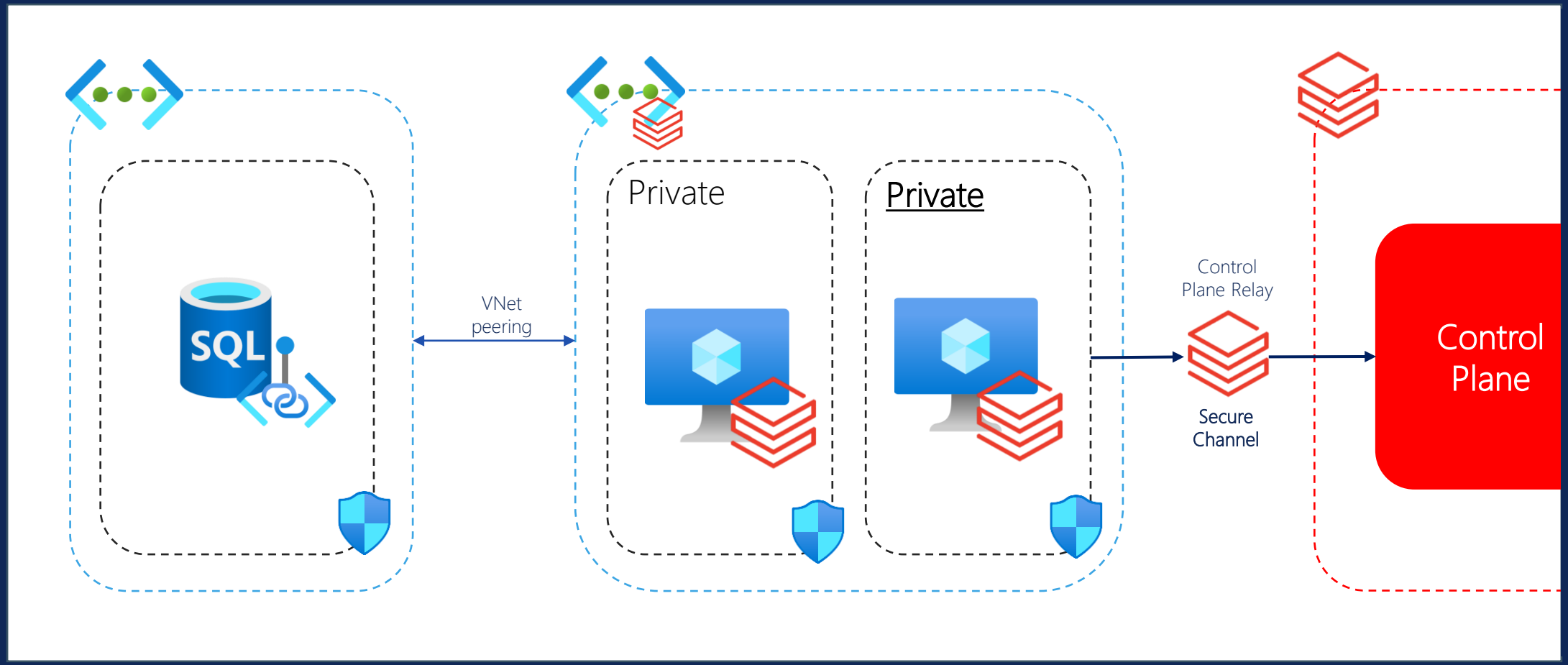
Secure Cluster Connectivity

Private Link

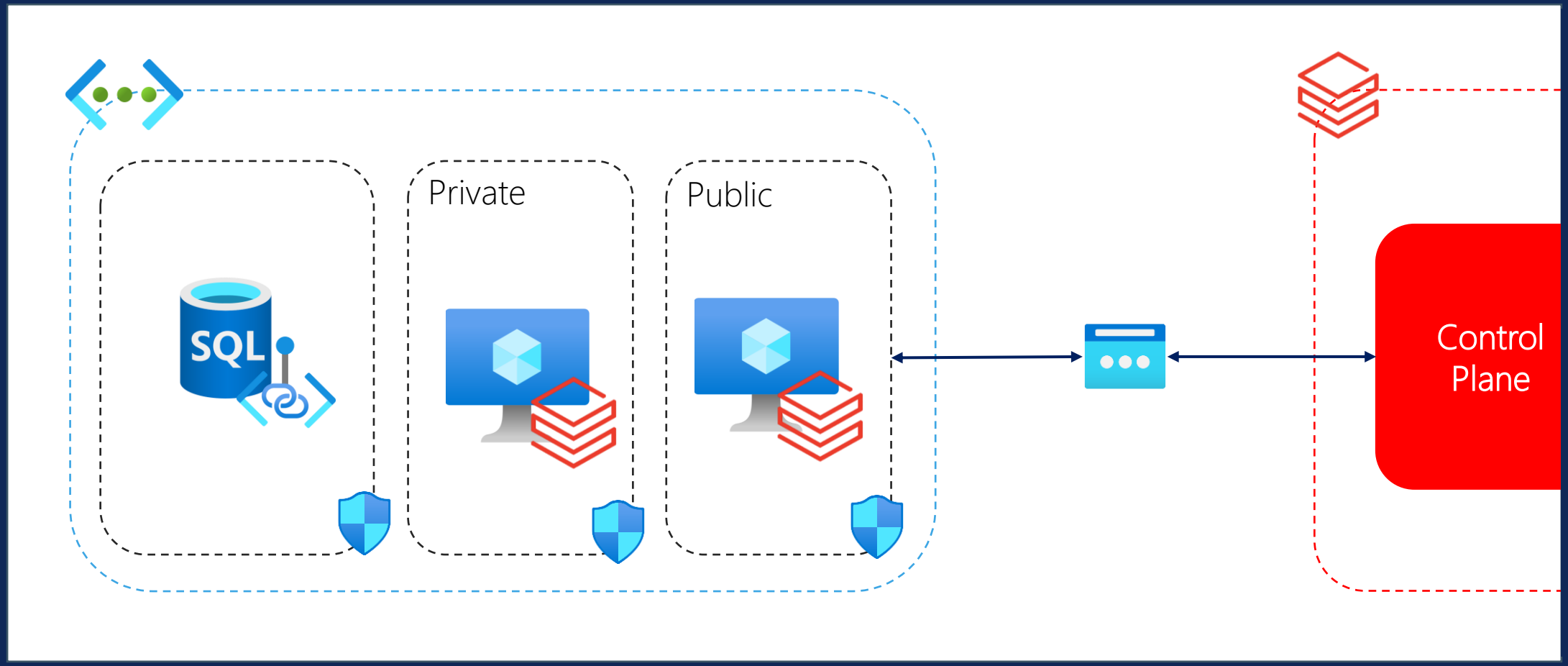
DATABRICKS: MANAGED VNET



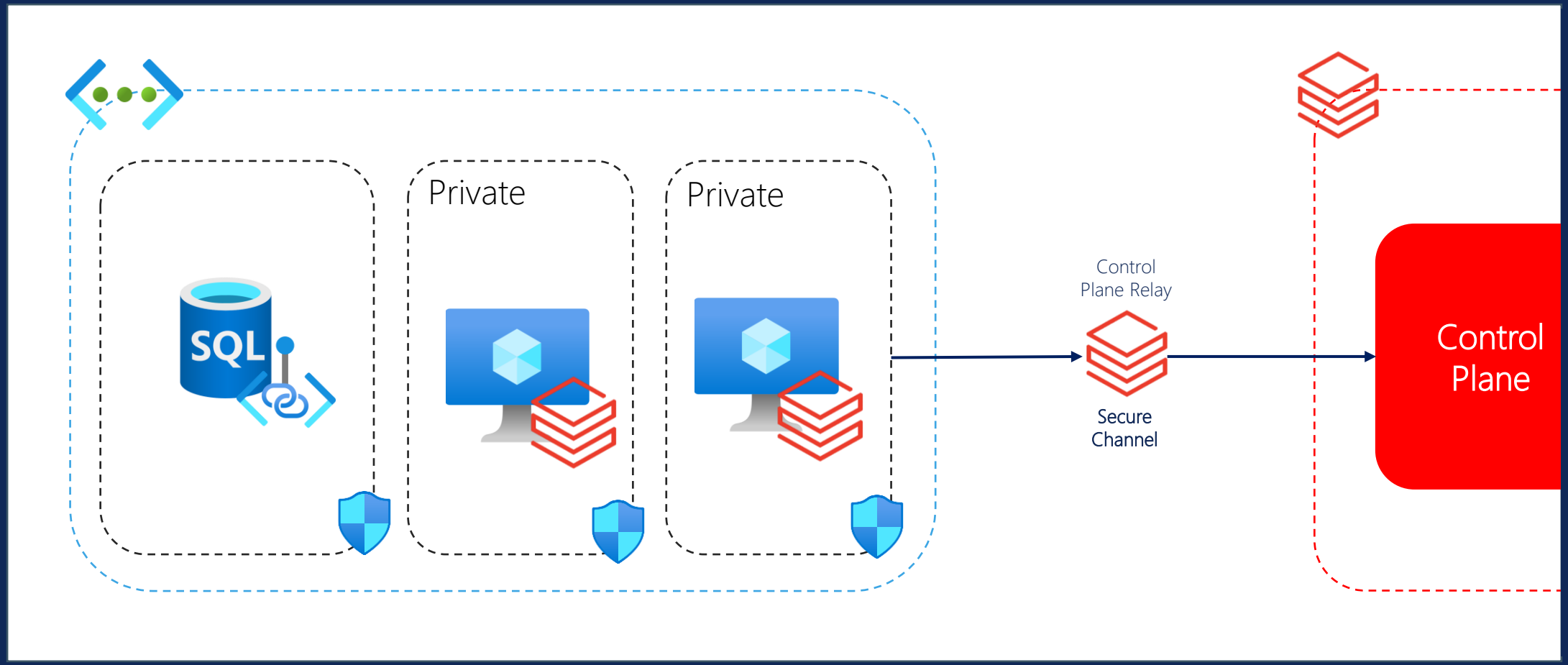
DATABRICKS: MANAGED VNET & SCC



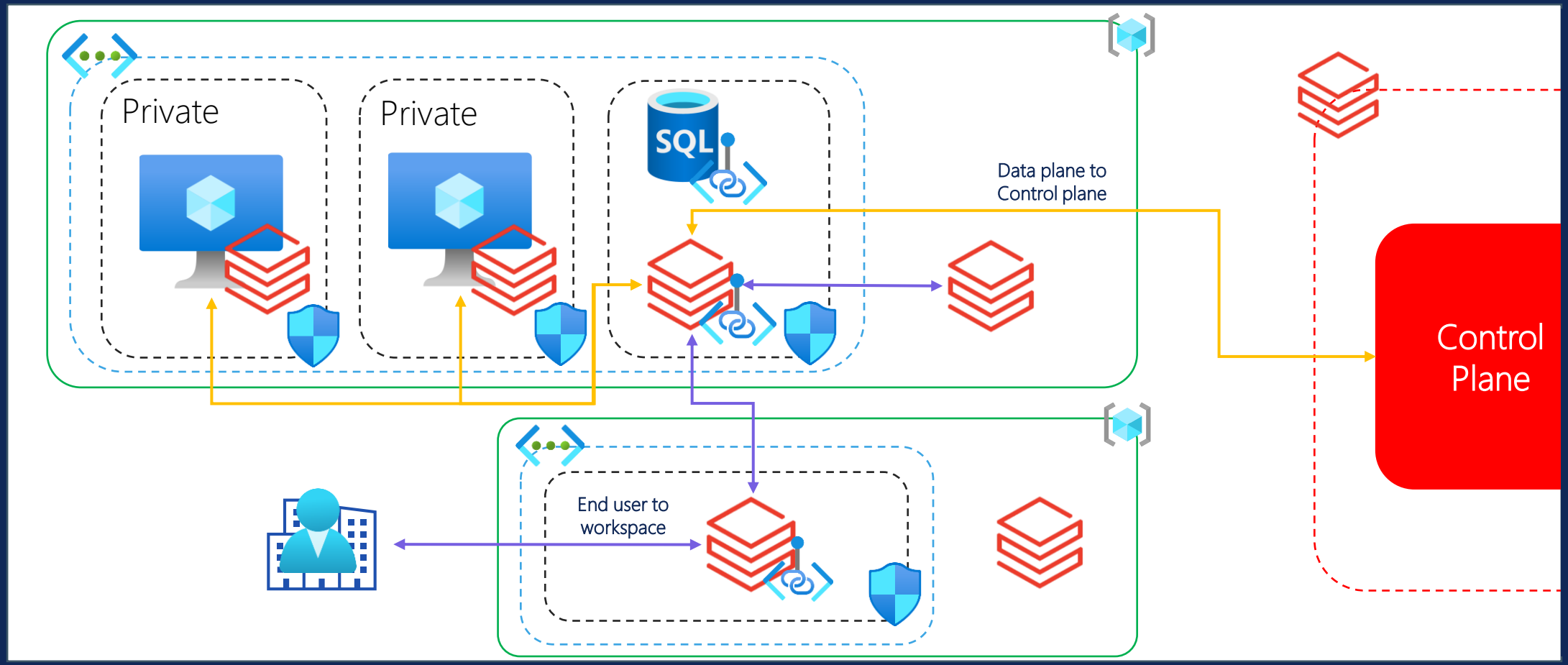
DATABRICKS: VNET INJECTION



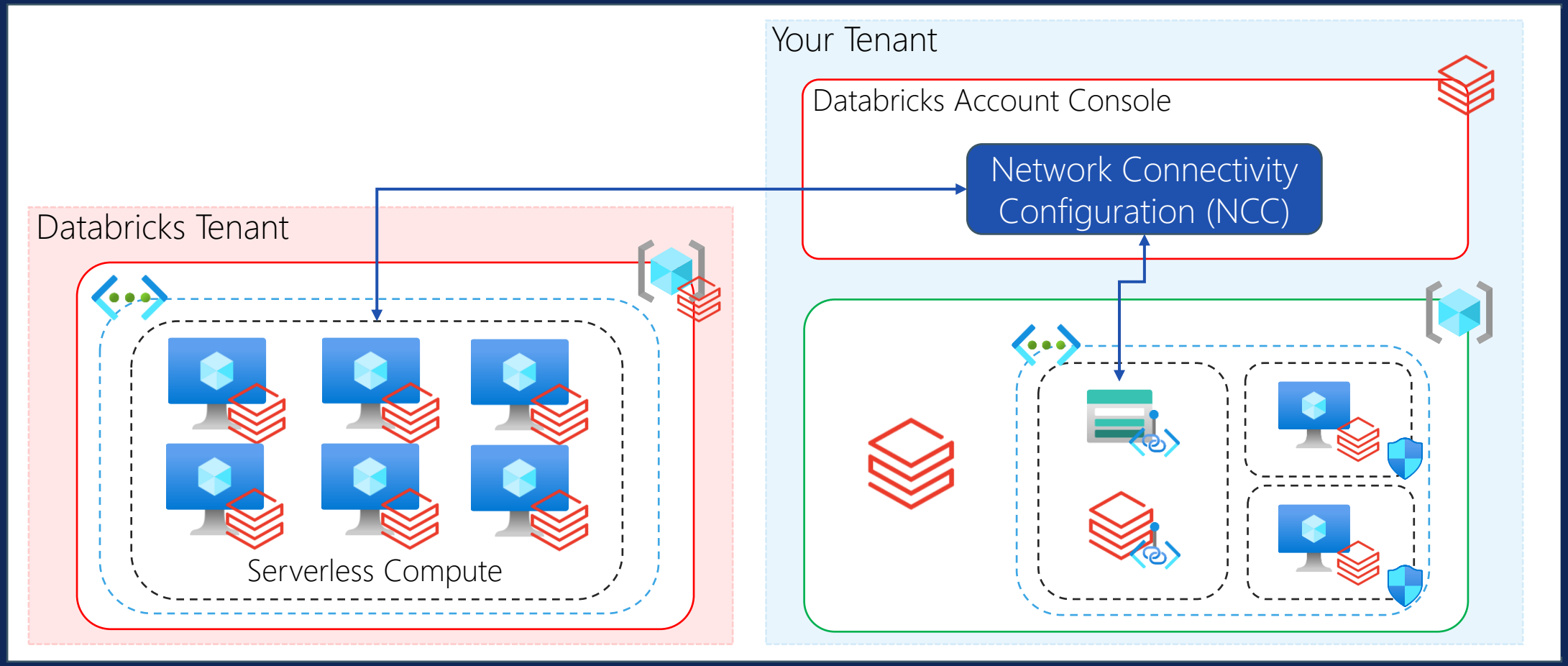
DATABRICKS: VNET INJECTION & SCC



DATABRICKS: VNET INJECTION & PRIVATE LINK



DATABRICKS: NCCS



Private Endpoints - YouTube

ADF IRs - Blog

Community Content - Github

THANK YOU

Any questions?

Grace O'Halloran (grace-o-halloran)

@graceaohalloran

grace@advancinganalytics.co.uk

www.thinkingacloud.co.uk

https://github.com/gracedev94/GraceOH-CommunityContent