



# Craig Porteous

Principal Consultant



**ADVANCING  
ANALYTICS**



<https://craigporteous.com>



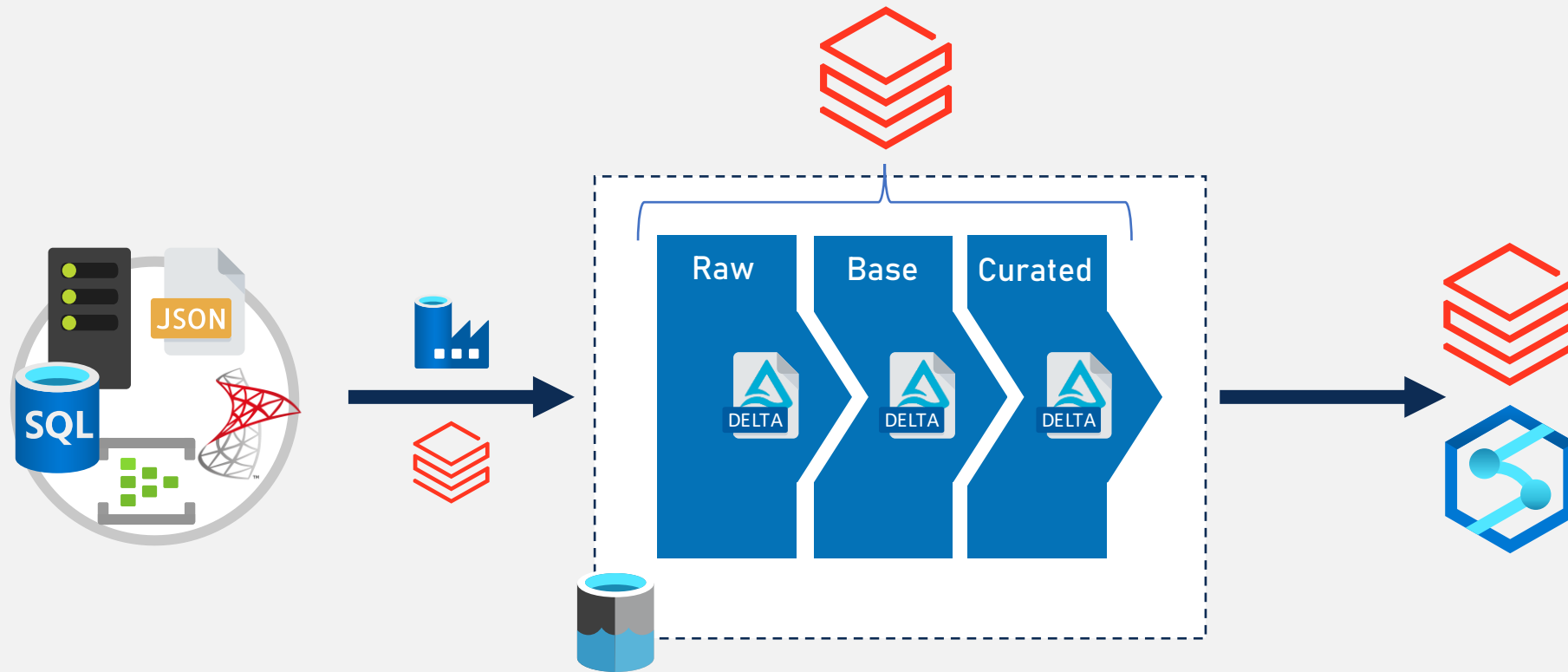
@cporteous



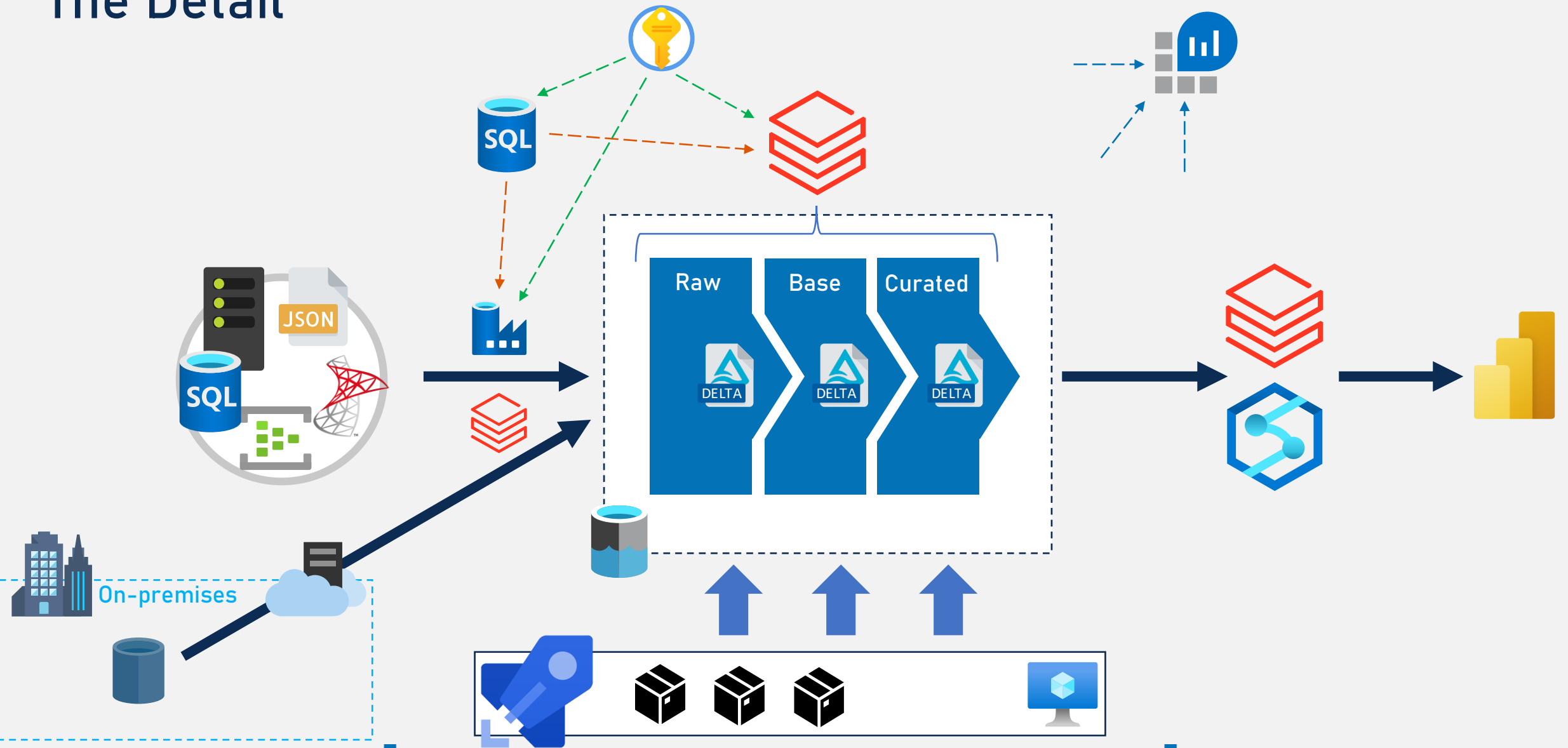
<https://github.com/cporteu>

## Designing Data Architectures that InfoSec will actually approve

# The Data Platform



# The Detail



# Public doesn't mean PUBLIC



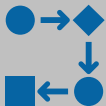
Networking



Resource Permissions



Active Directory



Governance / Process



Data Encryption

# How do we secure the “cloud”?



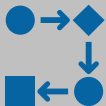
Networking



Resource Permissions



Active Directory



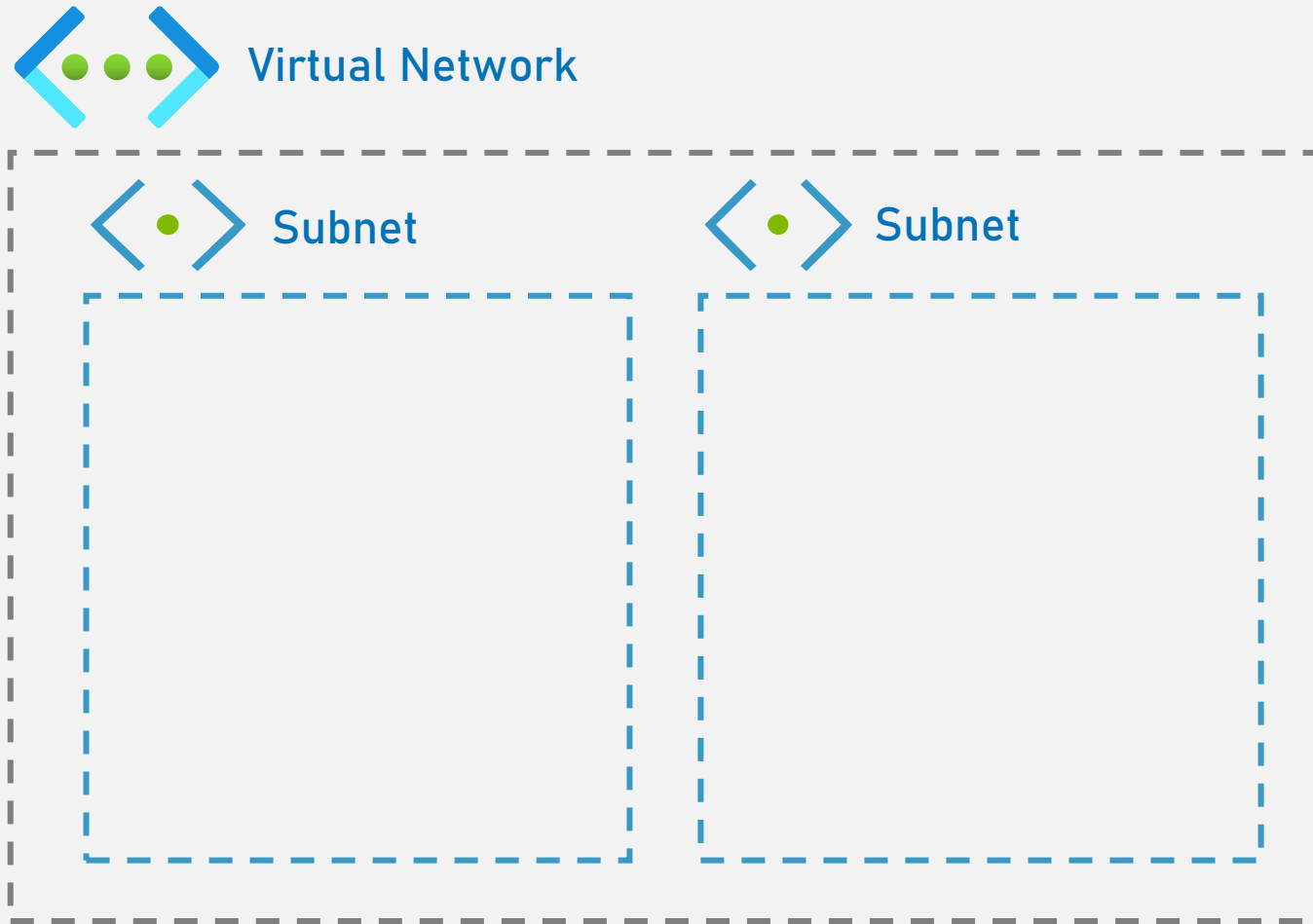
Governance / Process



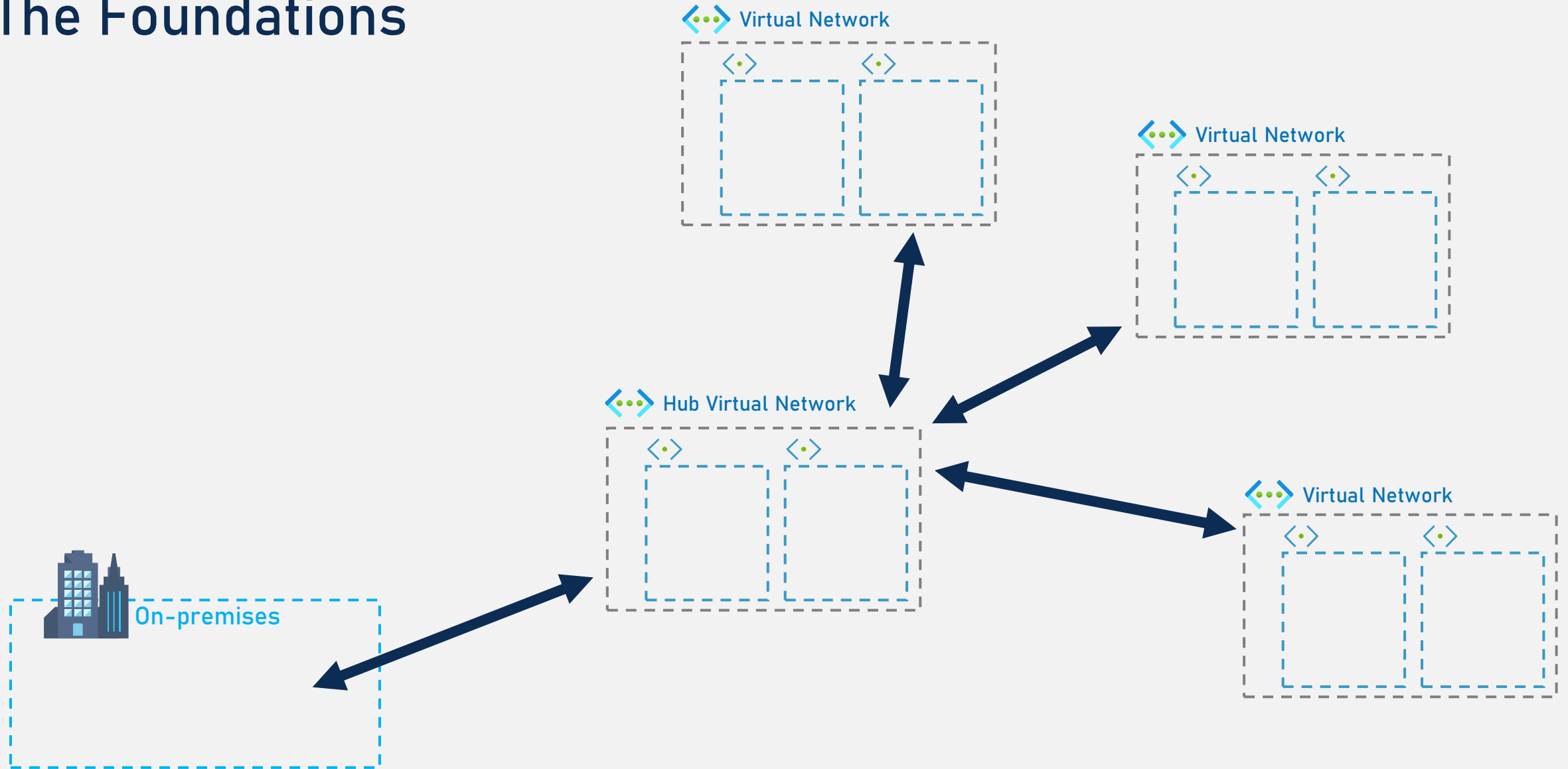
Data Encryption

# Networking

# The Foundations

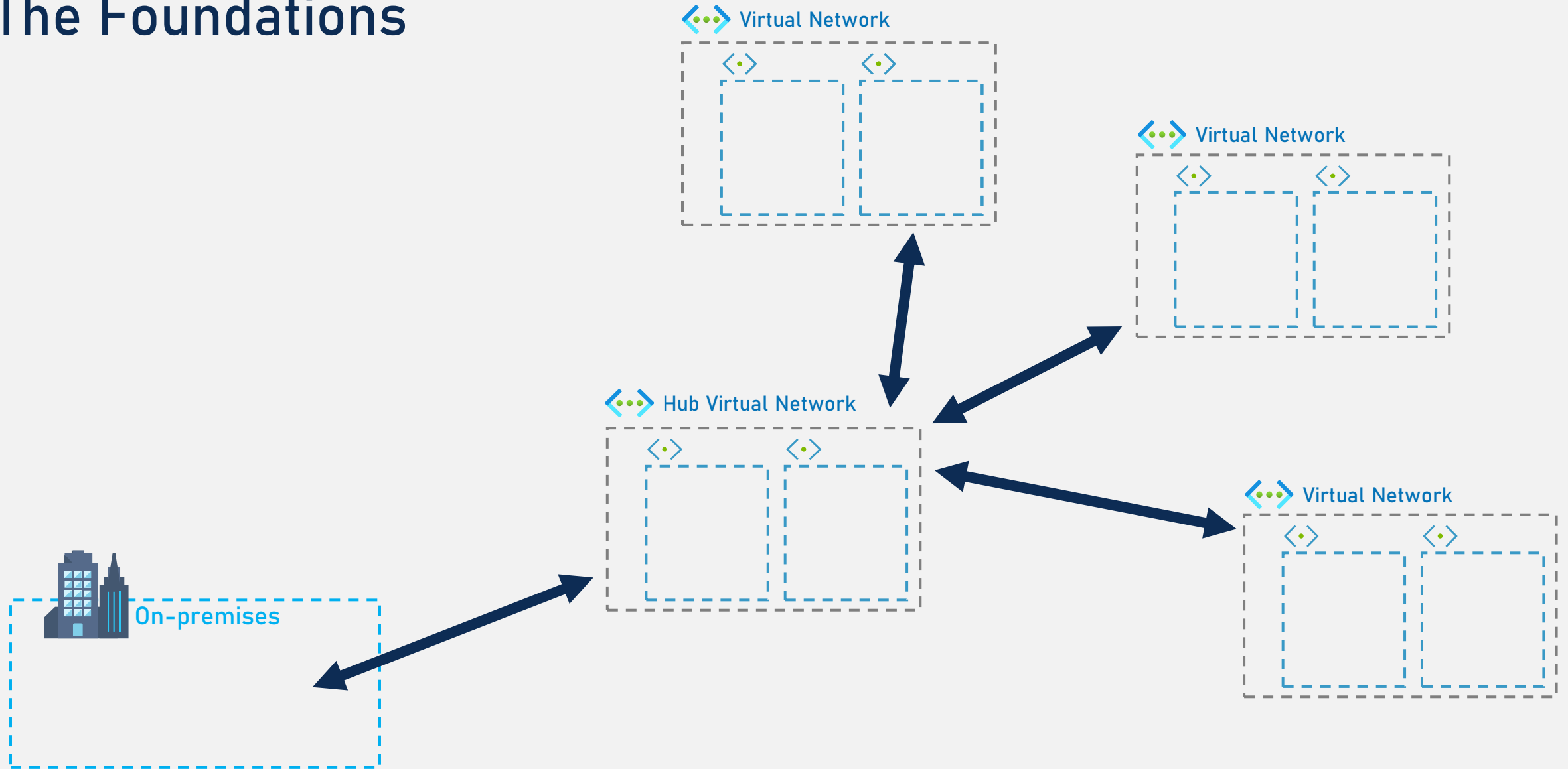


# The Foundations





# The Foundations



# Private Endpoints



 Virtual Network

 Subnet

10.X.X.X

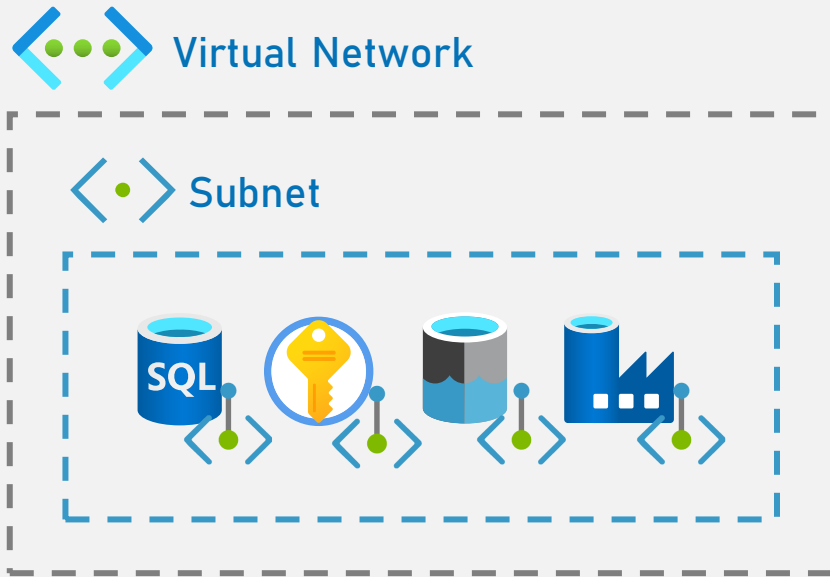
Private Link

SQL

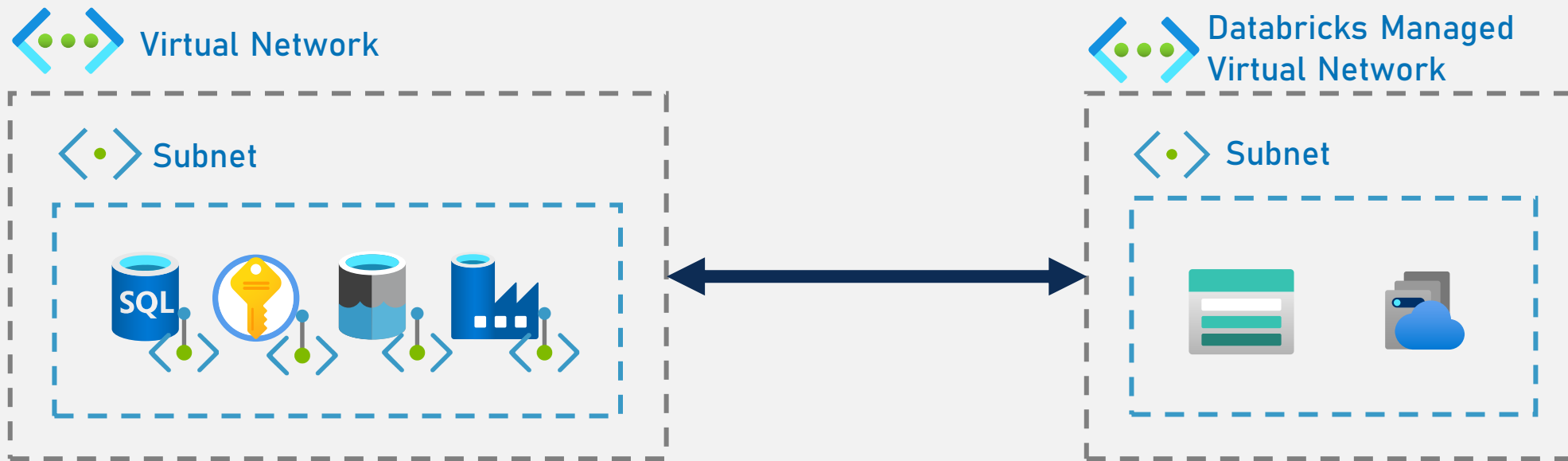
Azure

NOTE: Azure Private Link <> Azure Private Link [Service](#)

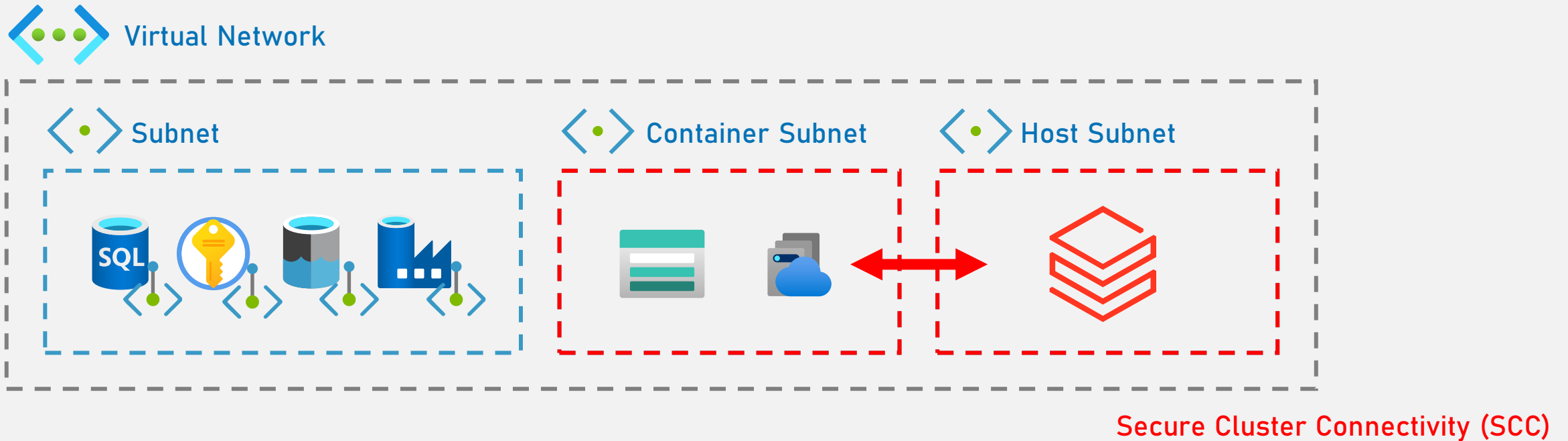
# Private Endpoints



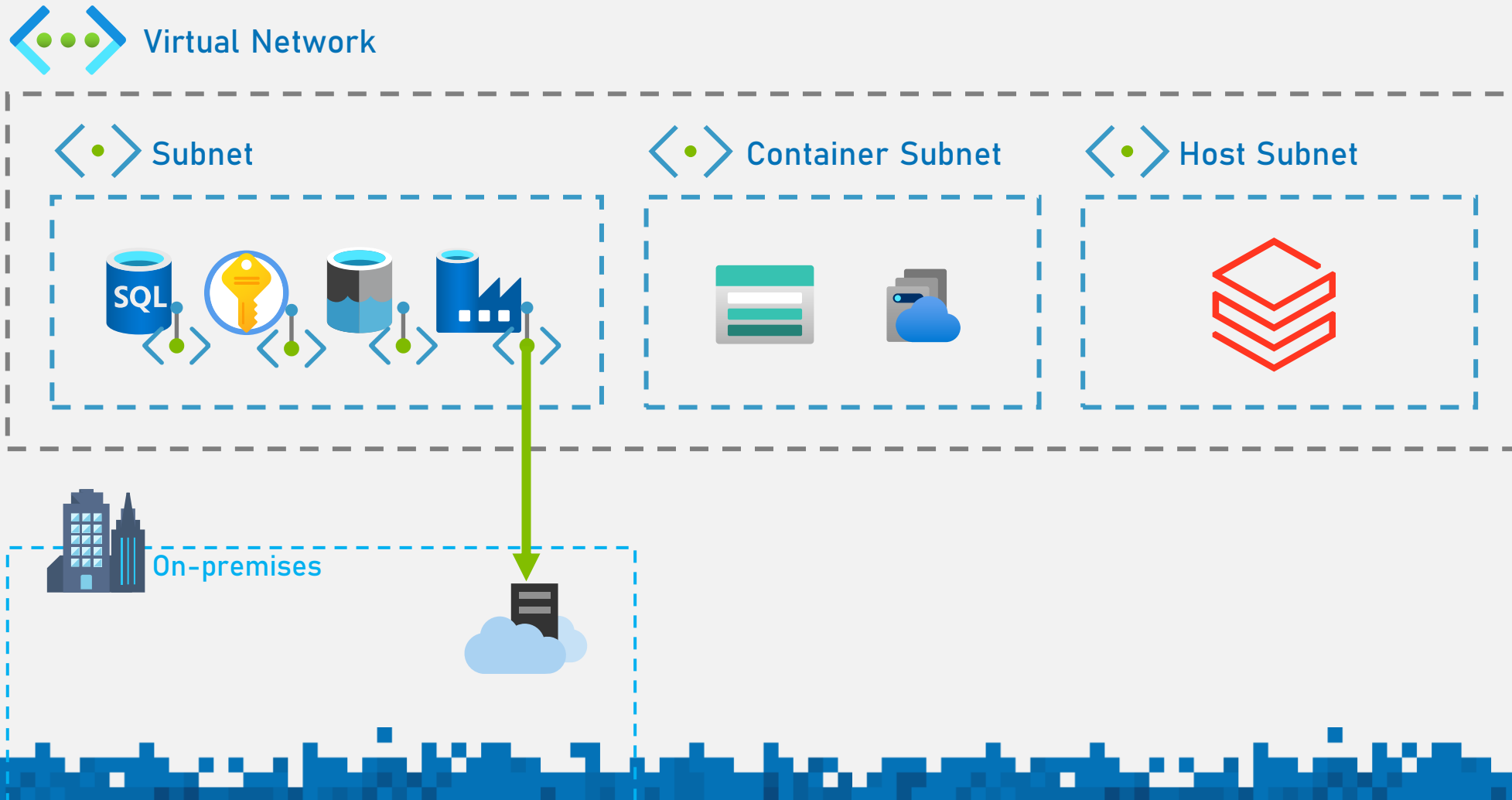
# Databricks VNet Peering



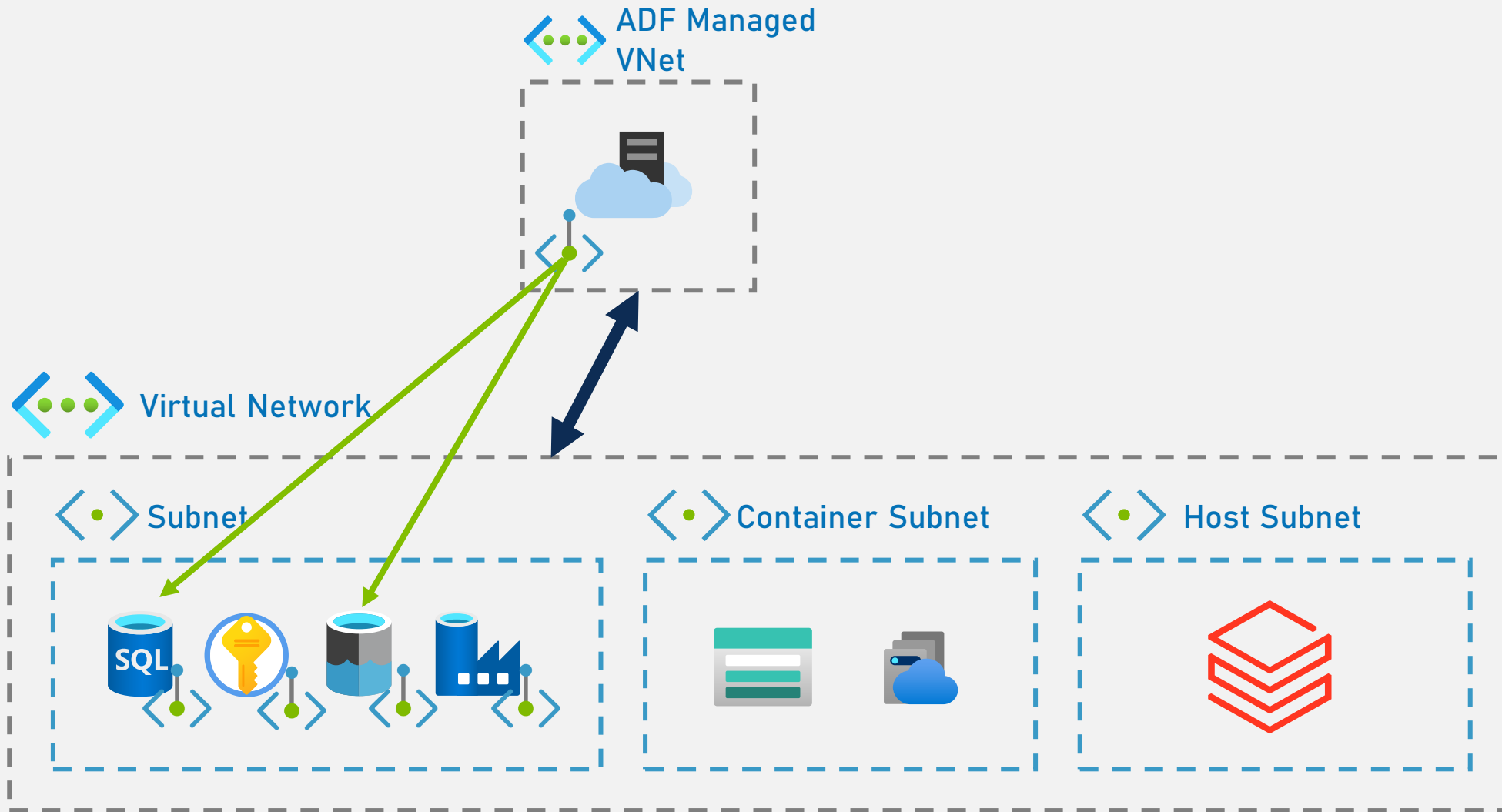
# Databricks VNet Injection



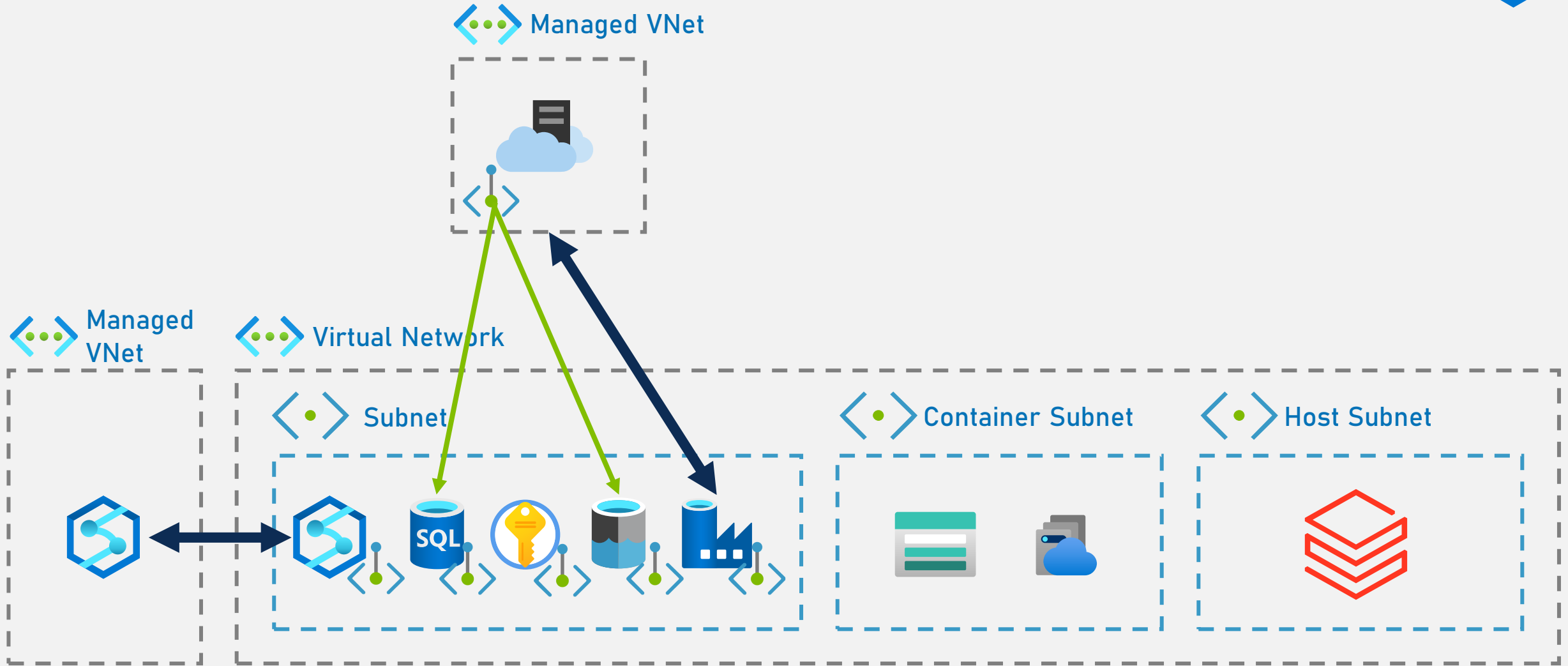
# Data Factory On-premises connectivity



# Data Factory

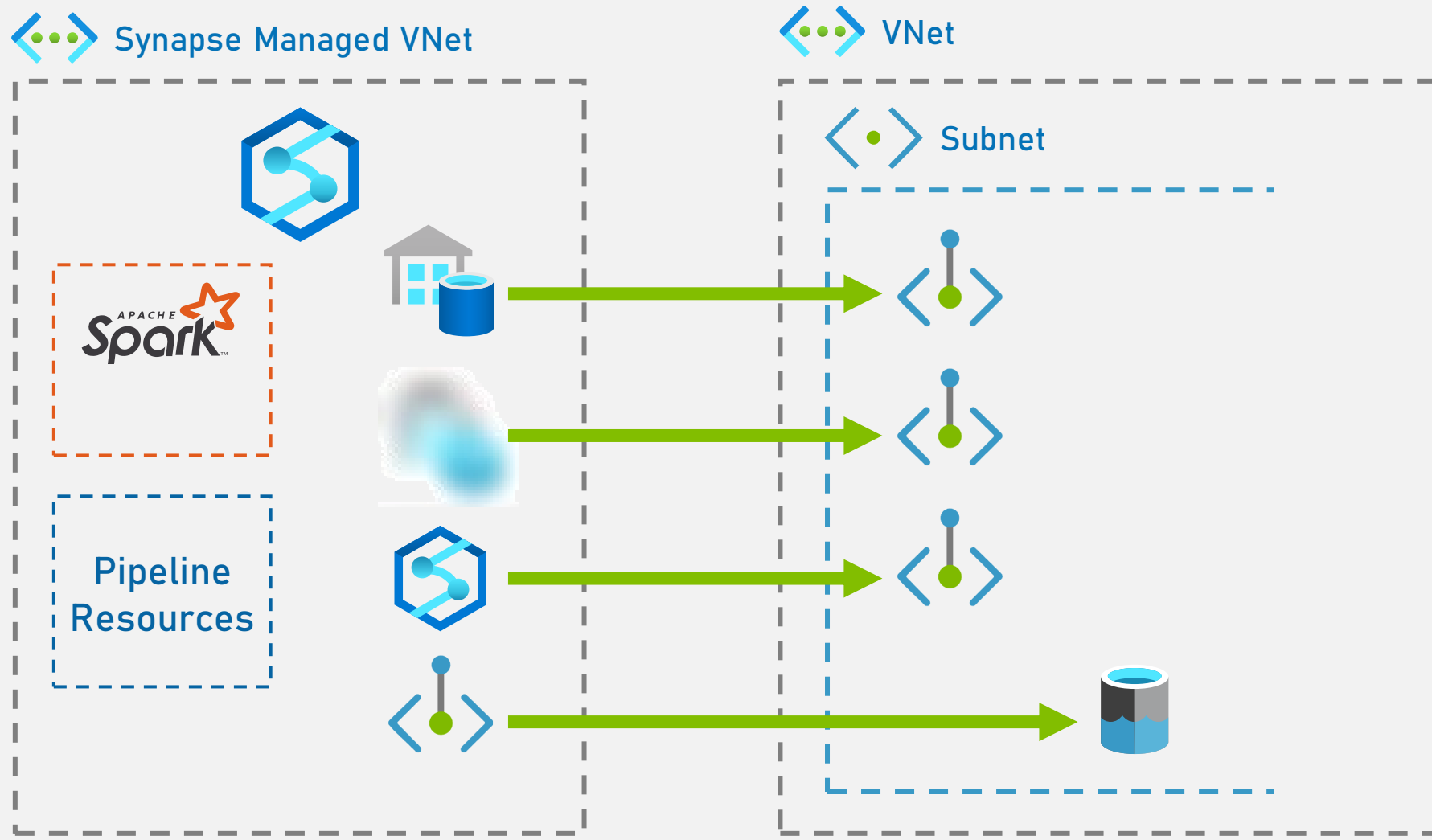


# What about Azure Synapse?





# What about Azure Synapse?



# Summary



## Virtual Network (VNet)



Traffic is privately secured within the organisation



Integrate with existing securely networked resources via Hub > Spoke



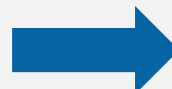
Microsoft-hosted Azure DevOps Agent



Self-hosted (IaaS)



Connectivity restrictions



Azure Bastion/Jump box (IaaS)



Increased costs



Don't underestimate complexity (NSGs, DNS zones, etc)

# Data Security Features

Just because you can,  
doesn't mean you should

# Security Features



Microsoft  
Managed  
Keys

Secure  
Cluster  
Connectivity

Premium  
Tier



Transparent  
Data  
Encryption

Azure SQL  
Auditing



Microsoft  
Managed  
Keys

Managed  
Identity

Managed  
Vnet on IR

SHIR to use  
Private  
Endpoint



Microsoft  
Managed  
Keys

No  
Anonymous  
Access

Disable Blob  
Public  
Access



Azure  
RBAC

Soft Delete

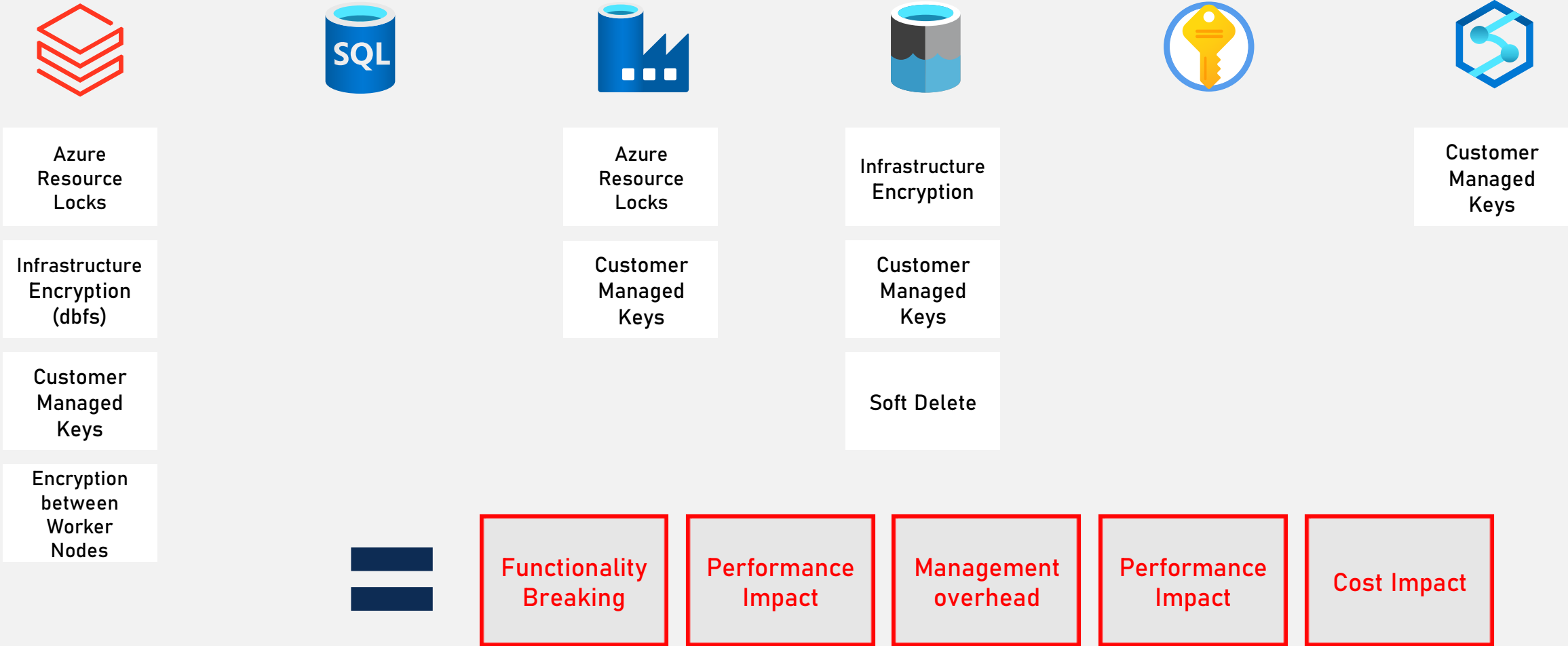


Microsoft  
Managed  
Keys

Azure SQL  
Auditing

Managed  
Vnet

# Use with Caution



# Other factors



Existing Azure  
Policies

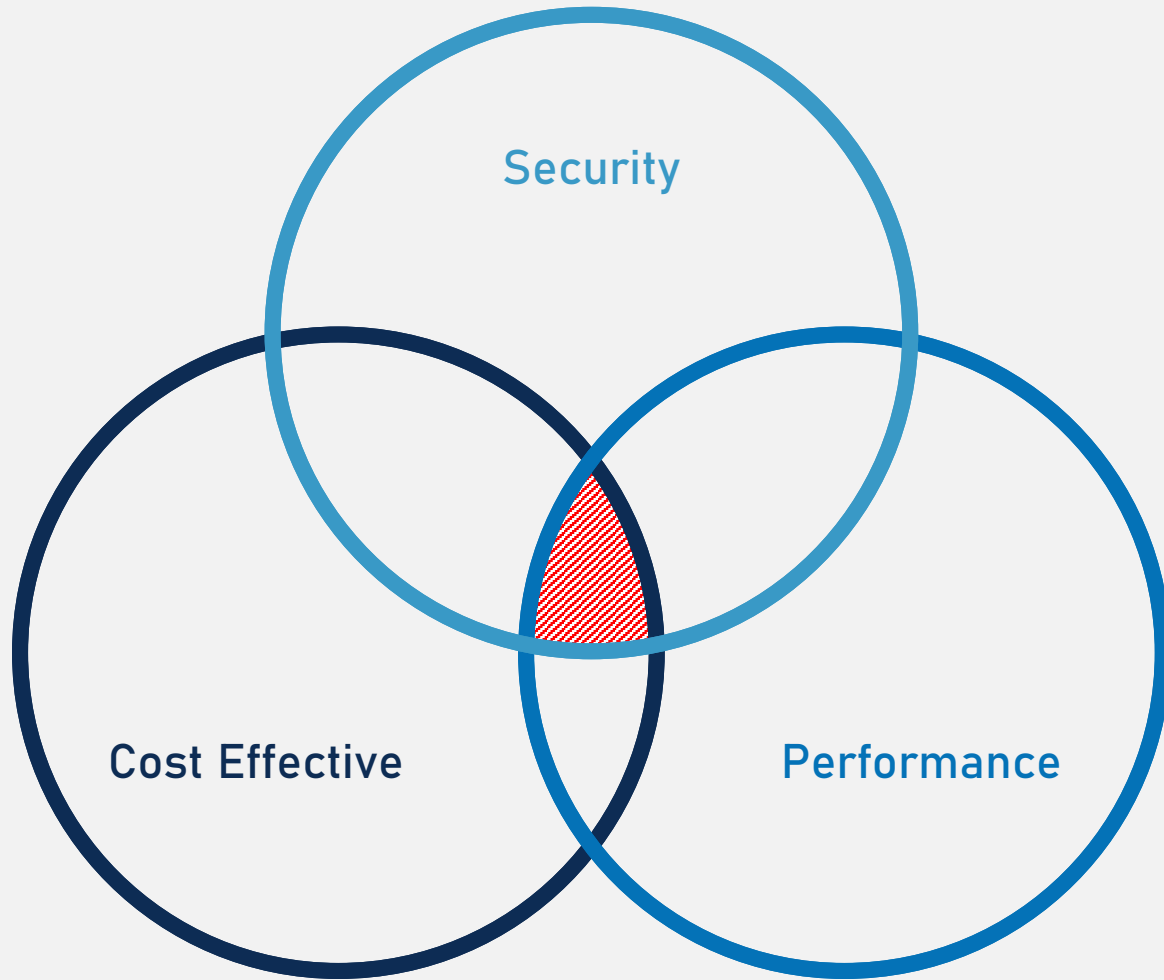


Some features  
don't make  
sense



Additional  
support  
overhead

# Summary

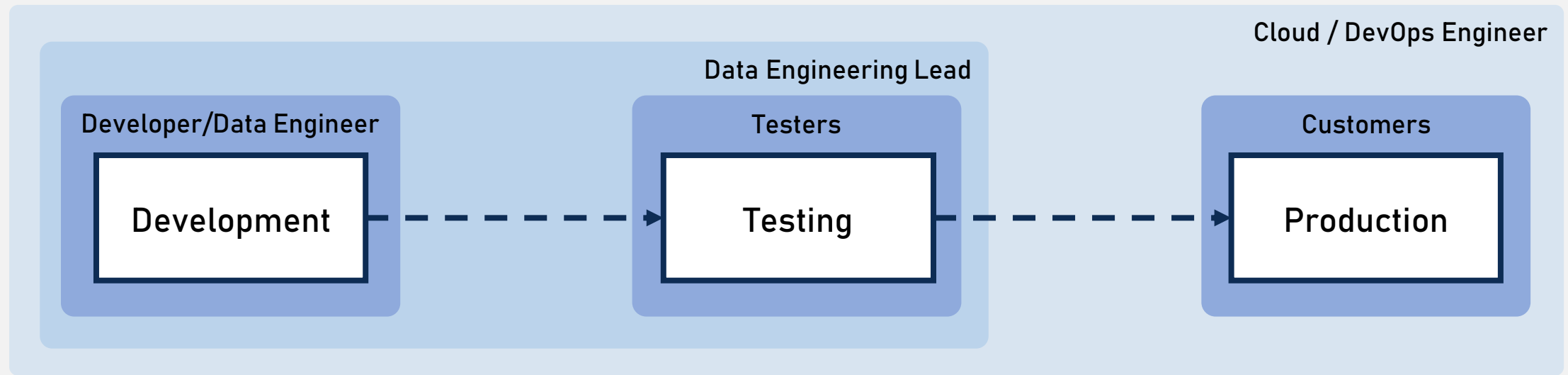


- Is there an established requirement?
- Does this deviate from best practices?
- Will this impact the business case for the platform?

# Data Governance

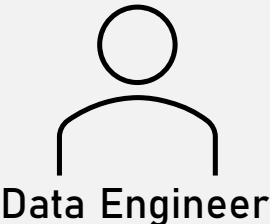














# Environment Security



# Data Security

 Read/Write access    Read access    No access



	Source	Raw Layer	Cleaned Layer	Presentation Layer
Dev				
Test				
Prod				

<https://craigporteous.com/mapping-data-personas-to-your-data-platform/>

# Data management



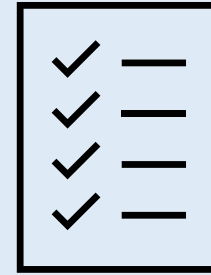
Azure Purview

- Classification of data
- Data ownership
- Management of PII Data

# InfoSec wish list



Azure  
Sentinel



CIS  
Benchmark



Design for  
Security



RBAC First



Defence in  
Depth



Don't just turn  
everything on



Don't  
underestimate  
complexity

evals.datagrillen.com

# Thank You



<https://craigporteous.com>



@cporteous



<https://github.com/cporteous>

# References

- [How to protect Data Exfiltration with Azure Databricks to help ensure Cloud Security](#)
- [Deploy Azure Databricks in your Azure virtual network \(VNet injection\) - Azure Databricks | Microsoft Docs](#)
- [Secure cluster connectivity \(No Public IP / NPIP\) - Azure Databricks | Microsoft Docs](#)
- [Azure Synapse Analytics security white paper: Network security - Azure Synapse Analytics | Microsoft Docs](#)
- [The Purpose and Pain of Azure Resource Locks - Craig Porteous](#)
- [Mapping Data Personas to your Data Platform - Craig Porteous](#)