Safe Haven Environment

High-Level Design Brief

December 2018

Prepared by

CORINIUMTECHNOLOGY

Contents

Overview

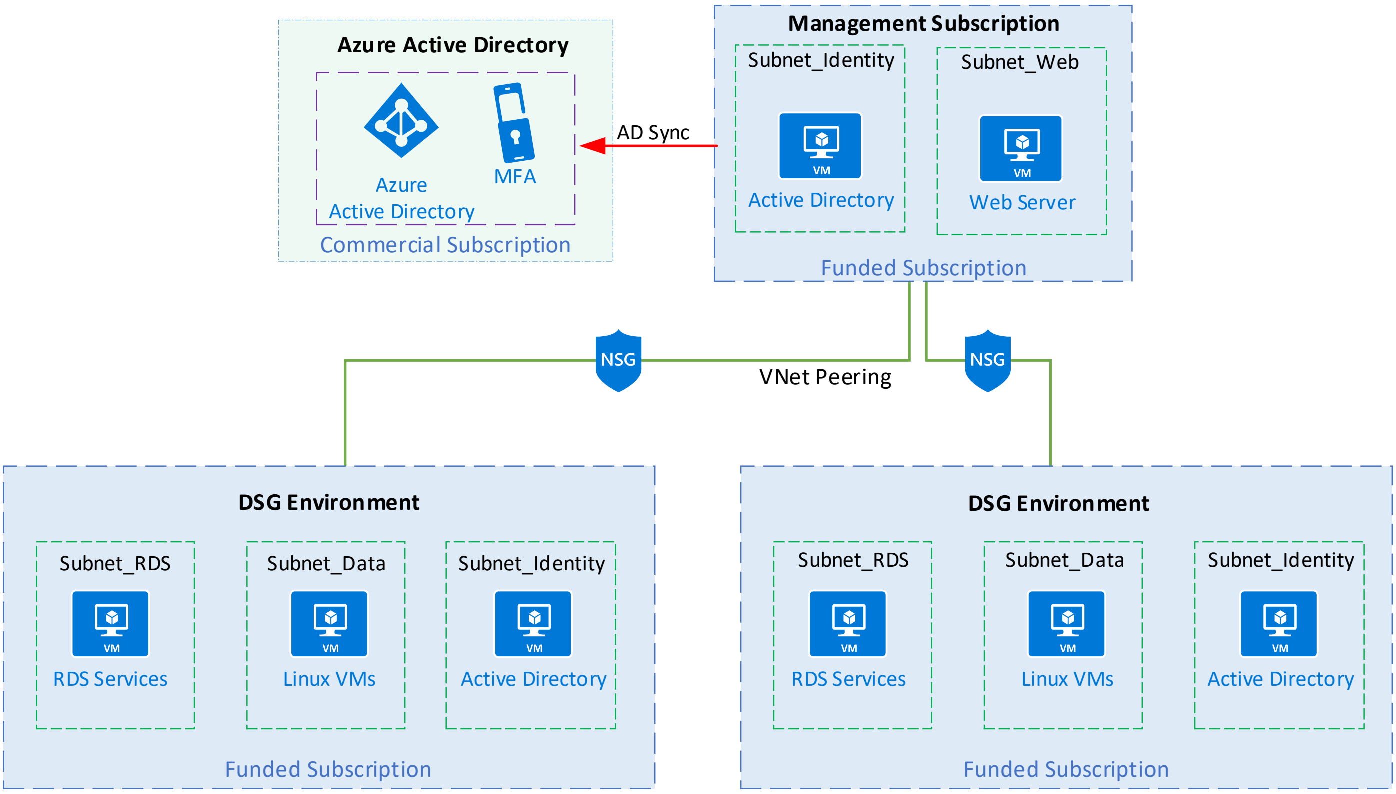
The Safe Haven Environment is designed and deployed into Microsoft Azure utilizing Azure networking and VM infrastructure. This document covers the topology design of the environment.

The Safe Haven is deployed using a dedicated Azure subscription that hosts the “management” services. These include the primary authentication providers, VM images and other services that contribute to the smooth running of the service.

In addition to this there are separate Data Study Group (DSG) subscriptions, these host the infrastructure required for running a DSG i.e. remote desktop services, customised VMs with data tools/utilities, collaboration systems (HackMD, Gitlab). The purpose of the separate subscriptions is to provide a management boundary and the ability to destroy an environment should rebuild/redeployment be required. External researchers connect to the DSG via a Microsoft Remote Desktop Service that is present in each DSG.

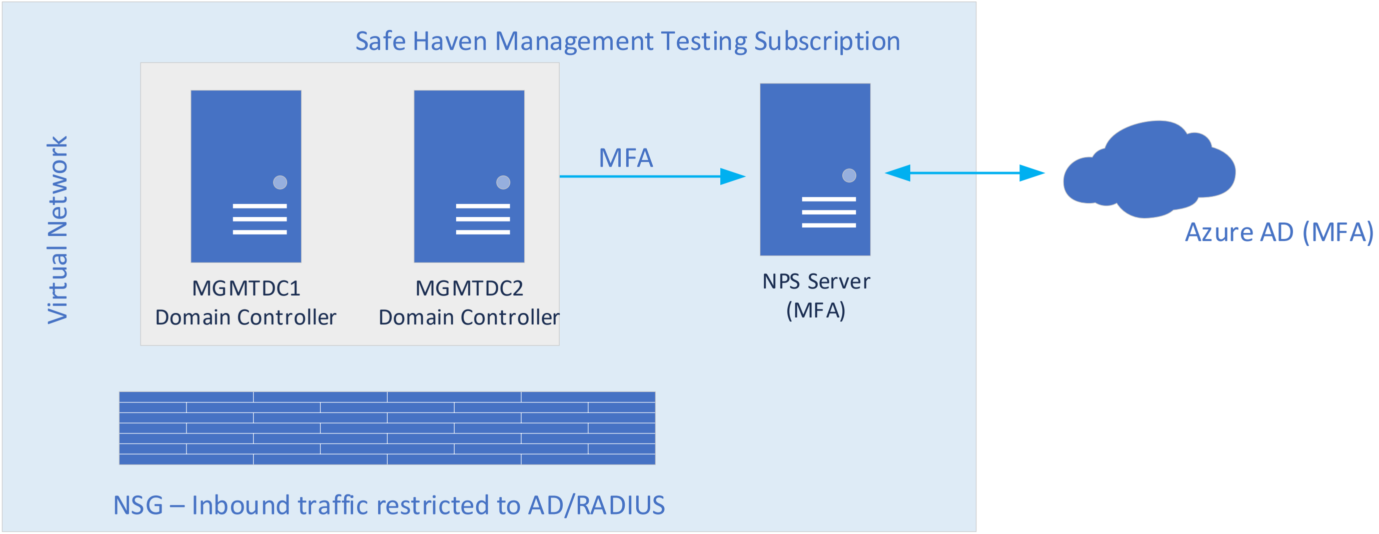
High Level Topology

The “Management” environment hosts the authentication providers for the infrastructure. The identity provider is Microsoft Active Directory, this in turn is synchronised with AzureAD to provide cloud and multifactor authentication into the Data Study Group (DSG) environments.



The Management environment is connected to the DSG environments using Azure Peer Network connections. This connection passes authentication traffic between the servers in the DSG to AD servers within the Management subscription. There is no connectivity between DSGs directly.

Management Environment

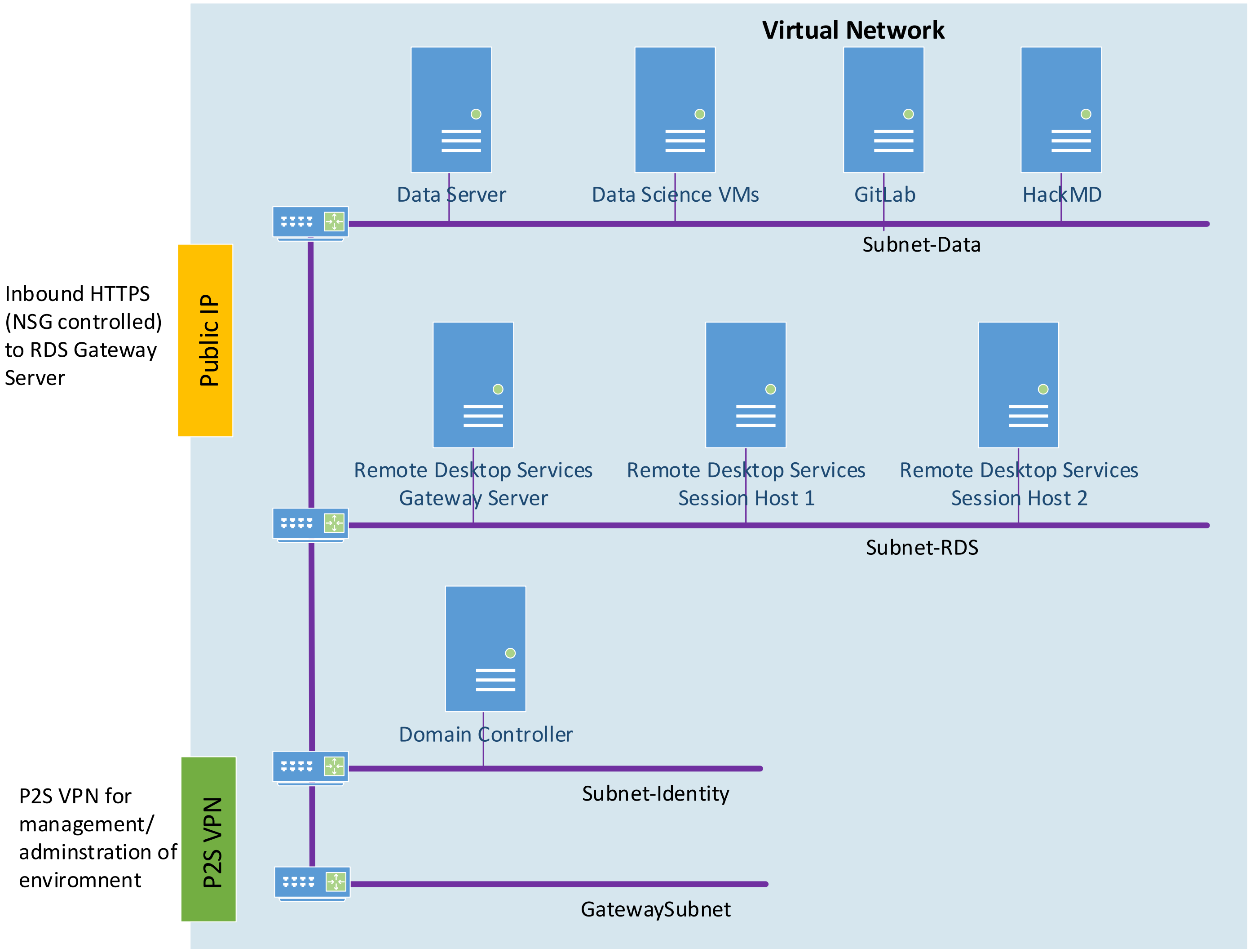


Within the Management environment there is a single virtual network that all authentication services are provisioned. The Windows Servers are running Active Directory and are acting as Domain Controllers. They are configured within an Azure availability set to ensure maximum up time.

In addition to the Domain Controllers there is a Windows Network Policy server, this server provides Multifactor Authentication services to the Remote Desktop Servers hosted within the DSG environments. The NPS server is also running Azure AD Connect to synchronise the Researchers user IDs into the AzureAD that is associated with the Management subscription.

Network security is provided by Azure Network Security Groups that ensure that inbound connections are limited to Active Directory and RADIUS traffic. This environment is not accessible from the internet.

Data Study Group Environments

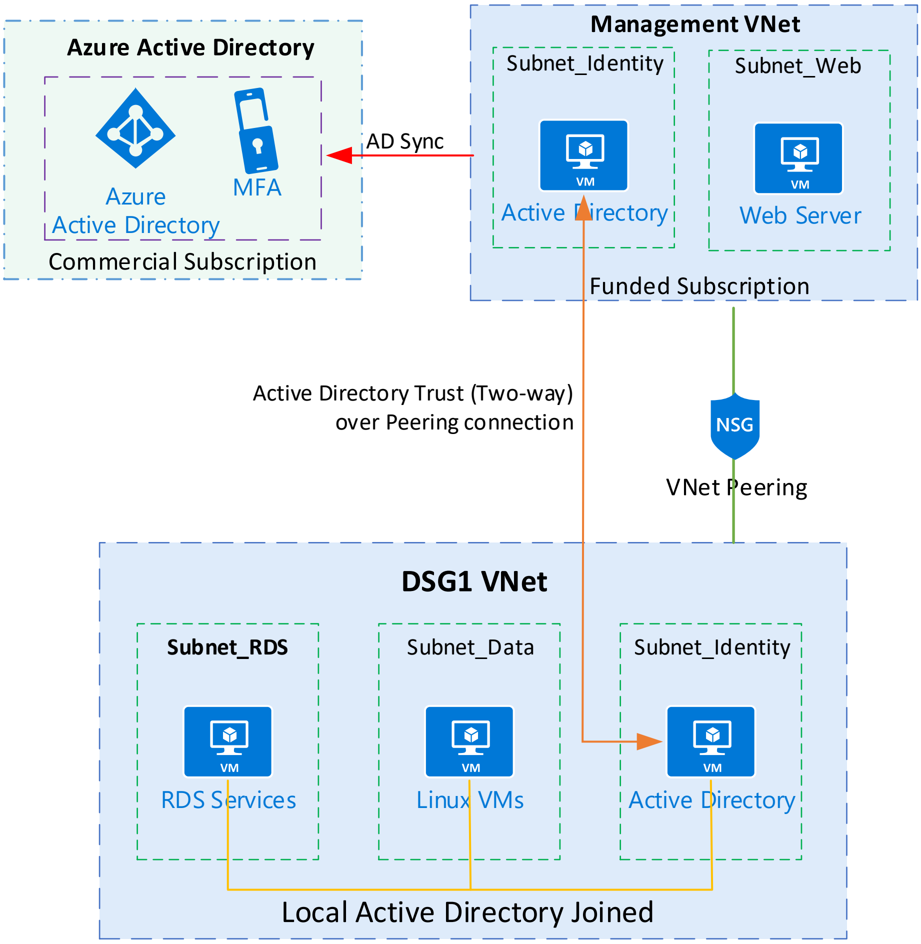

Virtual Network
Data Server
Data Science VMs
GitLab
HackMD
Subnet-Data
Inbound HTTPS (NSG controlled) to RDS Gateway Server
Public IP
Remote Desktop Services Gateway Server
Remote Desktop Services Session Host 1
Remote Desktop Services Session Host 2
Subnet-RDS
Domain Controller
Subnet-Identity
P2S VPN for management/ adminstration of environment
P2S VPN
GatewaySubnet

The DSG environments use Windows Remote Desktop services to provide a secure connection to the data science VMs within the environment. RDS prevents data bleed by restricting what users can copy in/out of the environment. On the higher tier environment access to the Internet is also blocked adding another layer of security.

The only way into the DSG environment from the Internet is via the RDS connection broker, this is a SSL secured connection that requires the user to authenticate using a Turing provided username and validated with MFA.

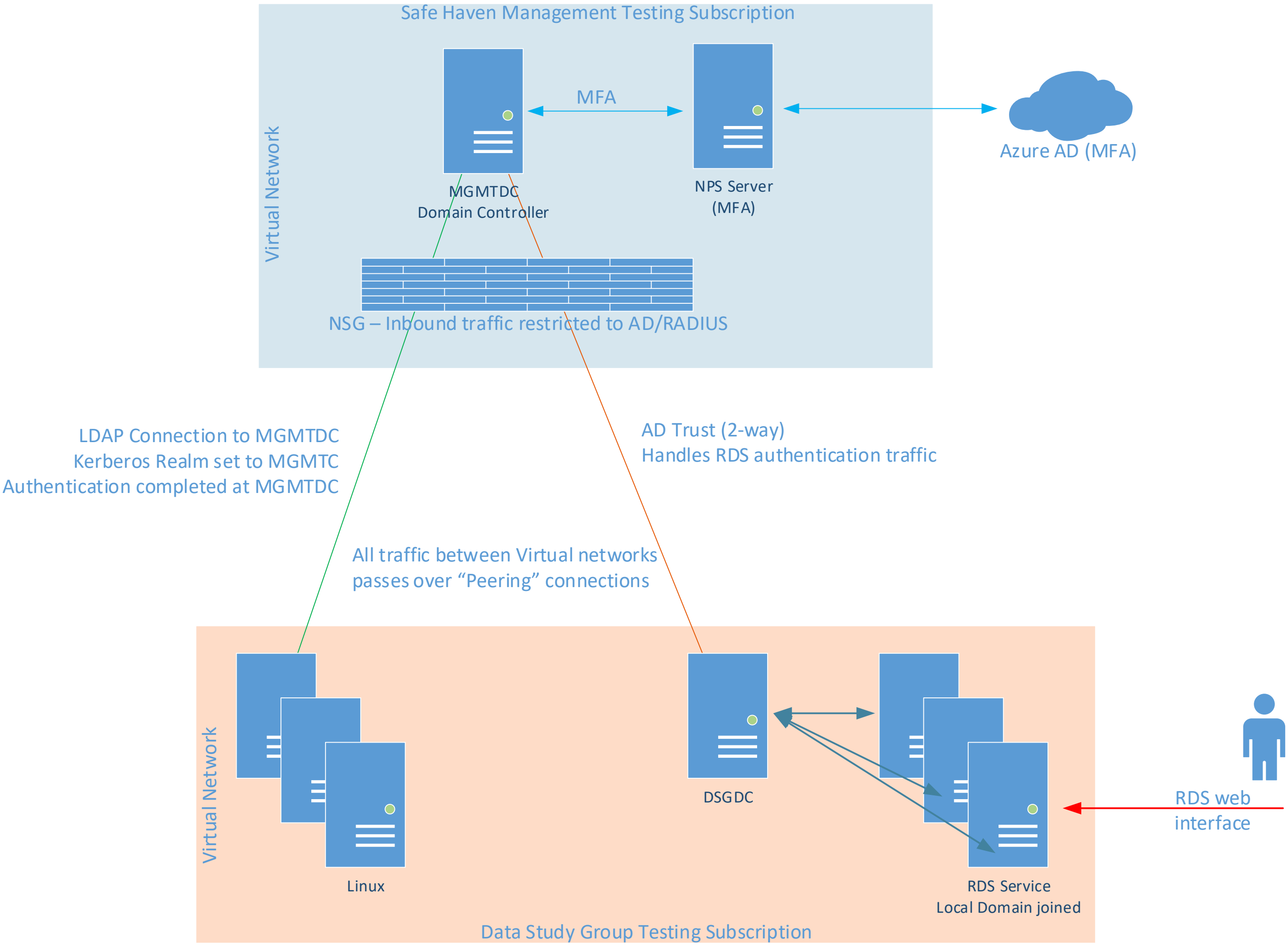
For management of the environment there is a P2S VPN service configured, access to this is limited to the Turing IT team only.

Authentication



Each of the DSG environment has a local Active Directory that is used for management of the RDS service. This local Active Directory domain has a Trust with the Active Directory domain within the Management environment. User accounts are created in the Management AD and added to security groups. These security groups are then applied to the RDS servers in the DSGs. This provides a central user management experience for the support staff and reduces the overhead of removing old users from expired DSG events.

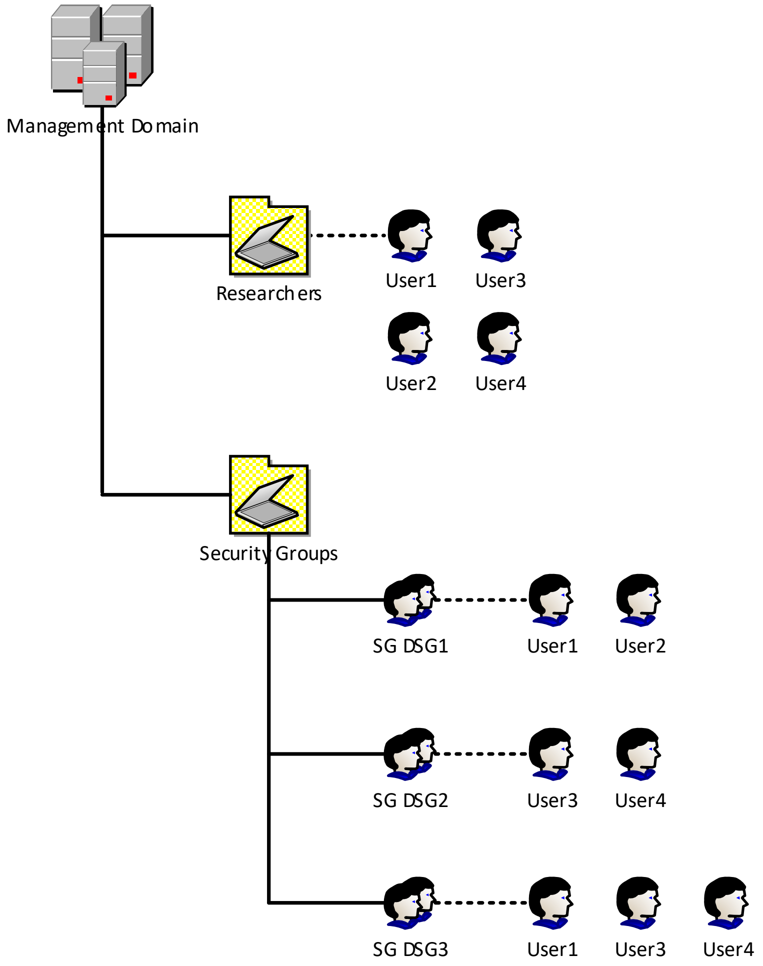
Topology Overview


Safe Haven Management Testing Subscription
Virtual Network
MFA
MGMTDC
Domain Controller
NPS Server
(MFA)
Azure AD (MFA)
NSG – Inbound traffic restricted to AD/RADIUS
LDAP Connection to MGMTDC
Kerberos Realm set to MGMTC
Authentication completed at MGMTDC
AD Trust (2-way)
Handles RDS authentication traffic
All traffic between Virtual networks
passes over "Peering" connections
Virtual Network
Linux
DSGDC
RDS Service
Local Domain joined
RDS web
interface
Data Study Group Testing Subscription

The above diagram shows the authentication paths between the Management and DSG environments. The Windows based servers use the Active Directory trust to validate users. Due to the limitations of Linux LDAP providers the Linux based server talk directly to the Management domain to authenticate users.

Active Directory Structure

The control of the users is performed by using both Active Directory organisational units (OUs) and security groups. No user accounts other than service accounts need be created on the active directories within the DSG environment.



The researchers OU contains all the researcher's user accounts, these accounts are automatically synchronised with Azure Active Directory ready for MFA enablement.

The security groups are used to control what access the user have to what environments. There is a security group for each DSG that is present/created, this security group is then populated with the users that need access to the DSG.