



Geo Redundant VPN in Azure

Arjen Gerritsen – Emergo IT

Demonstration Objectives

- Geo Redundant VPN, based on BGP routing
- Over 2 Azure Regions and 3 sites:
 - Head Quarters (HQ), On Premise Network, VPN enabled
 - West Europe, a hub-and-spoke configuration in The Netherlands
 - North Europe, a hub-and-spoke configuration in Ireland
- BGP Routing behaviour
- Setup monitoring for BGP Peers

Source available: <https://github.com/aggerritsen/georedundantvpn>

Classless Inter-Domain Routing (CIDR)

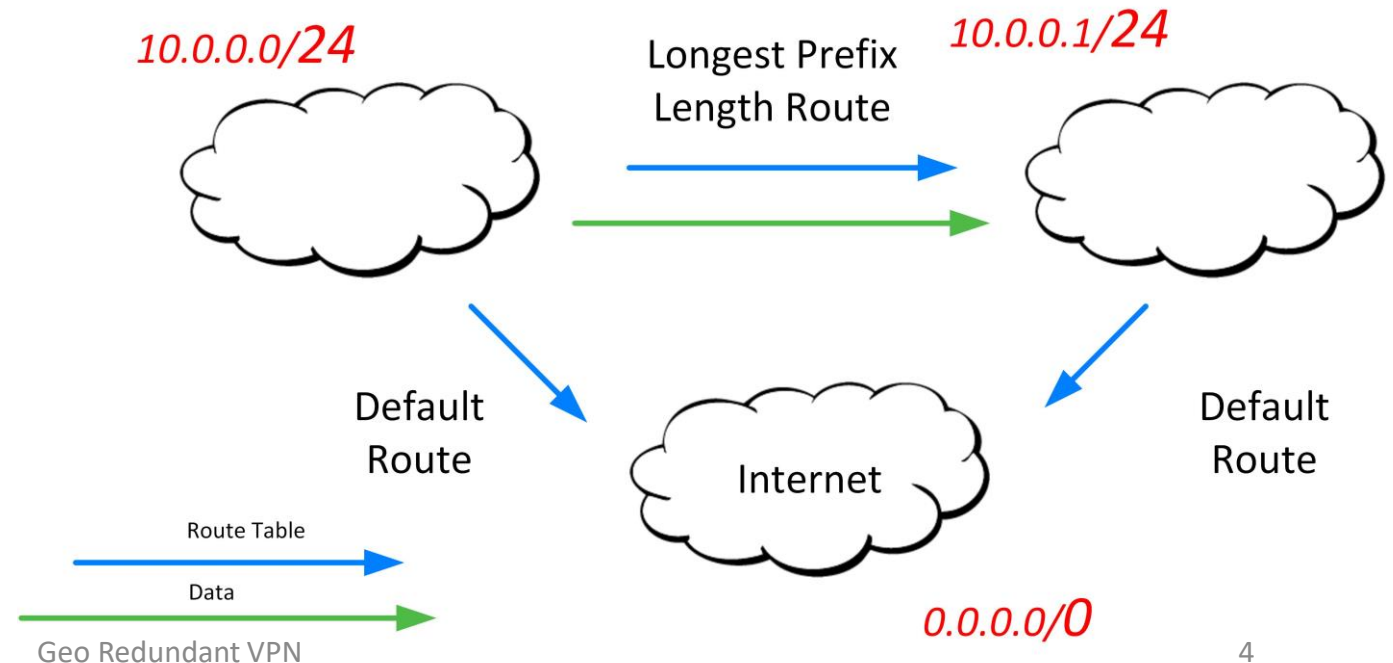
- IP addresses are described as consisting of two groups of bits in the address: the most significant bits are the **network prefix**, which identifies a whole network or subnet,
 - eg. **10.0.0.0/24**
defines network with 256 IP addresses between 10.0.0.1 and 10.0.0.254.
 - The /24 is the Network Prefix
 - See <https://cidr.xyz> for range calculation

Longest Prefix Length Routing Order

- This method defines a routing order based on :
*"When a particular destination IP address matches more than one route in a router's routing table, the router uses the most specific route - in other words, the route with the **longest prefix length**."*

Example:

a route defined for **10.0.0.0/24**
gets priority over
a *default route* defined
for **0.0.0.0/0**

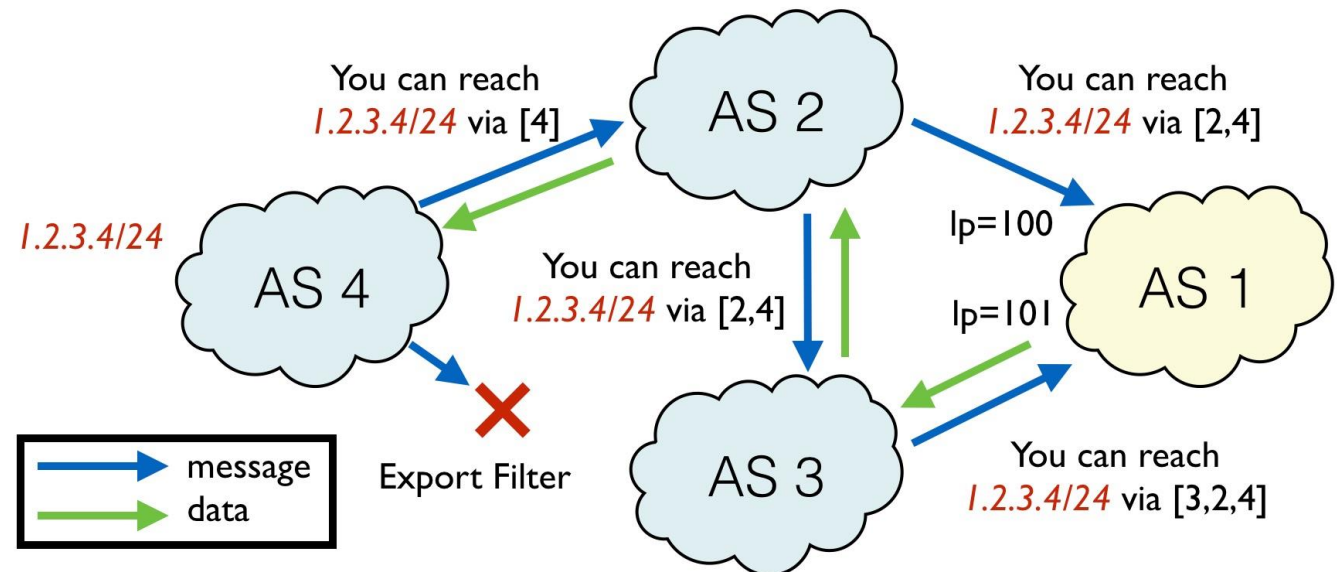


Azure Routing Priority

- Longest Prefix Length gets evaluated.
- If multiple routes contain the same address prefix, Azure selects the route type, based on the following priority:
 1. User-defined route
 2. BGP route
 3. System route
- System Routes are derived from VNets, VNet Peerings, Local Network Gateway Connections, Azure Service Endpoints and other Azure Backbone routes like Internet.

Border Gateway Protocol

- Border Gateway Protocol (**BGP**) is a standardized exterior gateway protocol designed to exchange routing and reachability information between autonomous systems (AS) on the Internet.



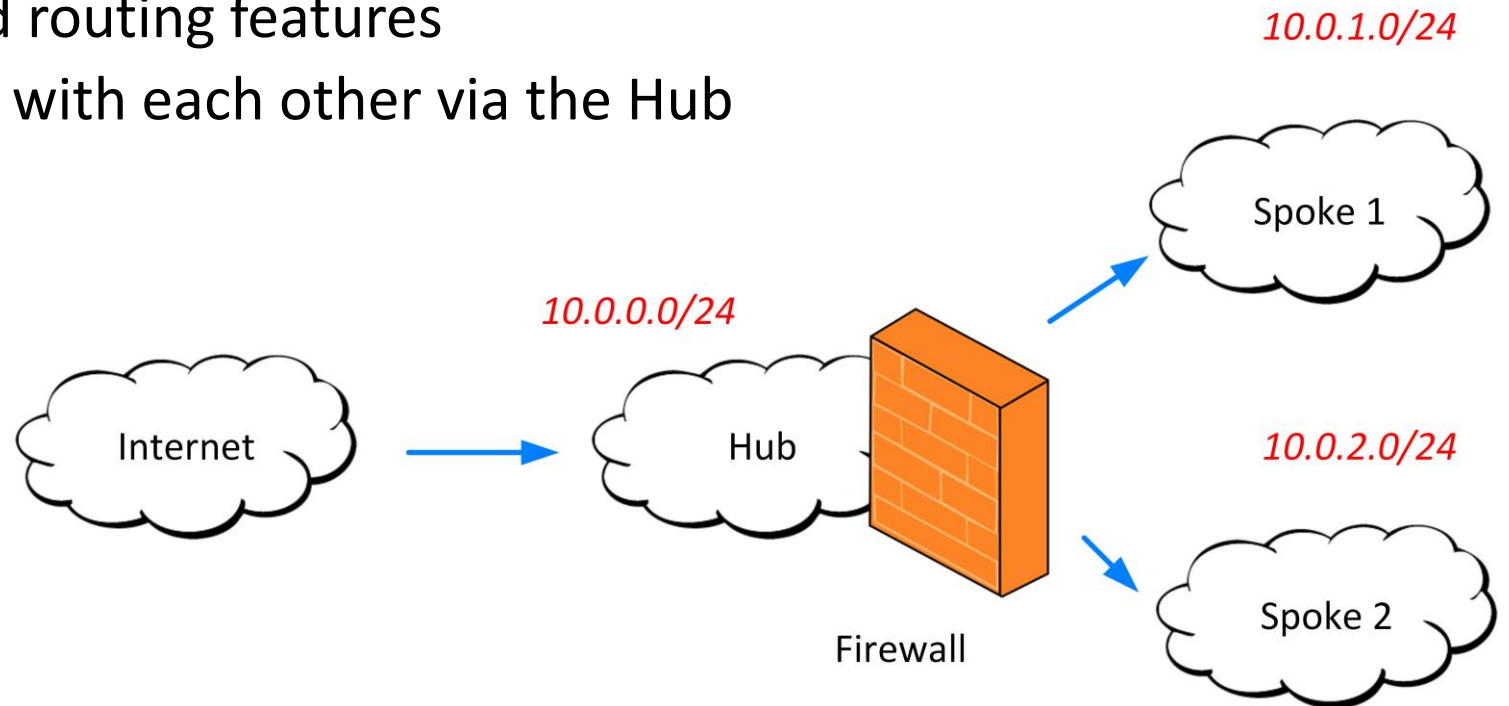
Azure Route Tables

- When you exchange routes with Azure using BGP, a separate route is added to the route table of all subnets in a virtual network for each advertised prefix.
- The route is added with *Virtual network gateway* listed as the source and next hop type.

See <https://docs.microsoft.com/en-us/azure/virtual-network/virtual-networks-udr-overview>

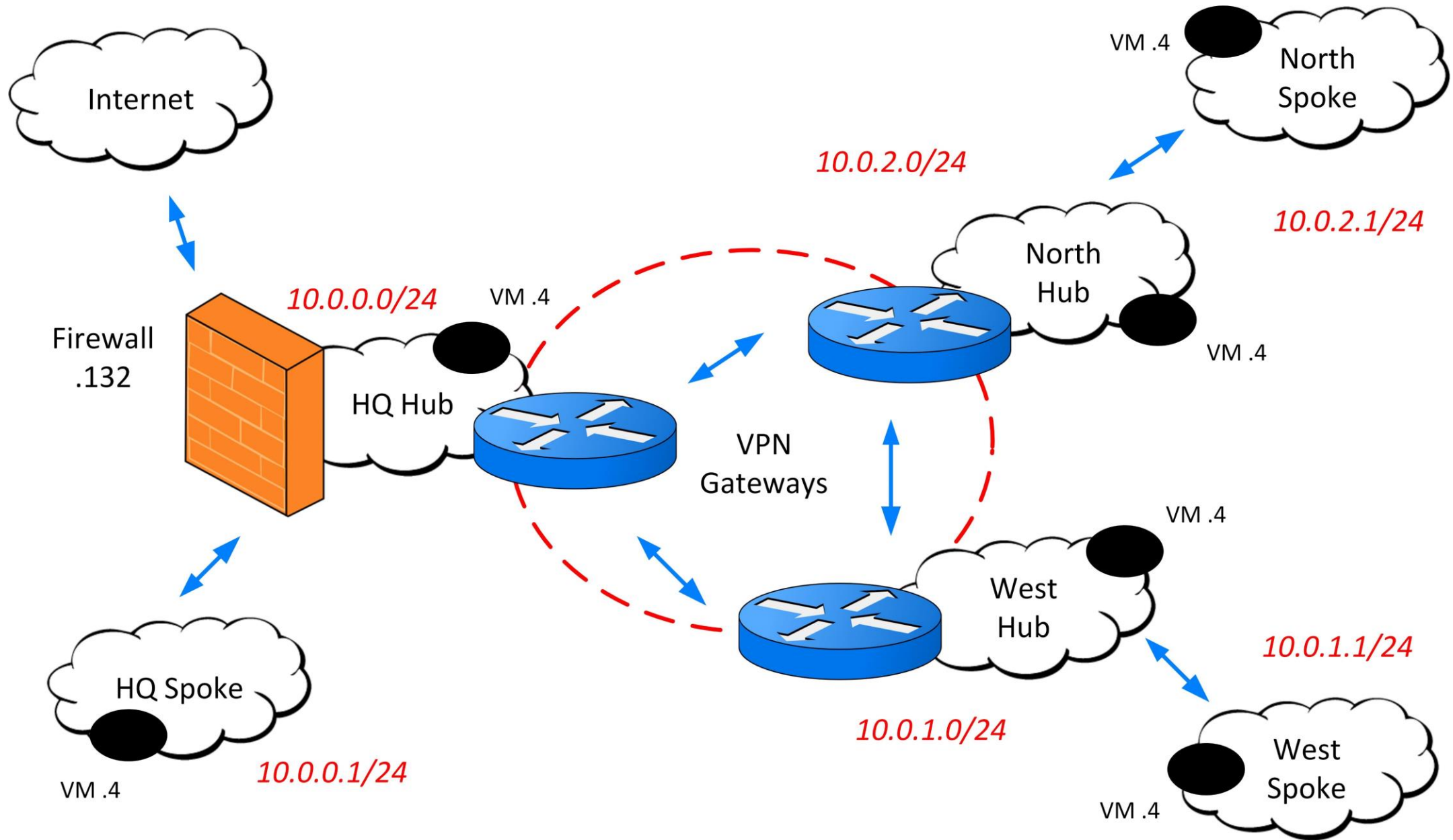
Hub-and-Spoke topology

- Hub-and-Spoke networks are a common architecture pattern in cloud
 - A Hub Network acts as 'edge' for a region or tenant
 - Hubs are used to connect other regions/tenants, providing firewall and routing features
 - Spokes communicate with each other via the Hub



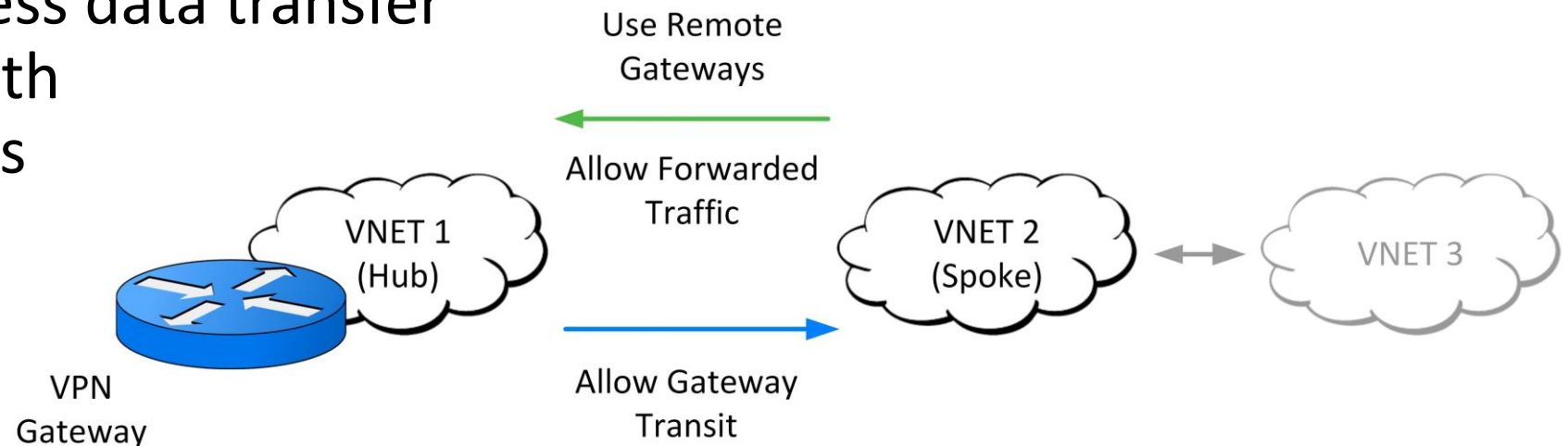
Demo environment

- We are going to use 6 Virtual Network
 - HQ-Hub: 10.0.0.0/24, HQ-Spoke: 10.0.1.0/24, with firewall setup
 - West-Hub: 10.1.0.0/24, West-Spoke: 10.1.1.0/24
 - North-Hub: 10.2.0.0/24, North-Spoke: 10.2.1.0/24
- Having each 1 or 2 subnets:
 - DefaultSubnet x.x.x.x./26, eg. 10.0.0.0/26 (64 IP addresses)
 - GatewaySubnet x.x.x.64/26, eg. 10.0.0.64/26 (64 IP addresses)
- Each subnet having a Virtual Machine
 - Jumphost-[XXX]-vm, with Internal and Public IP address, eg. JumphostHub-HQ-vm, 10.0.0.4 and 51.124.79.159



Virtual Network Peering

- Routing between two Virtual Network in the same region
 - Peering enables the use of the peered VNet's gateways
 - Allow or disallows forwarding from other peered networks
- Exist of a set of 2 peer connections
- Ingress and egress data transfer cost apply to both Virtual Networks



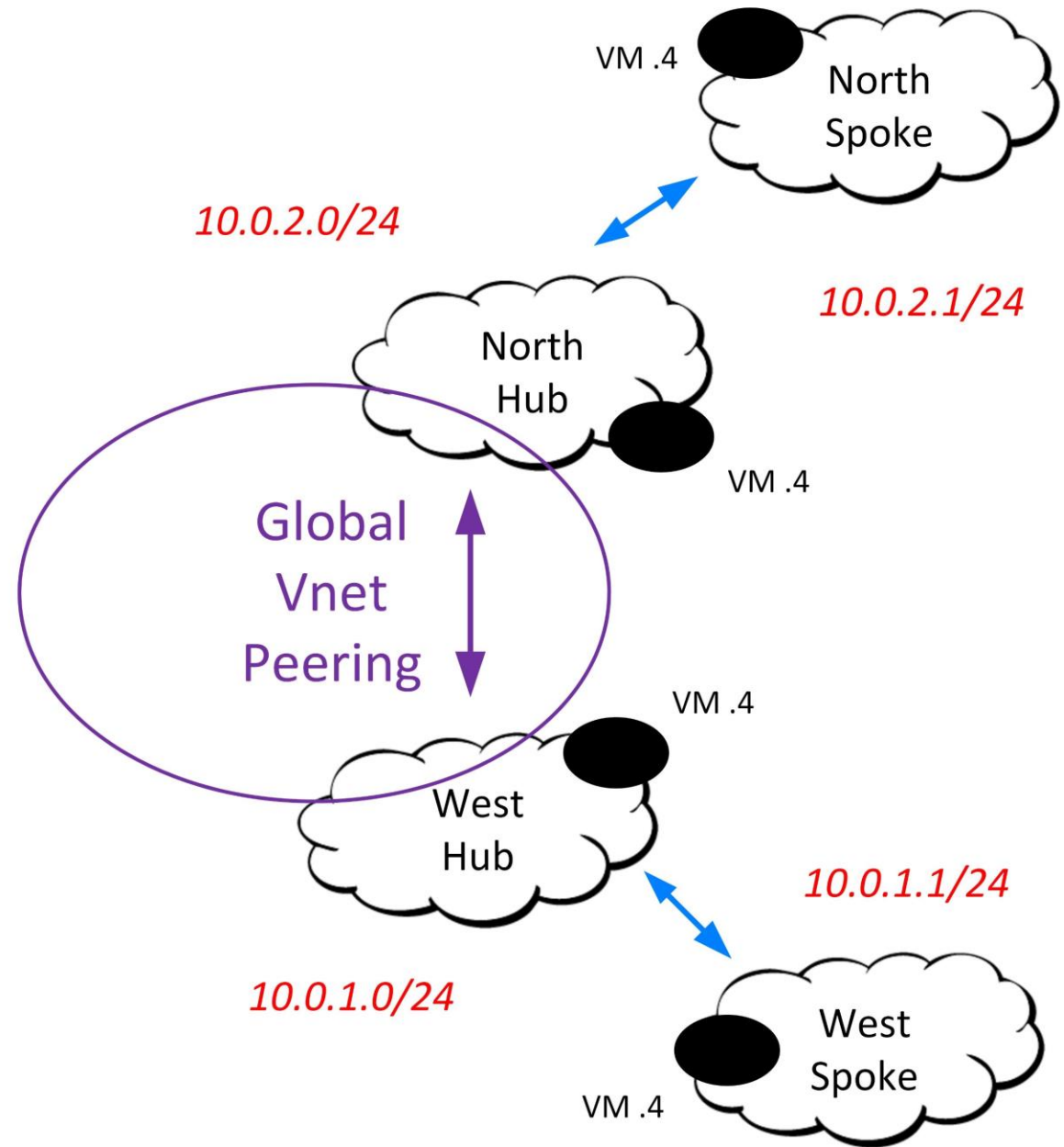
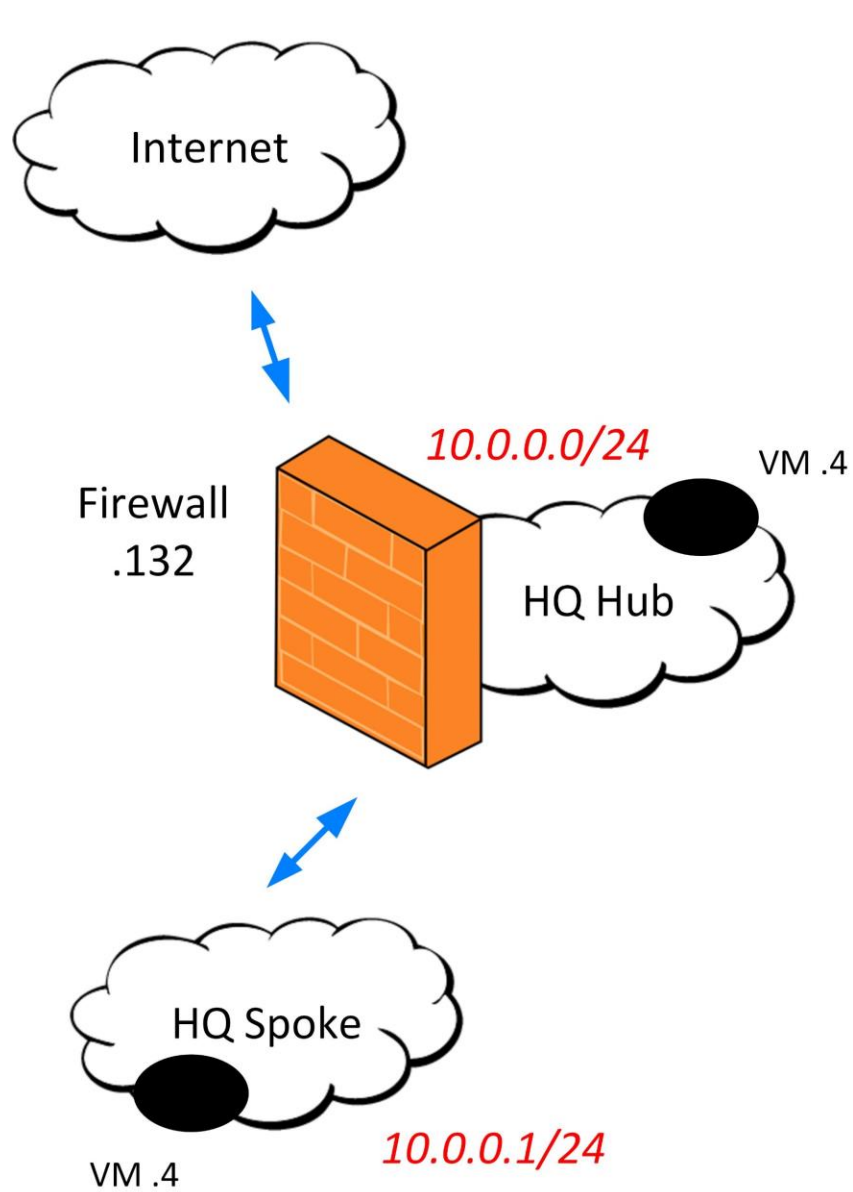
Options for connecting two regions

- Global VNet Peering
 - Runs over Azure backbone network
 - For peering two regions, use Global Vnet Peering
 - Peering two Azure Tenants, two accounts with permissions are required
- VNet-to-VNet Virtual Network Gateway
 - Runs over Azure backbone network
- Site-to-Site (IPSEC) Virtual Network Gateways
 - Runs over Internet
- Express Routes (not covered in this demonstration)
 - Runs over leased/dedicated lines, through listed service providers

Global VNet Peering

- Exist of a set of 2 connections
- Azure Backbone ensures connection with low and stable latency
- Ingress and Egress Data transfer costs are billed to each site
- Routing solely based on connected Virtual Networks, not on the spokes connected to the hub.
- Does not support BGP

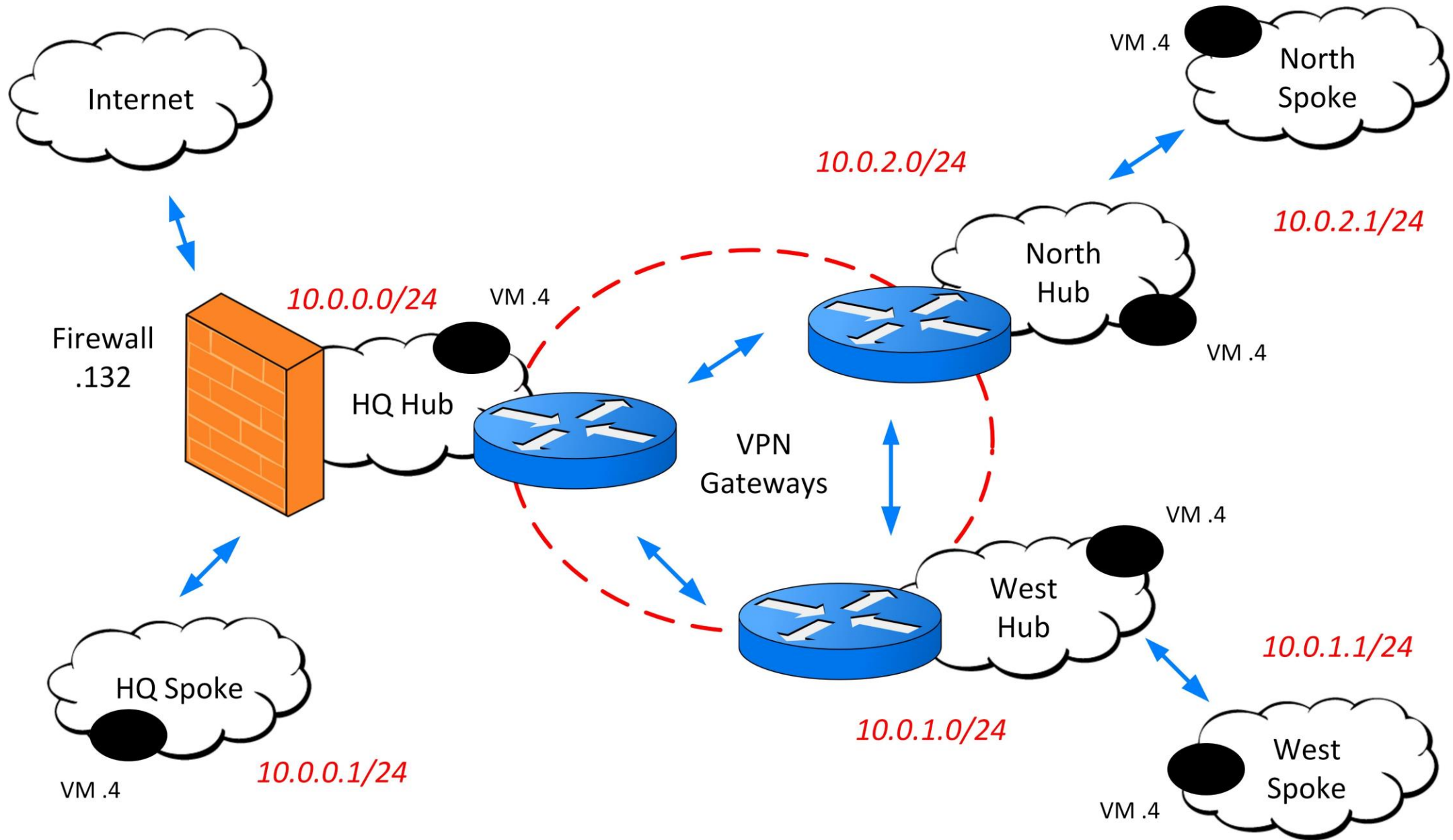
► Create Global VNet Peering (between West and North Hub)



VNet-to-VNet VPN

- Has paired connections using the same preshared key
- Uses Azure Backbone connection with low and stable latency
- In the same region, there are no transfer cost, between regions ingress and egress transfer cost are billed.
- Routing based on known System Routes of peers, including VNet Peering, but does not recognise Global VNet Peering

► Create Full Mesh VNet-to-VNet VPN



Redundancy issues

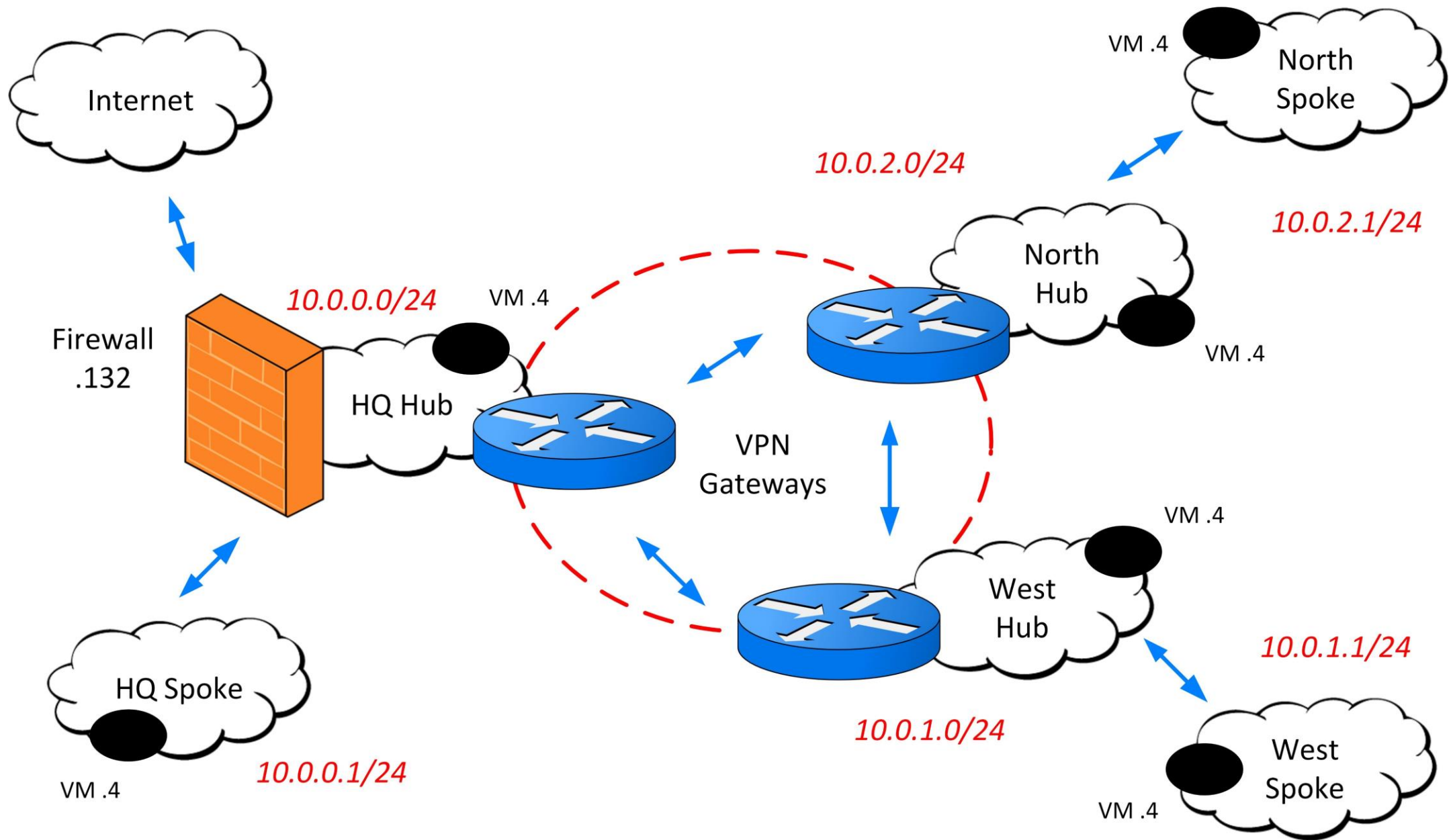
- Supports fully propagated routing and transparent connectivity
 - Fails to guarantee connectivity in case of loss of one link
-
- ▶ Start PING from West Spoke to HQ Spoke
10.1.1.4 (ifconfig) to 10.0.1.4
 - ▶ Break West European Connection (HQ-West)

Full Mesh with BGP on VNet-to-VNet VPN

- With BGP, routes are automatically added to route table
 - Routing to BGP Peer is added to route table
 - Guarantees connectivity by dynamic routing in case of loss of one link
-
- ▶ Enable BGP on connection (using resource manager)
 - ▶ Start PING from West Spoke to HQ Spoke (with Link broken)

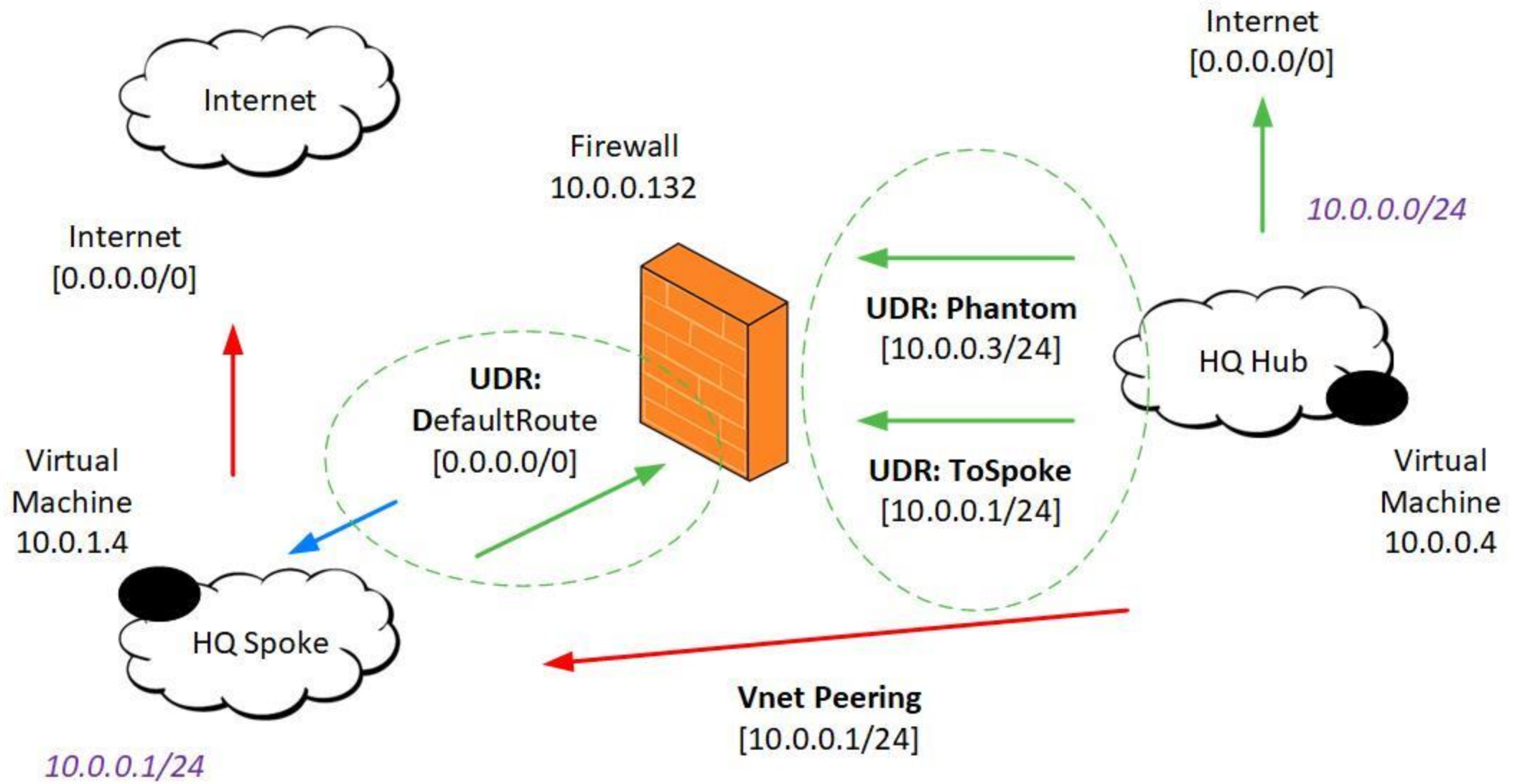
Site-to-Site VPN with BGP (S2S)

- Uses Local Network Gateways to connect to Non-Azure VPN, mandatory approach for connection On Premise networks
 - Connects via Internet with possible bandwidth and latency limitations
 - Transfer costs billed for egress data only, for every site.
 - S2S connection allows specific IPSEC settings for Non-Azure VPN
 - Same functionality as VNet-to-VNet VPN, with or without BGP
-
- ▶ Build full Mesh Site-to-site VPN using Local Network Gateways
 - ▶ Start PING from West Spoke to HQ Spoke, and break the link



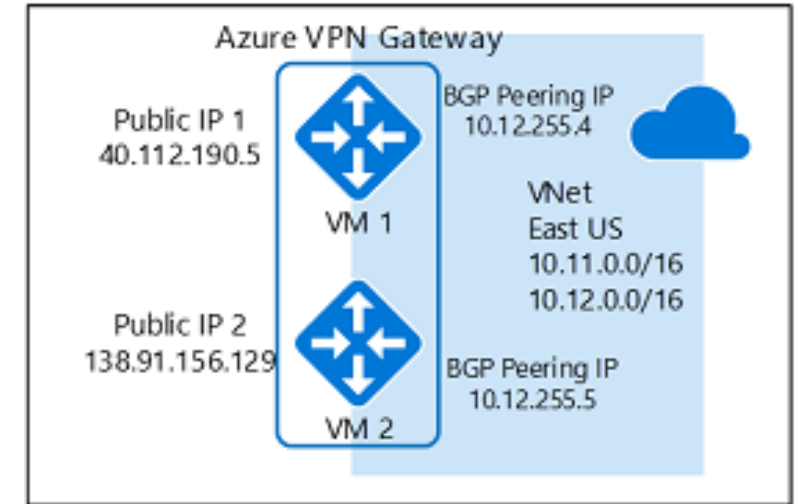
BGP and User Defined Routes (UDR)

- User Defined Routes take priority over System Routes in routing, but
 - User Defined Routes are not taken into account by BGP
-
- ▶ Demonstrate SNAT on HQ Spoke (from SSH internal IP)
 - ▶ Traceroute from West to HQ Hub and HQ Spoke
 - ▶ Add User Defined Route for Phantom IP range on Firewall



Key Considerations

- BGP offers redundant VPN connectivity, in a full mesh multi-site setup. For VPN to 1 site use Active-Active Virtual Network Gateways¹
- To connect to On Premise networks, use Local Network Gateways
- Between Azure Regions, VNet-to-VNet VPN is a better choice
- User Defined Routes are not propagated by BGP
 - When you need to propagate routes in your hybrid network, use Next Generation Firewall that can export and/or advertise specific routes
 - Even then the **default route** 0.0.0.0/0 is not propagated by BGP, split in 2 parts (0.0.0.0/1 and 128.0.0.0/1)



¹ <https://docs.microsoft.com/en-us/azure/vpn-gateway/vpn-gateway-activeactive-rm-powershell>

Monitoring BGP Peers

- Network Wachter monitor requires agent on Virtual Machine
- Peer IP's can be reached via TCP 179
- Monitoring on Round Trip Times may leads to too many false positives
- As BGP Peers keep being accessible via the redundant routes, alerts only occur when there is a total outage

► Create Network Watcher Monitoring