OVERVIEW

As a valued security partner for Microsoft’s Cloud services, Fortinet has worked diligently to build a complete cloud security solution that extends the Fortinet Security Fabric from the private data center into the cloud, with advanced security orchestration and unified threat protection.

The FortiGate-VM is a key part of this solution and provides more control and visibility by identifying and setting policy by user applications, device specs, IP, and network interfaces. Through these offerings, Fortinet delivers the highly optimized solution for Microsoft Azure where application workloads can be protected beyond native Azure security options.

When it comes to these deployments, high availability (HA) has been one of the biggest challenges for cloud providers to guarantee high service-level agreements. Constructing a highly available security system in the cloud has never been an easy task either. To this end, Fortinet provides end-to-end configuration templates for common cloud practices and makes them available in Fortinet’s Azure Marketplace listings. The current marketplace FortiGate HA option in Azure is a highly scalable HA design.

To deploy directly from your Azure portal, click on ‘+New’ and search for ‘FortiGate,’ then choose “FortiGateNGFW high availability (HA).”

In this architecture, inbound traffic comes through the load balancer and the flow of traffic is determined by the load balancer itself (unless a connection is already established, in which case the flow follows the existing path). For inbound traffic coming from the load balancer, both FortiGate appliances are in Active/Active configuration. Source NAT is recommended on the incoming policies in order to maintain path symmetry for return traffic. Fortinet doesn’t recommend Active/Passive deployments for this use case. Any manipulation of UDRs or public IPs for Active/Passive solutions takes 30 seconds to be applied after the failover is initiated. Azure load balancers can, however, send probes as often as every 5 seconds and will stop forwarding traffic after two failures. Thus, a failure is detected and mitigated within 6-10 seconds. The only drawback to Active/Active is that it relies heavily on source NAT or FortiGate Session Life Support Protocol (FGSP). FGSP can function, but asymmetric routing will reduce the security effectiveness of IPS and application control.
Outbound flows follow whichever User Defined Route (UDR) is currently installed, and all outbound traffic is initiated through the current active FortiGate. So, for outbound-initiated traffic from Azure Subnets, FortiGate appliances are in Active/Passive mode. We need to rely on Azure software-defined networking (SDN) principles with one or more worker nodes monitoring the Active FortiGate appliance. In case the FortiGate is not reachable, the script will switch to another FortiGate standby appliance.

FortiGate-A starts as active. The script monitors FortiGate-A until it stops responding. Then the route tables (UDRs) are changed to send traffic to FortiGate-B, at which point the script begins monitoring FortiGate-B until such time as FortiGate-B stops responding and the route tables are reverted to send traffic to FortiGate-A (and so on).

The worker nodes that run the monitor script need to be in the public subnet so they can communicate with Azure control to change routing if the primary FortiGate goes offline. We have sample bash scripts that can be tailored to suit the environment. Microsoft has recently developed a similar process using scripts running in a container cluster. Here are the links for that solution: https://channel9.msdn.com/Shows/Azure-Friday/Deploying-Network-Virtual-Appliances-for-High-Availability. The github repo is at https://github.com/mspnp/ha-nva

**DUAL LOAD BALANCER FORTIGATE HA SETUP IN AZURE**

Microsoft has updated their internal load balancer so it can be utilized as a next hop for these types of deployments. Fortinet can take advantage of this feature to develop FortiGate HA in Azure.

In this architecture, we leverage dual load balancers (public and internal) in Azure to achieve HA for the FortiGates. This design will remove the need for any sort of SDN scripting. The major difference is that an internal load balancer doesn’t have a public IP address. Instead, we associate an internal load balancer with a virtual network subnet and it gets a private IP address. All traffic will be routed to an internally load balanced service via this private IP address. Thus, the internal Azure load balancer functions as the UDR next hop for subnets not directly connected to the FortiGate appliances. To check availability of the FortiGate appliances, it will monitor the FortiGate appliances’ (inside/outside) NICs using probes and load balance connections.
For inbound flows from Internet/on-premises to Azure Subnets, the Azure public load balancer will load balance the connections between FortiGate(s) and both FortiGate(s) will be Active/Active, similar to Solution A. For outbound flows as well, with introduction of internal load balancers, this architecture provides Active/Active, similar to inbound flows. It is again recommended to use source NAT on FortiGate(s) for connections initiated outbound from Azure Subnets to maintain path symmetry.