Cloud NGFW for Azure

Disaster Recovery Guide

Version 1.0
June 2023

Introduction **3**

Disaster Recovery **5**

1. Zonal redundancy 5

A. VM Series Instance (or Azure VM) Failure 6

B. Azure Availability zone failures 7

2. Multi-Region Strategy 7

Conclusion **10**

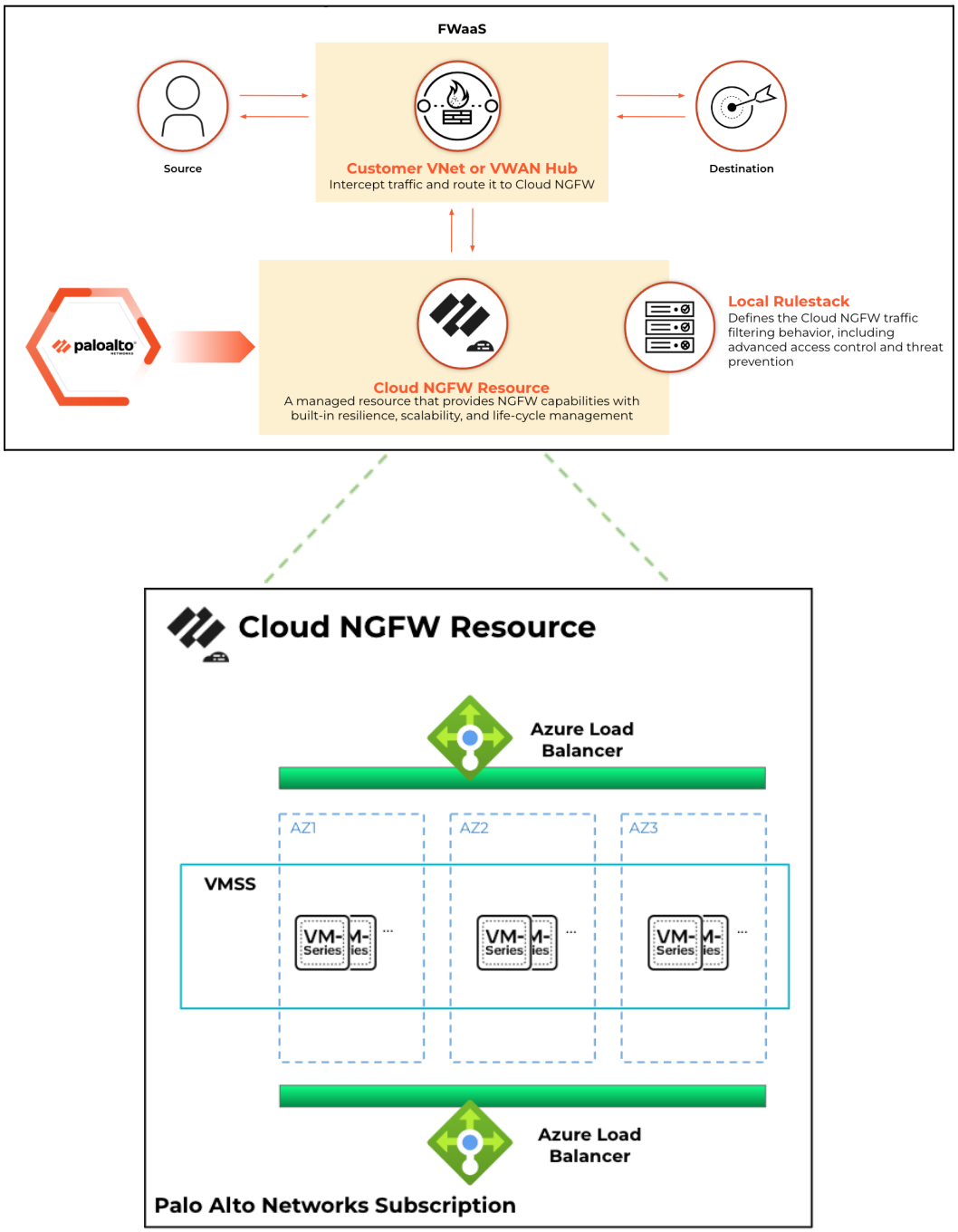
Introduction

Cloud NGFW for Azure is a regional service similar to other Azure zone-redundant services. This service is delivered on the Azure platform to protect your Azure Virtual Network (VNet), Virtual WAN (VWAN) Hub, Branch, VPN, and ExpressRoute traffic. A **Cloud NGFW resource** provides next-generation firewall capabilities without the need to manage the underlying infrastructure. This resource has built-in resiliency, scalability, and life-cycle management. A Cloud NGFW resource is deployed either into a Hub VNet or a VWAN Hub and uses the underlying Azure VNet infrastructure to inspect the traffic. In Azure, Virtual Networks span all availability zones in a region.

Under the hood, each Cloud NGFW resource includes a pair of Azure Load Balancer resources that are managed by Palo Alto Networks, with built-in resiliency, zone-affinity, scalability, and life-cycle management. Each Cloud NGFW resource includes a dedicated Virtual Machine Scale Set (VMSS) of Palo Alto Networks VM-Series virtual firewall nodes.

Cloud NGFW resource scales according to the traffic being processed by the service. The instances in the VMSS are continuously monitored by the Cloud NGFW service, and the cluster scales out by more underlying VM-series instances being created. As the traffic volume decreases, the VMSS automatically scales in.

A **Cloud NGFW rulestack** defines Cloud NGFW resource's advanced access control (App-ID, URL Filtering) and threat prevention behavior. A rulestack includes a set of security rules, associated objects, and security profiles. To use a rulestack, you associate the rulestack with one or more Cloud NGFW resources.


FWaaS
Source
Customer VNet or VWAN Hub
Intercept traffic and route it to Cloud NGFW
Destination
Local Rulestack
Defines the Cloud NGFW traffic filtering behavior, including advanced access control and threat prevention
Cloud NGFW Resource
A managed resource that provides NGFW capabilities with built-in resilience, scalability, and life-cycle management
Cloud NGFW Resource
Azure Load Balancer
AZ1
AZ2
AZ3
VMSS
VM-Series
Azure Load Balancer
Palo Alto Networks Subscription

Disaster Recovery

Disaster recovery is the process by which an organization anticipates and addresses technology-related disasters. IT systems in any company can go down unexpectedly due to unforeseen circumstances, such as power outages, natural events, or security issues. Disaster recovery includes a company's procedures and policies to recover quickly from such events.

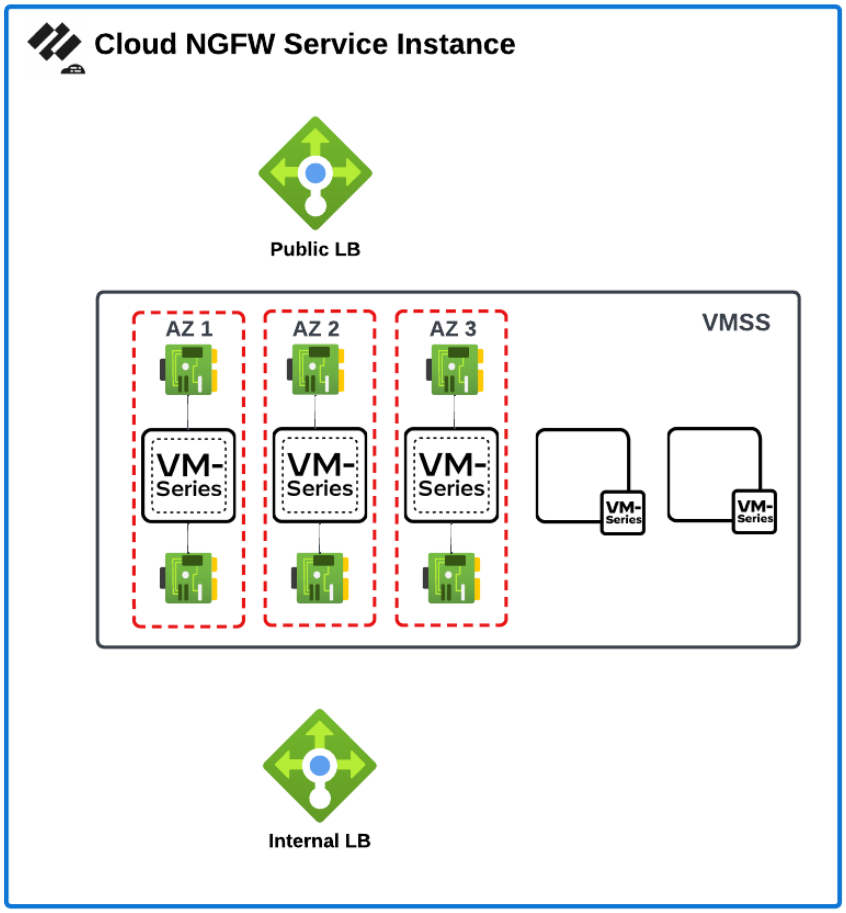
A disaster event can take down your workload. Therefore, your objective for disaster recovery should be bringing your workload back up or avoiding downtime altogether.

You may deploy application workloads across multiple Azure Availability Zones and regions for global availability or for reducing Recovery Point Objectives (RPO) and Recovery Time Objectives (RTO) as part of a disaster recovery (DR) plan. Azure recommends the following methods

1. Zonal redundancies
2. Regional redundancies

1. Zonal redundancy

Every Azure Region consists of multiple Availability Zones (AZs). Each AZ consists of one or more data centers in a separate and distinct geographic location. This Azure region architecture significantly reduces the risk of a single event impacting more than one availability zone.



You may deploy your application workloads across multiple availability zones for disaster recovery based on Azure recommendations. In this case, we have identified the following disaster events that can affect your Cloud NGFW resources. Cloud NGFW offers resiliency for these disaster events.

A. VM-Series instance (or Azure VM) failures
B. Azure Availability zone failures

A. VM Series Instance (or Azure VM) Failure

Cloud NGFW resource offers built-in **resiliency within an availability zone in an Azure region** by having a minimum of three VM-Series instances running simultaneously in a dedicated Virtual Machine Scale Set (VMSS) for high availability. Cloud NGFWs use the VMSSs running VM-Series to enable resiliency for VM-Series (or

VM) instance failures. The fine-grained health check configurations enable the Azure Load Balancer to detect faults in the VM-Series instances and immediately bring up a new VM-Series instance. Since the recovery heuristic is built into the product and does not require any action from your end, Palo Alto Networks will not notify you about this event.

B. Azure Availability zone failures

Cloud NGFW resource offers built-in **resiliency across Azure availability zones in an Azure region** by having a distinct Virtual Machine Scale Set (VMSS) which is configured to distribute the VMs across availability zones in a given region.

In a rare event of a complete Availability Zone failure, the blast radius within the Cloud NGFW resource is limited to the VM-Series instances provisioned in that specific availability zone.

The Cloud NGFW resource remains intact and protects traffic in other Azure availability zones using the VM-Series in those zones.

Suppose the entire Azure availability zone is down. In that case, all your application workloads and load balancers in that availability zone will also be down, and Cloud NGFW will receive no traffic in that availability zone.

When the Azure availability zone is back up, the Cloud NGFW resource automatically detects the change and immediately brings up the instances in that availability zone. Since the recovery heuristic is built into the product and does not require any action from your end, Palo Alto Networks will not notify you about this event.

2. Multi-Region Strategy

Based on Azure recommendations, you may deploy application workloads across multiple Azure regions for global availability and disaster recovery (DR) using different methods.

A. Azure Site Recovery
B. Storage replication
C. Traffic Manager

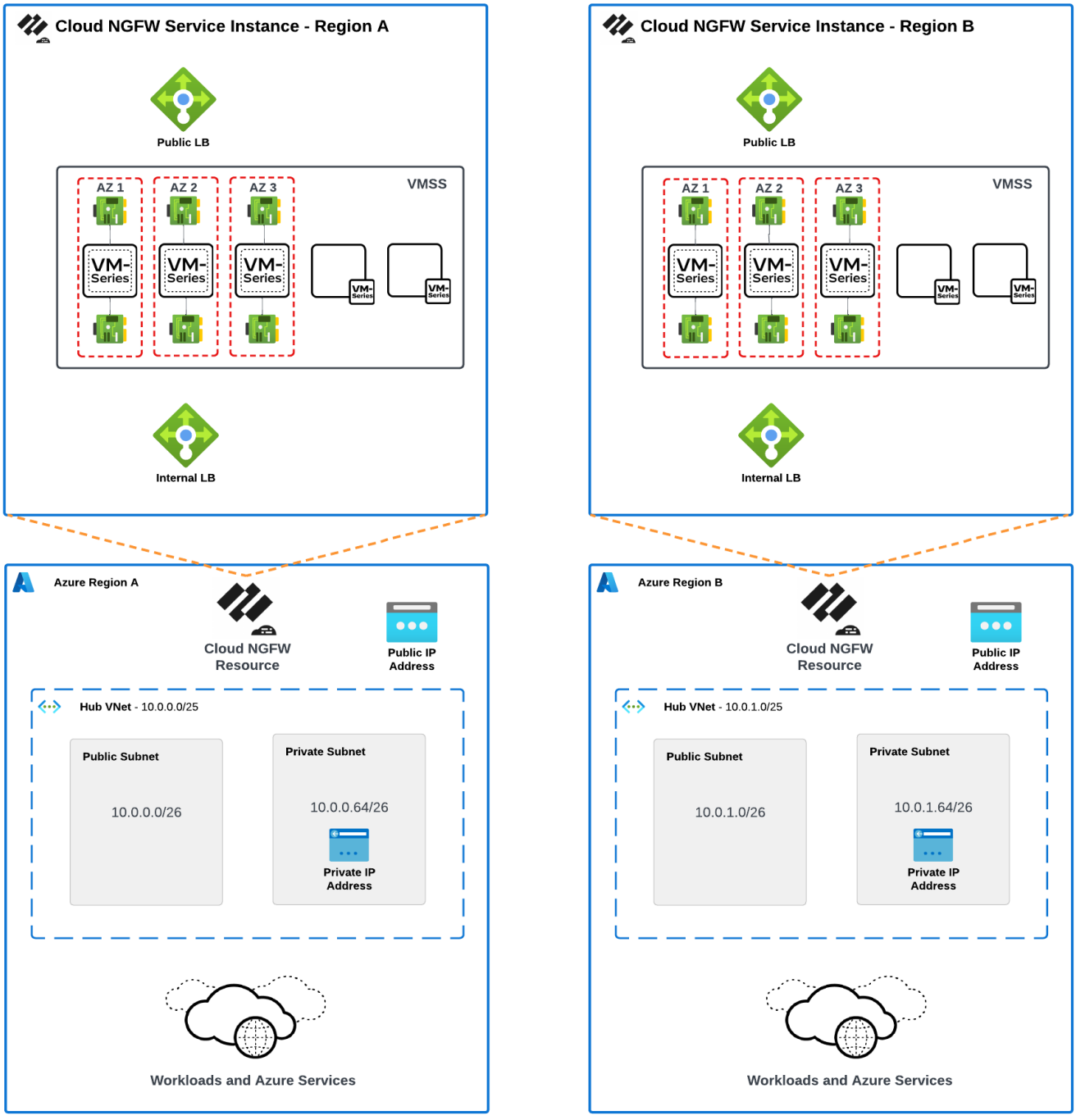
With all these services, you can protect the traffic by deploying the Cloud NGFW resource(s) in each Azure region where the applications are deployed. You will also need to synchronize the rules deployed on these Cloud NGFW resources using your preferred automation tools (such as ARM or Terraform).

In a rare event of a complete Azure regional failure, all scale sets and VM-Series instances powering your Cloud NGFW resource will be down.

If the entire Azure region is down, all your application workloads and VNets will also be down, along with your Cloud NGFW resources. There is no traffic in the region to secure during this type of outage. When the Azure region is back up, Cloud NGFW resource automatically detects the change. It immediately brings up the new VM-Series instances.

In this architecture, you will have Cloud NGFW resources deployed in both regions to secure the regional traffic.

For the VNet hub-and-spoke deployment, a Hub VNet in each region will have a Cloud NGFW resource associated with it and a user-defined route pointing to the Private IP address of the resource. Keep in mind that in the case of a partial region failure i.e. some service remaining operational, you have the option of pointing the user-defined route to the Cloud NGFW region running in the DR region provided the VNet peering between regions is in place and there's no impact to that traffic in Azure. Be aware that this will incur cross-region data transfer charges and introduce additional latency.



For the VWAN Hub deployment, a Hub in each region will have a Cloud NGFW SaaS associated with it and a routing intent configured to send the traffic via the Cloud NGFW resource. If the Virtual hub is operational during a partial regional outage, you

have the option of unconfiguring the routing intent and, as a result, bypassing the Cloud NGFW service. Custom route tables are not supported currently.

Conclusion

As mentioned above, the Cloud NGFW service is designed to withstand multifaceted failures. First, the Cloud NGFW resource is backed by a cluster of highly-available instances of the VM-series virtual firewalls. The service continuously monitors the health of the individual instances and automatically recovers the cluster should there be an issue with any VM-series that power up the service.

The Cloud NGFW is also protected against failures in the Azure data center by leveraging the Zonal redundancies and deploying the underlying VM-series virtual firewall instances into different Availability Zones. This guarantees that the Cloud NGFW will remain operational in case of a disaster taking place at a particular geographic location.

For the unlikely event of a full Azure region failure, Azure offers a suite of services that enable Enterprise-scale Disaster Recovery to ensure your workloads and data are replicated across the regions. It is recommended to deploy the Cloud NGFW in each region where the workloads are present not only from the latency perspective but also for resiliency purposes. Should the entire region be in jeopardy, the Cloud NGFW in the DR region will continue securing the applications and users there.