NAT gw Availability zones

Index

11:20

- #zonal-resource
- can be deployed and operate out of individual availability zones.

- Two options to deploy
  - no zone (region deployment)
  - specific zone within a region

- NAT gateway uses software defined networking to operate as a fully managed and distributed service.
- Has a built-in redundancy source.
- Once deployed, The zone of a NAT gateway cannot be changed.
- If you choose *no zone*. NAT gw is places in a zone by Azure, with no guarantee of which one.
- NAT gateway can provide out bound connectivity for VM from other availability zones in the same region and the VM's subnet should be configured to the NAT gateway resource to provide outbound connectivity.

- **Limitations**
  - NAT gateway can only serve subnets within the same region, regardless of zone placement.
  - Managing multiple zonal NAT gateways across vnets adds complexity compared to a single regional solution.
  - Zonal NAT gateways may be slightly more expensive than regional deployments

---

- When NAT gateway is placed in **no zone**, Azure places the resource in a zone for you.
- You won't have visibility into which zone Azure chooses for your NAT gateway.

Design - Considerations

Single Zonal NAT Gateway Resource for Zone-spanning Resources

- A single #zonal-NAT-gateway resource can be configured to either
  - A subnet that contains virtual machines that span across multiple availability zones.
  - or to multiple subnets with different zonal virtual machines.
- If the zone that NAT gateway is deployed in goes down, then outbound connectivity across all virtual machine instances associated with the NAT gateway will also go down.
- This set up doesn't provide the best method of zone-resiliency.

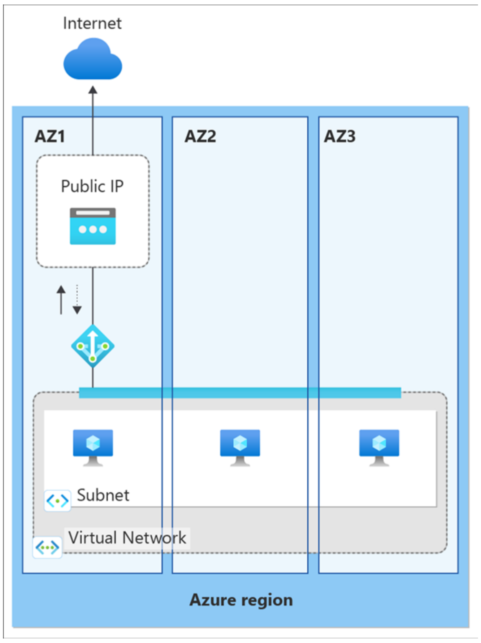

Figure: Single zonal NAT gateway resource for multi-zone spanning resources doesn't provide an effective method of zone-resiliency against outages.

Zonal NAT Gateway Resource for Each Zone in a Region to Create Zone-resiliency

- when a virtual machine instance using a NAT gateway resource is in the same zone as the NAT gateway resource and its public IP addresses.
- The pattern you want to use for zone isolation is creating a #NAT/zonal-stack per availability zone.
- This #NAT/zonal_stack consists of virtual machine instances, a NAT gateway resource with public IP addresses or prefix on a subnet all in the same zone.

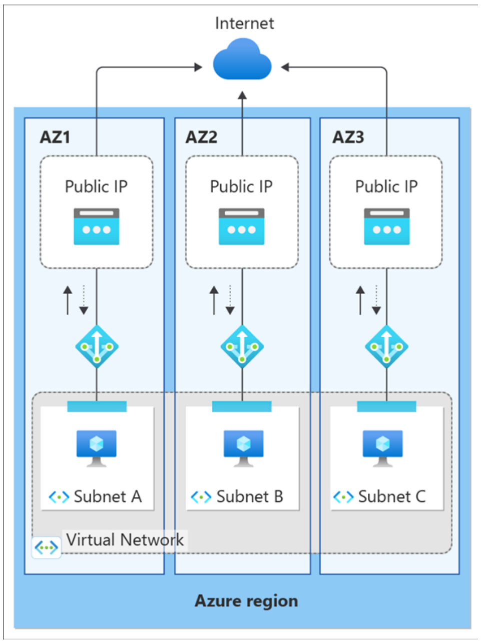

Figure: Zonal isolation by creating zonal stacks with the same zone NAT gateway, public IPs, and virtual machines provides the best method of ensuring zone resiliency against outages.

Note

- Creating zonal stacks for each availability zone within a region is the most effective method for building zone-resiliency against outages for NAT gateway.
- However, this configuration only safeguards the remaining availability zones where the outage did **not** take place.
- With this configuration, failure of outbound connectivity from a zone outage is isolated to the specific zone affected.
- The outage won't affect the other zonal stacks where other NAT gateways are deployed with their own subnets and zonal public IPs.

Option	Pattern	Example	Pro	Con
(1)	**Align** the inbound endpoints with the respective **zonal stacks** you're creating for outbound.	Create a standard load balancer with a zonal frontend.	Same failure model for inbound and outbound. Simpler to operate.	Individual IP addresses per zone may need to be masked by a common DNS name.
(2)	**Overlay** the zonal stacks with a cross-zone inbound endpoint.	Create a standard load balancer with a zone-redundant front-end.	Single IP address for inbound endpoint.	Varying models for inbound and outbound. More complex to operate.