Linkedin: https://www.linkedin.com/in/anand-rao-8aba1b123/

Blog site: https://alwayslearn.in

Community: https://www.meetup.com/azure-aks-chennai-user-group/

API Management with AKS and AI

Agenda

- EXPOSING AI ENDPOINTS HOSTED ON AKS AS API
- USING API MANAGEMENT TO SECURE THE API
- DIFFERENT TYPES OF NETWORKING IN API MANAGEMENT

What we are trying to achieve

- We are going to deploy a model from Azure AI and build an API to call that model from AKS
- This api can be used in custom applications (frontend for example)
- Once this API is hosted on AKS, we will secure this API using API management

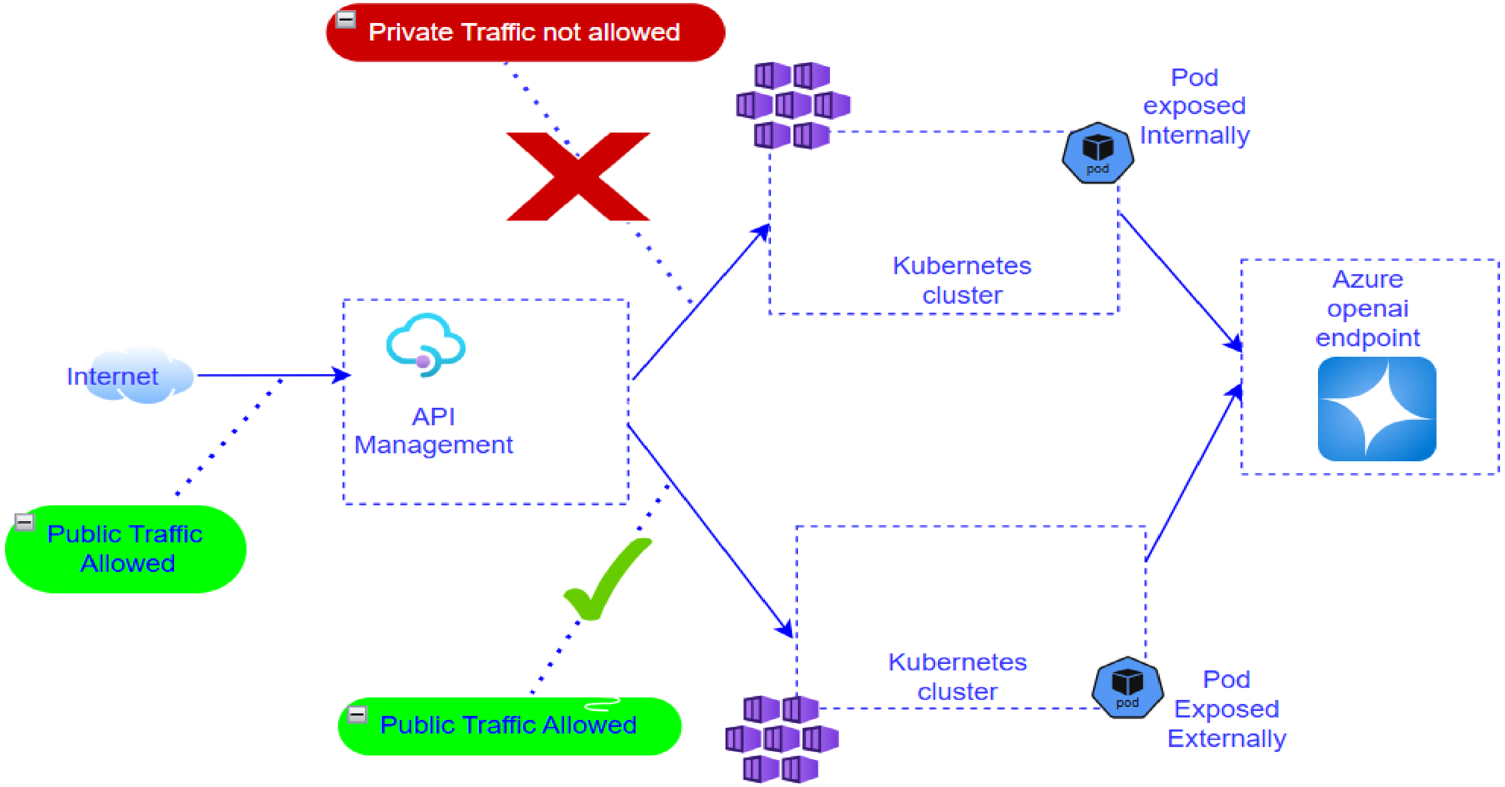
Why API Management

- **Security**
  - Authentication & authorization
  - Rate limiting & quotas
  - IP filtering through policies
  - Protect AKS APIs from direct exposure
- **AI Protection**
  - Token rate limiting
  - Cost control
  - Centralized logging & monitoring

Networking in API management

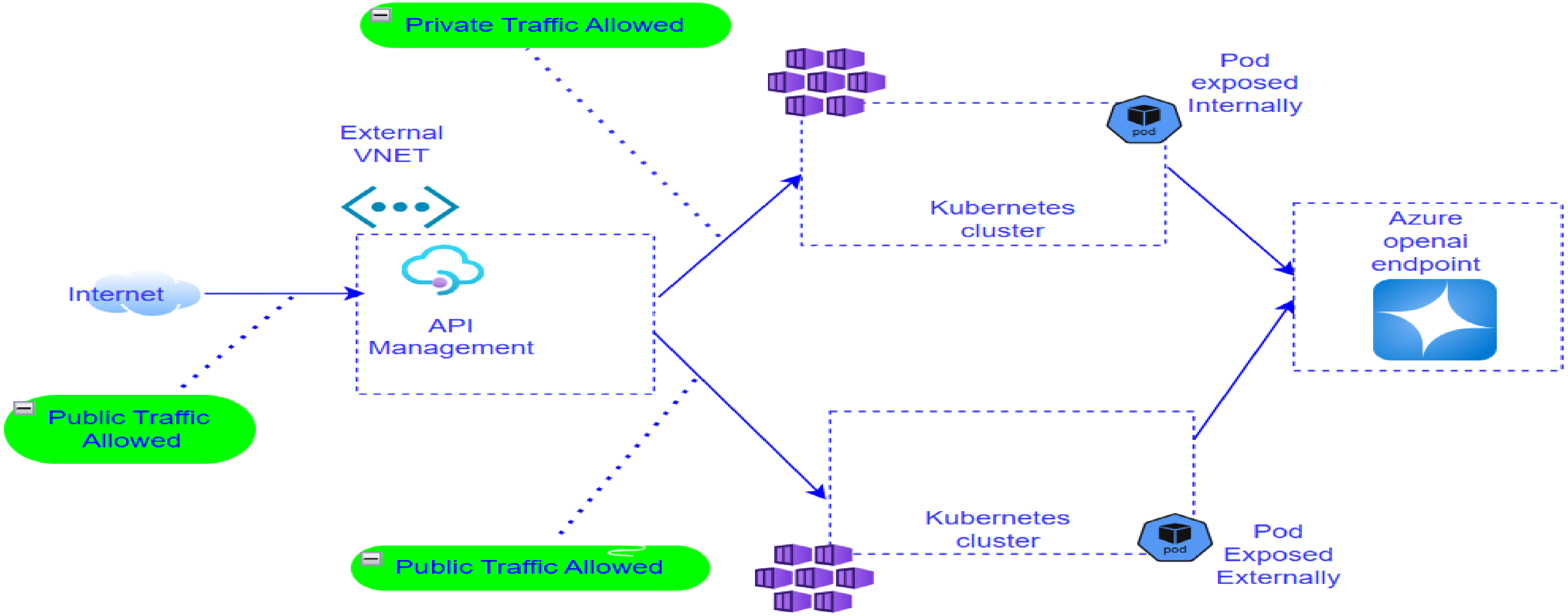
- API Management has few options in Networking
  - API Management completely public
  - API management with External mode
  - API management with Internal mode
  - API management with VNET integration

API Management public

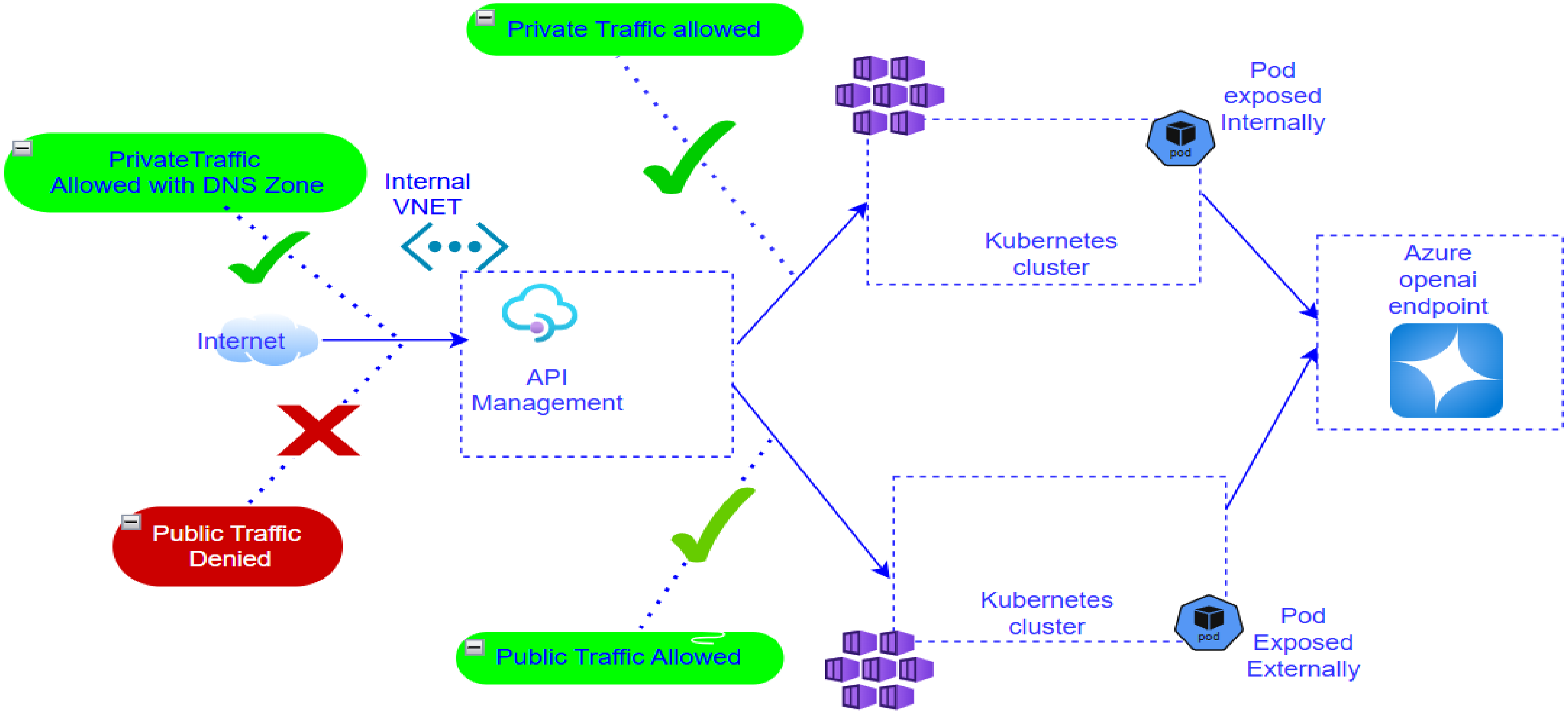


- Api endpoints are exposed publicly
- The backend public APIs can be called from the API management

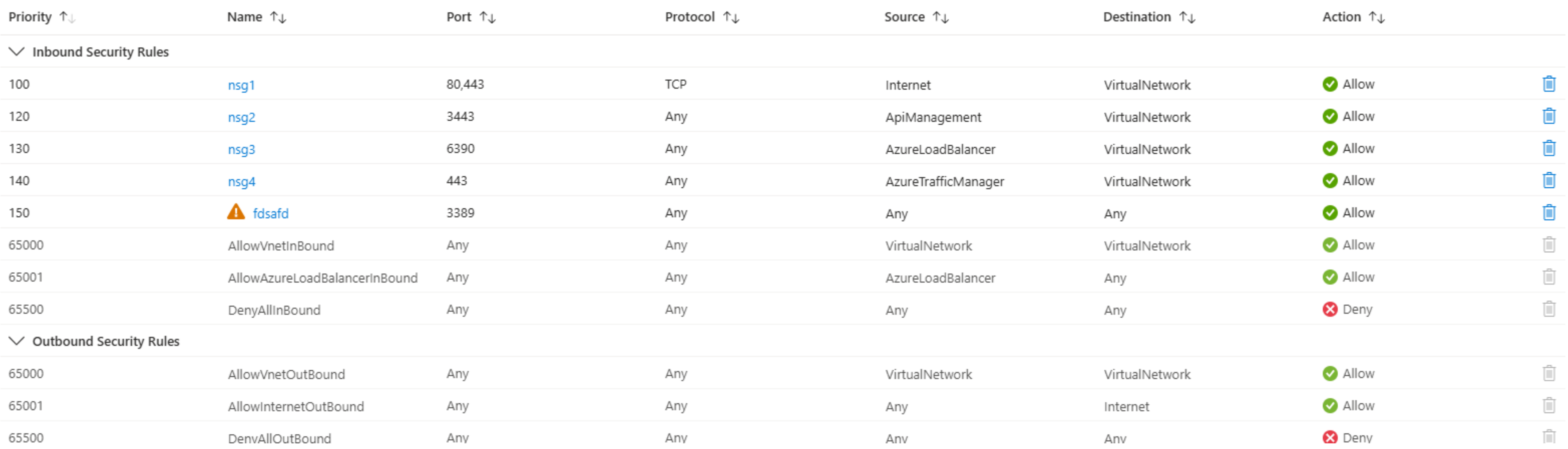
API management External



API management Internal

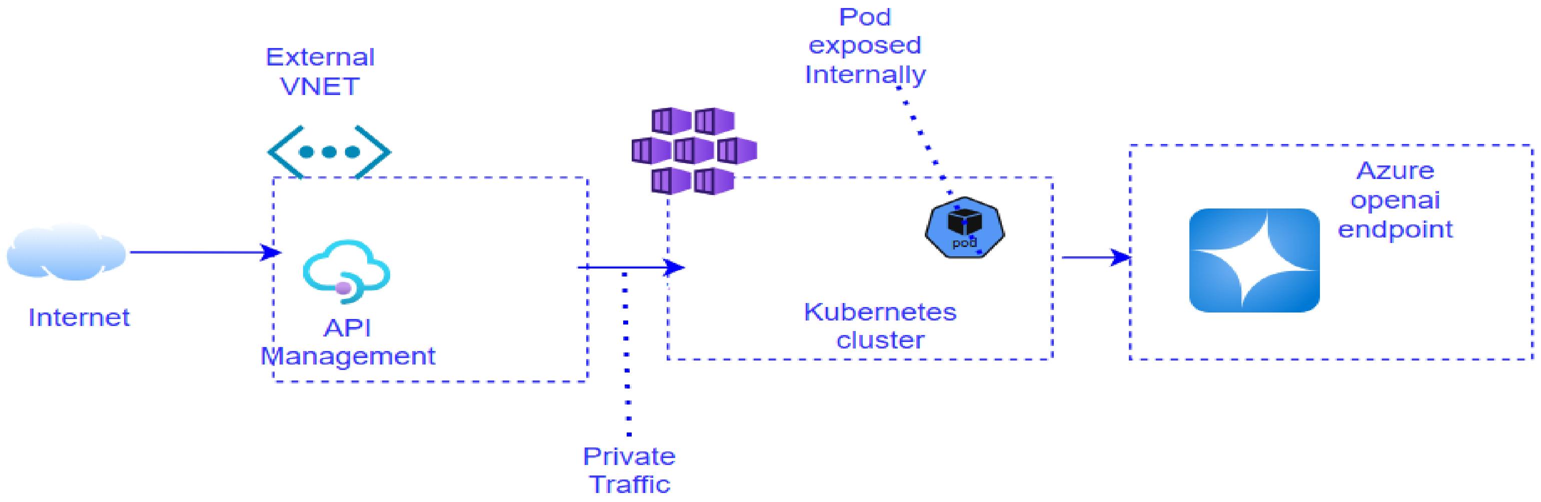


NSG Rules



Priority	Name	Port	Protocol	Source	Destination	Action
Inbound Security Rules						
100	nsg1	80,443	TCP	Internet	VirtualNetwork	✓ Allow
120	nsg2	3443	Any	ApiManagement	VirtualNetwork	✓ Allow
130	nsg3	6390	Any	AzureLoadBalancer	VirtualNetwork	✓ Allow
140	nsg4	443	Any	AzureTrafficManager	VirtualNetwork	✓ Allow
150	⚠ fdsafd	3389	Any	Any	Any	✓ Allow
65000	AllowVnetInBound	Any	Any	VirtualNetwork	VirtualNetwork	✓ Allow
65001	AllowAzureLoadBalancerInBound	Any	Any	AzureLoadBalancer	Any	✓ Allow
65500	DenyAllInBound	Any	Any	Any	Any	✗ Deny
Outbound Security Rules						
65000	AllowVnetOutBound	Any	Any	VirtualNetwork	VirtualNetwork	✓ Allow
65001	AllowInternetOutBound	Any	Any	Any	Internet	✓ Allow
65500	DenyAllOutBound	Any	Any	Any	Any	✗ Deny

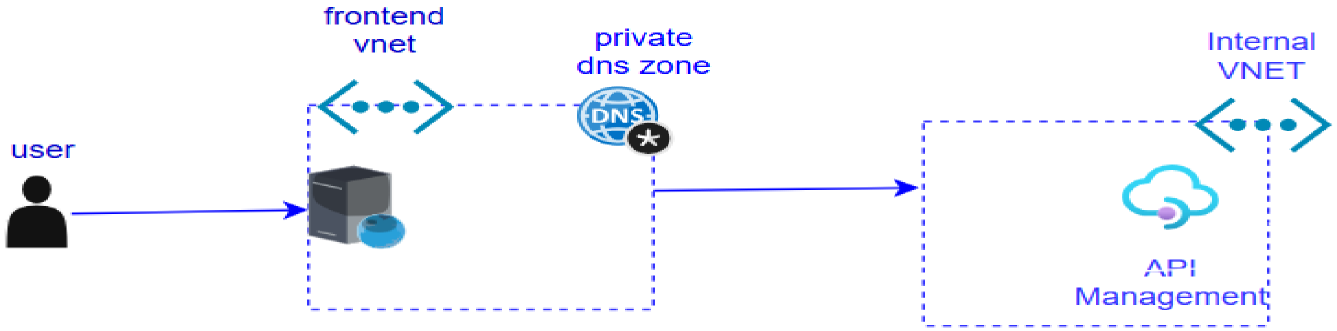
API management External



Additional security

- Add network policy on the backend pods to be accessible only from apim private ip (azure network policy should be enabled for cluster)
- Add JWT validation
- Add subscription keys
- Introduce application gateway in front of APIM
- Prevent APIM to be accessible only from specific IP’s
- Add cors policy in APIM

DNS in API Management



In internal vnet, the source making the API call should be able to resolve the API management url

- The source should be integrated with private dns zone

Repo to build the API

- Please use this repo to build API
  - anandazaugust/openaibot at testv4