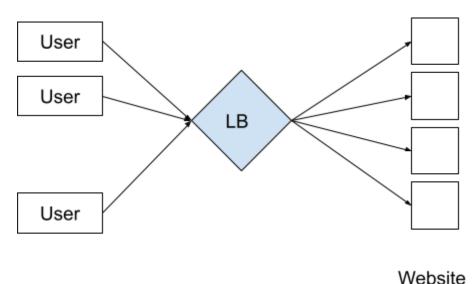
# **Azure Load Balancer**



Website

Un IP este asignat unui singur VM (nu putem sa avem doua VM-uri cu acelasi IP). Daca avem mai multe VM-uri pe care ruleaza aceeasi aplicatie, fiecare VM va avea un IP diferit. Utilizatorii trebuie sa stie toate IP-urile folosite de aplicatie. Daca inchidem un VM, trebuie sa actualizam lista IP-urilor.

Un Load Balancer va permite sa distributi request-urile utilizatorilor catre una sau mai multe masini virtuale. Asadar, aplicatia noastra va avea un singur IP si LB-ul va redirectiona traficul catre VM-uri (de exemplu folosind un algoritm de tipul round robin).

Pentru a testa Load Balancer-ul vom face doua masini virtuale care vor contine un site simplu (static). Atunci cand intram pe IP-ul public, acestea vor afisa un mesaj ca "Welcome to myapp. You are running on \$HOSTNAME", un \$HOSTNAME este numele VM-ului.

#### Website

- Creati o masina virtuala folosind Debian care sa permita accesul pe portul 80. Daca ati uitat sa dati acces cand ati creat VM-ul, intrati pe VM in Networking (in partea stanga), Apasati Add inbound port rule, la Service selectati HTTP si dati Add.
- Instalati nginx:

sudo apt-get update && sudo apt-get install nginx -y

 Verificati ca site-ul functioneaza: Intrati pe IP-ul public. Trebuie sa primiti o imagine similara cu cea de mai jos:

#### Welcome to nginx!

If you see this page, the nginx web server is successfully installed and working. Further configuration is required.

For online documentation and support please refer to <u>nginx.org</u> Commercial support is available at <u>nginx.com</u>.

Thank you for using nginx.

Modificati site-ul pentru a aparea mesajul mesajul:
 Welcome to myapp. You are running on VM #NR

echo "Welcome to myapp. You are running on VM #1" | sudo tee /var/www/html/index.nginx-debian.html

Intrati din nou pe IP-ul public si trebuie sa vedeti noul mesaj

## **Imagine**

 Daca am clona VM-ul in acest moment, am avea si pentru al doilea site acelasi mesaj. Vom crea un script care la pornirea VM-ului va modifica mesajul

sudo apt-get install cron -y

sudo crontab -e

@reboot echo "Welcome to myapp. You are running on VM `cat /etc/hostname`" | sudo tee

/var/www/html/index.nginx-debian.html

 Pentru a testa ca functioneaza, dati un reboot VM-ului sudo reboot

- Pentru a clona VM-ul, vom face o imagine care sa contina starea curenta.
- Inainte de a clona VM-ul trebuie sa pregatim masina (VM-ul trebuie sa nu refoloseasca acelasi nume / aceeasi cheie SSH a host-ului etc.). Rulati comanda

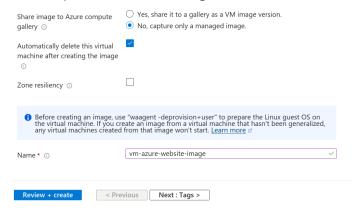
sudo waagent -deprovision+user

Dupa aceasta comanda nu veti mai putea sa intrati in VM.

 Pentru a crea imaginea, intrati din web browser pe VM si selectati Capture

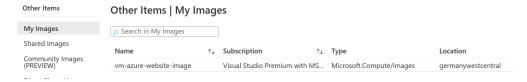


 Bifati ca vreti o imagine simpla, sa se stearga VM-ul si selectati un nume pentru noua imagine.



## Clonare

- Vom crea un VM care vor avea ca imagine de baza imaginea de mai sus (vm-azure-website-image).
- La imagini, selectam See all images si in partea stanga selectam My Images



 La conectare alegem sa folosim aceeasi cheie privata (Use existing key stored in Azure) sau copiem cheia publica.

- Permitem accesul pe portul 80 si apasam Review + create
- Intrati pe IP-ul public si vedeti daca primiti mesajul
  Welcome to myapp. You are running on VM <nume-vm>
- Dupa ce verificati ca functioneaza, puteti sterge VM-ul.

### Scale Set si Load Balancer

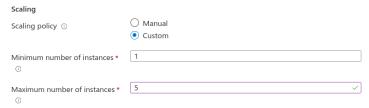
 Cream un Load Balancer care va distribui traficul catre una din cele doua VM-uri. In bara de Search scrieti Virtual machine scale set



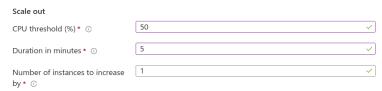
- Apasati Create; Alegeti un nume pentru scale set: vm-azure-ss
- La imagine alegeti imaginea creata de voi:



La meniul de scaling, lasati la Initial instance count: 2, dar modificati
 Scaling policy din Manual in Custom si adaugati maxim 5 instante:



La scalare in sus selectati perioada minima (5 minute):



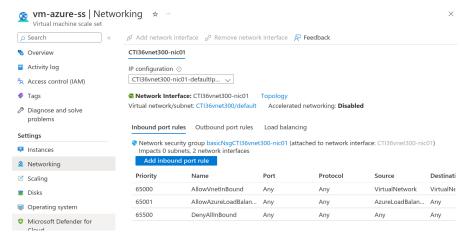
- Dati Review + Create;
- Daca intrati din web browser trebuie sa vedeti instantele initiale:



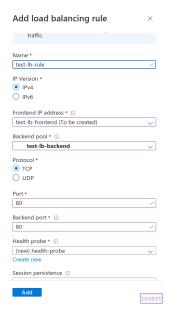
- Daca intrati pe VM-uri veti vedea ca acestea au IP-uri private (nu le putem accesa din afara). Cream un Load Balancer pentru a putea sa le accesam. Cauti in bara de Search Load Balancer si apasati Create. La Type selectati Public si dati Next.
- La Frontend IP apasati + Add a frontend IP configuration si creati un IPv4 nou; dupa apasati Add si Next:



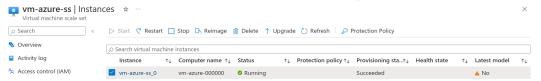
 La Backend pools apasati + Add a backend pool. Alegeti un nume si pentru a gasi Virtual network-ul intrati in Scale Set la partea de Networking (CTI36vnet300-nic01). Apasati pe Add inbound port rule si permiteti acces pe portul 80 (HTTP)



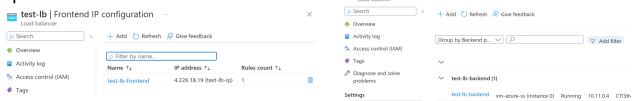
- La partea IP configurations din Add backend pool apasati Add, selectati Scale Set-ul vostru si dati Add. Dupa apasati Save.
- La partea de Inbound rules la Load balancing rule apasati
  - + Add a load balancing rule, selectati Frontend IP-ul si Backend pool-ul. De asemenea setati porturile la 80 si 80 si creati un nou Health probe:



- Sariti peste Outbound rules (aceastea sunt daca vreti sa permiteti VM-urilor sa acceseze internet-ul). Dati Review + Create.
- Dupa aceasta intrati in Scale Set si dati Upgrade la masini:



 Daca totul este in regula, trebuie sa intrati pe Load balancer in Frontend IP sa gasiti IP-ul LB-ului si sa verificati in Backend pools ca apar VM-uri din SS.

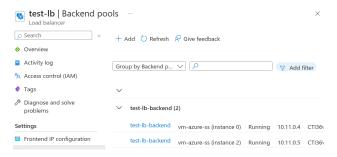


Daca intrati pe IP-ul public trebuie sa vedeti masina din SS



Pentru ca a durat mai mult timp sa cream LB-ul, a mai ramas un singur VM din cele 2 create initial. Sa modificam SS-ul sa aiba minim 2 VM-uri tot timpul. Intrati in SS si selectati din partea stanga Scaling. Jos la Instance limits selectati intre 2 si 5 VM-uri si dati Save; intrati in Instances si verificati ca exista 2 VM-uri cu status-ul Running.

 In Load Balancer la Backend pools trebuie sa vedeti cele 2 instante:



 Daca dati refresh de mai multe ori trebuie sa vedeti ca primiti raspunsuri diferite (de la cele doua masini):



Daca nu va merge din browser incercati comanda:

watch -n 0.1 curl -s <IP\_LB>

 Pentru a testa ca scalarea functioneaza, folositi un program care sa interogheze site-ul de foarte multe ori:

sudo apt-get install apache2-utils ab -c 100 -n 10000000 http://<IP LB>/

- Daca VM-ul folosit pentru teste nu este foarte puteric, nu se va produce scalarea. Folositi mai multe VM-uri pentru testare sau intrati in Scale Set -> Scaling -> Rules si selectati un procent pentru CPU mult mai redus (ex: scalare in sus la 10%, scalare in jos la 5%).
- Dupa ce CPU-ul e mai mare decat threshold-ul selectat pentru mai mult de 5 minute, o sa apara mai multe VM-uri in SS:



- Pentru a vedea si scalarea in jos opriti programul ab (apasati Ctrl + C).
- Dupa un timp, masinile virtuale vor disparea:



• Stergeti Load Balancer-ul, Scale Set-ul si imaginea.