When your VM is starting to run out of resources,

Azure will automatically add another VM

from a scale set.

This scale set is a pool of identical VMs,

(there's more here, as you can see)

that you have defined that are ready to go.

So if you have 90% usage in the first VM,

then effectively, you've just become 45% in two

instantly, well almost instantly.

When they start filling up,

another one can be added as well.

And when resource usage go down again,

The VM, or glasses, will be removed automatically as well.

Back to your scale set.

Scale sets are a way to ensure high availability

and application resiliency for your application.