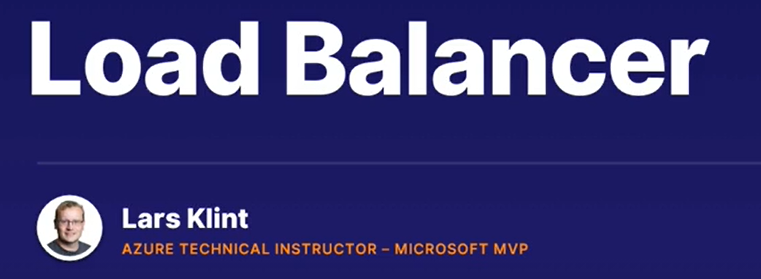
1. (Lars Klint) When you have more than one VM serving the application, how do you decide which VM gets a particular user?
2. A load balancer is the answer.



1. Imagine this scenario. Your online business for booking removals of

angry llamas is going super well. You're starting to get much more traffic

than you originally anticipated.

1. The VM that you use for processing some of the old data is getting overloaded at times.
2. You add a second VM, but how do you now manage the traffic flow and use both?

Diagram

Description automatically generated

1. Adding a load balancer in front of the two VMs to capture the traffic before it reaches them means you can manage where the traffic goes.

Diagram

Description automatically generated

1. As Microsoft puts it, load balancer distributes new inbound flows that arrive on the load balancer's frontend to backend pool instances, according to rules and health probes.
2. So let's break that down.
3. Inbound flows is any kind of traffic, whether from the internet or from a local VNet. Any inbound traffic will be received by the load balancer. The load balancer's frontend is the gatekeeper of all traffic to the VMs it manages.
4. In effect, whoever calls the service the VMs are running only sees one access point, which is the load balancer's frontend.
5. Backend pool instances are the VMs that are receiving the traffic from the load balancer.
6. And rules and health probes refers to the load balancer rules for directing the traffic.

More on that in just a sec.

1. A health probe is a service that makes sure a VM is ready to receive traffic

before the load balancer sends any. This ensures the load balancer always sends

traffic to a healthy VM.

Graphical user interface

Description automatically generated

1. Some of the scenarios where you would apply a load balancer includes load balancing incoming traffic - as we went through in the example at the start of the lecture - load-balanced traffic from internal networks across both Azure and hybrid scenarios, port forwarding of traffic - you can forward traffic to a specific port on a specific machine in a backend pool.
2. And a port is a way to specify a specific process. Each process on a computer listens for traffic and messages on one or more ports.
3. A load balancer can also provide outbound connectivity for the VMs in the backend pool should that be necessary.

A picture containing graphical user interface

Description automatically generated