Overcoming Challenges of Data Center with the help of Virtualization

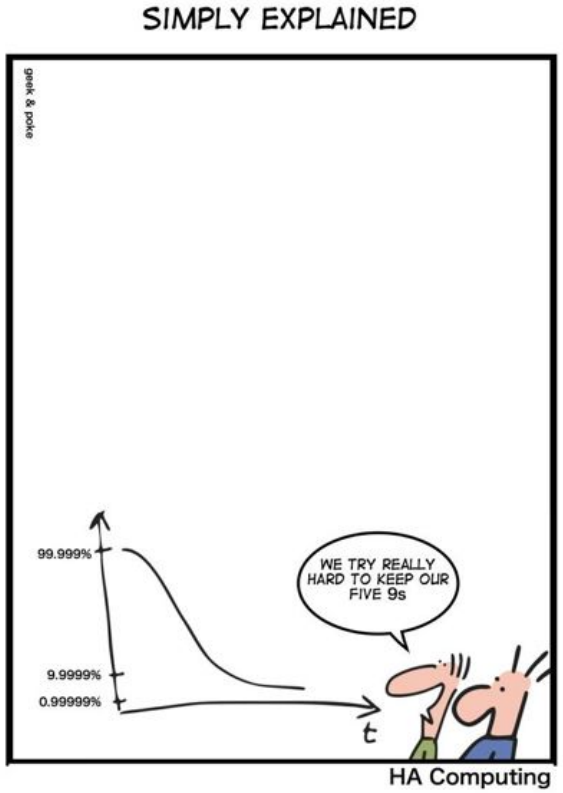
#Source VMware Docs

Asst. Prof. Ashwini Mathur

Major Challenges !!

- **Challenge** 1: Maintaining Availability and Uptime
- **Challenge** 2: Maximize the Fault tolerance
- **Challenge** 3: Improving Utilization of Capacity (Power, Cooling, Space)
- **Challenge** 4: Reporting Reduced Operating Expenses.
- **Challenge** 5: Managing Energy Usage & Costs.

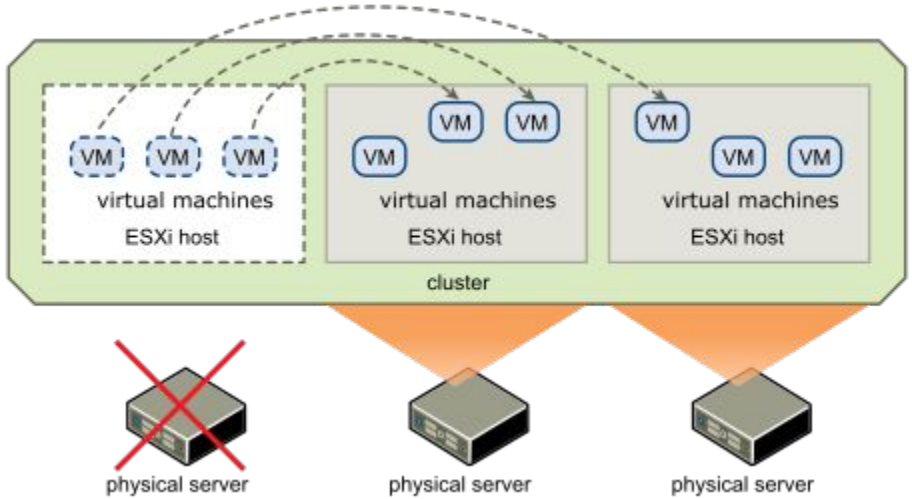
HIGH AVAILABILITY



Let’s understand, What is High Availability ?

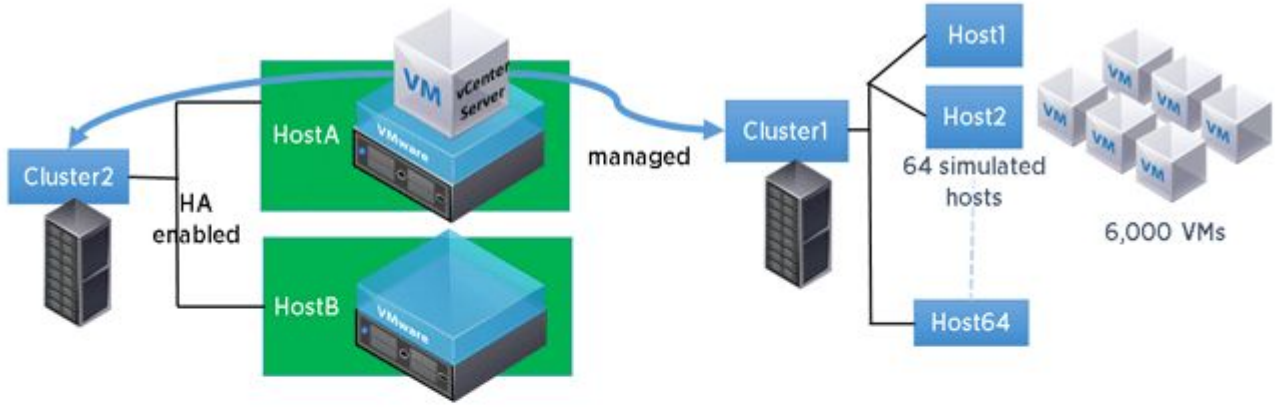
High Availability provides **uniform, cost-effective failover protection** against hardware and operating system outages within your virtualized IT environment. High Availability allows you to:

- Monitor Hosts and virtual machines to detect hardware and guest operating system failures.
- Restart virtual machines on other hosts server in the cluster without manual intervention when a server outage is detected.
- Reduce application downtime by automatically restarting virtual machines upon detection of an operating system failure.

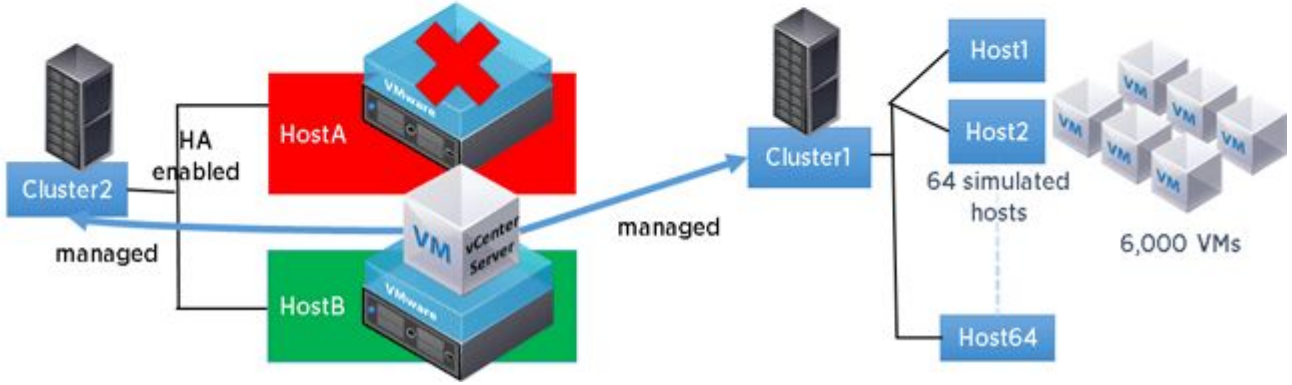

VM
VM
VM
virtual machines
ESXi host
VM
VM
VM
virtual machines
ESXi host
VM
VM
VM
virtual machines
ESXi host
cluster
physical server
physical server
physical server

High availability in the perspective of Data center

Before:



After:

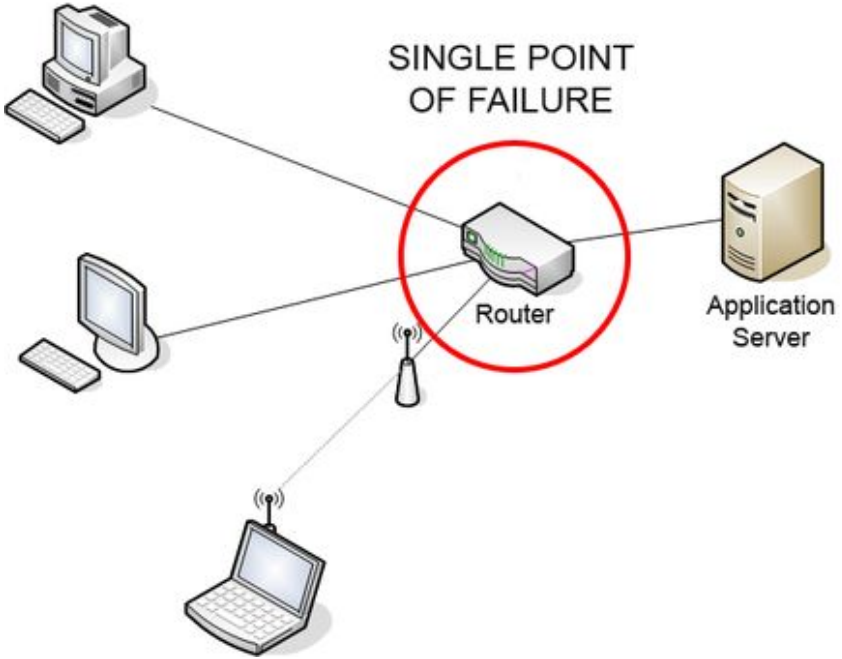


Fault Tolerance

Always remember “high fault tolerance is alway good”

Situation of low fault tolerance , k = 1

Nice !! example to understand low fault tolerance



Don’t situation

FT provides **continuous availability** for such a virtual machine by creating and maintaining another VM that is identical and continuously available to replace it in the event of a failover situation.

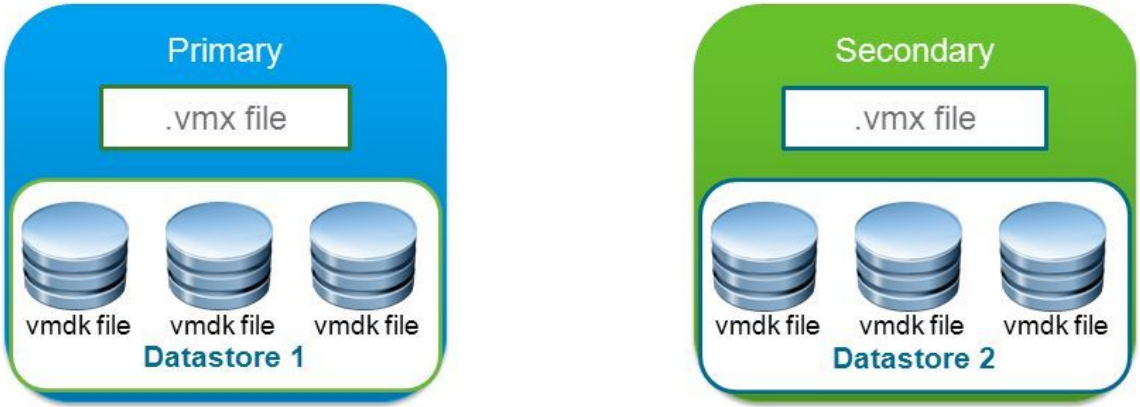
The **protected virtual machine** is called the **Primary VM**. The duplicate virtual machine, the Secondary VM, is created and runs on another host. The primary VM is continuously replicated to the secondary VM so that the secondary VM can take over at any point, thereby providing Fault Tolerant protection.

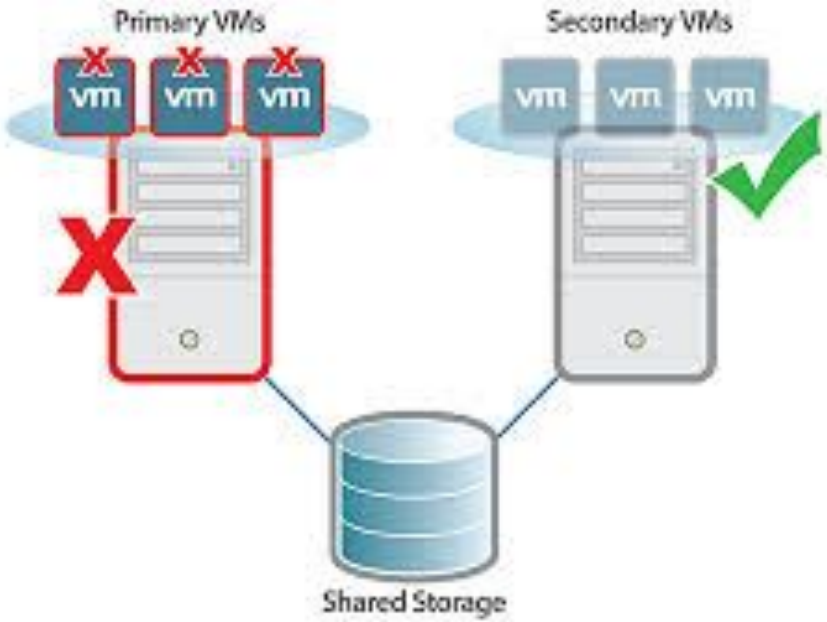
The Primary and Secondary VMs continuously monitor the status of one another to ensure that Fault Tolerance is maintained. A transparent failover occurs if the host running the Primary VM fails, in which case the Secondary VM is immediately activated to replace the Primary VM.

A new Secondary VM is started and Fault Tolerance redundancy is reestablished automatically. If the host running the Secondary VM fails, it is also immediately replaced. In either case, users experience no interruption in service and no loss of data.

A fault tolerant virtual machine and its secondary copy are not allowed to run on the same host. This restriction ensures that a host failure cannot result in the loss of both VMs.

Each virtual machine has its own `.vmx` configuration file and `.vmdk` files. Each of these virtual machines can be on a different datastore.




Primary VMs
vm
vm
vm
Secondary VMs
vm
vm
vm
Shared Storage

Now time to answer one simple question ?

Analyze the Difference between HA and FT ..

Analytical difference points

HIGH AVAILABILITY

A High Availability system is one that is designed to be available 99.999% of the time, or as close to it as possible. Usually this means configuring a failover system that can handle the same workloads as the primary system.

A Fault Tolerant system is extremely similar to HA, but goes one step further by guaranteeing zero downtime. HA still comes with a small portion of downtime, hence the ideal of a perfect HA strategy reaching “five nines” rather than 100% uptime.

The time it takes for the intermediary layer, like the load balancer or hypervisor, to detect a problem and restart the VM can add up to minutes or even hours over the course of yearly runtime.

FT ensures availability by keeping VM copies on a separate host machine.

Rest three challenges, we will briefly discuss in unit 4