# A basic introduction to VMware for Geeks

Preface:

VMware is the worlds largest provider of Virtualisation software with an annual turnover of over $5 Billion per year. It was initially founded by Diane Green and her husband in 2001 and the first version of what is now the ESX Hypervisor was released in 2001.

The components of VMware

**Type 1**

VMWare ESXi is a type 1 hypervisor. A type 1 hypervisor means its runs a bare metal hypervisor. There is very little load as there is only the thinest sliver of a system between the virtual guests and the hardware. Just enough to do the essentials to provide the services required to run the infrastructure.

**Type 2**

Type 2 hypervisors run within a fully fledged OS. An example of this would be using VMWare Fusion or Workstation on your desktop. All the commands that the virtual machine issues has to perform additional steps to get to the real physical hardware level as it has to talk to the OS and in turn the OS talks to the hardware on the virtual machines behalf, which obviously adds additional system load and delays.

The VMWare ESXi system has several components inside the hypervisor itself, which is free to download. There are CPU compute limitations though without a paid for license. You are limited to a single socket, but with unlimited cores. You also miss out on features such as CPU and RAM hot add facility, where supported by the guest OS.

The ESXi Hypervisor itself is logically divided into separate components. These are:

 **Compute:**

 RAM and CPU that can be consumed by the guest operating systems.. One very intelligent feature is memory de duplication. The VMware name for it is shared paging. If you have four virtual machines that run Windows 2008, a lot of the memory pages will be the same, so the machines keep one copy and use pointers for the rest of the machines with identical pages. This can dramatically reduce the amount of physical RAM required.

 **Storage:**

In a similar way to the way the memory is de duplicated, server disks have a lot of space that are not used, or very slowly over time. All this space that is not used is expensive, especially with external shared storage. A feature called thin provisioning is available. In essence it allows you to allocate the disk size you want to a guest but not allocate the storage at once. Put in simple terms it allows you to over allocate the storage.

 **Network:**

 Networks for the VMs as well as the management interface. The network infrastructure provides network capabilities by providing a layer two bridge for the traffic to move across. It can deal with multiple networks and works at the layer below IP so it doesn't care about the network.



Vcenter

On top of the hypervisor, you can run VMware Vcenter. Vcenter is an additional paid for product where all the cutting edge goodness happens (and you pay handsomely for it). Vcenter allows you to take the power and utility of a single host with several guest machines on it but create a cluster. When you run Vcenter and use it to manage your clusters you get access to additional features. You can create a cluster in less than 5 clicks! Some of the features (to name just a few) include:

 Centralised management

 You can use a single piece of glass to manage your entire infrastructure. You can even manage multiple sites using the same Vcenter. When you scale to 10's, 100's or 1000's of guests it becomes absolutely essential if only to keep your sanity!

 High Availability (HA)

 HA is a feature that allows you to utilise the power of a cluster of ESXi Hypervisors to ensure that should one of your machines suffer a hardware failure, the guest machines that resided on that host will be restarted on a surviving node. There will be downtime whilst the machine restarts.

 Distributed Resource Scheduling (DRS)

 DRS is a system that cleverly allows you to balance your virtual guest machines across the hardware in the cluster to maximise efficiency and even out the load. It will automatically evaluate the load on all the hosts and if any physical host machine is overloaded or a VM would be a better fit elsewhere on the cluster it will seamlessly migrate the machine to another cluster member whilst still maintaining the state of the machine, i.e. without any outages or blips. DRS is fully configurable in terms of sensitivity and also how much automation happens, as well as when. You can even put it in manual so no DRS happens, but you only want to do that in limited situations such as planned power downs.

 Fault Tolerance (FT)

 For those machines that you can't afford to be down for even a moment, you can use FT. FT effectively keeps a "warm" copy of the VMware Guest machine and if the host fails (or indeed a software failure is detected) the warm copy will instantly become live and take over serving the request. There are limitations. You cannot have more than 1 vCPU per FT VM. Also any resources used will be doubled as the shadow copy consumes the same resources although it is not used. There is also a limit to the number of FT machines than can run in single cluster.

 Templates and automated configurations

 What is exceptionally useful is the ability to create virtual machines (Windows and Linux) and turn them into templates. This allows an administrator to create a machine with all the required patches and configuration and then convert it into a template. You can then deploy several machines based on the same template. This can save hours of work per VM. It is also highly configurable whilst deploying the template including networks, disk size, cpu and RAM to name but a few.

There are additional components that ship with Vcenter (depending on licence type) but they are beyond the scope of this very simple overview.

How High Availability works:

In ESXi 5, there is a new purpose built watchdog service called Fault Domain Manager that communicates between hosts (and is managed by the Vcenter). It communicates between physical hosts in the cluster about the status of the physical hosts and the current load. Cleverly each host keeps a copy of physical loads and machines running on each node in the cluster.

In the event of a failure the surviving nodes have a good overview of remaining capacity and know which guests need to be restarted. Hosts will communicate with each other and restart the machines on the physical host with most free capacity. To stop the machines swamping a host DRS will step in and balance the loads as needed. Clever stuff this VMware!

HA works in conjunction with, but independently of Vcenter. It needs to be this way because there is always the possibility of the physical host failing that contains the Vcenter. If restart depended on Vcenter being up, the issue is obvious. There is a process called hostd that communicates between the host and the vCenter.

To be a bit more specific on how HA works (There are entire books written on the subject) but basically if a host fails the following happens.

Each cluster has a single master host and every other machine is a slave host. In the event of a slave host failing, the master node will control the distribution and restart of the failed guests on the available hosts. If the Master host fails, the suriving physical hosts hold an election, based on a number of criteria, and elect a new Master, which then assumes the role and starts the restart process.

The Virtual Machine Manager (VMM)

Sitting between the Hypervisor and the virtual guest is the VMM. Simply put the VMM acts as a sandbox environment for the VM. It prevents direct interaction between the host and the VM so that an errant VM cannot affect any of the other VMs on the infrastructure (At least directly by bad calls or actions) It also provides a standard predictable environment that hides the complexity of the hypervisor side from the guest. The guest just sees a real PC and doesn't know it is a virtual machine

.

Top tips for deploying VMware.

 Always make sure your guest OS has VMware tools installed (This is easily from the console to the VM). It will enable page sharing thereby reducing the RAM overhead, as well as RAM management strategies in case of shortage. It also provides highly optimised drivers to make your machine fly.

 Use Paravirtual disks where possible - Makes the disk perform much better, if supported by the guest OS

 Install VMNET3 network drivers. Once VMware tools is installed, you should remove the network from the VM and add a new VMXNET3 adapter. It is a very highly optimised networking driver that provides much improved performance over and above the E1000 network adapter.v