System & Service Management
Virtualisierung

Virtual Machines
System virtual machines (sometimes called hardware virtual machines) allow the sharing of the underlying physical machine resources between different virtual machines, each running its own operating system. The software layer providing the virtualization is called a virtual machine monitor or hypervisor. A hypervisor can run on bare hardware (Type 1 or native VM) or on top of an operating system (Type 2 or hosted VM).

The main advantages of VMs are:
- multiple OS environments can co-exist on the same computer, in strong isolation from each other
- the virtual machine can provide an instruction set architecture (ISA) that is somewhat different from that of the real machine
- application provisioning, maintenance, high availability and disaster recovery

The main disadvantages of VMs are:
- a virtual machine is less efficient than a real machine when it accesses the hardware indirectly
- when multiple VMs are concurrently running on the same physical host, each VM may exhibit a varying and unstable performance (Speed of Execution, and not results) , which highly depends on the workload imposed on the system by other VMs, unless proper techniques are used for temporal isolation among virtual machines.

Multiple VMs each running their own operating system (called guest operating system) are frequently used in server consolidation, where different services that used to run on individual machines in order to avoid interference are instead run in separate VMs on the same physical machine.

http://en.wikipedia.org/wiki/Virtual_machine

Operating System-level Virtualization
Operating System-level Virtualization is a server virtualization technology which virtualizes servers on an operating system (kernel) layer. It can be thought of as partitioning: a single physical server is sliced into multiple small partitions (otherwise called virtual environments (VE), virtual private servers (VPS), guests, zones, etc.); each such partition looks and feels like a real server, from the point of view of its users.

For example, Solaris Zones supports multiple guest OSes running under the same OS.

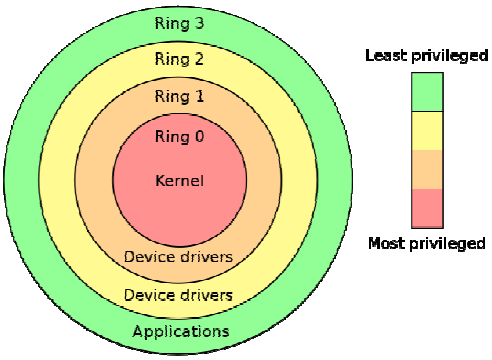
http://en.wikipedia.org/wiki/Virtual_machine
http://en.wikipedia.org/wiki/Logical_partition_(virtual_computing_platform)

Hypervisor
A hypervisor, also called virtual machine monitor (VMM), is one of many virtualization techniques which allow multiple operating systems, termed guests, to run concurrently on a host computer.

- Type 1 (or native, bare metal) hypervisors run directly on the host's hardware to control the hardware and to monitor guest operating systems. A guest operating system thus runs on another level above the hypervisor.
  (Examples: VMware ESXi or Microsoft Hyper-V)
- Type 2 (or hosted) hypervisors run within a conventional operating system environment. With the hypervisor layer as a distinct second software level, guest operating systems run at the third level above the hardware.

http://en.wikipedia.org/wiki/Hypervisor

Exkurs: Hierarchical Protection Domains
Hierarchical protection domains, often called protection rings, are a mechanism to protect data and functionality from faults and malicious behaviour.



On most operating systems, Ring 0 is the level with the most privileges and interacts most directly with the physical hardware such as the CPU and memory.

Special gates between rings are provided to allow an outer ring to access an inner ring's resources in a predefined manner, as opposed to allowing arbitrary usage. Correctly gating access between rings can improve security by preventing programs from one ring or privilege level from misusing resources intended for programs in another. For example, spyware running as a user program in Ring 3 should be prevented from turning on a web camera without informing the user, since hardware access should be a Ring 1 function reserved for device drivers.

http://en.wikipedia.org/wiki/Ring_(computer_security)

Software-based Virtualisation
In protected mode the operating system runs at a higher privilege ring such as 0, and applications at a lower privilege such as ring 3. Similarly, a host OS must control the processor while the guest OS' are prevented from direct access to the hardware. One approach used in x86 software-based virtualization is called ring deprivileging, which involves running the guest OS at a ring higher than 0. One possible technique is Binary Translation:

Binary translation is the emulation of one instruction set by another through translation of code. Sequences of instructions are translated from the source to the target instruction set. In some cases such as instruction set simulation, the target instruction set may be the same as the source instruction set, providing testing and debugging features such as instruction trace, conditional breakpoints and hot spot detection.

http://en.wikipedia.org/wiki/X86_virtualization

Paravirtualization
Paravirtualization is a virtualization technique that presents a software interface to virtual machines that is similar (but not identical) to that of the underlying hardware.

The intent of the modified interface is to reduce the portion of the guest's execution time spent performing operations which are substantially more difficult to run in a virtual environment compared to a non-virtualized environment.

Paravirtualization requires the guest operating system to be explicitly ported for the para-API. A conventional OS distribution which is not paravirtualization-aware cannot be run on top of a paravirtualizing VMM.

http://en.wikipedia.org/wiki/Paravirtualization