CS528

Intro to Cloud System

A Sahu

Dept of CSE, IIT Guwahati

Outline

- What is Cloud Computing?
- (HPC, Data Center, Grid) Vs Cloud
- Virtualization
- Advantage of Cloud System : User Prospects
- Dis-advantage of Cloud System : User Prospects

What is Cloud Computing?

- These platforms
  - hide the complexity and details of the underlying infrastructure from users and applications
  - by providing very simple graphical interface or API (Applications Programming Interface).

What is Cloud Computing?

- In addition, the platform provides on demand services, that are always on, anywhere, anytime and any place.
- Pay for use and as needed, elastic
  - Scale Up and Down in capacity and functionalities
- The H/W and S/W services are available to
  - general public, enterprises, corporations and businesses markets

Cloud Summary

- Cloud computing : an umbrella term used to refer to Internet based development and services
- A number of characteristics define cloud data, applications services and infrastructure:
  - **Remotely hosted**: Services or data are hosted on remote infrastructure.
  - **Ubiquitous**: Services or data are available from anywhere.
  - **Commodified**: The result is a utility computing model similar to traditional that of traditional utilities, like gas and electricity - you pay for what you would want!

Cloud Computing Characteristics

Common Characteristics:

- Massive Scale
- Resilient Computing
- Homogeneity
- Geographic Distribution
- Virtualization
- Service Orientation
- Low Cost Software
- Advanced Security

Essential Characteristics:

- On Demand Self-Service
- Broad Network Access
- Rapid Elasticity
- Resource Pooling
- Measured Service

Cloud Computing Characteristics

- **Scalability** Infrastructure capacity allows for traffic spikes and minimizes delays.
- **Resiliency** Cloud providers have mirrored solutions
  - To minimize downtime in the event of a disaster.
  - This type of resiliency can give businesses the sustainability they need during unanticipated events.
- **Homogeneity**: No matter which cloud provider and architecture an organization uses
  - An open cloud will make it easy for them to work with other groups,
  - even if those other groups choose different providers and architectures.

Cloud Computing Characteristics

- *On-demand self-service:* A consumer can unilaterally provision computing capabilities
  - Such as server time and network storage,
  - as needed automatically without requiring human interaction with each service's provider.
- *Broad network access.* Capabilities are available over the network and
  - Accessed through standard mechanisms
  - promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, PDAs).

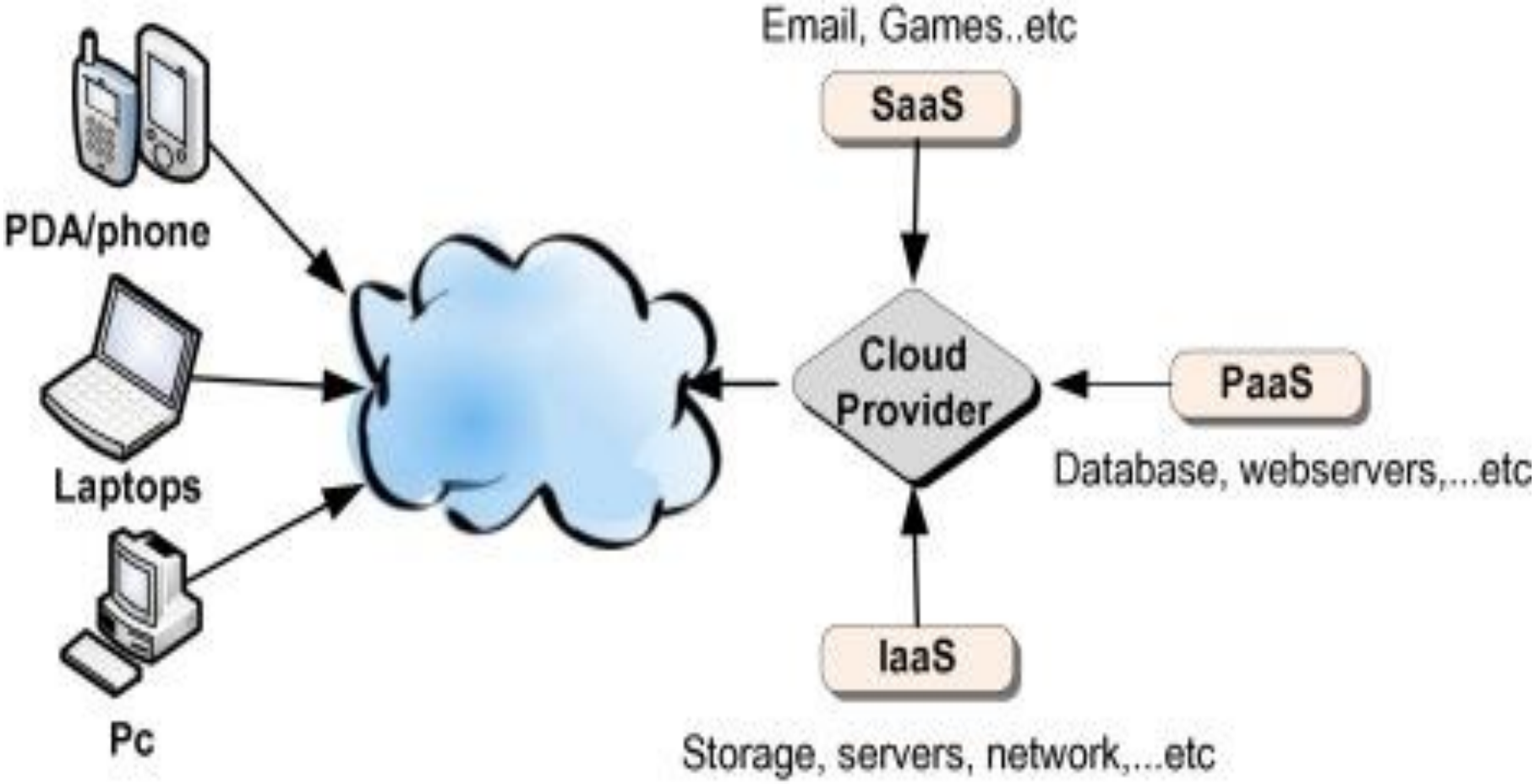
Cloud Computing Characteristics

- ***Resource pooling:*** **Multi-tenant model**
  - There is a sense of location independence in that
  - The customer generally has no control or knowledge over the exact location of the provided resources
  - but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter).
  - Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.
- ***Rapid elasticity.*** **Capabilities can be rapidly and elastically provisioned**
  - In some cases automatically
  - To quickly scale out and rapidly released to quickly scale in.
  - To the consumer, the capabilities available for provisioning often appear to be unlimited
  - can be purchased in any quantity at any time.

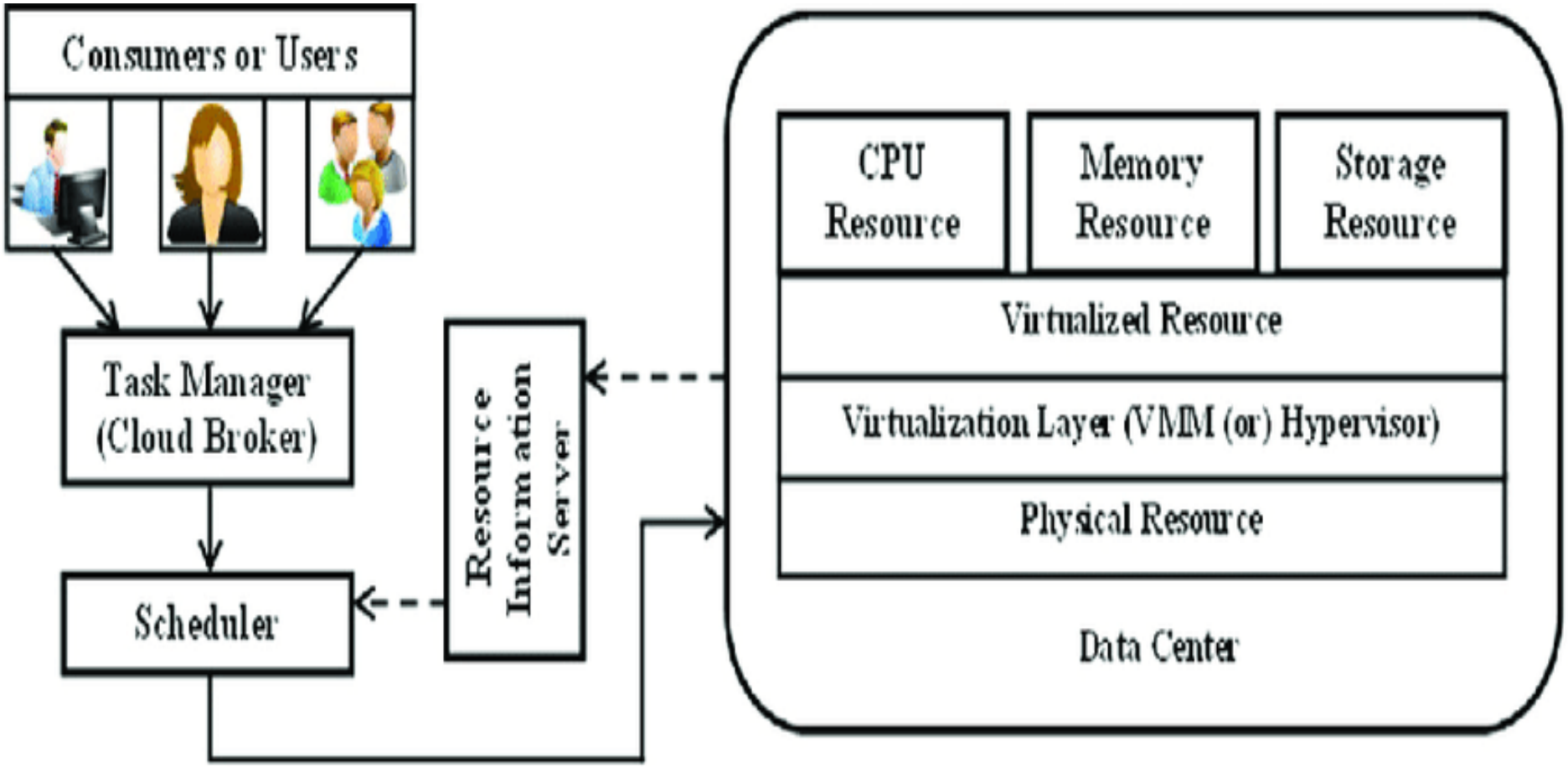
Cloud Computing Characteristics

- *Measured Service*
  - Cloud systems automatically control and optimize resource use
  - by leveraging a metering capability at some level of abstraction appropriate to the type of service
  - E.g., storage, processing, bandwidth, and active user accounts

Cloud System: HL



Cloud System: HL



Cloud Service Models

Software as a Service (SaaS)

Platform as a Service (PaaS)

Infrastructure as a Service (IaaS)

SalesForce CRM

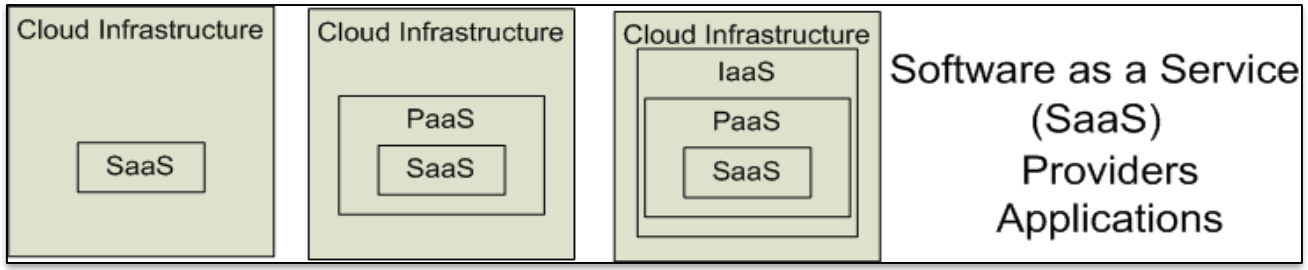
LotusLive

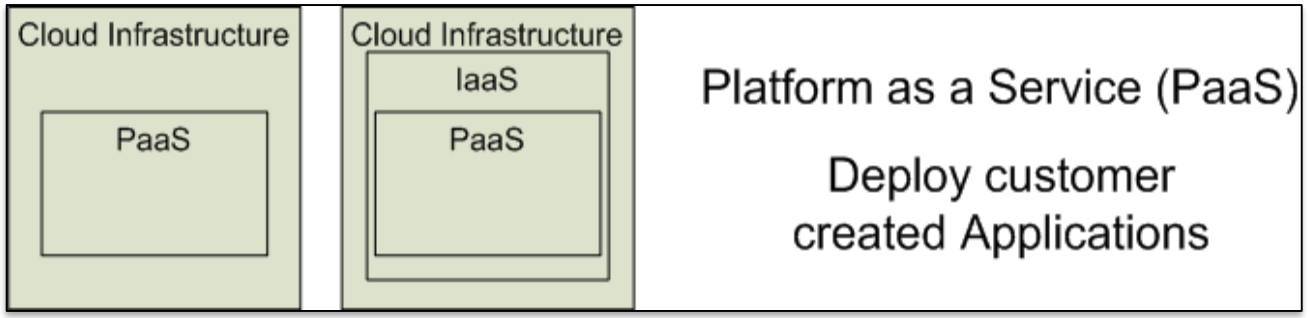
Google App

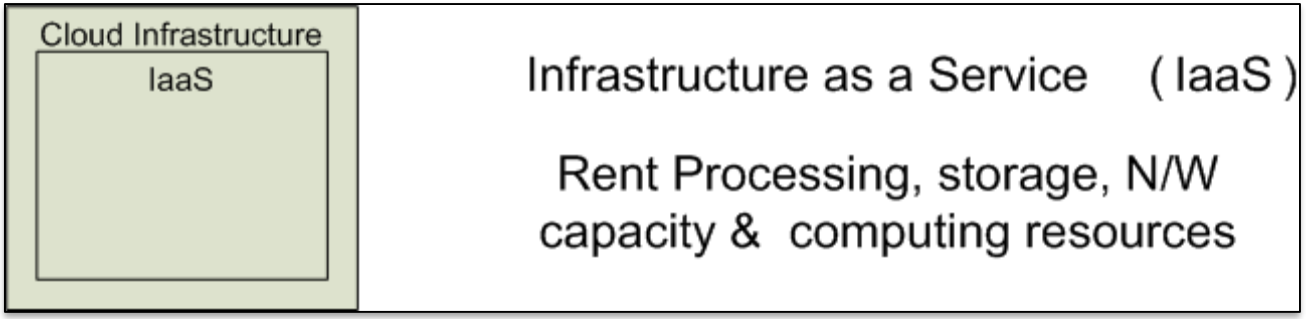
Windows Azure
The Future Made Familiar

amazon web services™

rackspace HOSTING

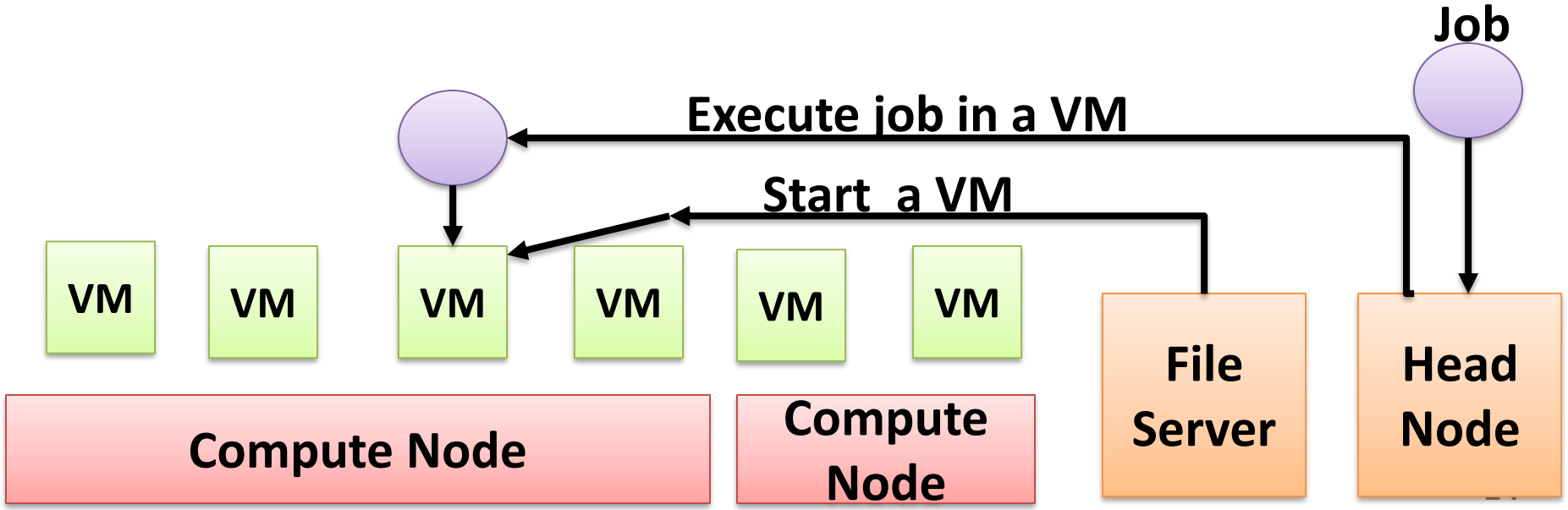






Cloud Computing

- Features of Clouds
  - Scalable, Enhanced Quality of Service (QoS)
  - Specialized and Customized, Cost Effective
  - Simplified User Interface



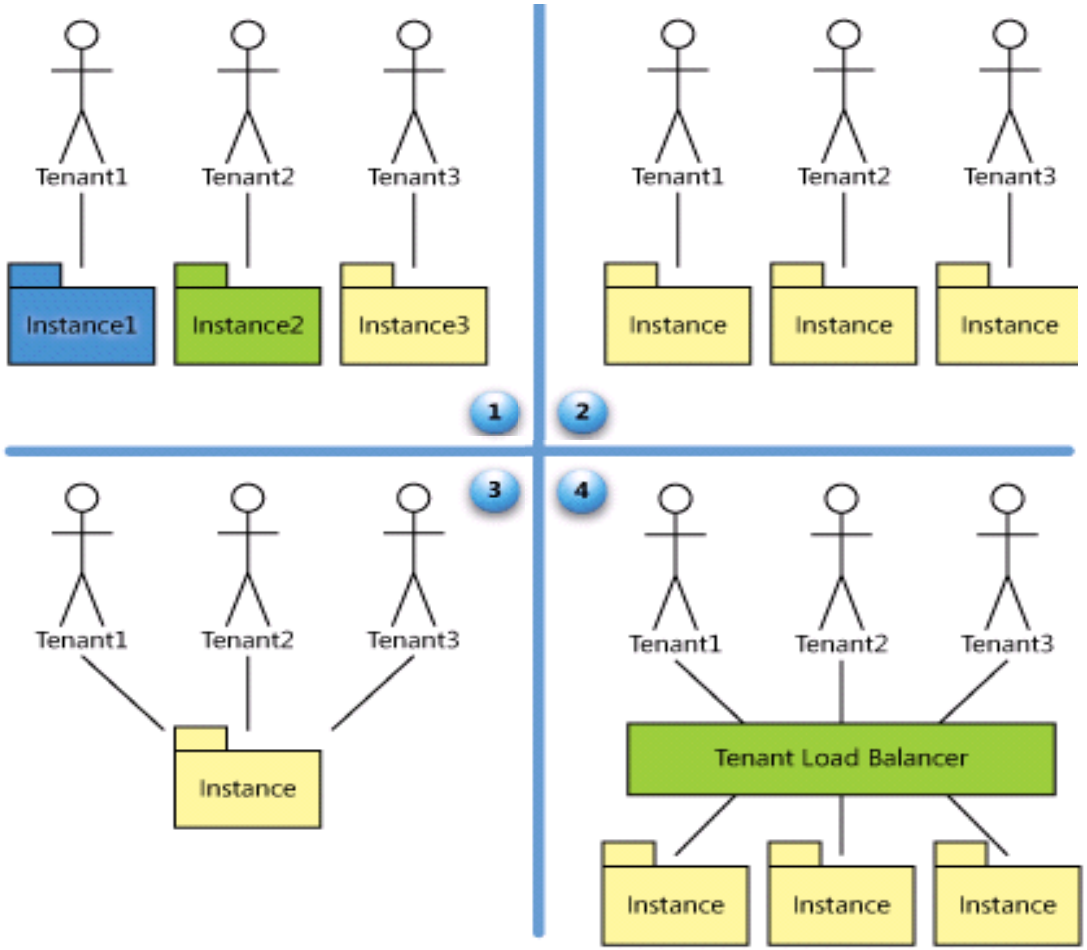
SaaS Maturity Model

Level 1: Ad-Hoc/Custom – One Instance per customer

Level 2: Configurable per customer

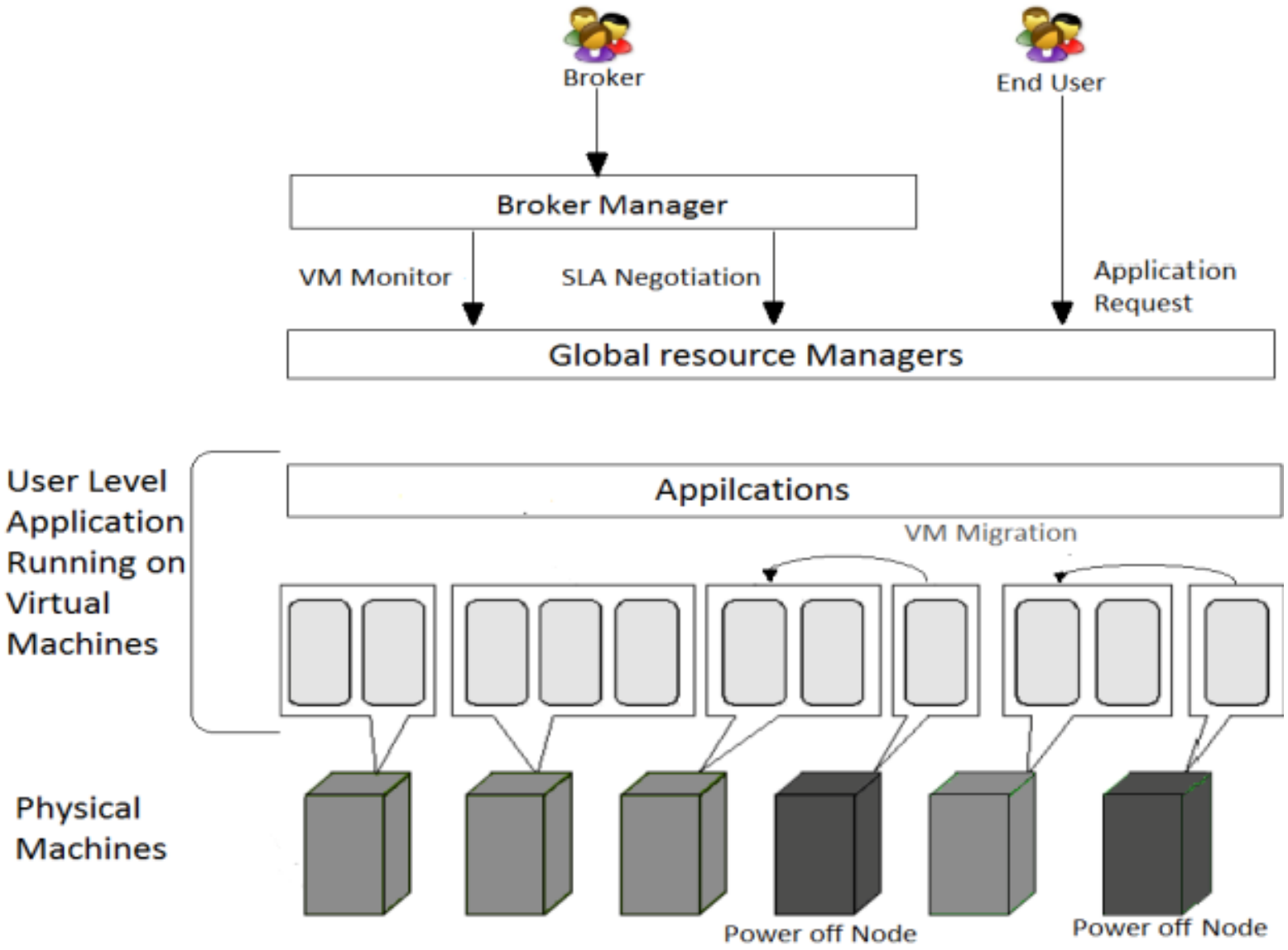
Level 3: configurable & Multi-Tenant-Efficient

Level 4: Scalable, Configurable & Multi-Tenant-Efficient

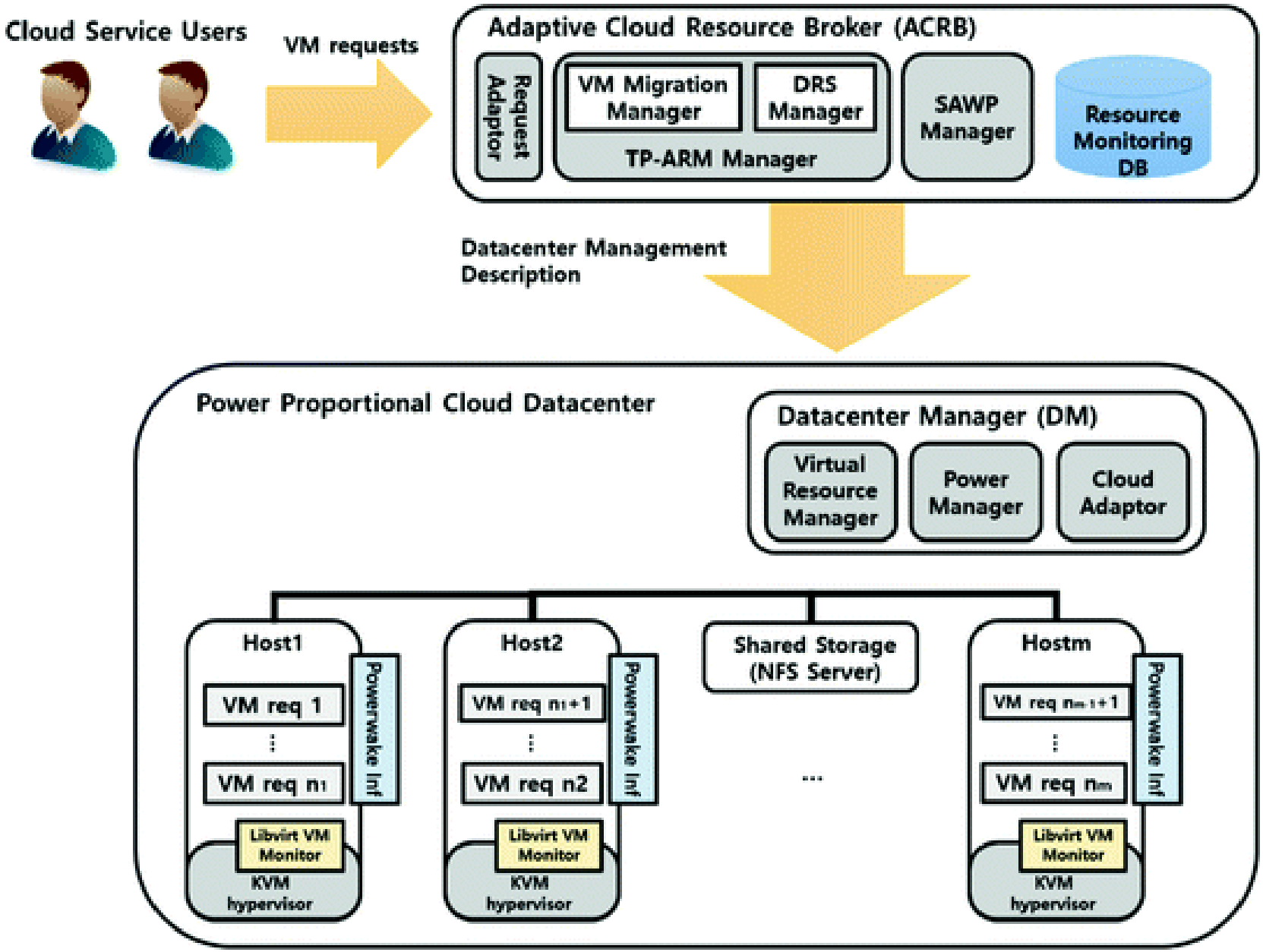


Cloud System: HL

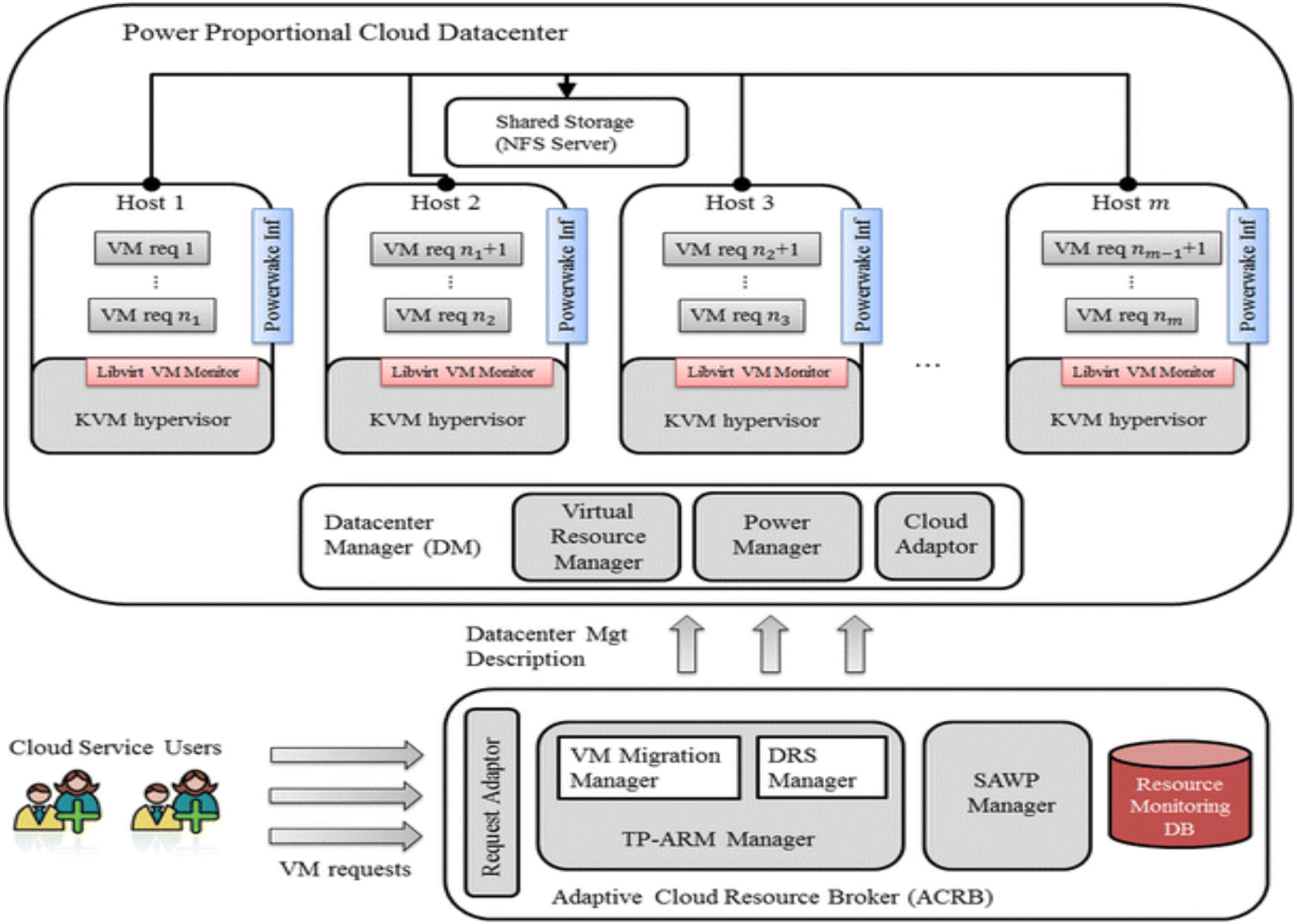
each node.



Cloud System: HL



Cloud System: HL



Cloud computing takes virtualization to the next step

- You don't have to own the hardware
- You "rent" it as needed from a cloud
- There are public clouds
  - e.g. Amazon EC2, and now many others (Microsoft, IBM, Sun, and others ...)
- A company can create a private one
  - With more control over security, etc.

Virtualization

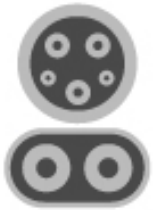
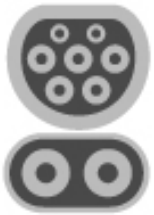
- Abstraction of computer resources.
- Virtualization hides the physical characteristics of computing resources
  - From their users, be they applications, or end users.

Virtualization

- Virtualization includes making a **single** physical resource
  - such as a server, an operating system, an application, or storage device
  - appear to function as **multiple** virtual resources
- Also include making **multiple** physical resources
  - such as storage devices or servers
  - appear as a **single** virtual resource

Virtualization Basic : Truck on Train

EV charger

	N. America	Japan	EU *and the rest of markets*	China	All Markets *except EU*
AC	 **J1772** (Type 1)	 **J1772** (Type 1)	 **Mennekes** (Type 2)	 **GB/T**	 **Tesla**
DC	 **CCS1**	 **CHAdeMO**	 **CCS2**	 **GB/T**	

Mobile Multi Charger

Virtualization Basic

- In OS, Classic example **FILE** as abstract virtual object
- File read/write:
  - fwrite: write data to File
  - fread: read data from File
- Underlying target **File** may be in HDD, Buffer, SSD, Network File, CDROM
- Internal may be diff but externally the same call

Virtualization Basic

- Example**: Virtual BOX , QEMU, Wine, Dalvik, JVM**
- **QEMU : ARM emulation on X86**
- **JVM and Dalvik : Java byte code and in Andriod**
- **Oracle Virtual BOX**
  - Running MS Window OS on Linux Host
  - Running Linux on MS Window Host
- **Cygwin:** Running Linux App on Window
  - Assume running shell script and GIMP in Window
- **Wine**
  - Running MS Window app on Linux, Running your favorite MS-Office in Linux

Virtualization Basic

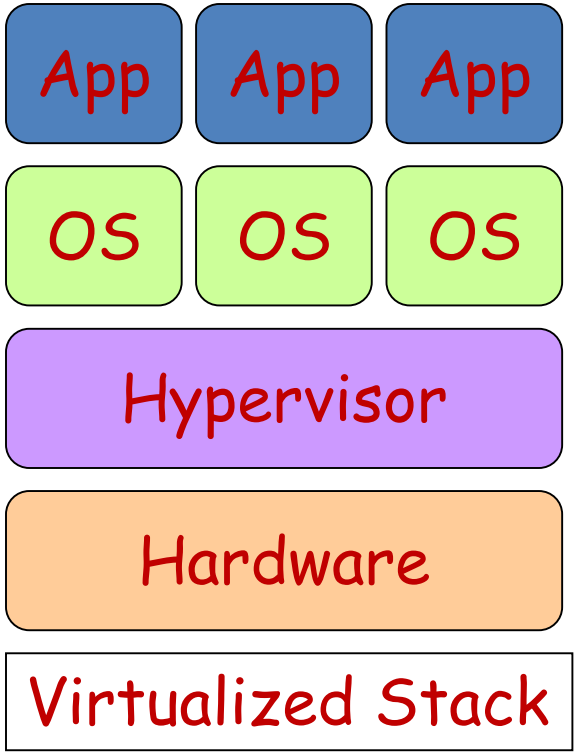
- In OS, Classic example **FILE** as abstract virtual object
- File read/write:
  - fwrite: write data to File
  - fread: read data from File
- Underlying target **File** may be in HDD, Buffer, SSD, Network File, CDROM
- Internal may be diff but externally the same call

Virtualization

- Virtual workspaces
  - An abstraction of an execution environment that can be made dynamically available to authorized clients by using well-defined protocols,
  - Resource quota (e.g. CPU, memory share),
  - Software configuration (e.g. O/S, provided services).

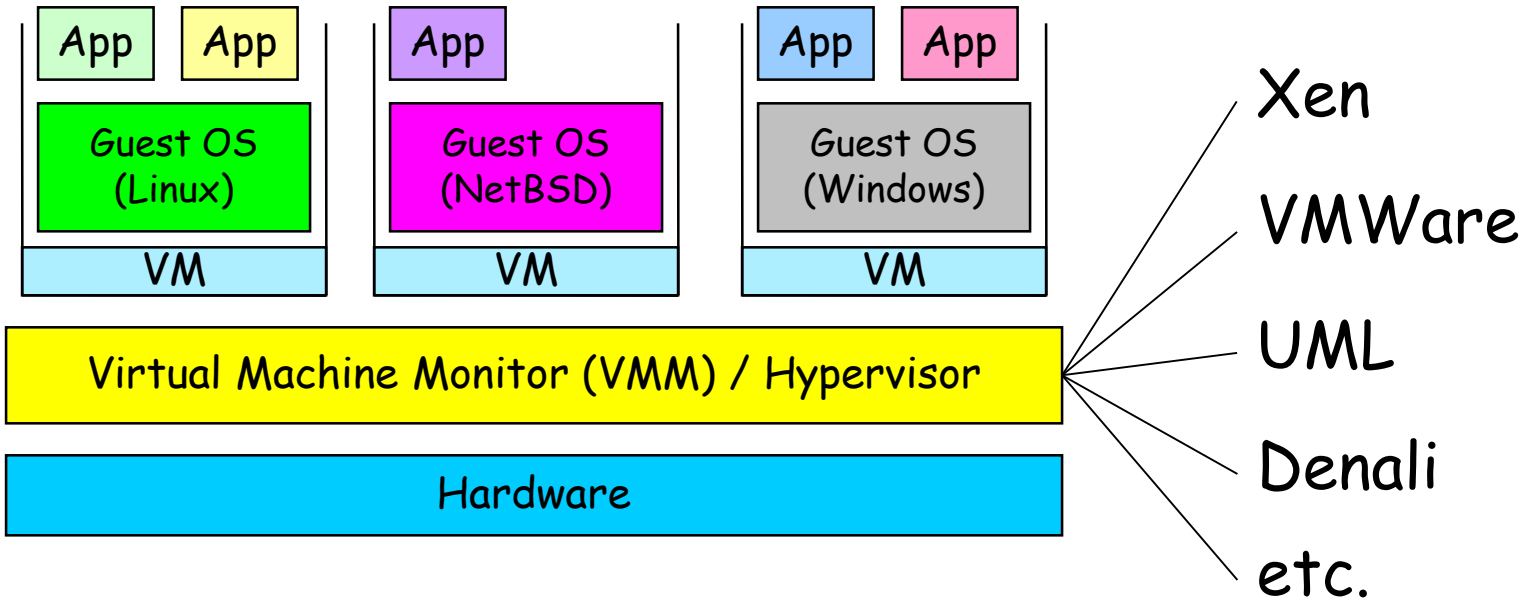
Virtualization

- Virtual WS Implement on Virtual Machines (VMs):
  - Abstraction of a physical host machine,
  - Hypervisor intercepts and emulates instructions from VMs, and allows management of VMs,
  - VMWare, Xen, etc.
- Provide infrastructure API:
  - Plug-ins to hardware/support structures



Virtual Machines

- VM technology allows multiple virtual machines to run on a single physical machine.

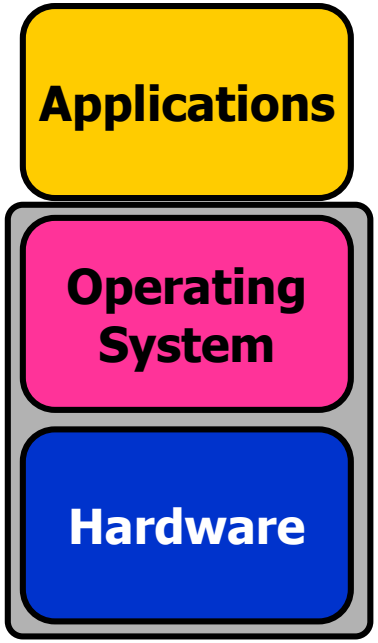


Performance: Para-virtualization (e.g. Xen) is very close to raw physical performance!

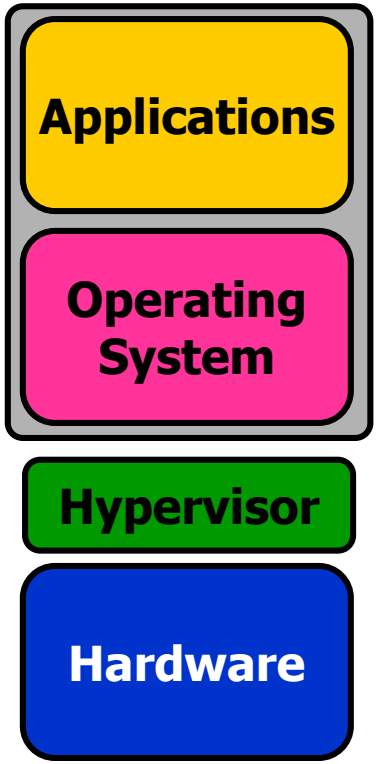
Virtual Machine

- What is Virtual Machine (VM)?
  - **VM** is a software implementation of a machine (i.e. a computer) that executes programs like a real machine.
- Terminology :
  - Host (Target) : The primary environment where will be the target of virtualization.
  - Guest (Source) : The virtualized environment where will be the source of virtualization.

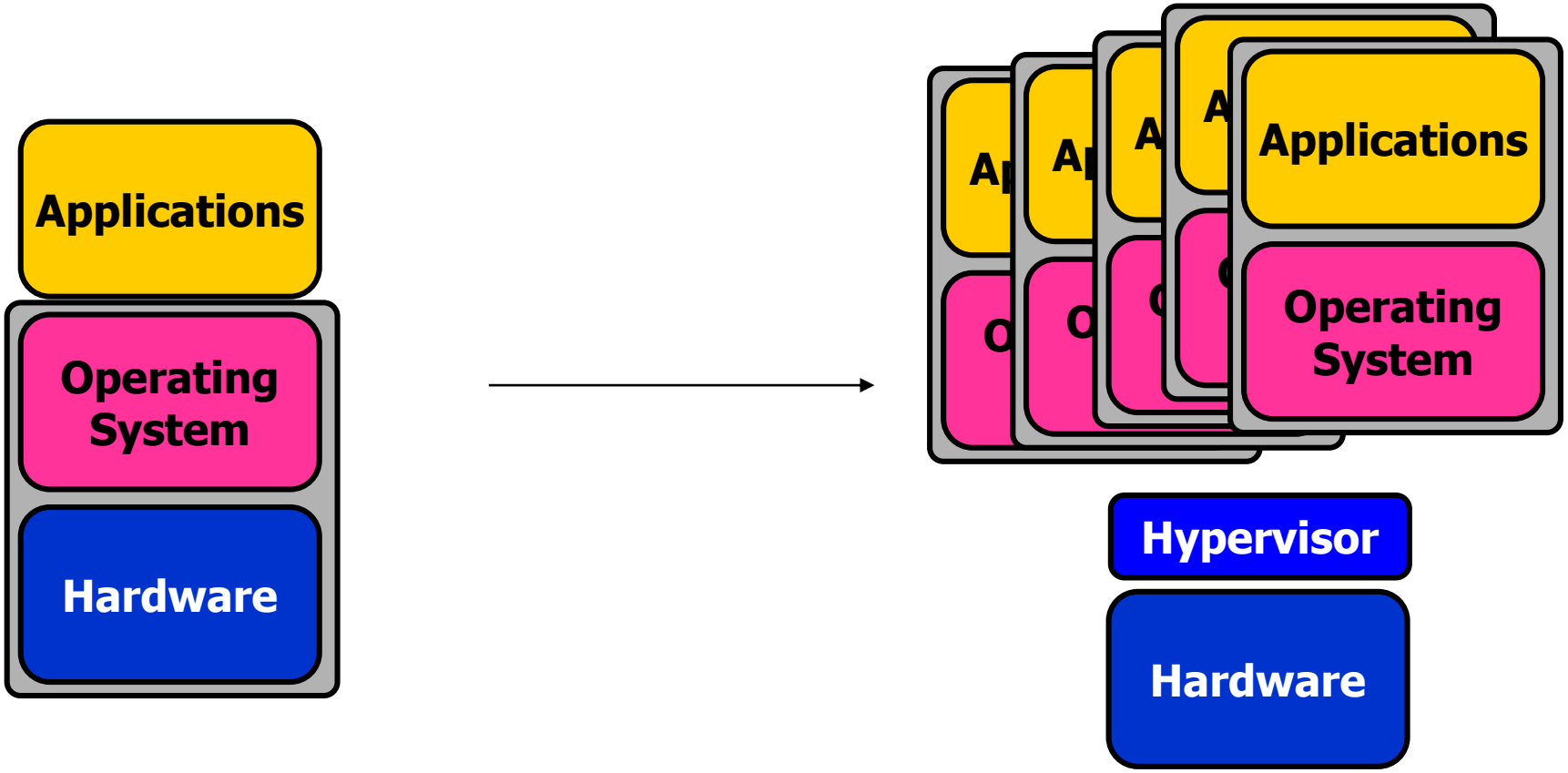
The Use of Computers



Virtualization



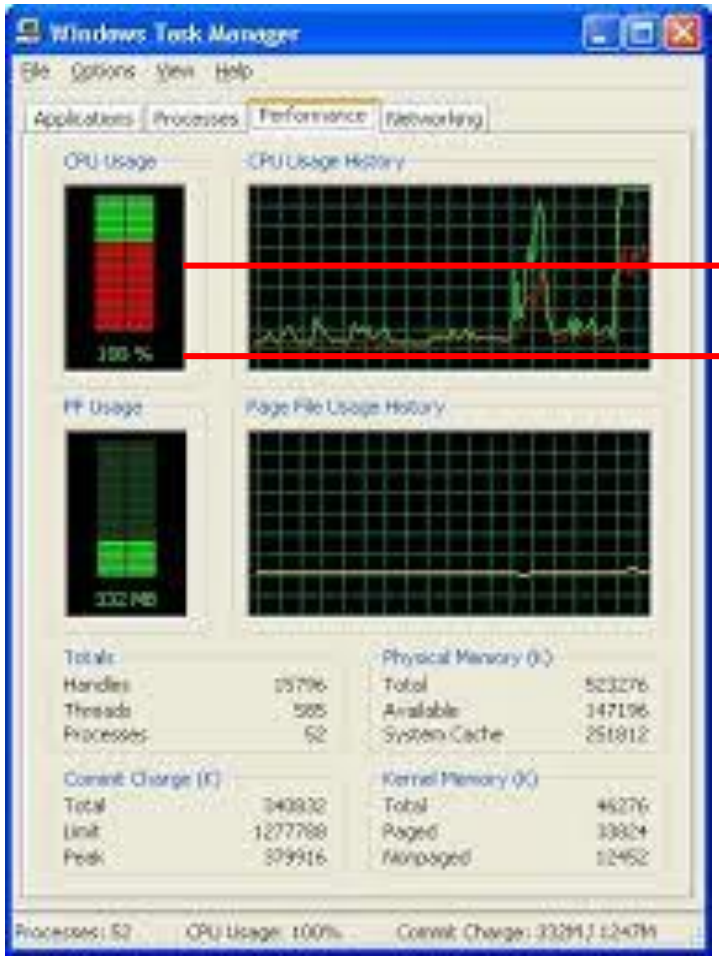
Virtualization -- a Server for Multiple Applications/OS



Virtualization -- a Server for Multiple Applications/OS

- **Hypervisor** is a software program
  - that manages multiple operating systems (or multiple instances of the same operating system)
  - on a single computer system.
- The hypervisor manages the system's
  - processor, memory, and other resources to allocate what each operating system requires.
- Hypervisors are designed for a particular processor architecture
  - and may also be called **virtualization managers**.

Capacity Utilization



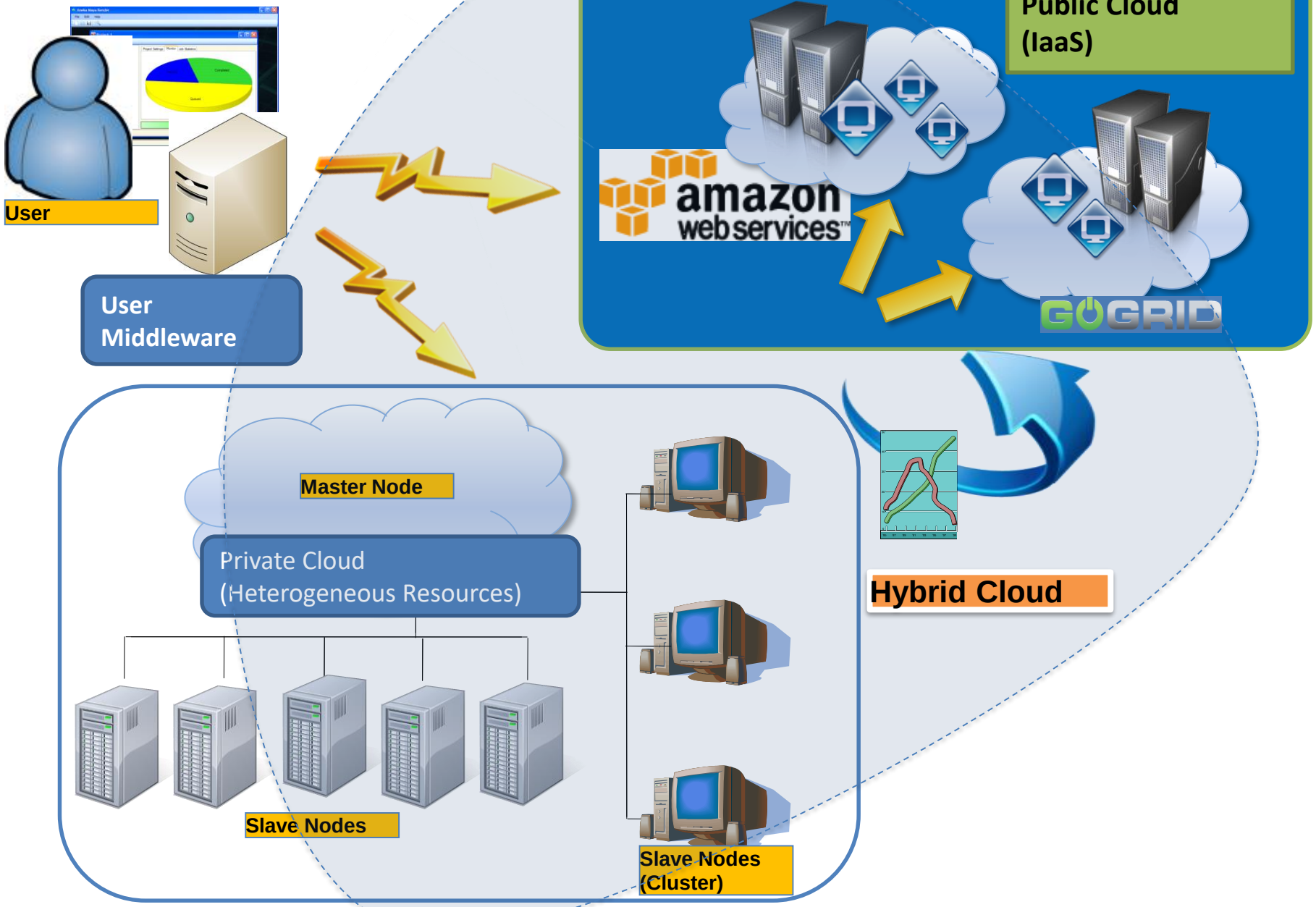
Virtualized system (high)

High utilized*

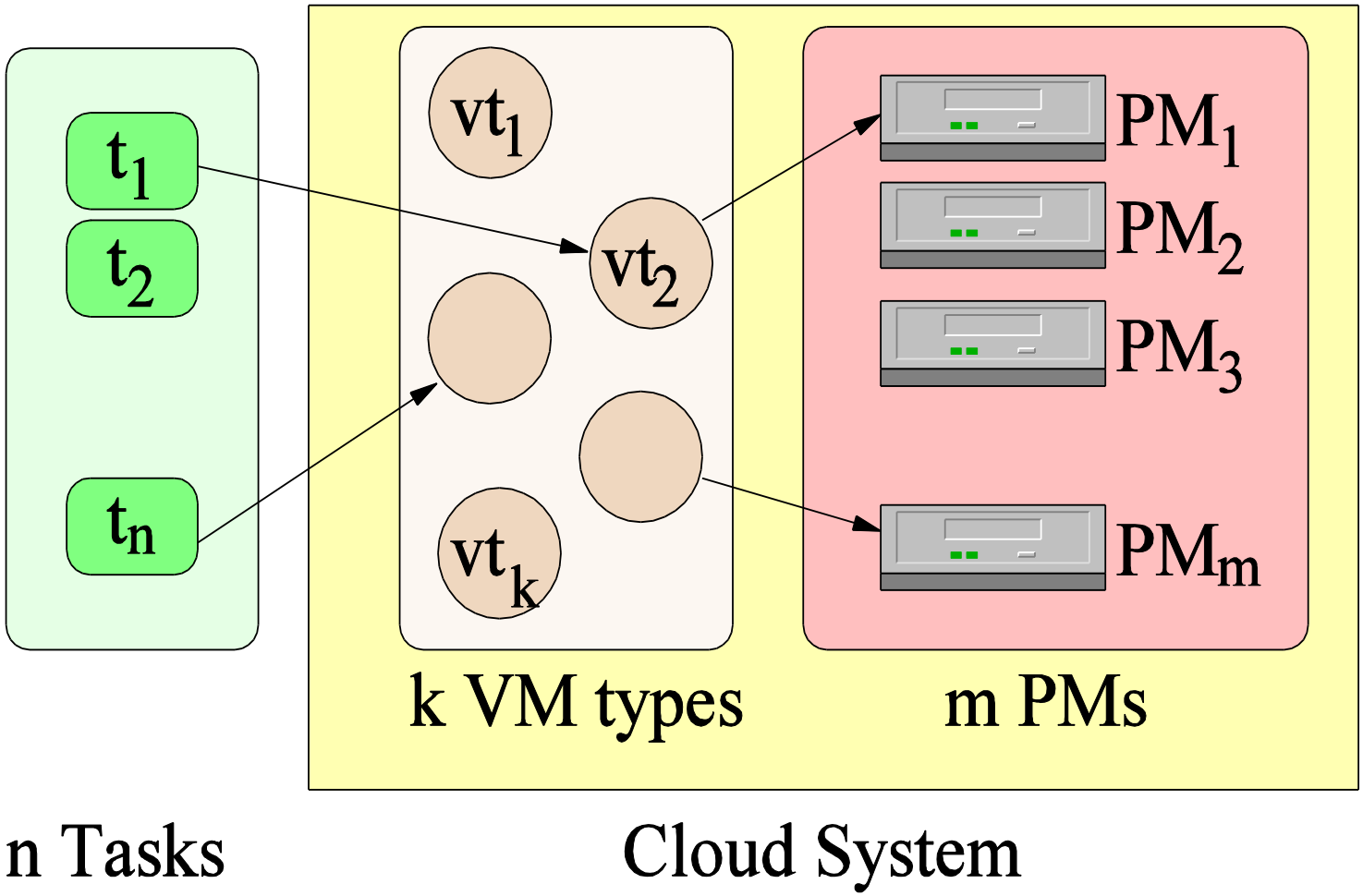
Low utilized

Stand alone system (low)

* But not overloaded…


User
User Middleware
Public Cloud (IaaS)
amazon web services
GOGRID
Master Node
Private Cloud (Heterogeneous Resources)
Slave Nodes
Slave Nodes (Cluster)
Hybrid Cloud

Logical view of Cloud System



Why now?

- 1960—1999
  - IBM, CP-40, CP/CMS, S/360-370, VM370, Virtual PC, VMware
- 2000—2005
  - IBM z/VM, Xen
- 2006
  - Intel VT-x
  - AMD's AMD-V
- 2008—

Hardware evolution

- Faster CPU clock than ever
  - Though almost hit its top
- More CPU cores in a single chip
  - 32/64-core CPUs already in the market
- Multi-core architectures make parallel processing more realizable
- Virtualization support on chip from CPU manufacturers (e.g., Intel, AMD)

Virtualization

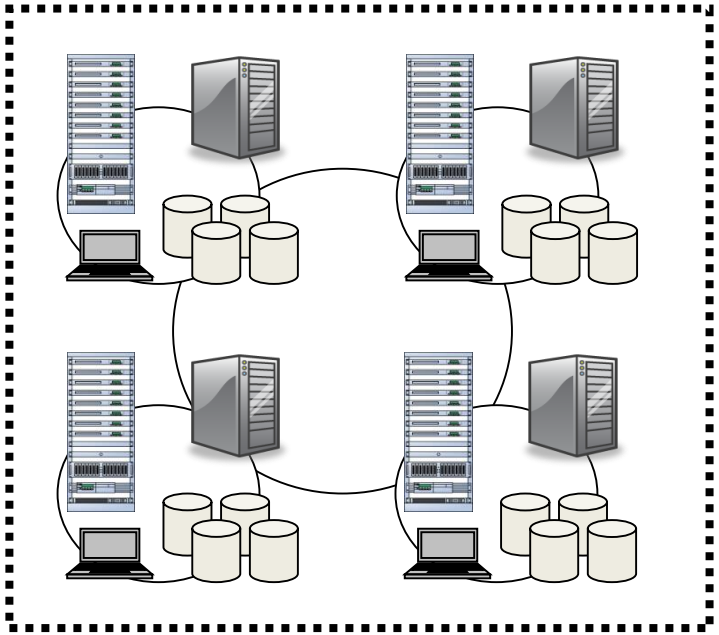
- Binary translation is the most established technology for full virtualization
- Hardware assist is the future of virtualization, but it still has a long way to go
- Para-virtualization delivers performance benefits with maintenance costs
  - Xen
  - VMWare, VBox
- OS level Virtualization: Container/Kubernetics

Issues in Virtualization for Cloud-Computing

- Aspects and expectation from
  - End-user
  - Operator/Manager

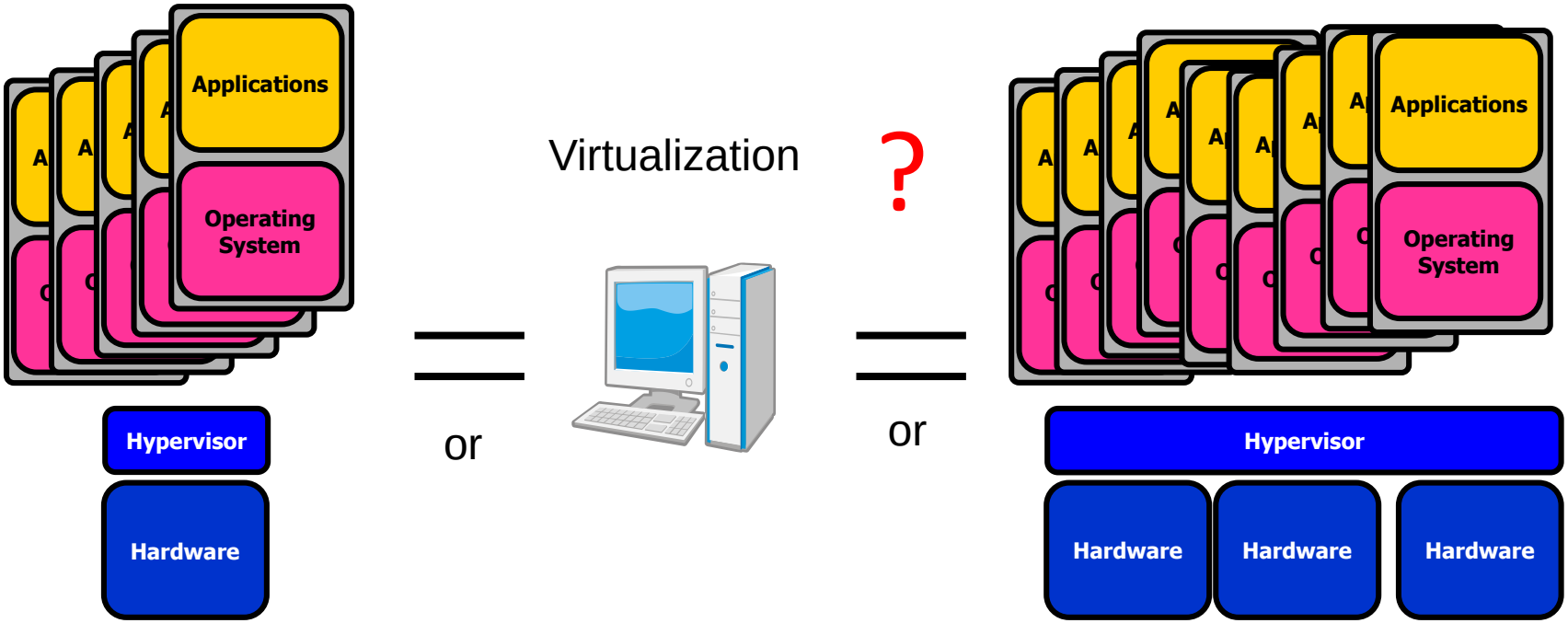
Virtualization

=



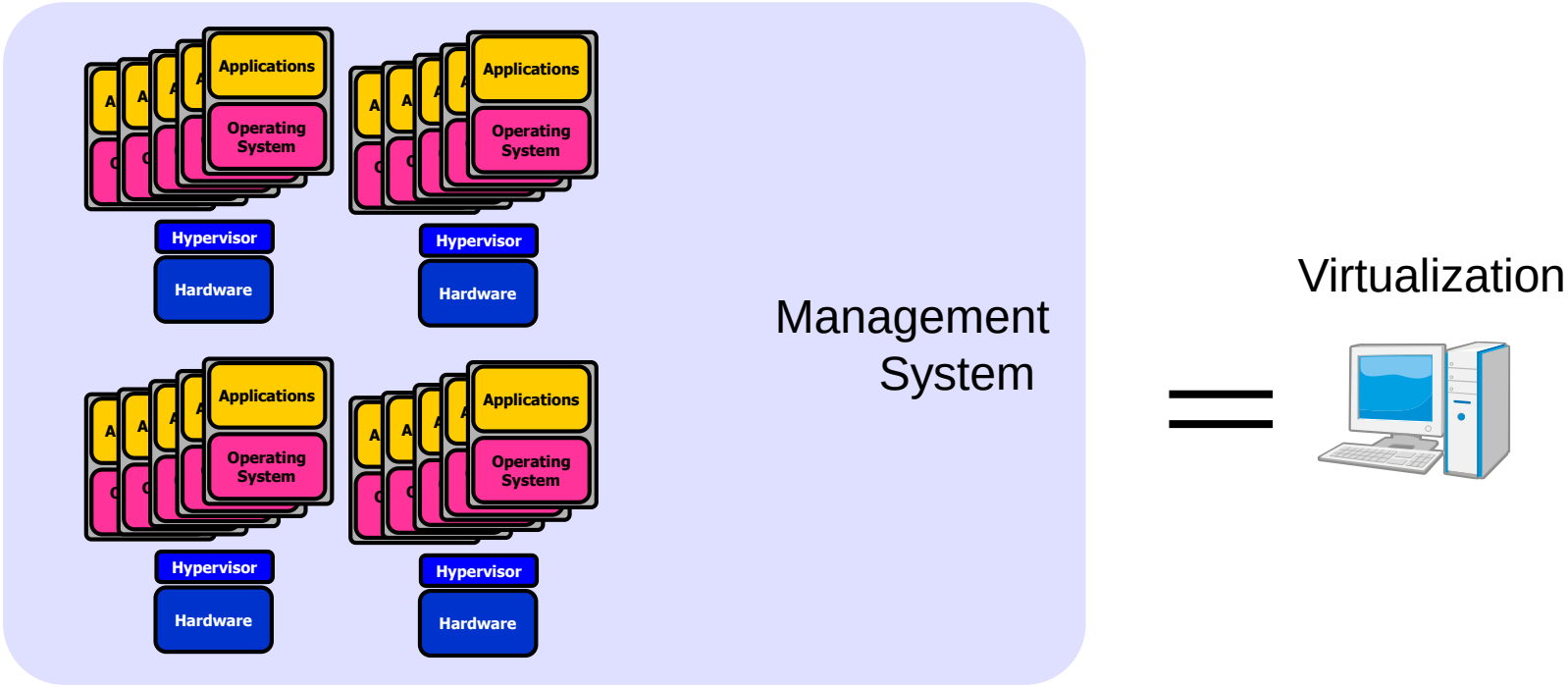
Issues in Virtualization for Cloud-Computing

- Virtualization implemented on
  - a single machine (with multi-core CPUs)
  - a cluster of machines (with multi-core CPUs)
- The state-of-the-art
  - Running a Xen or a cluster of Xens



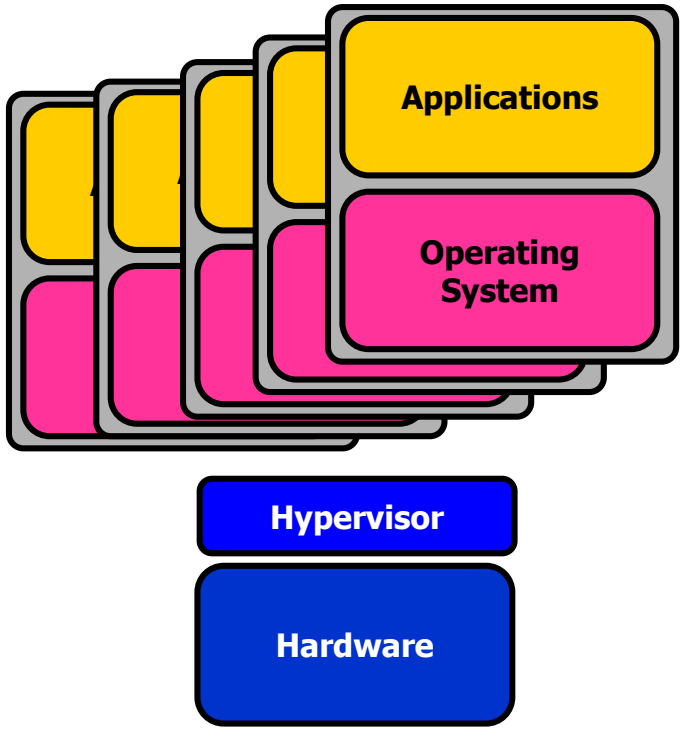
Issues in Virtualization for Cloud-Computing

- Abiquo/abicloud may provide partial solutions

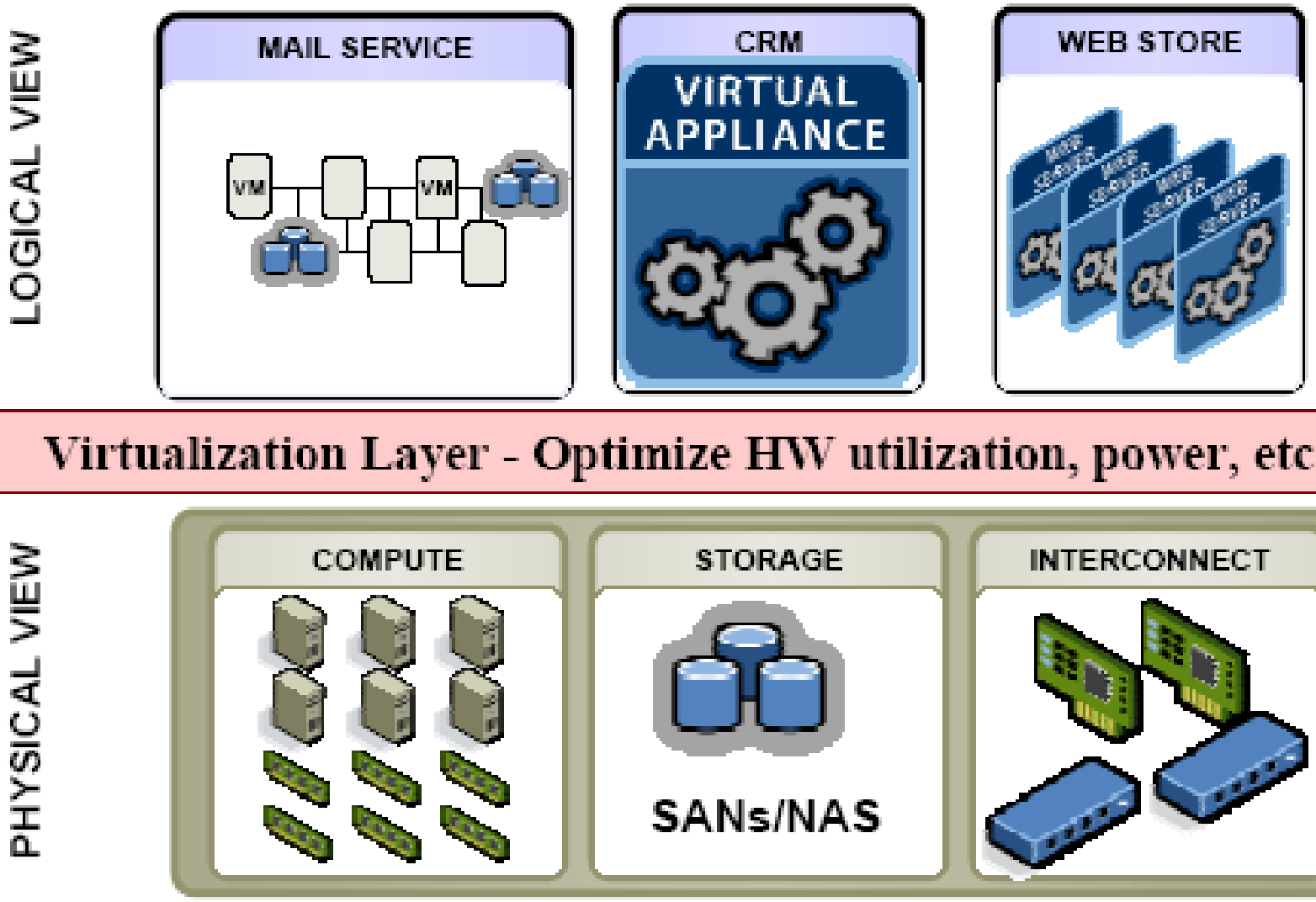


Running multiple OS and applications

- Virtualization: One physical hardware can run multiple OS and applications through a hypervisor.
- A hypervisor is the virtualization manager on a physical hardware.

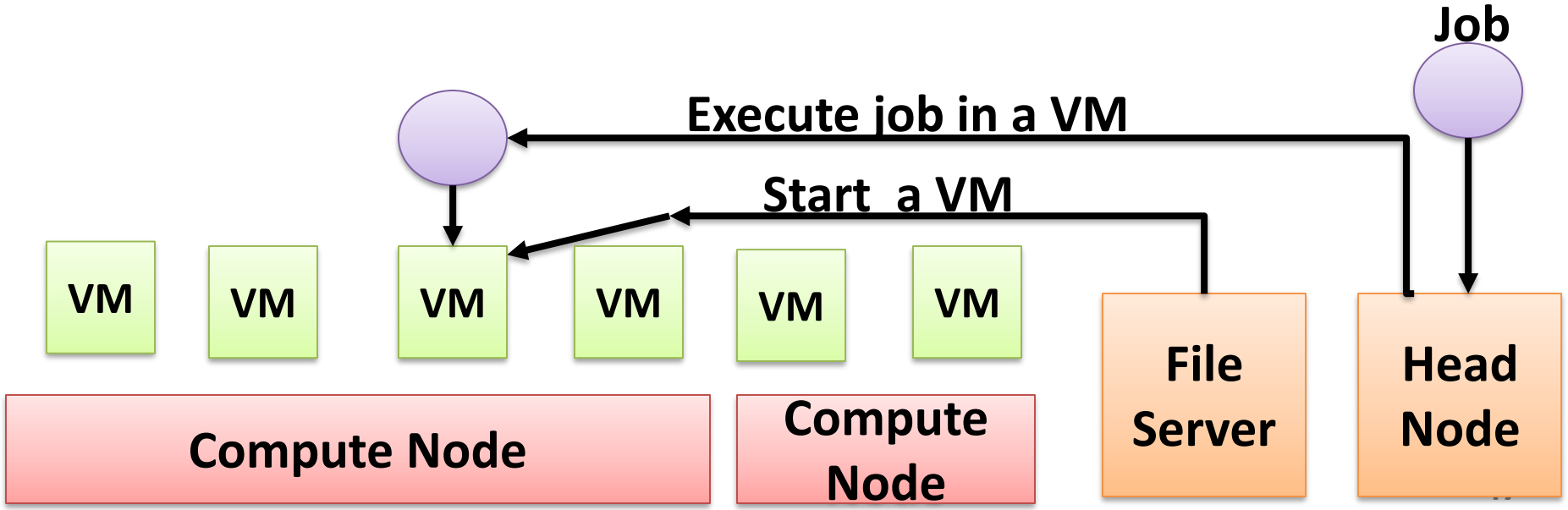


User View of Virtualization

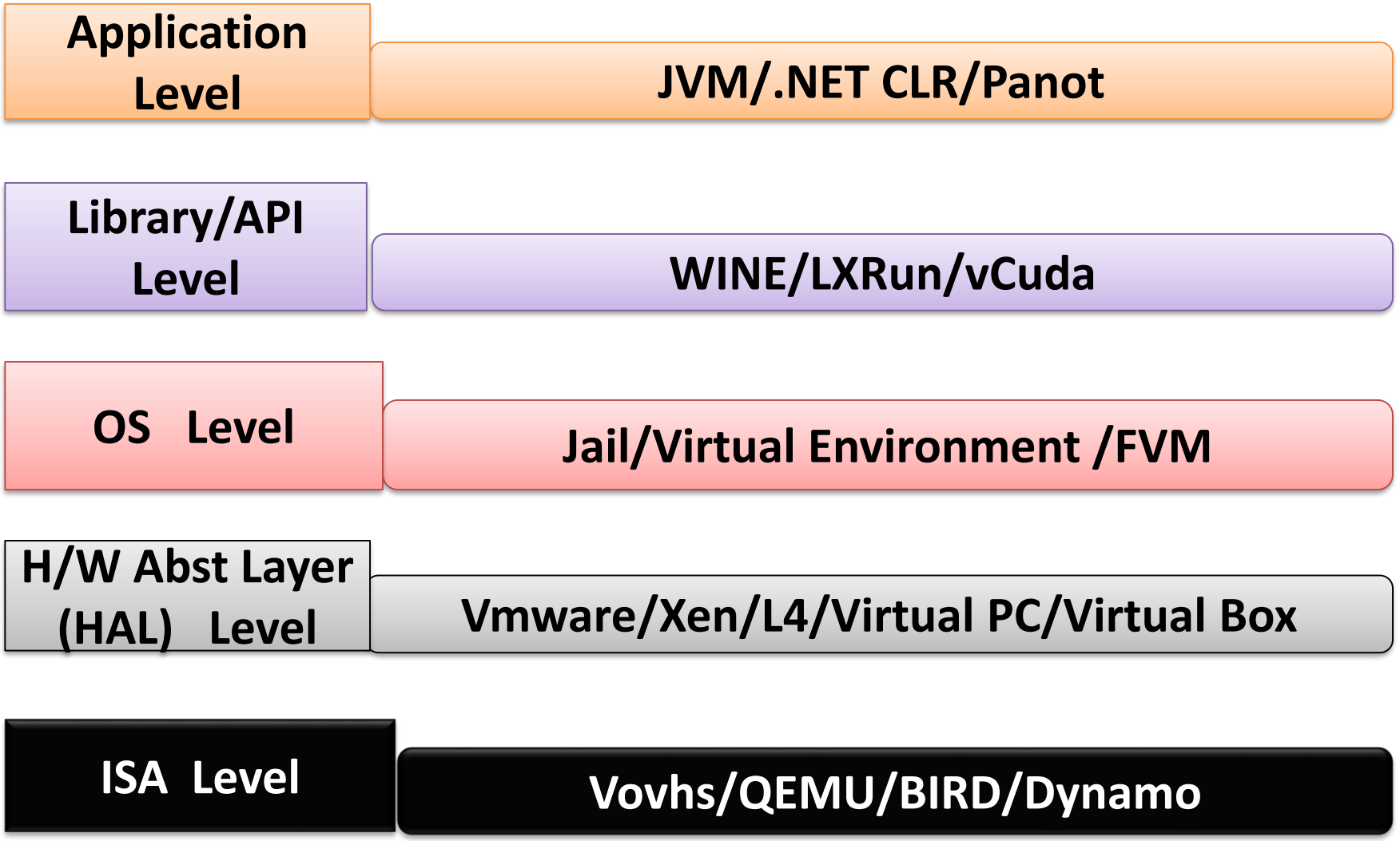


Cloud Computing

- Features of Clouds
  - Scalable, Enhanced Quality of Service (QoS)
  - Specialized and Customized, Cost Effective
  - Simplified User Interface



Virtualization in Five Abstraction Levels



Emulation vs. Virtualization

- **Emulation technique**
  - Simulate an independent environment where guest ISA and host ISA are different.
  - Example: Emulate x86 architecture on ARM platform.
- **Virtualization technique**
  - Simulate an independent environment where guest ISA and host ISA are the same.
  - Example : Virtualize x86 architecture to multiple instances.

Virtualization at ISA (Instruction Set Architecture) level

- With the help of ISA emulation
  - Example : MIPS binary code can run on an x-86 host
  - Typical systems: Bochs, Crusoe, Quemu, BIRD, Dynamo, Simic/Gems
- Advantage
  - It can run a large amount of legacy binary codes written for various processors on any given new hardware host machines
  - best application flexibility
- Shortcoming & limitation
  - One source instruction may require 10-100 of target instructions to perform its function, which is relatively slow.

Virtualization at Hardware Abstraction level

- Generates virtual hardware envts for VMs,
  - And manages the underlying hardware through virtualization.
  - Typical systems: VMware, Virtual PC, Xen, Virtual Box
- Advantage:
  - higher performance and good application isolation
- Shortcoming & limitation:
  - Very expensive to implement (complexity)

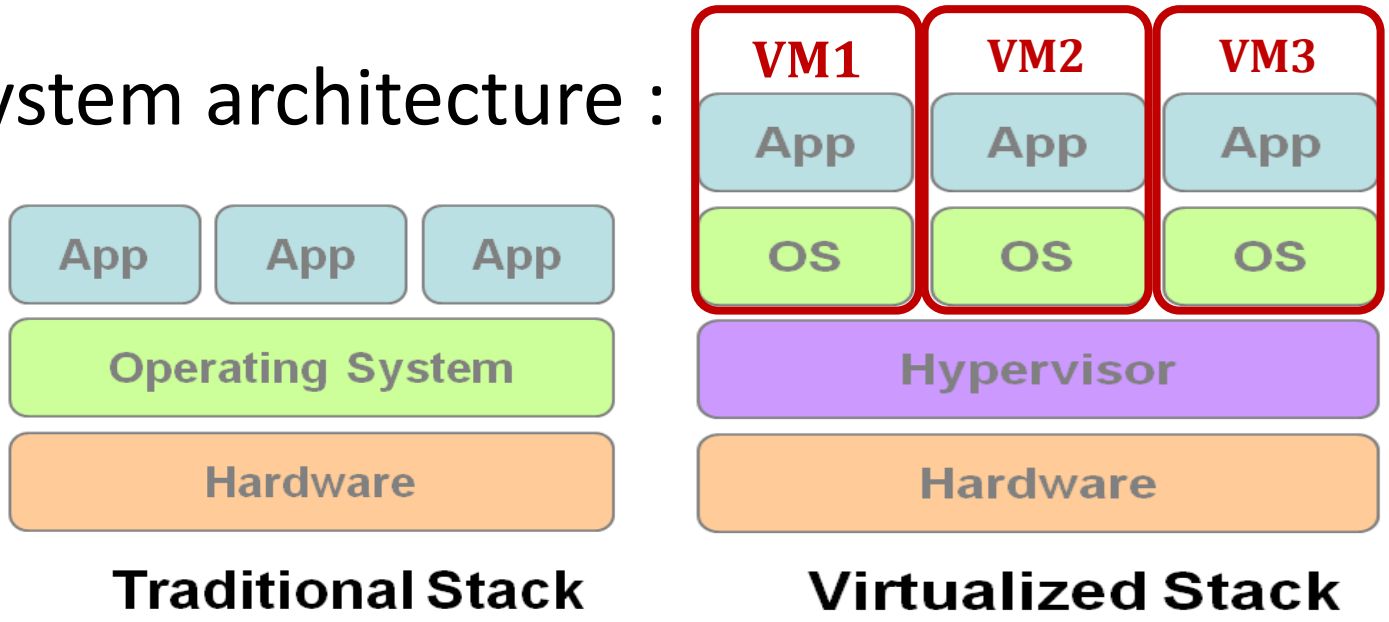
Virtualization at Operating System (OS) level

- This virtualization creates isolated containers on a single physical server and the OS-instance to utilize the hardware and software in datacenters.
  - Typical systems: Jail / Virtual Environment / FVM/
  - **Docker/Container/Kubernet**
- **Advantage**
  - Has minimal starup/shutdown cost, low resource requirement, and high scalability; synchronize VM and host state changes.
- **Shortcoming & limitation:**
  - All VMs at the operating system level must have the same kind of guest OS
  - Poor application flexibility and isolation.

Virtual Machine Monitor

- What's Virtual Machine Monitor (VMM) ?
  - **VMM** or **Hypervisor** is the software layer providing the virtualization.

- System architecture :



Virtualization Types

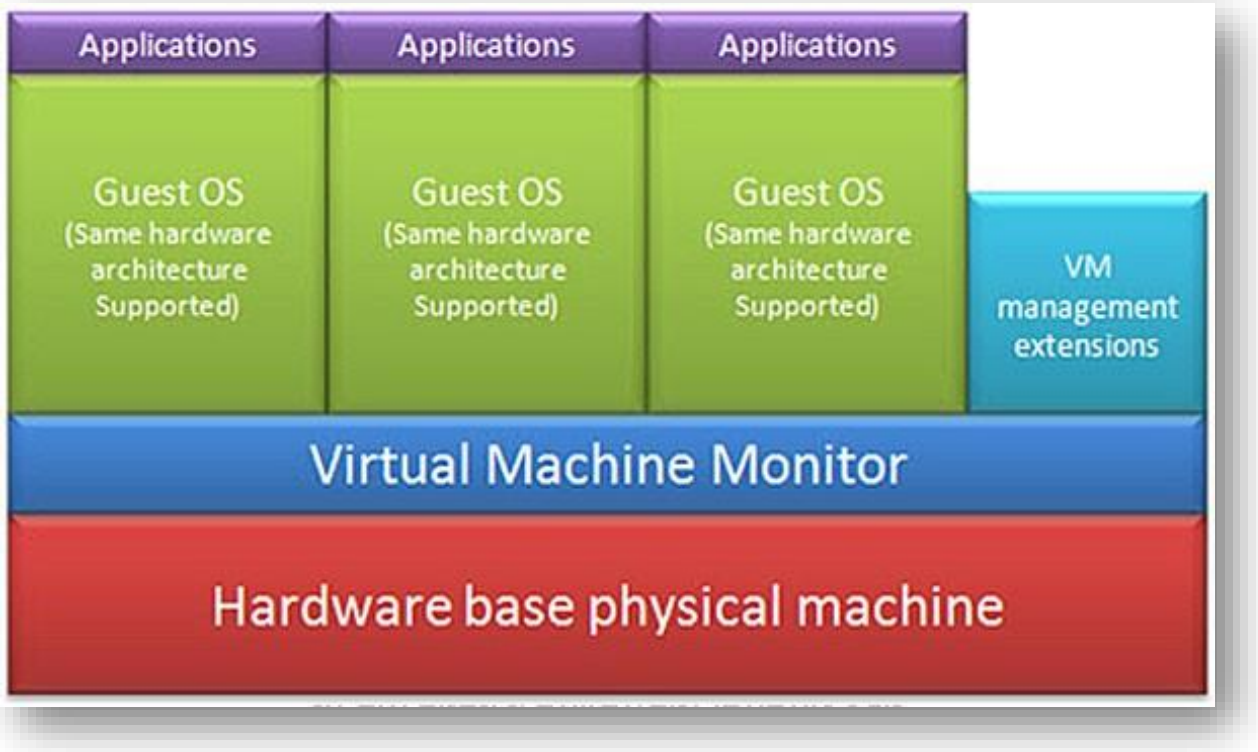
- Virtualization Types :
  - Type 1 – Bare metal
    - VMMs run directly on the host's hardware as a hardware control and guest operating system monitor.
  - Type 2 – Hosted
    - VMMs are software applications running within a conventional operating system.

Virtualization Approaches

- Virtualization Approaches :
  - Full-Virtualization
    - VMM simulates enough hardware to allow an unmodified guest OS.
  - Para-Virtualization
    - VMM does not necessarily simulate hardware, but instead offers a special API that can only be used by the modified guest OS.

Virtualization Approaches

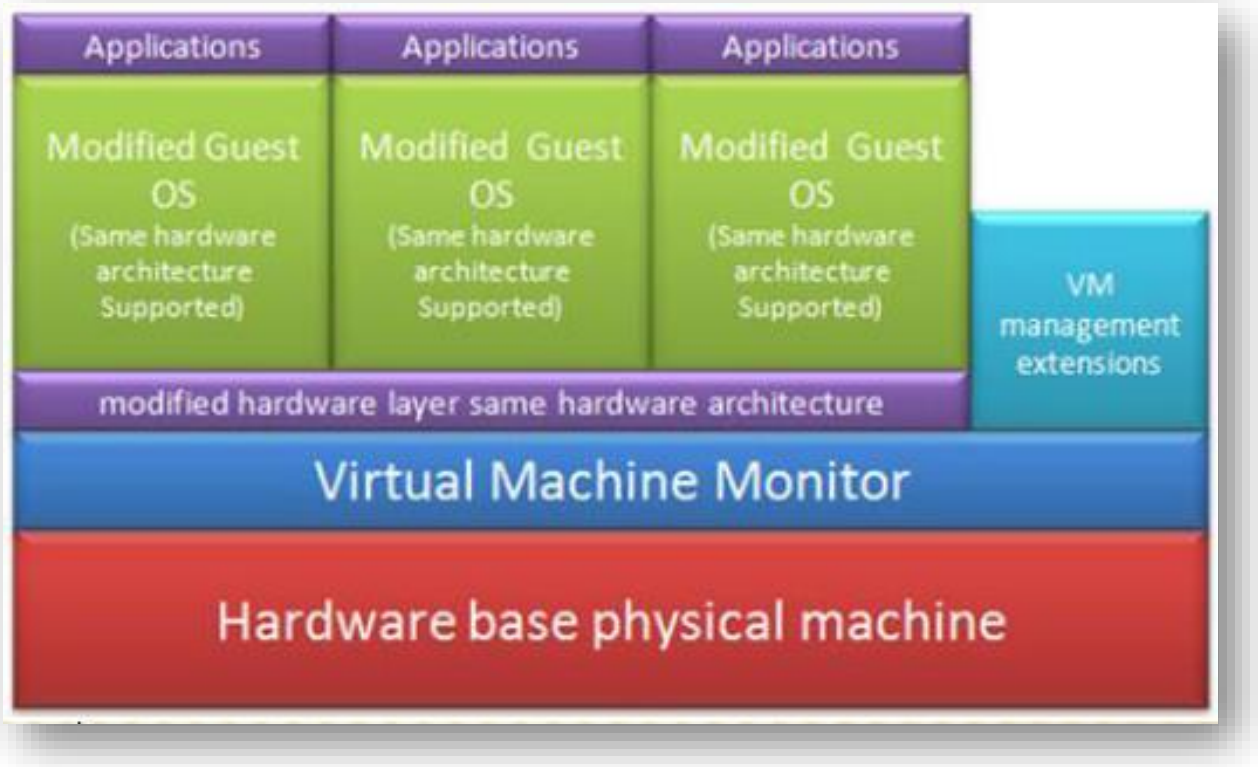
- Full-Virtualization



Pros	**Need not to modify guest OS**
Cons	**Significant performance hit**

Virtualization Approaches

- Para-Virtualization

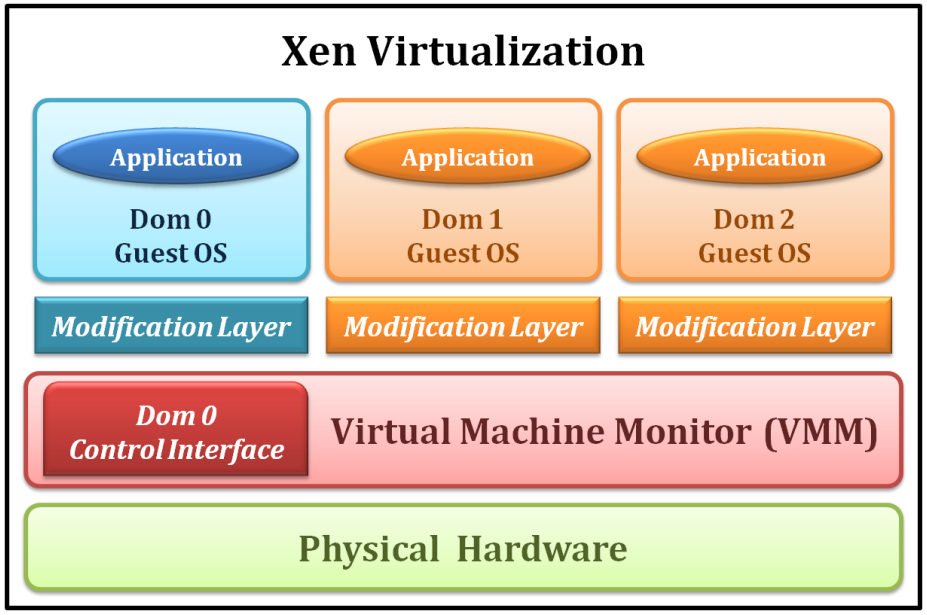


Pros	**Light weight and high performance**
Cons	**Require modification of guest OS**

Examples

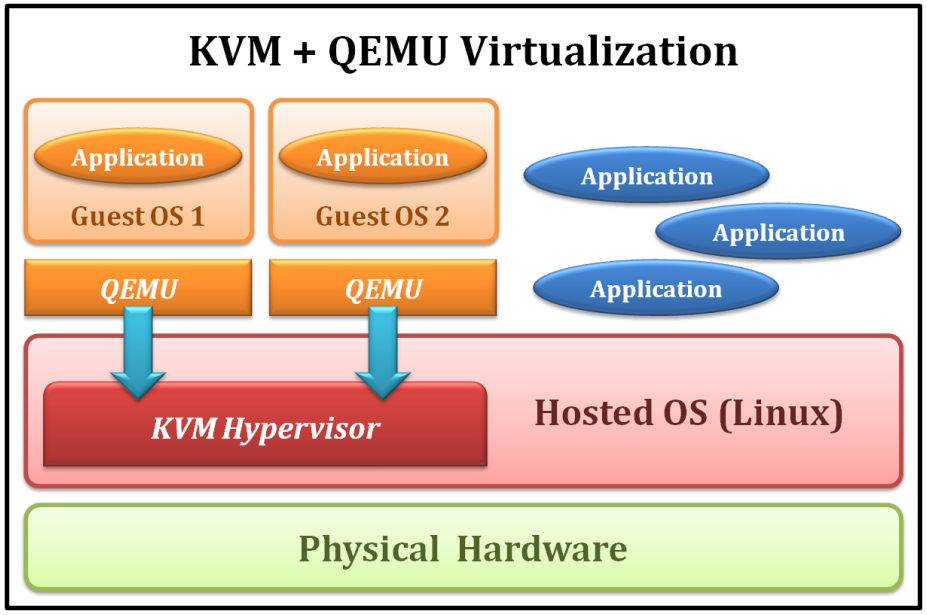
Xen

- Type 1 Virtualization
- Para-Virtualization



KVM

- Type 2 Virtualization
- Full-Virtualization



What is the purpose and benefits?

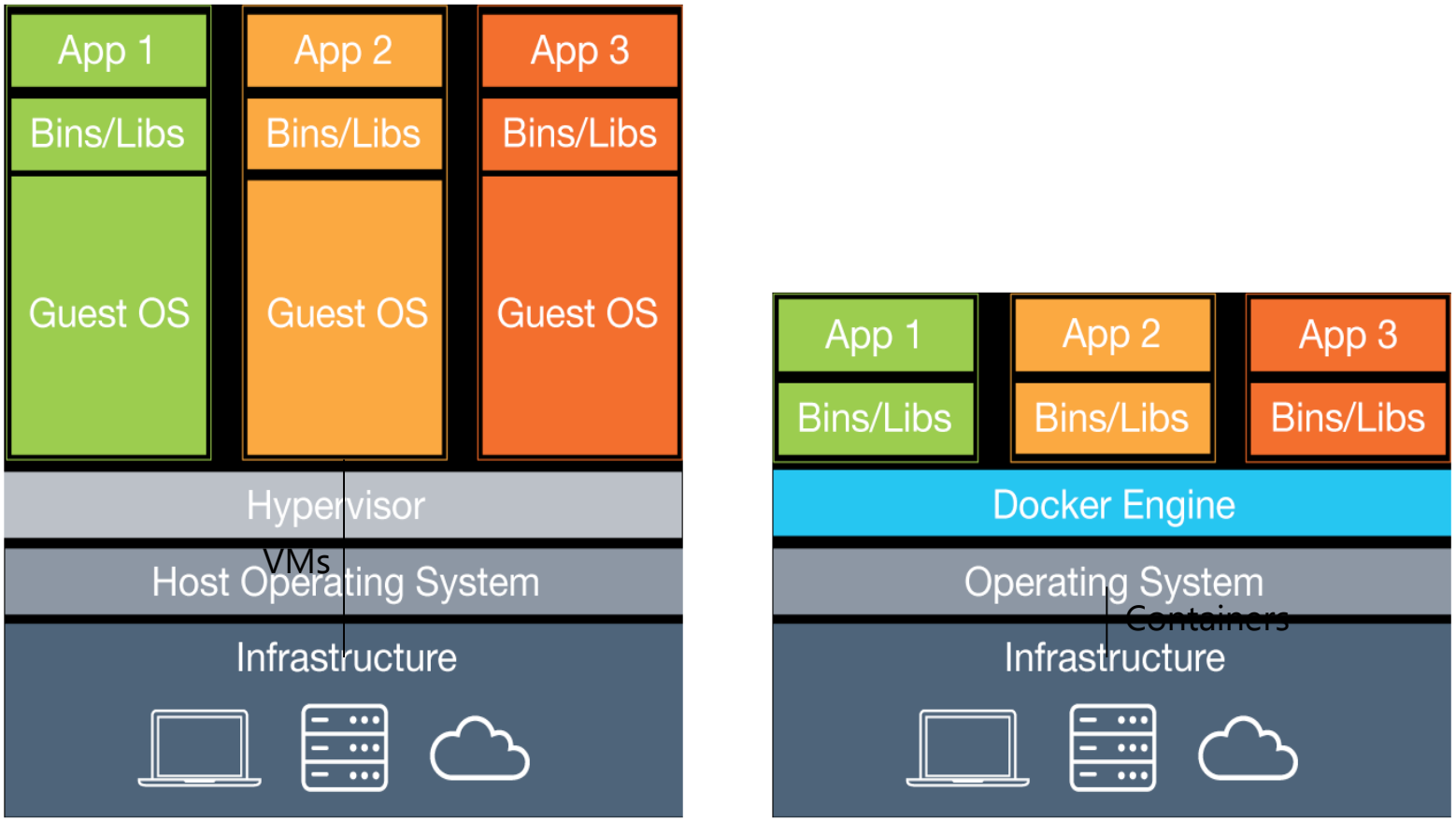
- Cloud computing enables companies and applications, which are system infrastructure dependent, to be infrastructure-less.
- By using the Cloud infrastructure on “pay as used and on demand”, all of us can save in capital and operational investment!
- Clients can:
  - Put their data on the platform instead of on their own desktop PCs and/or on their own servers.
  - They can put their applications on the cloud and use the servers within the cloud to do processing and data manipulations etc.

Containers : Docker

- What is a container?
  - a container consists of an application and all its dependencies which can be run in an isolated way
  - make uses of kernel features (cgroups, namespaces, ...)
- Benefits include
  - independence from host OS & libraries
  - can be run anywhere, regardless of kernel version or host Linux distribution

Containers vs Virtualization

- VM is complete OS
- Container is App with required Lib and tools



Singularity : isolation

- Project from Berkeley Labs (http://singularity.lbl.gov/)
- Designed to allow non-privileged users on HPC systems to provide their own OS
  - isolates filesystem & processes
  - no daemon, no UID switching
- Being pushed by Traceability & Isolation Working Group
  - seen as a (future) alternative to glexec
  - provides isolation but not traceability
    - payload cannot attack pilot or other payloads on same host

Kubernetes

- Open source container cluster manager, originally developed at Google
- Can be installed on-prem (bare metal or on a cloud), also available on public clouds
  - "click a button" on Google & Azure
  - straightforward to install on AWS, ...

Why Kubernetes?

- Standard open-source software (not HEP-specific)
- Using it as a means of abstracting differences between on-prem (bare metal) resources & different public clouds
- Eliminate vendor lock-in by avoiding any cloud or vendor specific APIs
  - No need to write software to talk to different cloud APIs, just use the Kubernetes API only
- Also has federation functionality making it easy
  - To deploy workloads across multiple clusters (new)

Container Verdicts

- Use of containers beneficial for both VOs and sites
  - jobs no longer depend on OS version or software installed on worker nodes
  - easier to provide a consistent environment at multiple sites
- Singularity seems to be a simple way for sites to run jobs in containers
- Container cluster managers (**Mesos**)
  - **C**an be used to provide an efficient platform for long-running services & multiple compute activities
  - Kubernetes can be used to provide portability between local resources & multiple public clouds

Pro and Cons
of
Cloud Computing

Opportunities of Use of Cloud

- It enables services to be used without any understanding of their infrastructure.
- Cloud computing works using economies of scale:
  - It potentially lowers the outlay expense for start up companies, as they would no longer need to buy their own software or servers.
  - Cost would be by on-demand pricing.
  - Vendors and Service providers claim costs by establishing an ongoing revenue stream.
- Data and services are stored remotely but accessible from “anywhere”.

Advantages of Cloud Computing : 1

- **Improved performance**
- With few large programs hogging your computer's memory, you will see better performance from your PC.
- Computers in a cloud computing system boot and run faster
  - Because they have fewer programs and processes loaded into memory…

Advantages of Cloud Computing :2

- **Reduced software costs**
- Instead of purchasing expensive software applications, you can get most of what you need for **free-ish!**
  - Most cloud computing applications today, such as the Google Docs suite.
- Better than paying for similar commercial software
  - which alone may be justification for switching to cloud applications.

Advantages of Cloud Computing :3

- **Universal document access**
- That is not a problem with cloud computing, because you do not take your documents with you.
- Instead, they stay in the cloud, and you can access them whenever you have a computer and an Internet connection
- Documents are instantly available from wherever you are

Advantages of Cloud Computing :4

- **Latest version availability**
- When you edit a document at home, that edited version is what you see when you access the document at work.
- The cloud always hosts the latest version of your documents
  - as long as you are connected, you are not in danger of having an outdated version

Advantages of Cloud Computing: 5

- **Unlimited storage capacity**
- Cloud computing offers virtually limitless storage.
- Your computer's current 1 Tbyte hard drive is small compared to the hundreds of Pbytes available in the cloud.

Advantages of Cloud Computing: 6

- **Increased data reliability**
- Unlike desktop computing, in which
  - if a hard disk crashes and destroy all your valuable data,
  - A computer crashing in the cloud should not affect the storage of your data.
  - if your personal computer crashes, all your data is still out there in the cloud, still accessible
- In a world where
  - Few individual desktop PC users back up their data on a regular basis,
  - Cloud computing is a data-safe computing platform!

Advantages of Cloud Computing: 7

- **Lower computer costs**
- To run cloud computing's web-based applications
  - You do not need a high-powered and high-priced computer
- Since applications run in the cloud, not on the PC
  - Your PC does not need high processing power or hard disk space demanded by traditional desktop software.
- When you are using web-based applications
  - Your PC can be less expensive, with a smaller hard disk, less memory, more energy efficient processor
  - PC does not even need a CD/DVD drive,
  - No software programs have to be loaded
  - No document files need to be saved.

Advantages of Cloud Computing : 8

- **Instant software updates**
- You are no longer faced with choosing between obsolete software and high upgrade costs.
- When the application is web-based
  - Updates happen automatically
  - Available the next time you log into the cloud.
- When you access a web-based application
  - You get the latest version
  - Without needing to pay for or download an upgrade.

Advantages of Cloud Computing : 9

- **Improved document format compatibility.**
- You do not have to worry
  - About the documents you create on your machine being compatible with other users' applications or OSes
- There are potentially no format incompatibilities
  - When everyone is sharing documents and applications in the cloud.

Advantages of Cloud Computing :10

- **Easier group collaboration**
- Sharing documents leads directly to better collaboration.
- Many users do this as it is an important advantages of cloud computing
  - multiple users can collaborate easily on documents and projects

Advantages of Cloud Computing : 11

- **Device independence**
- You are no longer tethered to a single computer or network.
- Changes to computers, applications and documents follow you through the cloud.
- Move to a portable device, and your applications and documents are still available.

Challenges in Using the Cloud

- In parallel there has been backlash against cloud computing
- Use of cloud computing means
  - Dependence on others and that could possibly limit flexibility and innovation
- The others are likely become the
  - Bigger Internet companies like Google and IBM, who may monopolise the market.
- Some argue that this use of supercomputers is
  - A return to the time of mainframe computing that the PC was a reaction against.

Disadvantages of Cloud Computing : 1

- **Security could prove to be a big issue**
- It is still unclear how safe out-sourced data is
- when using these services ownership of data is not always clear.

Disadvantages of Cloud Computing : 2

- **Issues relating to policy and access**
- If your data is stored abroad whose policy do you adhere to?
- What happens if the remote server goes down?
- How will you then access files?
- There have been cases of users being locked out of accounts and losing access to data.

Disadvantages of Cloud Computing : 3

- **Requires a constant Internet connection**
- Cloud computing is impossible if you cannot connect to the Internet.
- Since you use the Internet to connect to both your applications and documents
  - if you do not have an Internet connection you cannot access anything
  - Even your own documents.
- A dead Internet connection means no work
  - In areas where Internet connections are few or inherently unreliable
  - this could be a deal-breaker.

Disadvantages of Cloud Computing : 4

- **Does not work well with low-speed connections**
- Similarly, a low-speed Internet connection, such as that found with dial-up services, makes cloud computing painful at best and often impossible.
- Web-based applications require a lot of bandwidth to download, as do large documents.

Disadvantages of Cloud Computing : 5

- **Features might be limited**
- This situation is bound to change, but today many web-based applications simply are not as full-featured as their desktop-based applications.
- For example, you can do a lot more with Microsoft PowerPoint than with Google Presentation's web-based offering

Disadvantages of Cloud Computing : 6

- **Can be slow Even with a fast connection**
  - Web-based applications can sometimes be slower than Accessing a similar software program on your desktop PC.
- Everything about the program, from the interface to the current document
  - Has to be sent back and forth from your computer to the computers in the cloud.
- If the cloud servers happen to be backed up at that moment, or
- if the Internet is having a slow day, you would not get the instantaneous access

Disadvantages of Cloud Computing : 7

- **Stored data might not be secure**
- With cloud computing, all your data is stored on the cloud.
  - The questions is How secure is the cloud?
- Can un-authorised users gain access to your confidential data?

Disadvantages of Cloud Computing : 8

- **Stored data can be lost**
- Theoretically, data stored in the cloud is safe, replicated across multiple machines.
- But on the off chance that your data goes missing, you have no physical or local backup.
  - Put simply, relying on the cloud puts you at risk if the cloud lets you down.

Disadvantages of Cloud Computing : 9

- **HPC Systems and General concern**
  - Not clear that you can run compute-intensive HPC applications that use MPI/OpenMP!
  - Scheduling is important with this type of application
    - as you want all the VM to be co-located to minimize communication latency!
- General Concerns:
  - Each cloud systems uses different protocols and different APIs
    - may not be possible to run applications between cloud based systems
  - Amazon has created its own DB system (not SQL 92), and workflow system (many popular workflow systems out there)
    - so your normal applications will have to be adapted to execute on these platforms.