

Cloud Computing

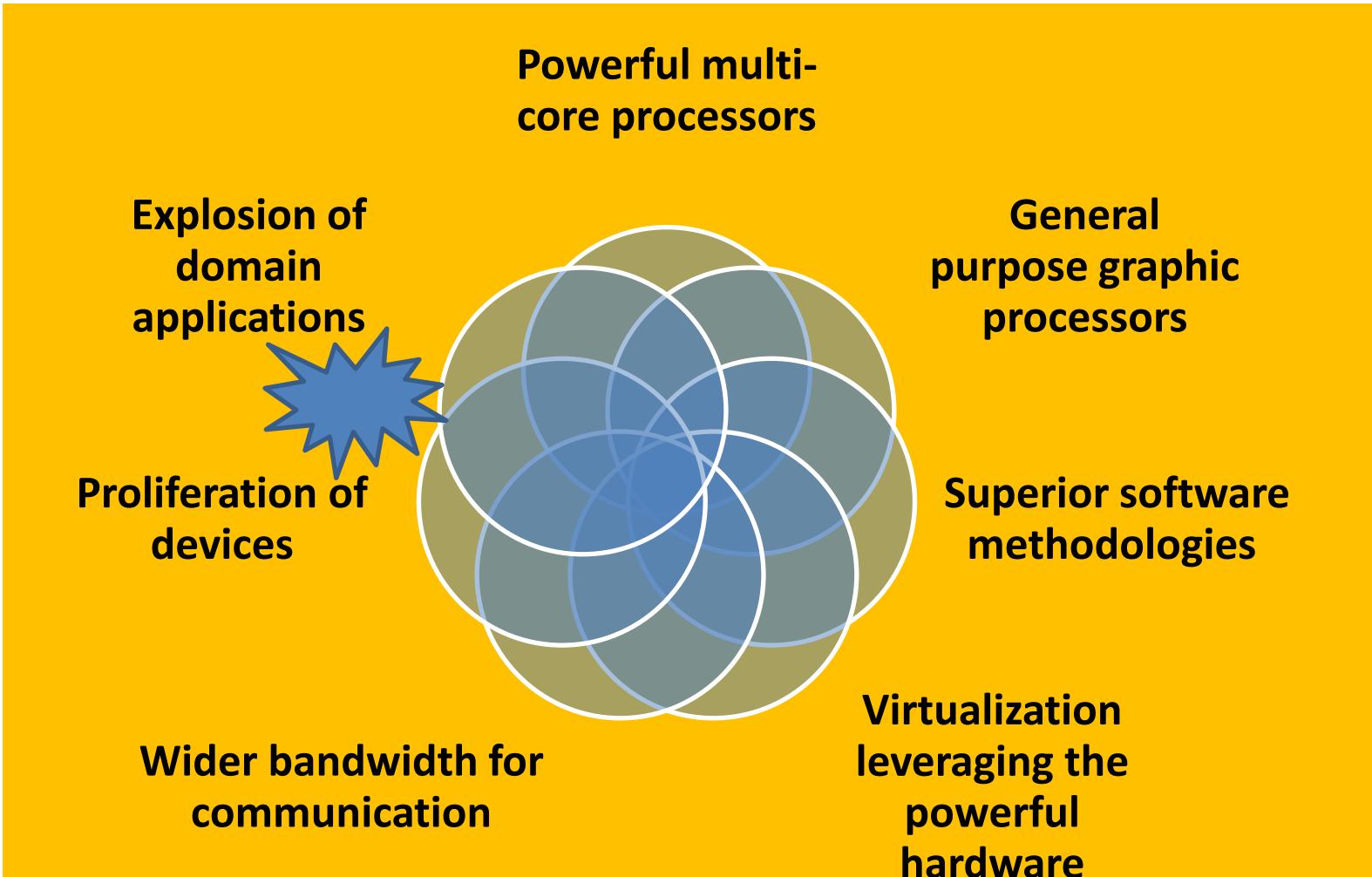
A Technical Overview

Chaiyaporn Khemapatapan, Ph.D

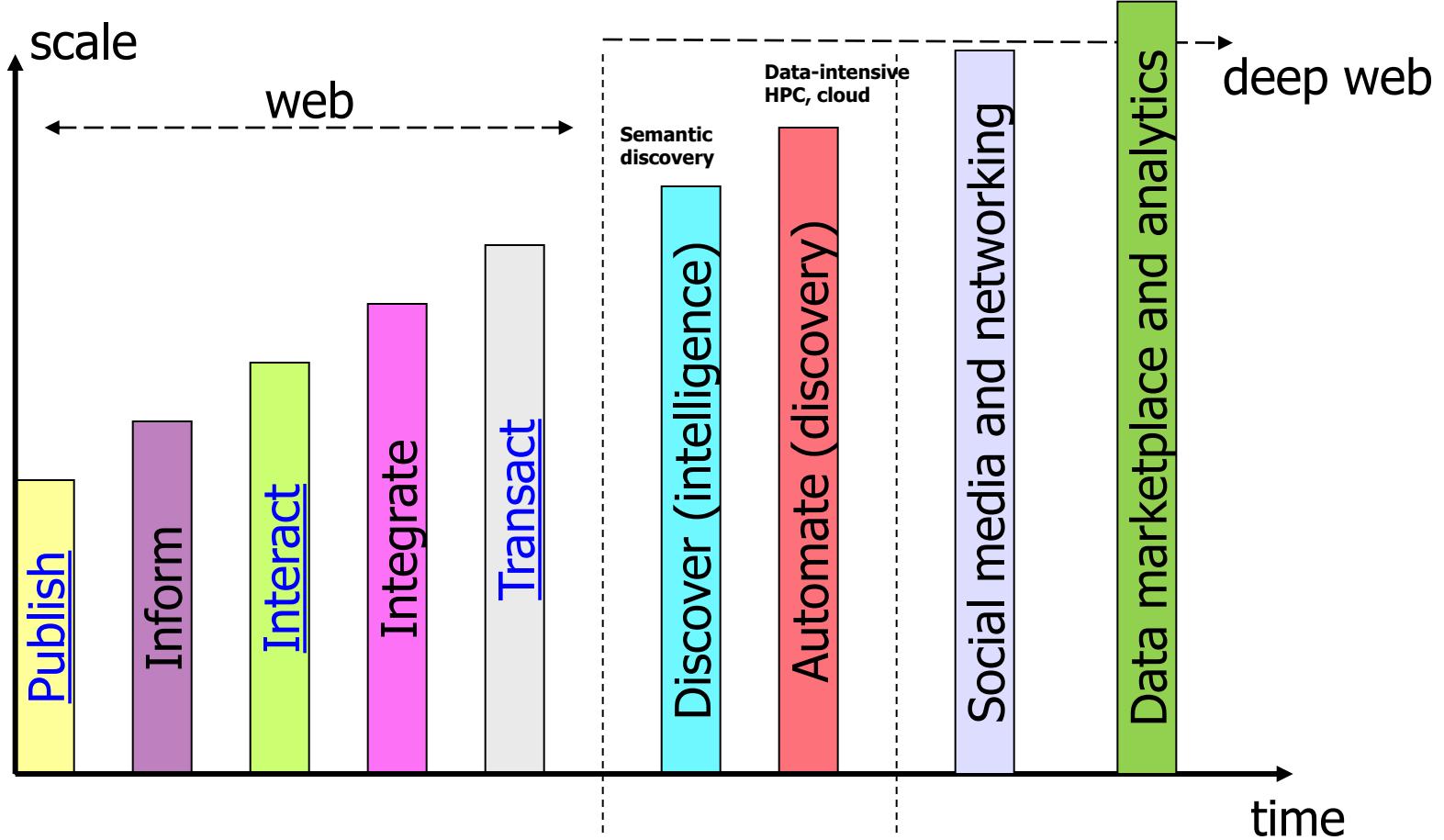
Learning Objectives

- To clearly know what is **Cloud Computing**
- To understand a brief technical concept of **Cloud Computing**
- To aware the importance of **Cloud Computing**
- To learn the architecture of **Cloud Computing**
- To learn how migration does
- To learn On-Premise Cloud Technology
- To know public cloud services

Introduction: A Golden Era in Computing

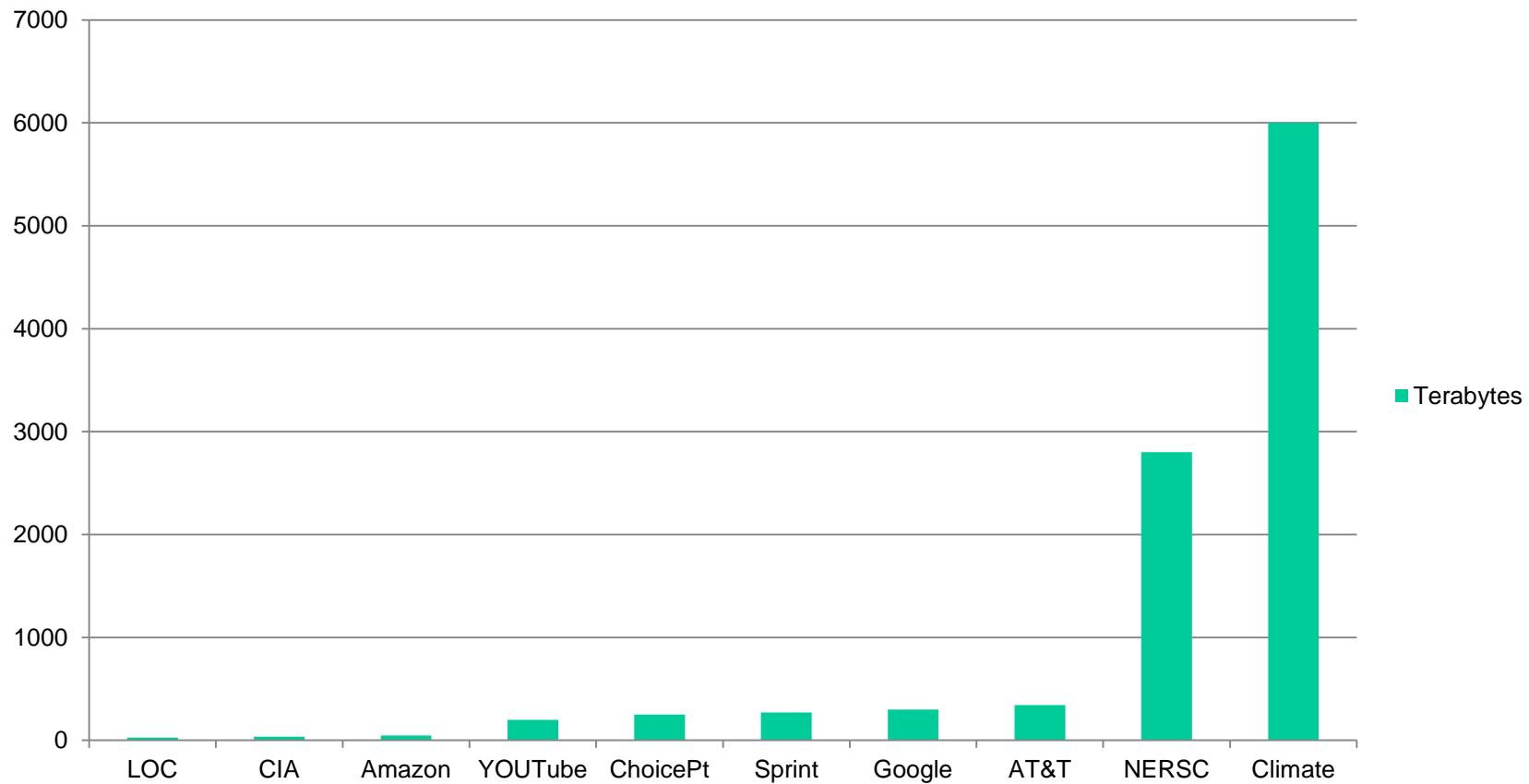


Evolution of Internet Computing



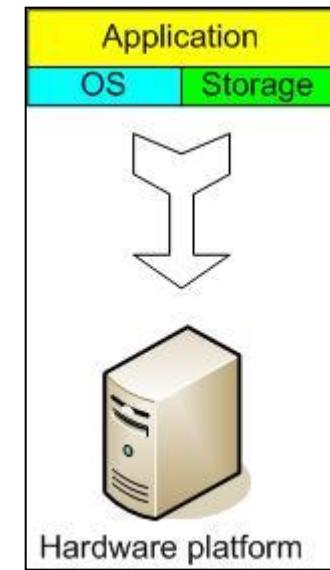
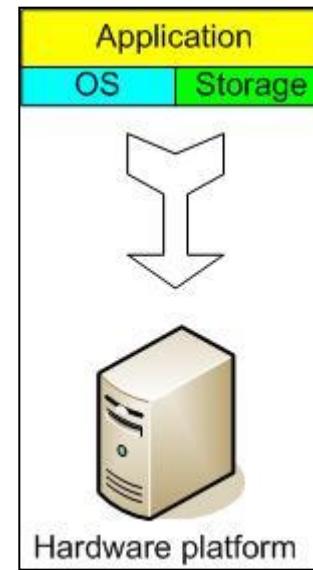
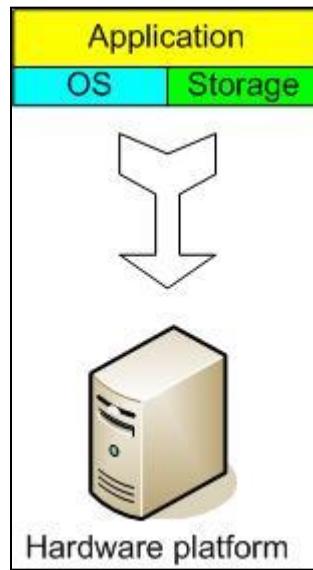
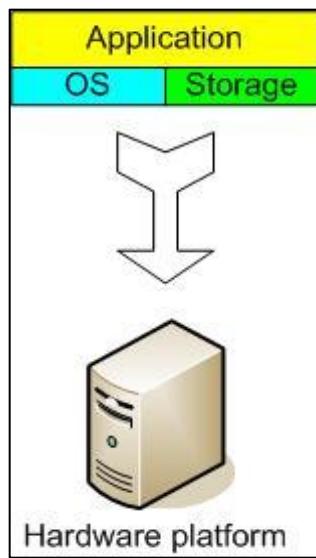
Top Ten Largest Databases

Top ten largest databases (2007)



Ref: <http://www.focus.com/fyi/operations/10-largest-databases-in-the-world/>

The Traditional Server Concept



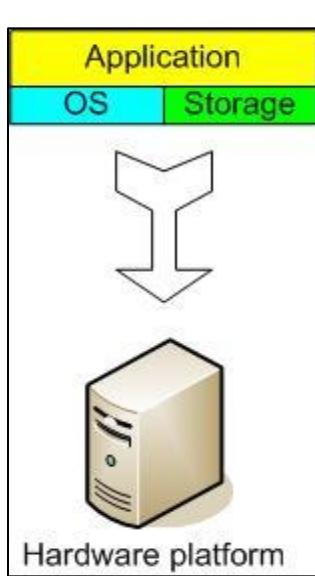
Web
Server
Windows
IIS

App Server
Linux
Glassfish

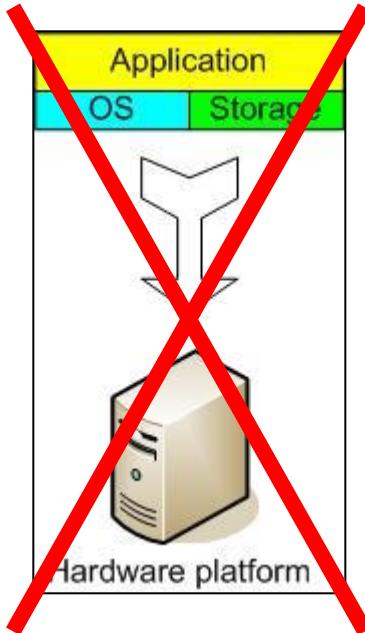
DB Server
Linux
MySQL

EMail
Windows
Exchange

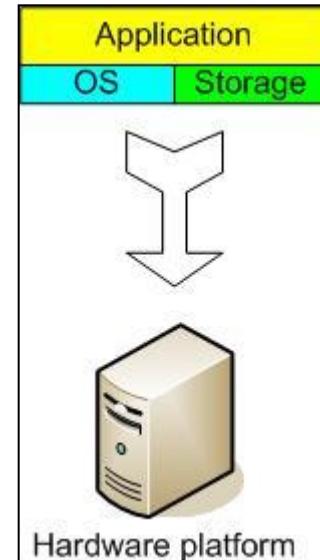
And if something goes wrong ...



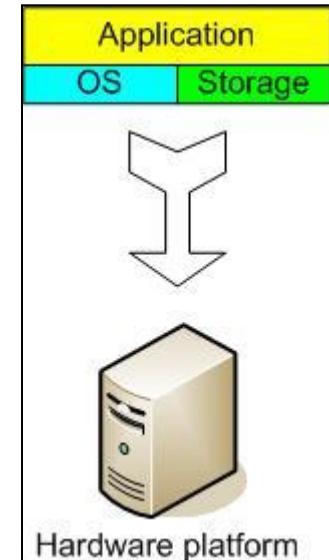
Web Server
Windows
IIS



App Server
DOWN!



DB Server
Linux
MySQL



EMail
Windows
Exchange

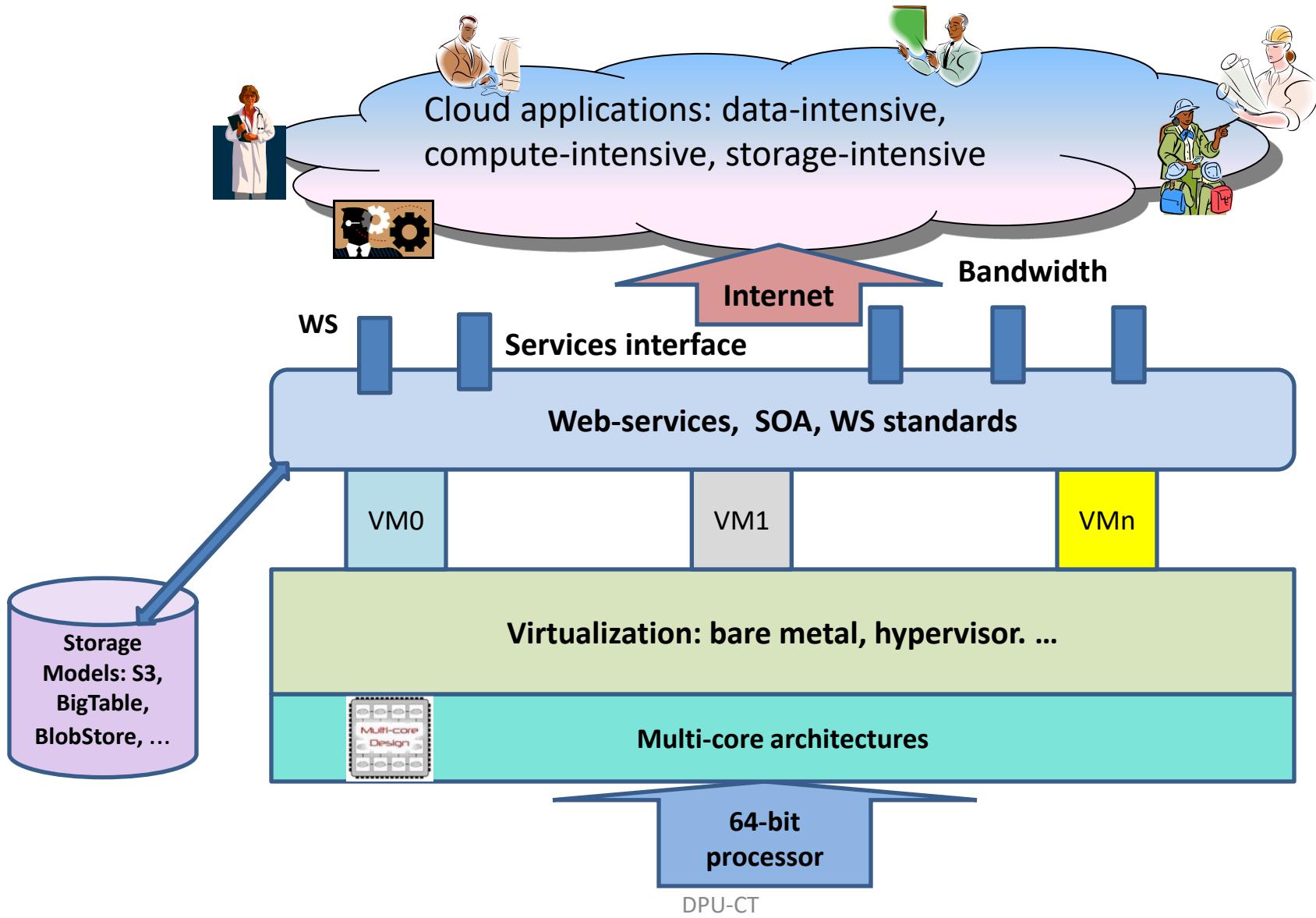
The Traditional Server Concept

- Pros
 - Easy to conceptualize
 - Fairly easy to deploy
 - Easy to backup
 - Virtually any application/service can be run from this type of setup
- Cons
 - Expensive to acquire and maintain hardware
 - Not very scalable
 - Difficult to replicate
 - Redundancy is difficult to implement
 - Vulnerable to hardware outages
 - In many cases, processor is under-utilized

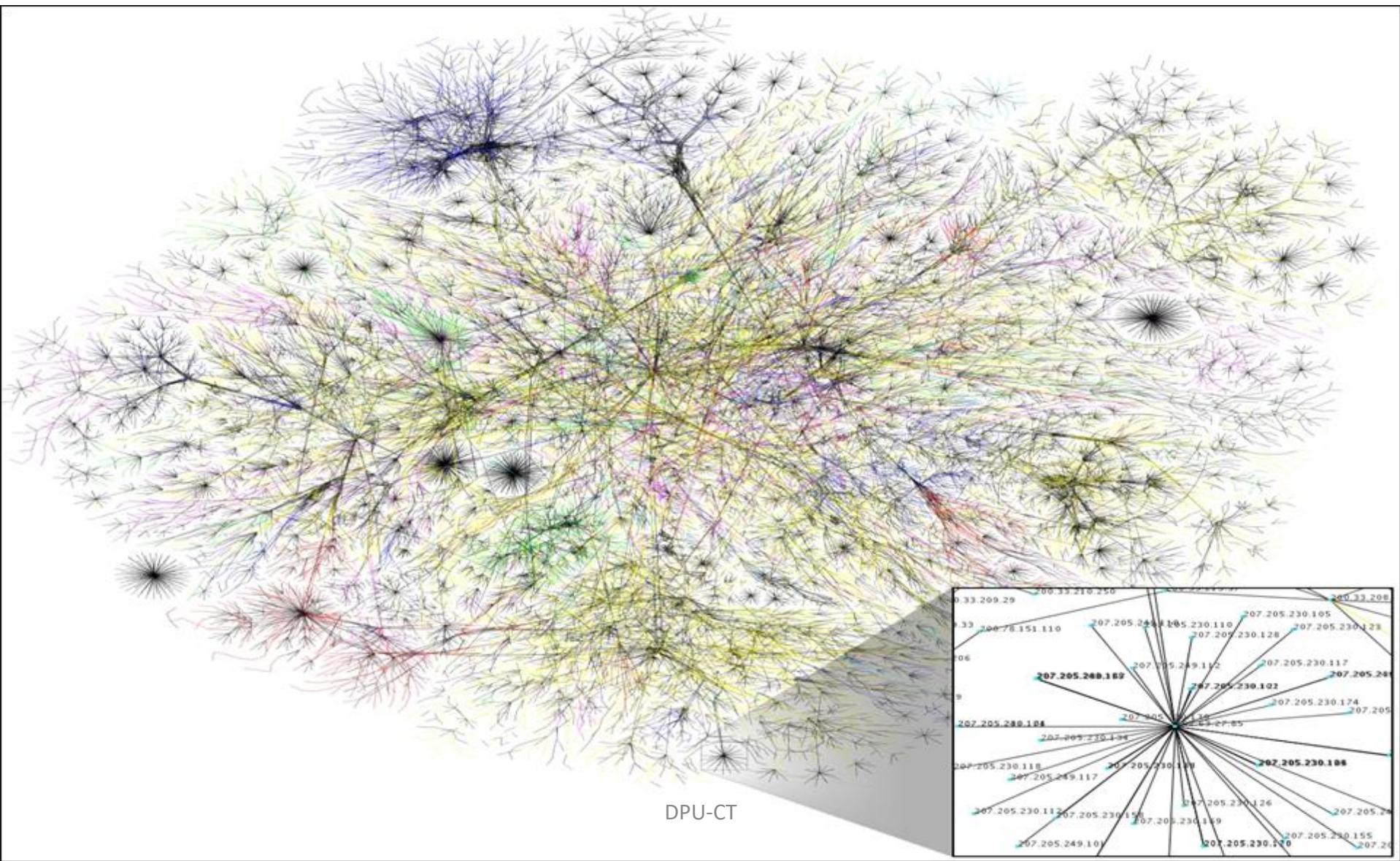
Enter the cloud

- **Cloud computing** is Internet-based computing, whereby shared resources, software and information are provided to computers and other devices on-demand, like the electricity grid.
- The cloud computing is a culmination of numerous attempts at large scale computing with seamless access to virtually limitless resources.
 - on-demand computing, utility computing, ubiquitous computing, autonomic computing, platform computing, edge computing, elastic computing, **grid computing**, ...

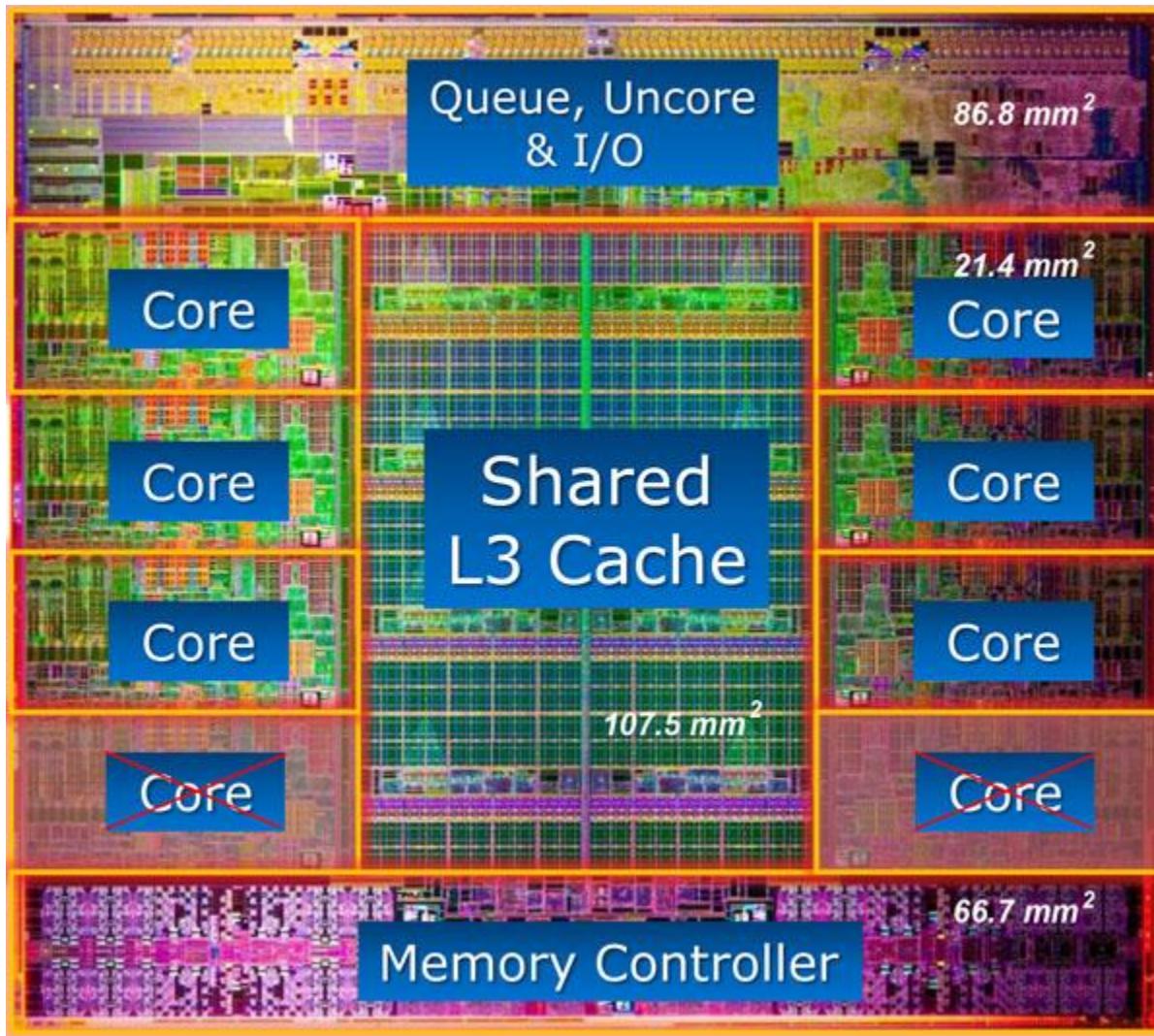
Enabling Technologies



Internet

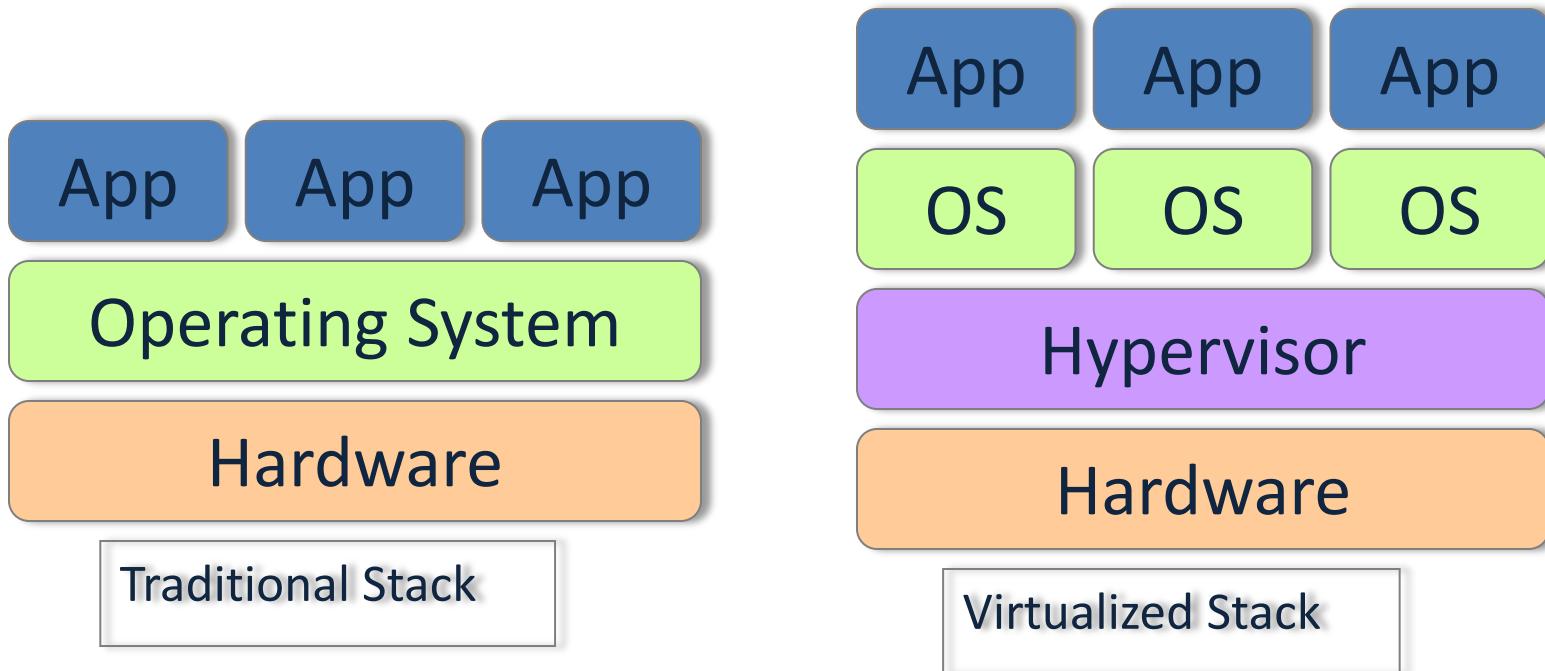


Key Technology: High Performance Computing



DPU-CT

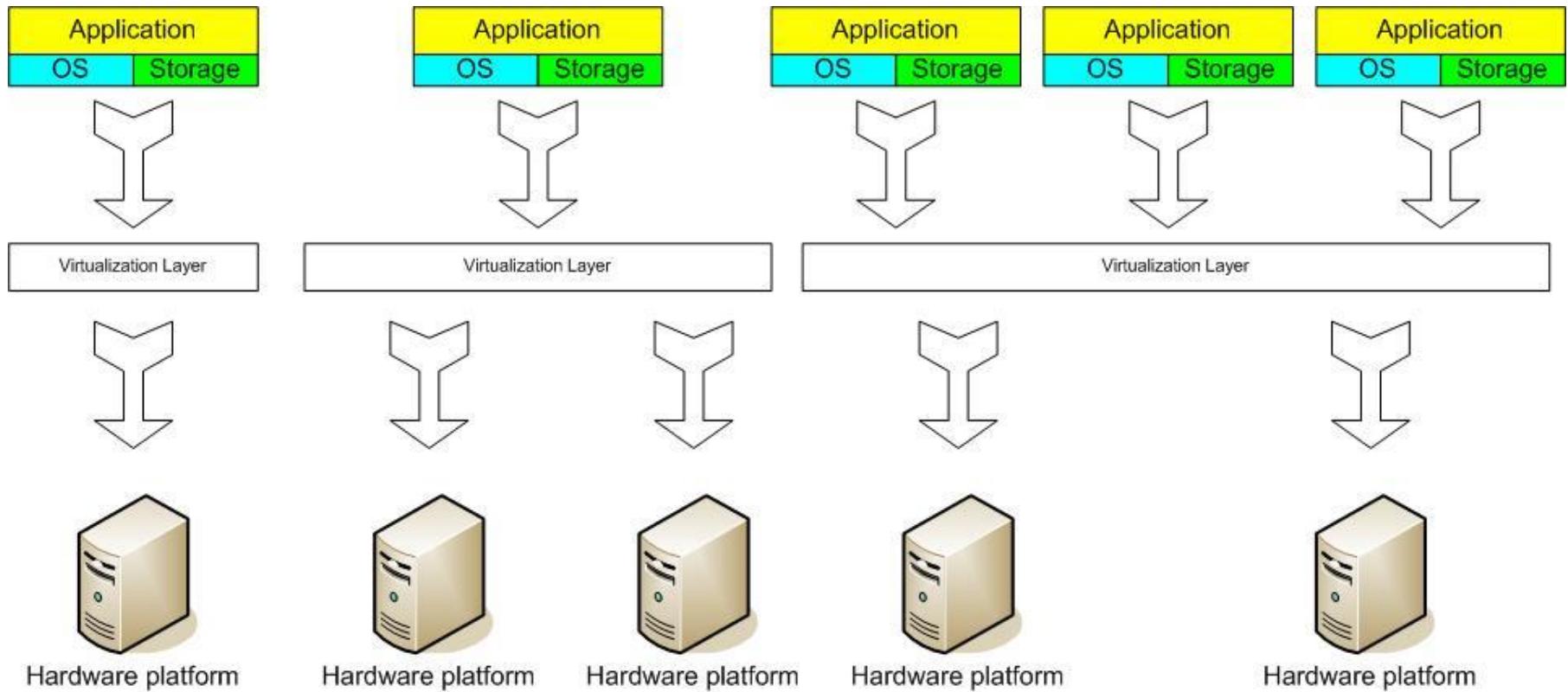
Key Technology: Virtualization



The Virtual Server Concept

- Virtual servers seek to encapsulate the server software away from the hardware
 - This includes the OS, the applications, and the storage for that server.
- Servers end up as mere files stored on a physical box, or in enterprise storage.
- One host typically house many virtual servers (**virtual machines or VMs**).
- A virtual server can be serviced by one or more hosts e.g. storage, services, etc

The Virtual Server Concept



Hypervisor layer between *Guest OS* and hardware

The Virtual Server Concept

- Pros
 - Resource pooling
 - Highly redundant
 - Highly available
 - Rapidly deploy new servers
 - Easy to deploy
 - Reconfigurable while services are running
 - Optimizes physical resources by doing more with less
- Cons
 - Slightly harder to conceptualize
 - Slightly more costly (must buy hardware, OS, Apps, and now the abstraction layer)

Cloud Computing?

- The cloud is Internet-based computing, whereby shared resources, software, and information are provided to computers and other devices on demand – pay per use.
- Cost-effective means of virtualising and making use of resources more effectively
 - Low start-up costs – pay for use helps to kick-start companies
 - Scaling is proportional to demand (revenue) so it's a good business model
- Vast range of Cloud Computing applications
 - Virtual private servers, Web hosting, data servers, fail-over services, etc

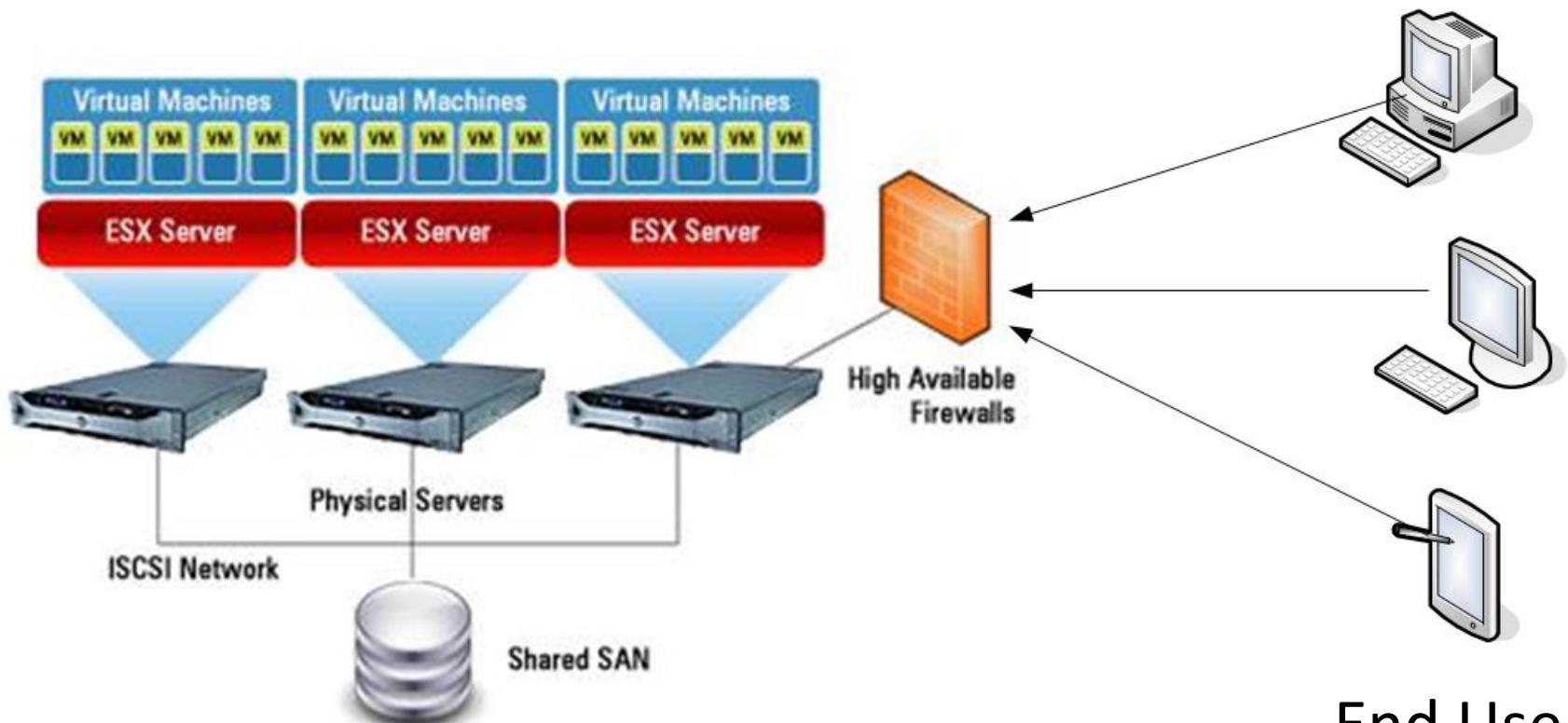


Hype Cycle of Cloud Computing 2012



Source: Gartner (August 2012)

Cloud as a User View



Public Vs Private



Public Cloud

- Hosted at a Service Provider Site
- Supports multiple customers
- Often utilizes shared infrastructure
- Supports connectivity over the internet
- Suited for information that is not sensitive
- Can be cheaper than private cloud

Private Cloud

- Hosted at an Enterprise or a Service Provider site
- Supports one customer
- Does not utilize shared infrastructure
- Connectivity over private network/ fiber or the internet
- Suited for information that needs a high level of security

Public Cloud Taxonomy

Infrastructure Services

Storage

- Amazon S3
- Amazon EBS
- CTERA Portal
- Mosso Cloud Files
- Nirvanix

Compute

- Amazon EC2
- Serve Path GoGrid
- Elastra
- Mosso Cloud Servers
- Joyent Accelerators
- AppNexus
- Flexiscale
- ElasticHosts
- Hosting.com CloudNine
- Terramark
- GridLayer
- ITRICITY
- LayeredTech

Services Management

- RightScale
- enStratus
- Scalr
- CohesiveFT
- Kaavo
- CloudStatus
- Ylastic
- Dynect
- CloudFoundry
- NewRelic
- Cloud42

Cloud Software

Data

- 10Gen MongoDB
- Oracle Coherence
- Gemstone Gemfire
- Apache CouchDB
- Apache HBase
- Hypertable
- TerraCotta
- Tokyo Cabinet
- Cassandra
- memcached

Compute

- Globus Toolkit
- Xeround
- Beowulf
- Sun Grid Engine
- Hadoop
- OpenCloud
- Gigaspaces
- DataSynapse
- Xeround

Cloud Management

- 3Tera App Logic
- OpenNebula
- OpenControlTier
- Enomaly Enomalism
- Altair Networks
- VMware vSphere
- OnPathTech
- CohesiveFT VPN Cubed
- Hyperic
- Eucalyptus
- Reductive Lbs Puppet
- OpenQRM
- Appistry

File Storage

- EMC Atmos
- ParaScale
- Zmanda
- CTERA
- Vordel

Appliances

- PingIdentity
- Symplyfied
- rPath
- Vordel

Platform Services

General Purpose

- Force.com
- Etelos
- LongJump
- AppJet
- Rollbase
- Bungee Labs Connect
- Google App Engine
- Engine Yard
- Caspio
- Qrimp
- MS Azure Services Platform
- Mosso Cloud Sites

Business Intelligence

- Aster DB
- Quantivo
- Cloud9 Analytics
- Blink Logic
- K2 Analytics
- LogiXML
- Oco
- Panorama
- PivotLink
- Sterna
- ColdLight Neuron
- Infobright
- Vertica

Integration

- Amazon SQS
- MuleSource Mule OnDemand
- Boomi
- SnapLogic
- OpSource Connect
- Cast Iron
- Microsoft BizTalk Services
- gnip
- SnapLogic SaaS Solution Packs
- Appian Anywhere
- HubSpan
- Informatica On-Demand

Development & Testing

- Keynote Systems
- Mercury
- SOASTA
- SkyTap
- Aptana
- LoadStorm
- Collabnet
- Dynamsoft

Database

- Google BigTable
- Amazon SimpleDB
- FathomDB
- Microsoft SDS

Software Services

Billing

- Aria Systems
- eVapt
- OpSource
- Redi2
- Zuora

Financials

- Concur
- Xero
- Workday
- Beam4d

Legal

- DirectLaw
- Advologix
- Fios
- Sertifi

Sales

- Xactly
- LucidEra
- StreetSmarts
- Success

Desktop Productivity

- Zoho
- IBM Lotus Live
- Google Apps
- Desktoptwo
- Parallels
- ClusterSeven

Human Resources

- Taleo
- Workday
- iCIMS

Content Management

- Clickability
- SpringCM
- CrownPoint

Backup & Recovery

- JungleDisk
- Mozy
- Zmanda Cloud Backup
- OpenRSM
- Syncplicity

CRM

- NetSuite
- Parature
- Responsys
- Rightnow
- Salesforce.com
- LiveOps
- MSDynamics
- Oracle On Demand

Document Management

- NetDocuments
- Questys
- DocLanding
- Aconex
- Xythos
- Knowledge TreeLive
- SpringCM

DPU-CT



SaaS

Software
as a Service

Email
CRM
Collaborative
ERP



PaaS

Platform
as a Service

Application Development
Decision Support
Web
Streaming



IaaS

Infrastructure
as a Service

Caching
Legacy
Networking
Security

File
Technical
System Mgmt

CONSUME



BUILD ON IT



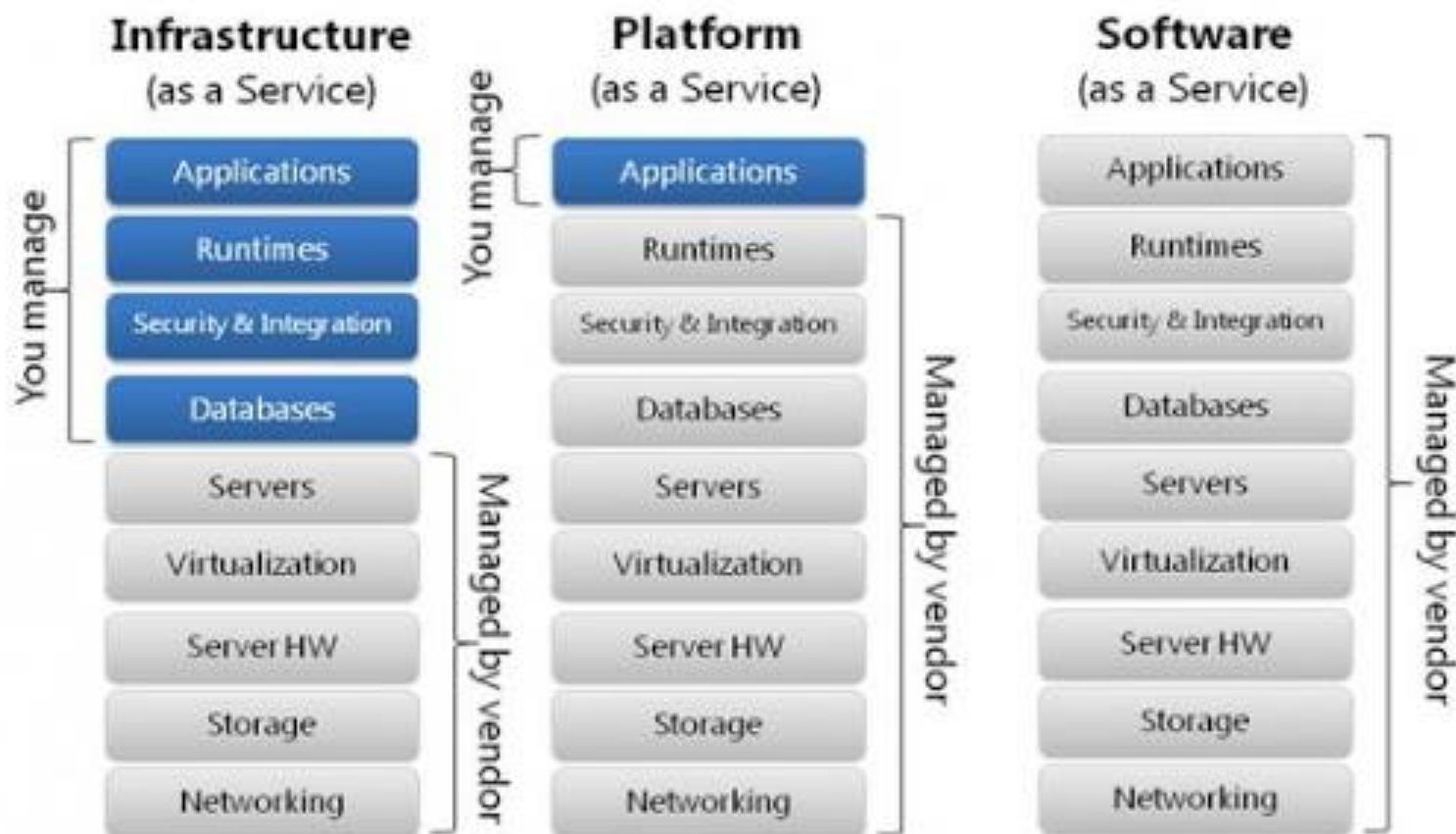
MIGRATE TO IT



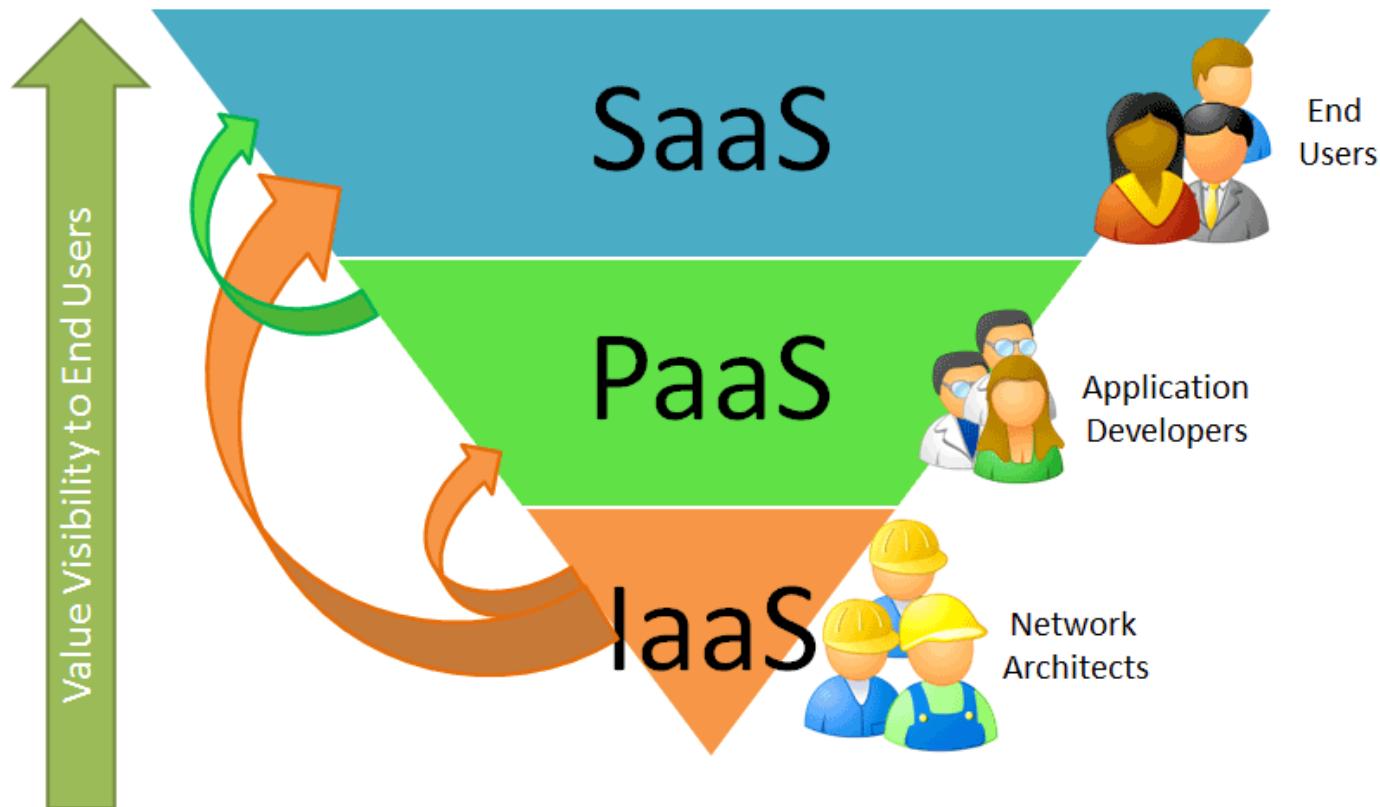
Public Cloud Services

	Amazon	Google	Microsoft	Salesforce
SaaS				
PaaS				
IaaS				

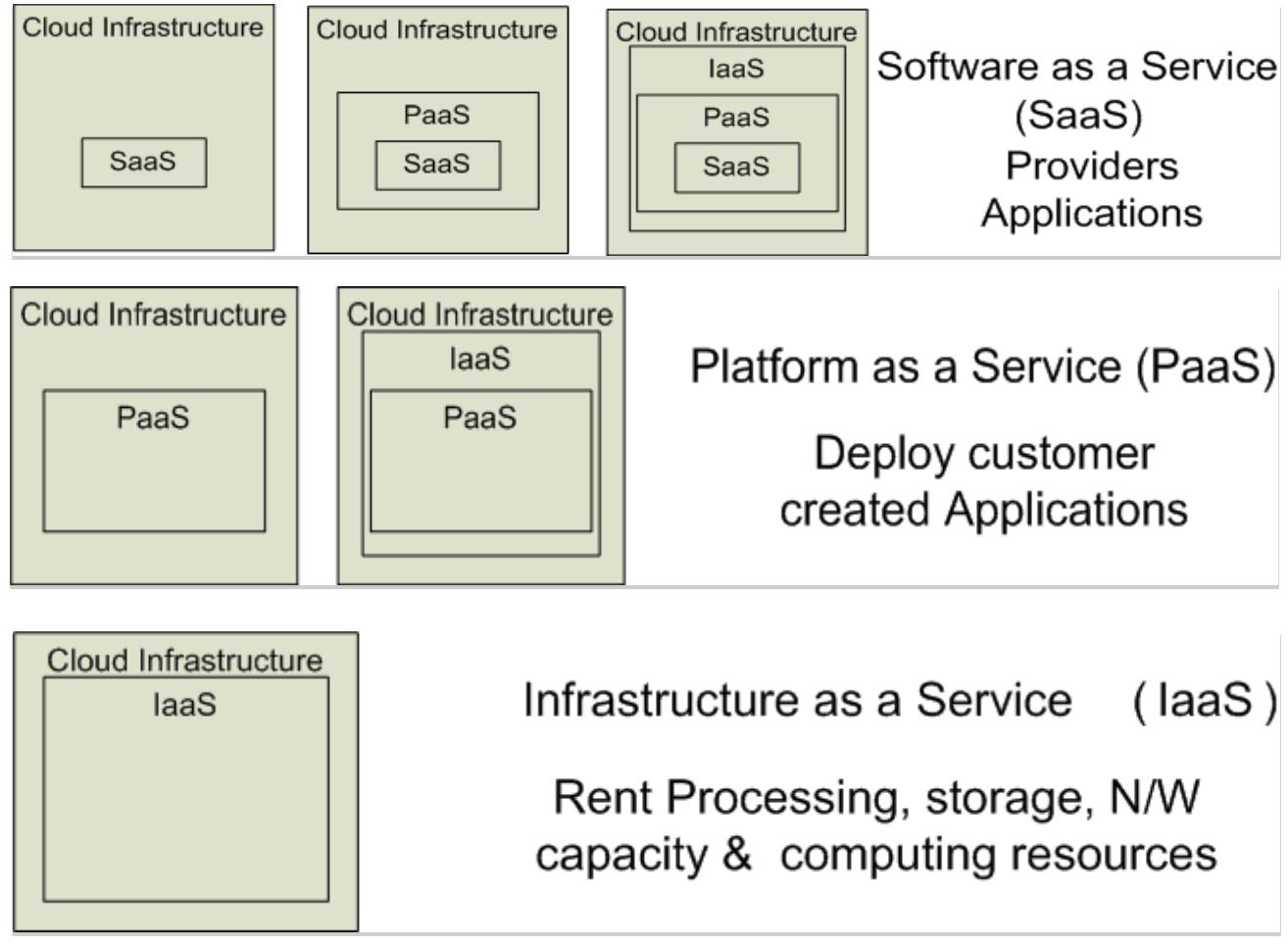
Where we can manage



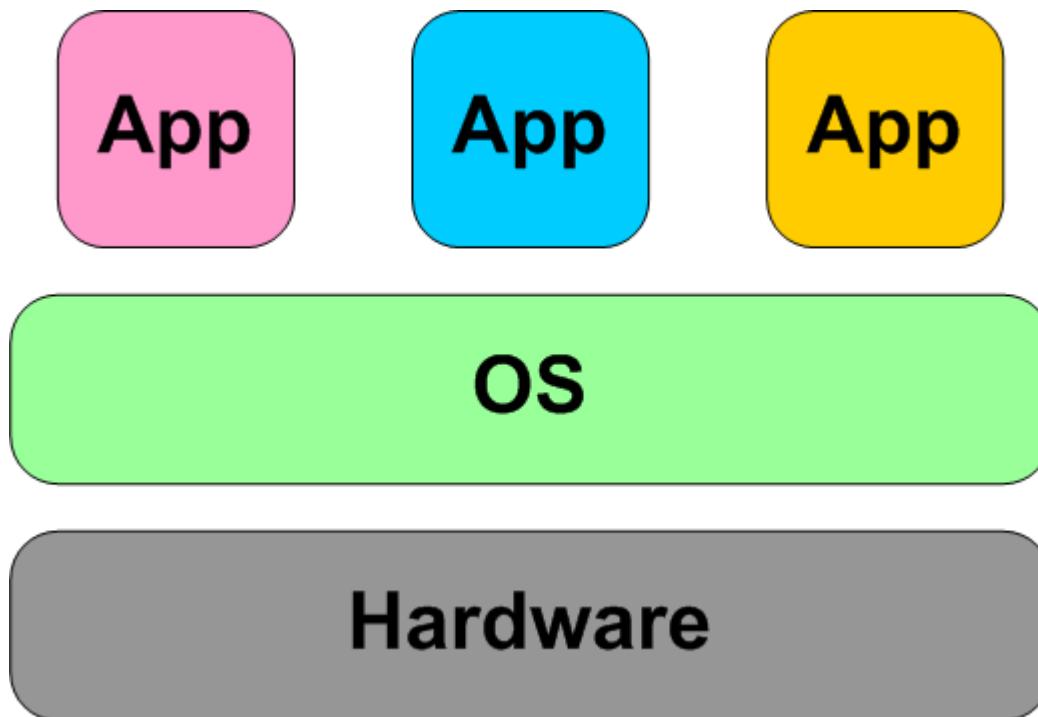
Value Visibility to End User



Starting Study



Traditional Computer



Applying Cloud on PC (Hosted Architecture)



Virtualization
(VM-Play / VM-Work / Hyper-V / VirtualBox)

OS

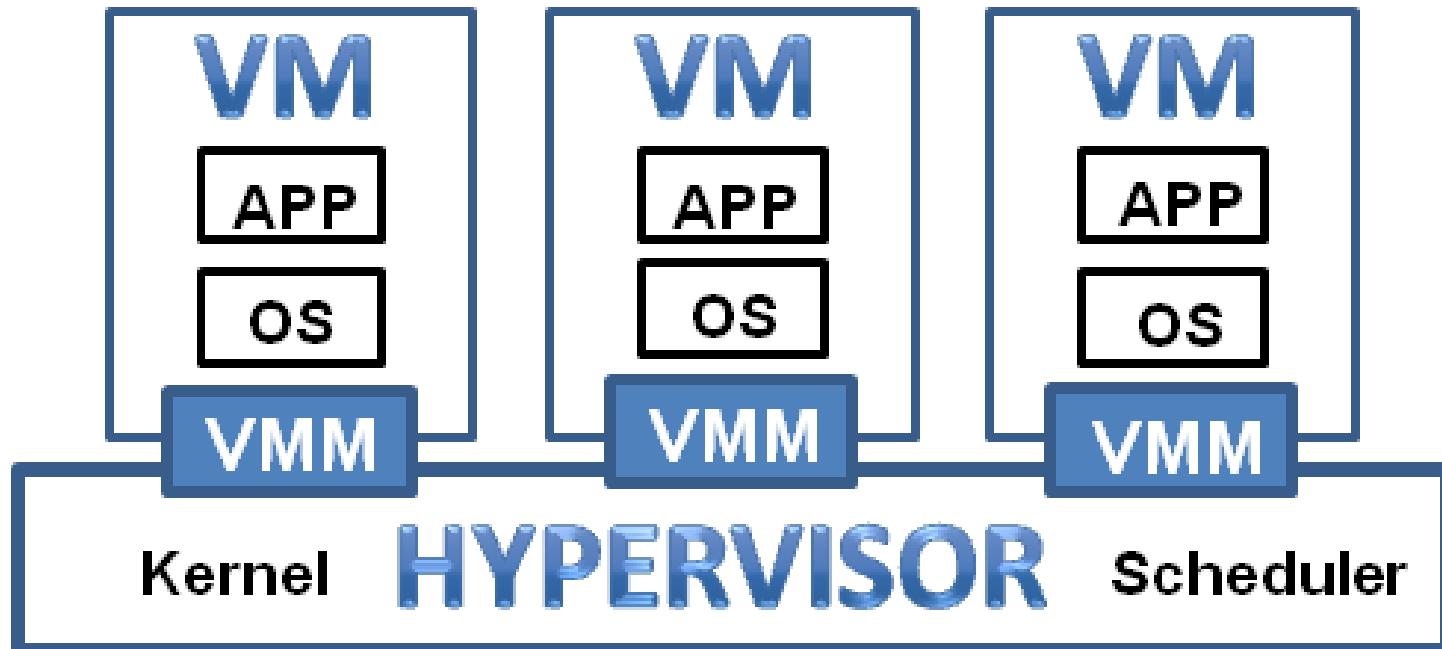
Hardware

Enterprise Cloud Computing

(Bare-Metal or Hypervisor Architecture)



Hypervisor



PHYSICAL HARDWARE

CPU

Storage

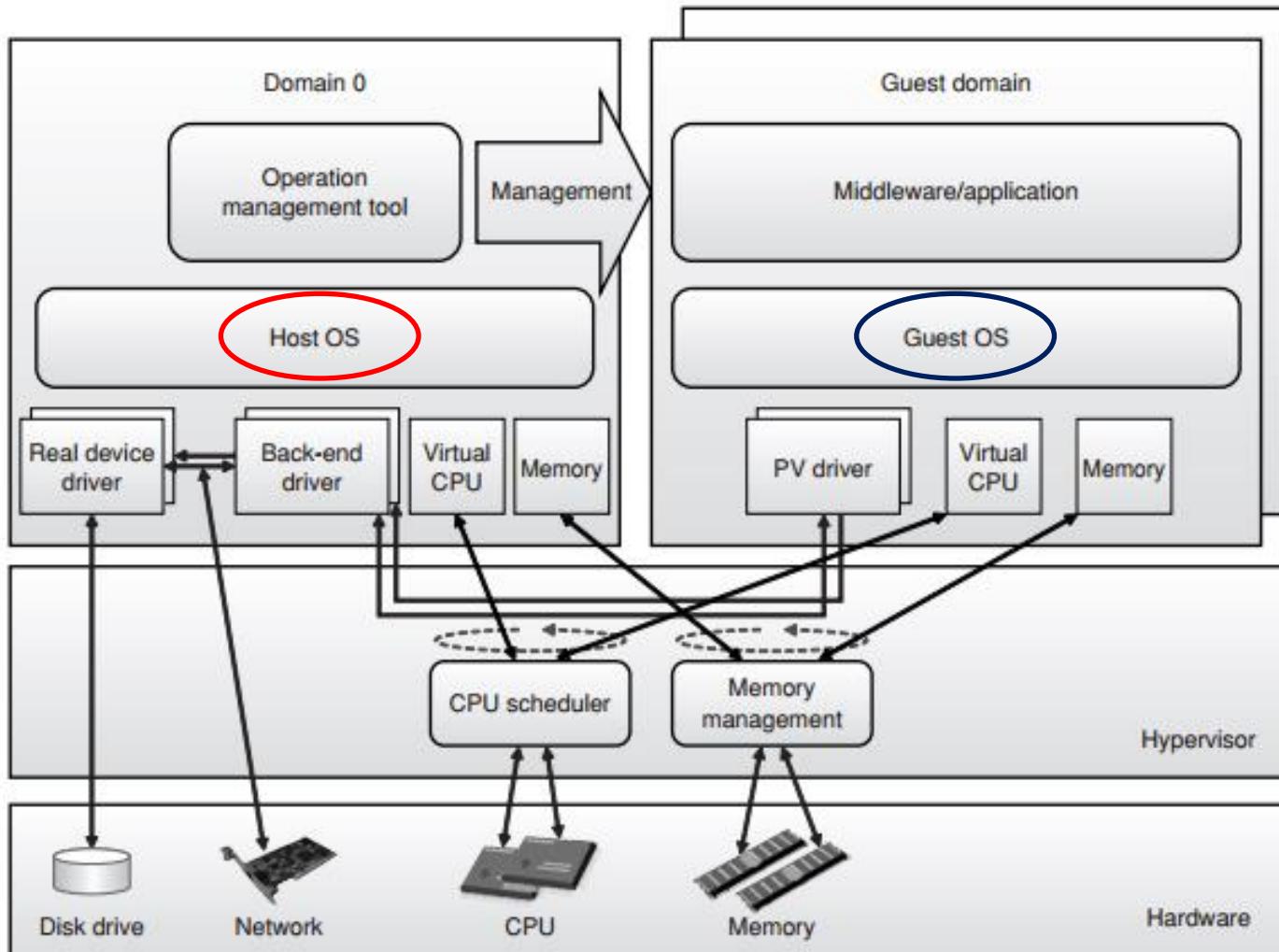
Memory

Networking

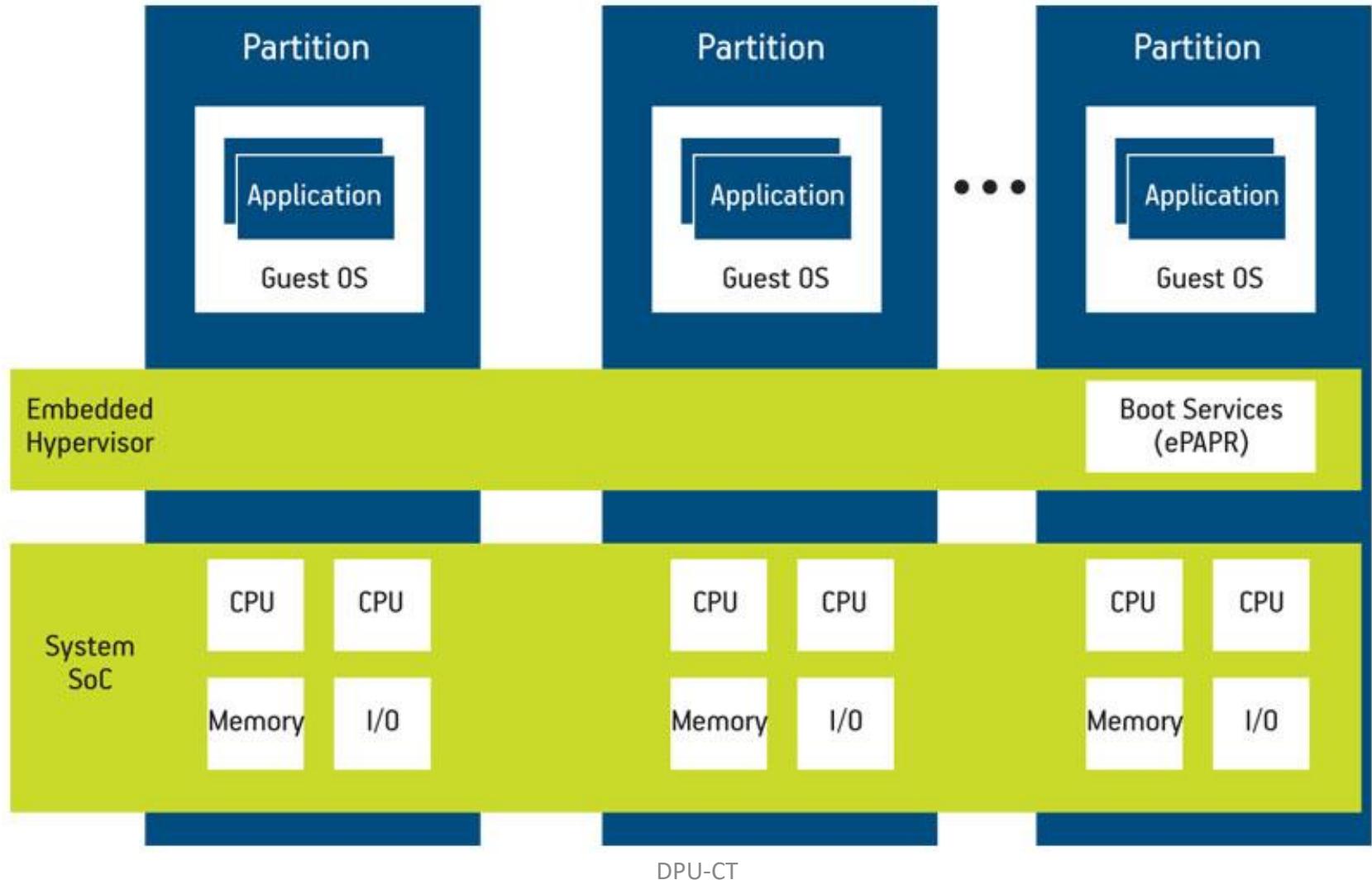
Hypervisors And Hosts

- A **hypervisor** is a piece of computer software, firmware or hardware that creates and runs virtual machines.
- A computer on which a hypervisor is running one or more virtual machines is defined as a **host machine**.
- Each **virtual machine** has a **guest operating systems**, which is managed by the hypervisor.
- Multiple instances of a variety of operating systems may share the virtualized hardware resources.

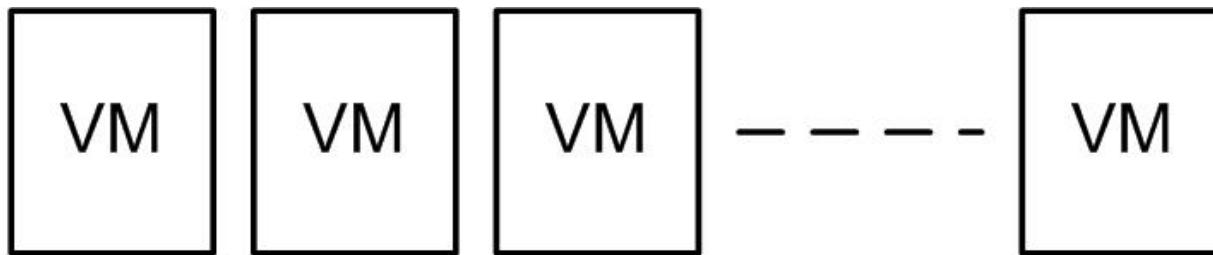
Virtualization / Hypervisor



Cloud Architecture



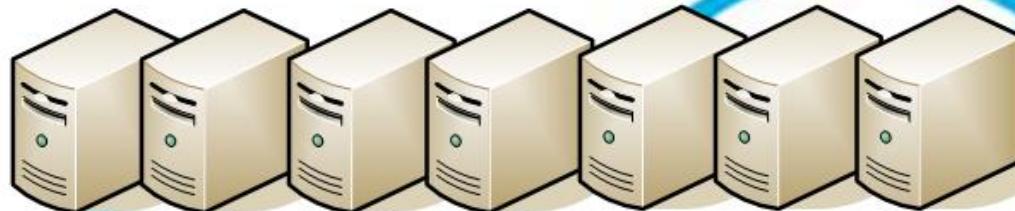
Enterprise Cloud Architecture (On Premise)



Virtual Technology Infrastructure

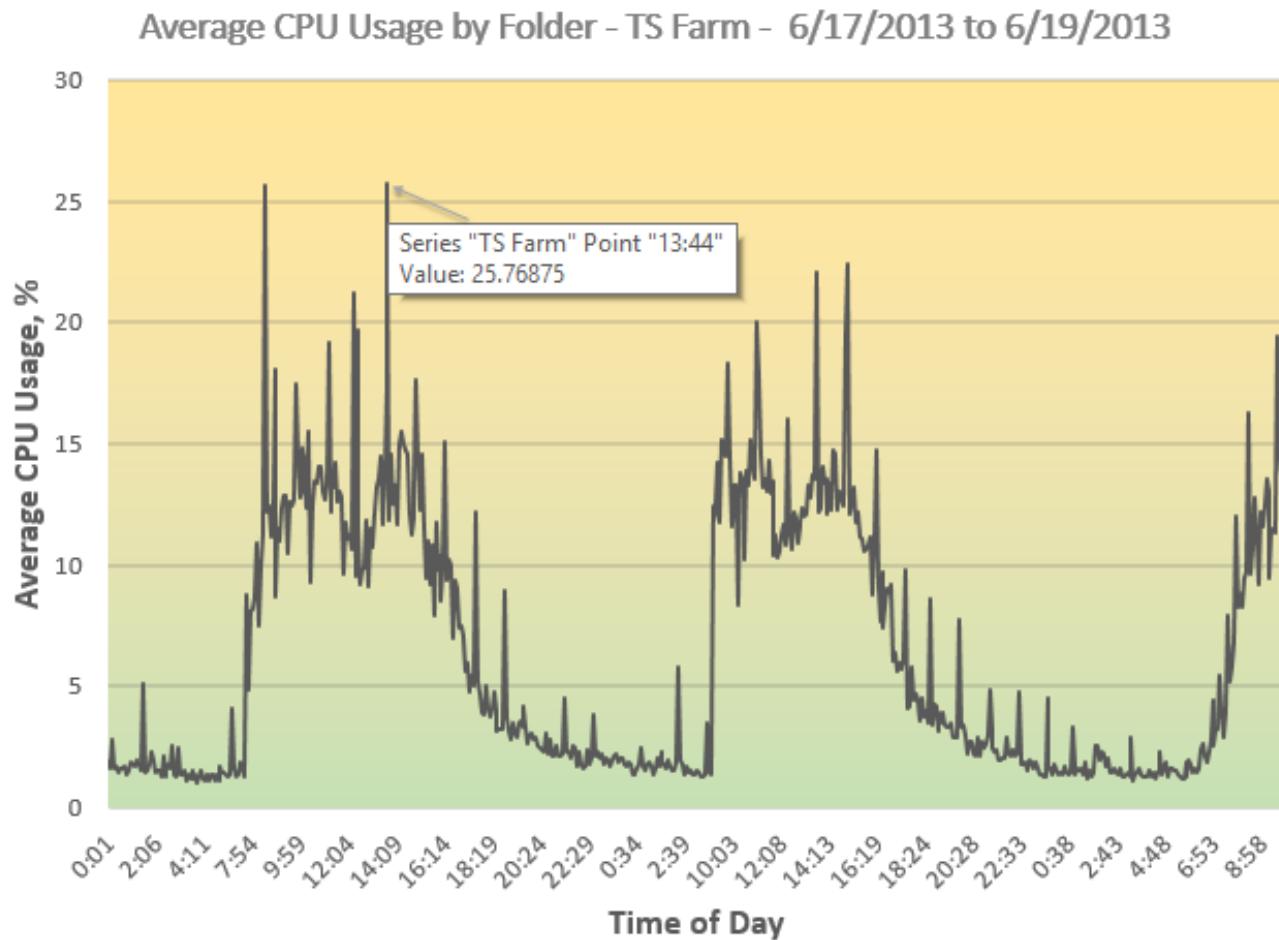
Hypervisor
(Linux KVM / Microsoft HyperV / VMWare SS,ESX / Citrix XEN)

Physical Infrastructure

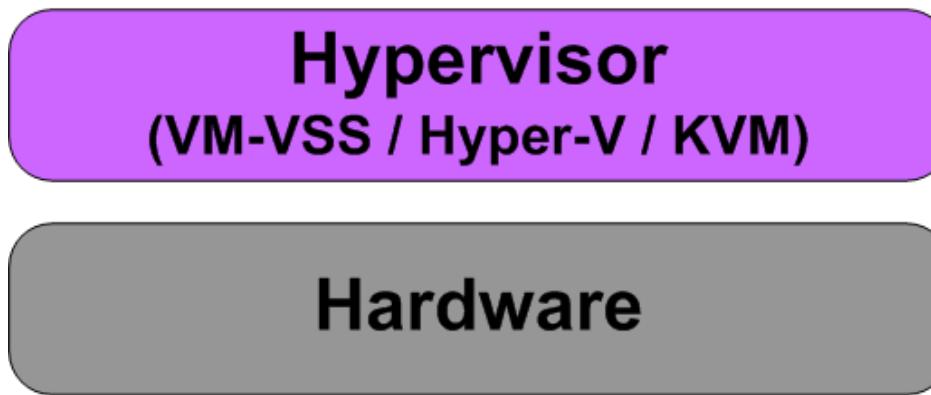
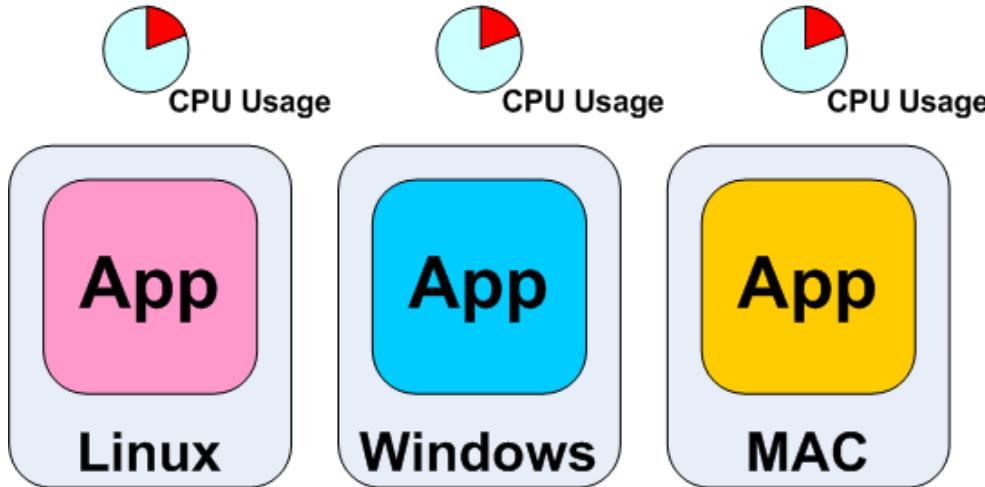


Network

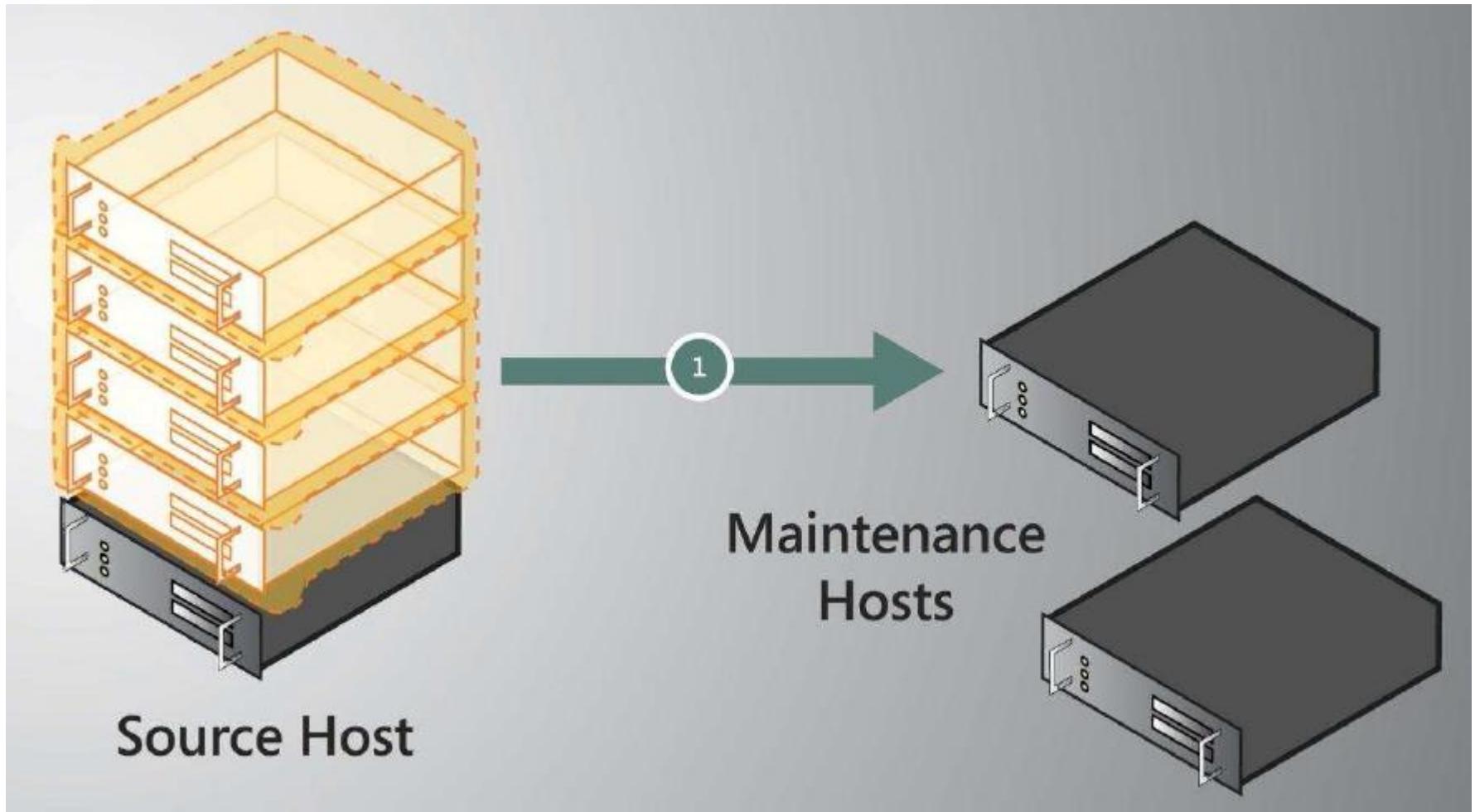
CPU Usage for each Application



Sharing CPU in Cloud Computing



Why migration !!!



VM Migration on Cloud Computing

Cold Migration

- Stop VM is needed
- Less complex
- Step
 - Stop VM on a current host
 - Resume VM on a new host
- Take a time to migrate about minutes !!

Hot Migration

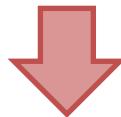
- VM needs to be non-stop
- Live migration
- More complex
- Step as in next slide
- Take a time in few seconds

Live Migration Step

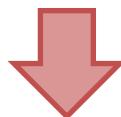
Configuration Data on Destination Host



Transferring memory pages



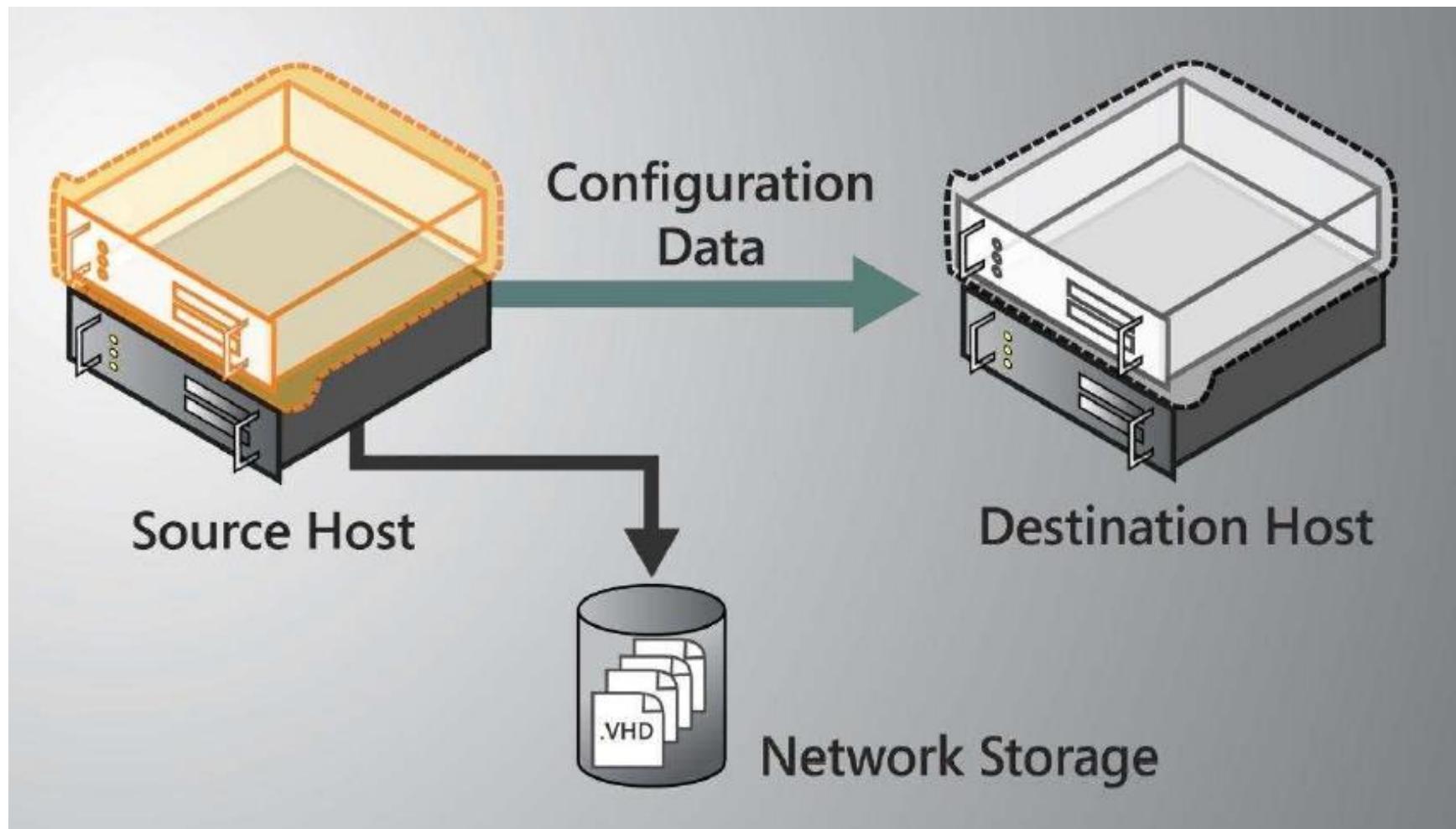
Move the storage handle



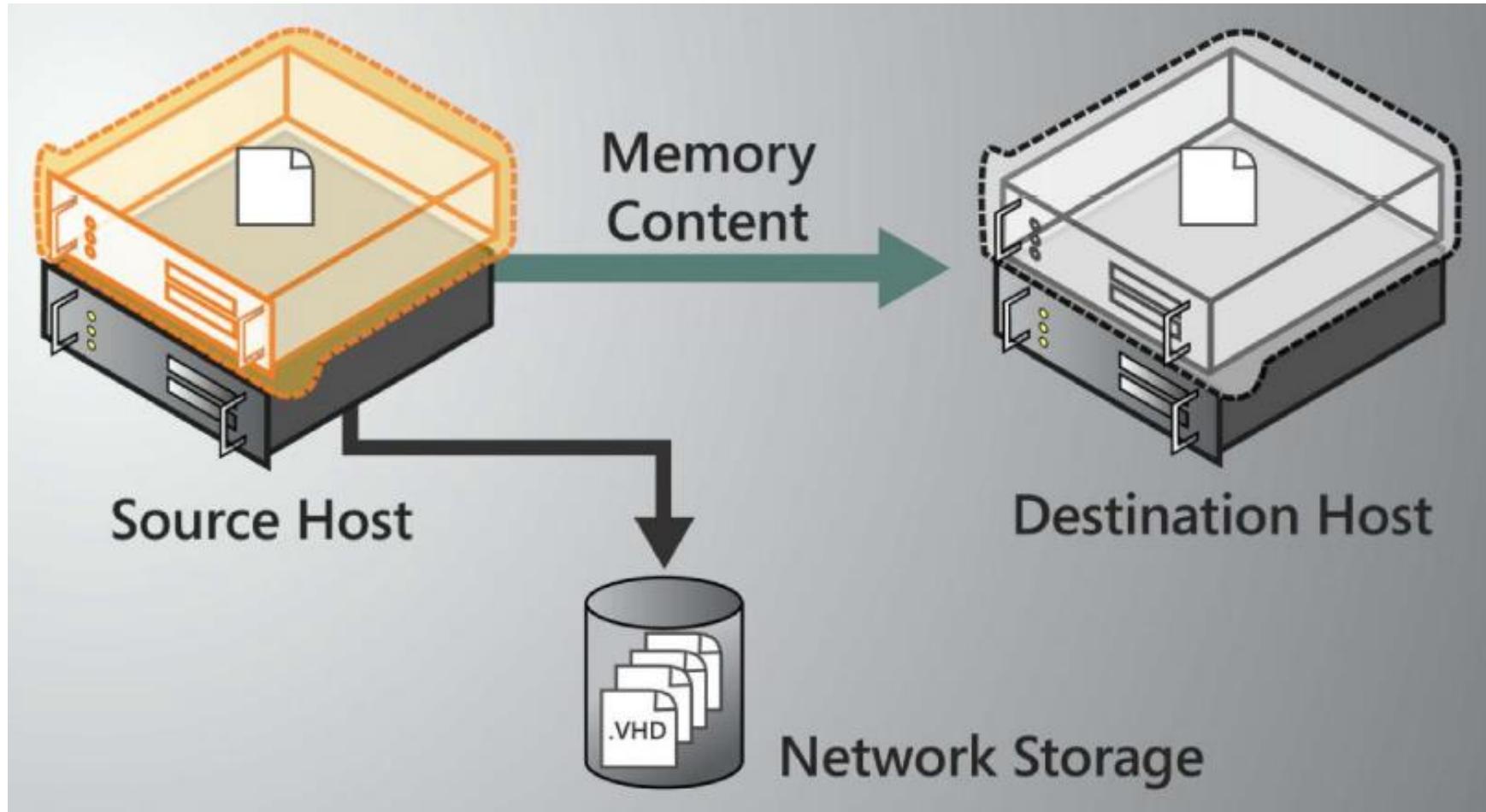
Resume VM online

0 to 10 seconds

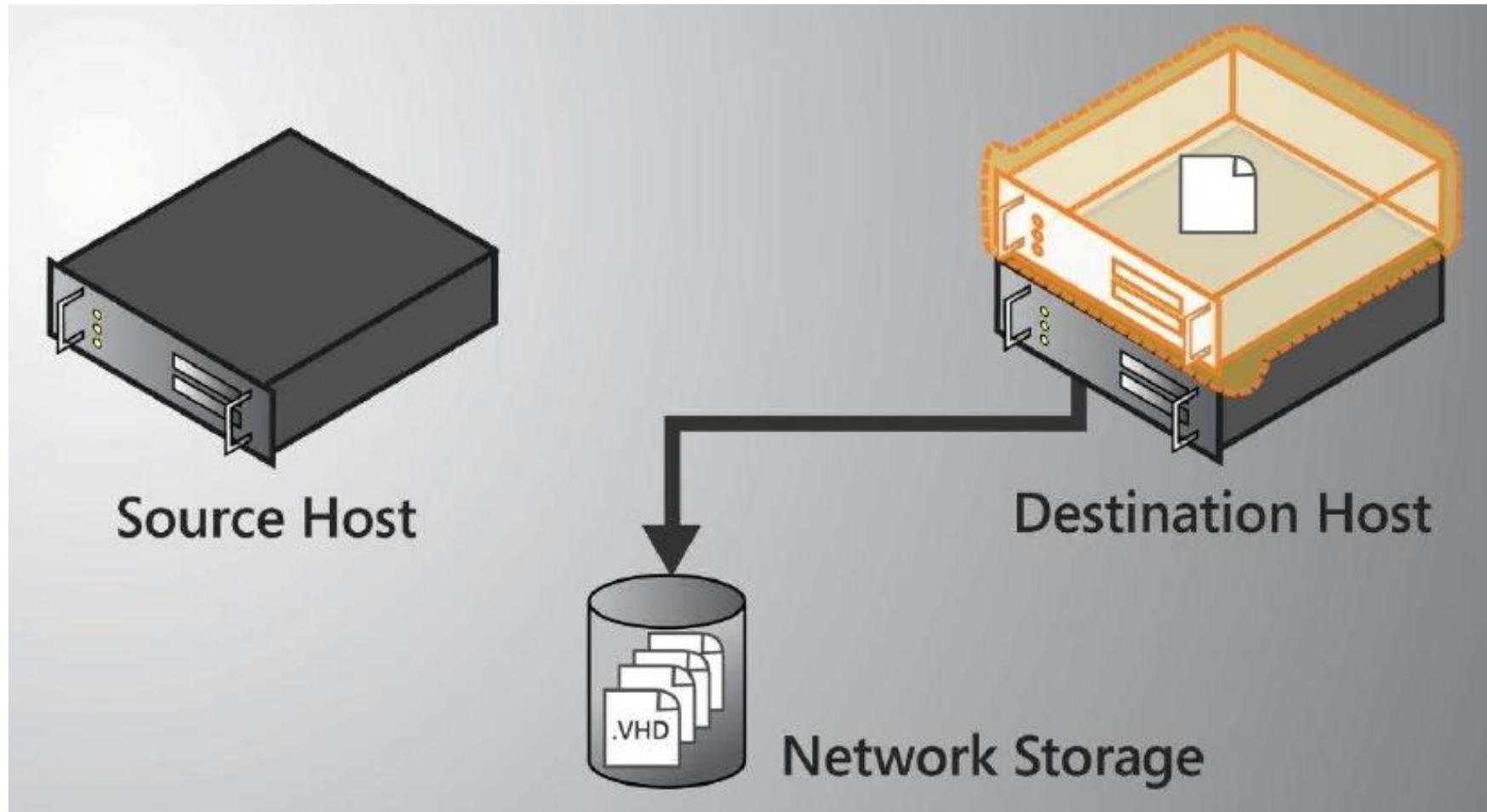
Configuration Data



Transferring memory

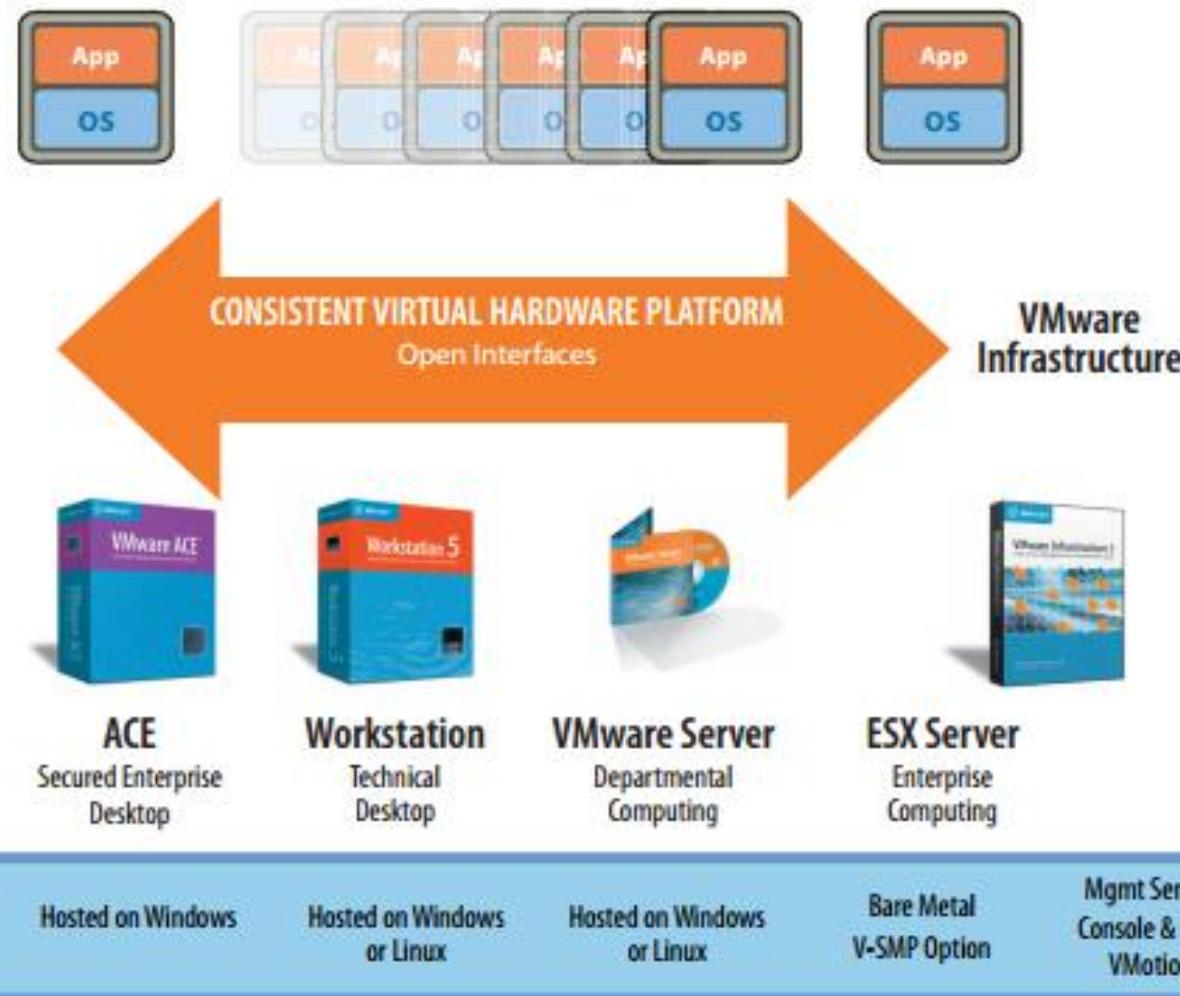


Handle Storage and Resume Online

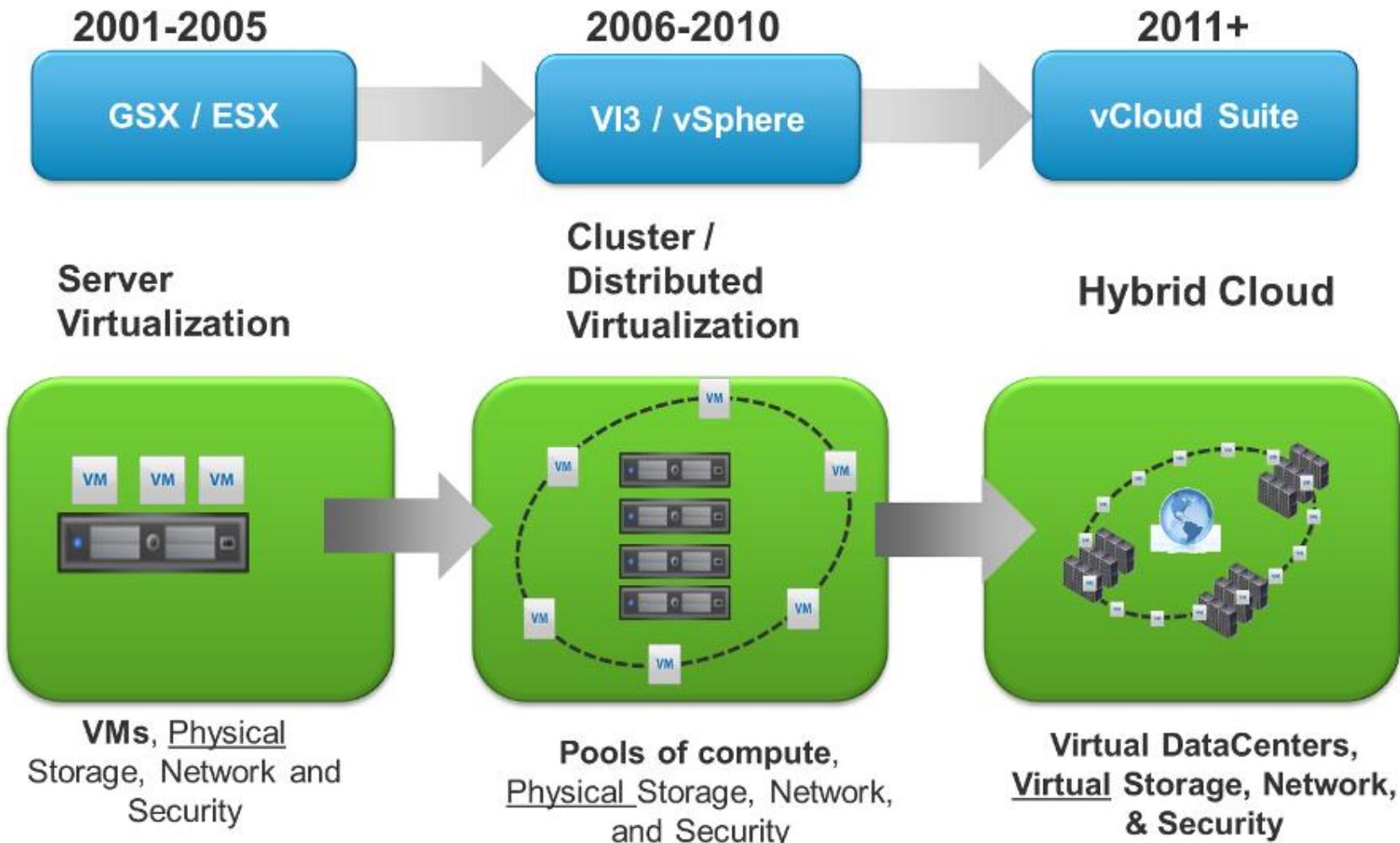


On Premise Battle

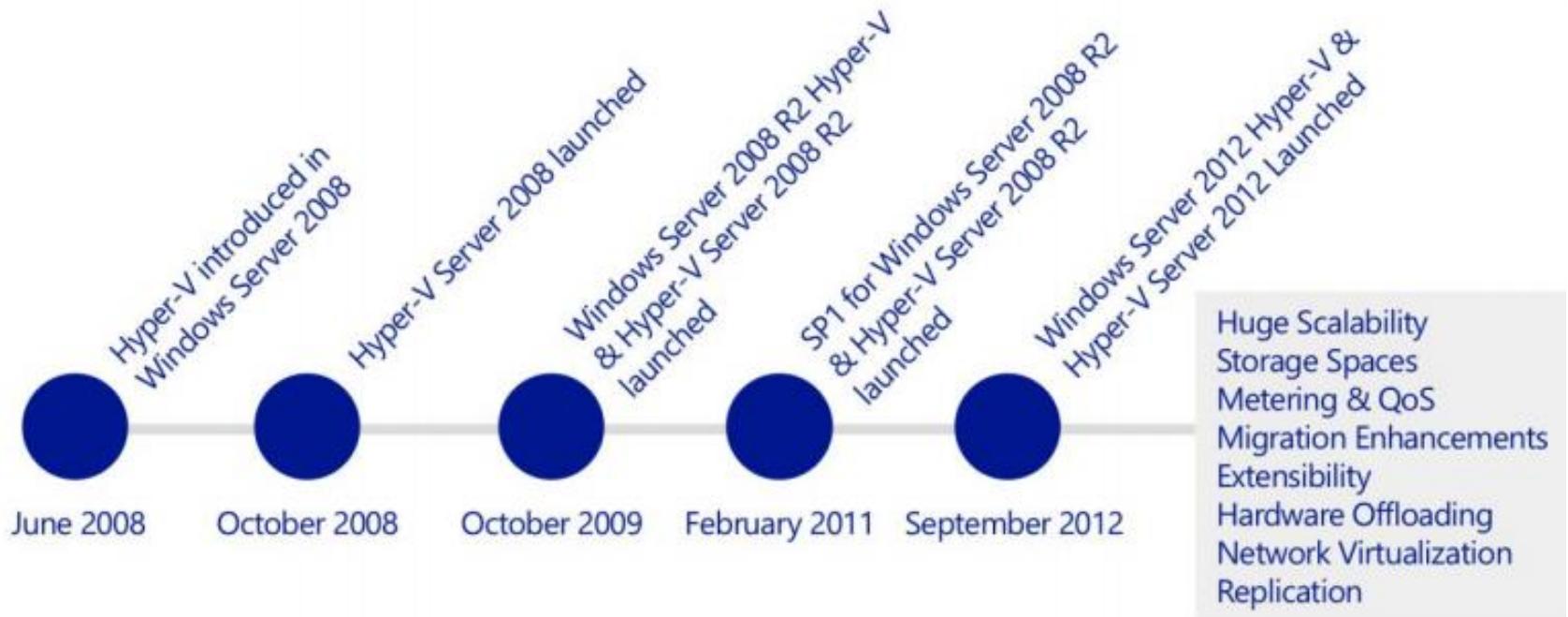
VMWare Cloud Computing Products



VMWare Evolution



Hyper-V Evolution



Hyper-V on WS2008R2 & WS2012R2

	Resource	Windows Server 2008 R2 Hyper-V	Windows Server 2012 R2 Hyper-V	Improvement Factor
Host	Logical Processors	64	320	5x
	Physical Memory	1TB	4TB	4x
	Virtual CPUs per Host	512	2,048	4x
VM	Virtual CPUs per VM	4	64	16x
	Memory per VM	64GB	1TB	16x
	Active VMs per Host	384	1,024	2.7x
	Guest NUMA	No	Yes	-
Cluster	Maximum Nodes	16	64	4x
	Maximum VMs	1,000	8,000	8x

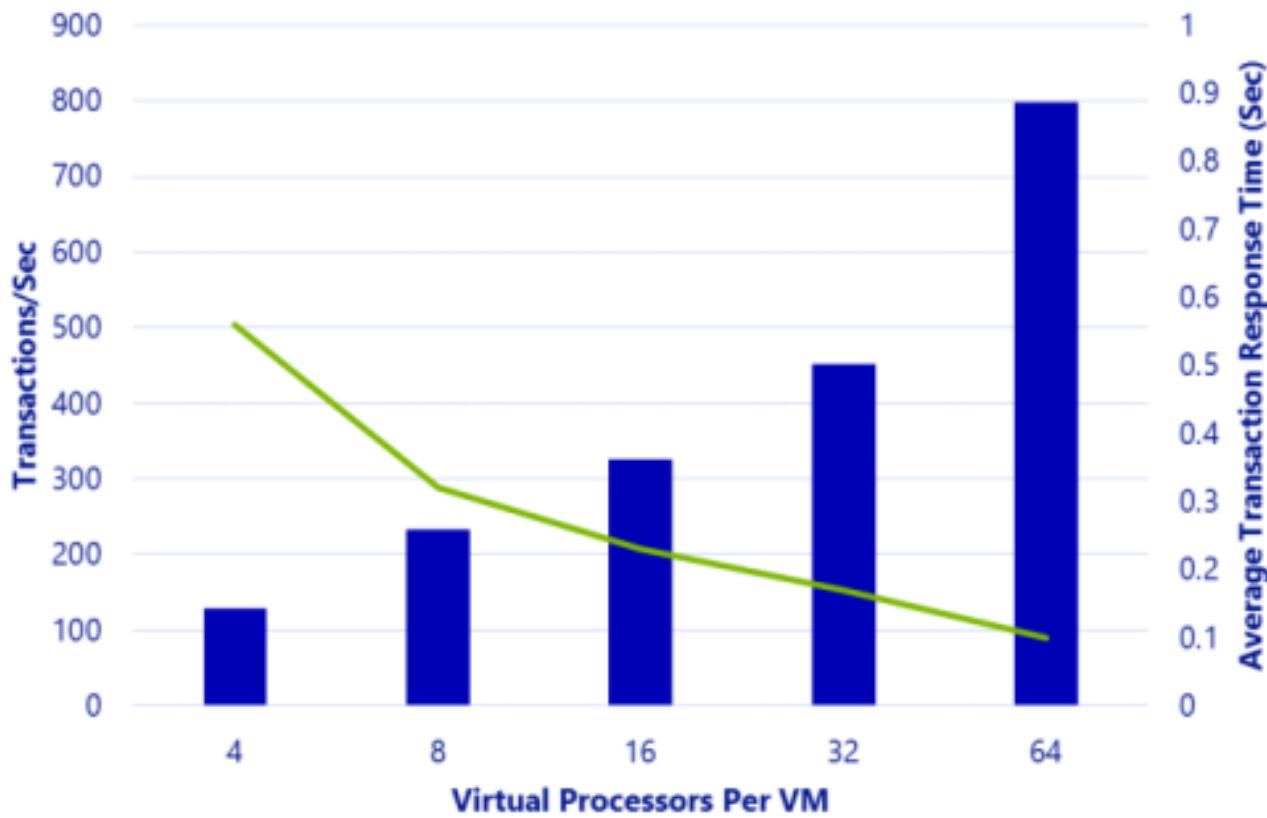
NUMA, or Non-Uniform Memory Access, inside a virtual machine.

NUMA refers to a computer architecture in multiprocessor systems, in which the required time for a processor to access memory depends on the memory's location relative to the processor

With NUMA, a processor can access local memory (memory attached directly to the processor) faster than it can access remote memory (memory that is local to another processor in the system). Modern operating systems and high-performance applications such as SQL Server have developed optimizations to recognize the system's NUMA topology and consider NUMA when they schedule threads or allocate memory to increase performance.

Hyper-V OLTP Performance

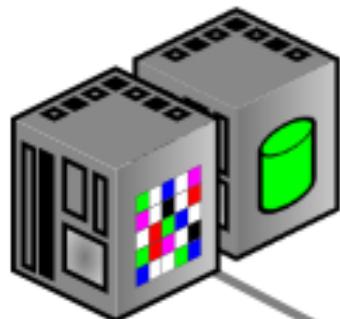
**Hyper-V Virtual CPU Scalability
with OLTP Workloads**



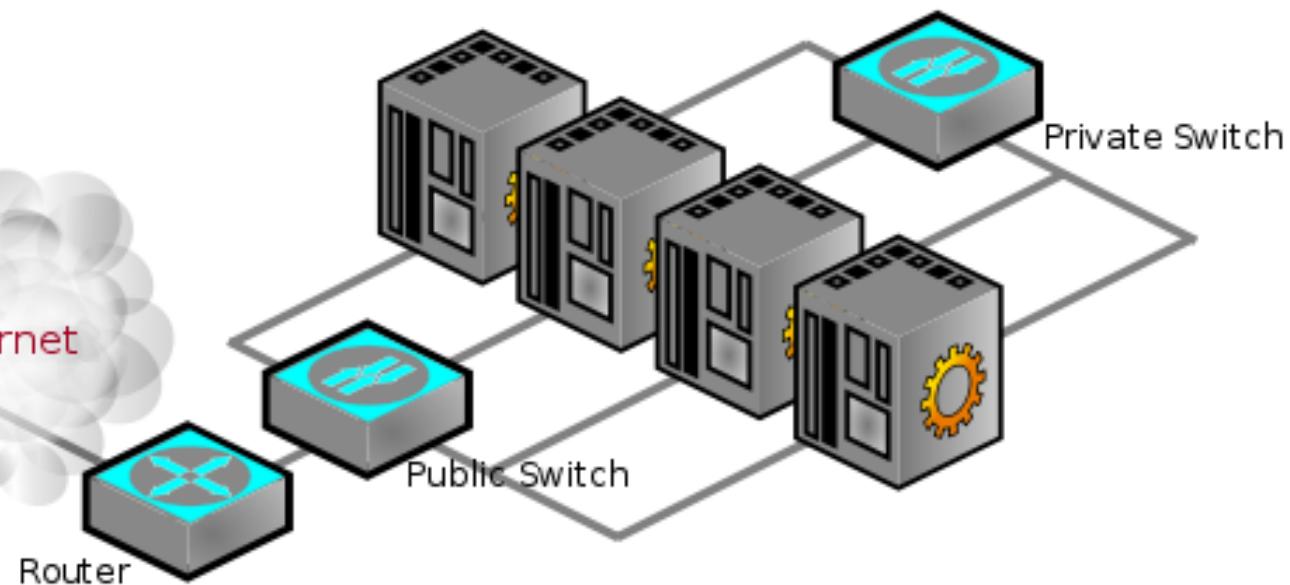
Windows Server 2012, SQL Server 2012, Single VM, 64GB of RAM

OpenStack

OpenStack Compute services
Database server on second node



Cloud of 2-4 virtual servers in one cluster
Self-contained storage of virtual images



Why and Why not OpenStack

Why

- Open Source
- Free to use
- Good Performance
- Many Software Suits
- Platform for Public Cloud
- Suitable for Data Center
- SUSE Cloud Platform

Why not

- Hard to use
- Hard to support
- Complexity
- Highly Technical
- High cost supporting

Feature	vSphere5			Hyper-V		
	Standard	Enterprise	Ent. Plus	Standard	Enterprise	DC
Max host processors	160	160	160	4	8	64
Max virtual SMP (guest)	8	8	32	4	4	4
Max host RAM (GB)	2048	2048	2048	32	2048	2048
Max RAM per VM	255	255	255	64	64	64
Failover nodes	32	32	32		16	16
Memory overcommit/dynamic mem.	✓	✓	✓	✓	✓	✓
Transparent page sharing	✓	✓	✓			
Live workload migration		✓	✓		✓	✓
Live storage migration		✓	✓			
Max guests per host	512	512	512	384	384	384
Distributed Resource Scheduler		✓	✓			
Distributed switch			✓			
Virtual instance rights (Windows)	0	0	0	1	4	Unlimited
Hypervisor licensing model	per proc	per proc	per proc	per host	per host	per proc

Sources: Configuration Maximums (VMware), Requirements and Limits for Virtual Machines and Hyper-V in Windows Server 2008 R2, Windows Server 2008 R2 Editions Comparison

Hyper-V Vs VSS

Resource		Windows Server 2012 R2 Hyper-V	VMware vSphere Hypervisor	VMware vSphere 5.5 Enterprise Plus
Host	Logical Processors	320	320	320
	Physical Memory	4TB	4TB	4TB
	Virtual CPUs per Host	2,048	4,096	4,096
VM	Virtual CPUs per VM	64	8	64
	Memory per VM	1TB	1TB	1TB
	Active VMs per Host	1,024	512	512
	Guest NUMA	Yes	Yes	Yes
Cluster	Maximum Nodes	64	N/A	32
	Maximum VMs	8,000	N/A	4,000

Live Migration Comparison

Capability	Windows Server 2012 R2 Hyper-V	VMware vSphere Hypervisor	VMware vSphere 5.5 Enterprise Plus
VM Live Migration	Yes	No	Yes
VM Live Migration with Compression	Yes	N/A	No
VM Live Migration over SMB/RDMA	Yes	N/A	No
1 GbE Simultaneous Live Migrations	Unlimited	N/A	4
10 GbE Simultaneous Live Migrations	Unlimited	N/A	8
Live Storage Migration	Yes	No	Yes
Shared-Nothing Live Migration	Yes	No	Yes
Live Migration Upgrades	Yes	N/A	Yes
VM Live Cloning	Yes	No	Yes
Network Virtualization	Yes	No	VXLAN/NSX
Network Virtualization Gateway	Yes	No	vCloud Suite
Flexible Linux Guest OS Support	Yes	Yes	Yes

Cluster Features Comparison

Capability	Windows Server 2012 R2 Hyper-V	VMware vSphere Hypervisor	VMware vSphere 5.5 Enterprise Plus
NIC Teaming	Yes	Yes	Yes
Integrated High Availability	Yes	No	Yes
Nodes per Cluster	64	N/A	32
VMs per Cluster	8,000	N/A	4,000
Max Guest Cluster Size (iSCSI)	64 Nodes	5	5
Max Guest Cluster Size (Fiber)	64 Nodes	5	5
Max Guest Cluster Size (File Based)	64 Nodes	5	5
Guest Cluster with Shared Virtual Disk	Yes	Yes	Yes
Guest Clustering with Live Migration	Yes	N/A	No
Guest Clustering with Dynamic Memory	Yes	No	No
Guest OS Application Monitoring	Yes	N/A	Yes
Cluster-Aware Updating	Yes	N/A	Yes
Failover Priority, Affinity & Anti-Affinity	Yes	N/A	Yes

Public Cloud Battle

Public Cloud Players

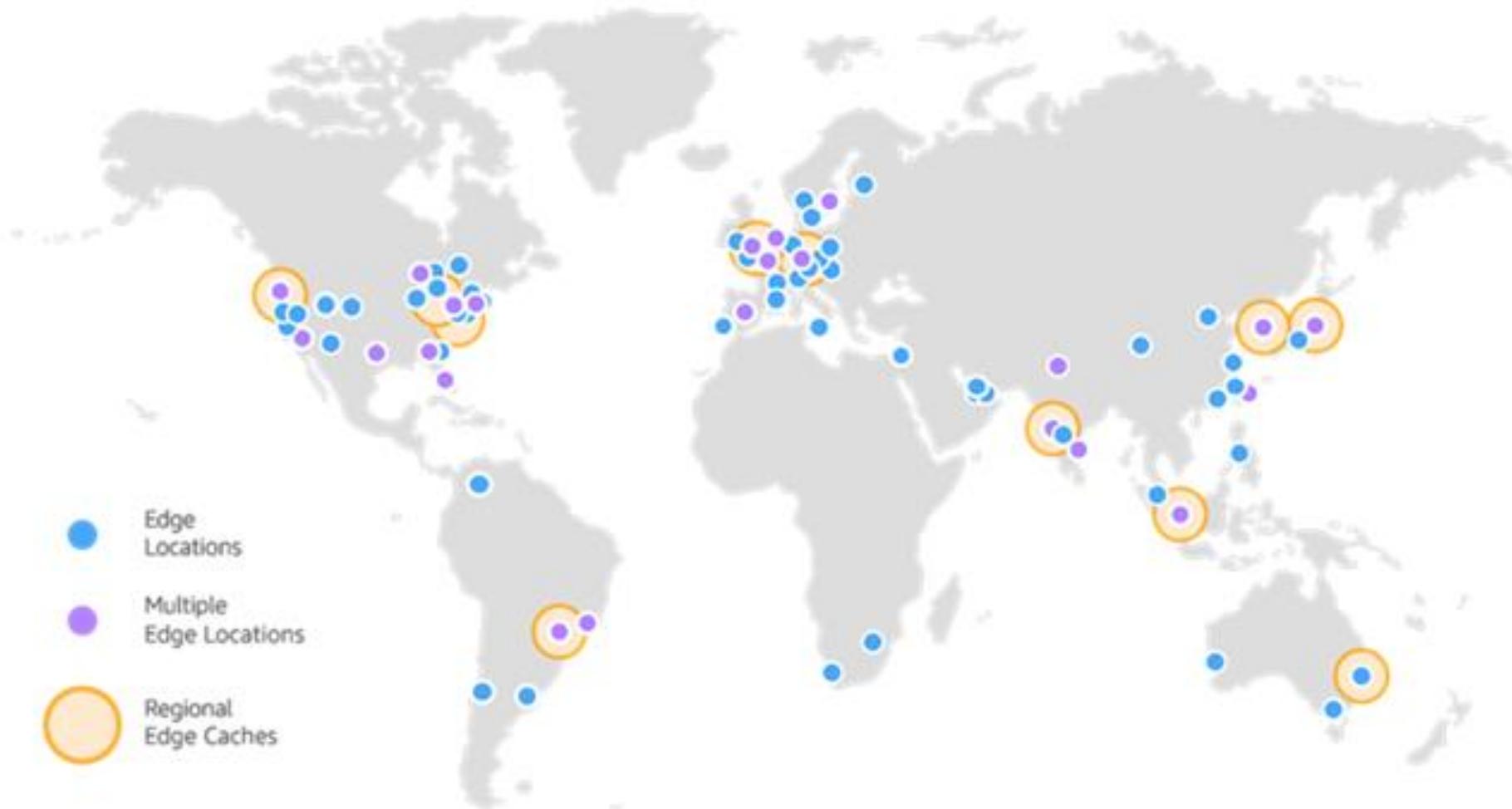


Google Cloud Platform



DigitalOcean

AWS Locations



AWS Products/Services

Compute

Amazon Elastic Compute Cloud (Amazon EC2)



Amazon Elastic MapReduce

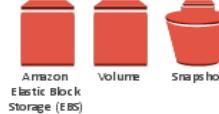


Storage

Amazon Simple Storage Service (Amazon S3)



Amazon Elastic Block Storage (Amazon EBS)



AWS Import/Export



AWS Storage Gateway Service

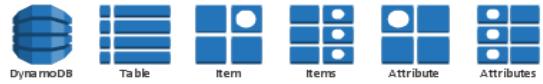


AWS Glacier



Database

Amazon DynamoDB



Amazon Relational Database Service (Amazon RDS)

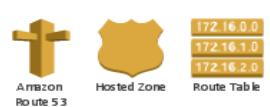


Amazon ElastiCache



Networking

Amazon Route 53



Amazon Elastic Load Balancing



AWS Direct Connect



Amazon Virtual Private Cloud (VPC)



Content Delivery

Amazon Cloudfront



Elastic Network Instance



Application Services

Amazon Simple Queue Service (SQS)



Amazon Cloudsearch



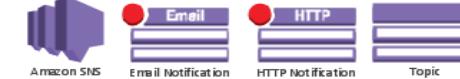
Amazon Simple Email Service (SES)



Amazon Simple Workflow (SWF)



Amazon Simple Notification Service (SNS)



Deployment and Management

Amazon Elastic Beanstalk



AWS Identity and Access Management (IAM)

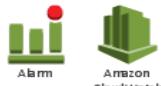


AWS CloudFormation



Monitoring

Amazon CloudWatch



Non-Service Specific

AWS Cloud



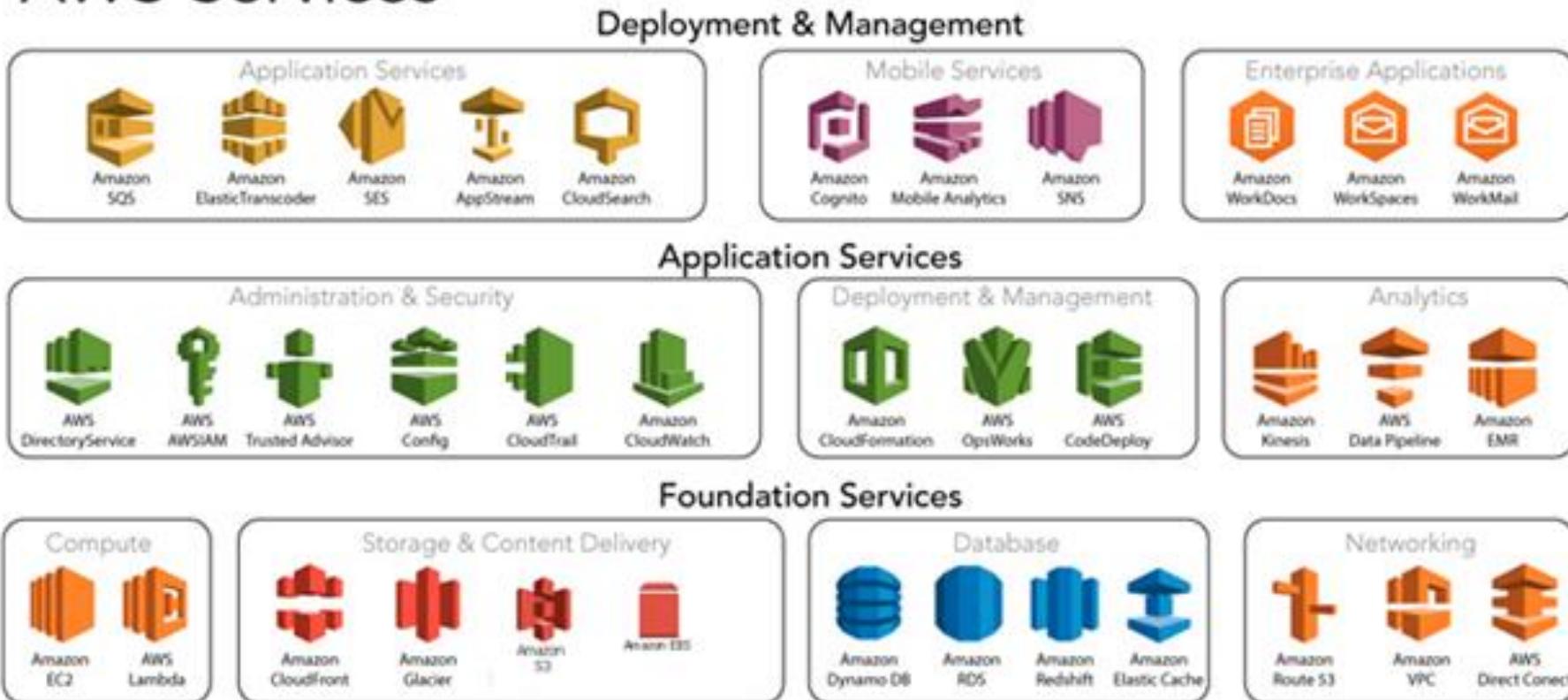
Groups



AWS Product Groups

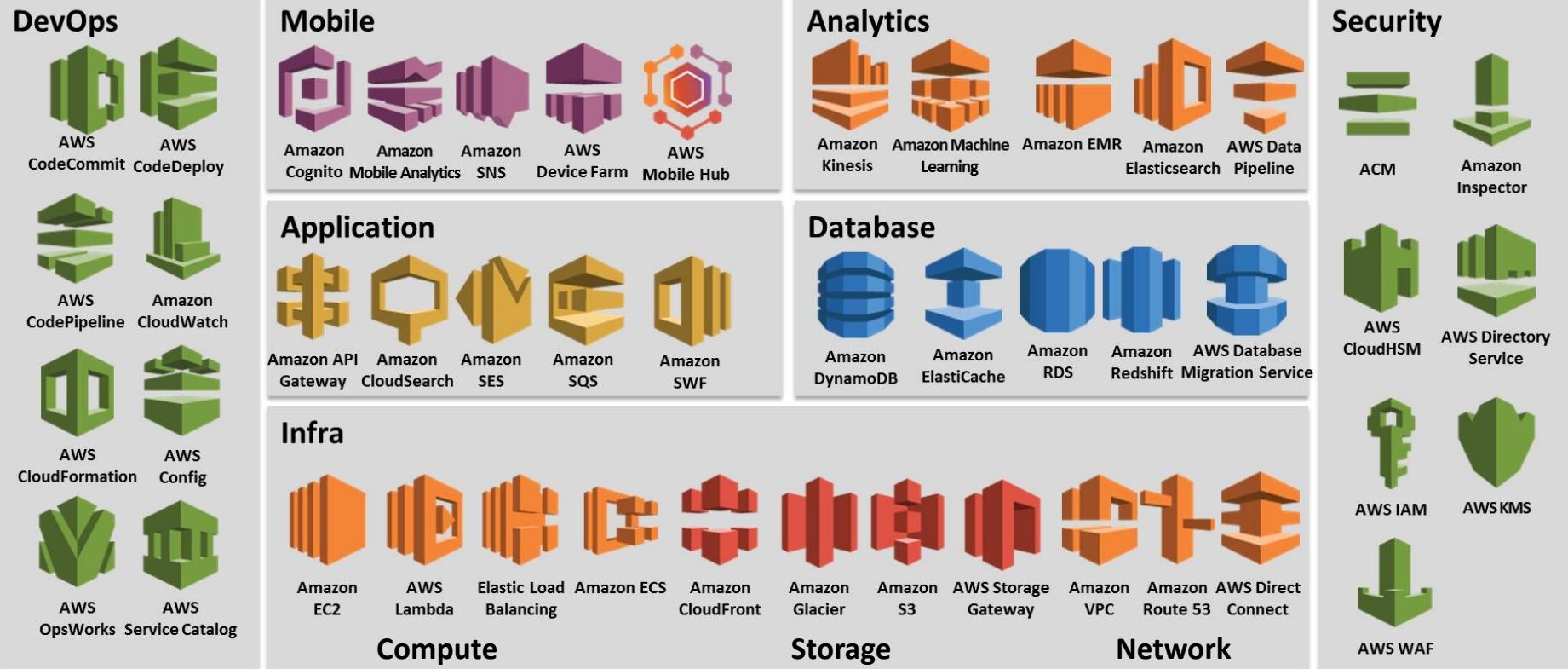
Services Level

AWS Services



AWS Product Groups

Service Groups



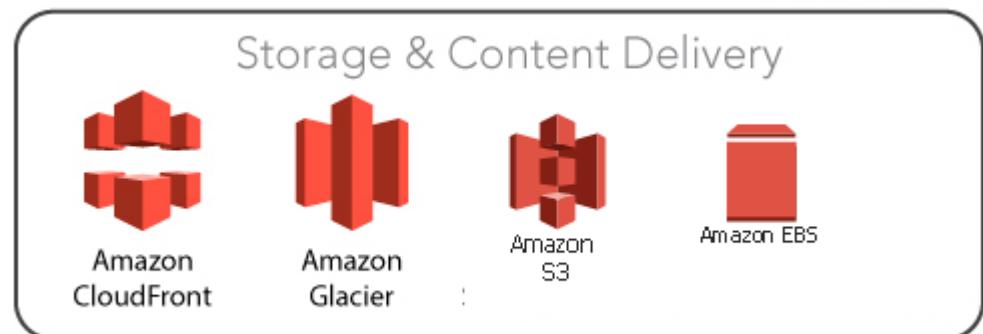
AWS Compute Services

- EC2 or Elastic Compute Cloud – Virtual machine for which user has OS level control.
- AWS Lambda – Run program functions in the cloud.



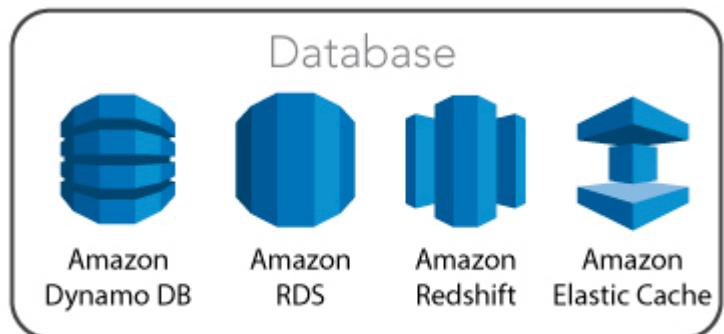
AWS Storage Services

- Amazon Glacier- Low-cost storage service for data archiving and backup.
- Amazon Elastic Block Store (EBS)- Provides block-level storage for Amazon EC2. EBS volumes are network-attached and are free from the life of an instance.
- Amazon S3 – S3 or simple storage service. It provides object storage.
- Cloud Front – It is a CDN or content delivery network to distribute the inert and active web content. For example, .html, .php files to clients



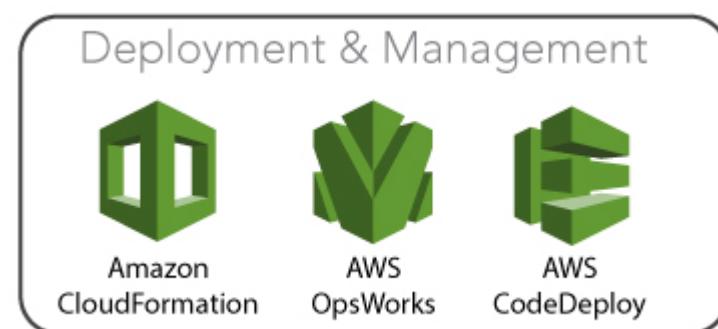
AWS Database Services

- Amazon RDS- Set up, operate, and scale a relational database in the cloud.
- Amazon DynamoDB- Fully managed NoSQL database service.
- Amazon ElastiCache- Deploy, operate, and scale an in-memory cache in the cloud.
- Amazon RedShift - Data warehousing solution to perform OLAP queries.



AWS DevOps

- CloudWatch — Monitor AWS environments like EC2, RDS instances, and CPU utilization. It also triggers alarms depends on various metrics.
- CloudFormation — It turns infrastructure into the cloud. Use templates to get whole production environment in minutes.
- OpsWorks — Automated Chef/Puppet deployments on AWS environment.



AWS Admin and Security

- IAM (Identity and Access Management) — IAM helps to manage users, assign policies, form groups to manage multiple users (Similar to Microsoft Active Directory)
- Certificate Manager — Offers free SSL certificates for your domains that are managed by Route53.
- WAF — Manage protection features for web applications.



GCP Products/Services

Management	Compute	Storage	Networking	Big Data	Machine Learning
 GOOGLE STACKDRIVER	 COMPUTE ENGINE	 CLOUD STORAGE	 VIRTUAL NETWORK	 BIGQUERY	 CLOUD ML
 IDENTITY AND ACCESS MANAGEMENT	 PREEMPTIBLE VMs	 NEARLINE	 LOAD BALANCING	 DATAFLOW	 SPEECH
	 CUSTOM MACHINE TYPES	 CLOUD SQL	 CDN	 DATAPROC	 NATURAL LANGUAGE
	 APP ENGINE	 DATASTORE	 DNS	 DATALAB	 VISION API
	 CONTAINER ENGINE	 BIGTABLE	 INTERCONNECT	 PUB/SUB	 TRANSLATE API

GCP Locations



* While some regions may launch with 2 zones, all regions are planned for a minimum of 3 zones.

Azure Products/Services

Compute

 Virtual Machines	 Virtual Machine Scale Sets
 Azure Container Service	 Azure Container Registry
 Functions	 Batch
 Service Fabric	 Cloud Services

Networking

 Virtual Network	 Load Balancer
 Application Gateway	 VPN Gateway
 Azure DNS	 Traffic Manager
 ExpressRoute	 Network Watcher

Storage

 Storage: Blobs, Tables, Queues, Files, Disks	 Data Lake Store
 StorSimple	 Azure Backup
 Site Recovery	

Monitoring & Management

 Azure Portal	 Azure Resource Manager	 Azure Advisor	 Azure Monitor	 Log Analytics	 Automation	 Scheduler
--	--	---	---	---	--	---

Web & Mobile

 Web Apps	 Mobile Apps
 Logic Apps	 API Apps
 Content Delivery Network	 Media Services
 Search	

Databases

 SQL Database	 SQL Data Warehouse
 SQL Server Stretch Database	 DocumentDB
 Redis Cache	 Data Factory

Intelligence & Analytics

 HDInsight	 Machine Learning
 Cognitive Services	 Azure Bot Service*
 Data Lake Analytics	 Power BI Embedded
 Azure Analysis Services	

Internet of Things & Enterprise Integration

 Azure IoT Hub	 Event Hubs
 Stream Analytics	 Notification Hubs
 BizTalk Services	 Service Bus
 Data Catalog	

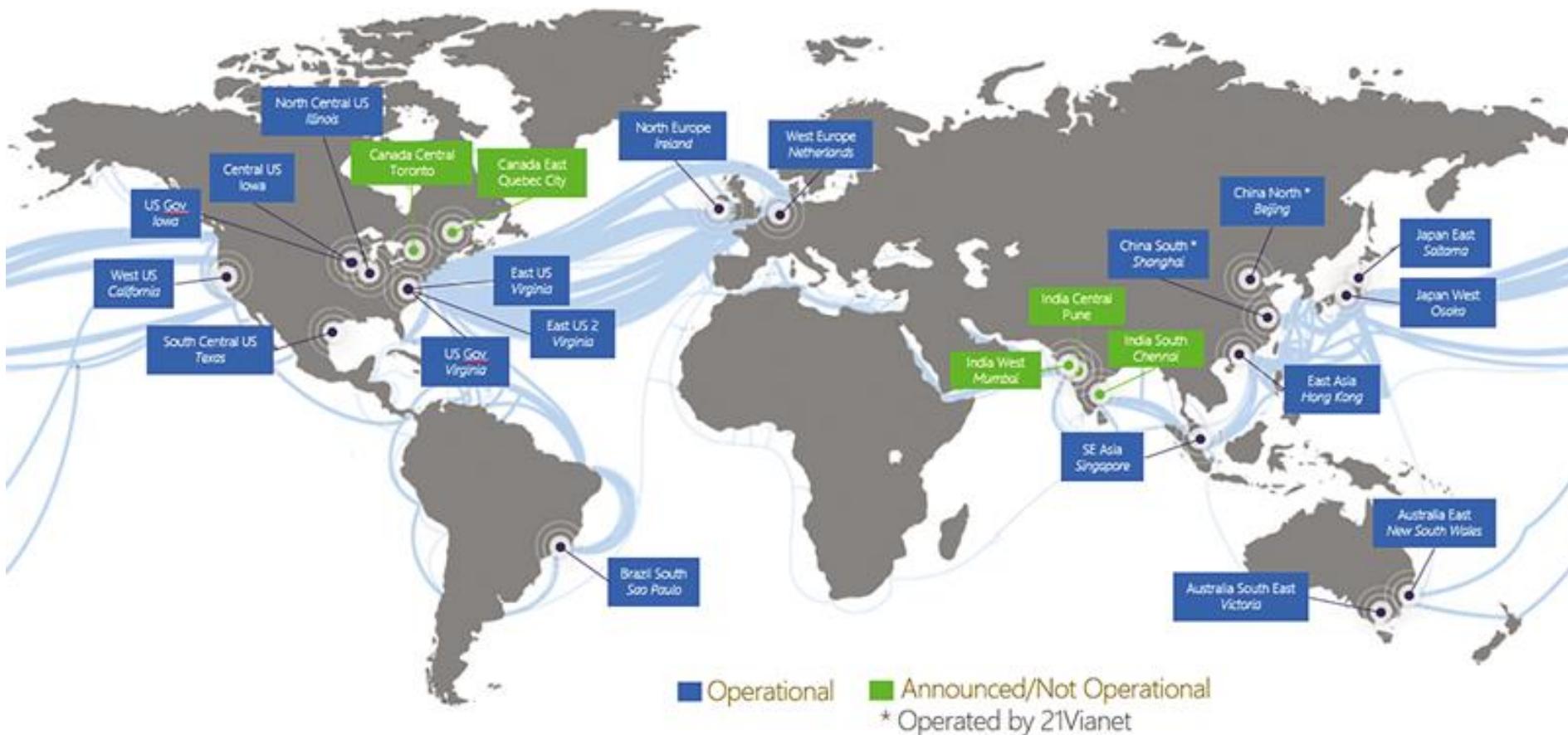
Security + Identity

 Security Center	 Key Vault
 Azure Active Directory	 B2C
 Domain Services	 Multi-Factor Authentication

Developer Services

 Visual Studio Team Services	 Azure DevTest Labs
 VS Application Insights	 API Management
 HockeyApp	 Developer Tools
 Service Profiler*	

Azure Locations



Public Cloud Comparison

Category	Service	 Amazon web services™	 Azure	 Google Cloud Platform	 IBM Cloud	 ORACLE®	 Alibaba Cloud
Compute	Virtual Server	 Amazon EC2 🔗	 Azure Virtual Machine 🔗	 Compute Engine 🔗	 Virtual Server 🔗	 Compute 🔗	 Alibaba ECS 🔗
Compute	Bare Metal Server	 None	 None	 None	 Bare Metal Servers 🔗	 Compute-Bare Metal Cloud Service 🔗	 Bare Metal Servers 🔗
Compute	Container Registration Service	 Amazon ECR Container Registry 🔗	 Azure Container Registry 🔗	 Container Registry 🔗	 IBM Cloud Container Registry 🔗	 Oracle Container Registry 🔗	 Alibaba Container Registry Archives 🔗
Compute	Container Management Service	 Amazon ECR Container Service 🔗	 Azure Container Service 🔗	 Container Engine 🔗	 IBM Cloud Container Service 🔗	 Container 🔗	 Alibaba Bluemix Container Service 🔗
Compute	Micro-Service App Development Platform	 AWS Lambda 🔗	 Azure Service Fabric 🔗	 Google Cloud Functions 🔗	 IBM Cloud Functions 🔗	 Oracle Functions 🔗	 Coming soon 🔗
Compute	Virtual Private Servers	 Amazon Lightail 🔗	 Azure App Service Environment 🔗	 Virtual servers 🔗	 Virtual servers 🔗	 Virtual servers 🔗	 Coming soon 🔗
Compute	Batch Jobs	 AWS Batch 🔗	 Azure Batch 🔗	 Preemptible VMs 🔗	 Batch Jobs 🔗	 Batch Jobs 🔗	 Batch Jobs 🔗
Compute	App Development/Deployment (.Net/C#/.NET Core/Python)	 AWS Elastic Beanstalk 🔗	 Azure Web Apps 🔗	 Google App-engine 🔗	 Bluemix Cloud Foundry app runtimes 🔗	 Application Container Cloud 🔗	 Application Container Cloud 🔗
Compute	Event Driven Computing	 AWS Lambda 🔗	 Azure Functions 🔗	 Cloud Functions 🔗	 IBM Cloud Functions 🔗	 Cloud Functions 🔗	 Cloud Functions 🔗
Storage	Object Storage	 Amazon Simple Storage Service (S3) 🔗	 Azure Blob Storage 🔗	 Cloud Storage 🔗	 Object Storage 🔗	 Object Storage 🔗	 Object Storage Service 🔗