Cloud Computing

Edi Surya Negara

Modified from Mark Baker

What is Cloud Computing?

- **Cloud Computing** is a general term used to describe a new class of network based computing that takes place over the Internet,
  - basically a step on from Utility Computing
  - a collection/group of integrated and networked hardware, software and Internet infrastructure (called a platform).
  - Using the Internet for communication and transport provides hardware, software and networking services to clients
- 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?

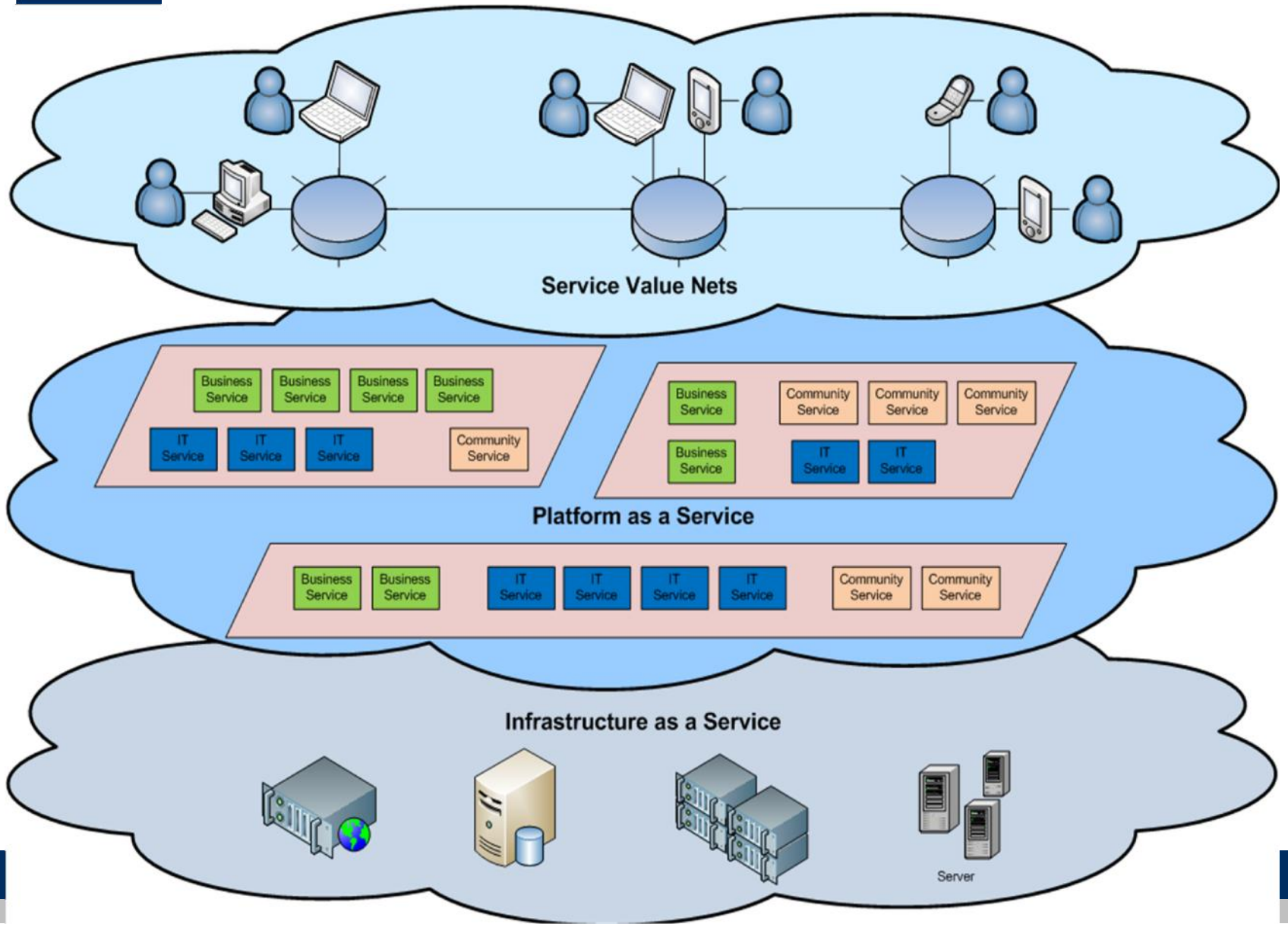
- In addition, the platform provides on demand services, that are always on, anywhere, anytime and any place.
- Pay for use and as needed, elastic
  - scale up and down in capacity and functionalities
- The hardware and software services are available to
  - general public, enterprises, corporations and businesses markets

Cloud Summary

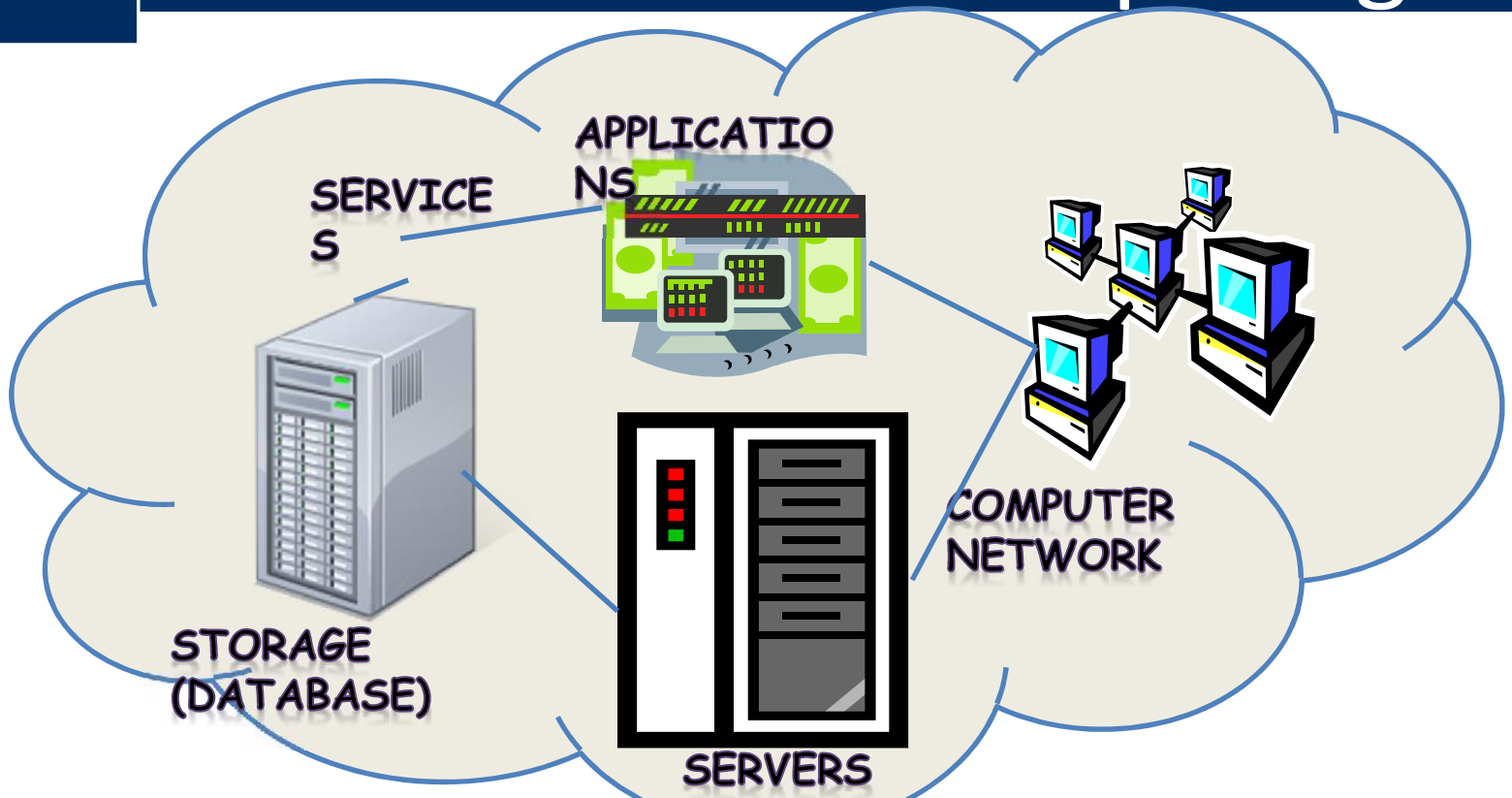
- Cloud computing is an umbrella term used to refer to Internet based development and services

- 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 Architecture



What is Cloud Computing



- Shared pool of configurable computing resources
- On-demand network access
- Provisioned by the Service Provider

Cloud Computing Characteristics

Common Characteristics:

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

Essential Characteristics:

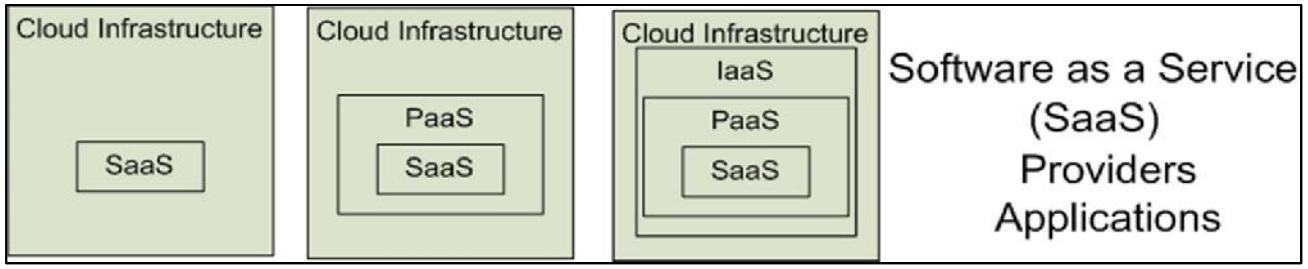
- On Demand Self-Service
- Broad Network Access
- Rapid Elasticity
- Resource Pooling
- Measured Service

Cloud Service Models

Software as a Service (SaaS) | Platform as a Service (PaaS) | Infrastructure as a Service (IaaS)

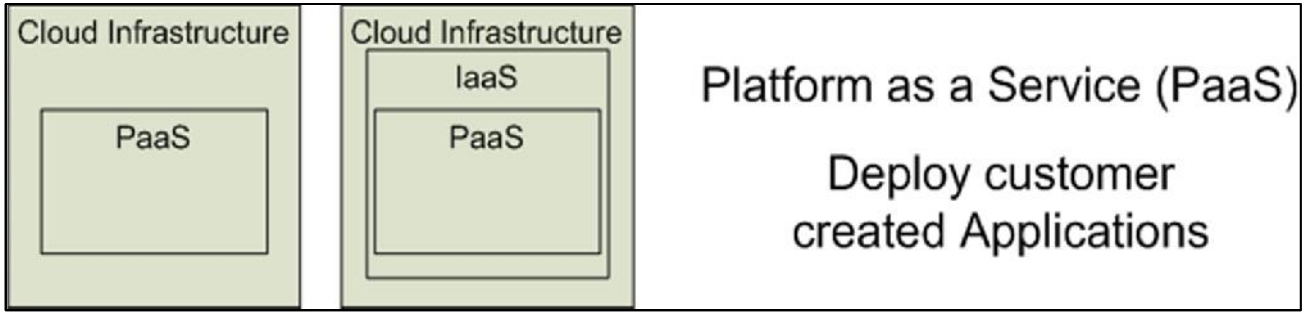
SalesForce CRM

LotusLive



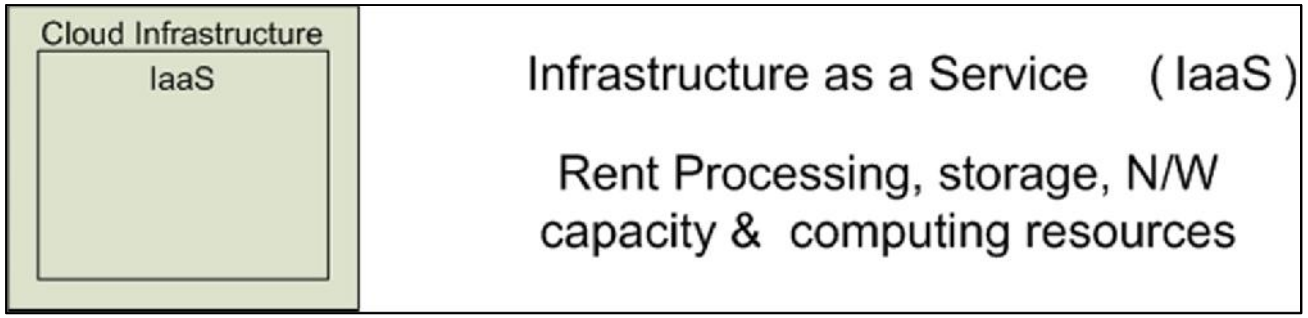
Google App

Windows Azure
The Future Made Familiar



amazon web services™

rackspace HOSTING



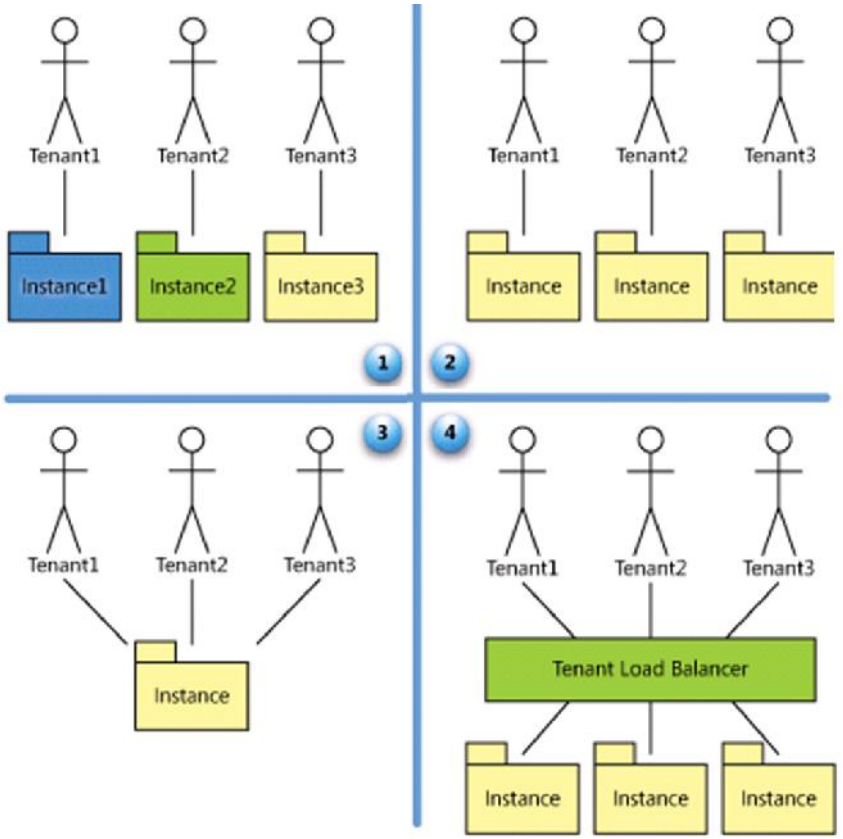
SaaS Maturity Model

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



Source: Frederick Chong and Gianpaolo Carraro, "Architectures Strategies for Catching the Long Tail"

Different Cloud Computing Layers

Application Service (SaaS)	MS Live/ExchangeLabs, IBM, Google Apps; Salesforce.com Quicken Online, Zoho, Cisco
Application Platform	Google App Engine, Mosso, Force.com, Engine Yard, Facebook, Heroku, AWS
Server Platform	3Tera, EC2, SliceHost, GoGrid, RightScale, Linode
Storage Platform	Amazon S3, Dell, Apple, …

Cloud Computing Service Layers

	Services	Description
Application Focused	Services	Services – Complete business services such as PayPal, OpenID, OAuth, Google Maps, Alexa
	Application	Application – Cloud based software that eliminates the need for local installation such as Google Apps, Microsoft Online
	Development	Development – Software development platforms used to build custom cloud based applications (PAAS & SAAS) such as SalesForce
Infrastructure Focused	Platform	Platform – Cloud based platforms, typically provided using virtualization, such as Amazon ECC, Sun Grid
	Storage	Storage – Data storage or cloud based NAS such as CTERA, iDisk, CloudNAS
	Hosting	Hosting – Physical data centers such as those run by IBM, HP, NaviSite, etc.

Basic Cloud Characteristics

- The "**no-need-to-know**" in terms of the underlying details of infrastructure, applications interface with the infrastructure via the APIs.
- The "**flexibility and elasticity**" allows these systems to scale up and down at will
  - utilising the resources of all kinds
    - CPU, storage, server capacity, load balancing, and databases
- The "**pay as much as used and needed**" type of utility computing and the "**always on!, anywhere and any place**" type of network-based computing.

Basic Cloud Characteristics

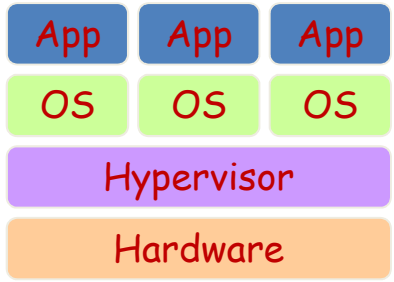
- Cloud are transparent to users and applications, they can be built in multiple ways
  - branded products, proprietary open source, hardware or software, or just off-the-shelf PCs.
- In general, they are built on clusters of PC servers and off-the-shelf components plus Open Source software combined with in-house applications and/or system software.

Software as a Service (SaaS)

- SaaS is a model of software deployment where an application is hosted as a service provided to customers across the Internet.
- Saas alleviates the burden of software maintenance/support
  - but users relinquish control over software versions and requirements.
- Terms that are used in this sphere include
  - **Platform as a Servic**e (PaaS) and
  - **Infrastructure as a Service** (IaaS)

Virtualization

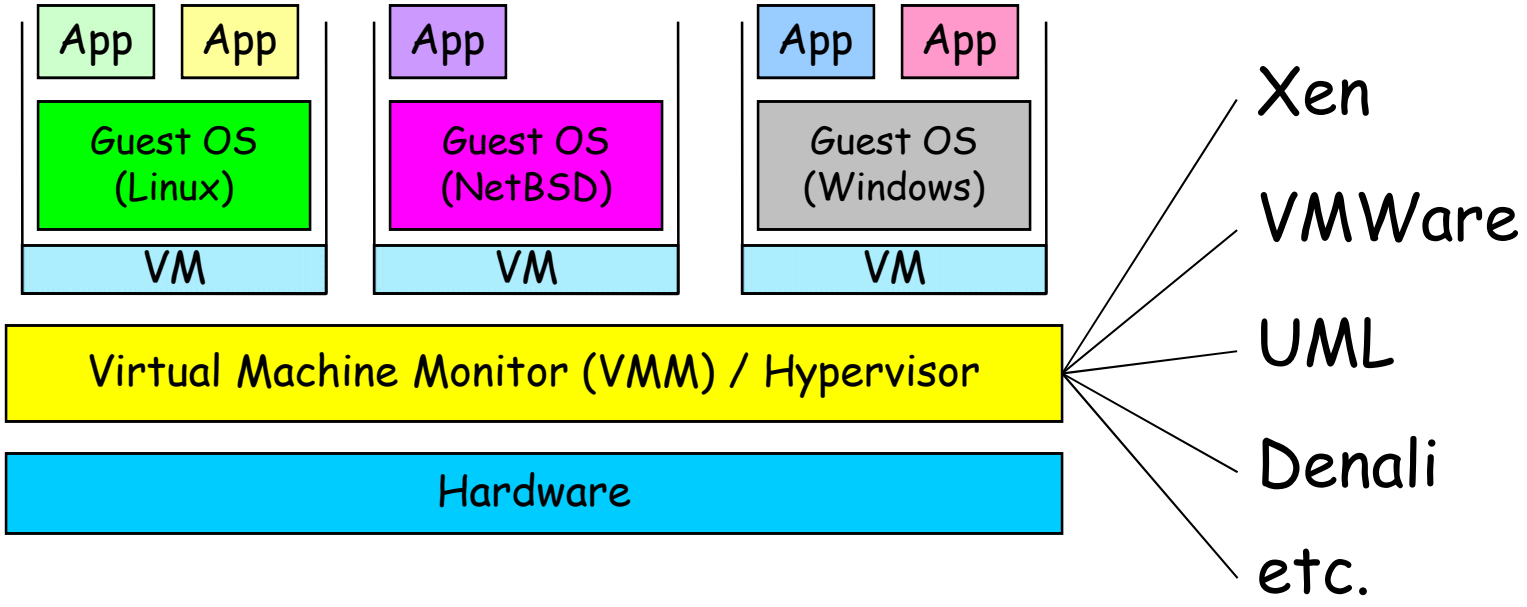
- Virtual workspaces:
  - An abstraction of an execution environment that can be made dynamically available to authorized clients by using well-defined protocols,
  - Resource quota (e.g. CPU, memory share),
  - Software configuration (e.g. O/S, provided services).
- Implement on Virtual Machines (VMs):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



Virtualized Stack

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!

Virtualization in General

- Advantages of virtual machines:
  - Run operating systems where the physical hardware is unavailable,
  - Easier to create new machines, backup machines, etc.,
  - Software testing using “clean” installs of operating systems and software,
  - Emulate more machines than are physically available,
  - Timeshare lightly loaded systems on one host,
  - Debug problems (suspend and resume the problem machine),
  - Easy migration of virtual machines (shutdown needed or not).
  - Run legacy systems!

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.

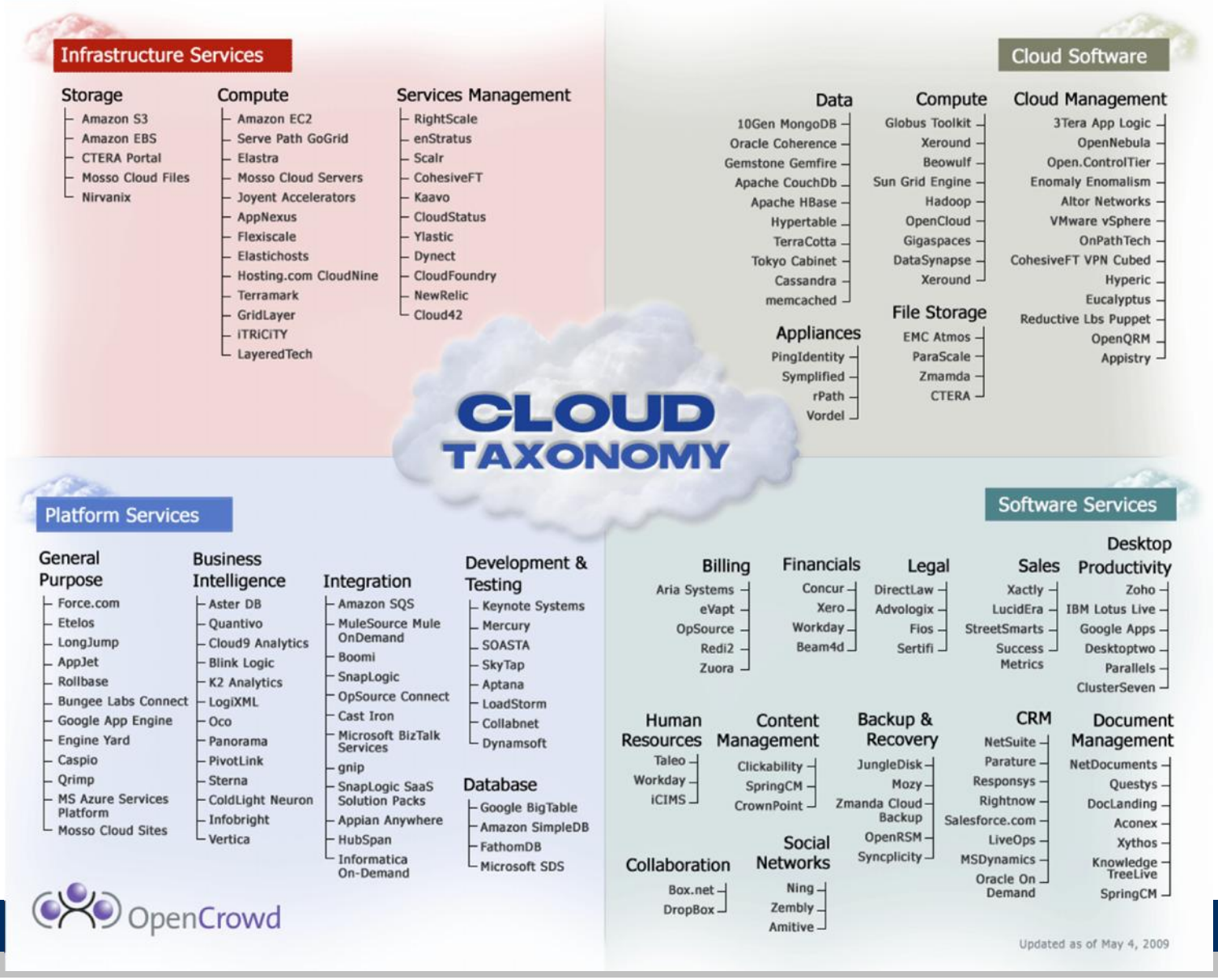
Cloud-Sourcing

- Why is it becoming a Big Deal:
  - Using high-scale/low-cost providers,
  - Any time/place access via web browser,
  - Rapid scalability; incremental cost and load sharing,
  - Can forget need to focus on local IT.
- Concerns:
  - Performance, reliability, and SLAs,
  - Control of data, and service parameters,
  - Application features and choices,
  - Interaction between Cloud providers,
  - No standard API – mix of SOAP and REST!
  - Privacy, security, compliance, trust…

Some Commercial Cloud Offerings

Amazon Elastic Compute Cloud (Amazon EC2) - *Beta*

Cloud Taxonomy



Cloud Storage

- Several large Web companies are now exploiting the fact that they have data storage capacity that can be hired out to others.
  - allows data stored remotely to be temporarily cached on desktop computers, mobile phones or other Internet-linked devices.

- Amazon's Elastic Compute Cloud (EC2) and Simple Storage Solution (S3) are well known examples
  - Mechanical Turk

Amazon Simple Storage Service (S3)

- Unlimited Storage.
- Pay for what you use:
  - $0.20 per GByte of data transferred,
  - $0.15 per GByte-Month for storage used,
  - Second Life Update:
    - 1TBytes, 40,000 downloads in 24 hours - $200,

Utility Computing – EC2

- Amazon Elastic Compute Cloud (EC2):
  - Elastic, marshal 1 to 100+ PCs via WS,
  - Machine Specs…,
  - Fairly cheap!
- Powered by Xen – a Virtual Machine:
  - Different from Vmware and VPC as uses “para-virtualization” where the guest OS is modified to use special hyper-calls:
  - Hardware contributions by Intel (VT-x/Vanderpool) and AMD (AMD-V).
  - Supports “Live Migration” of a virtual machine between hosts.
- Linux, Windows, OpenSolaris
- Management Console/AP

EC2 – The Basics

- Load your image onto S3 and register it.
- Boot your image from the Web Service.
- Open up required ports for your image.
- Connect to your image through SSH.
- Execute you application…

Opportunities and Challenges

- The use of the cloud provides a number of opportunities:
  - 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”.

Opportunities and Challenges

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

Advantages of Cloud Computing

- Lower computer costs:
  - You do not need a high-powered and high-priced computer to run cloud computing's web-based applications.
  - Since applications run in the cloud, not on the desktop PC, your desktop PC does not need the 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 efficient processor...
  - In fact, your PC in this scenario does not even need a CD or DVD drive, as no software programs have to be loaded and no document files need to be saved.

Advantages of Cloud Computing

- 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…
- 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

- Instant software updates:
  - Another advantage to cloud computing is that 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.

- 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

- 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.
- 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

- 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
- 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

- 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
- 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.

Disadvantages of Cloud Computing

- 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 and in areas where Internet connections are few or inherently unreliable, this could be a deal-breaker.

Disadvantages of Cloud Computing

- 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.
- 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

- 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 you might expect from desktop applications.

Disadvantages of Cloud Computing

- 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 unauthorised users gain access to your confidential data?
- 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

- HPC Systems:
  - 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.

The Future

- Many of the activities loosely grouped together under cloud computing have already been happening and centralised computing activity is not a new phenomena
- Grid Computing was the last research-led centralised approach
- However there are concerns that the mainstream adoption of cloud computing could cause many problems for users
- Many new open source systems appearing that you can install and run on your local cluster
  - should be able to run a variety of applications on these systems

Tentukan masalah dan solusi dari video "Why internet slows down when it's busy – Computerphile"

Yudy Pranata (192420001)
Mahasiswa Magister Teknik Informatika Universitas Bina Darma
Jl. Jend. A. Yani No. 12, Plaju, Palembang 30264
yudypranata26@gmail.com

Pada video "*Why internet slows down when it's busy – Computerphile*" yang telah diberikan, terdapat penjelasan yang menyebutkan bahwa penyedia jasa jaringan internet selalu menjanjikan kapasitas besar dan cepat dalam jaringan untuk menarik banyak peminat. Seperti didalam video dengan contoh penyedia jasa jaringan internet memiliki total kapasitas jaringan server sebesar 1 (satu) Gb/sec dan mereka memberikan "janji" kecepatan dan kapasitas jaringan internet kepada pengguna (konsumen) 100 Mb/sec. Tetapi yang sebenarnya mereka tidak bisa menjanjikan bahwa kapasitas tersebut dengan penuh dapat digunakan, karena pada saat menggunakan internet kapasitas jaringan yang diberikan tetap mengacu pada jumlah pengguna (konsumen) dan seberapa besar menggunakan kapasitas jaringan tersebut pada saat yang bersamaan.

Solusi

Berdasarkan pada penelitian *TIME-WAIT loading* [1] dalam percobaan dilakukan pada sebuah *workstation* yang terhubung ke Myrinet LAN sebesar 640 Mb/sec [2]. Dengan adanya 3 (tiga) penelitian yang pertama membuktikan bahwa beban TCB menurunkan kinerja server dan modifikasi yang digunakan akan menurunkan degradasi tersebut. Penelitian kedua menggambarkan dari kedua solusi TCP [3] dan HTTP [4] dapat meningkatkan kinerja server dibawah tolak ukur WebSTONE [5]. Dan penelitian yang terakhir menggambarkan bahwa modifikasi yang dibuat memungkinkan server untuk mendukung beban HTTP yang tidak dapat digunakan pada konfigurasi standar. Dapat dijelaskana bagaimana TCP berinteraksi dengan protokol aplikasi tertentu untuk memuat server dengan *TIME-WAIT* TCB dan menunjukkan dari tindakan ini. Interaksi tersebut adalah hasil langsung dari penyederhanaan asumsi bahwa titik akhir dalam penutupan koneksi tidak bergantung untuk selanjutnya membangun kembali koneksi tersebut. Dengan mangajukan usulan pada titik akhir yang dapat menahan *TIME-WAIT* selama pembentukan kembali koneksi. System tersebut akan memperluan fungsi TCP untuk mengalokasikan *TIME-WAIT* TCB ke ujung koneksi yang tepat tanpa mengganggu protokol aplikasi atau meninggalkan kelemahan pada TCP. Pada tahap implementasi dan menguji solusi pada TCP yang lebih sederhana dan solusi pada HTTP untuk menunjukkan beban

TIME-WAIT dari klien ke server dan dalam kondisi tersebut dapat dipastikan bahwa beban *TIME-WAIT* dapat mempengaruhi kinerja server dibawah SunOS 4.1.3 yang hasilnya dpaat berkurang sebanyak 50 %. Sementara itu, dengan menggunakan WebSTONE dapat dibuktikan bahwa HTTP pada klien dan server yang menggunakan system mereka memiliki hasil yang lebih besar dibandingkan dengan SunOS 4.1.3. Klien dan server yang telah menggunakan system konfigurasi mereka dapat mendukung tingkat koneksi yang lebih tinggi dari pada system yang belum dimodifikasi pada konfigurasi tertentu. Mereka juga telah menunjukkan bahwa sistemnya dapat dikombinasikan dengan koneksi HTTP yang persisten menggunakan lebih sedikit memori daripada system SunOS 4.1.3 yang tidak dimodifikasi menggunakan koneksi untuk beban klien tertentu.

	TCP with TIME-WAIT Negotation	TCP with Clien <RST>	CLIENT_CLOSE HTTP Extension
Reduces TIME-WAIT loading	Yes	Yes	Yes
Compatible with Current Protocols	Yes	Yes	Yes
Changes are Effective if only the client is modified	No	Yes	Yes
Allows system to prevent TIME-WAIT assassination	Yes	No	Yes
No changes to Transport Protocol	No	No	Yes
No changes to Application Protocols	Yes	Yes	No
Adds no packet exchanges to Modified Protocol	Yes	No	No
TIME-WAIT allocation is a requirement of connection establishment	Yes	No	No

Table 4: Summary of Proposed Systems [1]

Referensi

1. Faber Theodore, Joseph D. Touch, and W. Yue. The TIME-WAIT state in TCP and its effect on busy servers. In *Proceedings of IEEE Infocom* (March 1999).
2. Myricom, Inc., Nannette J. Boden, Danny Cohen, Robert E. Felderman, Alan E Kulawik, Charles L. Seitz, Jakov N. Selovic, and Wen-King Su, “Myrinet: A Gigabit-per-second Local Area Network.” *IEEE Micro,* pp.29-36, IEEE (February 1995).
3. Jon Postel, ed., “Transmission Control Protocol,” *RFC-793/STD-7* (September, 1981).
4. R. Fielding, J. Gettys, J. Mogul, H. Frystyk, and T. Berners-Lee, “Hypertext Transport Protocol – HTTP/1.1,” *RFC-2068* (January, 1997).
5. Gene Trent and Mark Sake, “WebSTONE: The First Generation in HTTP Server Benchmarking,” *white paper*, Silicon Graphics International (February, 1995).

TUGAS

Nama mahasiswa:

1. M DANIAL SENTOSA

Permasalahan : ketika user menggunakan internet diwilayah tertentu, kecepatan internet menurun, tidak seperti yang dijanjikan para provider, dikarenakan setiap provider membatasi pengguna pada suatu daerah, misal suatu daerah dibatasi 1 GB/s dan penggunanya ada 10 orang, maka 1 GB/s akan dibagikan kepada 10 orang, maka setiap orang mendapatkan kecepatan 100 Mb/s sehingga semakin banyak pengguna maka internet akan semakin lambat bahkan terjadi packet loss

Solusi : pada jurnal dengan judul *The TIME−WAIT state in TCP and Its Effect on Busy Servers* oleh Theodore Faber menjelaskan dengan menggunakan sistem TIME−WAIT

Jurnal ini telah menerapkan dan menguji solusi TCP yang lebih sederhana dan solusi HTTP untuk menunjukkan pergeseran waktu TIME-WAIT dari klien ke server. Kami telah menyajikan bukti eksperimental bahwa loading TIME-WAIT dapat mempengaruhi kinerja server di bawah SunOS 4.1.3. Dalam kondisi ini, throughput dapat dikurangi hingga 50%.

Dengan menggunakan system ini dan di kombinasikan dengan koneksi TCP persistent, hal ini menggunakan lebih sedikit memori dari pada Sun OS 4.1.3 yang tidak dimodifikasi menggunakan koneksi pertant, untuk beban klien tertentu. System yang digunakan berkoneksi terus-menerus, masalah pemuatan memori secara langsung. Ini dapat mengurangi biaya penyebaran server, yang dapat menjadi sangat penting server tertanam/baterai

Menurut theodore faber, pendistribusian TIME-WAIT TCBS ke client dengan memanfaatkan angka pertumbuhan mereka untuk mengurangi memory load pada server. Kelebihannya yaitu mendukung koneksi yang persistant dan berbagi manfaat lainnya seperti menghindari extra TCP 3 handshakes.

Pada jurnal ini percaya bahwa TCP pada akhirnya harus dimodifikasi untuk mendukung TIME−WAIT negotiation. Upgrade yang akan datang dari IP versi 4 ke versi 6 juga merupakan peluang untuk meninjau kembali implementasi dan desain TCP. Ini akan menjadi waktu yang tepat untuk memasukkan TIME−WAIT state negotiation.

Algoritma TIME-WAIT state negotiation yaitu mengizinkan busy server untuk menerima koneksi dari klien yang memilik TIME-WAIT oerhead. Penggunaan algoritma ini tidak harus mengubah penggunaan koneksi dekat dengan protokol mereka dan tidak mengalami perubahan performa yang terkait dengan pendistribusian TIME-WAIT TCBs.

Selain memodifikasi TCP, menurut thedore faber, memodifikasi HTTP juga diperlukan karena itu merupakan salah satu komponen terbesar pada internet traffic. Untuk mensupport koneksi yang persistent, indikasi end-of-connection dan end of transaction harus dipisahkan. Pemisahan ini mengizinkan kita untuk memodifikasi HTTP untuk mengizinkan klien mengaktifkan kembali close connections dan menahan TIME-WAIT state.