

Cloud Computing and Relevant Technique

Yuyang Zhang

What is clouding computing

What is cloud computing

First we have Internet.

Then broadband became very cheap. So we can run applications on remote machines.

This leads “software as a service”. Someone else owns the application and fast machines, users pay fixed subscription fee to access services using fast network connection and web browser.

What is cloud computing

Even some of them are for free.



What is cloud computing

Pos of “software as a service”:

Users don't have to worry about software and hardware update and maintenance.

Cons:

Services are the same to each user.

Not flexible enough for complex requirements.

Requirements

Flexibility in machine capacity.

Convenience and simplicity.

What is cloud computing

To solve this problem, we have visualization technique.

Pack application, database, middle-ware and OS. And run it in a large scale server cluster, or called

CLOUD

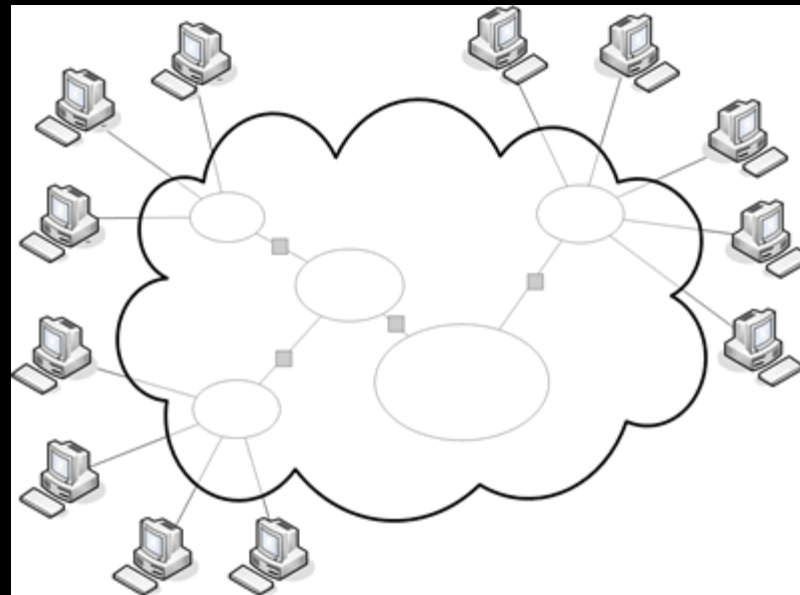
This approach is called

CLOUD COMPUTING

What is cloud computing

The term cloud computing probably comes from (at least partly) the use of a cloud image to represent the Internet or some large networked environment.

“We all know a network looks like this” -- Jan



What is cloud computing

Definition from Gartner:

A style of computing where massively scalable IT-related capabilities are provided 'as a service' across the Internet to multiple external customers using Internet technologies.

Definition from me:

Cloud computing is kind of rent service based on large scale server clusters and Internet. Customers is able to acquire this service any place any time. And the rent is very cheap. So users are freed from labors on maintaining software, OS, low-level utilities and hardware.

What is cloud computing

It is not cloud computing if...

You have to give service provider a 50 pages description of your requirement.

Cannot buy it with your credit card

They are selling you hardwares

You have to rebuild your system infrastructure

Cannot do provision and deprovision in 10 min

Know exactly where the server locates

you have those hardwares

What is cloud computing

Examples:

Amazon EC2

AMD Fusion Render Cloud

What is cloud computing

Amazon EC2

Amazon EC2's **simple web service interface** allows you to **obtain** and **configure** capacity with minimal friction. It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment. Amazon EC2 reduces the time required to obtain and boot new server instances to minutes, allowing you to **quickly scale capacity, both up and down**, as your computing requirements change. Amazon EC2 changes the economics of computing by allowing you to pay only for capacity that you actually use. Amazon EC2 provides developers the tools to build failure resilient applications and isolate themselves from common failure scenarios.

What is cloud computing

United States		Europe
Standard On-Demand Instances		
	Linux/UNIX Usage	Windows Usage
Small (Default)	\$0.10 per hour	\$0.125 per hour
Large	\$0.40 per hour	\$0.50 per hour
Extra Large	\$0.80 per hour	\$1.00 per hour
High CPU On-Demand Instances		
	Linux/UNIX Usage	Windows Usage
Medium	\$0.20 per hour	\$0.30 per hour
Extra Large	\$0.80 per hour	\$1.20 per hour

United States			
Linux/UNIX	One-time Fee		
Standard Reserved Instances	1 yr Term	3 yr Term	Usage
Small (Default)	\$325	\$500	\$0.03 per hour
Large	\$1300	\$2000	\$0.12 per hour
Extra Large	\$2600	\$4000	\$0.24 per hour
High CPU Reserved Instances	1 yr Term	3 yr Term	Usage
Medium	\$650	\$1000	\$0.06 per hour
Extra Large	\$2600	\$4000	\$0.24 per hour

What is cloud computing

AMD Fusion Render Cloud

A super computer with ultimate graphic process capability “to make HD cloud computing a reality”.

Deliver video games, PC applications and other graphically-intensive applications through the Internet “cloud” to mobile devices with a web browser without making the device rapidly deplete battery life or struggle to process the content.

What is cloud computing

AMD Fusion Render Cloud



Relative Techniques on Cloud Computing

Relative Techniques

Virtualization

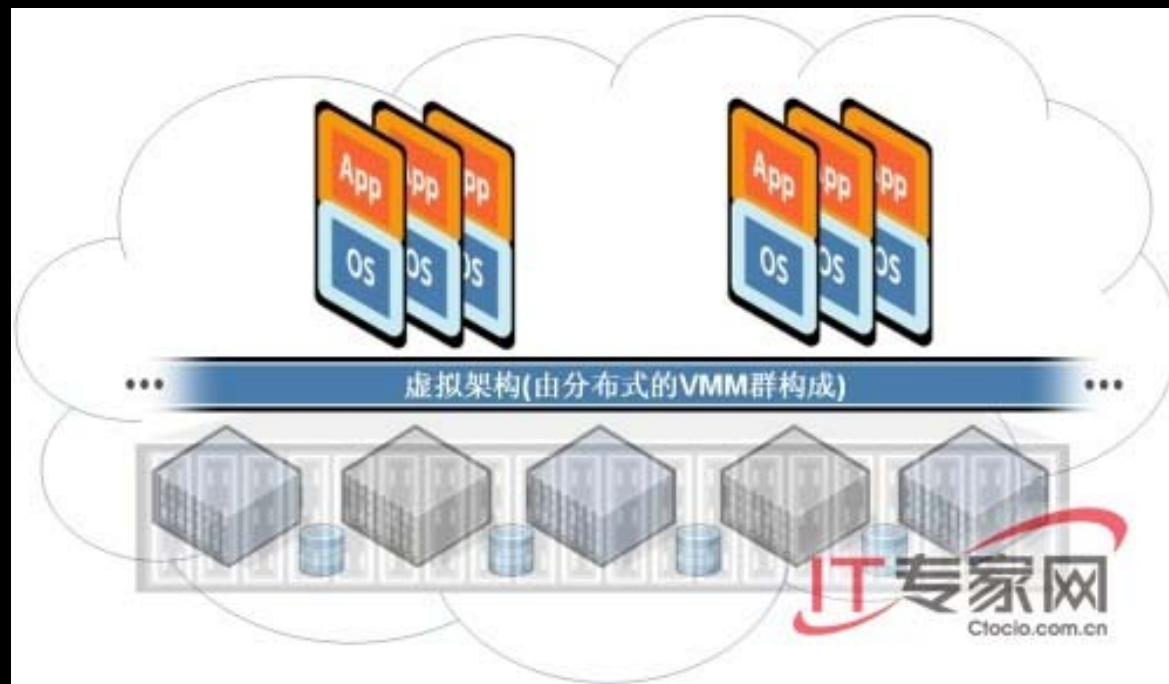
Rapid Real-time Allocation of Resources

Distributed computing

Service Migration

Virtualization

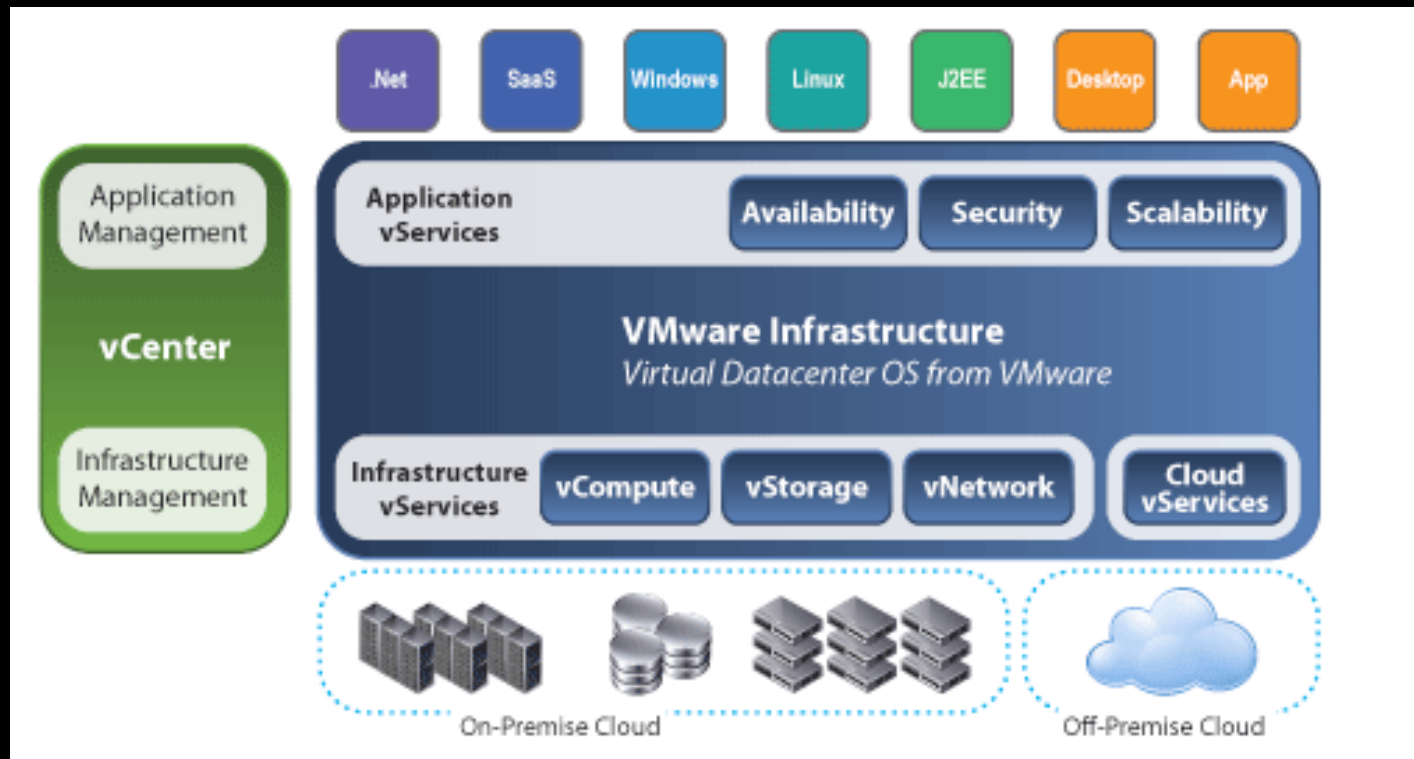
Virtualization for cloud computing is not for single server, but on a infrastructure level.



Virtualization

VMware Infrastructure 3, a virtual datacenter operating system.

Together goes its own cloud OS vSphere4.



Rapid real-time resource allocation

Work-conserving

Non work-conserving

eg. 2 virtual servers on 1 physical server. Each possess 50% resource. One virtual server is down.

For work-conserving, the remaining server takes 100% resource.

For non work-conserving, the remaining server still takes 50% resource.

Rapid real-time resource allocation

For CPU resource allocation, 3 algorithms are available.

- Borrowed virtual time

- Simple earliest deadline first

- Credit scheduler

Details presented in “*When Virtual is Harder than Real: Resource Allocation Challenges in Virtual Machine Based IT Environments*”

Distributed computing

Google's secret: MapReduce

A process to use a large number of computers solve distributable problems.

Map: take input, split it into smaller sub-problems, and distributes those to worker nodes.

Reduce: takes answers of sub-problems from worker nodes, combine them to get the output.

Distributed computing

A code example.

```
map(String name, String document):
```

```
    // key: document name
```

```
    // value: document contents
```

```
    for each word w in document:
```

```
        EmitIntermediate(w, 1);
```

```
reduce(String word, Iterator partialCounts):
```

```
    // key: a word
```

```
    // values: a list of aggregated partial counts
```

```
    int result = 0;
```

```
    for each v in partialCounts:
```

```
        result += ParseInt(v);
```

```
    Emit(result);
```

Live Service Migration

To move a running virtual machine from one physical server to another one.

User won't feel the migration. So service vendor can update hardware and software without interrupting users.

Latest live service Migration supports migration among machines running different OS.

Visions on Cloud Computing

Visions on cloud efficiency

Develop feature specializing servers. Cooperate them together

Xbox360 performs as well as top PC on game playing, but it costs much less.

Servers can also be feature specialized to increase performance and cut off costs.

Further virtualization approach have to be developed to aggregate different kind of feature specialized servers efficiently.

Visions on cloud efficiency

Efficient data and resource sharing among virtual servers.

Eg. Some virtual servers use same software such as Oracle 11g. There must be a better way than installing a separate copy on each virtual machine.

Visions for developer

JVM(java virtual machine) can help to run java code on multiple OS and machines with various capability.

Such a virtual machine might be useful, taking general computing requirement and allocate tasks to appropriate devices.

Reference

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

<http://www.kernelthread.com/publications/virtualization/>

<http://www.china-cloud.com/cloudcomputing.jsp?ID=121>

<http://www.vmware.com/products/vi/>

http://cio.ccw.com.cn/research/info/htm2008/20080911_500298.shtml

<http://aws.amazon.com/ec2/>

<http://www.techpluto.com/cloud-computing-characteristics/>

http://www.amd.com/us-en/Corporate/VirtualPressRoom/0,,51_104_543~129743,00.html

http://findarticles.com/p/articles/mi_m0BRZ/is_2_25/ai_n13798056

<http://www.parallels.com/products/virtuozzo/os/>

<http://en.wikipedia.org/wiki/Virtualization>

Brian Hayes, "cloud computing", Communications of the ACM, No7, 2008.07

White Paper - Top Ten Considerations For Choosing A Server Virtualization Technology, Version 1.0, www.parallels.com

First Steps Toward the Data Center of the Future - A META Group White Paper

Reference

Youtube – cloud computing in plain english

Ludmila Cherkasova, Diwaker Gupta¹, Amin Vahdat, *“When Virtual is Harder than Real: Resource Allocation Challenges in Virtual Machine Based IT Environments”*, HP LAB, HP

Jeffrey Dean and Sanjay Ghemawat, *“MapReduce: Simplified Data Processing on Large Clusters”*, google, Inc

NVIDIA_CUDA_Programming_Guide_2.0