Introduction to Cloud computing

Viet Tran

Type of Cloud computing

- **Infrastructure as a Service IaaS: offer full virtual machines via hardware virtualization tech.**
  - Amazon EC2, AbiCloud, ElasticHosts,…

- **Platform as a Service PaaS: offer platforms (e.g. API) where users can deploy services**
  - Amazon SQS, Microsoft Azure, Google App Engine, LAMP
  - Can be realized as IaaS+Platform (API)
  - But can be also realized without hardware virtualization (e.g. separation via user account or sandboxing)

- **Software as a Service: offer cloud/web applications**
  - Amazon RDS, Google Docs, …

Type of Cloud computing

- Public clouds
  - Sell/buy services
  - Amazon, AbiCloud, ElasticHosts, …
  - Limited customization, complex user management

- Private clouds
  - Internal use inside organization
  - VMWare VSphere, IBM Websphere cloud, Sun cloud
  - More like advanced or automatized virtual infrastructure
  - Wide customization, simple user management

Infrastructure as a Service

- IaaS = Virtualization + Standardization + Automation

- Can be considered as automation tool over virtualization software (Xen, KVM, VMWare)

- Nearly every large software/hardware vendor offer some middleware for IaaS (for public or private cloud)
  - IBM Websphere cloud, VMWare Vsphere, Sun Cloud, ….

- Opensource cloud middleware: Eucalyptus, OpenNebula, Nimbus
  - All support Amazon EC2 client beside native client

Amazon EC2

- Leader of IaaS, practically standard
- Running on XEN/Linux
- Using EC2:
  - Register, get certificate, install EC2 client
  - Creating own image if needed and upload
  - Run VM instances
  - Use VM instances

Creating images of VM

- Amazon offer wide range of ready-to-use images (major Linux distributions + basic software)
- They are XEN images with some strict configuration (user name, partitions)
- Users should download the provided images and install their own software if needed
  - Build images from scratch is not recommended because lack of documentations
  - Just simply download Xen images from Amazon, mount as loopback disk on local machine and copy software (or run the images in Xen)
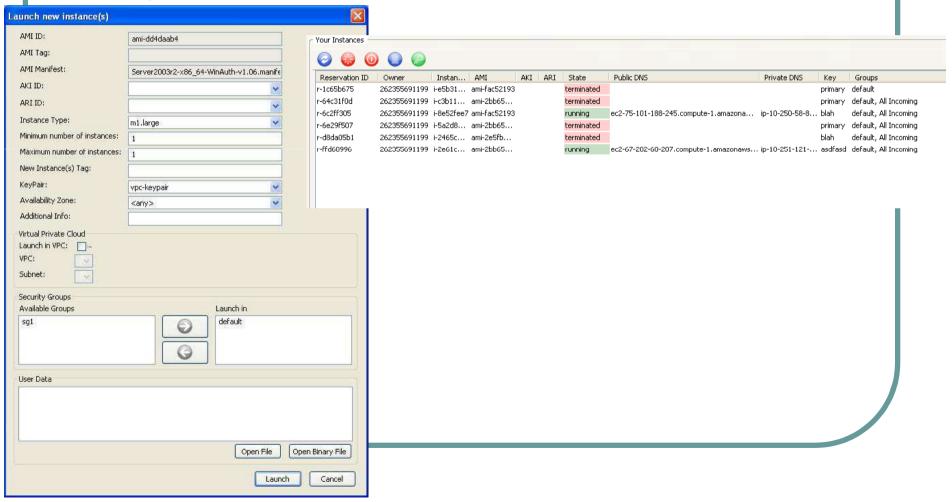
Upload and register images

- ec2-upload-bundle images
- ec2-register
- There are usually 3 files for a VM: kernel, ramdisk and disk images
- The images are stored in Amazon S3 service (splited to many small files), and accessible via ID (returned from rec2-register)

Run instances

- ec2-run-instances imageID -k mykey -t c1.medium
  - mykey: SSH key for login (login via password is not allowed)
  - c1.medium: type of virtual machine (CPU, RAM, …)
  - Return instance ID

- ec2-describe-instances instanceID
  - Return status of instance and IP/hostname
  - Just use SSH with private key to login to the creatd instance

Graphical interface: ElasticFox

- Plugin to Firefox

Notes

- EC2 does not save images after termination of instances
  - All modification (i.e. data created or software installed during running) will be lost if not save elsewhere
  - Use S3 or EBS to save data permanently

- Each instance has public (assigned) and private (real) IP
  - User connect to the instance via public IP
  - SNAT/DNAT used for translation to private IP
  - Can be integrate with company private network via VPN

Open source cloud middleware

- Eucalyptus
- OpenNebula
- Nimbus

- All support EC2 client
- OpenNebula and Nimbus have their own native clients

Eucalyptus

- Most close to EC2
- Support XEN and KVM
- Support elastic IP
- Excellent interface for user management and system configuration
- Very limited possibility for customization
- Excellent user docs, nearly no developer docs

=> Suitable for building public cloud

OpenNebula

- More close to virtual infrastructure
- Support XEN, KVM, VMWare (VirtualBox planned)
- No elastic IP, clumsy IP assignment
- Wide range of customization, e.g. store images after termination, easily to add new function
- Good develop. docs

=> Suitable for build private clouds

Nimbus

- Interface via Globus Toolkit 4
- Difficult to install (need to know about GT4), confusing docs
- Support cluster deployment at single command
- No elastic IP

=> Suitable for build on or integrate with grid infrastructure

Advanced cloud services at AWS

- Compute
  - **Amazon Elastic Compute Cloud** (EC2)
  - **Amazon Elastic MapReduce**
  - **Auto Scaling**
- Content Delivery
  - **Amazon CloudFront**
- Database
  - **Amazon SimpleDB**
  - **Amazon Relational Database Service**(RDS)
- E-Commerce
  - **Amazon Fulfillment Web Service** (FWS)
- Messaging
  - **Amazon Simple Queue Service** (SQS)

EC2 is only a small part of Cloud computing

- Monitoring
  - **Amazon CloudWatch**
- Networking
  - **Amazon Virtual Private Cloud** (VPC)
  - **Elastic Load Balancing**
- Payments & Billing
  - **Amazon Flexible Payments Service**(FPS)
  - **Amazon DevPay**
- Storage
  - **Amazon Simple Storage Service** (S3)
  - **Amazon Elastic Block Storage** (EBS)
  - **AWS Import/Export**
- Support
  - **AWS Premium Support**
- Web Traffic
  - **Alexa Web Information Service**
  - **Alexa Top Sites**
- Workforce
  - **Amazon Mechanical Turk**