Amazon Web Services

Un survol

PARTIE I

Notions de base

Quatre niveaux de l'infonuagique

Niveau 0 : On-Premise	Serveurs propriétaires
Niveau 1 : Iaas Infrastructure as a Service	AWS fournit machines, OS Client installle logiciel, BD, SDK,web serveur
Niveau 2 : Paas Platform as a Service	AWS fournit web serveur Client télécharge l'application
Niveau 3 : Saas Software as a Service	AWS fournit logiciel Client télécharge les données

Extensibilité (scalability)

- Extensibilité verticale : augmenter la puissance d'un processeur (en mémoire, CPU, etc.).
- Extensibilité horizontale, appelé aussi élasticité : répartir le travail entre plusieurs processeurs.
- Plusieurs services AWS sont «élastiques».

Services AWS

- Région : par pays et région de pays (fuseau horaire), par exemple us-east, us-west, ca-central
- Availability zones (AZ) : par exemple us-east-1, us-west-2, (min = 3, max = 6).
- Centre de données : au moins deux par AZ, distant d'au moins un km.

Services AWS

On choisit une région selon nos besoins :

- Proximité (temps de latence).
- Légal (politique)
- Services offerts par région
- Coût (en $, euros)

- Services globaux : IAM (A&A), Route53 (DNS redirect), Cloud Front (resource cache).
- Les autres sont des services régionaux (pas offerts dans toutes les régions).

Interaction avec AWS

- AWS Management Console : interface web
- AWS CLI (command line interface) : shell
- AWS SDK (software development kit) : programmatiquement

PARTIE II

Infrastructure as a Service

Infrastructure as a service

Authentification et autorisation

- IAM : Idendity and access management

Location de processeurs :

- EC2 : Elastic Compute Cloud (machine virtuelle)

Espace de stockage :

- EBS : Elastic Block Storage (disque dur virtuel)
- EFS : Elastic File System
- RDS : Relational Database Server

Load Balancers, Auto-scale groups

IAM : Gestion de comptes

IAM User : utilisateur

IAM User group : un utilisateur peut appartenir à n groupes

IAM Role : developpement, administration, AQ

(normalement un groupe → un role)

IAM Policy : ensemble de droits à une ressource

Authentification multi-facteurs

(applications MFA : Duo, Google Authentificator..)

Security keys, H/W TOTP tokens

IAM : Gestion de comptes

IAM outils de sécurité

- IAM credential reports : rapports des statuts utilisateurs en format CSV.
- IAM access advisor : permissions par utilisateurs, dernier accès, ..

IAM : Gestion de comptes

IAM Guidelines

- Compte root seulement pour opérations root
- Opérations root : facturation, création de comptes admin
- Least-privilege rule
- Un utilsateur physique par utilisateur AWS
- Politique de mots de passes, MFA

EC2 : Elastic Compute Cloud

Location d'une VM.

Choix de l'OS (Linux, Windows, Mac)

Choix du CPU, RAM (gratuit pour CPU, RAM de base)

EC2 instance type : m5.2xlarge

EC2 instance purchase options : on-demand, pay per second, reserved, saving plans (by $), spot instances

Security Group : accès aux ports (21 FTP, 80 HTTP, 443 HTTPS, ..), IP ranges, inbound/outbound

EC2 Instance Connect : shell interface in web browser

EC2 : Elastic Compute Cloud

AMI : Amazon Machine Image (customization of a EC2 instance) : create 100 Windows 8GB-RAM virtual machine in a batch w/ preloaded s/w (word, etc.)

EBS : Elastic Block Storage

Espace de stockage (disque virtuel)

Gratuit jusqu'à 30 GB

ESB snapshots → copie de sauvegarde

EC2 Instance store → meilleure performance que EBS

EFS : Elastic File System

Network File System, POSIX-based (similar to Linux)

EFS Storage Tiers

- Standard (comme disque dur interne)
- Infrequent Access (moins cher, comme disque externe)
- Archive (comme ruban magnétique)

RDS : Relational Database Server

- Stocké sur ESB (autoscaled)
- Read replicas → copies d'une DB source
  - SELECT permis,
  - INSERT UPDATE interdits
  - utiles pour rapports
  - Multi-AZ (Disaster Recovery)

ElastiCache : in-memory cache for DB

- Redis vs Memcached
- Lazy loading strategy (lazy population, cache-aside)
- Write Thru strategy
- Cache Evictions and Time-to-live (TTL)

ELB : Elastic Load Balancer

Gère le traffic à l'entrée vers 1..n instances EC2

- ELB ALB (Application, niveau 7) : HTTP, HTTPS
- ELB NLB (Network, niveau 4) : TCP, UDP
- ELB GLB (Gateway, niveau 3) : paquets IP
- ELB Sticky sessions : même cible (basé cookies)
- ELB cross-zone balancing (USA, Europe, ..)

ASG : Auto-scaling groups

Ajout (scale out) ou suppression (scale in) d'instances EC2

ASG policies

- Fixe (min capacity=2, desired=4, max=8)
- Dynamique (cible de 70% CPU)
- Scheduled (ouvert de 9AM 5PM)
- Predictive (nouvelle promotion)
- CloudWatch alarms (metrics comme pourcentage CPU)

Route 53 : serveur DNS

DNS : Domain Name System (annuaire d'adresses IP)

Un des rares services globaux de AWS

- DNS records (A, AAAA, CNAME réfère à A ou AAAA, NS name server)
- TLD : Top-level domains: .com .org
- Domain Registrar (GoDaddy, AWS Route 53)
- Route 53 parce que DNS port est le port 53.

VPC : Virtual Private Cloud

Subnet : public ou privé

NACL : Network ACL (access control list)

VPC flow logs : log all the traffic

Three-tier solution architecture :

Utilisateur → ELB (public subnet) → ASG (private subnet) → Data (private subnate)

Advanced infrastructure services

Amazon ECS (Elastic container service) : Docker

Amazon EKS (Elastic Kubernetes service) : Kubernetes

Amazon ECR (Elastic container registry) : Docker images

Amazon Fargate (Serverless container platform)

p.e. EventBridge : événement à chaque heure → exécute une tâche ESC → sauvegarde S3 (serverless)

PARTIE III

Platform as a Service

Platform as a service

Traitement :

- Serveurs web
- Fonctions Lambda (sans serveur web)
- API Gateway (REST API)
- Queuing System (cf ActiveMQ)
- Journalisation (logs)

Fonction de stockage :

- S3 : Simple Storage Service
- DynamoDB : BD NoSql (cf MongoDB)

Amazon Beanstalk

Services : serveur web, téléverse une application web (Java/Spring, python, etc.)

Stratégies de déploiements :

• All-at-one (+ rapide, -temps hors service), traffic splitting, p.e. 10% utilisateurs nouvelles versio (canary test), rolling, rolling w/ additional batches.

Politiques de cycle de vie (élimines anciennes versions)

Amazon Logging

CloudWatch : collect metrics, logs, send events, alarms.

CloudTrail : traces d'applications (haut niveau)

AWS X-Ray traces : pour architecture distribuée (micro-services)

- outils visuels d'analyse de traces;
- nécessite d'insérer des appels dans le code de l'application;
- diagramme pour suivre la requête à travaers les servives utilisés.

Amazon Queueing

AWS SQS : Simple Queue Service

Similaire à Apache ActiveMQ

Envoi de messages dans une queue; abonnement, dead-letter.

Pour applications découplées.

Un des plus anciens services fournis par AWS.

S3 : Simple Storage Service

- Objets (fichier), buckets (répertoire de niveau racine)
- S3 Bucket Policies :
  - CRR cross-region replication, SRR same-region replic.
- S3 Storage Classes :
  - Standard, infrequent-access (backups), glacier instant, flexible, deep.
- S3 Event Notifications :
  - Events for other AWS services (AWS EventBridge)
  - Access logs → dans un autre bucket !
- Encryption : SSE server-side enc. KMS key-management service.

Security

- Encryption en vol :
  - HTTPS (pour la partie publique utilisateur → load balancer)
  - AWS KMS (Key management service).
- Encryption au repos :
  - Côté serveur : géré par Amazon
  - Côté client : géré par l'utilisateur

Conclusion

Infrastructure as a servive :

- Amazon gère les machines, le système d'exploitation.
- Le développeur installe les applications, les SGBDs et en est responsable.
- Avantages : indépendance p/r Amazon.
- Inconvénients : plus de responsabilité.

Platform as a servive :

- En plus des machines et OS, Amazon gère les applications (services).
- Avantages : utilisations de services matures déjà pris en charge par Amazon, on bénéficie des améliorations faits par Amazon.
- Inconvénients : moins standard, coût de sortie plus important.