Haute disponibilité sur AWS : concevoir des infrastructures résilientes



Dans un monde où les interruptions de service peuvent nuire à la productivité, à la réputation
et aux revenus d'une entreprise, la **haute disponibilité** est devenue un élément
essentiel de toute infrastructure informatique moderne.Avec **Amazon Web Services (AWS)**, il est possible de concevoir des architectures
capables de résister aux pannes, de répartir la charge automatiquement et de maintenir les
applications accessibles même lorsqu'un composant rencontre un problème.Pour apprendre à concevoir ce type d'architecture, Doussou Formation propose la formation

AWS Solutions Architect Associate (SAA-C03) – Infrastructure, Serverless et Certification
.

Qu'est-ce que la haute disponibilité ?

La haute disponibilité, ou *High Availability*, consiste à concevoir une infrastructure
capable de continuer à fonctionner même si un serveur, une base de données,

un réseau ou une zone
de disponibilité rencontre une panne.

L’objectif est de :



Les Régions AWS

Une **Région AWS** représente une zone géographique complète dans laquelle AWS déploie plusieurs centres de données. Chaque région est isolée des autres afin d’améliorer la
sécurité, la résilience et la continuité des services.

Par exemple, une entreprise canadienne peut utiliser la région **Canada Central** afin d’héberger ses applications plus près de ses utilisateurs.

Les Zones de Disponibilité

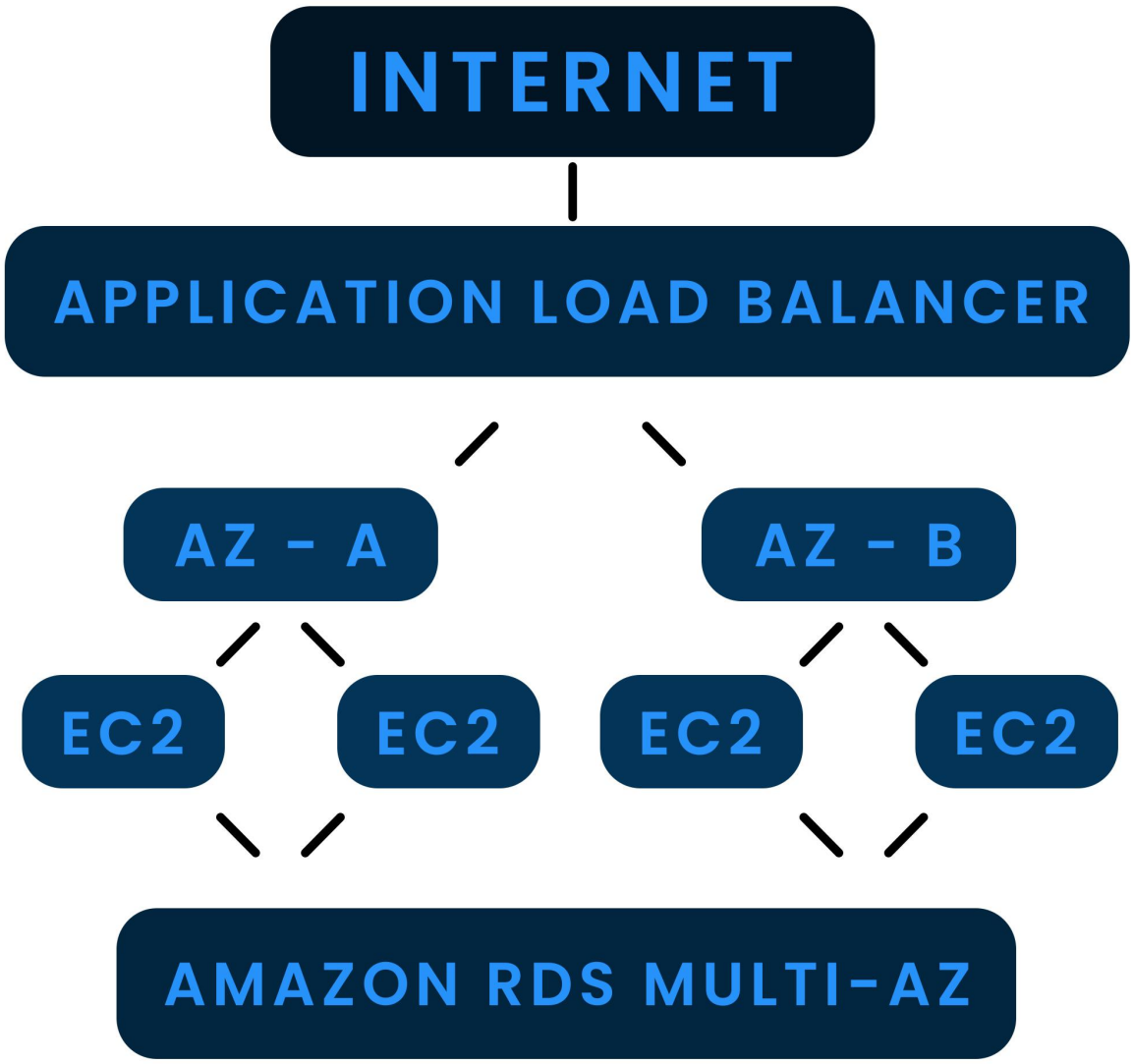
Chaque région AWS contient plusieurs **Zones de Disponibilité**, aussi appelées **Availability Zones** ou **AZ**.

Une AZ correspond à un ou plusieurs centres de données physiquement séparés, disposant de leur
propre alimentation électrique, réseau et système de refroidissement.

```
Région AWS
├── Availability Zone A
├── Availability Zone B
└── Availability Zone C
```

En répartissant les ressources sur plusieurs AZ, une application peut continuer à fonctionner
même si une zone rencontre une panne.

Exemple d’architecture haute disponibilité sur AWS



Dans cette architecture :

- le Load Balancer répartit le trafic entre plusieurs instances ;
- les instances EC2 sont déployées dans plusieurs zones de disponibilité ;
- la base de données utilise Amazon RDS Multi-AZ ;
- l'application reste disponible même si une instance ou une AZ tombe en panne.

Elastic Load Balancing

Elastic Load Balancing permet de distribuer automatiquement le trafic entrant vers plusieurs ressources, comme des instances EC2, des conteneurs ou des fonctions Lambda.

Les principaux types de Load Balancer sont :

- **Application Load Balancer** : adapté aux applications web HTTP et HTTPS ;
- **Network Load Balancer** : adapté aux performances élevées et au trafic TCP/UDP ;

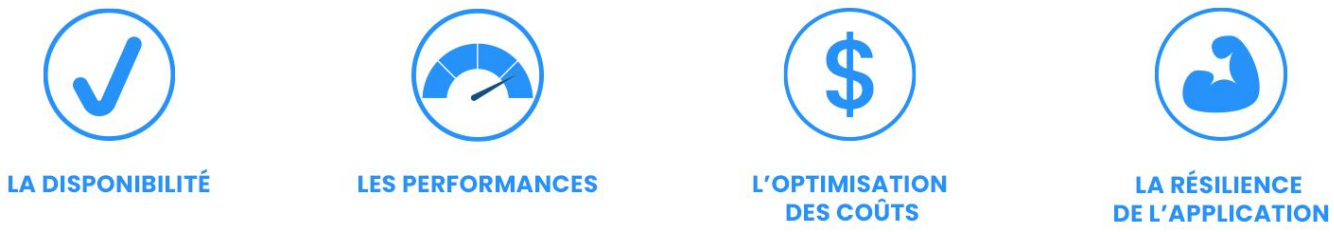
- **Gateway Load Balancer** : utilisé pour intégrer des appliances de sécurité réseau.

Auto Scaling : adapter automatiquement les ressources

Auto Scaling permet d'ajouter ou de supprimer automatiquement des instances EC2
selon la charge de l'application.

Par exemple, une application peut fonctionner avec deux serveurs en temps normal, puis passer à
dix serveurs lors d'un pic de trafic. Lorsque la charge diminue, AWS réduit automatiquement le
nombre d'instances.

Auto Scaling permet donc d'améliorer :



Amazon RDS Multi-AZ

Les bases de données sont souvent des composants critiques dans une architecture cloud.
Amazon RDS Multi-AZ permet de créer une base de données principale et une copie
synchronisée dans une autre zone de disponibilité.

En cas de panne de l'instance principale ou de son AZ, AWS peut effectuer un basculement
automatique vers l'instance de secours.

Cette approche permet de réduire les interruptions et de protéger les applications critiques.

Amazon S3 et la disponibilité du stockage

Amazon S3 est conçu pour stocker les données de manière hautement disponible et
durable. Les objets sont automatiquement répartis sur plusieurs équipements et zones de
disponibilité.

S3 est souvent utilisé pour :

- les sauvegardes ;
- les fichiers utilisateurs ;
- les sites web statiques ;
- l'archivage ;
- le stockage de données applicatives.

Serverless et haute disponibilité

Les services **serverless** d'AWS permettent de créer des applications sans gérer directement les serveurs.

Parmi les services importants, on retrouve :

- **AWS Lambda** ;
- **Amazon API Gateway** ;
- **Amazon DynamoDB** ;
- **Amazon SQS** ;
- **Amazon SNS**.

Avec le serverless, AWS gère automatiquement une grande partie de la disponibilité, de la
scalabilité et de l'infrastructure sous-jacente.

Haute disponibilité et tolérance aux pannes

La haute disponibilité et la tolérance aux pannes sont deux notions proches, mais différentes.

Concept	Objectif
Haute disponibilité	Réduire les interruptions de service
Tolérance aux pannes	Continuer à fonctionner malgré une panne importante

Une architecture répartie sur plusieurs AZ améliore la haute disponibilité. Une architecture
multi-région permet d'aller plus loin en renforçant la tolérance aux pannes et la reprise après
sinistre.

Architecture Multi-Région

Pour les applications très critiques, il est possible de répartir l'infrastructure sur plusieurs
régions AWS.

```
Région Canada Central  <----->  Région US East
```

Cette approche permet de :

- se protéger contre la panne complète d'une région ;

- mettre en place une stratégie de reprise après sinistre ;
- réduire la latence pour des utilisateurs répartis dans plusieurs pays ;
- améliorer la continuité des activités.

Services AWS essentiels pour une architecture résiliente

Service AWS	Rôle dans la haute disponibilité
Amazon EC2	Hébergement des serveurs virtuels
Elastic Load Balancing	Répartition du trafic entre plusieurs ressources
Auto Scaling	Ajout ou suppression automatique d'instances
Amazon RDS Multi-AZ	Base de données résiliente avec basculement automatique
Amazon Route 53	DNS hautement disponible et routage intelligent
Amazon CloudFront	Distribution mondiale du contenu
Amazon S3	Stockage durable et hautement disponible
AWS Lambda	Exécution serverless scalable
Amazon CloudWatch	Surveillance, métriques et alertes
AWS Backup	Sauvegarde centralisée des ressources AWS
AWS Backup	Sauvegarde centralisée des ressources AWS

Pourquoi apprendre la haute disponibilité sur AWS ?

Les entreprises recherchent des professionnels capables de concevoir des architectures cloud
robustes, sécurisées et évolutives.

Maîtriser la haute disponibilité permet de mieux comprendre :

- les architectures multi-AZ ;
- les stratégies de reprise après sinistre ;
- la répartition de charge ;
- la surveillance des applications ;
- l'optimisation des coûts ;
- la préparation à la certification AWS Solutions Architect Associate.

Se former avec Doussou Formation

La formation

AWS Solutions Architect Associate (SAA-C03) – Infrastructure, Serverless et Certification

permet d’apprendre à concevoir, déployer et sécuriser des architectures modernes sur AWS.

Elle couvre notamment :

- les services essentiels AWS ;
- les VPC et les réseaux cloud ;
- EC2, Auto Scaling et Load Balancing ;
- RDS, DynamoDB et les bases de données cloud ;
- S3 et les services de stockage ;
- IAM et la sécurité ;
- les architectures serverless ;
- CloudWatch et la surveillance ;
- les bonnes pratiques AWS Well-Architected ;
- la préparation à la certification AWS Certified Solutions Architect – Associate.

Cette formation s’adresse aux administrateurs systèmes, développeurs, techniciens TI,
responsables d’infrastructure et toute personne souhaitant développer ses compétences en cloud AWS.

Conclusion

La haute disponibilité est un pilier fondamental des infrastructures cloud modernes. Grâce à AWS,
il est possible de concevoir des environnements capables de résister aux pannes, de s’adapter à la
demande et de protéger les services critiques.

Comprendre les concepts comme les Availability Zones, le Load Balancing, Auto Scaling, RDS
Multi-AZ, S3, Route 53 et les architectures serverless est essentiel pour tout professionnel qui
souhaite évoluer dans le cloud computing.

Pour approfondir ces compétences, découvrez la formation

Formation AWS Solutions Architect Associate (SAA-C03)

proposée par Doussou Formation.