

Regiões e AZs (Availability Zones) na Infraestrutura Global da AWS

A AWS organiza sua infraestrutura em **regiões e Zonas de Disponibilidade (Availability Zones)** para fornecer serviços altamente disponíveis, escaláveis e com latência controlada ao redor do mundo.

01. Regiões

- **Regiões** são grandes áreas geográficas separadas, cada uma contendo várias **Zonas de Disponibilidade**.
- Cada região opera de maneira isolada e é independente de outras regiões em termos de conectividade e segurança, garantindo alta disponibilidade e permitindo que as empresas escolham onde hospedar seus dados e serviços, de acordo com requisitos de conformidade ou de desempenho.
- As regiões são identificadas por códigos como **us-east-1** (Virgínia do Norte), **eu-west-1** (Irlanda), entre outros.
- **URL para ver as regiões disponíveis:** [AWS Global Infrastructure - Regiões e Zonas de Disponibilidade](#)

Availability Zones (AZs)

As **Zonas de Disponibilidade (AZs)** são **datacenters fisicamente separados** dentro de uma mesma **região** da AWS. Cada AZ é projetada para operar de forma independente, com alimentação elétrica, rede e sistemas de refrigeração separados, garantindo um nível elevado de **isolamento físico**.

- **Conectividade de baixa latência:** As AZs dentro de uma mesma região são interconectadas por uma rede de **alta velocidade e baixa latência**, o que permite distribuir aplicativos entre várias AZs para alcançar **alta disponibilidade**, melhor desempenho e maior **tolerância a falhas**.
- **Isolamento de falhas:** Apesar de estarem conectadas, as AZs são projetadas para que falhas em uma zona não afetem as outras. Esse design cria uma camada adicional de **resiliência**, garantindo que sua aplicação continue funcionando mesmo se houver uma interrupção em uma zona específica.
- **Tolerância a falhas e recuperação:** Ao distribuir recursos e dados por várias AZs, você melhora a **tolerância a falhas** e a **recuperação de desastres**. Em caso de falha de hardware ou interrupção em uma AZ, as outras podem assumir rapidamente a carga de trabalho.
- **Distância entre AZs:** As AZs em uma mesma região estão separadas fisicamente por distâncias que podem chegar a **100 km**, garantindo que eventos locais, como desastres naturais, não afetem todas as zonas simultaneamente.
- **Disponibilidade de serviços:** Os serviços da AWS são frequentemente lançados primeiro nas **zonas de disponibilidade mais antigas**. Isso significa que pode haver diferenças na disponibilidade de serviços entre as AZs mais antigas e as mais recentes. Antes de utilizar uma AZ específica, é importante **verificar se o serviço que você deseja** está disponível nela.

Zonas Locais (Local Zones)

As **Zonas Locais** são **datacenters menores** da AWS localizados mais próximos dos usuários finais ou de áreas metropolitanas específicas, oferecendo uma infraestrutura de computação com **latência ultra baixa**.

- **Proximidade ao usuário:** Diferente das **Zonas de Disponibilidade (AZs)**, que ficam distribuídas em regiões maiores, as Zonas Locais são estrategicamente posicionadas mais próximas de áreas urbanas ou centros de negócios para reduzir a latência nas interações e melhorar o desempenho de aplicativos que exigem respostas rápidas.
- **Distância:** As Zonas Locais geralmente ficam a uma distância de até **100 km** dos usuários ou de centros de dados regionais, permitindo que empresas e desenvolvedores ofereçam experiências mais rápidas e responsivas.

Essas zonas são ideais para casos de uso como **streaming de vídeo, gaming, realidade aumentada/virtual**, ou aplicações sensíveis à latência.

Redundância e Alta Disponibilidade

- **Planejamento de Alta Disponibilidade:** Aplicações distribuídas por múltiplas AZs garantem que, caso ocorra uma falha em uma AZ, outras AZs possam continuar a operar sem interrupção.
- **Failover Automático:** Para serviços gerenciados como **Amazon RDS**, o AWS oferece suporte para failover automático em caso de falha da instância primária, migrando a operação para outra instância em uma AZ diferente.

Latência e Desempenho

- **Conectividade Ultra-rápida:** As AZs são interconectadas por uma rede de baixa latência com redundância, o que permite que aplicativos escalem horizontalmente com latência mínima.
- **Regiões e AZs:** Ao implantar serviços em diferentes AZs dentro da mesma região, você minimiza a latência, já que as AZs em uma região são projetadas para manter uma comunicação extremamente rápida entre elas.

Resiliência e Isolamento de Falhas

- **AZs Independentes:** Cada AZ é fisicamente separada e isolada das outras AZs em uma região, com infraestruturas de energia, resfriamento e rede distintas para evitar falhas em cascata.
- **Distância Estratégica:** AZs em uma região são geograficamente separadas (distâncias de até 100 km), para proteger contra falhas regionais ou desastres naturais, mas suficientemente próximas para permitir baixa latência.

Distribuição de Aplicações

- **Deploy Multi-AZ:** Ao utilizar serviços como EC2, RDS ou Elastic Load Balancer, você pode optar por uma arquitetura Multi-AZ, onde a replicação de dados ou a distribuição de tráfego é feita automaticamente entre AZs. Isso garante maior **tolerância a falhas e recuperação rápida** em caso de incidentes.
- **Auto Scaling Multi-AZ:** Serviços como EC2 Auto Scaling permitem distribuir instâncias entre múltiplas AZs, ajustando automaticamente a capacidade com base na demanda e garantindo que as instâncias em diferentes AZs compartilhem a carga.

Casos de Uso Relevantes

- **Aplicações Web de Alta Disponibilidade:** Ao configurar balanceamento de carga entre múltiplas AZs, sua aplicação pode continuar funcionando mesmo que uma das AZs fique indisponível.
- **Replicação de Dados em Multi-AZ:** Bancos de dados gerenciados como o RDS oferecem replicação síncrona entre AZs para garantir alta durabilidade e disponibilidade.

Regiões com Múltiplas AZs

- **Múltiplas AZs por Região:** Regiões AWS geralmente possuem múltiplas AZs (três ou mais) para permitir resiliência e balanceamento de carga, mas algumas regiões mais recentes podem começar com menos AZs.
- **Verificação de Suporte Regional:** Certifique-se de que o serviço desejado está disponível em todas as AZs da região que você está utilizando. Algumas AZs mais recentes podem não ter suporte para todos os serviços AWS no início.

Zonas Locais (Local Zones)

- **Zonas Locais:** AWS também oferece **Zonas Locais**, que são extensões de regiões AWS, projetadas para oferecer latência ultrabaixa ao trazer serviços mais próximos do usuário final. Isso é especialmente útil para aplicações que demandam tempos de resposta muito rápidos, como jogos online, processamento de vídeo em tempo real e realidade aumentada.

AWS Wavelength

O **AWS Wavelength** permite a implantação de aplicações que exigem **latência ultrabaixa** ao integrar os serviços da AWS diretamente nas **redes 5G** de provedores de telecomunicações. Ele oferece uma infraestrutura de computação da AWS próxima à borda da rede 5G, permitindo que o processamento e a entrega de dados sejam feitos de maneira mais eficiente e rápida.

Principais características e benefícios:

- **Latência ultrabaixa:** Ao integrar a infraestrutura AWS diretamente nas redes 5G, o Wavelength reduz drasticamente a latência em dispositivos móveis, essencial para aplicações que precisam de tempos de resposta extremamente rápidos.
- **Casos de uso:**
 - **Jogos online:** O Wavelength oferece suporte para jogos em tempo real, melhorando a experiência do jogador com respostas instantâneas.
 - **Realidade aumentada/virtual:** Minimiza o atraso em interações, permitindo experiências mais imersivas e suaves.
 - **IoT (Internet das Coisas):** Dispositivos conectados podem processar e responder a dados em tempo real, o que é fundamental para cidades inteligentes, automação industrial, carros autônomos, etc.
- **Proximidade ao usuário final:** Colocando a infraestrutura de computação da AWS mais próxima dos dispositivos que consomem os serviços, o Wavelength melhora o desempenho e a escalabilidade de aplicações sensíveis à latência.
- **Redes 5G:** Ele é projetado especificamente para aproveitar a alta largura de banda e a baixa latência proporcionadas pelas redes 5G, o que torna ideal para aplicações móveis inovadoras.

Como o AWS Wavelength funciona:

O Wavelength coloca zonas de infraestrutura AWS (chamadas de **Wavelength Zones**) nas bordas das redes de telecomunicações. Essas zonas permitem que os desenvolvedores de aplicações usem os mesmos serviços e APIs da AWS, como EC2, ECS e EKS, mas executando mais perto dos usuários finais, eliminando grande parte do tempo de trânsito na rede.

Redução da latência em dispositivos móveis:

A integração direta com redes 5G permite que os dados trafeguem menos, evitando saltos desnecessários pela rede de longa distância. Isso ajuda a reduzir a latência e oferece experiências de maior qualidade para os usuários móveis.

AWS Outposts

O **AWS Outposts** é um serviço que permite a execução de infraestrutura da AWS em **instalações locais**. Com ele, você pode trazer a mesma infraestrutura, serviços, APIs e ferramentas da AWS para seus próprios datacenters, colocation ou locais físicos. Isso é ideal para empresas que precisam de uma solução de **nuvem híbrida**, onde parte da infraestrutura está na AWS e outra parte em um ambiente on-premises.

Principais características e benefícios:

- **Infraestrutura local:** O AWS Outposts permite que você instale racks gerenciados da AWS diretamente no seu datacenter ou ambiente de colocation, trazendo a nuvem AWS para o local onde você já possui dados e aplicações.
- **Mesmas APIs e ferramentas da AWS:** Você utiliza as mesmas APIs, serviços, automações e ferramentas da AWS no Outposts, garantindo uma experiência de gerenciamento unificada entre a nuvem AWS e o ambiente on-premises.
- **Ideal para requisitos de baixa latência:** Para aplicações que precisam de processamento local por questões de latência ou conformidade de dados, o Outposts oferece uma solução robusta, mantendo a mesma consistência da AWS.
- **Execução de serviços locais da AWS:** É possível rodar serviços da AWS como EC2, ECS, EKS e RDS diretamente no seu Outpost, integrando de maneira contínua com o restante da sua infraestrutura em nuvem.
- **Nuvem híbrida:** Outposts facilita a criação de uma infraestrutura híbrida, combinando a flexibilidade da nuvem pública com o controle e a proximidade da infraestrutura local.

Como funciona:

- **Instalação no local:** O AWS Outposts é fisicamente instalado em seu datacenter ou local físico. A AWS envia racks pré-configurados que são integrados ao ambiente local, e a partir daí, a AWS gerencia remotamente toda a infraestrutura.
- **Gerenciamento completo pela AWS:** A AWS cuida de todo o gerenciamento, manutenção e atualizações de hardware e software, enquanto você continua usando os serviços da AWS como faria na nuvem pública.
- **Cenários de uso:** Empresas que necessitam manter dados localmente devido a regulamentos, ou que precisam de uma infraestrutura com baixa latência e alta performance, se beneficiam do AWS Outposts.

Benefícios principais:

- **Consistência entre ambientes:** A mesma infraestrutura e serviços da AWS, tanto no seu datacenter quanto na nuvem.
- **Baixa latência:** Ideal para casos onde a latência de rede para a nuvem pública pode ser um fator crítico.
- **Conformidade e localização de dados:** Mantém dados sensíveis ou críticos no local para atender exigências regulatórias.