

CLOUD COMPUTING

Feito com Marp plugin do VS Code

Notas: [Link](#)



O QUE É CLOUD COMPUTING?

- Cloud computing, conhecida também como computação em nuvem, é a tecnologia que permite o uso remoto de recursos da computação por meio da conectividade da Internet.
- Poucos termos foram tão populares no mundo da tecnologia quanto a computação em nuvem. Hoje em dia, parece que tudo acontece na “nuvem”.

- **Cloud computing** é uma tecnologia que usa a conectividade e a grande escala da Internet para hospedar os mais variados recursos, programas e informações.
- Dessa forma, a computação em nuvem permite que o usuário os acesse por meio de qualquer computador, tablet ou telefone celular.
- Tudo isso sem a necessidade de conectar-se a um computador pessoal ou servidor local.
- Se ainda não está claro, deixe-me exemplificar:
 - se você já editou um documento no Google Docs, ouviu música no Spotify ou assistiu a um filme na Netflix, teve contato direto com a computação em nuvem.
- Todos eles são serviços online que não exigem o download, instalação e acesso de qualquer tipo de software ou arquivo no seu computador. Você só precisa de um navegador e uma boa conexão com a Internet para usá-los.

VANTAGENS E PARA QUE SERVE A COMPUTAÇÃO EM NUVEM

- A nuvem surgiu como uma forma de democratizar informações e melhorar a experiência de quem depende de recursos tecnológicos em nível pessoal ou profissional.
- A computação em nuvem trouxe inúmeras vantagens. A opção por um serviço na nuvem permite:

- **A redução de custos com infraestrutura:** uma vez que elimina os gastos de capital com compras de hardware e software, instalação e manutenção;
- **A economia do espaço:** dado que os recursos permanecem armazenados online;
- **A centralização da informação:** impedindo que todos os dados sejam mantidos em diferentes programas, com diferentes tipos de formulários de autenticação e acesso;
- **O aumento ou diminuição de acordo com a necessidade do cliente:** graças ao poder de elasticidade que fornece a quantidade ideal de recursos, armazenamento e processamento;
- **Trabalho remoto:** funcionários da empresa podem acessar todos os dados através de qualquer computador, notebook, tablet ou smartphone, desde que estejam conectados à Internet.

Como funciona a Cloud?

- A computação em nuvem usa um servidor remoto para conectar dispositivos dos usuários a recursos centralizados. Um servidor remoto armazena todos os dados e programas de que você precisa e pode estar no seu próprio país ou do outro lado do mundo, por exemplo, no Japão.
- O fato é que você pode acessá-los on-line, de qualquer lugar, mesmo que estejam armazenados a quilômetros de distância.
- Existem 3 tipos de computação em nuvem: o **público**, o **privado** e o **híbrido**. Se você tem uma empresa e planeja usar um serviço na nuvem, tenha em mente que a escolha dependerá de fatores relacionados a custo, disponibilidade, desempenho e expectativas.

1. Nuvem pública

- São recursos computacionais, como servidores e armazenamento, fornecidos por terceiros e disponíveis para qualquer pessoa ou empresa que deseje contratá-los.
- Nesse modelo, o cliente é responsável pelo que será enviado para a nuvem, seja um backup, um aplicativo ou alguns arquivos, enquanto o provedor de nuvem está preocupado com a manutenção, segurança e gerenciamento de todos os recursos.
- **Na nuvem pública, tudo está disponível na web e compartilhado entre vários usuários que o usam simultaneamente** (mas separadamente), o que mantém os recursos padronizados.
- Ao oferecer soluções unificadas, esse modelo acaba sendo mais barato. Portanto, a nuvem pública é indicada para empresas que desejam economizar em investimentos.

2. Nuvem privada

- Nesse modelo, a empresa mantém a infraestrutura da nuvem em seu domínio interno e oferece acesso restrito a usuários selecionados, como funcionários e parceiros.
- **A nuvem privada oferece à empresa a possibilidade de personalizar as funções e o suporte às suas necessidades.** Como a nuvem é projetada exclusivamente para ela, todos os processos são direcionados para a realidade do negócio.
- Em geral, a nuvem privada é usada por organizações que devem seguir certos regulamentos e regras específicas sobre segurança e privacidade de dados e informações, como é o caso de algumas instituições financeiras e governamentais.

3. Nuvem híbrida

- Finalmente, quando falamos sobre a nuvem híbrida, nos referimos à união dos dois anteriores, ou seja, **a combinação que permite compartilhar dados e aplicativos entre os dois tipos de nuvem citados.**
- Assim, de acordo com a necessidade e a estratégia do negócio, alguns recursos são utilizados de forma privada, outros são utilizados publicamente ou estão ligados por meio de tecnologias.

Quais serviços a cloud computing possui?

- Atualmente, a computação em nuvem oferece uma ampla variedade de serviços, ferramentas e funcionalidades para atender às necessidades das empresas.
- Entre as principais ofertas, destacam-se 3 modelos:

SaaS (Software como serviço)

- O SaaS permite que você acesse o software sem comprar sua licença, usando-o gratuitamente na nuvem, geralmente com recursos limitados.
- No entanto, há também planos de pagamento em que é cobrada uma taxa fixa, um valor que varia de acordo com o uso ou mesmo sem custo, se o recurso oferece anúncios ou outra forma de monetização.
- **Nesse formato, a empresa acessa o software pela Internet**, sem a necessidade de se preocupar com a instalação, configuração e investimento em licenças.
- Muitos CRMs e ERPs funcionam através do SaaS. Os serviços de e-mail e outros aplicativos públicos, como Skype, Whatsapp, LinkedIn e Facebook também funcionam com esse tipo de modelo, e é por isso que é o mecanismo mais popular.

PaaS (Plataforma como serviço)

- Neste modelo de nuvem, é contratado um ambiente completo de desenvolvimento on demand, no qual é possível criar, modificar e otimizar softwares e aplicativos.
- A vantagem do modelo PaaS é que ele inclui sistemas operacionais, ferramentas de desenvolvimento, sistemas de gerenciamento de bancos de dados, serviços de Business Intelligence e muitos outros recursos, além de toda a infraestrutura necessária para executar ou aperfeiçoar aplicações Web ou móveis.
- A equipe de desenvolvimento só precisa se preocupar com a programação da aplicação, uma vez que a administração, manutenção e atualização da infraestrutura são deixadas para o provedor.

IaaS (Infraestrutura como serviço)

- **No modelo IaaS, os recursos de infraestrutura são alugados**, como servidores, routers, racks, datacenters, hardware e outras ferramentas que permitem a transmissão e o armazenamento de dados.
- Dependendo do provedor e do modelo escolhido, a empresa pode ser cobrada, por exemplo, pelo número de servidores usados e pela quantidade de dados armazenados ou trafegados.
- **No modelo IaaS, o que é necessário é utilizado de acordo com a demanda da empresa e somente o que é utilizado é pago.**
- **A grande vantagem deste modelo é a sua escalabilidade.** Afinal, em um mês, você pode exigir alguns servidores virtuais onde você armazenará poucos dados e terá pouco tráfego, enquanto, no mês seguinte, poderá solicitar o dobro dos recursos.

AWS

Feito com Marp plugin do VS Code

Notas: [Link](#)



O que é a hospedagem AWS?

- A Amazon Web Service, abreviada como AWS, é o serviço de hospedagem de sites oferecido pela Amazon, uma das maiores empresas de comércio digital do mundo.
- Por meio desses serviços, você tem **como criar um site e disponibilizar aplicativos online para seu público com mais facilidade.**

- Quando você contrata um serviço de hospedagem de site, não está pagando apenas por um espaço virtual, **mas também pelo servidor físico onde ele será armazenado.**
- Na hora de acessar a página, a distância entre o servidor de origem e o computador do usuário **afeta diretamente o tempo de processamento.** Por isso, alguns sites estrangeiros tendem a levar mais tempo para carregar.
- Isso significa que, quando você contrata um servidor de fora do país, precisará de mais tempo para fazer uma conexão adequada.
- Felizmente, a Amazon tem sistemas de hospedagem ao redor do mundo, sendo bem fácil encontrar uma localização otimizada para sua conexão e seu público-alvo.

Alta escalabilidade sob demanda

- Com o tempo, será necessário ampliar o seu serviço de cloud computing, seja adotando novas funcionalidades, seja simplesmente ampliando o tamanho do seu armazenamento de dados. **Uma situação familiar seria adicionar um novo disco rígido ao seu computador para ampliar sua memória.** Como você já deve saber, essa situação pode dar bastante trabalho.
- Para lidar com isso, o ideal é contar com um **serviço de hospedagem com alta escalabilidade**, ou seja, a possibilidade de ampliar suas funcionalidades e armazenamento sem fazer uma grande reestruturação.
- O AWS é uma ótima opção nesse quesito, pois todo o sistema é baseado em cloud computing, sem necessidade de alterar muito em sua empresa para melhorar o serviço.

Quais os serviços da hospedagem AWS?

- Junto à hospedagem em si, a Amazon também oferece outros serviços associados.

1. CloudFront

- Esse é o serviço de entrega de dados da Amazon, ou seja, ele garante que as informações armazenadas na nuvem cheguem até o usuário com o mínimo de espera. **Isso é o que permite que um site possa receber vários acessos simultâneos sem perda de qualidade.**

2. EC2

- O Elastic Computer Cloud é um serviço que administra a capacidade computacional dos processadores na nuvem. Com ele, é **possível solicitar e distribuir mais processamento para seu site** ou qualquer outro serviço de cloud computing.

3. RDS

- O Relational Database Service é um sistema que permite o gerenciamento de banco de dados em nuvem. Com ele, é **mais fácil fazer análises de larga escala e avaliar o desempenho do seu negócio.**

4. S3

- O Simple Storage Service, **é um sistema de armazenamento de dados e programas em larga escala**, como o nome diz. Com sua alta escalabilidade e segurança, é uma das opções mais seguras do mercado.

SEGURANÇA NA NUVEM AWS

- A AWS é projetada para lhe ajudar a construir uma infraestrutura segura, resiliente, eficiente e de alto desempenho para as suas aplicações.



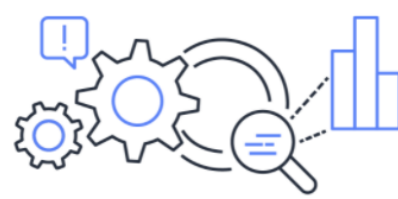
Prevenir

Definir medidas para permissões e identidades de usuários, proteção de infraestrutura e proteção de dados a fim de estabelecer uma estratégia de adoção na AWS suave e planejada.



Detectar

Adquirir visibilidade para a estratégia de segurança de sua organização a partir de serviços de monitoramento e registro em log. Ingerir estas informações em uma plataforma escalável para garantir o gerenciamento de eventos, testes e auditoria.



Resposta

Resposta e recuperação automatizadas a incidentes para ajudar a mudar o foco principal das equipes de segurança, de forma que possam se concentrar na análise da causa raiz.



Corrigir

Aproveitar a automação orientada por eventos para corrigir e proteger rapidamente seu ambiente da AWS, praticamente em tempo real.

Regiões e Zonas de Disponibilidade

- A Nuvem da AWS abrange possui uma infra-estrutura global com mais de 69 zonas de disponibilidade (AZs) em mais de 22 regiões geográficas em todo o mundo.
- Cada região da AWS fornece redundância, e contrário dos outros provedores de nuvem, que definem uma região como um único datacenter, as regiões da AWS consistem em várias zonas de disponibilidade, ou AZs (normalmente, três).

- Cada **zona** (AZ) é uma partição totalmente isolada da infraestrutura da AWS composta por **datacenters**.
- Cada datacenter conta com energia, redes e conectividade redundantes e está hospedado em instalações separadas.
- Imagine, por exemplo, que uma Região poderia ser São Paulo, que teria 3 data centers fisicamente isolados (AZs), um em Campinas, outro em São Bernardo do Campo e outro no bairro da Lapa.

- As zonas permitem alta disponibilidade, tolerância a falhas e escalabilidade em níveis superiores aos que um único datacenter pode oferecer.
- As zonas são fisicamente separadas por uma distância significativa (**100 quilômetros mais ou menos**) de todas as outras zonas.
- Todas as zonas são interconectadas por redes de banda larga e baixa latência, usando fibra dedicada e totalmente redundante para proporcionar redes de alto vazão e baixa latência.
- O desempenho da rede é suficiente para realizar a replicação síncrona entre as zonas.
- A AWS associa um data center físico a zona da sua conta exemplo:
- **us-east-1**, que não necessariamente é o mesmo data center físico de outra conta da AWS com esse mesmo nome, dessa forma a AWS faz um balançamento de carga entre os servidores.

Edge Locations

- A ideia básica é que você pode ter **caching, distribuição de conteúdo estático e até funções lambda** sendo executadas em **locais adicionais** que podem estar mais perto da request dos usuários gerando um latência reduzida para as requisições.
- Por exemplo, a AWS não tem (pelo menos em Novembro de 2021) uma Região com zonas no Rio de Janeiro, mas tem um Edge Location na cidade que pode usado por esses serviços para atender requests com menor latência.

VPC

- Uma VPC (Virtual Private Cloud) permite provisionar uma seção da Nuvem AWS isolada logicamente na qual é possível executar recursos da AWS em uma rede virtual que você mesmo define.
- Você tem controle total sobre seu ambiente de redes virtuais, incluindo a seleção do seu próprio intervalo de endereços IP, a criação de sub-redes e a configuração de tabelas de rotas e gateways de rede. Você pode usar IPv4 e IPv6.

- Você pode criar uma sub-rede pública para os servidores web que têm acesso à Internet e dispor os sistemas de back-end, como bancos de dados ou servidores de aplicativos, em uma sub-rede privada sem acesso à Internet.
- Quando você cria uma VPC, é necessário especificar um intervalo de endereços IPv4 para a VPC sob a forma de um bloco, por exemplo, 10.0.0.0/16.
- **Uma VPC abrange todas as zonas de disponibilidade na região.**
- Atualmente há um limite de 5 VPCs que podem ser criadas por região em cada conta da AWS.

- Para aprofundar em VPC e redes:

[Link](#)

DNS e Route 53

- A grosso modo, o DNS (Domain Name System) é o que permite que o seu domínio na Internet (ex: [marianne.com](#)) seja roteado para um determinado servidor que responde com o devido conteúdo ou serviço como envio e recebimento de e-mails, website, FTP, etc.

- Para isso o servidor de DNS deve armazenar uma série de registros. Existem diversos tipos de registros com diferentes finalidades, os principais são:
 - A (Host address)
 - AAAA (IPv6 host address)
 - ALIAS (Auto resolved alias)
 - CNAME (Canonical name for an alias)
 - MX (Mail eXchange)
 - NS (Name Server)
 - TXT (Descriptive text)

- Na AWS, o Route 53 é o serviço para você fazer a gestão de DNS, ele permite que você registre domínios, faça o roteamento dos domínios para recursos da AWS, e performa Health Checks.
- O nome desse serviço é **Route 53** porque a porta de serviço do DNS é a de número 53.
- Os registros de alias do Route 53 fornecem uma extensão específica à funcionalidade do DNS e permitem que você roteie o tráfego para recursos da AWS selecionados, como CloudFront, S3 e EC2.

- O Route 53 permite que você defina uma série de políticas para distribuição do tráfego, alguns deles são.
 - **Simple:** Simplesmente passa o tráfego para o destino.
 - **Failover:** Roteia se o health check da principal falhar.
 - **Geolocation:** Roteia para o mais próximo do client.
 - **Latency:** Roteia para a menor latência.
 - **Multivalue:** Distribui para múltiplos destinos (round-robin).
 - **Weighted:** Distribui para diferentes destinos de acordo um percentual pré-definido (ex: 70% – 30%).

Elastic Load Balancer

- O **Elastic Load Balancing** (ELB) distribui automaticamente o tráfego de entrada de aplicativos entre diversos destinos, como instâncias do EC2, contêineres, endereços IP e funções Lambda.
- O serviço pode lidar com a carga variável de tráfego dos aplicativos em uma ou em diversas zona de disponibilidade (AZs).

- O Elastic Load Balancing oferece três tipos de load balancers, todos eles com a alta disponibilidade, a escalabilidade automática e a segurança robusta necessárias para tornar os aplicativos tolerantes a falhas.
- O ALB, Application Load Balancer atua na Camada 7 (Aplicação), o Network Load Balancer na Camada 4 (Transporte), e o Classic Load Balancer em qualquer uma das duas.

- O Application Load Balancer é mais adequado ao balanceamento de carga de tráfego HTTP e HTTPS e oferece roteamento avançado de solicitações para a entrega de arquiteturas modernas de aplicativos, incluindo microsserviços e contêineres. O Application Load Balancer roteia tráfego a destinos dentro de uma VPC.
- O Network Load Balancer é mais adequado ao balanceamento de carga de tráfego TCP (Transmission Control Protocol), UDP (User Datagram Protocol) e TLS (Transport Layer Security) que exige performance muito alta.

- Operando no nível de conexão (camada 4), o Network Load Balancer roteia tráfego para destinos dentro da VPC e é capaz de lidar com milhões de solicitações por segundo, mantendo latências muito baixas.
- O ALB, dá suporte a Sticky Session e SNI, além de Path e Host Based Routing, o NLB por estar em camada mais baixa não, sendo o roteamento feito principalmente por número de porta. O NLB tem menos funcionalidades, mas naturalmente, oferece maior desempenho.
- Você pode usar o ELB para endpoints públicos, mas também para privados.