



ARQUITETURA DE SOLUÇÃO EM NUVEM

Solução IaaS

IaaS (Infraestrutura como Serviço)

É um modelo de serviço de computação na nuvem em que se fornecem todos os equipamentos de computação e infraestrutura para a criação de aplicativos e serviços. IaaS pode ser considerada ainda como a base prática para a computação na nuvem, pois é o serviço que fornece tudo o que é necessário para que a infraestrutura possa ser provisionada, como servidores e armazenamento. Uma solução de IaaS normalmente inclui componentes de hardware, como:

- Servidores virtuais.
- Instâncias de servidor.
- Sistemas de armazenamento de dados.
- Sistemas de gerenciamento de dados.
- Ferramentas de balanceamento de carga.

Além disso, uma solução de IaaS também pode incluir componentes de software, como plataformas de orquestração de nuvem, para gerenciar múltiplos provedores de serviços de nuvem em um único ambiente.

Dessa forma, as empresas podem obter os benefícios da escalabilidade, alta disponibilidade, elasticidade, balanceamento de carga, cenários multicloud e cloud híbrido, aproveitando todas as vantagens dos ambientes de IaaS. Vamos agora analisar cada um desses cenários.

Cenário de elasticidade

É um modelo de provisionamento de recursos que permite a um usuário solicitar rapidamente recursos e serviços adicionais em resposta às demandas do usuário. Utiliza-se para criar e ajustar o tamanho de recursos de computação, armazenamento e rede a fim de acomodar as demandas cambiantes. Esse recurso é útil para empresas que precisam de recursos adicionais de forma rápida e com custo reduzido, pois não precisam adquirir recursos permanentes.

A elasticidade no IaaS permite que os usuários adicionem rapidamente recursos, como servidores, memória e armazenamento, para lidar com picos de demandas.

A elasticidade no IaaS também permite aos usuários licenciar recursos adicionais de forma rápida e retirá-los quando não são mais necessários. A elasticidade permite que a empresa economize custos operacionais, pois ela só paga por recursos adicionais enquanto estiverem sendo usados. Outra vantagem é que a elasticidade no IaaS pode ser usada para escalar vertical ou horizontalmente a capacidade, a fim de permitir que a empresa aproveite melhor seus recursos.

Cenário de balanceamento de carga

Neste cenário, encontramos um sistema projetado para distribuir o tráfego dos usuários e garantir a alta disponibilidade dos servidores. Ele ajuda a balancear o tráfego entre múltiplos dispositivos, como servidores, links de rede, base de dados e serviços de armazenamento. O balanceamento de carga também aumenta a velocidade de resposta dos servidores, monitora o tempo de resposta dos servidores, otimiza o desempenho, evita o congestionamento e mantém os usuários conectados. Além disso, pode ser usado para evitar falhas críticas.

Um sistema de balanceamento de carga em camadas permite que as empresas dividam a carga entre vários servidores virtuais ou instâncias de servidor. Esses servidores virtuais ou instâncias de servidor também são gerenciados por um serviço de balanceamento de carga, como o Elastic Load Balancer da Amazon, que garante que todas as requisições sejam atendidas pelos servidores virtuais ou instâncias de servidor que possuem maior capacidade de processamento. Além disso, essa solução de IaaS de balanceamento de carga também pode oferecer suporte a vários serviços de armazenamento, como o S3 da Amazon, que garante a alta disponibilidade dos dados.

Cenário de alta disponibilidade

O foco deve estar em garantir o tempo de funcionamento máximo dos serviços fornecidos. Para isso, é necessário implementar medidas a fim de garantir altos níveis de redundância, alta escalabilidade e baixo tempo de latência. Vejamos alguns exemplos de ações que podem ser tomadas para permitir a alta disponibilidade:

Replicação de dados

É a prática de ter múltiplas cópias dos mesmos dados em diferentes locais. Isso assegura que, caso o sistema seja interrompido em um local, os dados ainda estarão disponíveis em outro lugar.

Balanceamento de carga

É o processo de distribuir o tráfego de solicitações entre vários servidores para garantir a alta disponibilidade e escalabilidade.

Uso de redundância

É responsável por duplicar os serviços, materiais e equipamentos para que, caso algo falhe, haja sempre uma cópia de backup a ser usada.

Monitoramento contínuo

É capaz de detectar e resolver problemas de desempenho antes que causem falhas no sistema.

Atualizações frequentes

É responsável por garantir que as vulnerabilidades e erros sejam corrigidos antes que eles possam afetar a operação.

Cenário multicloud

A solução de IaaS multicloud é ideal para empresas que desejam ter a liberdade de utilizar diversos provedores de serviços de nuvem em um único ambiente.

Essa solução é normalmente baseada em ferramentas de orquestração de nuvem, como o Kubernetes ou o OpenStack, que permitem que as empresas gerenciem de forma centralizada e unificada todos os serviços de nuvem oferecidos por vários provedores.

Cenário de cloud híbrido

A solução de IaaS de cloud híbrido é ideal para empresas que necessitam aproveitar os benefícios da nuvem pública e privada em um único ambiente. Normalmente, consiste em uma combinação de infraestrutura local (servidores físicos) e serviços de nuvem, como o S3 da Amazon e o Elastic Load Balancer.

Essa solução permite que as empresas aproveitem os recursos de nuvem de forma eficiente, fazendo com que servidores físicos sejam usados em locais estratégicos e serviços de nuvem sejam usados para lidar com a demanda crescente de forma escalonável. Além disso, oferece suporte a sistemas de gerenciamento de dados, como o MongoDB, para que as empresas possam armazenar, acessar e gerenciar seus dados de forma ágil e segura.

Edge computing

O que é edge computing?

É uma abordagem de computação de borda que permite que os dados sejam processados e armazenados localmente, em vez de forma remota. Isso diminui o tempo de latência e aumenta a confiabilidade da conexão. Ao executar os aplicativos do usuário diretamente nos dispositivos de borda, o edge computing oferece maior desempenho, segurança e escalabilidade.

Além disso, ajuda a minimizar os gastos de largura de banda, ao reduzir a quantidade de dados e instruções que precisam ser enviadas para servidores remotos. Isso é especialmente importante para aplicativos que precisam de uma resposta imediata, como os aplicativos de realidade aumentada e jogos. Para entendermos melhor essa tecnologia, vamos analisar cada cenário:

Internet das Coisas (IoT)

O processamento dos dados é realizado por meio de dispositivos responsáveis por reunir os dados coletados. A sua quantidade vai depender do equipamento, após a coleta, e armazenamento dos dados, que são enviados para a nuvem em um centro de processamento.

Edge computing

O processamento dos dados é feito de forma local e separa aqueles que podem ser processados ali mesmo, diminuindo o tráfego de dados e a necessidade de enviá-los. Os dados são processados nos extremos de uma rede e apenas uma parte é enviada para a nuvem, por isso o nome computação de borda.

Quando usar edge computing?

É usado para processar dados onde os usuários os fazem, ao invés de enviá-los para o cloud ou um datacenter remoto. Isso significa que a computação de borda permite processamento de dados mais rápido e eficiente, pois os dados são processados próximos aos usuários. É frequentemente usado para aplicativos de IoT, como dispositivos móveis e outras soluções inteligentes que necessitam de processamento local para tomar decisões rápidas e ações imediatas.

Quando os usuários têm aplicativos ou necessidades que exigem análise de dados em tempo real e processamento de dados no local para redes IoT inteligentes, edge computing pode ser a melhor solução. A computação de borda também pode ser usada para aplicações que exigem tempo de resposta

rápido para fornecer serviços em tempo real, como reconhecimento facial ou localização em tempo real.

Ainda, pode ajudar a resolver alguns problemas de largura de banda, já que os dados não precisam ser transferidos. Em geral, a computação de borda pode ser usada em qualquer aplicação que precise de processamento de dados rápido e eficiente, que se beneficie do tempo de resposta reduzido e habilidade de fazer decisões quase que instantaneamente.

Atenção!

Na hora de tomar a decisão sobre o uso de edge computing, alguns fatores são fundamentais, como: largura de banda de internet, energia, tamanho da infraestrutura, capacidade de processamento, tipo de serviço da sua empresa e custo de implantação.

É preciso levar em conta também a rede já existente e a facilidade de encontrar profissionais especializados para a manutenção, entre outros requisitos que a empresa já atenda, para avaliar se os benefícios da edge computing podem ser aplicados ao negócio.

Edge computing e serverless

Para entendermos melhor a tecnologia edge computing e assim poder explorar seus benefícios, as aplicações edge native devem ser capazes de executar a lógica de serverless computing quando e onde for necessário.

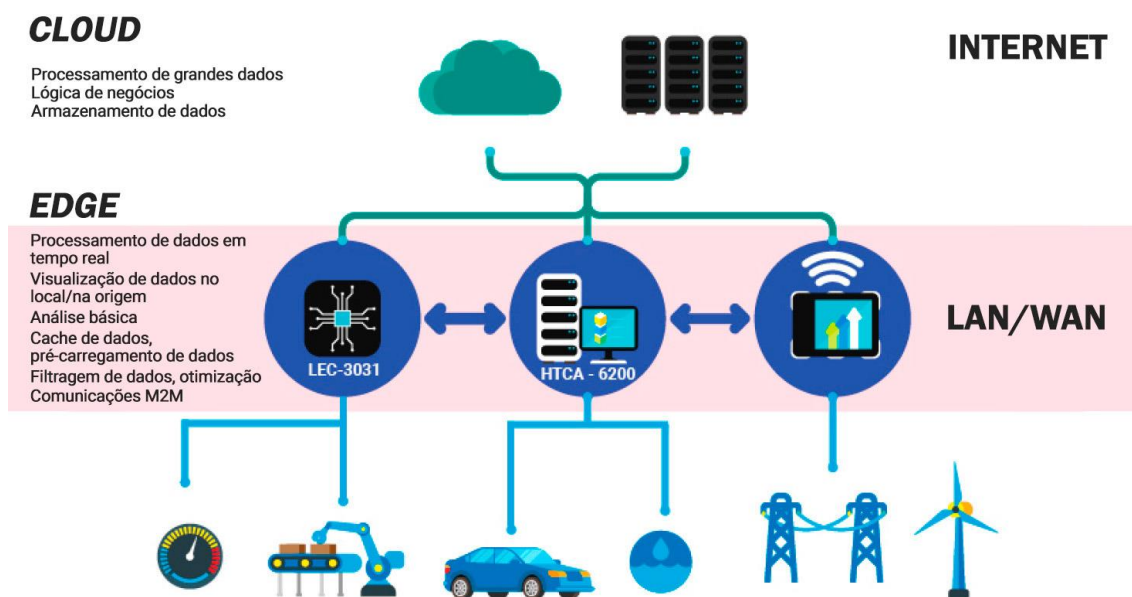
Serverless é uma tecnologia de computação que separa a infraestrutura de um back end, como servidores e outras configurações, de aplicações e serviços. Em vez de se preocupar com a configuração, a manutenção e a escalabilidade de uma infraestrutura, o usuário tem a capacidade de executar cargas de trabalho em ambientes gerenciados, sem se preocupar com a infraestrutura.

Os serviços gerenciados oferecem escalonamento horizontal automático, configuração de alta disponibilidade, monitoramento e alto desempenho, o que significa que o usuário pode se concentrar em suas cargas de trabalho em vez de gerenciar servidores. Sendo assim, o serverless computing pode ser visto como a evolução natural da tendência em direção à menor complexidade, ao uso de recursos mais eficientes e a serviços pay as you go.

Com serverless computing, os clientes não precisam operar ou gerenciar a infraestrutura subjacente, o que em uma arquitetura de rede altamente distribuída – como edge computing – seria quase impossível. Combinar edge computing e serverless é a única forma de permitir que os clientes façam uso

de recursos de edge computing de uma maneira muito fácil, com a mesma tecnologia altamente inovadora, colocando líderes de mercado e novos entrantes lado a lado.

Com serverless, as empresas podem construir e executar funções com edge computing, pagando apenas pelos recursos usados. Isso significa que os desenvolvedores só precisam se concentrar em escrever códigos, o que acelera o lançamento de novas aplicações e funcionalidades. Na imagem a seguir é possível visualizar a divisão de camadas da arquitetura usando a nuvem e o edge computing. Veja:



Arquitetura edge e cloud.

Migração para nuvem

Neste vídeo iremos apresentar as estratégias que podem ser utilizadas para realizar a migração para a nuvem.

Adotar a computação em nuvem não é uma decisão simples, porém atualmente tornou-se indispensável para se ter sucesso na era digital. Um número elevado de empresas está migrando para a nuvem corporativa, e assim incluindo no seu orçamento os gastos com TI em soluções cloud.

No entanto, apesar dos investimentos na nuvem, uma em cada três empresas não consegue perceber seus benefícios de acordo com o relatório Unisys Cloud Success. A estratégia de migração para a nuvem é um assunto complexo.

Quando migrar para a nuvem?

Vejam algumas respostas para esse questionamento:

- **Quando sua infraestrutura está ultrapassada:** Se sua infraestrutura de TI está ultrapassada ou não é mais suficiente para suas necessidades, a computação em nuvem pode ser uma opção viável.
- **Quando você precisa escalar rapidamente:** Se você precisa aumentar a capacidade de processamento ou armazenamento rapidamente, a computação em nuvem pode ser a solução ideal.
- **Quando você precisa acessar de qualquer lugar:** A computação em nuvem é ideal para aplicativos que precisam ser acessados de qualquer lugar, pois oferece acesso remoto e seguro aos seus dados e aplicativos.
- **Quando você precisa economizar dinheiro:** A computação em nuvem elimina muitos dos custos associados à infraestrutura de TI, como aquisição de hardware e licenciamento de software.
- **Quando você precisa de maior segurança:** A computação em nuvem oferece maior segurança dos dados e aplicativos, pois seu provedor de nuvem pode gerenciar backups e outras medidas de segurança.

Estratégia para migração

Inicia com um planejamento adequado e com uma boa seleção da abordagem de migração. Então, vamos conhecer as cinco primeiras estratégias de migração para a nuvem, também conhecidas como 5Rs (atualmente, foram inseridos mais 2Rs, evoluindo para os 7Rs).

Rehost (Re-hospedar)

É uma estratégia de migração para nuvem também conhecida como lift and shift. Ela consiste em mover aplicações, infraestrutura e dados para um ambiente de nuvem, mantendo intactas as suas configurações.

O rehost pode ser aplicado a ambientes on premises e bare metal, migrando os recursos de um ambiente físico para um ambiente de nuvem.

A vantagem do rehost é que, apesar de não alterar a arquitetura da aplicação, ele permite a execução em um ambiente de nuvem, aproveitando assim os benefícios da escalabilidade, alta disponibilidade, melhor performance e segurança da nuvem.

No entanto, esta estratégia não pode ser aplicada a todos os ambientes, pois existem alguns casos em que a aplicação não poderá ser executada diretamente na nuvem. Por isso, é importante avaliar todos os aspectos técnicos de uma aplicação antes de optar pelo rehosting.

A estratégia basicamente resume-se na movimentação de servidores físicos e virtuais existentes para uma solução IaaS. O modelo IaaS hospeda a infraestrutura normalmente encontrada em sites e oferece um ambiente virtualizado por meio de uma camada de hypervisor.

Replatform (Replataforma)

Nem todo sistema é facilmente migrado. No caso de sistemas legados, muitas organizações possuem sistemas estruturados demais para migrar para plataformas de nuvem IaaS. Nesse caso, a solução é emular as aplicações através de máquina virtual e assim garantir sua compatibilidade com as tecnologias de nuvem.

Na prática, esta estratégia permite que se façam algumas alterações de configuração para melhor se adequar ao ambiente de nuvem, sem alterar a arquitetura principal. Os desenvolvedores, geralmente, aplicam essa abordagem para alterar a maneira como se interage com o banco de dados para que possam ser executados em plataformas gerenciadas.

Repurchase (Recompra)

É conhecida como uma estratégia de marketing em que uma empresa incentiva seus clientes a comprar novamente seus produtos ou serviços. Ela envolve oferecer descontos, ofertas especiais, programas de fidelidade e outras formas de recompensar os clientes por sua lealdade.

A estratégia tem como objetivo aumentar a receita e a lucratividade da empresa, ao mesmo tempo em que ajuda a construir relacionamentos de longo prazo com os clientes.

A repurchase, também conhecida como estratégia drop and shop, substitui a aplicação on premise por um software nativo da nuvem. Normalmente, essa abordagem consiste em mudar para uma plataforma de SaaS (Software as a Service) com os mesmos recursos.

Refactoring/Re-architecting (Refatorar/Rearquitetar)

Consiste em desenvolver os sistemas do zero para torná-los nativos da nuvem. Essa estratégia permite aproveitar todo o potencial das tecnologias, como

arquitetura de microsserviços, sem servidor, contêineres e balanceadores de carga.

Por exemplo, você pode refatorar ativos ao mover seus ativos digitais de uma arquitetura monolítica local para uma arquitetura totalmente sem servidor na nuvem. Essas estruturas são escaláveis, ágeis, eficientes e retornam ROI, ou seja, um significativo retorno sobre o investimento a longo prazo, mesmo nos mercados mais competitivos.

Retire (Aposentar)

Esta estratégia está relacionada à desativação ou ao desligamento dos serviços (workloads) que não são mais necessários. Sua desativação permite que a organização se concentre em áreas que oferecem mais valor comercial, economizando recursos.

Retain (Reter)

Também conhecida como revisit, é a estratégia de refatorar algumas áreas críticas de seus ativos digitais antes de migrar para a nuvem.

Eventualmente, você pode descobrir que algumas aplicações são mais adequadas para ainda serem usadas localmente ou foram atualizadas recentemente e precisam ser mantidas. Em outros casos, os sistemas são retidos devido aos requisitos de latência, conformidade ou restrições regulatórias.

Relocate (Realocação)

Outra abordagem centrada em infraestrutura, mas, dessa vez, aproveita o mesmo conjunto de componentes básicos. Assim como na re-hospedagem, o software em execução nas máquinas virtuais migradas permanece inconsciente de que algo mudou.

Nesse caso, no entanto, as ferramentas e os processos operacionais existentes também podem ser mantidos, mesmo que dependentes de produtos de terceiros. Esse tipo de migração, geralmente, pode ser realizado por uma equipe de operações devidamente qualificada, sem envolvimento de desenvolvedores de aplicativos.