

# O que é Computação em Nuvem?

A computação em nuvem é um meio de alocar recursos de processamento, como capacidade de armazenamento para arquivos ou um banco de dados, ciclos de CPU para execução de transações e comunicação de internet, que são compartilhados. A alocação mais comum é feita em um provedor de nuvem, como a Microsoft Azure, Google Cloud, Amazon AWS ou IBM Cloud, para citar alguns dos principais provedores e funciona como um aluguel e os recursos, compartilhados por várias empresas e indivíduos, são cobrados apenas pelo seu uso.

Os serviços disponíveis variam de provedor para provedor, mas em geral você encontrará nos provedores serviços de:

**Computação** – como por exemplo, servidores Linux, Windows ou aplicativos Web, e mais recentemente serviços de GPU

**Armazenamento** – como arquivos e bancos de dados

**Rede** – conexões seguras, banda de acesso à internet

**Análise** – serviços de Aprendizado de Máquina e Análise de Dados

## Serviços de computação

Importante você notar que empregamos a palavra **serviço**. Esse é um aspecto essencial da Computação em Nuvem. Quando você contrata um recurso para uso, como exemplo ciclos de CPU ou 1TB de armazenamento, você pode definir características do *serviço* contratado (como horários de uso do processamento, tempo médio de acesso aos dados, tempos de recuperação etc.). Essas características do serviço devem ser garantidas pelo provedor, mas este tem liberdade de fornecer esses serviços com os recursos (hardware e software) que julga mais adequado. Desse modo será transparente, para você que contratou o serviço, os mecanismos que implementam e fornecem os serviços. Se os ciclos de CPU estão sendo fornecidos por uma única ou várias CPUs, a frequência e marca dos processadores (Intel, AMD etc.), ou se seu armazenamento de 1TB

está em um ou vários volumes físicos, em um disco sólido ou de rotação, é transparente, ou ainda, não acessível para você. É isso que diferencia você contratar um serviço e de você contratar um hardware ou dispositivo específico.

Como você verá mais adiante, essa é uma das vantagens do uso de serviços em nuvem. A meta da computação em nuvem, tendo como foco negócios e a computação comercial, facilitar e tornar mais eficiente a administração dos negócios e essa 'transparência' permite o negócio deixar vários detalhes de implementação dos serviços a cargo dos provedores.

Vamos ver alguns exemplos antes de prosseguir?

## **Exemplos de Serviços**

### **Computação**

Quando você faz transações bancárias, compras pela internet, envia ou email ou usa serviços como Wase, Spotify ou Netflix você provavelmente está interagindo com servidores em nuvem (de uma nuvem pública ou privada) que processam suas solicitações e retornam com uma resposta.

Quando você cria essas soluções várias formas de serviço podem ser adotadas. Por exemplo, uma alternativa para o processamento, é a criação de uma VM (máquina virtual) que 'emula' um computador. A VM aparece você como um computador físico com Windows ou Linux, embora nos bastidores essa máquina possa ser programa executando e alocando parte dos recursos (cpu, memória, espaço em disco e conectividade de rede) de uma ou mais máquinas físicas. Nessa máquina, você pode, então, instalar programas e softwares que desejar para construir seus serviços de negócio (por exemplo, seu website de vendas).

Veja que, nesta solução, você não precisa adquirir ou instalar um hardware e um sistema operacional. Embora, ao final exista um ou mais servidores físicos na implementação, é o provedor executa a máquina virtual em um de seus datacenters.

As VMs não são a única opção de computação, havendo ainda as opções de containers e de computação *serverless*, mas deixamos isso para mais adiante.

## Armazenamento

Os provedores de nuvem oferecem serviços para armazenar todos os tipos de dados: dados textos, dados estruturados (como tabelas de clientes, produtos etc.), imagens, fotos, e-mails etc. Cada tipo de dado, entretanto, tem requisitos diferentes de armazenamento, como tempo de acesso, tempo de recuperação, volume de dados. Enquanto dados texto (documentos) ou um vídeo, podem ser armazenados em um arquivo em disco comum, os dados de clientes e produtos para um sistema de vendas, são melhor armazenados de forma estruturada em um banco de dados.

O volume e a forma de acessar dados em formato *streaming* por exemplo, pode ser bastante diferente do volume e forma de acesso a dados de um catálogo de produtos. É comum, então, os provedores fornecerem diferentes tipos serviços de armazenamento para que você possa escolher o tipo de serviço mais adequado para sua necessidade.

O custo do armazenamento é então determinado pelas características do serviço. Por exemplo, tempos de acesso aos dados menores, maiores volumes e tempo de recuperação menores estão associados a custos maiores de armazenamento.

Uma das vantagens do armazenamento em nuvem é a possibilidade de escalar os recursos de acordo com a sua necessidade. Diferentemente de quando você compra um computador pessoal (imagine que você decide comprar um computador com 1TB de disco, você precisará liberar espaço, fazer um *upgrade* ou adquirir outro computador quando o espaço disponível terminar) os recursos podem ser expandidos e o custo estará sempre limitado ao usado (no computador pessoal do exemplo anterior, você pagará por 1TB mesmo não usando).

Agora que você já se familiarizou com os Serviços em Nuvem, que tal buscarmos algo mais formal para podermos se aprofundar nos seus conceitos?

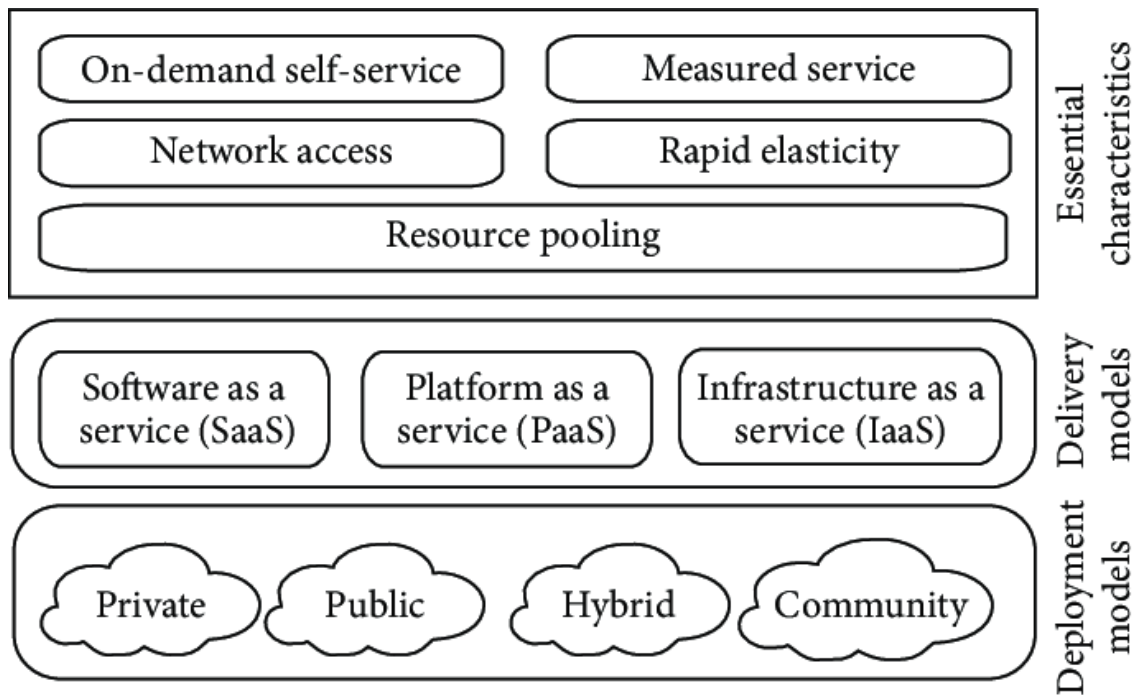
## Definição de computação em nuvem

O NIST – **National Institute of Standards and Technology** é um Órgão do Governo Americano que publica uma série de padrões empregados por inúmeras empresas americanas e no mundo. Ela publicou em 2011 uma versão de definição de computação em nuvem que traz o essencial da computação em nuvem e é amplamente aceita em todo mundo. Você vai conhecer aqui essa definição que permitirá você entender melhor os conceitos dos Serviços em Nuvem. Ela é uma referência para discussão dos temas de computação em nuvem e suas aplicações, e você pode achar interessante buscar as referências do NIST para outros temas relacionados como Big Data, Segurança da Informação etc.

### A definição do NIST de computação em nuvem

*“A computação em nuvem é um modelo para habilitar o acesso por rede ubíquo, conveniente e sob demanda a um conjunto compartilhado de recursos de computação (como redes, servidores, armazenamento, aplicações e serviços) que possam ser rapidamente provisionados e liberados com o mínimo de esforço de gerenciamento ou interação com o provedor de serviços”.*

É uma definição enxuta, mas que se você prestar bastante atenção traz tudo que é essencial para aquilo de denominamos computação em nuvem. O NIST ainda define um conjunto de Características Essenciais, as Modalidades de Serviço e de Distribuição que você vai explorar a seguir. Empregaremos as definições do NIST para apresentar diversos exemplos e aplicações dessas definições nos Serviços em Nuvem em geral.



**Figura 1. Modelo de Referência NIST de Computação em Nuvem.**

A Figura 1 representa os principais elementos do Modelo de Referência de Computação em Nuvem definido pelo NIST. Ela ajudará você a ter uma visão geral do modelo antes de entrarmos em cada tópico.

## Características essenciais

O NIST apresenta 5 características que são essenciais na Computação em Nuvem: Auto-serviço sob demanda, Amplo acesso por rede, Compartilhamento e recursos, Elasticidade rápida e Serviço mensurado. A ideia aqui é, que se algo pode ser chamado de Computação em Nuvem, ela deve ter todas essas características. De fato, você encontrará todas essas características nos principais provedores de nuvem.

1. **Auto-serviço sob demanda:** O usuário ou consumidor pode provisionar por conta própria, de forma autônoma, recursos de computação, como tempo de servidor e armazenamento em rede, automaticamente e

conforme necessário, sem necessitar intervenção humana dos provedores de serviços.

Comentário: note que isso é essencial e característico de como os serviços são fornecidos e consumidos. Você pode abrir sua conta na Amazon AWS ou a Google Cloud e passa a ter acesso a recursos de computação como Máquinas Virtuais e espaço em disco. É um serviço *self-service*, não tem interação com qualquer pessoa dos provedores. É esperado que o usuário seja capaz de encontrar sozinho tudo o que precisa no site que fornece os serviços.

2. **Ampla acesso por rede:** Os recursos estão disponíveis através da rede, vale dizer aqui Internet, e são acessados através de mecanismos padronizados que promovem o uso por dispositivos clientes leves ou ricos de diversas plataformas (como smartphones, tablets, laptops ou desktops).

Comentário: Todos os serviços são disponibilizados através da rede e a Internet é então essencial para os serviços em nuvem. Disso decorre uma série de características dos serviços em nuvem. Por exemplo, diferentemente de sistemas locais de armazenamento em que o acesso ao dado é feito através de operações e protocolo de I/O, muitas operações de acesso a dados podem ou precisarão ser feitas através de protocolos de rede. Você vai conhecer algumas dessas características mais adiante.

3. **Compartilhamento de recursos:** Os recursos de computação do provedor são agrupados em um *pool* de recursos compartilhados para atender a múltiplos usuário em um modo multiusuário, com recursos físicos e virtuais dinamicamente atribuídos e reatribuídos conforme a demanda. Há uma certa independência e transparência de localização geográfica, uma vez que o usuário em geral não controla ou conhece a localização exata dos recursos (como armazenamento, processamento, memória e comunicação de rede), mas pode especificar a localização em um nível de abstração mais alto (como país, região, estado ou datacenter).

Comentário: Esse é um recurso essencial que viabiliza o modelo de negócios de Computação em Nuvem. A ideia é simples e funciona como outros tipos de modelos de negócio com aluguel de recursos: os recursos não são compartilhados todos, ao mesmo tempo por todos usuários. Desse modo um recurso, como capacidade de processamento (ciclos de CPU ou GPU e capacidade de memória), alocada por um usuário, pode ficar disponível para outro usuário quando ele não estiver mais empregando. Para fazer uma analogia simples, imagine os recursos de CPU e memória do seu computador pessoal. E se eles pudessem ser utilizados à noite, enquanto você está dormindo e não está empregando o computador, por outro usuário? Imagine isso em milhares de computadores e você terá uma ideia de como esse mecanismo simples pode viabilizar negócios globais como o dos grandes provedores de nuvem. De fato, provedores de abrangência global fazem uso *time shift* para 'alugar' os mesmos recursos em momentos diferentes – recursos ociosos em um hemisfério durante (noite) podem ser empregados no hemisfério oposto (dia).

4. **Elasticidade rápida:** Os recursos podem ser provisionados e liberados elasticamente, em alguns casos automaticamente, para rapidamente aumentar ou diminuir de acordo com a demanda. Para o usuário os recursos disponíveis para provisionamento passam a ideia de recursos ilimitados e podem ser alocados em qualquer quantidade e a qualquer momento.

Comentário: Essa capacidade permite a expansão e retração dos recursos de acordo com a demanda. Torna o custo dos recursos mais flexível e permite viabilizar novos negócios exponenciais antes inviáveis em uma modelo de computação própria. Não são motivo é que você encontra a maior parte dos novos negócios e start ups empregando serviços em nuvem. Isso permite você iniciar negócios com uma infraestrutura mínima e crescer à medida que os negócios evoluem. Permite que você seja mais resiliente às crises, quando há retração dos negócios. Possibilita atender sazonalidades com períodos de alta demanda, como períodos de campanhas ou datas comemorativas, sem

pagar pelos recursos para o ano todo. Enfim, viabiliza muitos cenários de negócio importantes.

5. **Serviço mensurado:** Os sistemas na nuvem automaticamente controlam e otimizam o uso dos recursos através de medições em um nível apropriado para cada tipo de serviço (como armazenamento, processamento, comunicação de rede e contas de usuário ativas). A utilização de recursos pode ser monitorada, controlada e informada, gerando transparência para o fornecedor o usuário dos serviços.

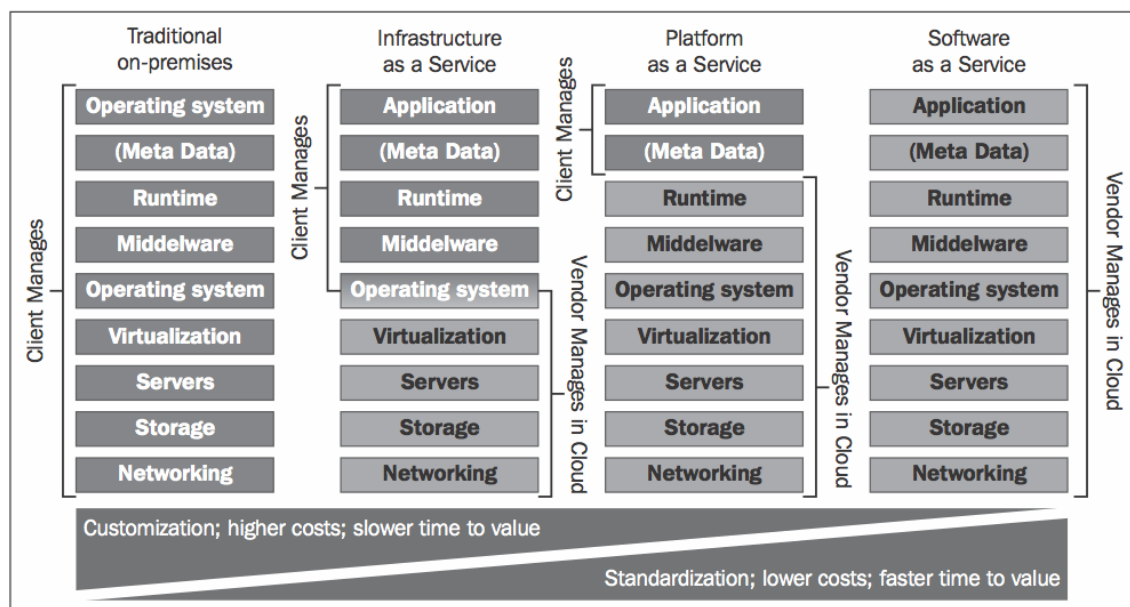
Comentário: Essa característica parece diferir muito das demais, mas é também uma característica essencial. Recursos compartilhados precisam ser gerenciados, e nada que não é medido pode ser gerenciado. É essencial que você meça a quantidade de ciclos de CPU para que saiba qual a capacidade ociosa disponível para outros usuários. É também essa medida que possibilita que os recursos sejam cobrados de acordo com o seu uso.

Mas para deixar claro o caráter ‘essencial’ desse item (e também dos demais), talvez você imagine que se tivermos uma nuvem privada (veremos mais adiante, aqui basta pensar em uma nuvem de uso por uma única empresa) então os serviços não precisariam ser mensurados. Afinal, os recursos são da empresa e não precisariam ser cobrados. Entretanto, mesmo em uma nuvem privada a medida dos serviços é uma característica essencial. A medida, e particularmente a ‘cobrança’ pelo uso dos recursos (mesmo quando não monetária), é o que permite dar eficiência ao uso dos recursos e viabiliza o modelo comercialmente. A ideia é que recursos que não são medidos e não são cobrados de alguma forma, tendem a ficar ociosos mas sem serem disponibilizados para uso por outros usuários. É como depois de ter alocado um recurso, como memória, você dissesse *‘esta memória é minha’* mesmo quando você não estivesse usando. Talvez você diga que não vê por que não liberaria a memória (afinal você é uma pessoa consciente), mas *com certeza* você não vai deixar de lembrar disso se tiver que pagar por esse recurso enquanto ele não estiver sendo usado.



## Modelos de serviço

Outro aspecto importante é o das modalidades de serviço oferecidas. Elas têm um grande impacto em como são feitas as ofertas de serviço e os negócios.



**Figura 2. Exemplos de modalidades de serviço.**

A Figura 2 traz de forma esquemática a distribuição dos serviços (modelo Tradicional *on-premises*) nas diferentes modalidades definidas pelo NIST.

1. **Software como Serviço (SaaS – Software as a Service):** O recurso fornecido ao usuário é o uso de aplicações do fornecedor executando em uma infraestrutura na nuvem. As aplicações podem ser acessadas por vários dispositivos clientes, como um web browser, um App ou um aplicativo cliente. O usuário não gerencia nem controla a infraestrutura subjacente aos serviços como rede, servidores, sistemas operacionais, armazenamento, ou recursos individuais da aplicação.

Comentário: Aqui, para o usuário, a infraestrutura empregada é totalmente transparente. De fato, se você estiver empregando um webmail na nuvem ou um aplicativo ERP, pelo Browser ou App, você não tem qualquer conhecimento ou controle de quantos servidores estão sendo empregados para processar aquele serviço ou mesmo o sistema

operacional ou linguagem que eles empregam. Mas, neste caso, você não está interessado nisso, mas quer enviar e receber os seus e-mails ou incluir e consultar pedidos no ERP.

2. **Plataforma como Serviço (PaaS – Platform as a Service):** Aqui o recurso fornecido permite ao usuário instalar na infraestrutura de nuvem aplicativos criados ou adquiridos pelo usuário, desenvolvidos com linguagens de programação, bibliotecas, serviços e ferramentas suportados pelo provedor de nuvem. O consumidor não gerencia nem controla a infraestrutura na nuvem subjacente, mas tem controle sobre as aplicações instaladas e possivelmente configurações do ambiente de hospedagem de aplicações.

Comentário: Você empregará essa modalidade quando desenvolve aplicações próprias para nuvem. Se você, por exemplo, cria uma aplicação .Net, um aplicativo que emprega uma plataforma de mensageria (tipo MQ, Rabbit etc.) ou algum outro framework, esse recursos precisam ser suportados pelo provedor para que você disponibilize seus serviços.

3. **Infraestrutura como Serviço (IaaS – Infrastructure as a Service):** O recurso fornecido ao consumidor é provisionar processamento, armazenamento, comunicação de rede e outros recursos de computação nos quais o usuário pode instalar e executar softwares em geral, incluindo sistemas operacionais e aplicativos. O consumidor não gerencia nem controla a infraestrutura de nuvem subjacente mas tem controle sobre os sistemas operacionais, armazenamento, aplicativos instalados e alguns componentes de rede (como firewalls).

Comentário: No nível mais baixo você pode contratar recursos para processamento (cpu, memória, armazenamento, rede) e ficar responsável por gerenciar toda as camadas superiores de serviço (SO, aplicativos etc.). É o caso de quando você contrata máquinas Virtuais com uma determinada capacidade de memória e cpu, ou ainda espaço de armazenamento.

## Modelos de Distribuição

Falamos até aqui de provedores como Amazon AWS e Google Cloud. Mas esses provedores representam apenas uma das formas como os serviços de nuvem são distribuídos. Provedores como Amazon AWS e Google Cloud são provedores de Nuvem Pública (atenção, público aqui é no sentido de que qualquer um pode utilizar os serviços, no mesmo sentido que os serviços de água, telefonia e transporte são fornecidos e cobrados por concessionárias privadas) e você pode encontrar outras formas de distribuição.

1. **Nuvem privada:** A infraestrutura na nuvem é provisionada para uso exclusivo por uma única organização composta de diversos consumidores (como unidades de negócio).
2. **Nuvem comunitária:** A infraestrutura na nuvem é provisionada para uso exclusivo por uma determinada comunidade de consumidores de organizações que têm interesses em comum (de missão, requisitos de segurança, políticas, observância de regulamentações).
3. **Nuvem pública:** A infraestrutura na nuvem é provisionada para uso aberto ao público em geral.
4. **Nuvem híbrida:** A infraestrutura na nuvem é uma composição de duas ou mais infraestruturas na nuvem (privadas, comunitárias ou públicas) que permanecem entidades distintas, mas são interligadas permitindo a comunicação de dados e portabilidade de aplicações.

Nos últimos anos uma ênfase maior vem sendo dado às distribuições híbridas, particularmente para grandes empresas e, devido aos custos, é cada vez menos frequente distribuições de comunidade.

## Sumário

Aqui você aprendeu alguns dos principais conceitos sobre Computação e Serviços em Nuvem. Explorou algumas de suas Características Essenciais, e como elas determinam os serviços em nuvem e seu modelo de negócios, e conheceu as diferentes Modalidades de Serviço e de Distribuição.

## Referências

**Cloud computing defined: Characteristics & service levels.** Disponível em: <https://www.ibm.com/blogs/cloud-computing/2014/01/31/cloud-computing-defined-characteristics-service-levels/>. Acesso: 08.02.2021.

**Conceitos de nuvem – Princípios da computação em nuvem.** Disponível em: <https://docs.microsoft.com/pt-br/learn/modules/principles-cloud-computing/>. Acesso: 08.02.2021.

**The NIST Definition of Cloud Computing.** Disponível em: <https://csrc.nist.gov/publications/detail/sp/800-145/final>. Acesso: 08.02.2021.