

INFRAESTRUTURA DE TI

Juliane Adélia Soares



SOLUÇÕES
EDUCACIONAIS
INTEGRADAS



Fundamentos de computação em nuvem

Objetivos de aprendizagem

Ao final deste texto, você deve apresentar os seguintes aprendizados:

- Definir os fundamentos de computação em nuvem.
- Identificar as características essenciais de computação em nuvem.
- Analisar a relação da computação em nuvem com redes, virtualização, armazenamento e *data centers*.

Introdução

A computação em nuvem é uma inovação no setor da tecnologia da informação (TI), representando a solução para problemas de sobrecargas de computação e de necessidade de altos investimentos para montar uma infraestrutura completa nas organizações. Essa solução é apresentada pela oferta de *software* e de infraestrutura como serviços, com base na internet, trazendo mudanças significativas na maneira como o *hardware* pode ser projetado e adquirido.

Este capítulo tem o objetivo de introduzir a computação em nuvem, apresentando seus fundamentos, seus princípios, sua relação com redes, virtualização, armazenamento e *data centers* e, por fim, os benefícios e desafios de adotar um ambiente de serviços em nuvem.

1 Fundamentos

A computação em nuvem (*cloud computing*) refere-se ao fornecimento de serviços de computação via internet, com o intuito de oferecer inovações mais rápidas, recursos flexíveis e economias de escala, os quais incluem servidores, armazenamento, banco de dados, rede e *software*, geralmente pagando-se apenas pelo uso.

De acordo com Erl, Mahmood e Puttini (2013), a ideia de computação em nuvem remete-se às origens da computação de utilidade, conceito proposto pelo cientista da computação John McCarthy em 1961: “Se os computadores do tipo que defendi se tornarem os computadores do futuro, a computação poderá algum dia ser organizada como um utilitário público, assim como o sistema telefônico é um utilitário público. [...] O utilitário de computadores pode se tornar a base de uma nova e importante indústria” (MCCARTHY, 1961 *apud* ERL; MAHMOOD; PUTTINI, 2013, p. 49, tradução nossa).

O cientista Leonard Kleinrock, chefe da Rede de Agências de Projetos de Pesquisa Avançada e do projeto Advanced Research Projects Agency Network (ARPANET), que semeava a internet, declarou em 1969: “A partir de agora, as redes de computadores ainda estão engatinhando, mas à medida que crescerem e se tornarem sofisticadas, provavelmente veremos a disseminação de ‘utilitários de computador’” (KLEINROCK, 1969 *apud* ERL; MAHMOOD; PUTTINI, 2013, p. 49, tradução nossa).

Desde meados dos anos 1990, os usuários já vêm alavancando formas de utilitários de computador com base na internet por meio de mecanismos de buscas, serviços de *e-mail* e mídias sociais, serviços que validaram e popularizaram os conceitos da computação em nuvem atual.

Segundo Bhowmik (2017), o termo “computação em nuvem” teve origem em 1990, quando engenheiros de rede desenhavam diagramas de rede representando diferentes dispositivos e conexões entre eles, e redes externas com o símbolo de uma nuvem, já que desconheciam os detalhes dessas redes — naquele período, isso ficou conhecido como nuvem de rede, ou apenas nuvem.

No final do século XX, as principais empresas de *software* deram início ao fornecimento de aplicativos pela internet, momento em que os serviços de *e-mail* começaram a crescer na medida em que as empresas ofereciam esse recurso aos seus usuários. Em 1999, a Salesforce.com foi a empresa que teve a iniciativa mais notável, oferecendo aplicativos de negócios pela internet. Até então, toda essa evolução fez parte do desenvolvimento de instalações de computação de serviços públicos.

Em 2006, o termo “computação em nuvem” surgiu no mercado com o significado que utilizamos hoje, diferentemente do que se deu em 1990, tendo sido empregado pela primeira vez em um fórum oficial pelo então CEO da Google, Eric Schmidt, durante uma conferência. Durante esse período, diversas empresas, como Amazon, Microsoft e IBM, começaram a divulgar seu próprio serviço de computação em nuvem.

De acordo com Sousa Neto (2015, p. 39):

A arquitetura computação em nuvem significa mudar fundamentalmente a forma de operar a TI, saindo de um modelo baseado em aquisição de equipamentos para um modelo baseado em aquisição de serviços. A computação em nuvem, com a virtualização, teoricamente permite obter o melhor dos mundos: otimização do uso dos recursos e flexibilidade para o usuário.

Antes da computação em nuvem, para terem acesso à computação, os usuários precisavam investir em recursos como *hardware*, *software*, rede e armazenamento, além de dispor de espaço físico para essas instalações, mantê-las e torná-las operacionais: isso tudo gerava um custo muito elevado. A computação em nuvem é uma opção para quem não deseja, ou não pode, ter esse tipo de gasto alto, já que é prático e fácil obter a energia e os recursos necessários de um provedor, além de exigir um investimento razoável, ou seja, optando pela computação em nuvem, o usuário ou as empresas podem obter os recursos de computação como e quando necessário pagando apenas pelo uso. Portanto, Chandrasekaran (2014, p. 10) define que “a computação em nuvem é um mecanismo de levar a contratação ou a obtenção de serviços da capacidade ou infraestrutura de computação a um nível organizacional ou individual, na extensão exigida e pagando apenas pelos serviços consumidos”. Assim, é possível eliminar um grande investimento em infraestrutura sem comprometer o uso da computação no nível do usuário a um custo operacional.

Esse modelo ainda apresenta conveniência e a confiabilidade, visto que, antes, para poder transportar um arquivo, por exemplo, era necessário um *pen drive*, um HD externo ou até mesmo gravá-lo em um CD, e hoje existem aplicativos em nuvem que permitem salvar os arquivos desejados e acessá-los de qualquer local com acesso à internet.

A computação em nuvem está muito presente no dia a dia, mesmo que as pessoas geralmente não saibam disso: *e-mails*, Netflix, redes sociais e aplicativos do Google são alguns exemplos de serviços de nuvem muito utilizados atualmente.

Além disso, trata-se de um modelo de acesso à rede sob demanda, em que se armazenam os dados na nuvem e se pode compartilhar informações e recursos computacionais em segurança, e que promove fácil gestão e interação com o provedor de serviços, visto que utiliza recursos compartilhados no lugar de servidores locais.

O National Institute of Standards and Technology (NIST) define a computação em nuvem como (MELL; GRANCE, 2011, documento *on-line*):

A computação em nuvem é um modelo para permitir acesso onipresente, conveniente e de rede sob demanda a um conjunto compartilhado de recursos de computação configuráveis (por exemplo redes, servidores, armazenamento, aplicativos e serviços) que podem ser rapidamente provisionados e liberados com o mínimo esforço de gerenciamento ou interação do provedor de serviços.

Visto que o termo “nuvem” representa uma metáfora da internet, computação em nuvem significa armazenar e acessar dados e programas pela internet a partir de um local ou computador remoto, o que pode ser feito em qualquer lugar e em qualquer momento. E, como dito, o custo do uso ou da acessibilidade deve se basear no nível de uso, de acordo com as necessidades e demandas dos clientes; quanto maior a necessidade, mais recursos são disponibilizados pelo provedor. Outro ponto importante reside no fato de que a manutenção dos recursos é de responsabilidade do provedor (CHANDRASEKARAN, 2014).

Ainda, a computação em nuvem pode ser definida a partir das perspectivas descritas a seguir.

1. **Como um serviço** — pela computação em nuvem, tudo pode ser disponibilizado como serviço: infraestrutura de TI, plataformas de desenvolvimento, banco de dados, aplicações. Buyya, Vecchiola e Selvi (2013) afirmam que o conceito de “tudo como um serviço” é conhecido como XaaS (X as a Service), em que a letra X pode ser substituída de acordo com o que está sendo ofertado: S para *software*, P para plataforma e I para infraestrutura. Essa abordagem tem um papel significativo, visto influenciar a maneira como o *software* é construído, implementado e disponibilizado, bem como a infraestrutura de TI é projetada e os custos para as necessidades de TI são alocados. Assim, abrange tanto as necessidades de um usuário que hospeda documentos na nuvem quanto as das empresas que decidem implementar parte ou toda a sua infraestrutura na nuvem.
2. **Como uma plataforma** — basicamente, o termo “plataforma” refere-se ao suporte no qual os aplicativos são executados ou dão resultados aos usuários. A internet desempenha um papel fundamental na computação em nuvem, pois por meio dela os serviços são fornecidos e disponibilizados. Alguns aplicativos que eram executados apenas localmente, hoje funcionam também pela nuvem, como o Microsoft 365. Com isso, a *web* tem se tornado uma plataforma.

No modelo de serviços em nuvem, as operações do *data center* são separadas das tarefas de computação do usuário final, residindo em alguma extremidade remota e sendo gerenciados pelo provedor do serviço, o que deixa os usuários livres para se concentrar exclusivamente em suas tarefas específicas.

Considerando as questões econômicas, os usuários pagam taxas de serviço enquanto o usam, sem a necessidade de grandes gastos iniciais de aquisição para construir uma infraestrutura de computação, o que é benéfico para empresas iniciantes ou até mesmo para grandes organizações que desejam fazer alguma mudança nas tecnologias da empresa, obtendo maior flexibilidade, já que pode reduzir e ampliar facilmente sua infraestrutura de TI de acordo com a necessidade, sem precisar investir em infraestrutura física. Essa redução de custos está relacionada à diminuição dos gastos com ativos, servidores, licenças de *software* e energia. De acordo com o NIST, “se a computação em nuvem reduz ou não os custos gerais de uma organização, depende de uma análise cuidadosa de todos os custos de operação, conformidade e segurança, incluindo os custos para migrar e, se necessário, migrar de uma nuvem” (BADGER *et al.*, 2012, p. ES-1, tradução nossa).

Em relação à segurança, muitos clientes podem ficar receosos ao analisar a contratação da computação em nuvem, já que o controle, o armazenamento e a segurança dos dados são de total responsabilidade de terceiros. A esse respeito, o NIST analisa (BADGER *et al.*, 2012, p. ES-2, tradução nossa):

Como sistemas de redes complexos, as nuvens são afetadas por problemas tradicionais de segurança de computadores e redes, como a necessidade de fornecer confidencialidade, integridade e disponibilidade do sistema. Ao imporem práticas de gerenciamento uniformes, as nuvens podem melhorar alguns problemas de atualização e resposta de segurança. As nuvens, no entanto, também têm potencial para agregar uma quantidade e variedade sem precedentes de dados de clientes em *data centers* em nuvem. Essa vulnerabilidade em potencial requer um alto grau de confiança e transparência para que os provedores de nuvem possam manter os dados dos clientes isolados e protegidos. Além disso, os usuários e administradores da nuvem dependem muito dos navegadores da *web*, portanto falhas na segurança do navegador podem levar a violações da segurança da nuvem. A privacidade e a segurança da computação em nuvem dependem principalmente de se o provedor de serviços em nuvem implementou controles de segurança robustos e uma política de privacidade sólida desejada por seus clientes, a visibilidade que os clientes têm de seu desempenho e o quão bem ele é gerenciado.

Em outras palavras, o consumidor deve buscar conhecer todos os controles de segurança implementados pelo provedor antes de contratar o serviço da nuvem, a fim de minimizar os riscos de perda ou vazamento de dados, criando-se uma relação de confiança entre cliente e fornecedor.



Fique atento

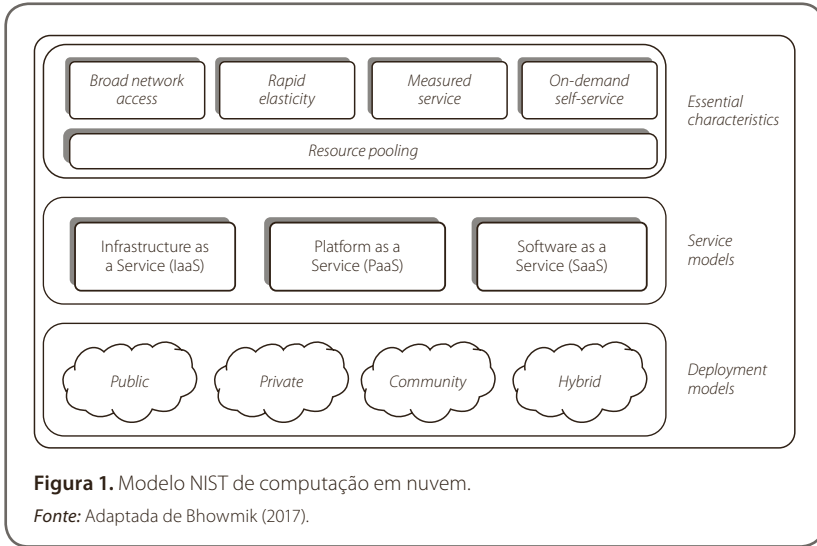
O acordo de nível de serviço (*service level agreement* — SLA) compreende um contrato entre fornecedor de TI e um cliente que define os serviços prestados pelo fornecedor, sendo um dos mais importantes ao se contratar a computação em nuvem. Esse documento pretende gerir as expectativas dos clientes em relação à qualidade do serviço entregue, medindo e validando o cumprimento dos parâmetros previamente acordados e a capacidade do fornecedor em atender a essas expectativas.

Nas cláusulas, devem constar alguns parâmetros considerados cruciais, como disponibilidade do serviço, taxa de transferência de dados, tempo de inatividade do serviço, tempo de resposta, etc. Os contratos de SLA precisam ser personalizados de acordo com os interesses particulares dos clientes, levando-se em conta que, caso alguma das cláusulas do contrato não seja cumprida, devem ser aplicadas penalidades.

É preciso se atentar, pois muitos provedores de serviços não conseguem oferecer SLA realmente substanciais, alguns nem o ofertam e outros oferecem, mas com garantias de serviço inadequadas para muitas organizações. Então, torna-se muito importante analisar a SLA ofertada pelo provedor antes da contratação dos serviços para evitar surpresas.

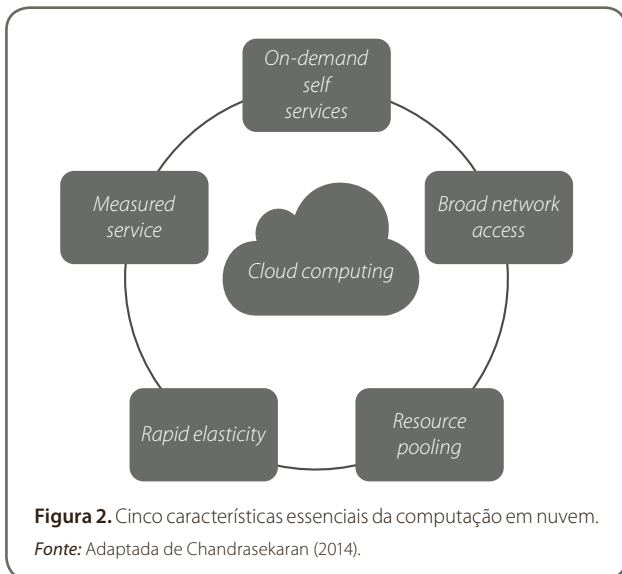
2 Princípios

A definição do NIST estabelece que o modelo de nuvem é composto por cinco características essenciais, três modelos de serviços e quatro modelos de implantação (Figura 1), como visto a seguir.



Características essenciais

Para que um aplicativo seja considerado uma verdadeira implementação de nuvem, deve seguir as cinco características essenciais da Figura 2, definidas pelo NIST.



1. Autoatendimento sob demanda (*on-demand self-service*)

Um consumidor pode provisionar recursos de computação unilateralmente, como tempo do servidor e armazenamento em rede, conforme o necessário e de modo automático, sem exigir interação humana em cada provedor de serviço, ou seja, a implementação do autoatendimento do usuário possibilita que os clientes adquiram e acessem rapidamente os serviços desejados, agilizando a obtenção e o monitoramento de recursos. Ainda, um consumidor pode solicitar e receber acesso a uma oferta de serviço sem que um administrador ou alguém da equipe precise atender à solicitação manualmente. Nos ambientes tradicionais, as solicitações poderiam levar dias ou até mesmo semanas para serem atendidas, o que causava atrasos nos projetos. Já nos ambientes em nuvem, com a implementação do autoatendimento sob demanda, os usuários ficam liberados de solicitar o acesso a recursos adicionais por *e-mail* ou telefone, o que também reduz a carga administrativa para o provedor.

Essas implementações podem ser difíceis de construir, mas são muito vantajosas, tanto para os usuários quanto para os provedores. O próprio usuário pode solicitar serviços conforme sua necessidade por meio de uma interface, normalmente formulários da *web*, a partir do qual *web* os recursos ficam disponíveis instantaneamente. Assim, o serviço de computação em nuvem atende automaticamente aos picos de demanda, sem a necessidade de novos servidores. Também não exige que a alteração do contrato de fornecimento de serviços sempre que o usuário desejar modificar os recursos disponíveis, permitindo que os recursos sejam alocados, liberados ou reconfigurados conforme a demanda.

Um exemplo simples e conhecido desse autoatendimento para usuários finais é o Google Play Filmes, um serviço de *streaming* de filmes da Google: o cliente consegue acessá-lo por meio de usuário e senha, cadastrando seu cartão de crédito para as transações, de qualquer lugar — *smartphone*, *tablet*, *notebooks*, *smart TV* — e podendo escolher o filme que deseja assistir, optando pela compra ou apenas pelo aluguel (para acesso temporário), sendo o valor creditado no cartão informado e com acesso instantâneo à sua escolha. Todo esse processo é possível em virtude do autoatendimento implementado, em que os serviços são escolhidos e utilizados sem interação humana por parte do provedor, facilitando o seu emprego e tornando-os mais eficientes.

2. Amplo acesso à rede (*broad network access*)

Os recursos estão disponíveis na rede e acessados por mecanismos-padrão que promovem o uso por plataformas heterogêneas de *thin* ou *tick client* (p. ex., *smartphones*, *tablets*, *notebooks*, *desktops*). Os usuários têm acesso à nuvem e aos seus recursos de qualquer lugar por meio de um dispositivo conectado à internet.

A computação em nuvem fornece vantagem econômica para os usuários, pois, como as aplicações são acessadas via internet, não se torna necessária a configuração de *data centers* internos caros. Os serviços são desenvolvidos e instalados na estrutura dos provedores e acessados remotamente pelos clientes pela rede. Bhowmik (2017) explica que, para que esse objetivo seja atendido, deve haver uma infraestrutura de rede forte para fornecer os serviços de computação de maneira rápida e sem esforço. Os *links* de comunicação de alta largura de banda espalhados pela área de serviço constituem os atributos essenciais da computação em nuvem.

Porém, fica uma questão: apesar da vantagem de ter acesso ao serviço de qualquer lugar e sempre que necessário, desde que haja algum dispositivo conectado à internet, nem sempre isso se dará de maneira eficiente. No caso de uma largura de banda do usuário baixa, as aplicações podem apresentar lentidão e dificuldade de acesso aos dados.

De todo modo, as vantagens superam essa desvantagem, por exemplo, a possibilidade de algumas empresas oferecerem o trabalho *home office*, visto que utilizam aplicações em nuvem. Então, se o funcionário dispor de uma conexão com a internet e um dispositivo que promova o acesso, conseguirá exercer suas funções fora das dependências da empresa, uma característica que torna a computação em nuvem onipresente.

3. Agrupamento de recursos (*resource pooling*)

Os recursos de computação do provedor são agrupados para atender a vários consumidores usando um modelo de vários inquilinos, com diferentes recursos físicos e virtuais dinamicamente atribuídos e reatribuídos de acordo com a demanda do consumidor. Geralmente, o cliente não tem controle ou conhecimento sobre a localização exata dos recursos fornecidos, o que causa uma sensação de independência em relação à localização, mas ele pode especificar uma localização em um nível mais alto de abstração (país, estado, *data center*). Armazenamento, processamento, memória e rede compreendem alguns exemplos de recursos.

Na computação em nuvem, todos esses recursos de computação são mantidos em locais remotos pelo provedor com a utilização de uma única infraestrutura por vários clientes, devendo ser razoavelmente grandes, flexíveis e capazes de suportar muitos usuários de modo simultâneo, sem falhas, e de distribuir esses recursos de acordo com a necessidade de cada usuário.

Essa característica ajuda a economizar custos e permite maior flexibilidade no lado do provedor, pois se baseia no fato de que os clientes não terão uma necessidade constante de dispor de todos os recursos disponíveis para eles; então, quando os recursos não estão sendo empregados por um cliente, podem ser utilizados por outro, de modo que não fiquem ociosos, possibilitando que os provedores atendam a um maior número de usuários, o que não seria possível se cada um dos clientes exigisse recursos dedicados (ROUNTREE; CASTRILLO, 2014).

4. Rápida elasticidade (*rapid elasticity*)

Os recursos podem ser provisionados elasticamente e liberados, em alguns casos de modo automático, para escalar rapidamente para expansão e redução, conforme a demanda. Para o consumidor, os recursos disponíveis para provisionamento geralmente parecem ilimitados e podem ser apropriados em qualquer quantidade e a qualquer momento, ou seja, um ambiente de nuvem tem a capacidade de se ampliar facilmente para satisfazer à demanda do usuário.

Geralmente, a rápida elasticidade é obtida pelo uso de automação, pela qual se pode aumentar a eficiência dos negócios com a redução ou substituição de interação humana por sistemas de TI, de orquestração, que corresponde à configuração, ao gerenciamento e à coordenação automatizada de serviços, aplicações e sistemas de computação. Quando o uso de recursos atinge determinado ponto, a capacidade se expande; quando o uso diminui, a capacidade diminui novamente, conforme o necessário, para garantir que os recursos não sejam desperdiçados. Por meio de sistemas escalonáveis, os provedores conseguem criar uma sensação de recursos infinitos ou imediatamente disponíveis, por exemplo, caso os clientes precisem de maior armazenamento ou memória, o provedor deve estar preparado para a alocação desses recursos de acordo com a demanda.

As implantações em nuvem já devem dispor de uma infraestrutura necessária para expandir a capacidade de serviço, o que significa que as aplicações também precisam ser elásticas. Se o sistema for projetado corretamente, será possível adicionar mais recursos de computação. Com a implantação da elasticidade, apesar de os recursos estarem disponíveis, não são utilizados até que

sejam necessários, o que também possibilita uma economia para o provedor em relação aos custos de consumo, o que inclui energia e refrigeração.

E, como o provedor pode não saber quando e a quantidade de recursos que os usuários necessitarão, é extremamente importante que o ambiente da computação crie uma impressão de repositório ilimitado de recursos, para que o usuário consuma qualquer volume de recurso a qualquer momento. Rountree e Castrillo (2014, p. 5) afirmam que “o recurso de rápida elasticidade das implementações em nuvem é o que lhes permite lidar com a capacidade de ‘explosão’ necessária para muitos de seus usuários. A capacidade de ruptura é uma capacidade necessária apenas por um curto período de tempo”.

Por exemplo, uma companhia aérea faz uma promoção “relâmpago”, com passagens a preços baixíssimos. Nesse curto período, o acesso ao sistema da companhia e as compras das passagens têm um crescimento elevado. Então, para que tudo funcione sem maiores problemas, o provedor precisa expandir os recursos utilizados por essa aplicação para que a carga seja suportada. Ao final da promoção, quando os acessos retornam ao seu fluxo normal, os recursos reduzem proporcionalmente. Com a utilização da computação em nuvem, não é mais necessário ter essa capacidade disponível internamente e por tempo integral.

5. Serviço medido (*measured service*)

Os sistemas em nuvem controlam e otimizam automaticamente o uso de recursos, aproveitando um recurso de medição em algum nível de abstração apropriado ao uso de serviço. O uso de recursos pode ser monitorado, controlado e relatado, fornecendo transparência ao provedor e ao consumidor do serviço utilizado, normalmente com base no pagamento ou na cobrança por uso.

O provedor deve empregar algum mecanismo para medir o consumo real de cada usuário ou organização, em que o uso do *pool* de recursos seja calculado e cobrado de cada usuário com base apenas no consumo real de recursos da nuvem. Cada componente é medido para fornecer um produto configurado com precisão ao consumidor, como o desempenho da CPU, a utilização da memória, os aplicativos disponíveis, a largura de banda da rede e qualquer aspecto da solução na nuvem. Depois que se identifica uma métrica (método de medição) apropriada, determina-se uma taxa, por meio da qual será determinado quanto um cliente será cobrado. Assim, os usuários não são mais cobrados por nenhum recurso ocioso, e sim com base nos seus níveis de consumo dos serviços disponibilizados pelo provedor. É muito importante que

o cliente conheça e entenda quais serviços estão sendo medidos e cobrados para que não tenha nenhuma surpresa indesejada, como custos inesperados.

Como a computação em nuvem se fundamenta em cargas de trabalho variadas, torna-se econômica em virtude do serviço medido e do modelo de pagamento por uso envolvido. Pense em uma empresa que faz uso de *e-commerce*, vendas *on-line* de roupas, calçados e acessórios femininos: geralmente, a demanda de vendas segue um padrão, embora próximo a datas comemorativas, como o Dia das Mães, as vendas tendam a crescer muito, o que significa que os recursos de que a empresa necessita nessa época são superiores aos utilizados normalmente (maior processamento, largura de banda, memória, etc.). Então, em vez de cobrar um valor fixo por essa utilização, já que na maior parte do tempo os recursos disponíveis ficam ociosos, as taxas são cobradas de acordo com o aumento da demanda e a utilização dos recursos entregues pelo provedor. Do mesmo modo, se a empresa tiver uma queda brusca de vendas e a demanda do sistema diminuir, causando a diminuição de utilização de recursos de computação, o valor cobrado pelos serviços também diminuirá.

A proposta da computação em nuvem consiste em criar uma ilusão de que o recurso computacional é infinito e, ao mesmo tempo, permitir a eliminação do comprometimento antecipado da capacidade, além de possibilitar o pagamento pelo uso real dos recursos (SOUSA NETO, 2015).

Modelos de serviços

Os recursos de computação em nuvem então disponíveis para os usuários em três modelos diferentes, conforme descrito a seguir.

1. *Software como serviço (SaaS — software as a service)*

Nesse modelo, a capacidade fornecida ao consumidor é de utilizar os aplicativos do provedor em execução em uma infraestrutura de nuvem, acessíveis a partir de dispositivos com acesso à internet, como um navegador da *web* ou uma interface de programa. O usuário não gerencia ou controla a infraestrutura de nuvem subjacente, incluindo rede, servidores, sistemas operacionais, armazenamento ou até mesmo recursos de aplicativos individuais, sendo possíveis apenas definições limitadas de configurações de aplicativos específicos do usuário. O público-alvo desse modelo é o usuário final. São exemplos a Salesforce.com e o Gmail.

2. Plataforma como serviço (PaaS — *platform as a service*)

A capacidade fornecida ao consumidor nesse modelo de serviço consiste em implantar na infraestrutura de nuvem os aplicativos criados ou adquiridos pelo usuário, a partir de linguagens de programação, bibliotecas, serviços e ferramentas suportadas pelo provedor. O consumidor não é responsável por gerenciar nem por controlar a infraestrutura de nuvem subjacente, incluindo rede, servidores, sistemas operacionais ou armazenamento, além, possivelmente, das definições de configuração para o ambiente de hospedagem de aplicativos. O público-alvo desse modelo é o desenvolvedor. São exemplos de provedores de PaaS o Google App Engine e a Microsoft Azure Services.

3. Infraestrutura como serviço (IaaS — *infrastructure as a service*)

Nesse modelo, o usuário deve fornecer processamento, armazenamento, redes e outros recursos computacionais fundamentais, nos quais consiga implantar e executar um *software* arbitrário, que pode incluir sistemas operacionais e aplicativos. O consumidor não é responsável pelo gerenciamento e pelo controle da infraestrutura de nuvem subjacente, mas tem controle sobre sistemas operacionais, armazenamentos e aplicativos implantados, além, possivelmente, do controle limitado de componentes de rede selecionados, como *firewalls* de *host*. O público-alvo desse modelo é o arquiteto de TI. É um exemplo de provedor IaaS a Amazon Web Services (AWS).

Analisando a relação entre os três modelos, o IaaS fornece recursos computacionais para o PaaS, que, por sua vez, oferece recursos, tecnologias e ferramentas para desenvolvimento e execução dos serviços implementados a serem disponibilizados como SaaS. Apesar dessa interação, um provedor não precisa disponibilizar os três modelos obrigatoriamente.

Modelos de implantação

Descrevem as maneiras pelas quais os serviços em nuvem podem ser implantados ou disponibilizados para seus clientes, de acordo com a estrutura organizacional e o local de provisionamento. Segundo a definição do NIST de computação em nuvem, são descritos quatro modelos diferentes, como veremos a seguir.

1. Nuvem privada (*private cloud*)

A infraestrutura de nuvem é provisionada para uso exclusivo por uma única organização, composta por vários consumidores (p. ex., unidades de negócios). Pode ser de propriedade, gerenciada e operada pela organização, por terceiros ou por uma combinação deles, existindo dentro ou fora das instalações.

2. Nuvem comunitária (*community cloud*)

A nuvem comunitária é fornecida para uso exclusivo por uma comunidade específica de consumidores de organizações que compartilham as mesmas preocupações quanto, por exemplo, a missão, requisitos de segurança, política e considerações de conformidade. Esse modelo pode ser de propriedade, gerenciado e operado por uma ou mais organizações da comunidade, terceiros ou uma combinação delas, existindo dentro ou fora das instalações.

3. Nuvem pública (*public cloud*)

Nesse modelo, a infraestrutura de nuvem é fornecida para uso aberto pelo público em geral. Pode ser de propriedade, gerenciada e operada por uma empresa, organização acadêmica ou governamental, ou alguma combinação deles, dentro das instalações do provedor de nuvem.

4. Nuvem híbrida (*hybrid cloud*)

Esse modelo é uma composição de duas ou mais infraestruturas de nuvens distintas (privadas, comunitárias ou públicas) que permanecem entidades únicas, mas são unidas por tecnologia padronizada ou proprietária que permite a portabilidade de dados e aplicativos.

3 Computação em nuvem: relações, benefícios e desafios

A seguir, trataremos das relações da computação em nuvem com *data center*, rede, armazenamento e virtualização, além dos benefícios e desafios de adotar o modelo de serviços em nuvem.

Computação em nuvem e suas relações

Como seu nome já diz, o *data center* é o centro de dados das empresas. No modelo de computação tradicional, esse departamento e todos os sistemas residiam em um local físico. Com a inovação da computação e o crescimento do modelo de computação em nuvem, os *data centers* se tornaram o centro de dados de um provedor, uma central na qual se concentram todos os equipamentos e recursos essenciais para manter as soluções virtualizadas sempre disponíveis, íntegras e seguras. A TI é centralizada em grandes pontos de armazenamento e processamento, mas conserva a estrutura de interligação em redes. Sousa Neto (2015, p. 40) explica que “a nuvem, na verdade, é um conjunto de grandes pontos de armazenamento, processamento de dados e provimento de informação”.

Quando usada como *data center*, a nuvem oferece uma escala muito maior de gestão de serviços, pois a ideia central consiste em oferecer níveis de serviços de acordo com a criticidade do aplicativo, buscando atender à demanda de negócios de maneira efetiva, reduzir custos e flexibilizar o negócio. A centralização da nuvem em *data center* é possível em razão da atual oferta de banda. Com isso, as organizações centralizam o processamento e o armazenamento de dados em grandes estruturas, propiciando ganho de escala, ao mesmo tempo que obtêm flexibilidade pela integração em rede. Essa centralização e a implementação de políticas de infraestrutura de rede, o que inclui a segurança dos dados, permitem que os provedores forneçam níveis adequados de serviços para as empresas clientes. Já os *data centers* e suas conexões formam a infraestrutura da nuvem.

Na computação em nuvem, utiliza-se uma rede de servidores remotos, hospedados na internet, com os objetivos de armazenar, gerenciar e processar dados. A rede baseada na nuvem permite que a infraestrutura, os componentes associados a ela e os usuários externos se comuniquem entre si, tendo um papel fundamental no ambiente em nuvem e sendo a responsável pela mudança na maneira como os serviços de TI são fornecidos, já que é por meio dela que os usuários têm acesso aos recursos computacionais do fornecedor dos serviços e às aplicações contratadas a qualquer momento e de acordo com a necessidade. Para o cliente, é como se estivesse localmente, com memória, armazenamento e processamento individuais, porém os componentes, os dispositivos e as operações estão na nuvem, podendo ser facilmente expandidos ou reduzidos conforme a necessidade, assim como as aplicações configuradas diretamente via *web*, sem precisar de interações humanas do provedor. A conectividade, a segurança, a administração e o controle da rede são hospedados na nuvem e entregues aos clientes como um serviço.

Na computação em nuvem, os dados são armazenados fisicamente nos servidores dos provedores, e não mais em discos rígidos locais; assim, estão disponíveis a qualquer hora e em qualquer local, podendo ser acessados via internet desde que o usuário disponha das informações de *login* e senha. Esse recurso é vantajoso tanto para as empresas, que não ficam presas ao ambiente local e têm maior flexibilidade para executar suas demandas, quanto para os usuários finais, que não precisam se preocupar em armazenar seus arquivos em HD externos ou *pen drive* para poder acessar em outros lugares. Isso também confere maior segurança aos dados, pois, se estiverem salvos localmente e o dispositivo utilizado para acesso apresentar algum problema, todas as informações podem ser perdidas. Outro ponto a se considerar em relação ao armazenamento em nuvem reside na escalabilidade oferecida pelos fornecedores dos serviços, pois o armazenamento é disponibilizado de acordo com a demanda do cliente — por exemplo, caso a empresa necessite de um maior armazenamento, o recurso é disponibilizado pelo provedor sem que a empresa precise investir em infraestrutura interna.

O *pool* de recursos, uma das características essenciais da computação em nuvem, é criado no *data center*, e uma camada de abstração é construída sobre os *pools* de vários tipos de recursos físicos utilizando a virtualização. Todos os recursos do *data center* são virtualizados e os consumidores de serviços em nuvem têm acesso apenas aos recursos de computação virtual.

A virtualização do *data center* representa uma das bases da computação em nuvem e influencia o dimensionamento dos demais serviços oferecidos. A partir dessa tecnologia, é criado o modelo de múltiplos inquilinos para um ambiente em nuvem, ou seja, ela otimiza e flexibiliza o compartilhamento de uma única infraestrutura física de um provedor entre vários usuários, aumentando a utilização de recursos e oferecendo suporte para que os provedores de serviço atendam a um número maior de clientes. Ela fornece um caminho para que o sistema em nuvem gerencie recursos com eficiência, em razão do mapeamento dinâmico dos recursos virtuais para os recursos físicos subjacentes. Bhowmik (2017) explica que o comportamento dinâmico do sistema se dá porque o mapeamento dos recursos de computação pode ser realizado facilmente com base nas mudanças de condições e requisitos, o que significa que o sistema mede de maneira automatizada o uso dos recursos e faz a distribuição e a redistribuição de armazenamento, processamento, memória e largura de banda de acordo com a necessidade e a utilização variável de cada um dos clientes. O processo é rápido, pois qualquer alteração pode ser realizada instantaneamente. E o acesso aos recursos agrupados é fornecido usando a virtualização de recursos.

Ainda, de acordo com uma pesquisa realizada pela empresa de TI Gartner, a virtualização constitui o principal facilitador da maioria dos atributos fundamentais da computação em nuvem. O compartilhamento dos recursos é possível pela virtualização, já que os usuários não conhecem os recursos físicos reais e não podem ocupar nenhuma unidade de recurso específica enquanto não realizam um trabalho produtivo. As máquinas e demais recursos virtuais exibem elasticidade e flexibilidade, visto que, com os recursos virtualizados, a capacidade subjacente dos recursos reais pode ser alterada para atender às cargas de serviço variáveis. Por fim, a medição precisa do consumo de recursos é possível pelo uso de recursos virtualizados.

Benefícios da computação em nuvem

A computação em nuvem representou muitas mudanças para o escopo da computação: ela é ofertada como serviço e sob demanda, pode ser medida em volume ou tempo de uso, os recursos são gerenciados pelos provedores e tem a flexibilidade de permitir que os usuários tenham o máximo e o mínimo sempre que necessário. Segundo Bhowmik (2017) e Chandrasekan (2014), seus benefícios são os descritos a seguir.

Acessibilidade

Todos os arquivos que estiverem salvos na nuvem podem ser acessados e usados a qualquer momento, em qualquer lugar, por meio de um dispositivo com acesso à internet.

Poder e recursos de computação

A computação em nuvem permite estabelecer com precisão a quantidade certa de poder e recursos de computação a serem usados para aplicativos. Os provedores do serviço fornecem serviços relacionados à computação como um pacote de poder de computação e os distribuem sob demanda.

Pagamento apenas pelo que for usado

Os clientes utilizam apenas a partir do que necessitam e são cobrados pelo tempo de uso e pelo que foi utilizado, o que gera economia para o cliente.

Escalabilidade

Quando for preciso um maior poder de computação, os serviços de nuvem conseguem fornecer acesso instantâneo exatamente ao que é requisitado e retornar ao menor uso assim que não haja mais necessidade.

Redução do custo de aquisição

Como implementação da infraestrutura física e a aquisição de *software* são de responsabilidade do provedor, os consumidores têm uma redução drástica no investimento inicial, pois só precisam organizar os sistemas clientes para acessar os serviços de nuvem.

Redução de custos com infraestrutura interna

Manter servidores, *data centers* e toda a infraestrutura de TI em funcionamento 24 horas por dia é de responsabilidade total do provedor, garantindo redução de custos aos clientes, visto não haver despesas de capital.

Redução da responsabilidade do gerenciamento do sistema

Na computação em nuvem, a maioria das tarefas de infraestrutura e gerenciamento de sistema é de responsabilidade do fornecedor do serviço. Assim, o usuário pode se concentrar exclusivamente em suas atividades, sem se preocupar com as camadas de computação subjacentes.

Disponibilidade

A tecnologia da virtualização permite que o *software* de nuvem de um fornecedor mova dados automaticamente de um *hardware* com defeito ou seja colocado *offline* para uma seção do sistema ou *hardware* que está funcionando ou operacional; portanto, minimiza a possibilidade de interrupção do serviço.

Confiabilidade

Sistemas de *backup* separados, com estratégias de recuperação de desastres na nuvem, fornecem outra camada de segurança e confiabilidade, sendo de responsabilidade do fornecedor.

TI verde

Empresas que adotam a computação em nuvem têm vários benefícios em relação à sustentabilidade pelo fato de essa tecnologia não exigir muitos equipamentos próprios, economizando energia e espaço, com menor necessidade de sistemas de refrigeração e um número menor de equipamentos obsoletos para descarte, o que reduz o lixo eletrônico, trazendo benefícios para o meio ambiente.

Rápida implantação

É de rápida implantação pelo provisionamento de recursos ser rápido e automático no ambiente em nuvem.

Atualizações automáticas de *software*

No ambiente da computação em nuvem, as aplicações são atualizadas automaticamente. Os fornecedores de nuvem sempre entregam a versão mais recente disponível dos *software* para os seus usuários.

Sem aquisição de licença

Na computação em nuvem, os usuários não precisam adquirir nenhuma licença periódica para usar aplicativos.

Desafios

Apesar dos diversos benefícios apresentados pela computação e nuvem, ela também traz desafios, como os descritos por Bhowmik (2017) e Chandrasekaran (2014).

Dependência da internet

Caso a conexão com a internet se perca, perde-se o *link* com a nuvem e, por consequência, os dados e aplicativos ficam inacessíveis.

Portabilidade limitada entre provedores de nuvem

Diferentes fornecedores criam recursos de computação em nuvem para uso público, proprietário de várias extensões. Torna-se difícil mudar aplicativos desenvolvidos nessas nuvens proprietárias para outra plataforma em nuvem pelo bloqueio dos fornecedores, o que limita a portabilidade de aplicações.

Segurança de dados

A segurança também representa uma preocupação, pois todo o trabalho com dados e aplicativos depende do poder da computação dos provedores.

Controle reduzido

A computação em nuvem suporta a escalabilidade, porém não permite o controle desses recursos, já que estes não são de propriedade dos clientes.

Problema de interoperabilidade

Outra grande barreira reside na interoperabilidade de aplicativos, a capacidade de dois ou mais aplicativos necessários para suportar uma necessidade comercial de trabalharem em conjunto, compartilhando dados e outros recursos relacionados a negócios. Na nuvem, isso pode não acontecer, já que esses aplicativos podem estar disponíveis em dois fornecedores diferentes e não cooperarem entre si, visto seguirem padrões diferentes.

Conformidade multirregional e questões legais

Os fornecedores de computação em nuvem constroem *data centers* em locais que lhes convêm, tanto geográficos quanto econômicos. Como os clientes acessam pela internet a computação em nuvem, podem não estar cientes da localização real dos recursos que consomem, o que acarreta preocupações legais, já que esses *data center* podem estar localizados até mesmo em outros países ou regiões, com regras diferentes em relação à privacidade ou à conformidade dos dados armazenados. A maioria das estruturas regulatórias reconhece organizações de consumidores em nuvem responsáveis pela segurança, pela integridade e pelo armazenamento de dados, mesmo quando são mantidos por fornecedores externos de nuvem.

Custo da largura de banda

Embora não se trate de um grande desafio, não podemos deixar de destacar que, embora o modelo da computação em nuvem de pagar apenas pelo uso promova uma redução de custos, o custo da largura de banda da rede é uma despesa adicional na computação em nuvem.

A computação em nuvem ganha cada vez mais espaço tanto nas organizações quanto para os usuários finais, estando presente em diversas aplicações na era de internet, imediatismo, necessidade de facilidades e redução de custos que vivemos hoje. Segundo o NIST, o conceito de computação apresenta cinco características essenciais para que uma aplicação seja considerada uma verdadeira implementação de nuvem, além de seus modelos de serviços e implantações. Deve-se definir o modelo de implantação de acordo com as estruturas e as necessidades da organização que deseja aderir aos serviços de nuvem.

Um estudo de desempenho global sobre computação em nuvem realizado pela Business Software Alliance (BSA) com 24 países que lideram o mercado de TI mundial mostra que o Brasil tem investido muito em tecnologia na nuvem, pois passou da 22ª posição em 2016 para a 18ª em 2018 (WAKKA, 2018, documento *on-line*). Embora ainda existam desafios, os benefícios da computação em nuvem com certeza superaram as dificuldades encontradas, o que deve ser levado em consideração na hora de decidir quanto à sua adoção.



Referências

BADGER, L. *et al.* *Cloud computing synopsis and recommendations: recommendations of the National Institute of Standards and Technology*. Gaithersburg: Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, 2012. 81 p. (Special Publication 800-146). Disponível em: <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-146.pdf>. Acesso em: 26 maio 2020. ERL, T.; MAHMOOD, Z.; PUTTINI, R. *Cloud computing: concepts, technology & architecture*. Upper Saddle River: Prentice Hall, 2013. 487 p.

MELL, P.; GRANCE, T. The NIST Definition of Cloud Computing. *Computer Security Resource Center/NIST*, Gaithersburg, Sep. 2011. Disponível em: <https://csrc.nist.gov/publications/detail/sp/800-145/final>. Acesso em: 26 maio 2020.

ROUNTREE, D.; CASTRILLO, I. *The basics of cloud computing: understanding the fundamentals of cloud computing in theory and practice*. Waltham: Syngress, 2014. 172 p.

SOUSA NETO, M. V. *Computação em nuvem*. Rio de Janeiro: Brasport, 2015. 192 p.

WAKKA, W. Brasil sobe quatro posições no ranking de computação em nuvem. *Canaltech*, São Bernardo do Campo, 9 mar. 2018. Disponível em: <https://canaltech.com.br/computacao-na-nuvem/brasil-sobe-quatro-posicoes-no-ranking-de-computacao-em-nuvem-109594/>. Acesso em: 26 maio 2020.



Fique atento

Os *links* para *sites* da *web* fornecidos neste capítulo foram todos testados, e seu funcionamento foi comprovado no momento da publicação do material. No entanto, a rede é extremamente dinâmica; suas páginas estão constantemente mudando de local e conteúdo. Assim, os editores declaram não ter qualquer responsabilidade sobre qualidade, precisão ou integralidade das informações referidas em tais *links*.

Encerra aqui o trecho do livro disponibilizado para esta Unidade de Aprendizagem. Na Biblioteca Virtual da Instituição, você encontra a obra na íntegra.

Conteúdo:



SOLUÇÕES
EDUCACIONAIS
INTEGRADAS