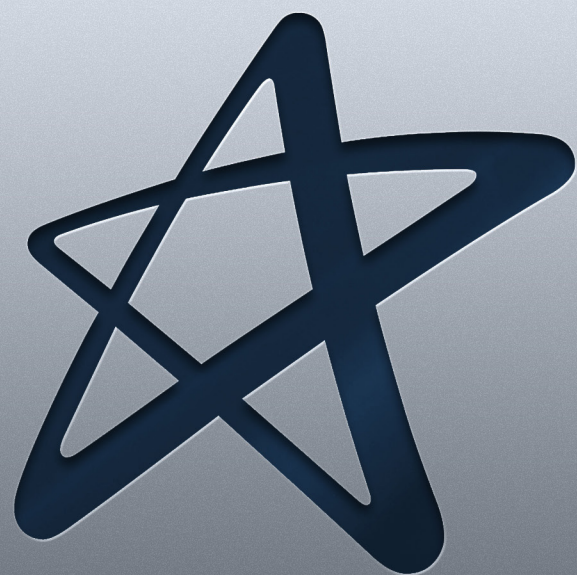


Big Data



Cruzeiro do Sul Virtual
Educação a distância

Material Teórico



Big Data com Computação em Nuvem

Responsável pelo Conteúdo:

Prof. Dr. Alberto Messias

Revisão Textual:

Prof.^a Dr.^a Selma Aparecida Cesarin

UNIDADE

Big Data com Computação em Nuvem



- O que é Computação em Nuvem e a Relação com *Big Data*;
- Características Essenciais à Computação em Nuvem;
- Benefícios da Computação em Nuvem;
- Modelos de Computação em Nuvem e *Big Data*.



OBJETIVO DE APRENDIZADO

- Compreender a relação entre Big Data e Computação em Nuvem;
- Compreender o que é Computação em Nuvem, suas características essenciais, os benefícios do uso e os modelos de Computação em Nuvem.



Orientações de estudo

Para que o conteúdo desta Disciplina seja bem aproveitado e haja maior aplicabilidade na sua formação acadêmica e atuação profissional, siga algumas recomendações básicas:



Assim:

- ✓ Organize seus estudos de maneira que passem a fazer parte da sua rotina. Por exemplo, você poderá determinar um dia e horário fixos como seu “momento do estudo”;
- ✓ Procure se alimentar e se hidratar quando for estudar; lembre-se de que uma alimentação saudável pode proporcionar melhor aproveitamento do estudo;
- ✓ No material de cada Unidade, há leituras indicadas e, entre elas, artigos científicos, livros, vídeos e *sites* para aprofundar os conhecimentos adquiridos ao longo da Unidade. Além disso, você também encontrará sugestões de conteúdo extra no item **Material Complementar**, que ampliarão sua interpretação e auxiliarão no pleno entendimento dos temas abordados;
- ✓ Após o contato com o conteúdo proposto, participe dos debates mediados em fóruns de discussão, pois irão auxiliar a verificar o quanto você absorveu de conhecimento, além de propiciar o contato com seus colegas e tutores, o que se apresenta como rico espaço de troca de ideias e de aprendizagem.

O que é Computação em Nuvem e a Relação com *Big Data*

A adoção da Computação em Nuvem está aumentando de forma significativa nas Organizações; ela é vista como uma Tecnologia de ponta “disruptiva” na próxima década.

A Nuvem está dirigindo otimização e inovação dos modelos de negócios nas Organizações, permitindo tendências como mobilidade, *Big Data* e Mídias Sociais, que também estão influenciando a adoção da Nuvem.

A Computação em Nuvem trata-se de uma coleção de recursos de TI acessíveis por meio da Rede; consiste, essencialmente, em *pools* compartilhados de recursos de *hardware* e *software* instalados em *datacenters*.

O Modelo de Nuvem permite que os consumidores contratem um provedor de recursos de TI como um serviço, com um modelo que permite conveniência, acesso ubíquo sob demanda, com *pools* compartilhados de recursos computacionais configuráveis, (por exemplo, servidores, armazenamento, Redes, aplicações e serviços), que podem ser rapidamente provisionados e liberados com um esforço mínimo de gerenciamento ou prestador de serviços interação.

O modelo de Nuvem é semelhante a um serviço de fornecimento, por exemplo, de energia elétrica, no qual um consumidor conecta um aparelho elétrico a uma tomada e o liga. O cliente não necessariamente precisa conhecer como a eletricidade é gerada ou distribuída e deve pagar somente pela quantidade usada.

Essa analogia pode ser aplicada à Computação em Nuvem ou à abstração dos serviços de TI contratados: os clientes pagam apenas pelos serviços que utilizam, com base no consumo de recursos.

Muitas Organizações agora vêem a Nuvem como uma extensão de sua estratégia de aquisição de recursos de TI e pode se tornar a forma predominante como as Organizações vão passar a adquirir e a usar a Tecnologia de Computação no futuro.

Por meio da Computação em Nuvem, Empresas pequenas podem obter recursos de TI necessários e competir, o que antes não era possível, pois eram caras, chegando muitas vezes a um custo proibitivo para o negócio.

A Figura a seguir ilustra as características iniciais da Computação em Nuvem.

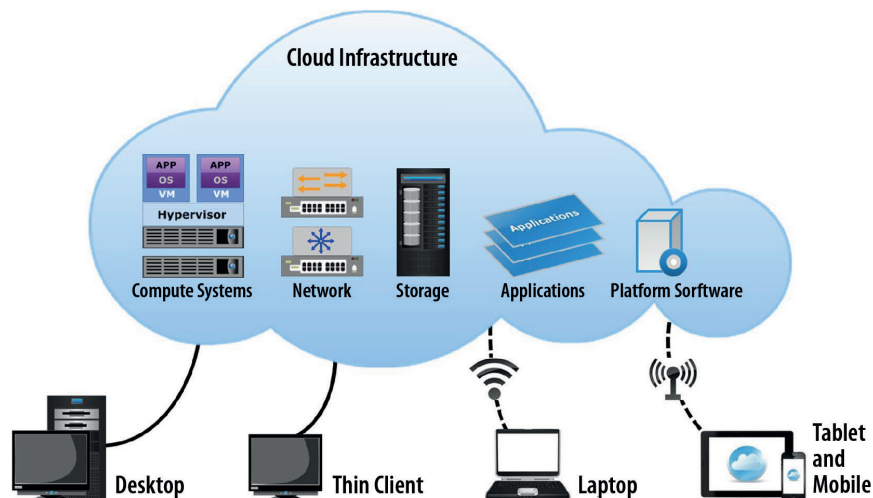


Figura 1

A Figura ilustra um ambiente hipotético de Computação em Nuvem, no qual vários tipos de serviços em Nuvem são acessados por clientes de diferentes dispositivos e em diferentes tipos de Rede.

O termo “Nuvem” se origina da bolha semelhante à Nuvem, que é comumente usada em diagramas de Arquitetura Técnica para representar um Sistema, como a *Internet*, uma Rede ou um *cluster* de computação.

Uma infraestrutura de computação só pode ser classificada como uma Nuvem se tiver algumas características essenciais específicas.

Características Essenciais à Computação em Nuvem

A Computação em Nuvem possui algumas características principais. São elas: *On-demand self-service*, acesso amplo à Rede, *pool* de recursos, elasticidade rápida e o serviço medido, descritos a seguir:

- **On-demand Self-service:** os consumidores usam um portal de autosserviço baseado na *web* para visualizar um catálogo de serviços e solicitar serviços em Nuvem. Isso permite aos consumidores o provisionamento de serviços em

Nuvem de uma forma simples e flexível e reduz o tempo necessário para a implantação de recursos novos ou adicionais de TI. Um consumidor pode unilateralmente provisionar recursos de computação, tais como tempo de servidor e armazenamento em Rede, conforme sua necessidade e, automaticamente, sem a necessidade de interação humana com cada provedor de serviço;

- **Acesso através de Rede:** os consumidores podem acessar os serviços em Nuvem de qualquer dispositivo, cliente ou lugar, através da *Internet* e clientes que implementem o modelo *OSI* e protocolos *TCP/IP*, clientes de qualquer plataforma, como celulares, *tablets*, *laptops* e estações de trabalho;
- **Pooling de Recursos:** permite aos provedores, de modo a melhorar a utilização de recursos e à prestação de serviços de forma flexível, recuperar recursos de computação, que estão reunidos para servir a múltiplos clientes, usando um modelo *multi-tenant*, com diferentes recursos físicos e virtuais atribuídos e realocados dinamicamente de acordo com a demanda do cliente. Há um senso de independência local em que o cliente, geralmente, não tem controle ou conhecimento sobre a localização exata dos recursos disponibilizados, mas pode ser capaz de especificar o local num nível mais alto de abstração (por exemplo, país, estado ou *datacenter*). Exemplos de recursos incluem armazenamento, processamento, memória e largura de banda de Rede;
- **Elasticidade Rápida:** os consumidores podem se adaptar a variações em cargas de trabalho e manter os níveis de desempenho requeridos. Os clientes podem ser capazes de evitar custos excessivos, ou excesso de provisionamento de recursos, pois esses recursos podem ser elasticamente provisionados e liberados, em alguns casos, automaticamente, para escalar rapidamente recursos mensuráveis de acordo com a demanda. Para o consumidor, os recursos disponíveis para realizar o provisionamento, muitas vezes, parecem ser ilimitados e podem ser selecionados ou provisionados em qualquer quantidade, a qualquer momento;
- **Serviço Medido:** os Sistemas em Nuvem controlam e otimizam o uso de recursos automaticamente, por meio de um recurso de medição em nível de abstração, apropriado ao tipo de serviço (por exemplo, armazenamento, processamento, largura de banda e contas de usuário ativas). O uso de recursos pode ser monitorado, controlado e relatado, de maneira transparente, tanto para o provedor, quanto para o consumidor do serviço.

Uma infraestrutura de Nuvem possui um Sistema de Medição que gera contas para os consumidores com base nos serviços utilizados. O Sistema de Medição monitora continuamente a utilização de recursos pelo consumidor e fornece diversos relatórios quanto ao uso de recursos. Por exemplo, o Sistema de Medição monitora a utilização do tempo do processador, da largura de Banda da Rede e da capacidade de armazenamento. Ele fornece informações sobre a demanda atual na Nuvem e ajuda os provedores com planejamento de capacidade e serviços. O monitoramento do uso de recursos ajuda a identificar quando recursos adicionais precisam ser provisionados dinamicamente para atender às cargas dos clientes.

Benefícios da Computação em Nuvem

Os principais benefícios da Computação em Nuvem são os seguintes:

- **Agilidade dos negócios:** num ambiente tradicional, o processo de aquisição de recursos novos ou adicionais de TI, tradicionalmente, são aprovações burocráticas. Como resultado, o processo pode demorar muito tempo, o que, por sua vez, pode atrasar as operações e pode aumentar o tempo de colocação no Mercado. A Computação em Nuvem fornece a capacidade de prover recursos de TI rapidamente e a qualquer momento, reduzindo consideravelmente o tempo necessário para a implantação de novos Sistemas e Serviços. Isso permite que as Empresas reduzam o tempo de colocação no Mercado e respondam mais rapidamente às mudanças nas condições do Mercado. A agilidade também possibilita um rápido desenvolvimento e experimentação que, por sua vez, facilita a inovação que é essencial para a pesquisa e o desenvolvimento, para a descoberta de novos mercados e oportunidades de receita, para a criação de novos segmentos de clientes e para o desenvolvimento de novos produtos;
- **Custos de TI Reduzidos:** num ambiente tradicional, os recursos são frequentemente adquiridos e dedicados a aplicações empresariais específicas. À medida do orçamento, os recursos são provisionados para acomodar os requisitos máximos estimados ou de pico de uso da aplicação. Essas práticas frequentemente resultam em maiores custos iniciais, na criação de *datacenters*, na subutilização de recursos e no aumento do consumo de energia. A Computação em Nuvem permite que os consumidores contratem quaisquer recursos de TI necessários, com base no pagamento por uso ou no preço de assinatura. Isso reduz os gastos de capital de TI de um consumidor (CAPEX), uma vez que o investimento é necessário apenas para os recursos necessários para acessar os serviços em Nuvem. Além disso, o consumidor contrata apenas os recursos da Nuvem que são necessários, eliminando, assim, *datacenters* e recursos inicialmente subutilizados. As despesas associadas à configuração de infraestrutura de TI, gerenciamento, espaço, energia e resfriamento são reduzidas. Assim, a adoção da Nuvem tem o potencial de reduzir o Custo Total de Propriedade (TCO) para um consumidor;
- **Alta disponibilidade:** a Computação em Nuvem tem a capacidade de garantir a disponibilidade de recursos em diferentes níveis, dependendo da política do consumidor e da prioridade do aplicativo. Componentes de infraestrutura redundantes (Sistemas de Computação, Caminhos de Rede e equipamentos de armazenamento, juntamente com o *software* em *cluster*) permitem tolerância a falhas para implantações em Nuvem. Essas técnicas podem abranger vários *datacenters* localizados em diferentes regiões geográficas, o que impede a indisponibilidade de dados devido a falhas regionais;

- **Continuidade do Negócio:** a indisponibilidade dos serviços de TI pode ocorrer por diversas causas, tais como desastres naturais, erros humanos, falhas técnicas e manutenção planejada. Isso pode levar a perdas financeiras significativas para as Organizações e pode afetar sua reputação. No entanto, ter serviços secundários e remotos para recuperação de desastres envolve despesas adicionais e administrativas. Por meio do uso de soluções de continuidade de negócios em Nuvem, uma Organização pode mitigar o impacto do tempo de inatividade e pode se recuperar de interrupções que afetam negativamente as operações comerciais. Por exemplo, uma Organização pode usar *backup* baseado em Nuvem para manter cópias adicionais de seus dados, que podem ser recuperados no caso de uma interrupção. Além disso, uma organização pode economizar nas despesas de capital necessárias para implementar uma solução de *backup* para sua infraestrutura de TI;
- **Escalabilidade:** as Organizações podem ter a necessidade de recursos de TI adicionais em momentos em que as cargas de trabalho são maiores. No entanto, eles não querem investir na compra de Sistemas de Computação adicionais e, em seguida, ter Sistemas ociosos. Eles querem liberar os recursos de computação após a conclusão da tarefa. Na Computação em Nuvem, os consumidores podem escalar unilateralmente e automaticamente recursos de TI para atender à demanda da carga de trabalho; isso é significativamente mais econômico do que comprar novos recursos de TI, que serão usados apenas por um curto período de tempo ou apenas em períodos específicos;
- **Flexibilidade de Acesso:** num ambiente tradicional, os recursos de TI são acessados a partir de dispositivos dedicados, como um *desktop* ou um *laptop*, um aplicativo tem de ser instalado para ser usado. Nesse ambiente, geralmente, não é possível acessar o aplicativo se o usuário estiver longe do dispositivo em que ele está instalado. Na Computação em Nuvem, os aplicativos e os dados residem centralmente e são acessados de qualquer lugar de uma Rede a partir de qualquer dispositivo, como *desktop*, celular, *thin client* etc. Isso elimina a dependência de um consumidor num dispositivo de ponto final específico. Isso também possibilita *Bring Your Device* (BYOD), pelo qual os funcionários podem usar dispositivos não pertencentes à Empresa. BYOD e *thin clients* criam uma oportunidade para reduzir custos de aquisição e operacionais;
- **Desenvolvimento e Teste de Aplicativos:** o teste de novos aplicativos em ambiente de produção ou dentro da Organização pode trazer impactos; sendo assim, o uso da Computação em Nuvem permite o desenvolvimento e teste de aplicativos numa escala maior e testes em múltiplas plataformas e fora do ambiente operacional da Organização, reduzindo, assim, o impacto de erros ou testes desastrosos dentro das Organizações. Isso pode reduzir, significativamente, o tempo total de entrega de novos Sistemas e Aplicativos;

- **Gerenciamento Simplificado de Infraestrutura:** num ambiente tradicional, o Departamento de TI de uma Organização precisa gerenciar grande quantidade de recursos de *hardware* e *software* e tarefas que envolvem configuração, aplicação de *patches* e atualizações e substituições. As cargas de trabalho e os requisitos de mão-de-obra aumentam com o tamanho da infraestrutura de TI. Quando uma Organização usa serviços em Nuvem, suas tarefas de gerenciamento de infraestrutura são reduzidas para gerenciar apenas os recursos necessários para acessar os serviços em Nuvem. A infraestrutura da Nuvem é gerenciada pelo provedor de serviços em Nuvem e tarefas como atualizações de *software* e renovações são tratadas pelo provedor. Ele garante que a infraestrutura da Nuvem permaneça moderna e atualizada com as necessidades do consumidor;
- **Mascaramento da Complexibilidade:** a Computação em Nuvem permite que uma Organização melhore os serviços sem aumentar o nível de conhecimento necessário para usar o serviço. Por exemplo, uma Organização pode usar a Nuvem para implementar um serviço de impressão de documentos, permitindo que os usuários imprimam documentos de qualquer local sem precisar configurar o serviço e as impressoras. Além disso, o serviço pode ser gerenciado e atualizado sem a participação do usuário final, diminuindo, assim, a complexidade de utilização e operação do serviço como um todo.

Modelos de Computação em Nuvem e *Big Data*

Um modelo de serviço em Nuvem especifica os serviços e os recursos que são fornecidos aos consumidores, no documento SP 800-145, o NIST classifica as ofertas de serviços em Nuvem nos três modelos principais listados a seguir:



- **Software como um serviço (SaaS):** segue o *link* para leitura complementar, no qual a Microsoft mostra a definição de SAAS: <https://goo.gl/S7AFkZ>.

Os diferentes modelos de serviço oferecem diferentes capacidades e são adequados para diferentes consumidores e objetivos de negócios.

Muitos modelos de serviços em Nuvem alternativos baseados em IaaS, PaaS e SaaS são definidos em várias publicações e por diferentes grupos da Indústria. Esses modelos de serviço são específicos para determinados serviços e recursos especializados em Nuvem que (eles) fornecem. Esses modelos de serviços em Nuvem são o *Backup* como Serviço (BaaS), Rede como Serviço (NaaS), Caso como Serviço (CaaS), Área de Trabalho como Serviço (DaaS), Ambiente de Teste como um serviço (TEaaS), Recuperação de Desastres como um Serviço (DRaaS), e assim

por diante. No entanto, esses modelos, eventualmente, pertencem a um dos três principais modelos de serviço em Nuvem.

Segue a especificação das três categorias principais:

- **Infraestrutura como Serviço:** o Serviço fornece processamento, armazenamento, Redes e outros recursos de computação fundamentais nos quais o consumidor é capaz de implantar e executar seus *softwares*, que pode incluir Sistemas Operacionais e Aplicativos. O consumidor não gerencia ou controla a infraestrutura de Nuvem subjacente, mas tem controle sobre Sistemas Operacionais, armazenamento e aplicativos implantados; e, possivelmente, controle limitado de componentes de Rede selecionados (por exemplo, *firewalls* de *host*) “- NIST.

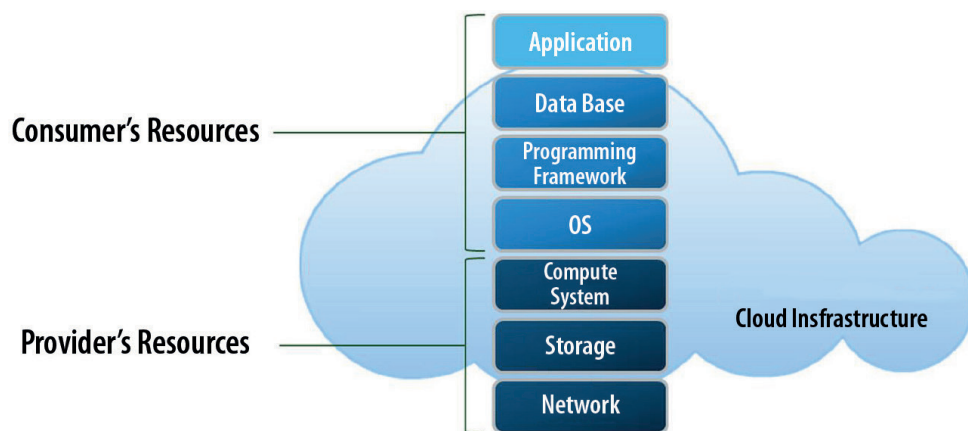


Figura 2

- **Plataforma como um Serviço:** o **serviço** oferecido é implementar na infraestrutura da Nuvem aplicações criadas pelo consumidor ou criadas usando Linguagens de Programação, Bibliotecas, Serviços e Ferramentas suportadas pelo provedor. O consumidor não gerencia, nem controla a infraestrutura de Nuvem subjacente, incluindo Rede, Servidores, Sistemas Operacionais ou Armazenamento, mas tem controle sobre os aplicativos implantados e, possivelmente, configurações para o ambiente de hospedagem de aplicativos. “- NIST. No modelo PaaS, um serviço em Nuvem, geralmente, inclui recursos de computação, armazenamento e Rede, juntamente com *software* de plataforma, incluindo um Sistema Operacional, um Banco de Dados, uma Estrutura de Programação, *Middleware* e ferramentas para desenvolver, testar, implantar e gerenciar aplicativos. O PaaS permite aos desenvolvedores de aplicativos projetar e desenvolver aplicativos baseados em Nuvem usando as Linguagens de Programação, as Bibliotecas de Classes e as Ferramentas suportadas pelo Provedor. Os produtos WISA (Windows, *Internet Information Services*, SQL Server e ASP.NET) e LAMP (Linux, Apache, MySQL e PHP/Python/Perl) são exemplos de pilhas de soluções fornecidas por meio do PaaS para desenvolver e implementar aplicações em Nuvem.

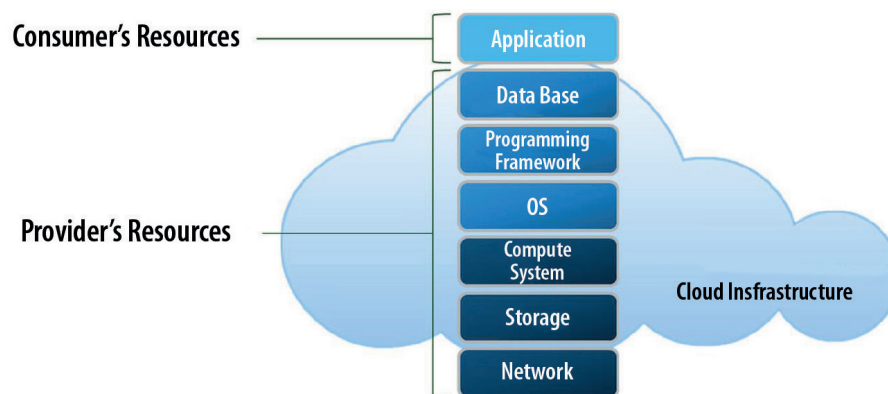


Figura 3

- **Software como Serviço:** o **serviço** oferecido é usar os aplicativos do provedor em execução numa infraestrutura em Nuvem. As aplicações são acessíveis a partir de vários dispositivos cliente através de uma interface de *thin client*, como um navegador da Web (por exemplo, *email* na Web) ou uma interface de programa. O consumidor não gerencia, nem controla a infraestrutura de Nuvem subjacente, incluindo Rede, Servidores, Sistemas Operacionais, Armazenamento ou mesmo Recursos de Aplicativos Individuais, com a possível exceção de configurações limitadas de configuração de aplicativos específicas do usuário. “- NIST”. No modelo SaaS, um provedor hospeda um aplicativo centralmente na Nuvem e oferece aos vários consumidores para uso como um serviço. Os consumidores não possuem nem gerenciam qualquer aspecto da infraestrutura da Nuvem. Em SaaS, uma determinada versão de um aplicativo, com uma configuração específica (*hardware* e *software*), normalmente, fornece serviço a vários consumidores, particionando suas sessões e dados individuais. Esses são alguns exemplos de *software* como serviço: *Customer Relationship Management* (CRM), *e-mail*, *Enterprise Resource Planning* (ERP) e *Apache HTTP*.

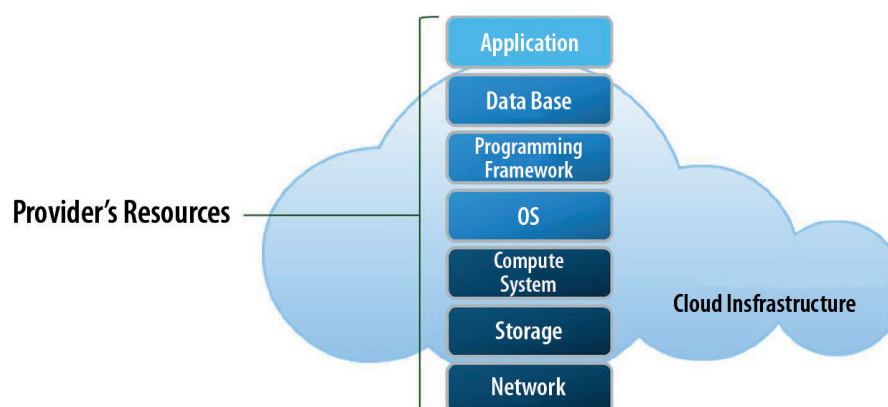


Figura 4

O provedor de serviços deverá ter uma maneira de monitoramento, mediação e tarifação de todos os serviços utilizados pelo cliente; para isso existe o *Cloud Services Brokerage* (CSB).

Um modelo de implantação em Nuvem fornece uma base para como a infraestrutura de Nuvem é criada, gerenciada e acessada. No documento SP 800-145, o NIST especifica os quatro modelos principais de implantação em Nuvem, listados a seguir:

- **Nuvem pública:** a infraestrutura da Nuvem é provisionada para uso aberto pelo público em geral. Pode ser de propriedade, gerenciado e operado por uma Empresa, Acadêmica ou Organização Governamental, ou alguma combinação entre eles. Ele existe nas instalações do provedor de Nuvem;
- **Nuvem privada:** a infraestrutura da Nuvem é provisionada para uso exclusivo por uma única organização composta por vários consumidores (por exemplo, unidades de negócios). Pode ser de propriedade, gerenciado e operado pela Organização, um terceiro, ou alguma combinação deles, e pode existir dentro ou fora das instalações. A infraestrutura pode ser criada internamente à Organização “On-premise” ou em *data centers* externos “Externally-hosted”;
- **Nuvem de Comunidade:** a infraestrutura da Nuvem é provisionada para uso exclusivo por uma comunidade específica de consumidores de Organizações que têm preocupações compartilhadas (por exemplo, missão, requisitos de segurança, política e considerações de conformidade). Ela pode ser de propriedade, gerenciada e operada por uma ou mais Organizações da comunidade, por um terceiro, ou por alguma combinação deles, e pode existir dentro ou fora das instalações. Nesse caso, a infraestrutura também pode ser criada internamente à organização “On-premise” ou em *datacenters* externos “Externally-hosted”;
- **Nuvem Híbrida:** a infraestrutura de Nuvem é uma composição de duas ou mais infraestruturas de Nuvem distintas (privadas, comunitárias ou públicas), que permanecem como entidades exclusivas, mas que são unidas pela Tecnologia padronizada ou proprietária que permite a portabilidade de dados e aplicativos (por exemplo, balanceamento de carga entre nuvens).

Note que a Tecnologia de *Big Data* está intimamente ligada às Tecnologias de Computação em Nuvem, devido à escalabilidade e à elasticidade da Nuvem, tendo em vista que os serviços de processamento de *Big Data* podem ser encontrados nas principais plataformas de Computação em Nuvem, como:

- Site da IBM Bluemix, uma das grandes plataformas de Computação em Nuvem: <https://goo.gl/aveu7L>;
- Explicação do que é Computação em Nuvem pela Microsoft, a plataforma Azure: <https://goo.gl/bMMov0>;
- Site Amazon WS, uma das grandes plataformas de Computação em Nuvem: <https://goo.gl/Q1xU4j>;
- Uma visão geral dos produtos existentes no Google Cloud Services ou serviço de Nuvem do Google: <https://goo.gl/5g3nZr>.

Material Complementar

Indicações para saber mais sobre os assuntos abordados nesta Unidade:

Sites

IBM

Analise a suite completa de serviços da Nuvem BlueMix IBM.

<https://goo.gl/aveu7L>

Azure Microsoft

Descreve com linguagem simples o que é a Computação em Nuvem e as características nesta Unidade também mencionadas.

<https://goo.gl/bMMov0>

Sales Force

Empresa Sales Force que ilustra também as definições sobre Computação em Nuvem.

<https://goo.gl/ABt9zX>

Livros

Public Pass for dumies

Ebook da Oracle: ***Public Pass for dumies***, importante descrição da Oracle para as definições e produtos que oferece. Este *link* pode requerer um cadastro simples no *site* da Oracle.

<https://goo.gl/qAzByo>

Referências

AGUADO, D. *et al.* **Practical Approach to Cloud IaaS with IBM SoftLayer: Presentations Guide**. IBM RedBook. 2016. Disponível em: <<http://www.redbooks.ibm.com/abstracts/sg248350.html?Open>>.

MELLAND, Peter M.; GRANCE, Timothy. 2011. SP 800-145. **The NIST Definition of Cloud Computing. Technical Report**. NIST, Gaithersburg, MD, United States. Disponível em: <<http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>>.

ZHU, W. *et al.* **Building Big Data and Analytics Solutions in the Cloud**. IBM RedBook. 2014. Disponível em: <<https://www.redbooks.ibm.com/redbooks.nsf/RedbookAbstracts/redp5085.html?Open>>.



Cruzeiro do Sul
Educatonal