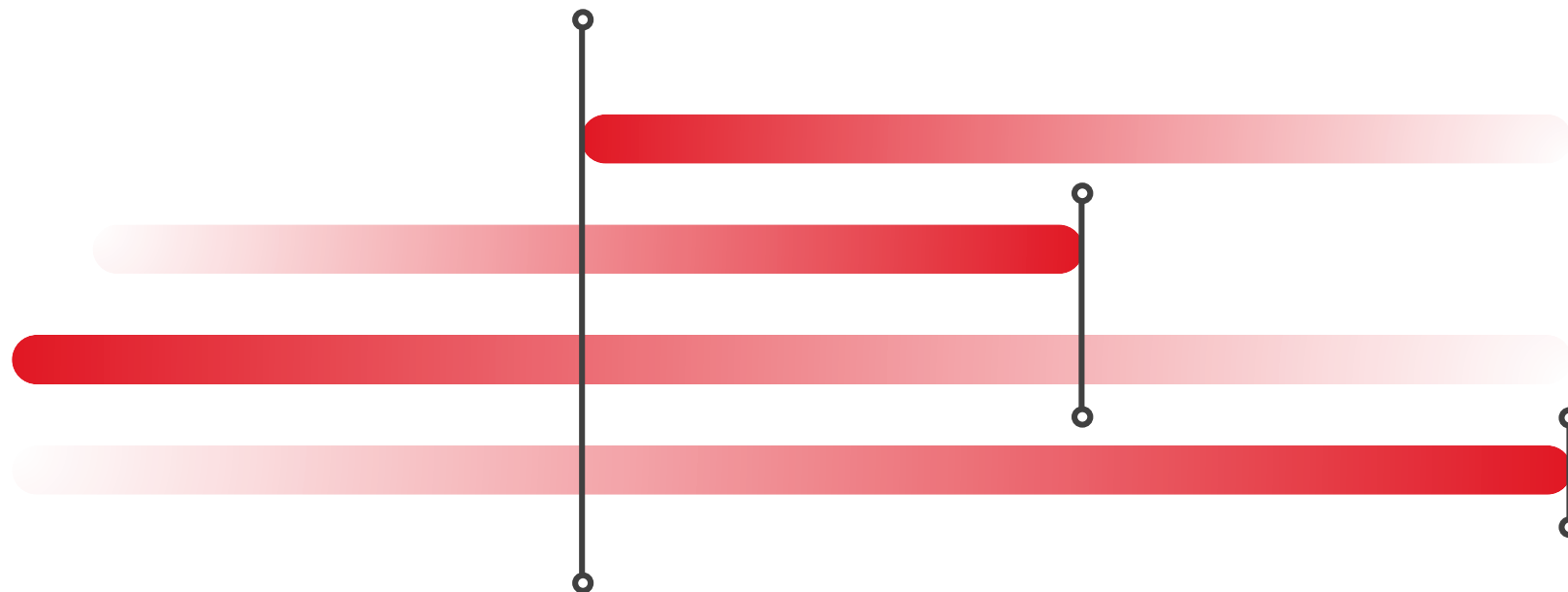




**uniesp**

Centro Universitário



ENGENHARIA  
**DE DADOS**

| **LAYSA**  
BELICHI



# ENGENHARIA DE DADOS

## TÓPICOS ABORDADOS

### Introdução à Computação em Nuvem:

- Definição de computação em nuvem.
- Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.
- Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).

### O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

### Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

### Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

### Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.



CLOUD COMPUTING

# CLOUD COMPUTING

## INTRODUÇÃO



É a entrega de recursos computacionais por demanda, ou seja conforme são necessários.

Essa escala é entregue através de plataformas, como por exemplo:

- AWS
- Azure
- IBM Cloud
- Oracle OCI
- Google Cloud Platform

Cloud abrangem tudo desde aplicativo até datacenter. E é pago conforme o uso.

Esse é o conceito de PaaS (Platform as a Service), onde ao invés do desenvolvedor demandar tempo para preparo de infraestrutura, isso já é fornecido em um modelo padrão e configurável.

# CLOUD COMPUTING

## INTRODUÇÃO

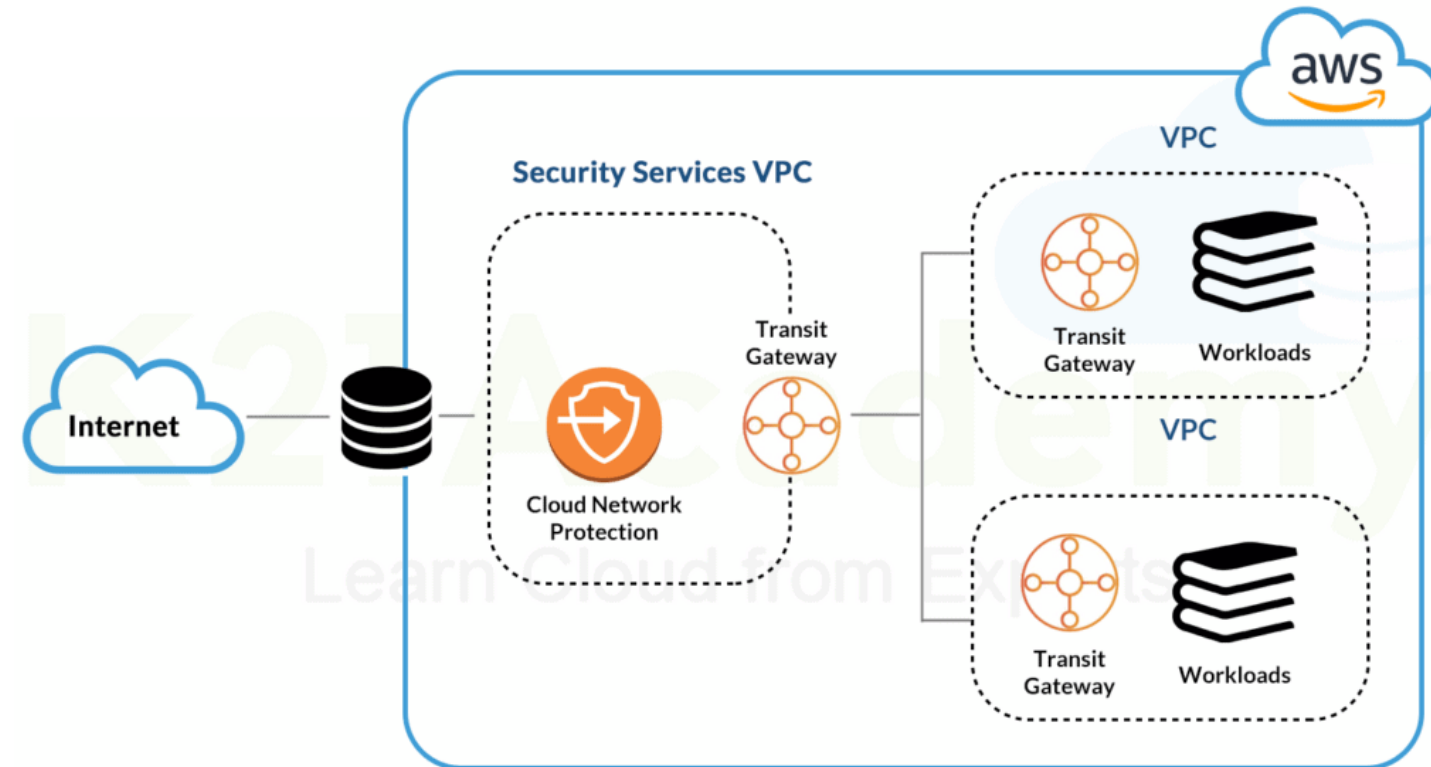


Exemplos básicos de recursos computacionais que operam em Cloud Computing e podem ser rapidamente provisionados são:

- Redes (Networks)
- Servidores (Servers)
- Armazenamentos (Storages)
- Aplicações (Applications)
- Serviços (Services)

# CLOUD COMPUTING

## NETWORKS



**Networks:** redes se referem às infraestruturas de comunicação que conectam todos os recursos e usuários dentro de um ambiente de computação em nuvem.

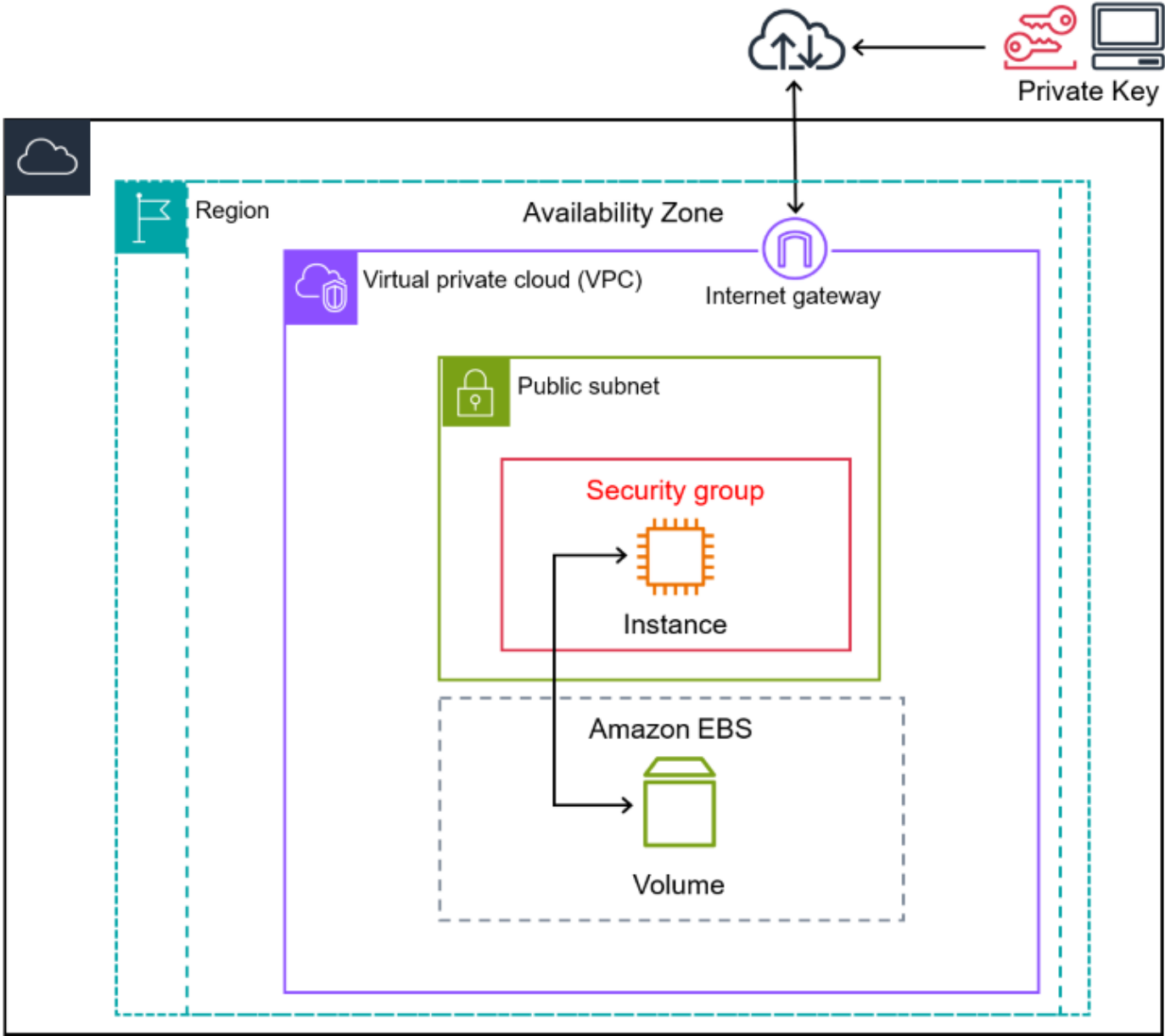
### Exemplo:

Amazon Virtual Private Cloud (VPC) é um serviço da AWS que permite criar uma rede virtual na nuvem, isolada logicamente, onde você pode lançar recursos da AWS, como instâncias EC2 (servidores virtuais) e RDS (banco de dados relacional).

Dentro de uma VPC, os usuários podem criar sub-redes, que são segmentos lógicos da rede que podem ser configurados com diferentes regras de acesso e políticas de segurança. Além disso, é possível configurar rotas de rede personalizadas para direcionar o tráfego entre sub-redes e para a internet.

# CLOUD COMPUTING

## SERVICES



**Servers:** Servidores na nuvem são instâncias de computação virtual que executam aplicativos ou serviços para os usuários.

Exemplo: Amazon EC2 (Elastic Compute Cloud) é um serviço de computação em nuvem da AWS que oferece capacidade de computação escalável na nuvem. Os usuários podem lançar instâncias de servidores virtuais sob demanda, sistema operacional, quantidade de CPU, memória, armazenamento e outras características de configuração.

O Amazon EC2 oferece recursos de escalabilidade e resiliência incorporados. Os usuários podem dimensionar horizontalmente suas instâncias EC2 conforme necessário para lidar com picos de carga de trabalho e podem aproveitar as opções de armazenamento e backup para garantir a resiliência dos dados e aplicativos.



# CLOUD COMPUTING

## STORAGES



### Amazon S3

Durable object storage for all types of data



### Amazon Glacier

Archival storage for infrequently accessed data



### Amazon EBS

Block storage for use with Amazon EC2



### Amazon EFS

File storage for use with Amazon EC2

#### Economics

Pay as you go  
No upfront investment  
No commitment

#### Easy to Use

Self service administration  
SDKs for simple integration

#### Reduce risk

Durable and Secure  
Avoid risks of physical media handling

#### Agility, Scale

Reduce time to market  
Focus on your business, not your infrastructure

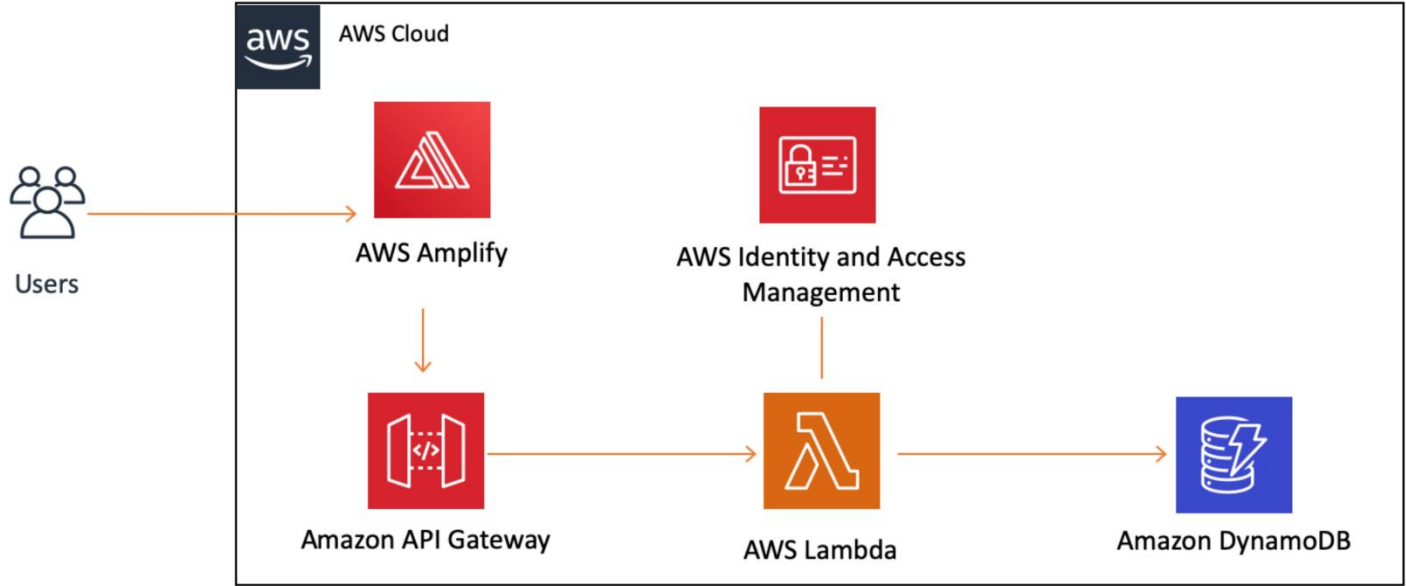
**Storage:** Armazenamento em cloud computing refere-se aos recursos de armazenamento de dados disponíveis na nuvem para os usuários, permitindo o armazenamento e a recuperação de dados de forma segura e escalável.

Exemplo: Amazon S3 (Simple Storage Service) é um serviço de armazenamento de objetos da AWS, projetado para armazenar e recuperar grandes quantidades de dados de forma durável e altamente disponível.

O Amazon S3 se integra perfeitamente com outros serviços da AWS, permitindo aos usuários construir soluções completas e integradas na nuvem. Por exemplo, o S3 pode ser usado para armazenar arquivos de mídia para serem acessados por aplicativos hospedados em instâncias EC2, ou pode ser usado como origem de dados para serviços de análise de dados, como Amazon Athena e Amazon Redshift.

# CLOUD COMPUTING

## APPLICATIONS

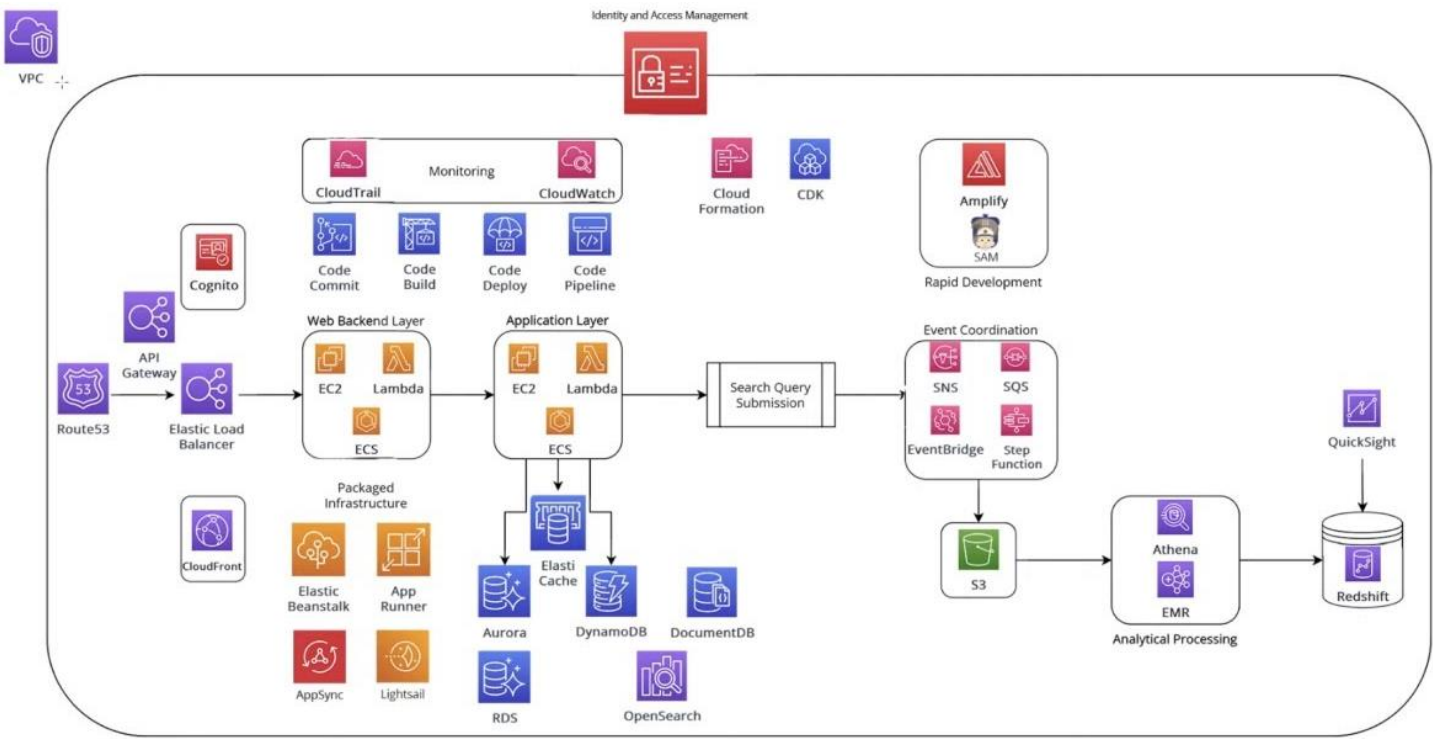


**Applications:** Aplicativos na nuvem são programas ou software que são executados em infraestrutura de computação em nuvem e são acessíveis aos usuários pela internet.

Exemplo: O Amazon WorkMail é um serviço de e-mail e calendário gerenciado na nuvem que oferece uma alternativa ao Gmail. Ele permite que você crie e gerencie caixas de correio de forma escalável e segura, com integração com outros serviços da AWS, como o Amazon Simple Storage Service (S3) para armazenamento de anexos.

Ao combinar esses serviços da AWS, é possível criar uma suíte de aplicativos de produtividade baseada em nuvem semelhante ao Google Workspace. Cada serviço é executado na infraestrutura de nuvem da AWS e pode ser acessado pelos usuários através de navegadores web, proporcionando uma experiência de colaboração eficiente e escalável.

# CLOUD COMPUTING SERVICES



**Service:** Serviços na nuvem são recursos ou funcionalidades disponibilizadas como serviços pela internet, permitindo que os usuários consumam esses recursos sob demanda.

Exemplo: AWS Lambda é um serviço de computação serverless da AWS que permite executar código sem provisionar ou gerenciar servidores. É uma ferramenta poderosa e flexível que permite aos desenvolvedores executar código de forma eficiente e econômica na nuvem, sem se preocupar com a infraestrutura subjacente. Ele oferece escala automática, pagamento por uso e integração perfeita com outros serviços da AWS, tornando-o uma escolha popular para uma ampla variedade de cargas de trabalho e aplicativos na nuvem.

# CLOUD COMPUTING SERVICES



## IaaS

Infrastructure-as-a-Service

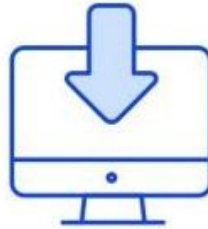
**host**



## PaaS

Platform-as-a-Service

**build**



## SaaS

Software-as-a-Service

**consume**

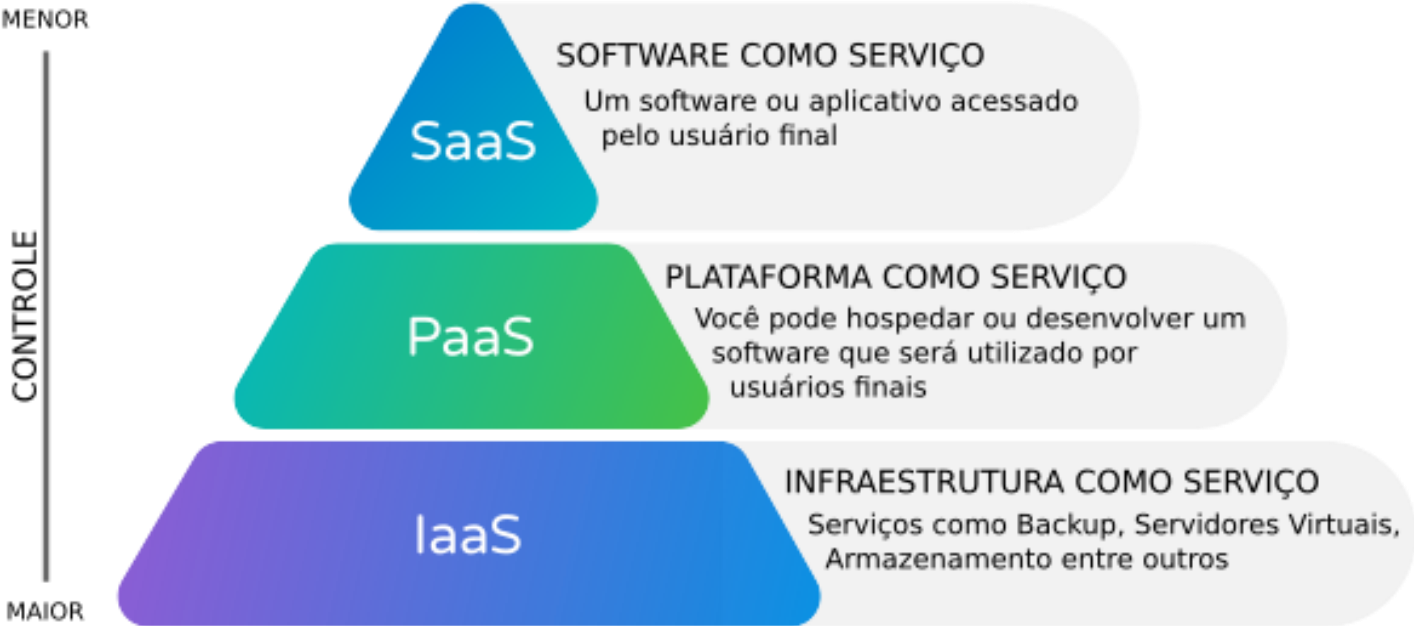
Os serviços mais comuns disponíveis na nuvem são:

- **IaaS** (Infrastructure as a Service);
- **SaaS** (Software as a Service);
- **PaaS** (Platform as a Service);
- **CaaS** (Container as a Service);

Como essas ferramentas da plataforma são acessadas livremente pela Internet, e não através de um sistema operacional ou pacote instalado em uma máquina local, os desenvolvedores não precisam se preocupar com a logística de montar algo que será instalado em um computador do cliente, pois tudo é mantido na própria nuvem. Qualquer pessoa com um navegador da web pode acessar o aplicativo.

# CLOUD COMPUTING

## IAAS (INFRASTRUCTURE AS A SERVICE)



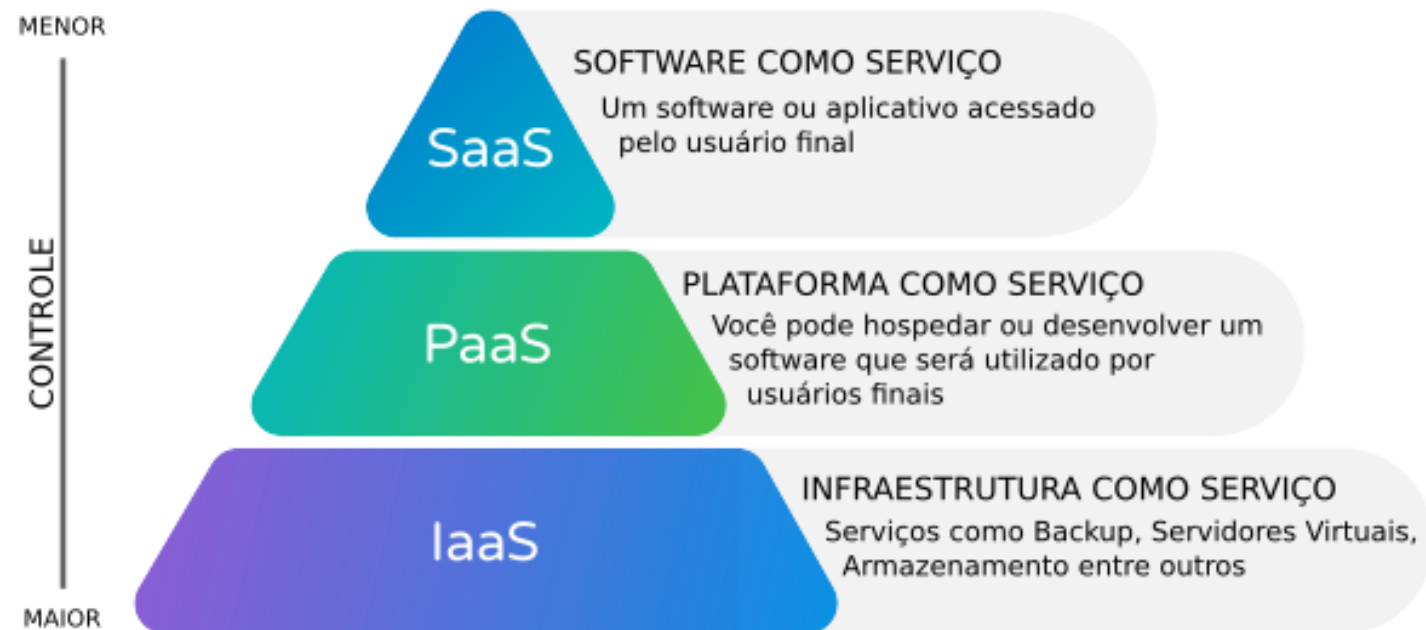
A **Infraestrutura como Serviço**, como o próprio nome sugere, fornece infraestrutura de computação em nuvem, incluindo servidores, rede, sistemas operacionais e armazenamento por meio da tecnologia de virtualização.

Esses servidores em nuvem geralmente são fornecidos à organização por meio de um painel ou uma API, proporcionando aos clientes controle total sobre toda a infraestrutura. O IaaS fornece as mesmas tecnologias e recursos que um data center tradicional, sem a necessidade de manter ou gerenciar fisicamente tudo.

Os clientes de IaaS ainda podem acessar seus servidores e armazenamento diretamente, mas tudo é terceirizado por meio de um “datacenter virtual” na nuvem.

# CLOUD COMPUTING

## IAAS (INFRASTRUCTURE AS A SERVICE)



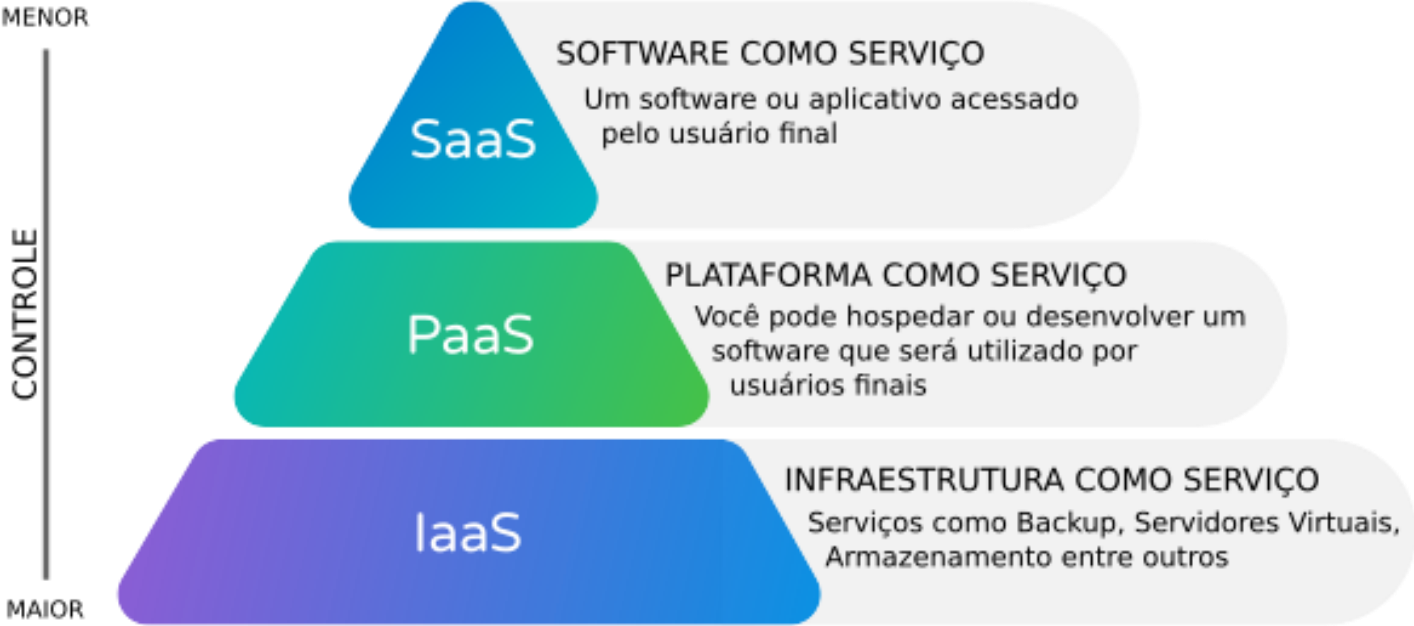
A interligação entre IaaS e Engenharia de Dados pode ser feita de várias maneiras para suportar efetivamente o fluxo de trabalho de engenharia de dados. Aqui estão algumas abordagens:

- I. *Provisionamento de Infraestrutura;*
- II. *Implantação de Ferramentas de Engenharia de Dados;*
- III. *Escala Automática;*
- IV. *Armazenamento de Dados;*
- V. *Segurança e Conformidade;*

Podemos construir uma infraestrutura flexível, escalável e segura para suportar todas as etapas do ciclo de vida dos dados, desde a ingestão até a análise e a tomada de decisões.

# CLOUD COMPUTING

## IAAS (INFRASTRUCTURE AS A SERVICE)



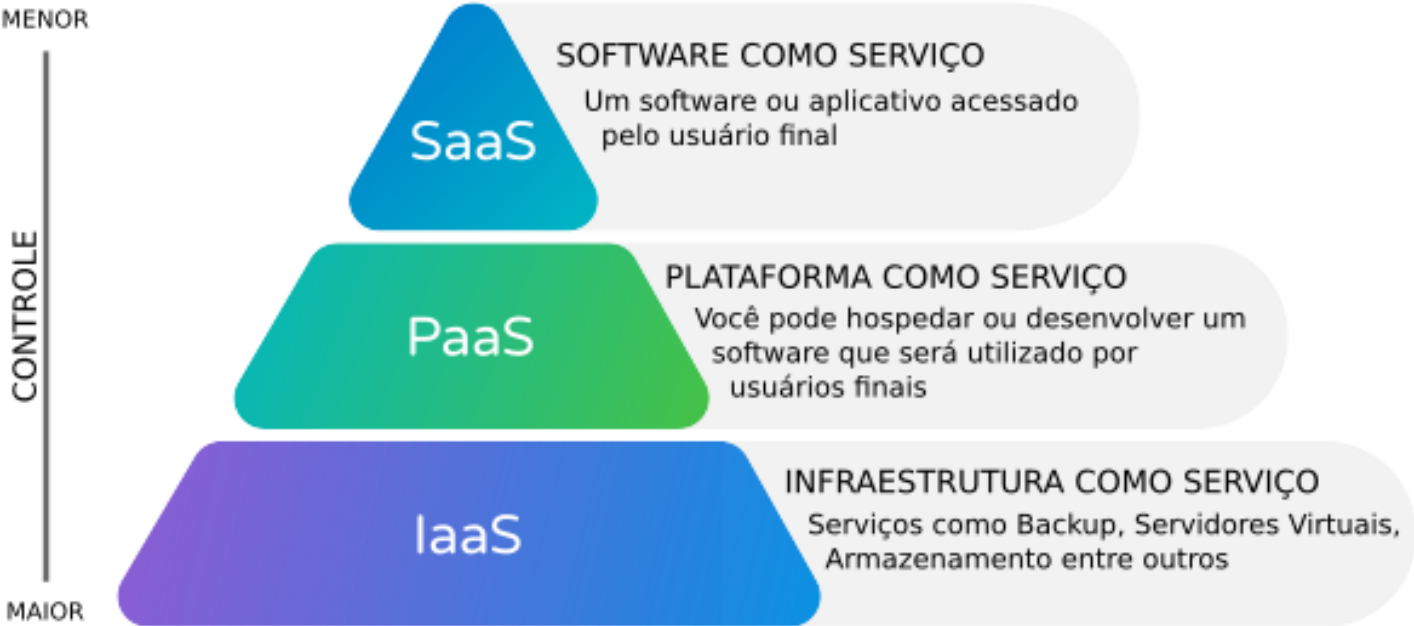
**Cenário:** Uma empresa de comércio eletrônico precisa lidar com grandes volumes de dados de transações de clientes e deseja construir um sistema de análise de dados para entender o comportamento do cliente e otimizar as recomendações de produtos.

**Aplicação de IaaS:** A empresa pode usar serviços de IaaS, como Amazon Web Services (AWS) ou Microsoft Azure, para provisionar recursos de computação e armazenamento conforme necessário. Eles podem implantar máquinas virtuais para executar clusters de processamento de big data, como Apache Hadoop ou Apache Spark. Além disso, eles podem usar serviços de armazenamento escalável, como Amazon S3 ou Azure Blob Storage, para armazenar grandes conjuntos de dados brutos e processados.



# CLOUD COMPUTING

## SAAS (SOFTWARE AS A SERVICE)



O **Software como Serviço** é a opção mais utilizada para empresas no mercado de nuvem.

O SaaS utiliza a Internet para entregar aplicativos aos clientes, sendo gerenciados por um fornecedor terceirizado.

A maioria dos aplicativos SaaS é executado diretamente no navegador de web, o que significa que eles não exigem nenhum download ou instalação por parte do cliente.



# CLOUD COMPUTING

## SAAS (SOFTWARE AS A SERVICE)



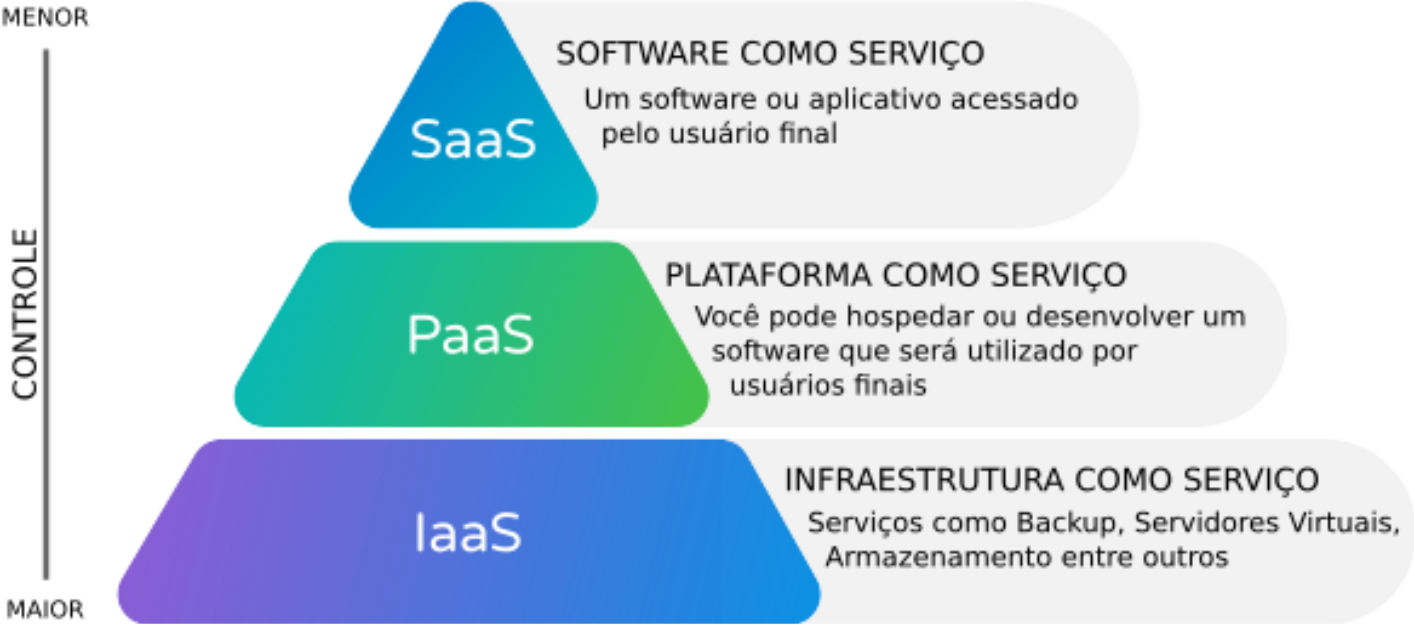
SaaS e Engenharia de dados se interligam de várias maneiras, onde são aproveitados recursos para facilitar processos.

Ao interligar SaaS com engenharia de dados, você pode aproveitar as vantagens dos aplicativos SaaS para coletar, gerenciar e analisar dados, melhorando assim a eficiência operacional e impulsionando a tomada de decisões baseada em dados.

- I. *Integração de Dados;*
- II. *Análise de Dados;*
- III. *Automação de Processos;*
- IV. *Gerenciamento de Dados;*
- V. *Colaboração e Compartilhamento de Dados;*

# CLOUD COMPUTING

## SAAS (SOFTWARE AS A SERVICE)

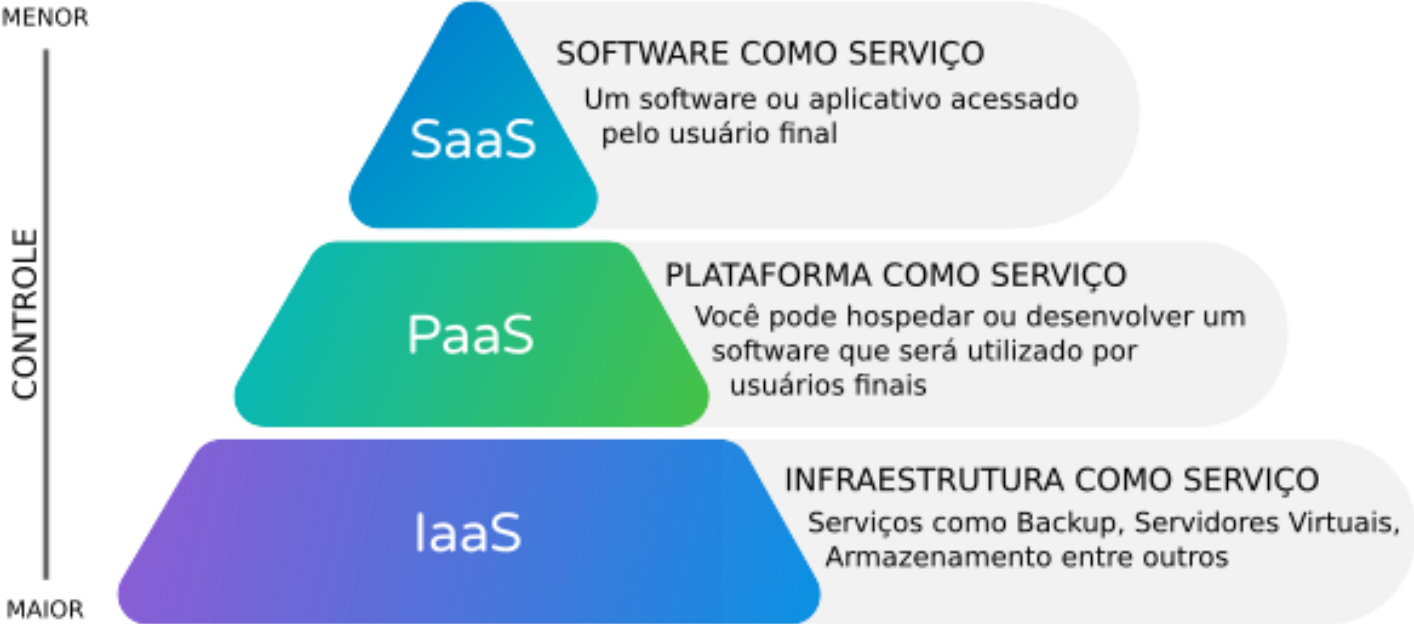


**Cenário:** Uma empresa de marketing precisa analisar o desempenho de suas campanhas de marketing em várias plataformas de mídia social e anúncios online.

**Aplicação de SaaS:** A empresa pode utilizar plataformas de análise de marketing como Google Analytics, HubSpot ou Adobe Analytics, que são exemplos de SaaS. Essas plataformas oferecem recursos poderosos de análise de dados, permitindo que a empresa acompanhe métricas-chave, como tráfego do site, conversões, ROI de campanhas e engajamento nas redes sociais. Os dados podem ser facilmente acessados através da interface web do SaaS, sem a necessidade de configuração de infraestrutura adicional.

# CLOUD COMPUTING

## PAAS (PLATAFORMA AS A SERVICE)



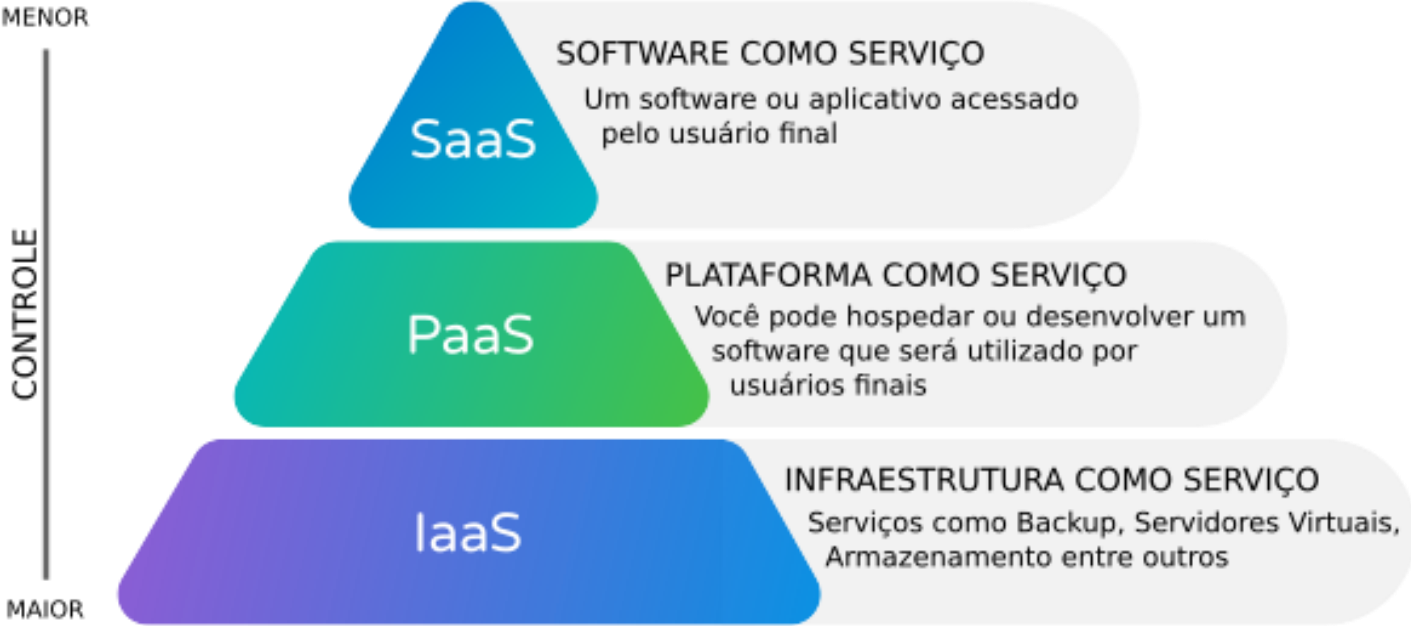
A **Plataforma como Serviço** fornece componentes em nuvem para determinado software enquanto são usados principalmente para aplicativos.

O PaaS fornece uma estrutura para desenvolvedores desenvolver e usar para criar aplicativos personalizados. Todos os servidores, armazenamento e rede podem ser gerenciados pela empresa ou por um provedor de terceiros, enquanto os desenvolvedores podem manter o gerenciamento dos aplicativos.

O modelo do PaaS é semelhante ao SaaS, exceto que, ao invés de entregar o software pela Internet, o PaaS fornece uma plataforma para criação de software. Essa plataforma é fornecida via web, oferecendo aos desenvolvedores a liberdade de se concentrar na criação do software sem se preocupar com sistemas operacionais, atualizações de software, armazenamento ou infraestrutura.

# CLOUD COMPUTING

## PAAS (PLATAFORMA AS A SERVICE)



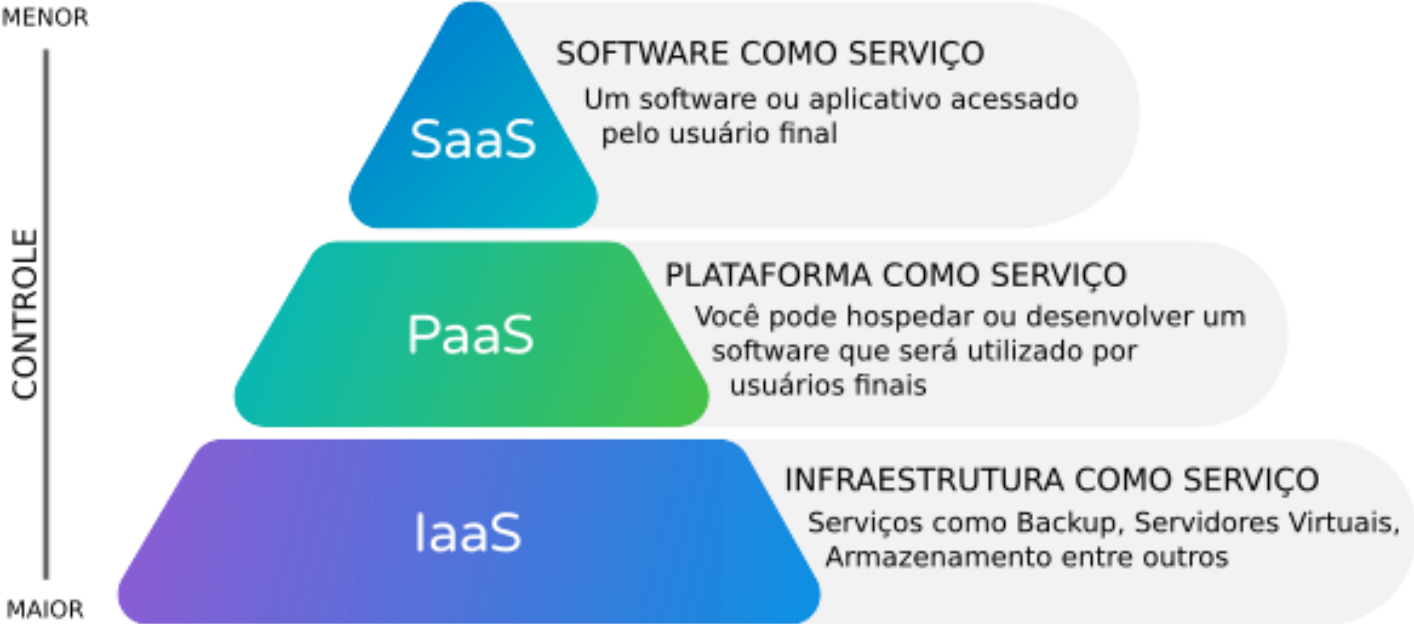
PaaS fornece uma plataforma completa para desenvolver, implantar e gerenciar aplicativos, o que pode ser aproveitado para facilitar o trabalho de Engenharia de Dados:

- I. Ferramentas de Processamento de Dados;
- II. Bancos de Dados Gerenciados;
- III. APIs para Integração de Dados;
- IV. Automação de Processos;
- V. Desenvolvimento de Aplicações de Análise;
- VI. Colaboração e Compartilhamento de Dados;

Ao interligar PaaS com Engenharia de Dados, podemos criar um ambiente integrado e eficiente para desenvolver, implantar e gerenciar pipelines de dados, análises e aplicativos de dados personalizados. Isso pode ajudar a melhorar a eficiência, acelerar o ciclo de desenvolvimento e fornecer insights valiosos para o seu negócio.

# CLOUD COMPUTING

## PAAS (PLATAFORMA AS A SERVICE)



**Cenário:** Uma startup de tecnologia está desenvolvendo um aplicativo de análise de dados para processar e visualizar grandes conjuntos de dados de sensores IoT (Internet das Coisas).

**Aplicação de PaaS:** A startup pode usar uma plataforma PaaS, como Microsoft Azure IoT Hub ou AWS IoT Core, para gerenciar a ingestão e processamento de dados de sensores IoT. Essas plataformas fornecem serviços gerenciados para coleta, armazenamento e análise de dados de dispositivos IoT. Além disso, a startup pode utilizar serviços de análise de dados em tempo real, como Azure Stream Analytics ou AWS Kinesis, para processar e analisar os dados dos sensores. Eles também podem usar serviços de banco de dados gerenciados, como Azure Cosmos DB ou AWS DynamoDB, para armazenar os dados processados e permitir consultas rápidas e escaláveis.

CLOUD COMPUTING

VANTAGENS

SERVIÇO	VANTAGEM	APLICAÇÃO
IaaS	O IaaS é o modelo de Computação em Nuvem mais flexível, pois os recursos de hardware podem ser adquiridos conforme a necessidade. Além disso, o modelo é fácil de automatizar e é altamente escalável. Os clientes, portanto, mantêm o controle completo de sua infraestrutura.	AWS EC2 Google Compute Engine (GCE) Huawei Kunpeng Cloud Services.
SaaS	Devido ao seu modelo de entrega na web, o SaaS elimina a necessidade de a equipe de TI fazer o download e instalar aplicativos em cada computador. Com o SaaS, os fornecedores gerenciam todos os possíveis problemas técnicos como dados, middleware, servidores e armazenamento, resultando em uma manutenção e suporte otimizados para o seu negócio.	G Suite, Salesforce, Mailchimp, Google Analytics, Dropbox.
PaaS	Os benefícios do PaaS incluem desenvolvimento e implementação simples e econômica, que é altamente disponível. Além disso, os desenvolvedores podem focar na personalização dos aplicativos.	AWS Beanstalk, AWS RDS, GCP App Engine.

# ENGENHARIA DE DADOS

## TÓPICOS ABORDADOS

### Introdução à Computação em Nuvem:

- ~~Definição de computação em nuvem.~~
- ~~Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.~~
- ~~Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).~~

### O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

### Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

### Implementação do DaaS:

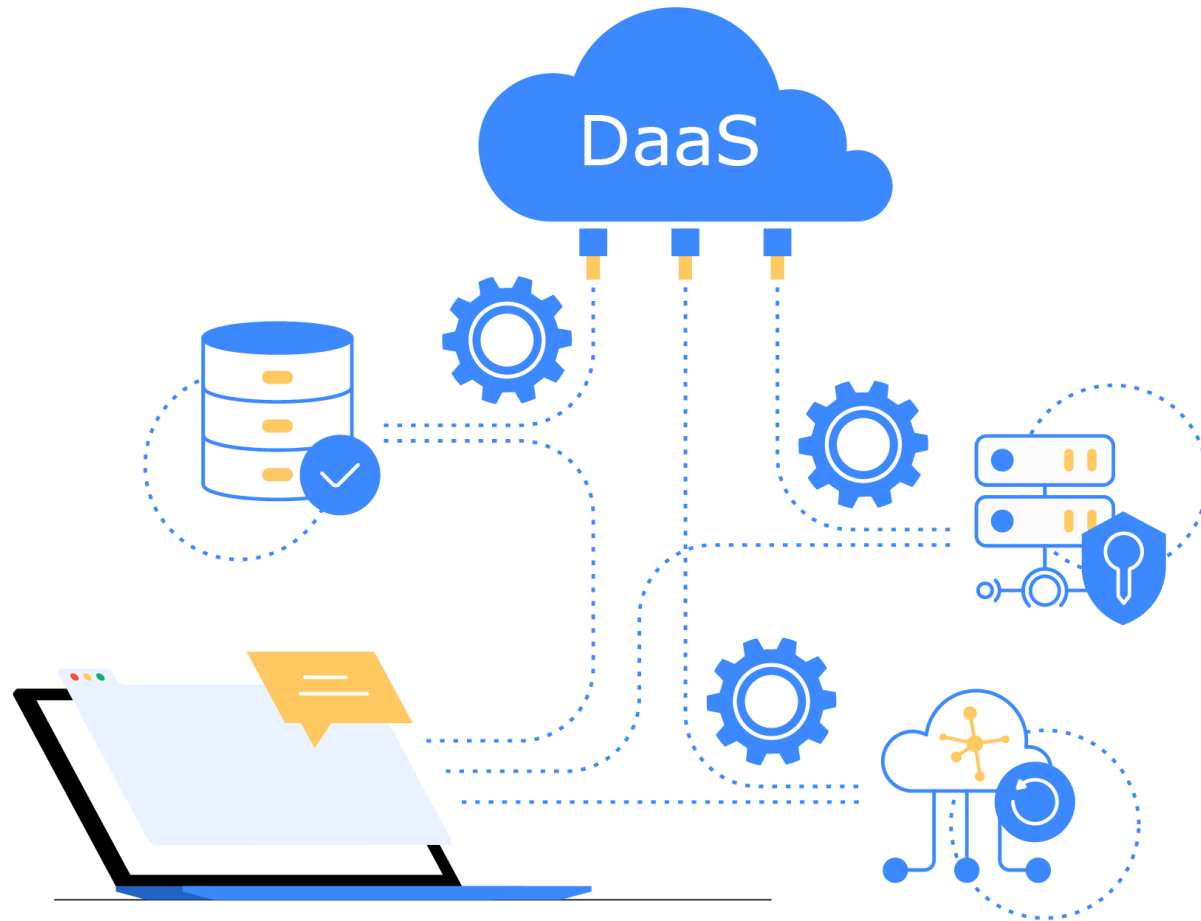
- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

### Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.

# DATA AS A SERVICE

## DEFINIÇÃO



**Dados como serviço** é uma variedade de soluções baseadas em nuvem usadas para trabalhar com dados. Essas soluções são oferecidas por fornecedores de DaaS, que disponibilizam dados de negócios prontamente sob demanda pela Internet.

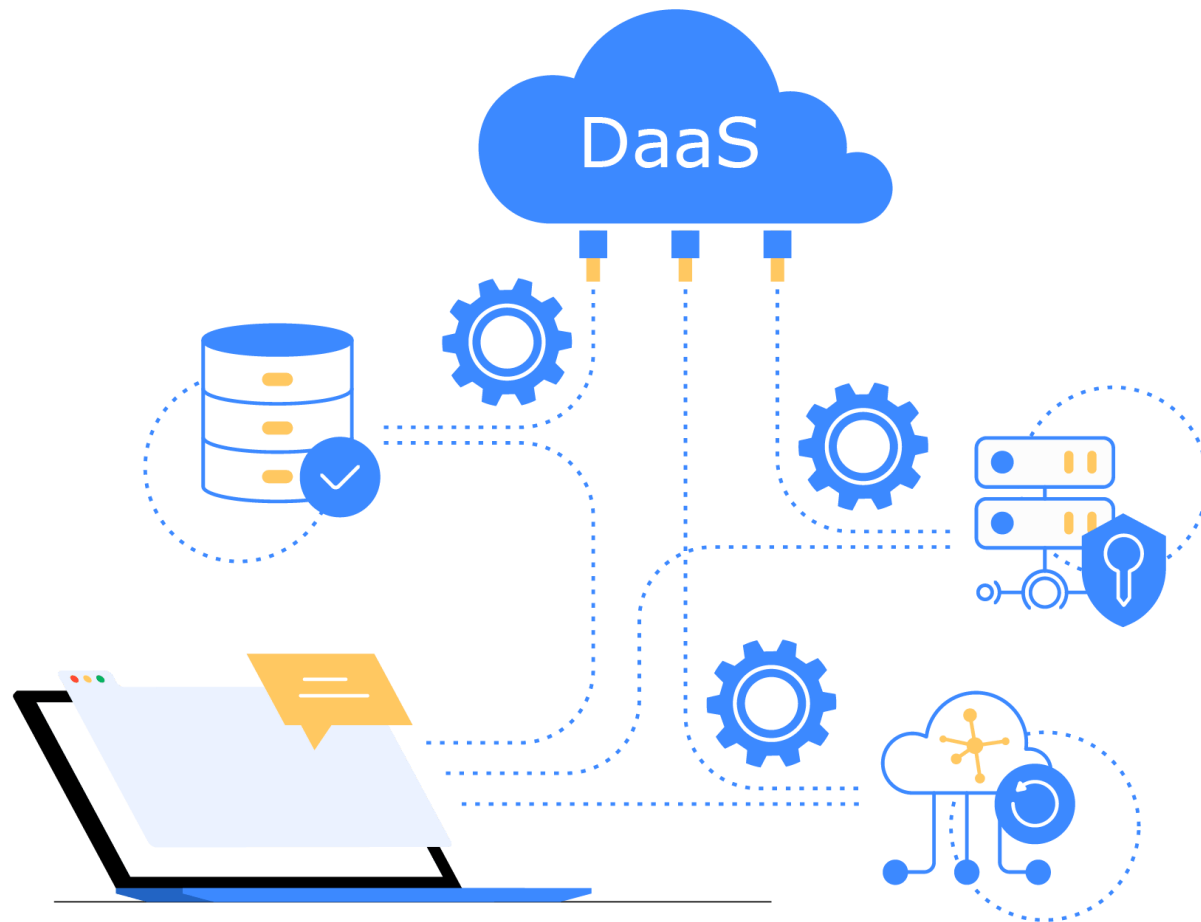
DaaS é frequentemente descrito como uma estratégia de gerenciamento de dados que emprega a nuvem para fornecer armazenamento de dados, realizar monitoramento, processar dados e assim por diante.

Semelhante ao SaaS (Software as a Service), o DaaS elimina a necessidade de baixar e gerenciar software localmente. Essa característica torna o serviço em nuvem popular entre empresas que desejam migrar seu armazenamento de dados de servidores locais para a nuvem.



# DATA AS A SERVICE

## DEFINIÇÃO



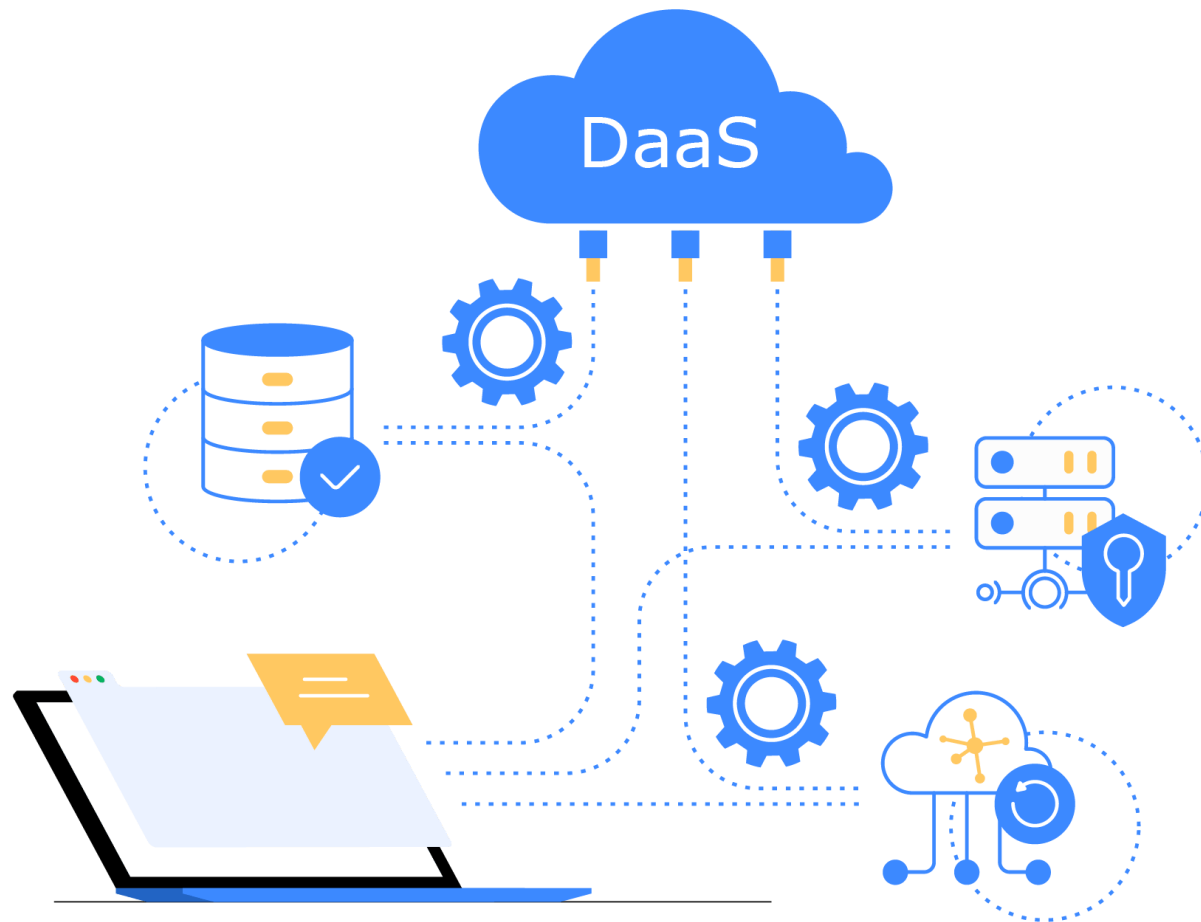
Numerosos benefícios do DaaS que podem influenciar significativamente o desempenho dos negócios atraem grandes e pequenas empresas em todo o mundo.

**Maior acessibilidade** – As empresas que adotam DaaS não estão vinculadas a sistemas e servidores de armazenamento locais, uma vez que armazenam dados remotamente. Portanto, podem colaborar com parceiros globais de forma eficiente. Simultaneamente, os funcionários podem acessar dados sem precisar ir ao escritório, economizando tempo e energia.

**Produtos melhores** – o DaaS ajuda as empresas a perceber os pontos fracos e fortes de seus produtos. Com base nisso, eles podem minimizar os pontos fracos e aproveitar os pontos fortes, levando a melhorias no produto.

# DATA AS A SERVICE

## DEFINIÇÃO



**Custo-benefício** – Quem entende perfeitamente o que é DaaS enfatiza que sua vantagem mais considerável é o custo-benefício. Com o DaaS, não há necessidade de investir recursos financeiros na manutenção e gerenciamento de dados porque tudo é feito pelos fornecedores de DaaS.

**Pouco tempo de configuração** – As empresas podem começar a armazenar e processar dados imediatamente ao empregar DaaS. Por outro lado, as empresas que mantêm os dados localmente gastam muito mais tempo na configuração.

**Experiência do usuário aprimorada** – As empresas que estão cientes do que é DaaS tomam decisões baseadas em dados que melhoram a experiência do cliente. Portanto, eles observam melhorias na retenção de clientes e no valor da vida útil do cliente.

DATA AS A SERVICE

TRADICIONAL X MODERNO

ABORDAGEM	TRADICIONAL	DAAS
Acesso aos Dados:	Nos métodos tradicionais, os dados são armazenados localmente em servidores ou infraestrutura de TI da própria organização. O acesso aos dados geralmente requer uma infraestrutura de hardware e software dedicada dentro da organização.	Os dados são disponibilizados como um serviço através da nuvem. Isso significa que os usuários podem acessar os dados remotamente pela internet, sem a necessidade de manter uma infraestrutura de TI local. Os dados são hospedados e gerenciados por um provedor de serviços em nuvem, que fornece acesso sob demanda aos dados conforme necessário.
Armazenamento de Dados:	Nos métodos tradicionais, os dados são armazenados localmente em bancos de dados, servidores de arquivos ou sistemas de armazenamento próprios da organização. A capacidade de armazenamento é limitada pela infraestrutura de TI disponível.	Os dados são armazenados na nuvem pelo provedor de serviços. Isso permite uma escalabilidade quase ilimitada do armazenamento, conforme a demanda, sem a necessidade de investimentos em infraestrutura adicional por parte da organização.
Gerenciamento e Manutenção:	A organização é responsável pelo gerenciamento e manutenção dos seus próprios dados. Isso inclui tarefas como backup, atualizações de software, segurança e monitoramento de desempenho.	O provedor de serviços em nuvem é responsável pelo gerenciamento e manutenção dos dados. Isso inclui garantir a disponibilidade, segurança, conformidade e desempenho dos dados hospedados na nuvem.
Custos e Flexibilidade:	Os custos de infraestrutura de TI própria podem ser significativos, incluindo a compra de hardware, licenças de software e custos operacionais contínuos. A escalabilidade pode ser limitada pela capacidade da infraestrutura existente.	Custos são geralmente baseados no uso, o que significa que as organizações podem pagar apenas pelo que realmente utilizam. Além disso, a escalabilidade é facilitada, permitindo que as organizações aumentem ou diminuam facilmente a capacidade de armazenamento e processamento de dados conforme necessário.

## DATA AS A SERVICE

### BENEFÍCIOS PARA EMPRESAS



Os benefícios do DaaS para organizações são significativos e podem transformar a maneira como lidam com dados.

**Acesso Fácil a Conjuntos de Dados:** Com o DaaS, as organizações podem acessar uma ampla variedade de conjuntos de dados de forma fácil e rápida, muitas vezes por meio de APIs simples ou interfaces web amigáveis.

**Exemplo de Uso:** Uma empresa de entrega de utiliza as API do Google Maps Platform para integrar recursos de mapeamento em seu aplicativo. Eles usam a API de Geocodificação para converter endereços em coordenadas geográficas, a API de Roteamento para calcular rotas otimizadas e a API de Visualização de Mapa para exibir mapas interativos em seu app. Com isso a empresa fornece aos entregadores e clientes informações precisas de localização em tempo real, melhorando a eficiência e a experiência do usuário.



**Eliminação de Custos de Infraestrutura:** O DaaS elimina a necessidade de investimento em infraestrutura de TI dedicada para armazenamento, processamento e análise de dados.

**Exemplo de Uso:** Uma empresa de comércio eletrônico usa o Amazon Redshift para armazenar e analisar grandes volumes de dados de transações de clientes, histórico de compras e comportamento de navegação na web. Eles podem executar consultas complexas e gerar insights valiosos sobre padrões de compra, preferências do cliente e tendências de mercado sem precisar investir em infraestrutura de data warehousing dedicada. Isso permite que a empresa tome decisões de negócios mais informadas e eficazes para melhorar a experiência do cliente e impulsionar as vendas.



**amazon**  
**REDSHIFT**

O Amazon Redshift é um serviço de data warehousing totalmente gerenciado que permite às organizações armazenar e analisar grandes volumes de dados de maneira eficiente e escalável na nuvem da AWS.

Ele oferece capacidade de armazenamento expansível, alta disponibilidade e desempenho rápido para cargas de trabalho de análise de dados.

## DATA AS A SERVICE

### BENEFÍCIOS PARA EMPRESAS



**Agilidade e Flexibilidade:** As organizações podem facilmente adicionar novos conjuntos de dados, alterar os requisitos de processamento ou análise e integrar-se a novas ferramentas ou plataformas conforme necessário.

**Exemplo de Uso:** Uma empresa de varejo online usa o Snowflake Data Cloud para armazenar e analisar dados de vendas, estoque, comportamento do cliente e marketing. Eles podem facilmente adicionar novos conjuntos de dados, como dados de mídia social, análises de sentimentos ou dados de localização, conforme necessário para melhorar suas análises e insights. Além disso, eles podem integrar-se a novas ferramentas de análise de dados e visualização de dados para fornecer uma visão abrangente de suas operações e desempenho de negócios.



O Snowflake Data Cloud é uma plataforma de dados em nuvem que oferece armazenamento, processamento e compartilhamento de dados de maneira escalável e flexível.

Ele permite que as organizações armazenem e analisem uma ampla variedade de dados, incluindo dados estruturados, semiestruturados e não estruturados, em uma arquitetura de dados unificada.



# ENGENHARIA DE DADOS

## TÓPICOS ABORDADOS

### Introdução à Computação em Nuvem:

- ~~Definição de computação em nuvem.~~
- ~~Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.~~
- ~~Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).~~

### O que é DaaS:

- ~~Definição de DaaS.~~
- ~~Diferença entre dados tradicionais e DaaS.~~
- ~~Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.~~

### Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

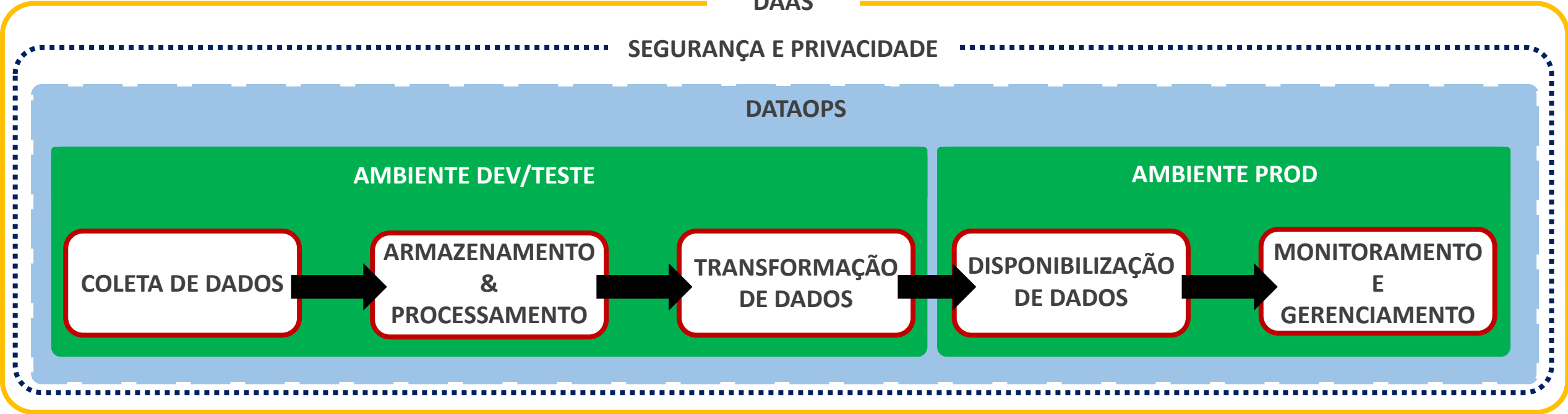
### Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

### Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.

DATA AS A SERVICE  
ARQUITETURA





## DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

### Definição de Requisitos:

- O processo começa com a definição clara dos requisitos de dados da organização. Isso pode incluir os tipos de dados necessários, os volumes esperados, os requisitos de segurança e privacidade, entre outros.

### Planejamento de Infraestrutura:

- Com base nos requisitos definidos, é feito um planejamento da infraestrutura necessária para armazenar, processar e disponibilizar os dados. Isso pode incluir a seleção de tecnologias de armazenamento, processamento e segurança adequadas.



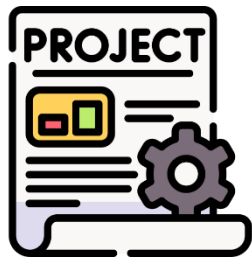
## DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

### Provisionamento de Recursos:

- Nesta etapa, os recursos de hardware e software são provisionados de acordo com o planejamento feito anteriormente. Isso pode envolver a compra ou alocação de servidores, sistemas de armazenamento, software de banco de dados, ferramentas de processamento de dados, entre outros.

### Configuração e Instalação:

- Uma vez provisionados, os recursos são configurados e instalados conforme as especificações do projeto. Isso pode incluir a configuração de servidores, a instalação e configuração de software, a configuração de redes e firewalls, entre outros.



AMBIENTE DEV/TESTE

COLETA DE DADOS

ARMAZENAMENTO  
&  
PROCESSAMENTO

TRANSFORMAÇÃO  
DE DADOS

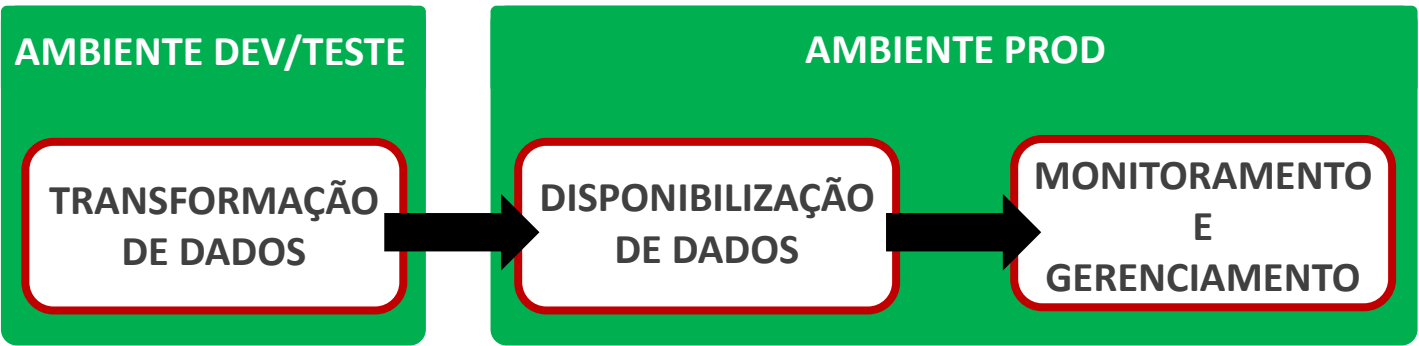
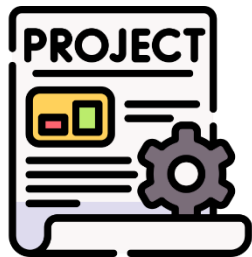
DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E  
GERENCIAMENTO DE SERVIÇOS DE DADOS

Integração de Dados:

- Os dados são integrados a partir de diversas fontes, como bancos de dados internos, sistemas legados, feeds de dados externos, etc. Isso pode envolver o desenvolvimento de pipelines de dados para extrair, transformar e carregar (ETL) os dados para a infraestrutura de armazenamento.

Implementação de Segurança:

- As medidas de segurança são implementadas para proteger os dados contra acesso não autorizado, perda de dados e outros riscos de segurança. Isso pode incluir a configuração de políticas de acesso, criptografia de dados, monitoramento de atividades suspeitas, entre outros.



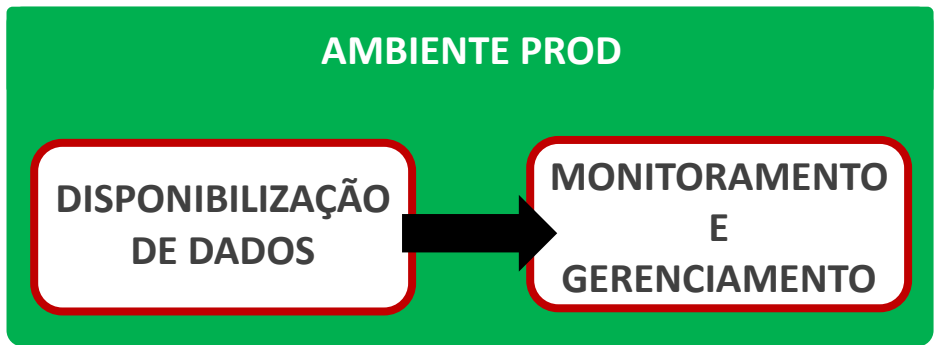
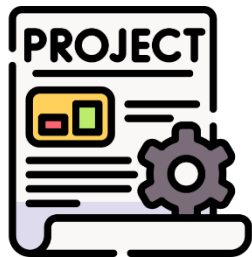
## DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

### Testes e Validação:

- Antes de disponibilizar os serviços de dados para uso, são realizados testes e validações para garantir que tudo esteja funcionando conforme o esperado. Isso pode incluir testes de integridade de dados, testes de desempenho, testes de segurança, entre outros.

### Implantação e Disponibilização:

- Uma vez que os testes sejam bem-sucedidos, os serviços de dados são implantados e disponibilizados para uso pelos usuários finais. Isso pode envolver a configuração de interfaces de consulta, APIs, ferramentas de visualização de dados, entre outros.



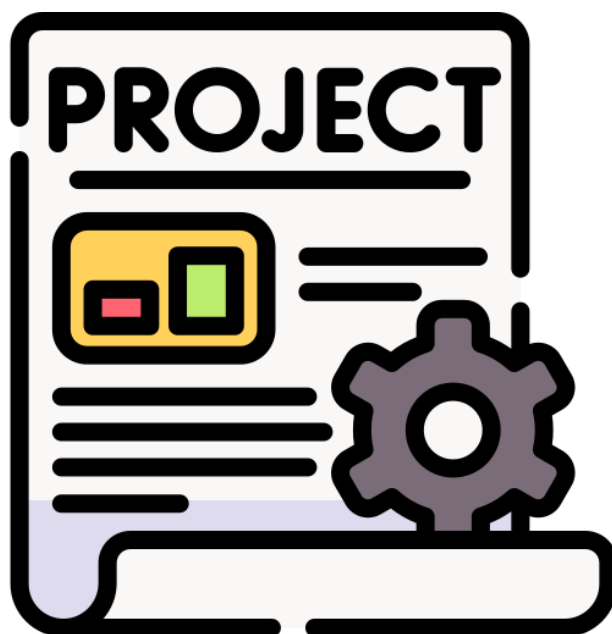
## DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

### Monitoramento e Manutenção:

- Após a implantação, os serviços de dados são monitorados continuamente para garantir seu desempenho, disponibilidade e segurança. São realizadas atividades de manutenção preventiva, correção de problemas e atualizações conforme necessário.

### Otimização Contínua:

- O processo de provisionamento e gerenciamento de serviços de dados é um ciclo contínuo. À medida que os requisitos e as tecnologias evoluem, são feitas otimizações contínuas na infraestrutura, nos processos e nos serviços de dados para garantir que atendam às necessidades em constante mudança da organização.



## DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

1. Definição de Requisitos;
2. Planejamento de Infraestrutura;
3. Provisionamento de Recursos;
4. Configuração e Instalação;
5. Integração de Dados;
6. Implementação de Segurança;
7. Testes e Validação;
8. Implantação e Disponibilização;
9. Monitoramento e Manutenção;
10. Otimização Contínua;



# ENGENHARIA DE DADOS

## TÓPICOS ABORDADOS

### Introdução à Computação em Nuvem:

- Definição de computação em nuvem.
- Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.
- Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).

### O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

### Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

### Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

### Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.

# ENGENHARIA DE DADOS

## VERDADEIRO OU FALSO

- 1 - O modelo DaaS permite aos usuários acessar dados sob demanda através da internet.
- 2 - DaaS é exclusivamente implementado no modelo de nuvem pública.
- 3 - No modelo DaaS público, os dados são armazenados e processados em infraestrutura própria da organização.
- 4 - DaaS Privado oferece maior controle sobre a segurança e privacidade dos dados.
- 5 - DaaS Híbrido combina elementos de infraestrutura local e serviços de nuvem pública.
- 6 - DaaS é exclusivamente utilizado por grandes empresas de tecnologia.
- 7 - O modelo DaaS público é mais adequado para organizações que precisam de controle total sobre seus dados.
- 8 - DaaS pode incluir serviços de integração de dados, limpeza de dados e enriquecimento de dados.
- 9 - No modelo DaaS Privado, os recursos de dados são compartilhados entre várias organizações.
- 10 - O modelo DaaS Híbrido oferece flexibilidade ao combinar infraestrutura local e serviços de nuvem pública.