

Computação em Nuvem

Conceitos Gerais

1. Introdução

- A **computação em nuvem** trata de um conjunto de recursos virtuais de fácil utilização e acesso, incluindo hardware, software, plataformas de desenvolvimento e serviços.
- Esse recursos podem ser dinamicamente reconfigurados para se ajustarem a uma carga de trabalho variável, buscando uma otimização do uso, e por consequência, do custo.
- Geralmente é utilizado um modelo “pague-pelo-uso”, ou seja, o cliente paga por determinadas quantias de recursos e pode aumentar ou diminuir de uma maneira simples (elasticidade).
- Além disso, é comum que o provedor do serviço de nuvem ofereça garantias ao cliente, conforme o pacote (ou plano) escolhido. Essa garanti apode ser de backup, disponibilidade, entre outras.

1. Introdução

- Um dos principais benefícios da computação em nuvem é a elasticidade. Em um servidor próprio, a empresa sempre precisa ter capacidade de atender a demanda máxima, mesmo que ela seja esporádica.
- Os datacenters que oferecem serviços de computação em nuvem possuem alta escala, com poder computacional muito grande, permitindo ajustes rápidos sempre que o cliente precisa.
- Por exemplo, um site que realiza em média 1000 vendas por dia resolve fazer uma grande promoção em um único dia e as vendas sobem para 50000.
- Nessa situação mais recursos (processamento, memória, banda de rede etc.) podem ser adquiridos apenas pelo período da promoção.

2. Características da Computação em Nuvem

- As características mais importantes da computação em nuvem são:



2. Características da Computação em Nuvem

- **Serviço sob demanda:** funcionalidades são providas automaticamente, conforme a solicitação do consumidor.
- **Amplo acesso aos serviços de rede:** recursos disponíveis através da Internet e podendo ser acessado por qualquer dispositivo (computador, celular, etc.)
- **Pool de recursos:** recursos computacionais (físicos ou virtuais) de um provedor são utilizados para servir diversos usuários, com alocação dinâmica, de acordo com a demanda do usuário;
- **Elasticidade rápida:** as funcionalidades devem ser rápidas e elasticamente providas e/ou liberadas, dando a impressão ao usuário que há recursos ilimitados
- **Medição de serviços:** cada tipo de serviço (processamento, armazenamento, largura de banda etc.) é monitorado e controlado, de forma clara e transparente ao provedor e ao consumidor.

3. Modelos de Serviço

- Basicamente existem três modelos de serviço (embora outros tenham surgido). Mas vamos focar nos três básicos definidos pelo NIST (National Institute of Standards and Technology):

1 - **SaaS (Software como Serviço):** É um modelo em nuvem onde o provedor hospeda e gerencia aplicações, disponibilizando-as pela internet. O usuário não compra licenças ou instala o programa no computador. O usuário paga periodicamente uma assinatura e acessa tudo diretamente pelo navegador ou aplicativo. A lógica é a mesma de um serviço de streaming. Exemplo: Office 365, google drive, Conta Azul (ERP online), etc.

2 - **PaaS (Platform as a Service ou Plataforma como Serviço):** é uma categoria de computação em nuvem que fornece um ambiente completo sob demanda para desenvolver, testar, gerenciar e entregar aplicativos.

Em vez do programador se preocupar com o gerenciamento da infraestrutura (servidores, redes, sistemas operacionais e armazenamento), ele paga pelo uso contínuo da plataforma e foca exclusivamente na criação do seu código.

3. Modelos de Serviço

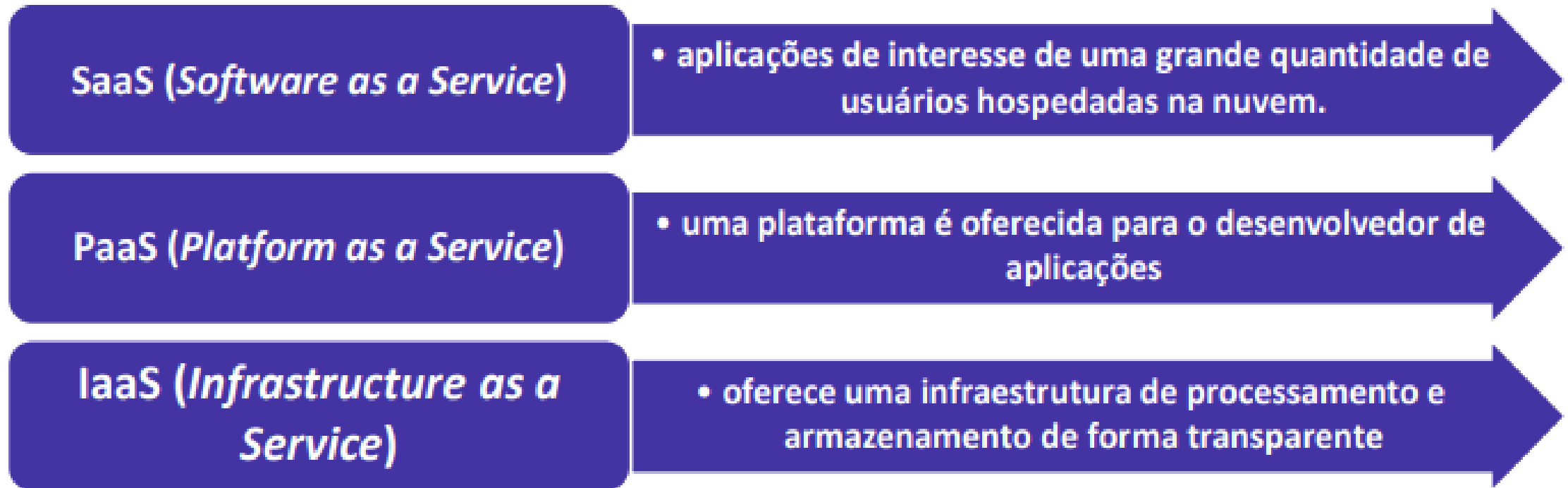
3 - **IaaS (Infrastructure as a Service ou Infraestrutura como Serviço)** é a categoria mais flexível da computação em nuvem, onde o usuário aluga os recursos fundamentais de computação diretamente de um provedor.

Em vez de comprar servidores físicos, espaço em datacenters ou equipamentos de rede, o usuário contrata esses recursos computacionais de forma virtualizada e paga apenas pelo que usar.

Diferente do PaaS (onde a plataforma já vem pronta), no IaaS o provedor entrega o "hardware virtualizado" em branco. O usuário recebe um servidor limpo e é o responsável absoluto por escolher e instalar o SO, configurar firewalls, redes e segurança, instalar bancos de dados e demais serviços necessários.

3. Modelos de Serviço

- A imagem abaixo mostra de forma resumida os três modelos de serviços:



3. Modelos de Serviço

- Como visto anteriormente, quanto mais “alto nível” (SaaS), menos o usuário sabe o que está sendo utilizado (S.O, redes etc.), e quanto mais “baixo nível”, mais controle pode se ter sobre o que utilizar. Essa é a ideia da pirâmide abaixo (lado esquerdo).
- No lado direito podemos ver uma pirâmide invertida, que mostra a quantidade de provedores dos três modelos de serviço, ou seja, quanto mais alto nível, existem mais provedores.



4. Modelos de Implantação da Nuvem

- Os modelos de implantação definem as opções de implantação de uma nuvem, além de quem poderá fazer uso dos serviços. São basicamente 3 modelos:

1 - **Nuvem privada:** infraestrutura de nuvem operada e quase sempre gerenciada pela organização do cliente. Os serviços são utilizados internamente pela organização e a configuração é realizada pela equipe de TI da empresa. A empresa decide quais usuários têm permissão para usar a infraestrutura.

2 - **Nuvem pública:** disponibilizada publicamente, modelo “pague-por-uso”. É oferecida por grandes grupos empresariais que possuem grande capacidade de processamento e armazenamento, como por exemplo Amazon, Microsoft, Google, Oracle, etc.

3 - **Nuvem híbrida:** duas ou mais nuvens (privadas, públicas ou comunitárias). Continuam sendo entidades únicas, porém são conectadas de forma que propiciem a portabilidade de dados e aplicações.

4. Modelos de Implantação da Nuvem

- Um outro modelo que vem ganhando popularidade é a **nuvem comunitária**. Trata-se de um modelo onde um grupo de empresas monta uma nuvem e dividem a responsabilidade e os custos de sua implantação.
- Em outros casos, uma empresa maior monta sua própria nuvem, e aluga os serviços excedentes para outras empresas menores.
- O objetivo da nuvem comunitária é reduzir custos, tendo em vista que o grande problema da computação em nuvem é seu elevado custo

5. Backup

- De forma simples, podemos definir backup como uma cópia de segurança. Essa cópia pode ser aquela manual que o usuário copia para uma mídia externa, por exemplo, ou algo mais elaborado, de forma automática e de acordo com alguns critérios pré-estabelecidos.
- Seguindo as três propriedades básicas da segurança da informação (CID – Confidencialidade, Integridade e Disponibilidade), o backup é um mecanismo que se encaixa na Disponibilidade, pois se os dados de um servidor são perdidos, eles podem ser recuperados.
- Com relação a periodicidade do backup, não existe uma resposta ideal, dependerá do valor dos dados. Se os dados forem extremamente valiosos e não admitirem nenhuma perda, o backup deve ocorrer em intervalos curtos (custo computacional alto).

5. Backup

- Outro aspecto importante a ser decidido é a **retenção**, ou seja, o tempo no qual os dados do backup precisam ficar armazenados. Se o tempo de retenção for grande, mais meios de armazenamento (HDs, fitas, SSDs etc.) devem ser adquiridos, aumentando o custo.
- Entretanto, não adianta ter o backup sem testar se ele realmente funciona. Por isso é fundamental que testes periódicos de **restore** (recuperação) sejam realizados, senão haverá a falsa sensação de segurança.
- Grandes empresas (que não podem parar) utilizam muito a técnica de **hot-site(ou site de contingência quente)** como backup. Essa técnica consiste em manter uma infraestrutura de TI espelhada e pronta para uso.
- Ele replica seus dados e sistemas em tempo real, permitindo que a operação seja assumida quase instantaneamente se a estrutura principal falhar.

5. Backup

- Também é muito comum em empresas que possuem grande quantidade de dados o uso de técnicas de rotação de mídia de armazenamento do backup.
- Uma das técnicas mais comuns é o **GFS (Grandfather-father-son – Vô-pai-filho)**. Vamos pensar assim: os filhos são mais ágeis, então os backups ocorrem mais frequentemente, os pais um pouco menos frequentes e os avós menos ainda. A configuração mais comum é:
 - 1 - Backups diários = “filhos” (7 mídias que armazena o backup diferencial).
 - 2 - Backups semanais = “pais” (4 mídias– 1 backup completo por semana).
 - 3 - Backups mensais = “avós” (12 mídias – uma para cada mês)

5. Backup

- Outro conceito importante dentro do contexto de backup é o **plano de contingência**. Ele consiste em uma importante ferramenta para minimizar o risco de inoperância dos serviços essenciais.
- O plano de contingência é um documento composto por uma sequência de procedimentos necessários para fazer com que os serviços voltem a funcionar o mais rápido possível.
- Este plano deve indicar as pessoas responsáveis por resolver o problema, locais onde estão os dados, observações técnicas, passo a passo dos procedimentos necessários, entre outras informações importantes.
- Outra recomendação importante é ter backup dos dados em **locais distantes** do servidor (prédios e até cidades diferentes). Além disso, deve-se levar em consideração o tempo de **restore**. Um exemplo de restauração rápida é através o uso de hot-site.

6. Deduplicação

- A **deduplicação** (alguns autores chamam de desduplicação) é o processo de analisar, identificar e remover a duplicidade nos dados. Ou seja, salva uma única cópia de dados idênticos e substitui todas as outras por apontadores (ou ponteiros).
- Imagine uma situação em existem 20 cópias do mesmo arquivo espalhados pelo disco. Quando for feito um backup de todo o disco, ele conterá apenas um arquivo e 19 apontadores para ele. Importante destacar que não se trata de compressão de dados.
- Com isso é possível economizar no uso do disco, tráfego de rede, processamento, tempo de backup e, principalmente, tempo de restore.

6. Deduplicação

- Sistemas de deduplicação se integram facilmente aos softwares de backup, mas não é todo sistema de backup que possui essa técnica. Existem 3 níveis de deduplicação:

1 - **Nível de arquivo:** uma "assinatura" (hash) é atribuída ao arquivo inteiro. É bom quando há interesse em comparar apenas arquivos inteiros.

2 - **Nível de blocos:** é o tipo mais comum. Possui um overhead maior que a deduplicação de arquivos, conseguindo comparar pedaços menores de dados (blocos).

3 - **Nível de byte:** apresenta o maior overhead, mas também o maior retorno. Analisa sequências de bytes (não atrelados a blocos predefinidos). Consegue obter o melhor resultado, substituindo tamanhos diferentes de sequências de dados repetidos por ponteiros.

6. Deduplicação

- Uma outra classificação possível é a deduplicação in-line x pós-processamento, conforme veremos a seguir:

1 - **In-line**: elimina os dados redundantes durante o processo de backup, antes de serem gravados (geralmente mais demorado);

2 - **Pós-processamento**: o backup completo é realizado e a análise e remoção das redundâncias são efetuadas em um processo posterior separadamente do backup (mais rápido). Necessita de mais espaço de armazenamento.

6. Deduplicação

- A deduplicação pode ocorrer na origem ou no destino:

1 - **Deduplicação na origem:** remove os dados redundantes em ambiente de produção, antes de serem enviados para o servidor de backup. Utiliza softwares que se comunicam com o servidor de backup, comparando os novos dados com os dados já armazenados.

Vantagem: um menor tráfego da rede e melhor desempenho no armazenamento.

Desvantagem: utiliza o poder de processamento do próprio servidor de produção, podendo afetar o desempenho de outras aplicações.

2 - **Deduplicação no destino:** o processamento e remoção dos dados redundantes são realizados após o seu envio pela rede. Causa muito menos impacto no processamento do ambiente de produção. Utiliza muita largura de banda. É ideal para ambientes de produção com altas cargas de trabalho ou que possuam uma janela de backup limitada. Ex.: servidores de banco de dados.

7. Snapshot

- Um **snapshot** é como uma *foto do sistema em um momento específico*. É como tirar um "print" em um dado instante do sistema ou disco. Com isso, é possível voltar o sistema exatamente como ele estava naquele momento.
- Por exemplo, as 10:00 o sistema estava funcionando e foi feito um snapshot. Por volta de 11 horas um usuário cometeu um erro e comprometeu o sistema. Basta voltar o snapshot das 10 horas.
- O snapshot **não copia tudo sempre**. Ele normalmente guarda só o que mudou para economizar espaço em disco.

8. Tipos de Backup

- Basicamente existem três tipos de backup, são eles:

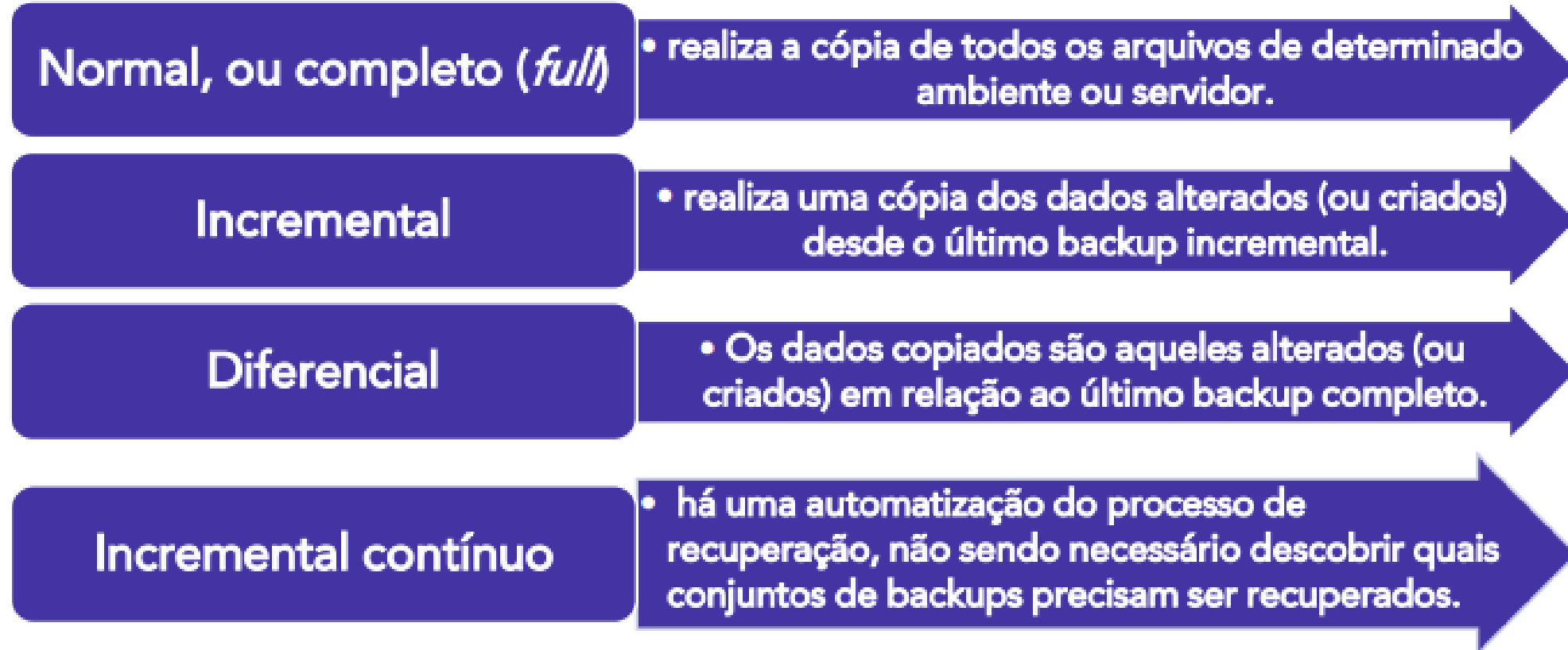
1 - **Normal, ou completo (full)**: realiza a cópia de todos os arquivos de determinado ambiente ou servidor, mesmo que menos de 1% dos dados tenha sido alterado desde o último backup.

2 - **Incremental**: realiza uma cópia dos dados alterados (ou criados) desde o último backup incremental. Em caso de restore, precisará dos arquivos do backup completo e todos os incrementais até a data que os dados foram perdidos.

3 - **Diferencial**: Os dados copiados são aqueles alterados (ou criados) em relação ao último backup completo. Em caso de restore, precisa apenas do backup completo e do último backup diferencial.

8. Tipos de Backup

- Existe um quarto tipo conhecido como Incremental contínuo, que oferece uma automatização do processo de recuperação, não sendo necessário descobrir quais conjuntos de backups precisam ser recuperados. A imagem abaixo resume os tipos de backup:

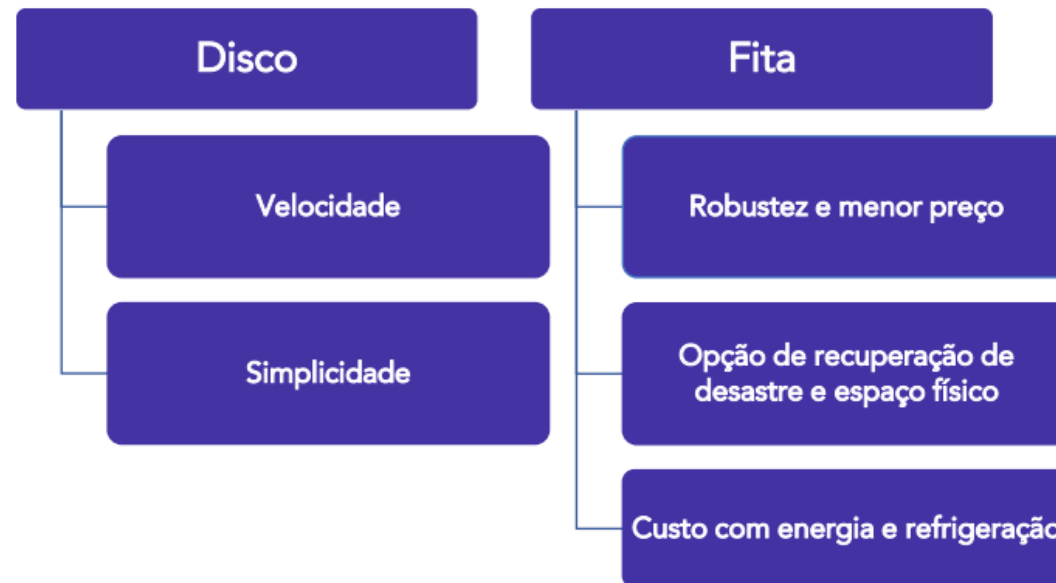


9. Meios de Arquivamento de Backup

- Os meios de armazenamento mais utilizados para backup e suas devidas vantagens são:

1 - **Disco**: velocidade, simplicidade;

2 - **Fita**: robustez, menor preço, opção de recuperação de desastre, espaço físico, custo com energia e refrigeração.



9. Meios de Arquivamento de Backup

- As fitas tipicamente utilizadas para backup são as LTO (Linear Tape-Open). Trata-se de um padrão aberto de tecnologia de armazenamento de dados em fita magnética.
- Segundo alguns fabricantes, elas possuem vida útil de até 30 anos (com um manuseio e ambiente adequados).
- Grandes empresas (bancos, governos) utilizam fitas para armazenamento de backup por grandes períodos, principalmente devido ao seu baixo custo e alta durabilidade.

