



CLOUD COMPUTING

Texto base

3

Virtualização de armazenamento

Prof. Me. Rodolfo Riyoei Goya

Resumo

A “Computação em Nuvem” baseia-se em diversas tecnologias para diversas finalidades, dentre elas o armazenamento de dados. Diversas tecnologias são usadas para realizar a virtualização de armazenamento. Abordam-se aqui os principais princípios e conceitos básicos para se entender estas tecnologias.

3.1. Introdução

Como a “Computação em Nuvem” lida com o armazenamento de dados? Como as máquinas virtuais oferecem armazenamento de modo virtual?

Informações armazenadas em meio persistente são tão ou até mesmo mais importantes que o próprio processamento.

Atualmente, as tecnologias de armazenamento mais utilizadas são baseadas em discos magnéticos (HDD) ou em memórias flash de semicondutores (SSD). Nos sistemas de virtualização de armazenamento, uma dada capacidade física de armazenamento pode ser disponibilizada na forma de blocos de dados virtuais. Há, basicamente, quatro principais arquiteturas para implementar essa virtualização: Direct Attached Storage - DAS, Storage Area Network – SAN, Network Attached Storage – NAS e armazenamento de objetos através da rede.

3.2. Direct Attached Storage – DAS

A forma mais básica de armazenamento virtual usa parte do armazenamento interno da mesma máquina física (em inglês: Direct Attached Storage - DAS) como local onde a máquina virtual é criada e executada (Figura 1). Cada local é reservado e protegido das demais máquinas virtuais.

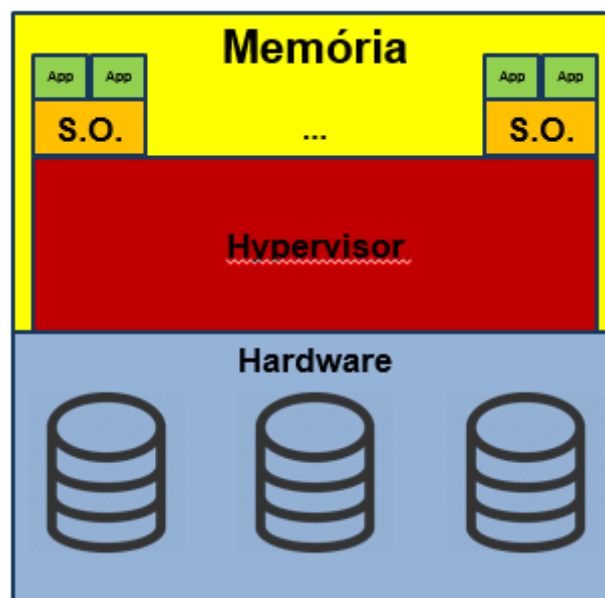
O tamanho da área alocada pode ser redimensionado na medida da necessidade, mas é limitado ao tamanho do armazenamento físico (aumentar acima deste limite exige desligar o servidor e está limitado à capacidade interna dele). Comparada as demais

formas de virtualização (que usam armazenamento externo) esta tem maior desempenho, mas tem um compartilhamento limitado por negar acesso a máquinas virtuais executadas em máquinas físicas diferentes.

Ao se terminar e reiniciar uma instância de máquina virtual, nos provedores de serviço de computação em nuvem, não é comum que uma nova máquina, mesmo com a mesma imagem, seja criada na mesma máquina física. Deste modo, os dados em um DAS em armazenamento local virtual não sobrevivem ao término da máquina virtual. De fato, é comum que toda a área de armazenamento local seja apagada antes de ser disponibilizada para uma nova máquina virtual para proteger os dados dos processamentos anteriores.

Deste modo, o caso de uso mais comum para o armazenamento local é como local para armazenamento temporário de arquivos e dados.

Figura 1. Armazenamento virtualizado em armazenamento local

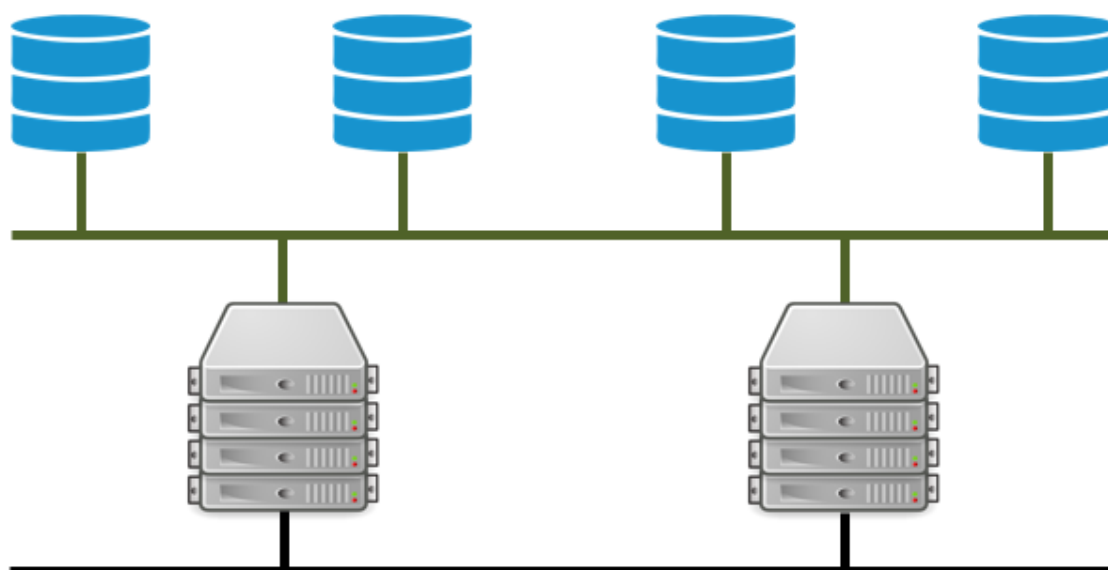


Fonte: do autor, 2022.

3.3. Storage Area Network – SAN

No Storage Area Network (SAN), o armazenamento é feito em um conjunto de dispositivos externos (de modo diferente do Armazenamento Local, são montados fora dos servidores) conectados entre si e ao conjunto de servidores para os quais são compartilhados (Figura 2).

Figura 2. Arquitetura SAN



Fonte: do autor, 2022.

Esta estrutura permite compartilhamento do recurso físico entre máquinas virtuais e manutenção e ampliação simples por ser “hot swappable”, ou seja, podem ser substituídos sem a necessidade de desligamento (Figura 3). Este arranjo permite capacidade de expansão escalável e praticamente ilimitada (diferentemente do DAS, limitado pelo espaço interno dos servidores).

Figura 3. HDs externos montados em um rack



Fonte: ACNODES, s.d.

3.3.1. Padrões de interconexão SAN

A conexão entre os dispositivos físicos de armazenamento é feita de modo padronizado, com os padrões SCSI (costuma-se dizer “scuzi”), FCP, iSCSI, AoE e FCoE sendo os mais usados.

Os meios físicos mais usados são os pares metálicos e fibras ópticas. A maioria destes padrões diferem dos de rede local, mas alguns estão aderindo aos padrões usados pelo Gigabit Ethernet e seus sucessores.

Os padrões existentes permitem conexões da ordem de poucos metros e não dispõem de suporte para protocolos equivalentes à camada de rede, o que restringe seu uso a redes LAN dentro de um datacenter.

3.3.2. Organização

O armazenamento é feito em meio magnético (HDD) ou flash (SDD) e organizado em blocos que constituem a unidade de movimentação e armazenamento. Quando alguma informação é modificada, o bloco inteiro é tido como modificado (para backup, por exemplo). Quando se usa espelhamento ou RAID para desempenho e tolerância a falhas, a unidade de implementação é o bloco. Cada arquivo armazenado ocupa um número inteiro de blocos: arquivos diferentes não compartilham blocos.

Cada área de armazenamento virtual alocada para uma máquina virtual é constituída de um número inteiro de blocos distribuídos entre diversas mídias físicas para paralelizar o acesso, aumentando a vazão.

3.3.3. Sistema de arquivo

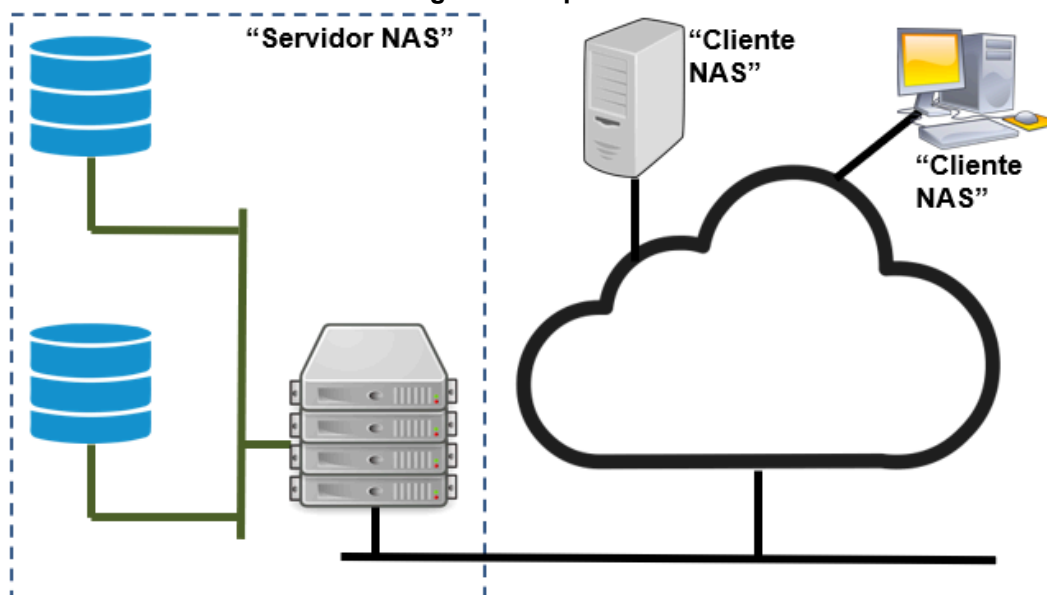
Cada área criada como armazenamento virtual é vista como se fosse um equipamento virtual. Com isso, cabe a máquina virtual que recebe o armazenamento fazer a formatação e definir o sistema de arquivos a usar.

Uma área de armazenamento virtual em SAN pode, inclusive, conter área de boot e configuração de sistemas operacionais, podendo ser desconectada de uma máquina virtual e conectada a outra e ser facilmente replicada para produzir máquinas virtuais idênticas.

3.4. Network Access Storage – NAS

No serviço de virtualização de armazenamento em nuvem baseado em Network Access Storage – NAS, um servidor compartilha sua grande capacidade de armazenamento oferecendo serviço de armazenamento virtualizado para seus clientes.

Figura 4. Arquitetura NAS



Fonte: do autor, 2022.

3.4.1. Interconectividade

O serviço é oferecido através de protocolos de camada de aplicação TCP/IP. Com isso, as máquinas que usam o serviço podem estar geograficamente distantes dos dispositivos físicos de armazenamento.

Os protocolos de aplicação mais usados são o Network File System – NFS (Unix/Linux) e o Server Message Block – SMB (Windows).

3.4.2. Organização e sistema de arquivos

A infraestrutura física dos NASs é muito parecida com a dos SANs. Com dispositivos físicos montados externamente em racks, apresenta um escalabilidade com crescimento praticamente ilimitado e fácil manutenção por serem “hot swappable”. A principal diferença é que os clientes não interagem diretamente com os dispositivos físicos: esse acesso é intermediado por um processador.

Por ser um protocolo de alto nível, o NAS comumente oferece áreas de dados montadas nos clientes como pastas ou diretórios que já estão configuradas para algum sistema de arquivos. Estas áreas estão protegidas a nível de permissões de grupo e usuário e podem ser usadas de modo compartilhado entre múltiplos clientes.

3.5. Armazenamento de Objetos

Nos chamados serviços de armazenamento de objetos, uma plataforma de Software as a Service (SaaS) oferece serviços de valor agregado sobre armazenamento de objetos. Há muitos exemplos comercialmente conhecidos de armazenamento de objetos, tais como o Google Drive (Google), Dropbox, OneDrive (Microsoft) e S3 (AWS).

Tais serviços frequentemente fazem interface tanto com o usuário final como com programas (através de Application Programming Interfaces – APIs) usando protocolos padronizados como, por exemplo, HTTP, SOAP e REST. Isso permite que sejam oferecidos pela Internet para clientes de muitos tipos diferentes.

Como exemplo de serviços de valor adicionado, podem ser citados: backup automático, controle de acesso discricionário por usuário e grupo, controle de versões, proteção em repouso com criptografia e controle de ciclo de vida.

Por ser oferecido tipicamente na forma de Software as a Service, o gerenciamento da infraestrutura e plataforma onde é executado é transparente para o cliente. Assim aspectos como provisionamento de instância, mídia, escalonamento de processadores, balanço de carga e alta disponibilidade são cuidados pelo provedor sem que o cliente precise se preocupar.

3.6. Que tipo de armazenamento de nuvem usar?

A Tabela 1 resume as características e os casos de uso de cada tipo de armazenamento em nuvem.

Tabela 1. Variáveis consideradas na escolha de armazenamento em nuvem.

Tipo	Característica	Caso de uso
DAS	Baixo custo	Dados de uso temporários e que podem ser descartados ao término da instância
	Excelente desempenho	Sem compartilhamento com outras máquinas virtuais em outros servidores físicos
	Dados perdidos no desligamento da instância	Escalabilidade limitada ao volume que cabe em uma instância física de servidor.
SAN	Baixo custo	Pode ser usado como partição de boot, configuração, swap de Sistemas Operacionais.
	Armazenamento em rede local	Sem compartilhamento com outras máquinas
	Sobrevive a desligamento de instâncias	Altamente escalável
NAS	Armazenamento em WAN	Dados compartilhados em grupo
	Afetado pelo desempenho de rede	Segurança por usuários e grupos
	Sobrevive a desligamento de instâncias	Altamente escalável
Objetos	Acesso por protocolo de aplicação	Acesso vindo de diferentes plataformas, programas e pelo cliente final
	Serviços adicionais	Arquivos que requerem segurança discricionária por objeto, usuários e grupos
	Gerenciamento delegado ao provedor. Não requer instância	Altamente disponível e escalável

Fonte: do autor, 2022.

3.7. Vamos praticar?

3.7.1. Níveis gratuitos

A plataforma de Computação em Nuvem AWS oferecem nível gratuito de uso para diversos serviços, inclusive armazenamento. Além do DAS nas instâncias de EC2,

o nível gratuito inclui um ano de uso de 5GB de NAS (EFS), 30GB de SAN (EBS) e 5GB de armazenamento de objetos (S3). Navegue pelos links abaixo, ative o filtro para armazenamento e veja o que é possível fazer nestas plataformas e aproveite.

<https://aws.amazon.com/pt/free/>

3.7.2. Quanto custa?

A maioria das plataformas comerciais de Computação em Nuvem oferecem serviços de armazenamento na nuvem. Navegue pelo link abaixo, escolha uma região, defina um tamanho, deixe o resto com os defaults e veja quanto custaria mensalmente ativar um serviço de armazenamento de bloco na AWS.

<https://calculator.aws/#/>

Referências

- ACNODES. **RMS4836**. Acnodes, s.d. Disponível em: <<https://www.acnodes.com/rms4836.htm>>. Acesso em: 17 jan. 2022.
- ANDREWS, Joshua; HALL, Jon. **VMware certified professional data center virtualization on vSphere 6.7 study guide**: Exam 2V0-21.19. New Jersey: Sybex, 2020. 640p.
- MARSHALL, Nick; BROWN, Mike; BLAIR FRITZ, G.; JOHNSON, Ryan. **Mastering VMware vSphere 6.7**. New Jersey: Sybex, 2019. 848p.
- OFFICIAL AMAZON WEB SERVICES (AWS) DOCUMENTATION. **Amazon simple storage service**: user guide. Amazon, s.d.a. 1.287p. Disponível em: <<https://docs.aws.amazon.com/s3/>>. Acesso em: 17 jan. 2022.
- OFFICIAL AMAZON WEB SERVICES (AWS) DOCUMENTATION. **Amazon elastic block store (Amazon EBS)**. Amazon, s.d.b. 1.774p. Disponível em: <<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html>>. Acesso em: 17 jan. 2022.
- OFFICIAL AMAZON WEB SERVICES (AWS) DOCUMENTATION. **What is Amazon elastic file system?** Amazon, s.d.c. 379p. Disponível em: <<https://docs.aws.amazon.com/efs/latest/ug/whatisefs.html>>. Acesso em: 17 jan. 2022.
- SANTOS, Tiago. **Fundamentos da computação em nuvem** (Série Universitária). São Paulo: Editora Senac, 2018. 211p.
- TAURION, Cezar. **Cloud computing**: computação em nuvem: transformando o mundo da tecnologia da informação. Rio de Janeiro: Brasport, 2009.
- VELTE, Anthony T.; VELTE, Toby J.; ELSENPETER, Robert. **Cloud computing**: a practical approach. EUA:McGraw-Hill, 2010.