

Megadados

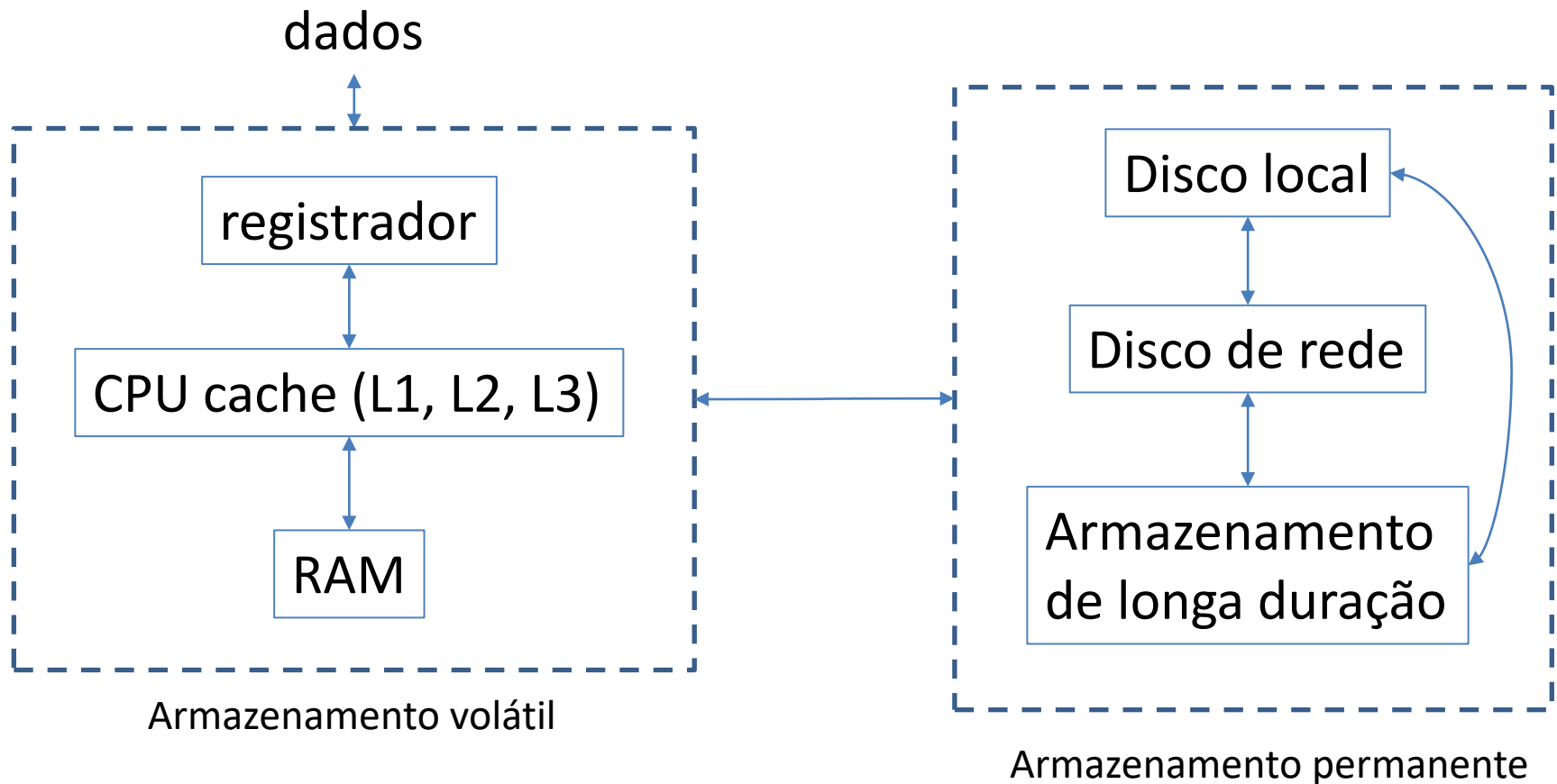
Aula 17 – Sistemas de armazenamento

2020 – Engenharia

Fábio Ayres <fabioja@insper.edu.br>

Armazenamento persistente

A volta da “Hierarquia de Memória”!



Disco local

| | |
|------------|--|
| Uso | Sistema operacional, arquivos locais, cache de dados de rede |
| Exemplo | HD, SSD, CD, USB drive |
| Tamanho | Baixo – alguns TB |
| Latencia | baixa |
| Throughput | alto |
| Custo | alto |

Array de discos locais

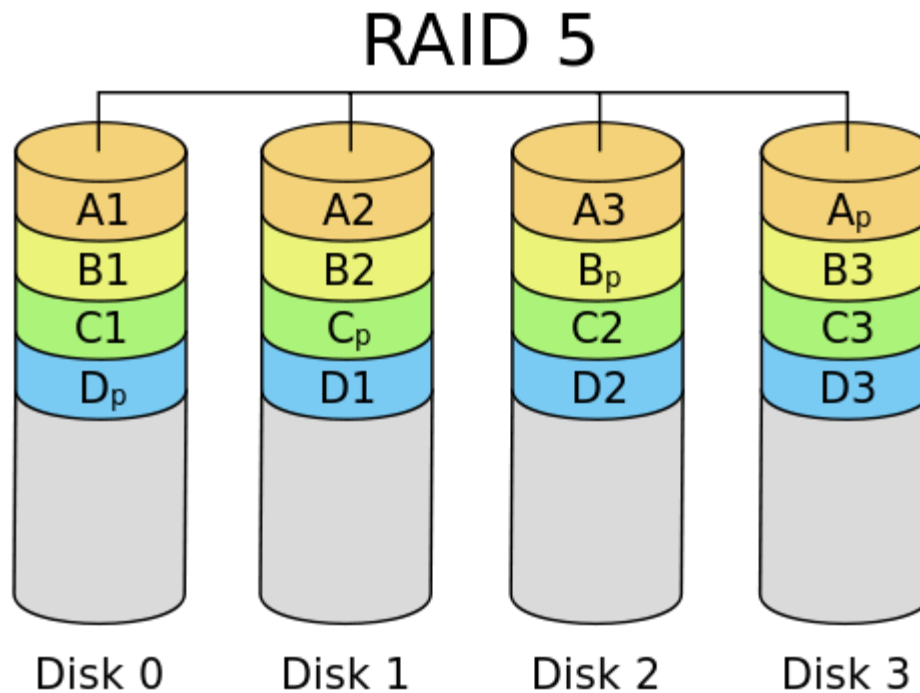
| | |
|------------|--|
| Uso | Multimedia, alto desempenho, alta confiabilidade |
| Exemplo | RAID |
| Tamanho | Medio – centenas de TB |
| Latencia | baixa |
| Throughput | alto |
| Custo | alto |

RAID

- Redundant Array of Independent Disks
 - Vários discos físicos combinados em um único disco lógico
- Várias configurações diferentes possíveis
- Objetivos: aumentar performance e confiabilidade

<https://en.wikipedia.org/wiki/RAID>

Exemplo



https://en.wikipedia.org/wiki/Standard_RAID_levels

Network File System

| | |
|------------|--|
| Uso | Informações compartilhadas em rede local |
| Exemplo | NFS, Amazon EFS |
| Tamanho | medio-alto / centenas de TB (exceto serviços elásticos como AWS EFS – pode escalar monstro) |
| Latencia | media |
| Throughput | medio |
| Custo | Medio-alto |

Network File System (NFS)

- Permite que arquivos armazenados remotamente sejam acessados como se fossem arquivos locais
- Desenvolvido originalmente pela Sun Microsystems
 - Os mesmos autores do Java
 - Venda para a Oracle em 2010



Server Message Block (SMB)

- Outro sistema de arquivos de rede, geralmente associado a sistemas Windows
- Conhecido no mundo Linux como Samba, a implementação open-source do SMB

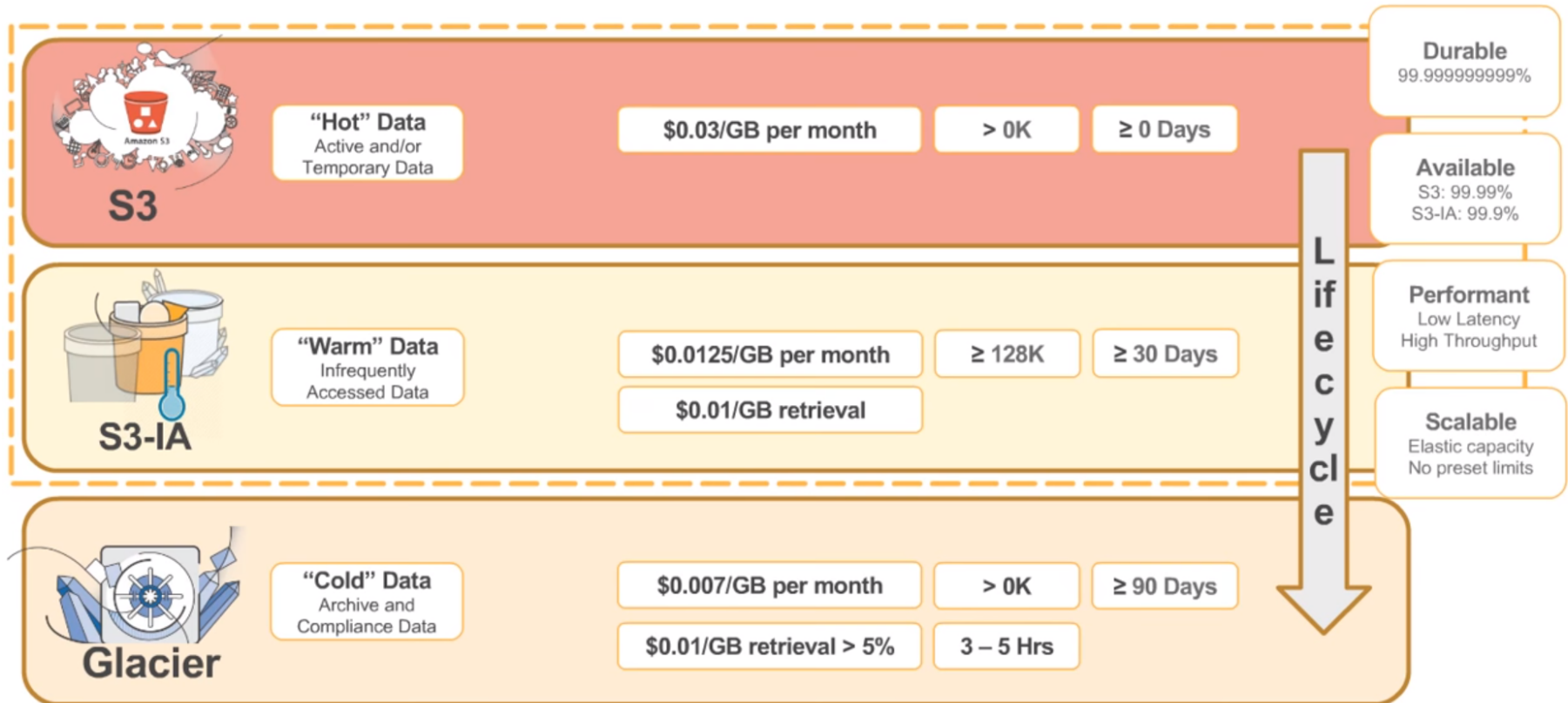
Armazenamento em nuvem

| | |
|------------|---|
| Uso | Informações compartilhadas em núvem |
| Exemplo | Amazon S3 |
| Tamanho | Monstro - de vários petabytes a exabytes! |
| Latencia | Depende: alta para fora do datacenter, media-baixa dentro do datacenter |
| Throughput | Depende: alto para fora do datacenter, media-baixa dentro do datacenter |
| Custo | Médio-baixo |

Armazenamento de longa duração (vaults)

| | |
|------------|---|
| Uso | Backup gigante |
| Exemplo | Amazon Glacier |
| Tamanho | Monstro - de vários petabytes a exabytes! |
| Latencia | Altíssima – de minutos a horas |
| Throughput | N/A – uso offline |
| Custo | baixo |

Storage tiered to your requirements



<https://www.youtube.com/watch?v=3iCSrkkXwal>

AWS EBS

- Amazon Web Services: Elastic Block Storage
- Um disco virtual, não formatado, para uso na AWS Elastic Compute Cloud (EC2)
- Pode ser “mounted” nas instâncias EC2

AWS EFS

- AWS Elastic File System
- O NFS da Amazon, com elasticidade
- Pode ser “mounted” nas instâncias EC2

AWS S3

- AWS Simple Storage Service
- Sistema de armazenamento em núvem
- Acesso via API

AWS Glacier

- Sistema de armazenamento de longa duração
- Voltado para backups, armazenamento por demandas regulatórias, etc.
- Acesso via API

Exercício

(Prova Final 2019-2)

Em um *cluster* temos vários arquivos enormes, de tamanho médio 1 GB. Todos os arquivos são do tipo *write-once-read-many*: uma vez que o arquivo é criado, ele pode apenas ser removido por completo, nunca modificado parcialmente. Esse é o modo de operação do S3 da AWS.

Os arquivos são divididos em blocos de 64 MB e estão armazenados de modo redundante, com fator de replicação 3 (este é o modo padrão do HDFS). Ou seja, cada bloco é armazenado 3 vezes, em máquinas diferentes.

Desejamos armazenar 10000 arquivos destes, em máquinas com 4 discos de 1TB SSD cada, nas quais queremos que os dados ocupem no máximo 75% do espaço.

Toda semana parte dos arquivos será processada usando Spark para alguma atividade de extração de informação.

Exercício

- a) (0,5 pts) Quais as vantagens e desvantagens da replicação em máquinas diferentes, neste contexto?
- b) (0,5 pts) Porque *write-once-read-many* é desejável aqui?
- c) (0,5 pts) Quantas máquinas este cluster deve ter, no mínimo, para acomodar esses dados?
- d) (0,5 pts) Mais dados chegam constantemente ao sistema, que eventualmente ficará sem capacidade! Mas uma vez que a pipeline Spark é executada, os dados não são mais necessários de imediato. Proponha uma solução simples para não perder dados e não ter que aumentar indefinidamente o cluster.

Exercício

(Prova Final 2018-2)

(1 pt) Descreva um sistema de armazenamento adequado para edição de vídeo, onde temos grandes massas de dados e a necessidade de transferi-los em alta velocidade numa rede local.

- Estime capacidade e velocidade de transferência para editar vídeos full HD (1920x1080, 3 quadros por segundo, 24-bit RGB) com taxa de compressão típica de 30% (ou seja, o tamanho comprimido é 30% do tamanho original).
- Cite tecnologias adequadas para este sistema.

Exercício

(Prova Final 2017-2)

(contexto: nesta prova estávamos falando de uma rede social fictícia da start-up na qual você trabalha)

g) (1 pt) A rede social está fazendo um sucesso inacreditável! Estima-se que daqui a 6 meses a rede terá 2 milhões de usuários, cada um postando uma média de 10kB por dia de conteúdo (fotos, áudio, texto, etc). Pelas regras desta rede social, todo o material postado nos últimos 12 meses estará rapidamente acessível, (em questão de milissegundos). A empresa deseja manter registros históricos de todo o material mais antigo (após 12 meses), indefinidamente. Projete uma solução de armazenamento para a empresa, indicando o tamanho esperado do(s) espaço(s) de armazenamento.

Insper

www.insper.edu.br