

System Design

RAID e Storage



\$ whoami

Matheus Fidelis

Engenheiro de \$RANDOM

@fidelissauro

<https://fidelissauro.dev>

<https://linktr.ee/fidelissauro>



OBJETIVOS

-  REFORÇAR TEORIA
-  REPRESENTAR CONCEITUALMENTE
-  FALAR SOBRE PROVIDERS
-  DETALHAR TECNOLOGIAS DE FORMA NOMINAL
-  DETALHAR FERRAMENTAL *

Detalhar os conceitos com a ótica de System Design, abstraindo suas implementações.

1

STORAGE EM SYSTEM DESIGN

Definições e
Conceitos

STORAGE

- Persistência de dados
- Recuperados e acessados
- Armazenamento por longos períodos
- Segura e performática
- Desempenho, segurança, latência e escalabilidade
- Capacidade de expansão
- HDD's, SSD's, Sistemas de Rede e etc



2. Storage

MÉTRICAS E DIMENSÕES

Definições e
Conceitos

Throughput em Storage

- Operações / Tempo
- Dados transferidos para a unidade de armazenamento
- Quantidade de leitura e escrita
- Megabytes por segundo (MB/s)
- Gigabytes por segundo (GB/s)

Bandwidth em Storage

- Largura de Banda
- Quantidade máxima de dados do canal de comunicação
- Representa o teto do Throughput
- Megabytes por segundo (MB/s)
- Gigabytes por segundo (GB/s)

I/O e IOPS em Storage

- Input/Output Operations Per Second
- Operação de escrita e leitura
- Entre um sistema e o volume de armazenamento ao qual o mesmo tenha acesso
- Persistir arquivos ou linhas de dados para serem tratados em disco
- Recuperar dados no storage para processar ou exibir



3. Storage

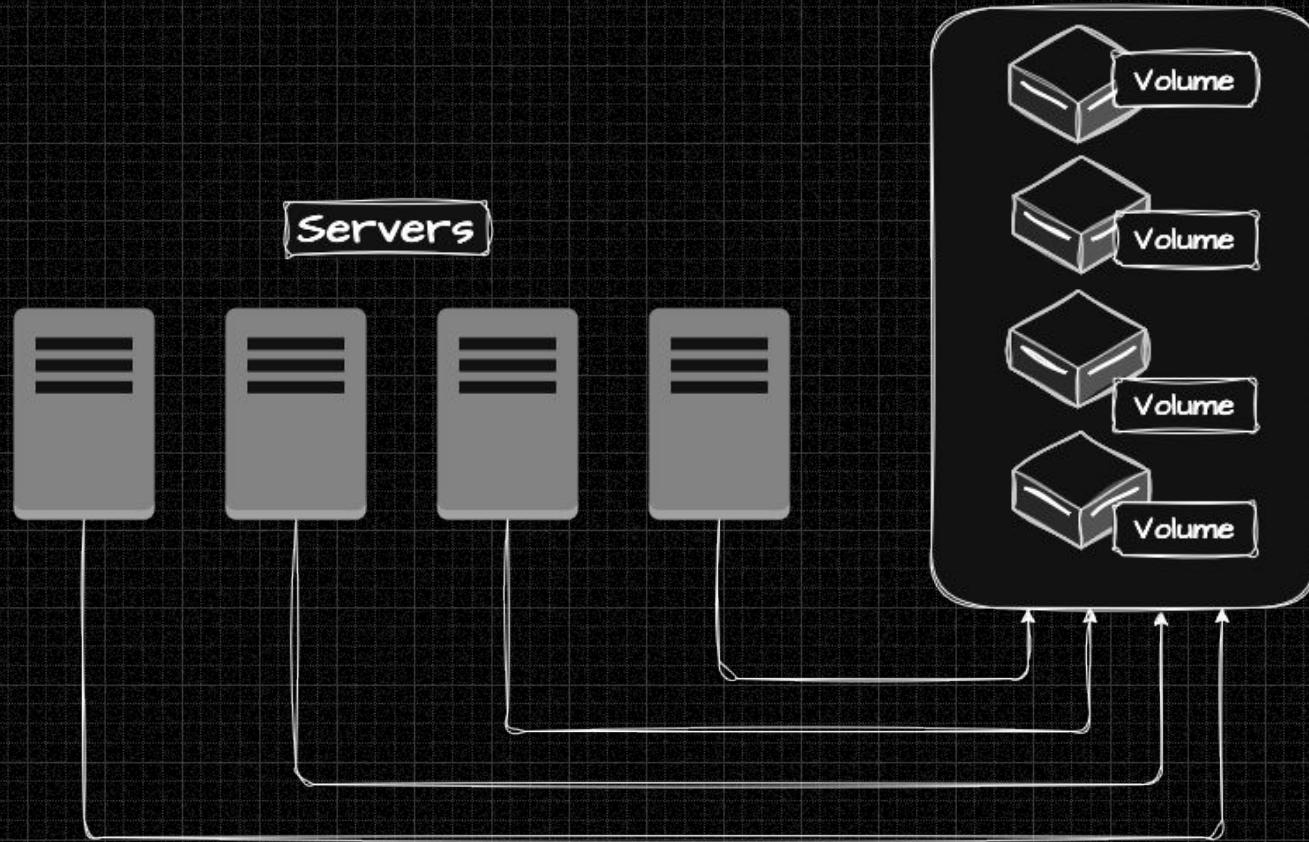
MODELOS DE STORAGE

Definições e
Conceitos

DAS - Direct-Attached Storage

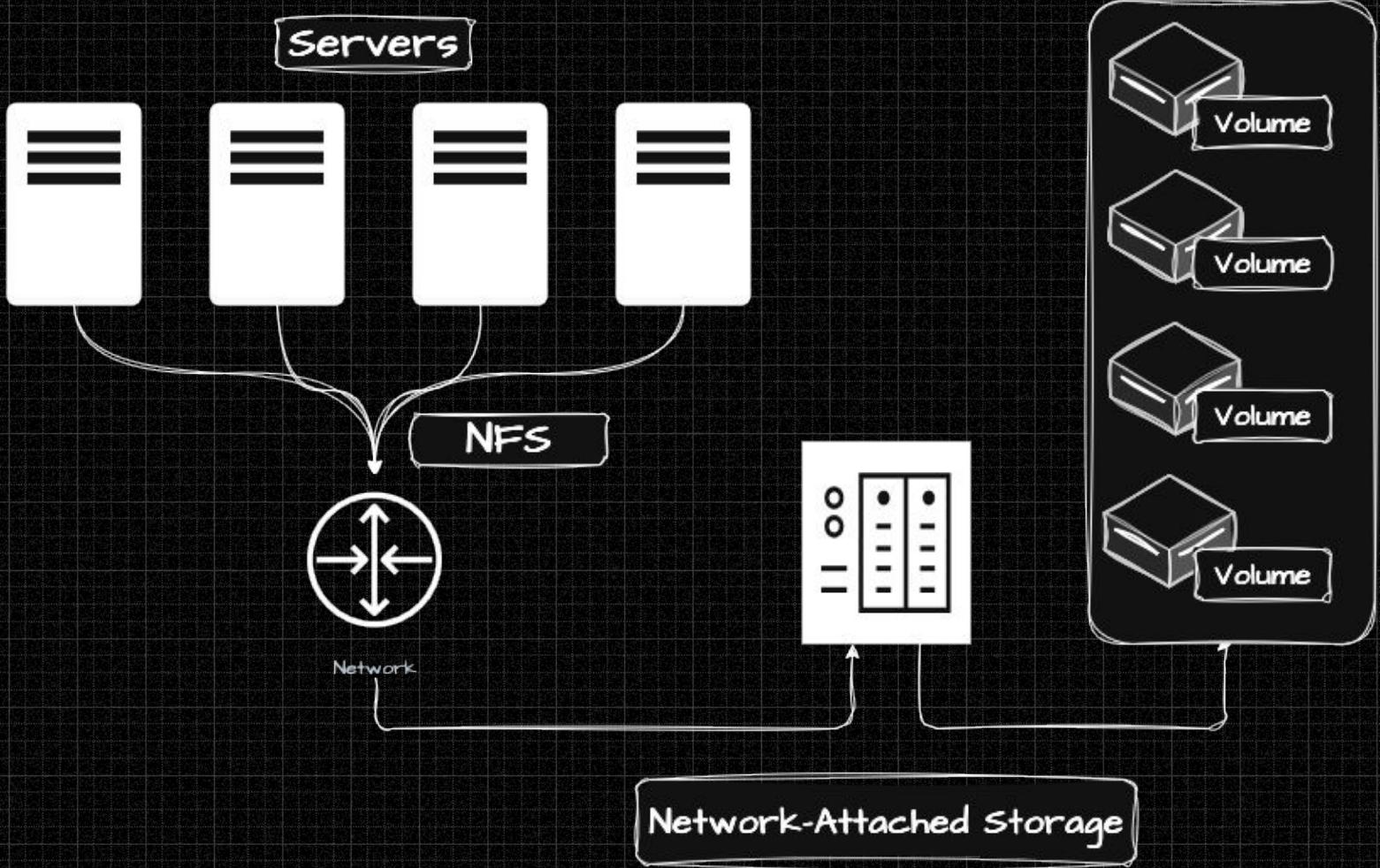
- Modelo mais tradicional
- Discos são alocados diretamente no host
- Interfaces SATA, USB ou NVMe
- Acesso direto
- Sem latência adicional
- Sem intermediações de dispositivos de rede ou troca de protocolos complexos
- Acessos frequentes e intensos no filesystem

Direct-Attached Storage



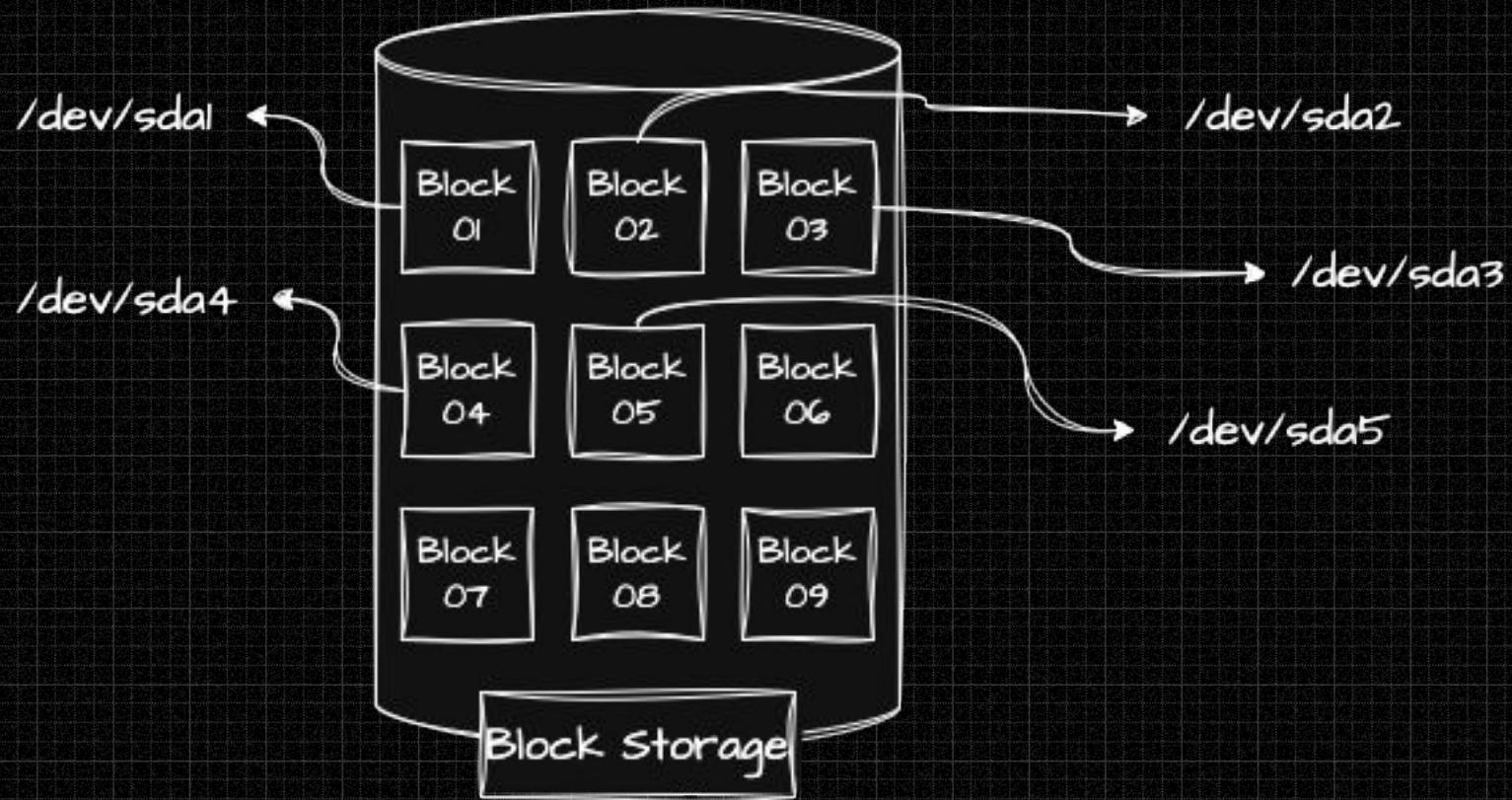
NAS - Network-Attached Storage

- NFS (Network File System)
- Dispõem de seus dados na rede local
- Múltiplos clientes se conectam ao mesmo volume
- Modifiquem os mesmos dados de forma simultânea
- Compartilhamento de diretórios corporativos (SMB)



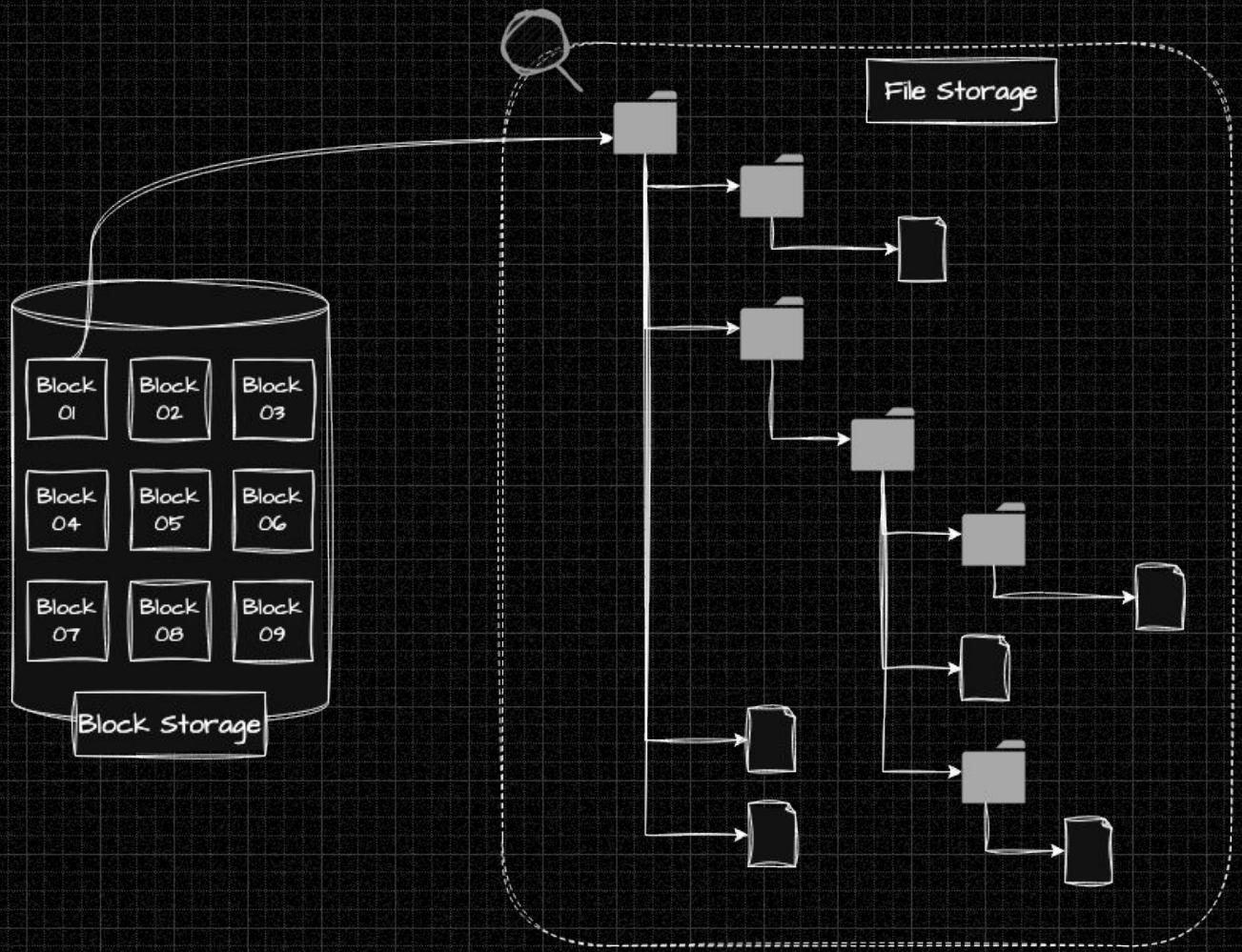
Block Storage

- Permitem armazenar informações dispersas em forma de blocos por todo o volume
- Acesso diretamente pelo SO e montagem como um drive
- O servidor responsável por gerenciar os blocos pode formatá-los e utilizá-los como sistemas de arquivos
- Endereçado de forma única e organizado de forma individual
- Disco isolado, porém podendo ser tratado de forma virtual



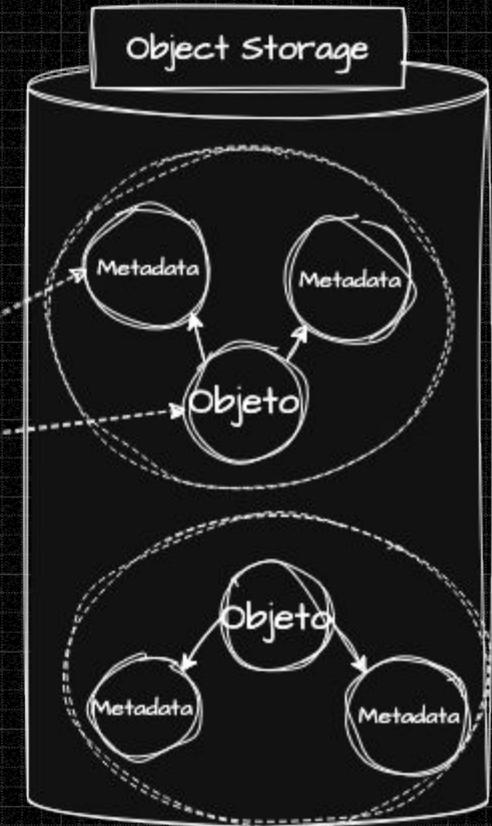
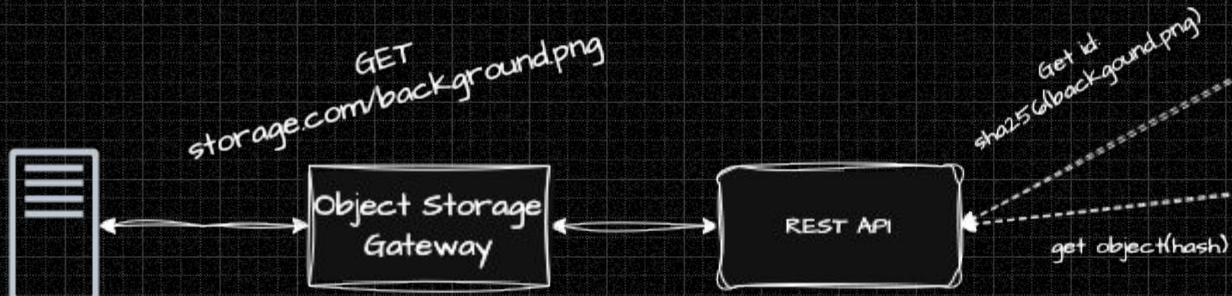
File Storage

- File-level ou file-based storage
- Sistemas de estruturas hierarquicas de diretórios e arquivos
- Nome do arquivo com a estrutura hierárquica da pasta forma o identificador único
- Metadados importantes associados
- Esses sistemas são configurados por meio de aplicações de RAID, e são anexados a sistemas NAS (Network-Attached Storage)



Object Storage

- Layer 7
- Armazenamento de Objetos
- Abordagem altamente escalável (API's)
- Abordagem altamente abstraída
- Grandes quantidades de dados de forma totalmente desacoplada da aplicação
- Conteúdo e metadados
- Cloud Providers
- Backup, Replicação, Lifecycle transparente
- Arquiteturas cloud native
- Alto desacoplamento





4. Storage

MODELOS DE RAID

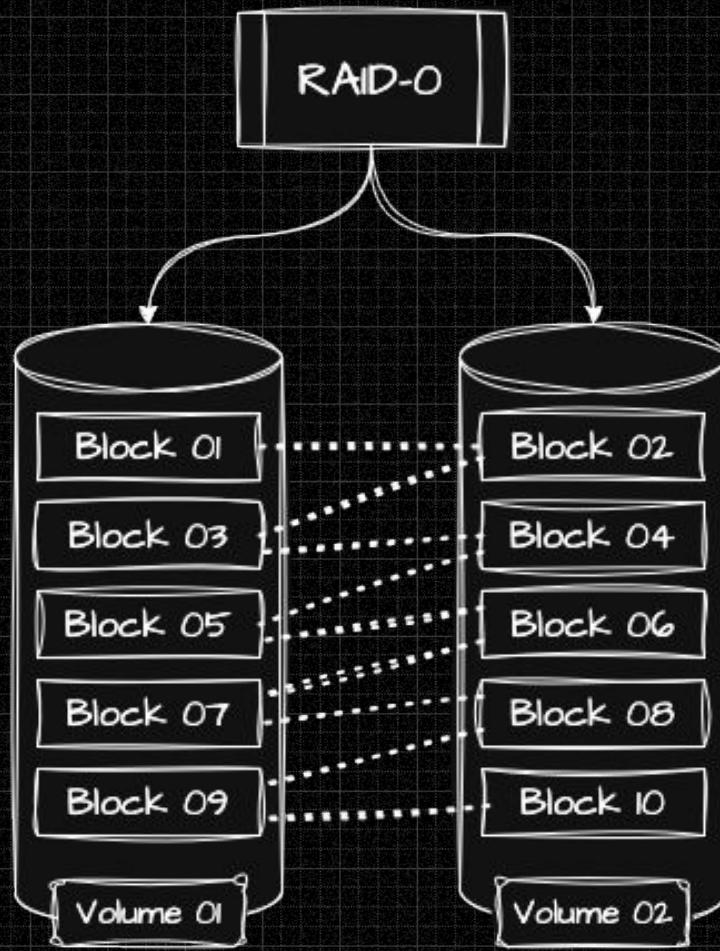
Definições e
Conceitos

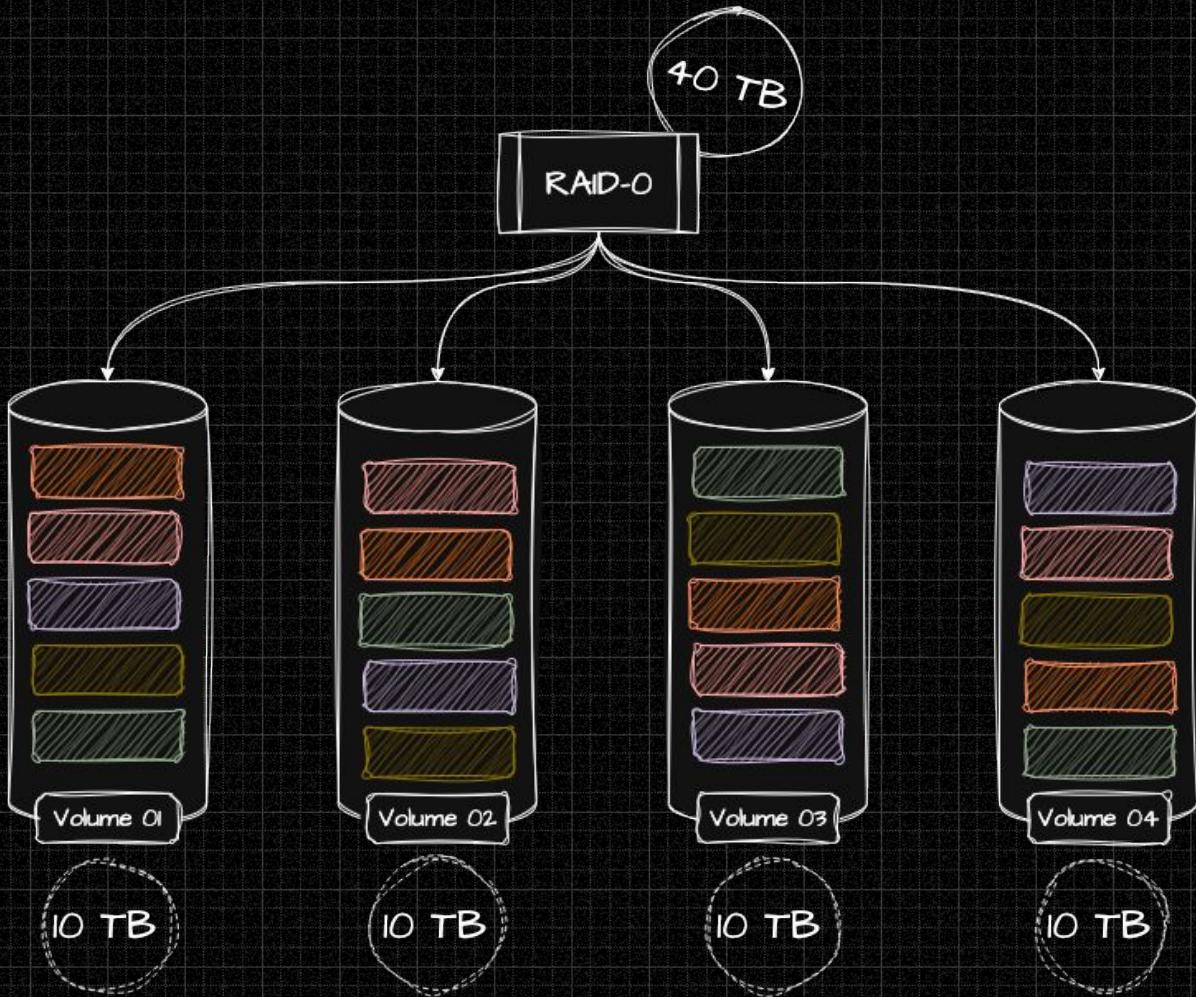
RAID

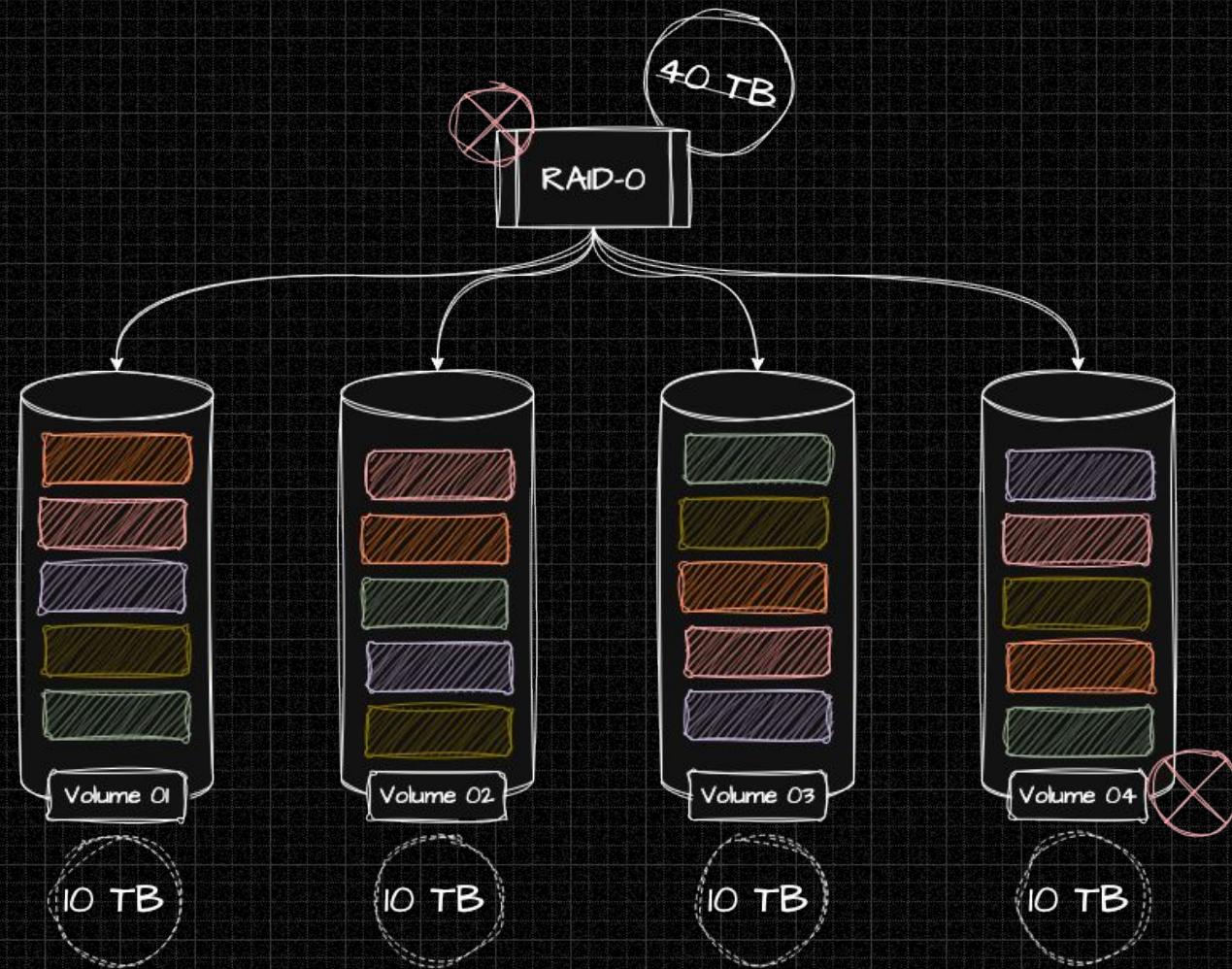
- Redundant Array of Independent Disks
- Combinar múltiplos volumes de discos físicos em um único sistema lógico
- Tradeoff entre resiliência, tolerância a falhas, desempenho e integridade dos dados
- Raid 0, 1, 5, 6 e 10 (1+0)

RAID 0 (Striping)

- Foca em espaço útil e performance
- Peca em disponibilidade e resiliência
- Otimização em termos de escrita e leitura
- Dados são distribuídos igualmente entre dois ou mais discos ou volumes
- Soma dos IOP's dos discos
- Se um único disco do RAID falhar, todos os dados são perdidos
- Inadequado para cenários onde temos dados críticos de longa duração





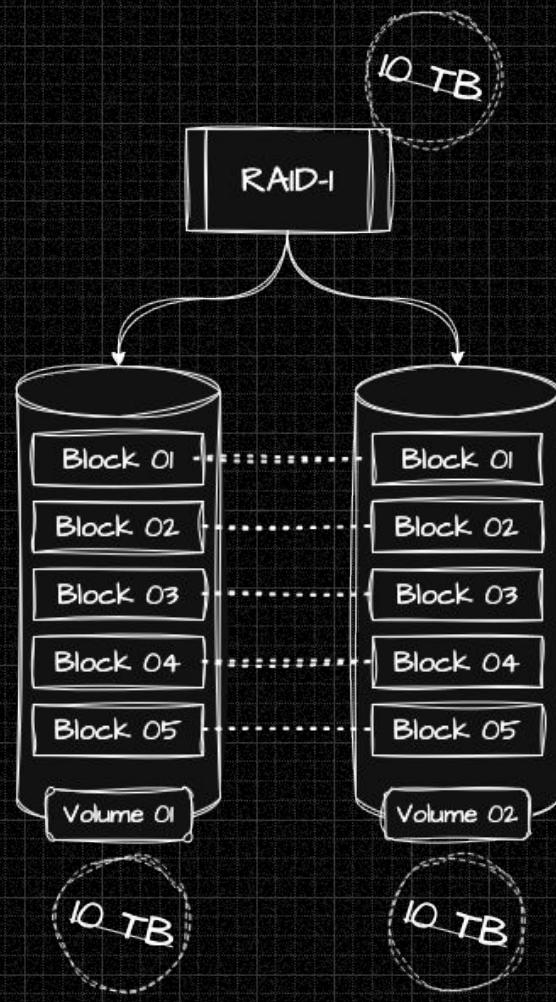


RAID 0 - Casos de Uso

- Clusters de processamento de dados temporários (ex: Hadoop local)
- Ambientes de cache persistente de curta duração
- Servidores de renderização de vídeo/imagem
- Clusters de CI/CD com builds efêmeros
- Clusters Stateless de Alto I/O de transição
- Evitar workflows críticos e transacionais
- Performance é prioritário
- Dados descartáveis ou redundantes

RAID 1 (Mirroring)

- Foca em disponibilidade e redundância
- Aplica o conceito de espelhamento
- Cada disco possui um espelho exato
- Cópia de todos os seus dados em outro disco
- Caso um dos discos apresente problemas, o outro assume imediatamente, sem nenhum tipo de interrupção ou perda
- Contempla metade do storage por conta da replicação
- 50% de espaço útil

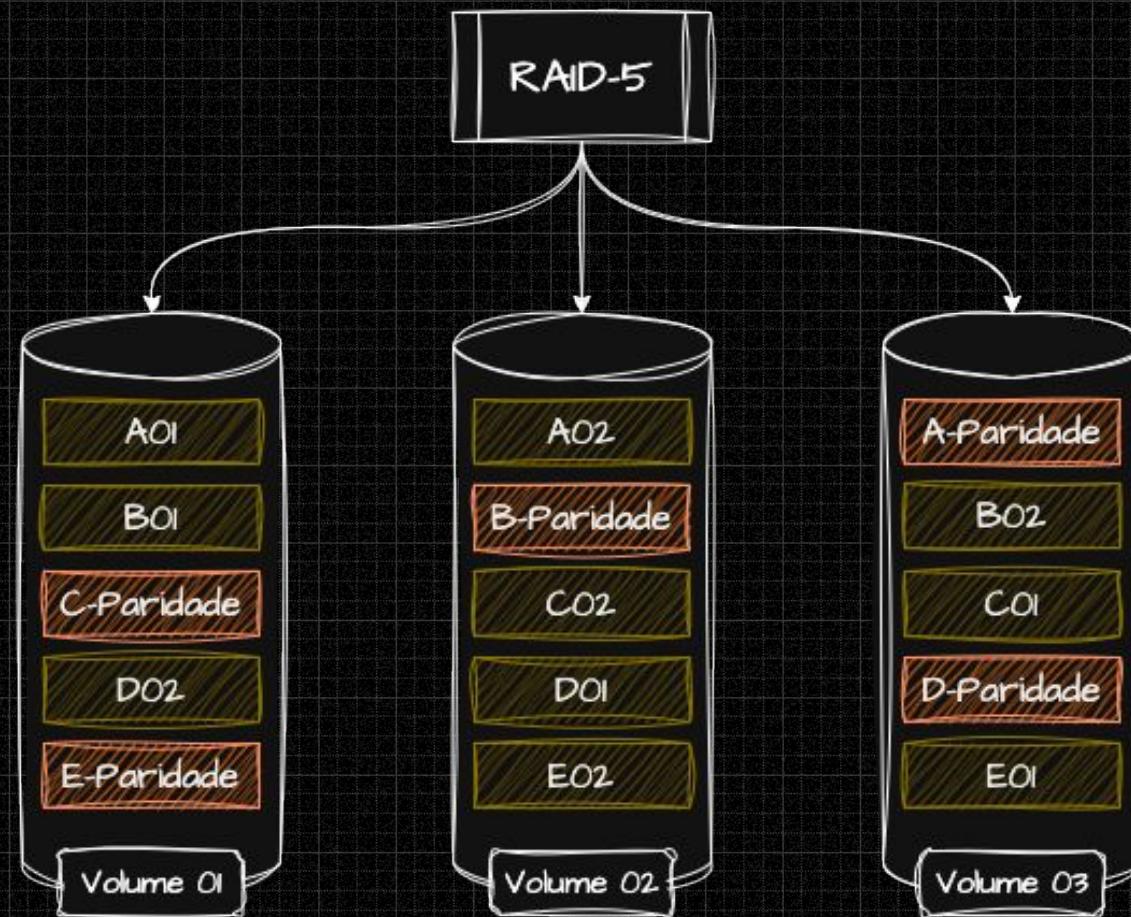


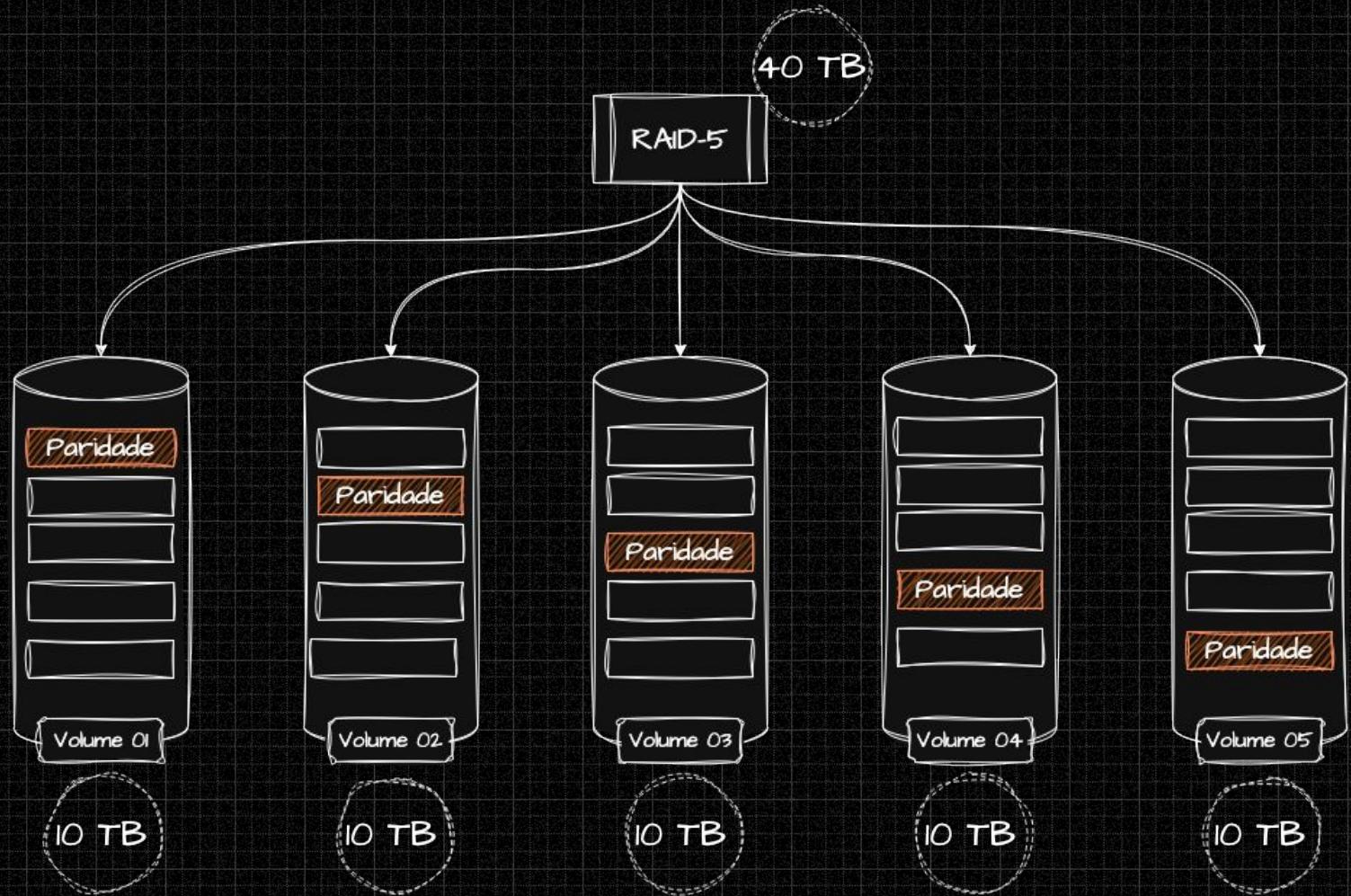
RAID 1 - Casos de Uso

- Alta disponibilidade e simplicidade
- Cluster Management
- Etcd, Gestores de Estado, Quorum Commanders
- Storage com Edge Nodes
- Servidores de Arquivos
- Sistemas Críticos mas que de preferência com dados que podem ser reconstituídos de forma externa se preciso

RAID 5 (Striping com Paridade Distribuída)

- Trade-offs do RAID 0 e RAID 1
- Desempenho de escrita e leitura sem sacrificar disponibilidade e segurança
- Realizando suas escritas de forma distribuída entre os volume
- Mantém metadados de paridade distribuídos entre todos os volumes
- Em caso de falha de um parcela dos mesmos
- Informação rapidamente reconstruída e restaurada
- Mínimo 3 discos



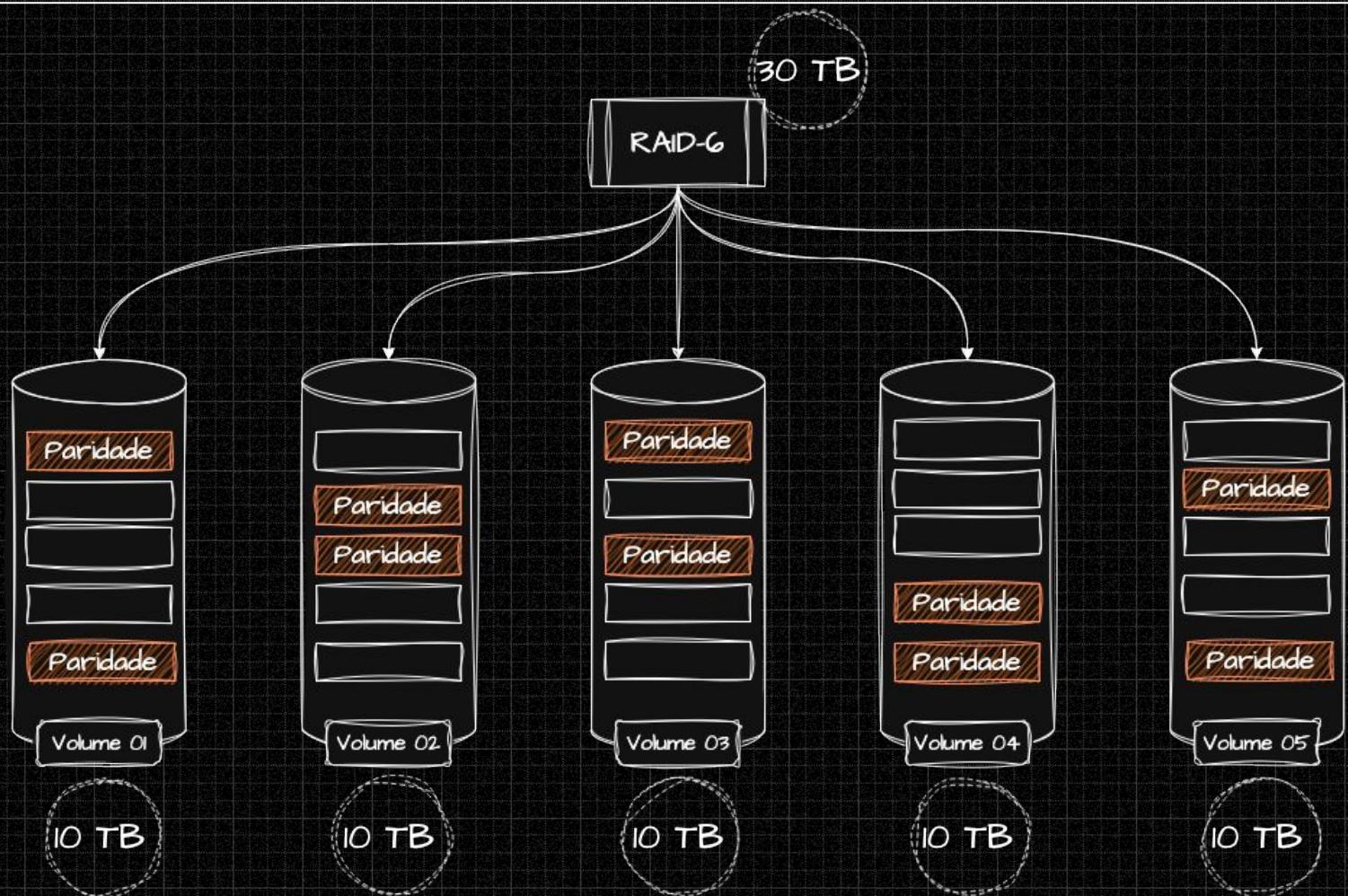


RAID 5 - Casos de Uso

- Tolera 1 falha de disco
- Reconstrução Lenta
- Bom uso de espaço em disco
- Dados frios / pouco mutáveis
- Volume de leitura intenso em servidores frios
- Repositórios de consulta a longo prazo
- Prometheus TSDB
- Armazenamento de Logs
- GlusterFS
- Ceph e etc

RAID 6 (Striping com Dupla Paridade)

- Distribuição de paridade semelhante ao R5
- Camada adicional de paridade distribuída
- Dois discos possam falhar simultaneamente sem perda de dados
- Desempenho são razoavelmente reduzidos
- RAID 6 precisa, ao invés de um, de dois volumes adicionais

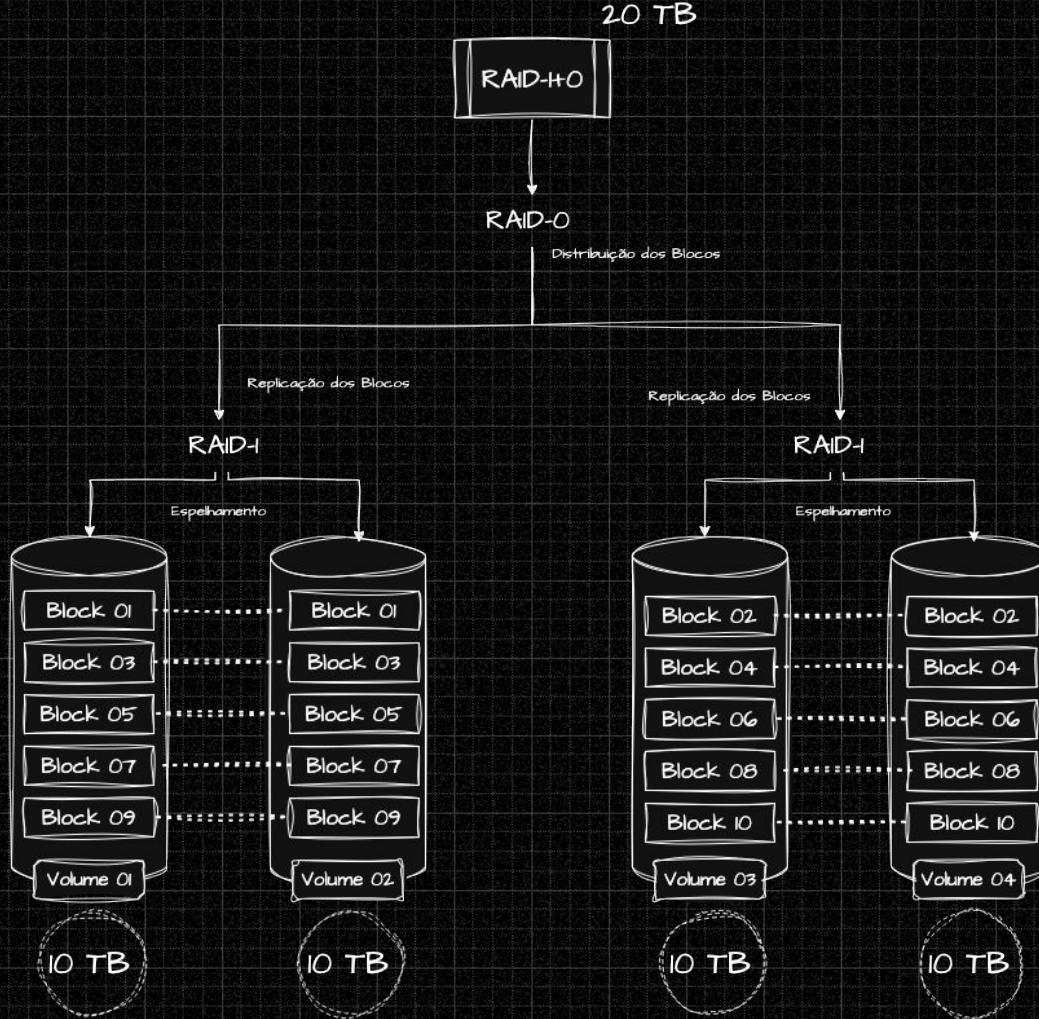


RAID 6 - Casos de Uso

- Tolera 2 falhas simultâneas
- Escrita mais lenta, overhead de paridade dobrado
- Sistemas de Alta Ingestão e Retenção
- Replicação Geográfica
- Data Lakes e Data Warehouse
- Padrões de Datacenter

RAID 10 (Combinação de RAID 1 com RAID 0)

- RAID 10, RAID 1+0
- Combinação de Striping com Mirroring
- Distribuição de blocos entre vários discos
- Podem cada disco possuir seu espelho
- Alta disponibilidade
- Tolera falhas simultâneas de disco
 - Se não forem nos pares
- Capacidade total é reduzida pela metade, pois 50% dos volumes são utilizados para redundância
-



RAID 6 - Casos de Uso

- Databases Críticos
- PostgreSQL, MySQL, Oracle e etc
- MongoDB, Cassandra e etc
- Logs Críticos e Garantias Transacionais

RAID Comparação

RAID	Tolerância a Falhas	Performance	Espaço Útil
0	Nenhuma	Alta	100%
1	1 Disco	Média	50%
5	1 Disco	Boa	~80%
6	2 Discos	Regular	~65%
10	1 Por Par	Alta	50%

\$ whoami

Matheus Fidelis

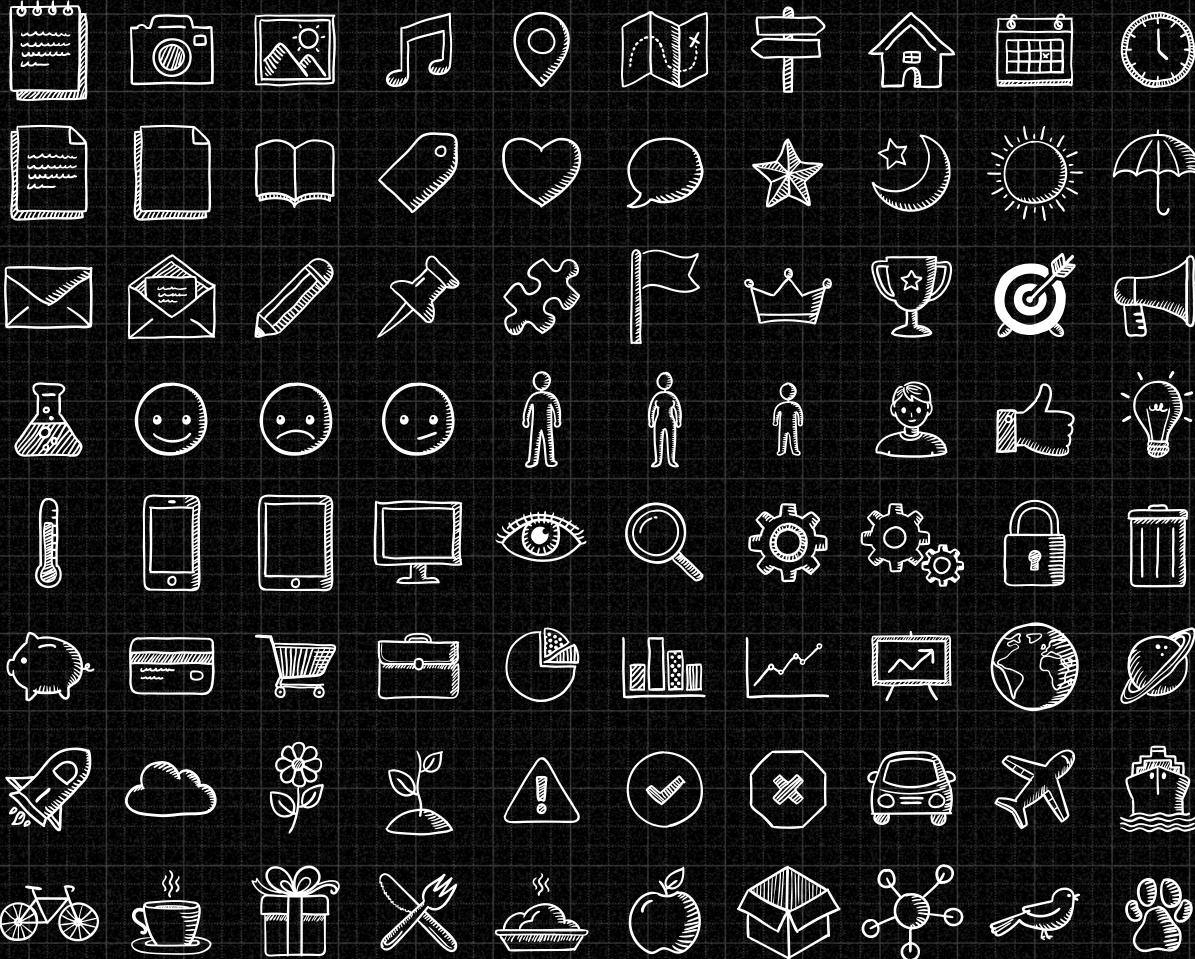
Engenheiro de \$RANDOM

@fidelissauro

<https://fidelissauro.dev>

<https://linktr.ee/fidelissauro>





SlidesCarnival icons are **editable shapes**.

This means that you can:

- Resize them without losing quality.
- Change fill color and opacity.

Isn't that nice? :)

Examples:

