



CURSO

Ajuste (Tuning) de Servidores e SGBDs

Aula 2



departamento
de engenharia
elétrica



Memória RAM - Definição

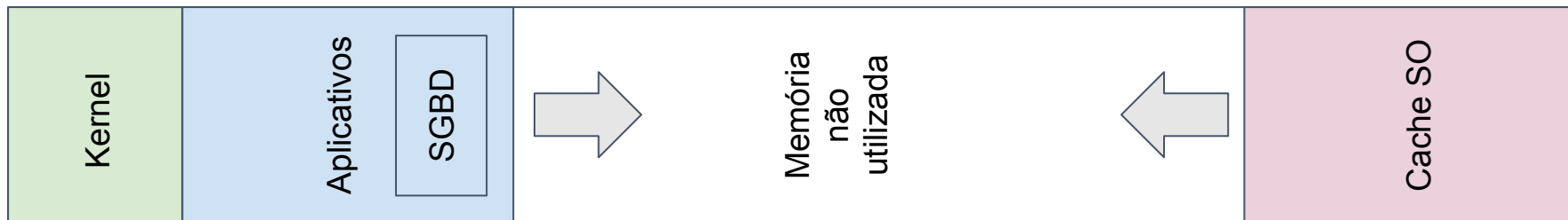
- Dispositivo de armazenamento de dados permitindo leitura e escrita;
- Muito mais rápido que outros dispositivos de armazenamento;
- É volátil: quando a alimentação de energia é interrompida, todo o conteúdo na memória RAM é perdido;
- Custo por Byte é muito maior que outros dispositivos;

Devido à sua velocidade de acesso, é intensamente utilizada como área temporária (*cache ou buffer*) para armazenamento de dados evitando-se assim que o processador tenha que ler ou gravar dados de outros dispositivos bem mais lentos com HDs, SSDs, Rede, etc.

Distribuição da Memória RAM no Sistema Operacional (SO)

- Kernel - Parte da memória necessária ao funcionamento do SO;
- Aplicativos - Memória alocada para cada uma das aplicações em execução;
- Cache do SO:
 - SO aproveita a memória livre utilizando como cache/buffer de bloco de dados lidos/gravados no disco;
 - Se for necessária mais memória para os aplicativos, o SO libera o buffer/cache

Memória RAM



Memória RAM - Cache do Sistema Operacional (SO)

- Funcionamento:
 - Ao ler um bloco de dados do disco, o SO mantém uma cópia do bloco na memória;
 - Leituras subsequentes no mesmo bloco acessam apenas a memória;
 - Em caso de alteração do bloco, o mesmo é alterado na memória e enviado para disco quando solicitado (*fsync*);
 - Caso não haja mais blocos de dados disponíveis no cache, o SO descarta blocos mais antigamente utilizados, liberando espaço para novos blocos

Com isso, evita-se o acesso a disco aumento o desempenho do sistema operacional como todo.

Comparação de velocidades entre tecnologias:

Tecnologia	Latência (microsegundos)	Capacidade oferecida
Armazenam. em Nuvem	60.000	Praticamente Ilimitada
HDs alta capacidade	12.000	4.000 GB
HDs alto desempenho	7.000	600 GB
SSD	300 (write) 45 (Read)	128 GB
Memória RAM	0,005	16 GB

Porque a memória é RAM é fundamental para desempenho do SGBD?

- Devido à grande diferença de velocidade de acesso a disco comparado à memória RAM, é fundamental que o sistema gerenciador de banco de dados tenha estratégias para evitar o acesso ao disco;

Regra geral para otimização do desempenho:

1. Ao responder uma consulta, o banco de dados procura ler a menor quantidade de bloco de dados possível (menor quantidade de leitura lógica);
2. Dos blocos a serem lidos, o banco procura manter os mais acessados na memória, evitando ir ao disco (menor leitura física);

Array de Discos (RAID):

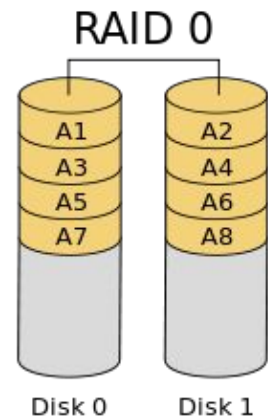
- RAID (Redundant Array of Inexpensive Disks):
- Devido à natureza mecânica, o HD é um dos componentes mais lentos de um computador;
- Embora a capacidade de armazenamento tenha crescido bastante no decorrer dos anos, os HDs ainda apresentam baixa performance;
- O sistema de array de discos RAID busca contornar o problema de desempenho dos HDs ao mesmo tempo em que procura melhorar a segurança dos dados contra falhas de mídia;

Array de Discos (RAID):

RAID 0 (striping):

Divide o bloco em partes iguais a serem distribuídos sobre os discos do sistema.

- Boa performance tanto na escrita como na leitura (distribuição da carga sobre os discos);
- Piora confiabilidade (perdeu um disco, perdeu tudo);
- Preserva o espaço para armazenamento (espaço é igual a soma de todos os disco);

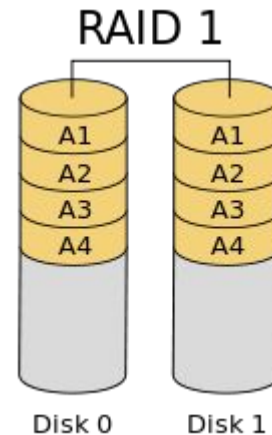


Array de Discos (RAID):

RAID 1 (mirroring):

Bloco de dados é copiado em todos os discos do sistema.

- Performance ruim para gravação (Tem que aguardar todos os discos)
- Boa performance de leitura (Leitura distribuída);
- Melhora confiabilidade (Pode perder todos os discos menos 1);
- Espaço resultante é equivalente ao tamanho do menor HD



Array de Discos (RAID):

RAID 10 (mirror/strip):

Divide parte dos discos para fazer mirroring e parte para striping.

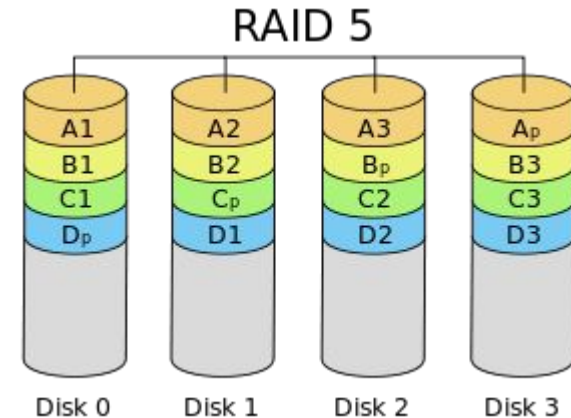
- Boa performance de leitura e gravação
- (Leitura e gravação distribuídas entre os discos);
- Melhora confiabilidade (Pode perder mais de um disco, dependendo da disposição);
- Espaço resultante é equivalente à metade do espaço total (Quando usando 4 discos. O espaço pode mudar dependendo da disposição);

Array de Discos (RAID):

RAID 5:

Divide o bloco em partes iguais a serem gravadas nos discos reservando um disco para a paridade. A cada bloco gravado, a paridade é gravada em disco diferente.

- Boa performance de leitura e gravação (Leitura e gravação distribuída);
- Boa confiabilidade (Pode perder 1 disco);
- Espaço resultante é equivalente ao espaço total menos o espaço de um HD;

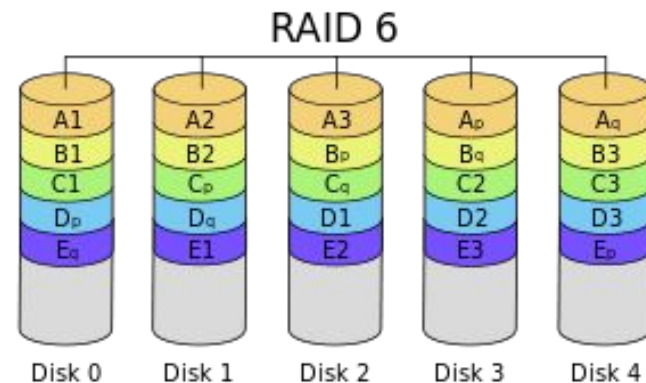


Array de Discos (RAID):

RAID 6:

Semelhante ao RAID 5, com a diferença que usa dois discos para paridade em vez de um.

- Boa performance de leitura e gravação;
- Melhora ainda mais a confiabilidade;
- Espaço resultante é equivalente ao espaço total menos o espaço de dois HDs;



Array de Discos (RAID):

Observações:

- Procure distribuir os discos nas controladores disponíveis;
- Não protege contra erros humanos;
- Para sistemas críticos é recomendada a utilização de *hotspares*;
- O RAID pode ser implementado via software ou via hardware. Este último sendo mais caro, porém mais performático.



Storage:

DAS (Direct Attached Storage):

- Se conecta ao computador por uma porta USB ou eSATA;
- Pode ser um dispositivo de baixo custo ou dispositivos mais complexos com RAID implementados via hardware.



Storage:

NAS (Network Attached Storage):

- Funciona como um servidor de arquivos;
- Possui sistema operacional próprio;
- Pode ser um dispositivo de baixo custo ou dispositivos mais complexos com RAID implementados via hardware;
- Geralmente são administrados via interface web;
- Disponibilizam os dados através de protocolos de rede como CIFS, NFS, FTP e SFTP;
- Controle de acesso por autenticação



Storage:

SAN (Storage Area Network):

- Possui unidade controladora complexa que trata do acesso aos dados, RAID e outras funções;
- Conecta-se ao servidor por interface dedicada;
- Possui alta capacidade de armazenamento e de recursos de redundância;
- Controle de acesso por interface



Storage:

SAN (Storage Area Network):

- Componentes:
 - Storage Rack
 - Switch (de acordo a interface para comunicação com os servidores);
 - Cabos correspondentes à interface;
 - Servidor.



Storage:

SAN (Storage Area Network):

- É comum a utilização de dois switches de forma redundante;
- Pode ser expandido com acréscimo de Storage Racks e Switchs;
- Alto custo justificado por maior confiabilidade e maior capacidade de armazenamento



Storage:

SAN (Storage Area Network):

- Pode ser particionado em frações de espaço menores chamadas LUN (Logica Unit Number) para atender diversos serviços/servidores;
- Como vantagem, centraliza-se o armazenamento de dados em uma SAN, reduzindo gerenciamento com backup, por exemplo;



Storage:

SAN (Storage Area Network):

- Normalmente possui uma boa porção de memória sustentada por baterias servido como cache para agilizar gravações e leituras;
- Alguns fornecedores implementam funcionalidades como clone de LUNs, espelhamento para backup, etc.;

PostgreSQL - Principais áreas de memórias alocadas em SGBD

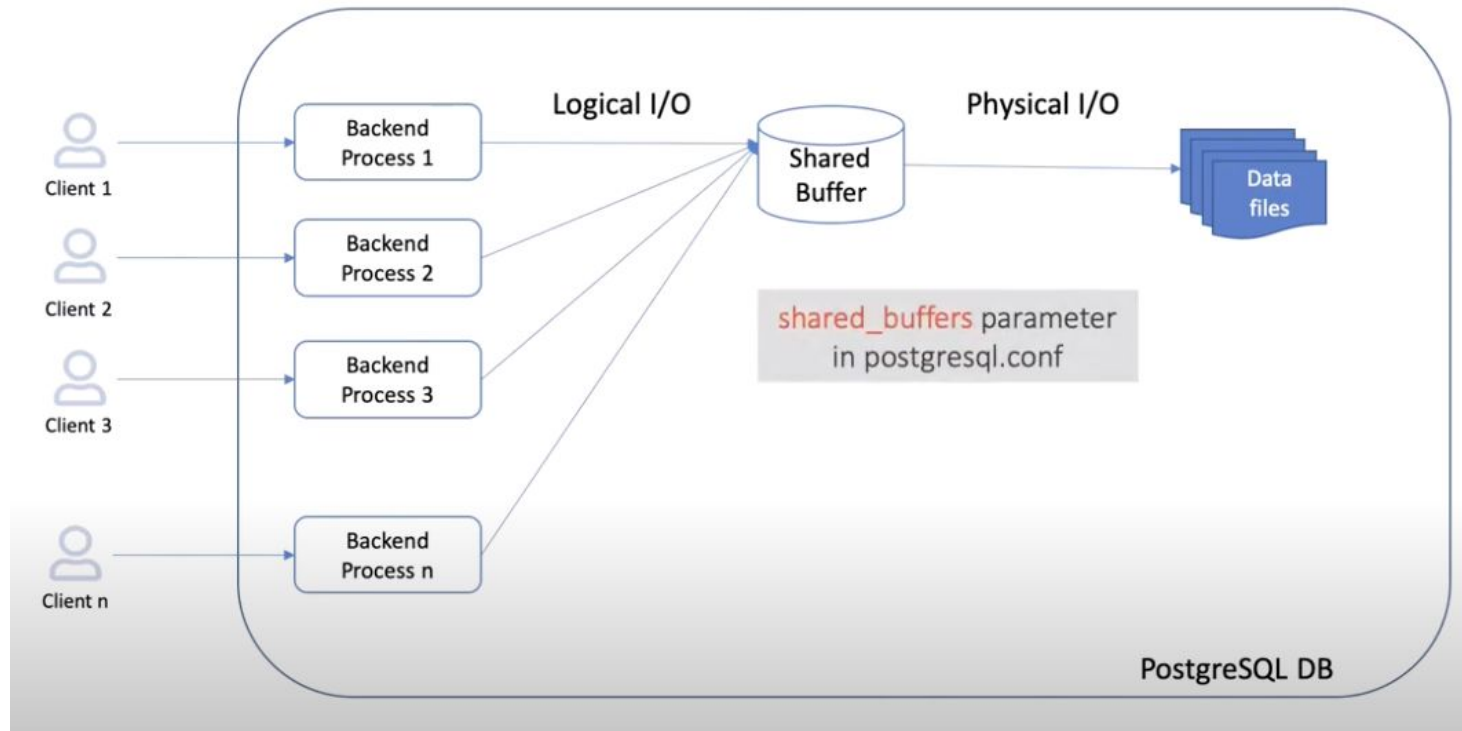
- Da área de memória alocada para o SGBD, as mais importantes são:
 - Buffer de Dados (*shared_buffer*)
 - Conexões (*max_connections*)
 - Ordenação/Classificação (*work_mem*)
 - Manutenção (*maintenance_work_mem*)
 - Planos de acesso

PostgreSQL - Principais áreas de Memórias em SGBD - Buffer de Dados

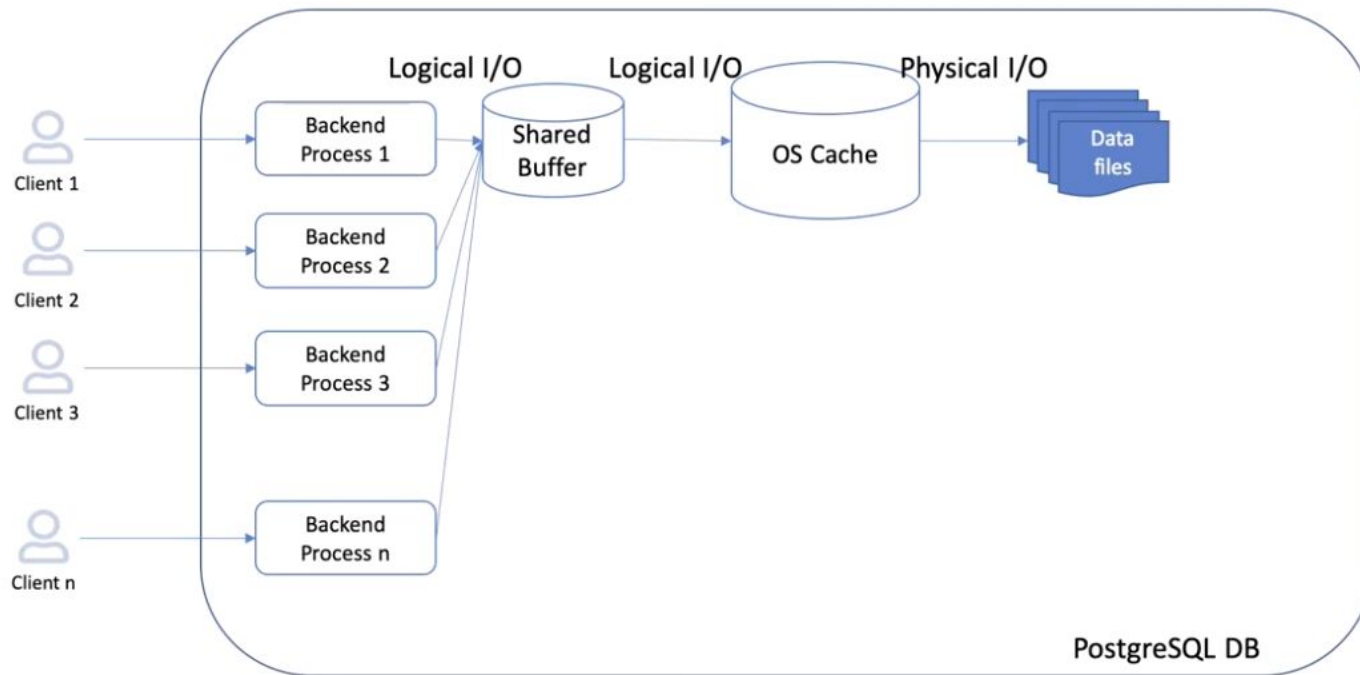
- Buffer de Dados (*shared_buffer*)
 - Área onde os blocos de dados mais acessados são armazenados evitando-se o acesso a disco (semelhante ao cache do SO);
 - A principal memória a ser configurada no SGBD, pois impacta diretamente o tempo de resposta das consultas mais utilizadas;
- **Leituras lógicas: Ocorre cada vez que o SGBD precisa ler um bloco de dados para responder a um comando;**
- **Leituras físicas: Ocorre quando o bloco lido não está no *buffer* de dados e obriga o SGBD a ler o bloco do disco através do sistema operacional;**
- ***Hit Ratio*: Percentual de vezes que os blocos lidos pelo SGBD é encontrado no buffer de dados.**

Para um sistema transacional, Hit Ratio deve ser maior que 95%.

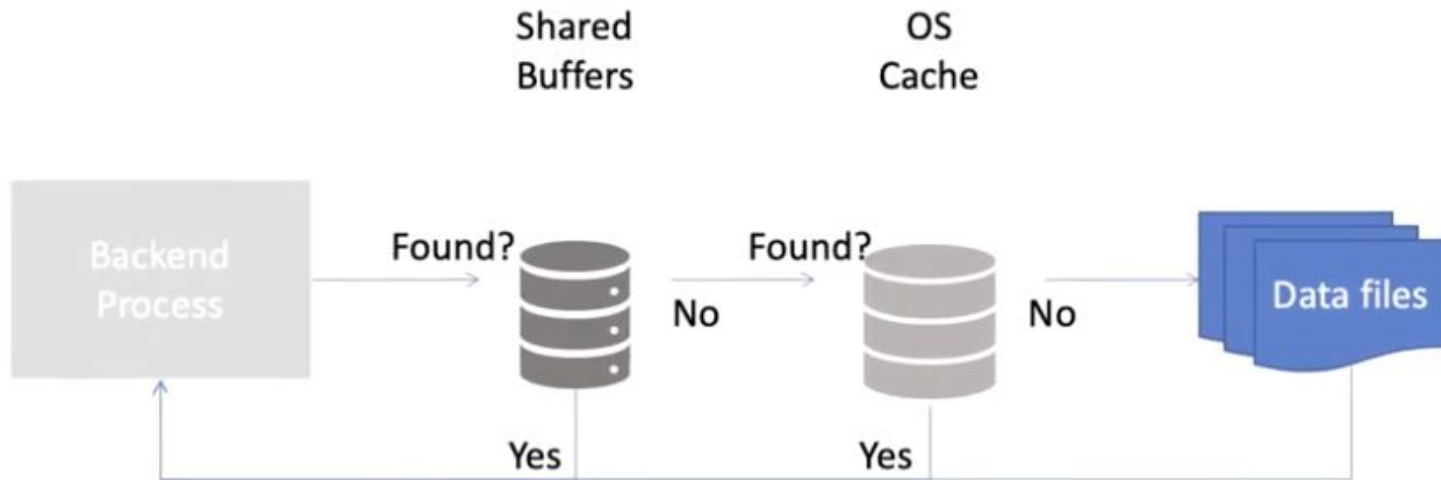
PostgreSQL - Principais áreas de Memórias em SGBD - Buffer de Dados



PostgreSQL - Principais áreas de Memórias em SGBD - Buffer de Dados



PostgreSQL - Principais áreas de Memórias em SGBD - Buffer de Dados



- Conexões (*max_connections*):
 - Cada conexão física do cliente no SGBD reserva uma quantidade de memória para o processo dedicado àquela conexão;
 - Quanto mais conexões no banco de dados, mais memória é alocada para tratamento de conexões;
 - Como forma de prevenção, a quantidade máxima de conexões deve ser estipulada na configuração do SGBD;
 - Pool de conexões na aplicação pode ajudar a reduzir conexões no SGBD;
 - Memória utilizada: $(1800 + 270 * \text{max_locks_per_transaction}) * \text{max_connections}$
 - 100 conexões ~ 2 Mbytes

Ordenação/Classificação

- Ordenação/Classificação (*work_mem*)
 - Memória utilizada para ordenação de registros antes da entrega do resultado ao usuário;
 - Cláusulas que requerem ordenação de registros: ORDER BY, GROUP BY, DISTINCT, UNION;
 - Essa memória é individual para cada conexão;
 - Quando esta memória não é suficiente, o SGBD utiliza espaço em disco para prover a ordenação.

- Manutenção (*maintenance_work_mem*)
 - Memória utilizada para operações de:
 - Vacuum
 - Criação de índices
 - Criação de *Foreign Keys*
 - Como a criação de índices e FKs são atividades mais raras no banco de dados, eventuais ações como esta poderão ser suportadas através do aumento temporário da memória de manutenção;
 - Com relação ao processo de Vacuum, este deve ter seu desempenho acompanhado e diferentes valores para a configuração desta memória devem ser testados a fim de atingir o valor mais efetivo.

PostgreSQL - Principais áreas de Memórias em SGBD - Planos de Execução

- Planos de execução
 - Memória dedicada ao armazenamento de planos de execução prontos evitando o reprocessamento/análise de consultas repetitivas;
 - Nem todos os SGBDs possuem configuração para memória dedicada ao armazenamento de planos de execução. É o caso do PostgreSQL;
 - Depende de substituição de expressões para que o banco de dados identifique variáveis e planos de execução realmente semelhantes.

