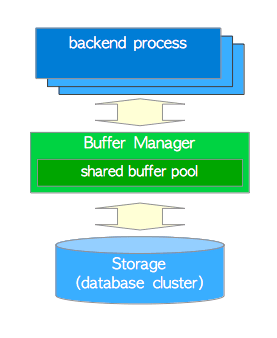
**Postgresql缓存管理**

## 一．缓存和缓冲区的概念

## 缓冲区是PostgreSQL用来读取和写入数据的内存区域。当从磁盘读取数据时，PostgreSQL会将读取的数据存储在缓冲区中，以提高后续访问的速度。这样，当需要再次访问相同的数据时，PostgreSQL可以直接从缓冲区中获取数据，而不必再次访问磁盘。

缓存是PostgreSQL用来存储查询结果和执行计划的内存区域。当执行一个查询时，PostgreSQL会将查询结果和执行计划存储在缓存中，以便下次执行相同的查询时可以直接使用缓存中的数据，而不必重新计算。

**二．缓冲区**



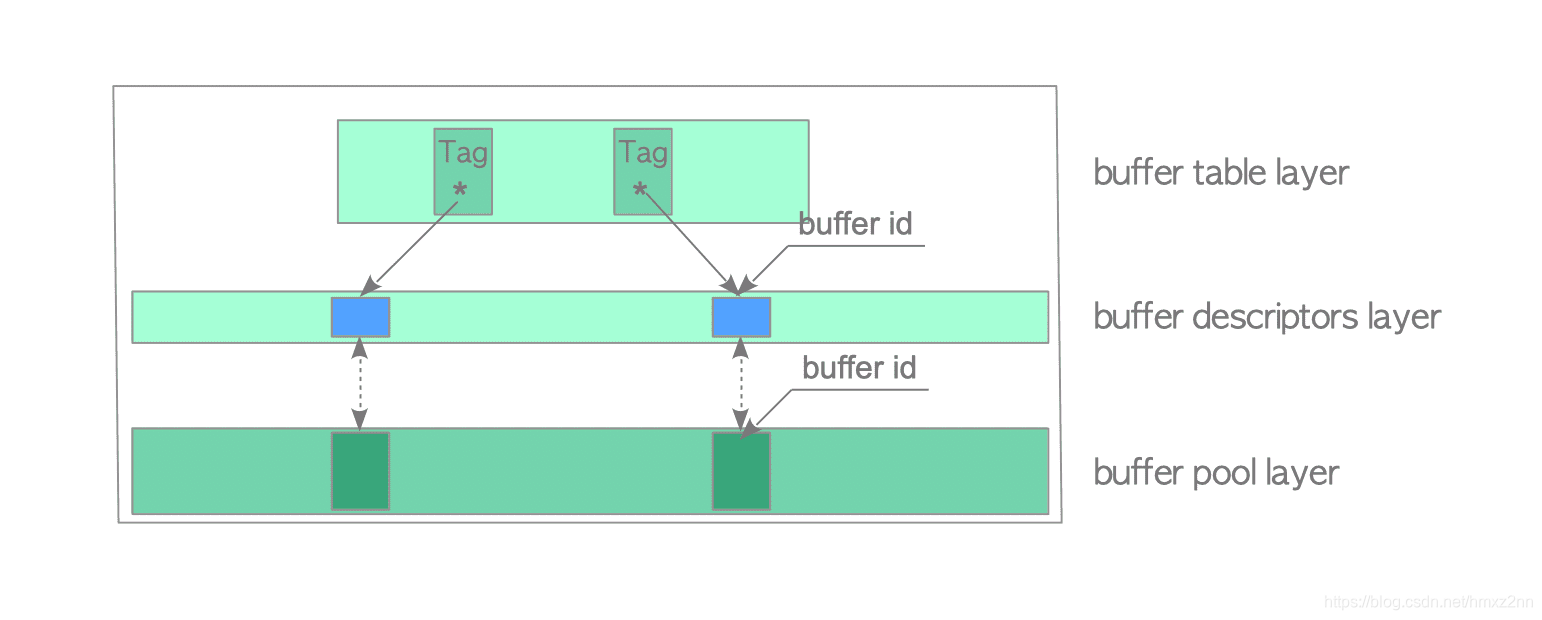
我们知道所有数据库在访问数据块时，都会先把数据块读到内存。因为磁盘IO比起内存中的访问速度慢得多。大多数都是随机io。因此会在在存储层上面添加了一层缓存。缓存可以存储频繁读取的数据和索引，但后端进程在需要数据时，可以直接从缓存里面取。从而减少从磁盘读取操作的次数。减少磁盘IO操作，提高响应速度

内存缓冲的内容分为很多种，本地缓存与共享缓存是两个大区分。先主要说一下共享缓冲池和本地缓冲池

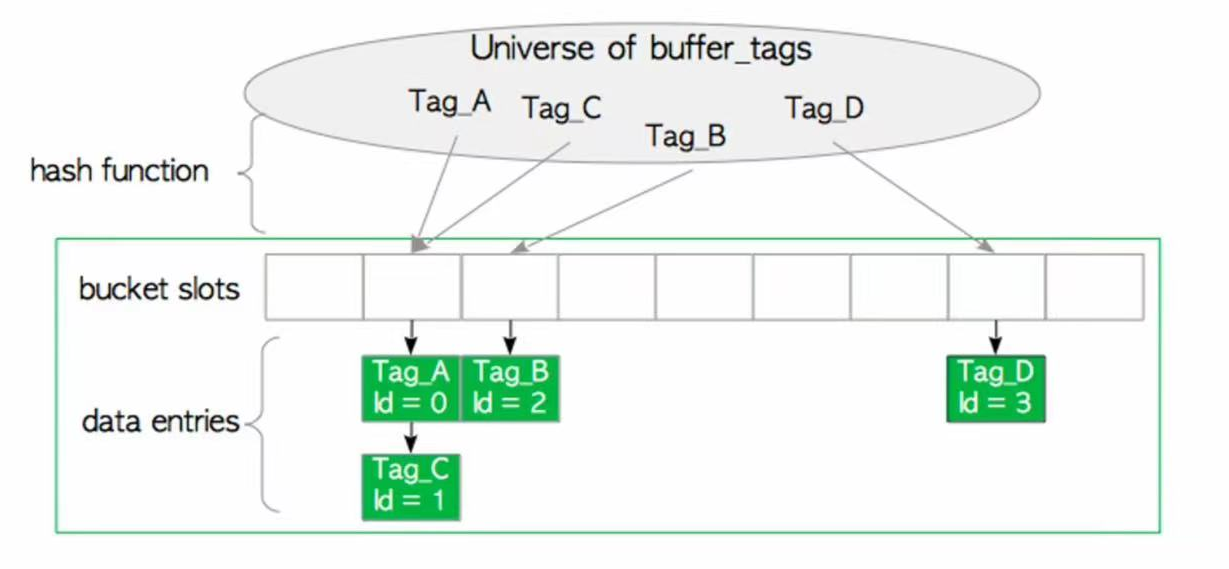
共享缓存会被所有进程访问，一般普通的表和索引都会使用它。而本地缓存只有自身进程可以访问到，仅仅只有临时表使用。

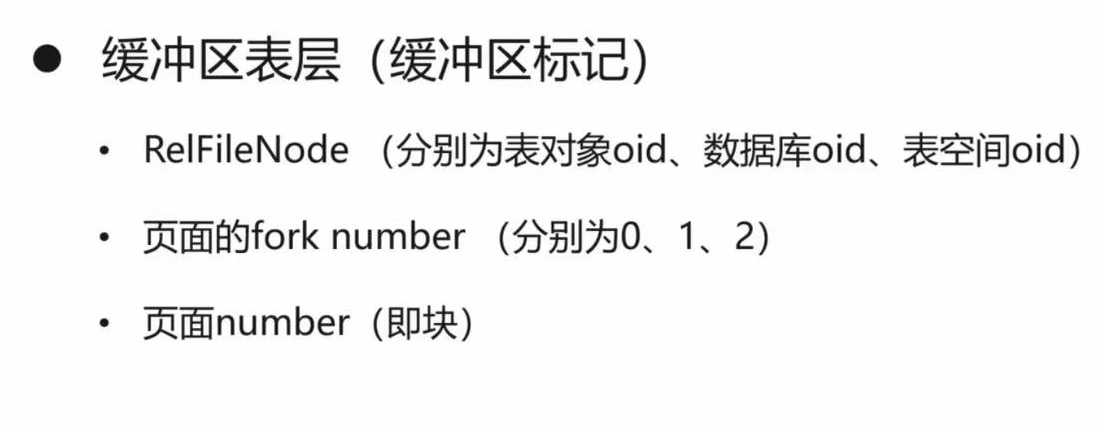
无论是共享缓存还是本地缓存，postgresql 都是使用一个固定大小的 buffer 数组来存储。当 buffer 不够用时会使用算法剔除一些不常用的元素。

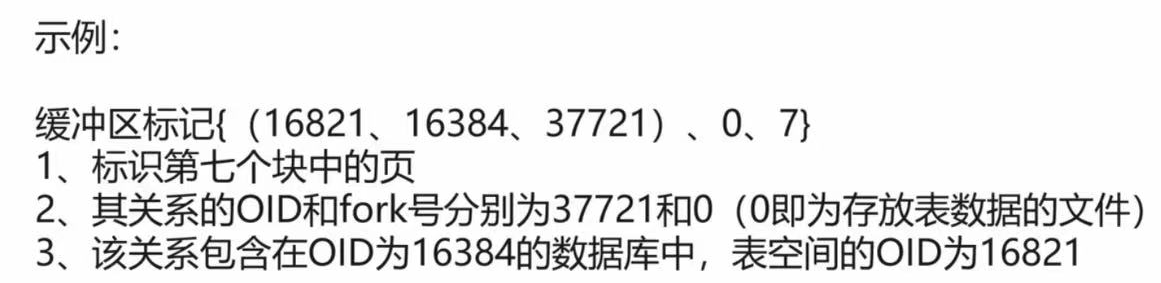
二．Buff 管理结构



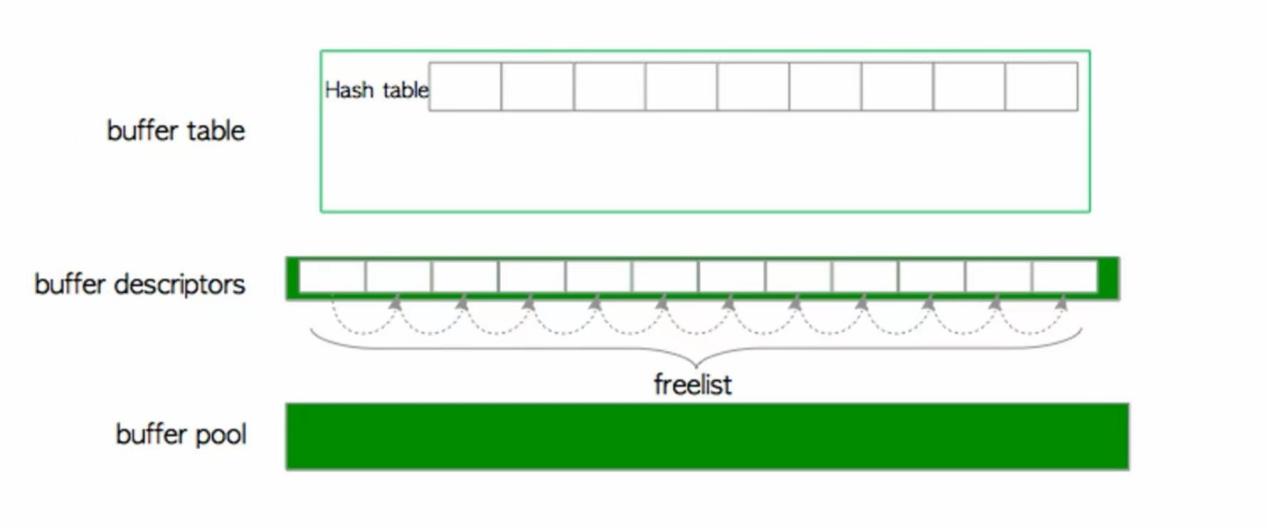
缓冲区表层

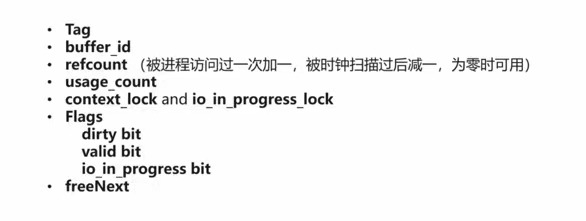




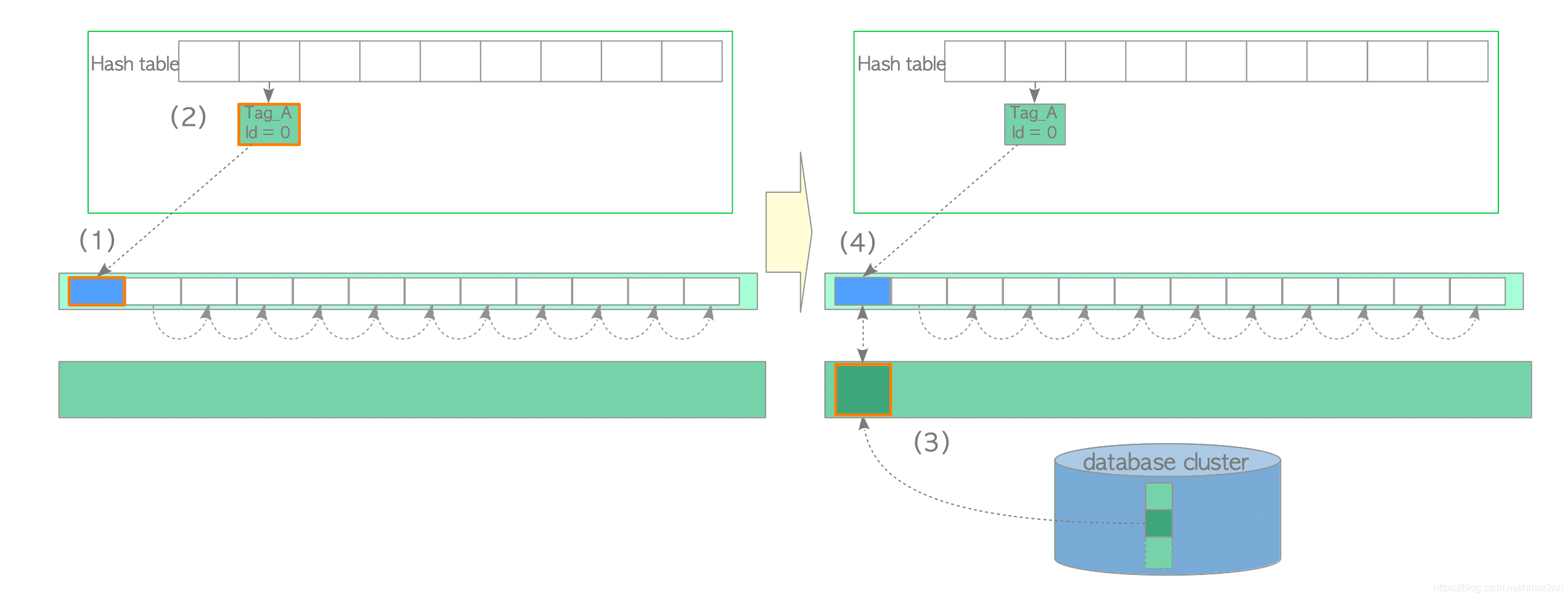


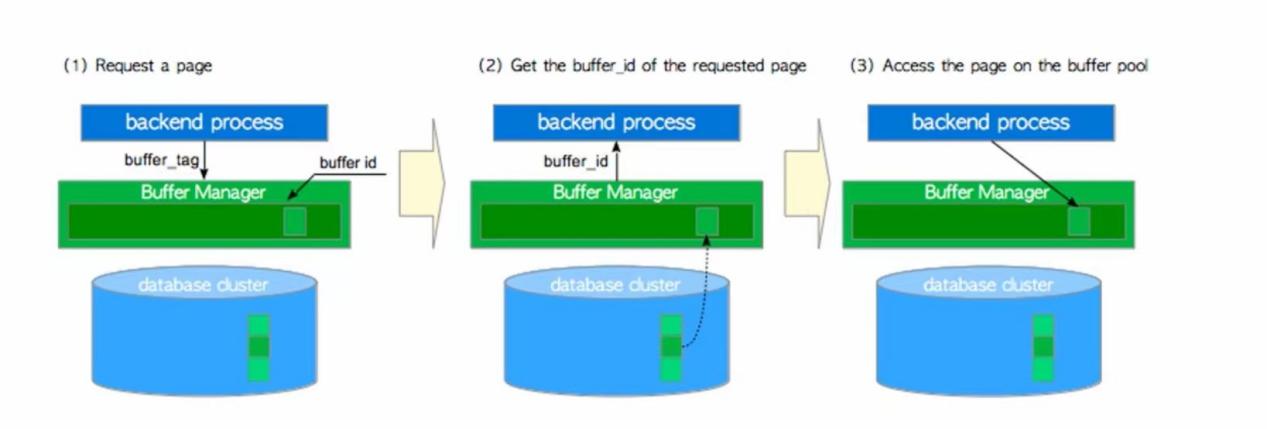
缓冲区描述层

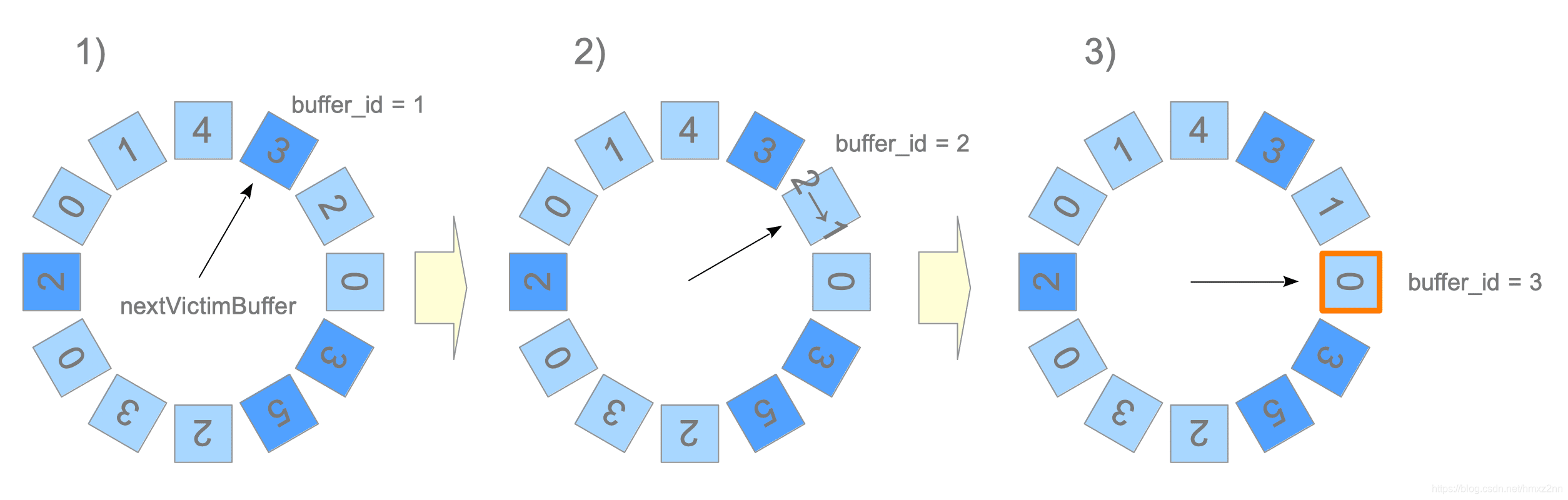




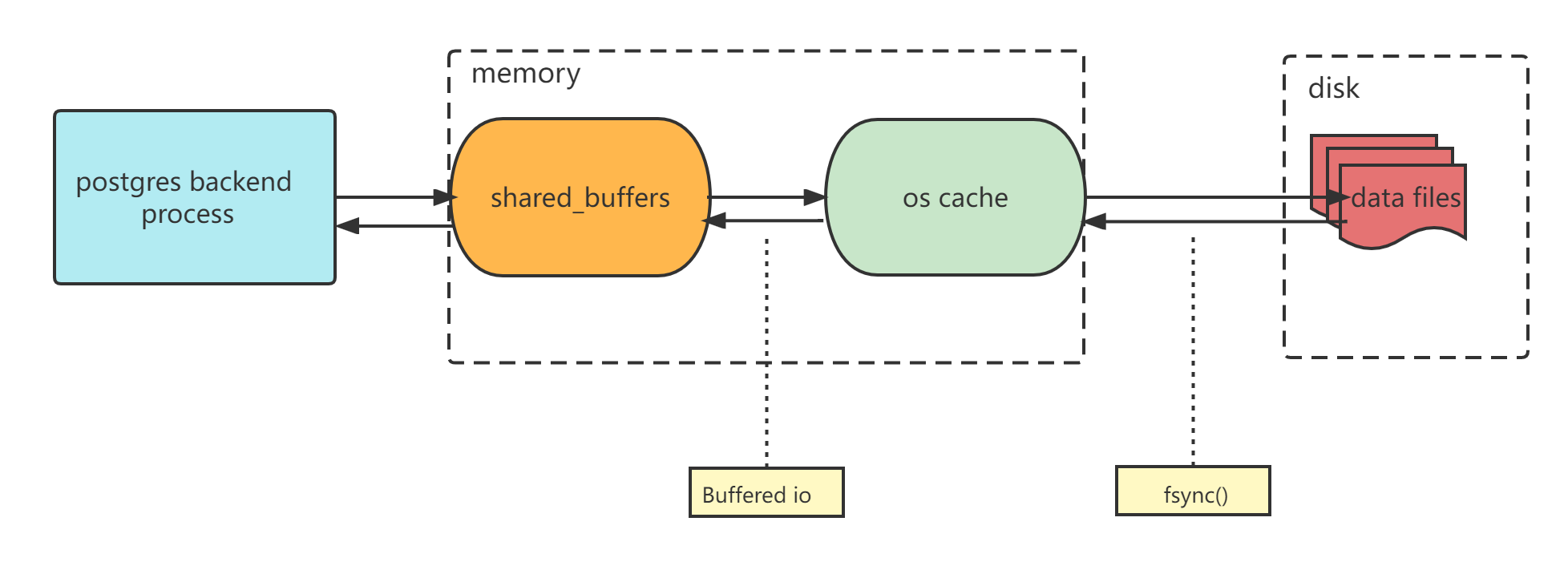
加载数据块到缓冲区







**三.Postgresql的双缓存**



1. **介绍**
2. **与mysql与oracle的不同 dio**
3. **优点**
4. **Dig 插件**
5. **内存浪费**
6. **读写**
7. **大小**
8. **实验**

**四．对于Buffer的认识，与实验相关源代码解读**

1.三个重要结构体

BufferTag 用来表示缓存的内容是哪个文件的block

BufferDesc 包含了buffer的头部信息

BufferLookupEnt 缓冲表哈希key

**typedef** **struct**{

BufferTag key; */\* Tag of a disk page \*/*

**int** id; */\* Associated buffer ID \*/*

} BufferLookupEnt; *//缓冲表哈希entry*

1. **函数**

#### ReadBuffer\_common 函数

BufferAlloc函数

**五．缓冲区命中率**

当数据库需要访问某个数据页时，它首先检查该页是否已在共享数据缓冲区中；如果是，直接从内存中读取数据，否则从磁盘加载数据页到缓冲区再进行访问。所以命中率高，自然可以提高性能。

共享数据缓冲区的大小是可配置的，通过调整 shared\_buffers 参数来实现，这个参数定义了缓冲区分配的内存量。理想的缓冲区大小取决于系统的总内存、数据库的工作负载以及其他内存需求。正确配置共享数据缓冲区对于优化数据库性能至关重要，因为它直接影响到数据检索的速度。

如果给sharebuff太大内存，相应操作系统缓存（OS Cache）分配的值就少了

一旦页面被标记为dirty，它就会被刷新到OS缓存中，然后写入磁盘。在这里，OS可以更自由地根据传入的流量进行I/O。如果操作系统缓存的大小更小，那么它就不能重新排序写操作和优化I/O。这对于编写繁重的工作负载尤其重要。所以操作系统缓存大小也很重要。

PostgreSQL建议你将25%的系统内存分配给共享缓冲区，并且可以根据环境随时修改共享缓冲区的值。

**六．实验相关**

数据库是冷数据，开始的每一次查询都要到磁盘去读取，然后写入数据库缓存。如果当数据库启动时，就有大量的数据库访问，可能造成数据库堵死。

那个实验三第一个优化应该就是把磁盘内容刷到内存，而内存又不占用pg的buffer。本实验中每次都要把全部缓存初始化，所以性能都不高

实验中要控制缓存，要通过重启postgres和echo 3 > /proc/sys/vm/drop\_caches。分别初始化sharebuff和os缓存

PostgreSQL 在pg\_statio\_user\_tables表中，为数据库中的所有表提供了[缓存命中率统计信息](https://www.rockdata.net/zh-cn/docs/14/monitoring-stats.html)。该表中有两个很有用的列：heap\_blks\_read定义为“从此表读取的磁盘块数”，heap\_blks\_hit定义为“此表中的缓冲区命中数”。

所以每次改变share\_buff大小都要清空一次这个表

SELECT pg\_stat\_reset();

数据库服务重启时，postgresql.auto.conf 配置优先

执行 ALTER SYSTEM … 命令时，改变会被自动写入 postgresql.auto.conf 文件，而不是 postgresql.conf文件。方便看到参数变化

