



PostgreSQL

Разработка серверной части приложений PostgreSQL 16 **(dev-1)**



Архитектура Буферный кеш и журнал

Posgres PROFESSIONAL



Устройство буферного кеша

Алгоритм вытеснения

Журнал предзаписи

Контрольная точка

Процессы, связанные с буферным кешем и журналом

Буферный кеш

Массив буферов

страница данных (8 Кбайт)

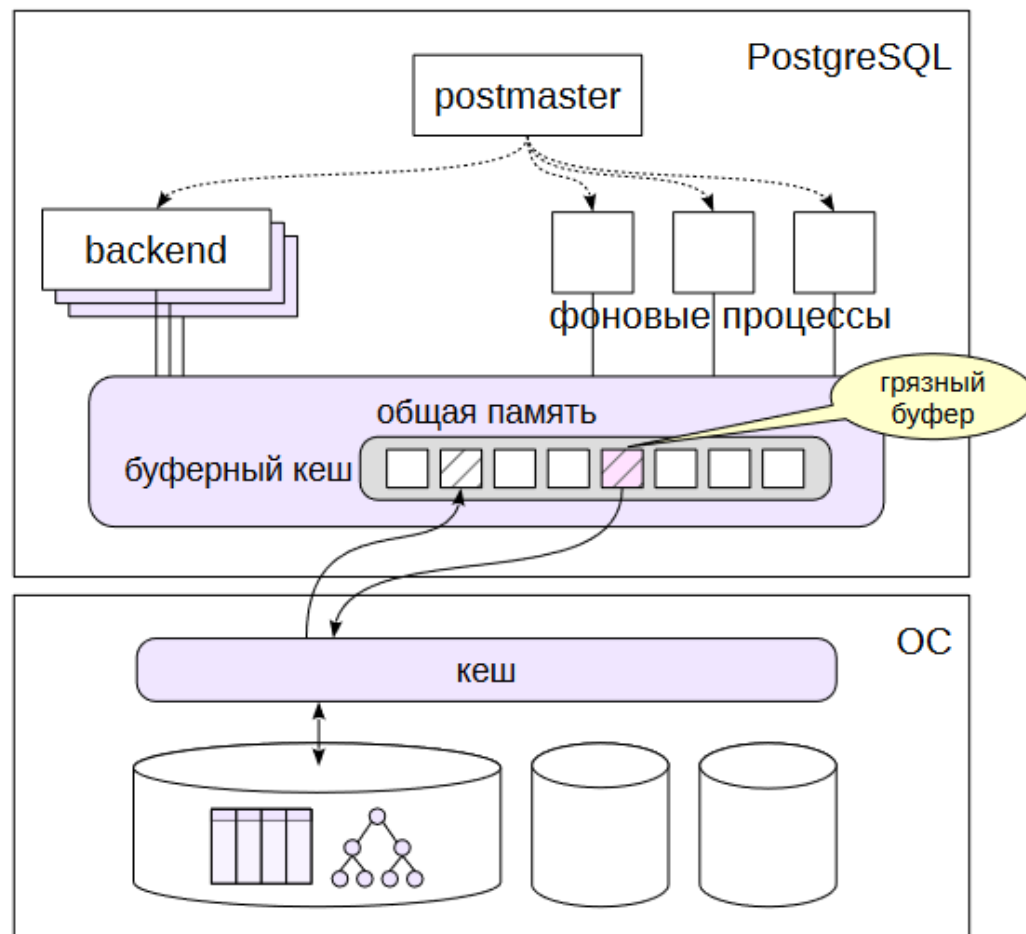
доп. информация

«Грязные» буферы

асинхронная запись

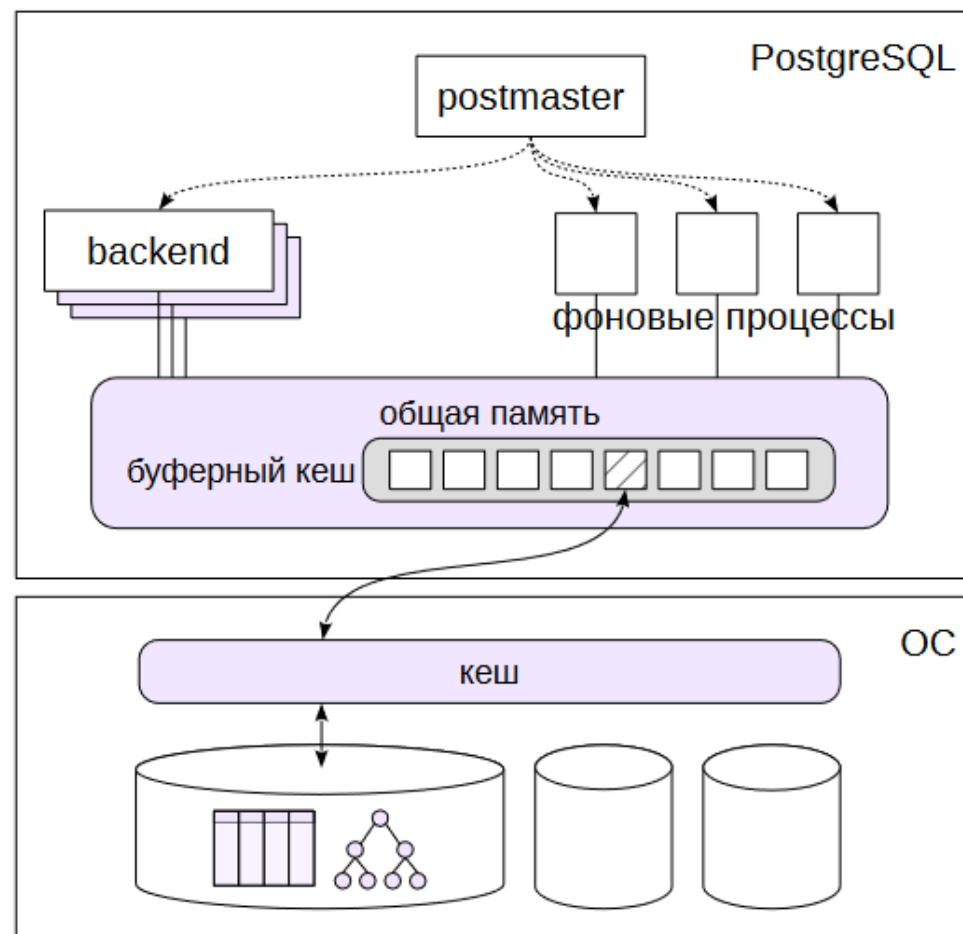
Блокировки в памяти

для совместного доступа



Вытеснение редко используемых страниц

грязный буфер
записывается на диск
на освободившееся место
читается другая страница



Проблема: при сбое теряются данные из оперативной памяти, не записанные на диск

Журнал

поток информации о выполняемых действиях,
позволяющий повторно выполнить потерянные при сбое операции
запись попадает на диск раньше, чем измененные данные

Журнал защищает

страницы таблиц, индексов и других объектов
статус транзакций (clog)

Журнал не защищает

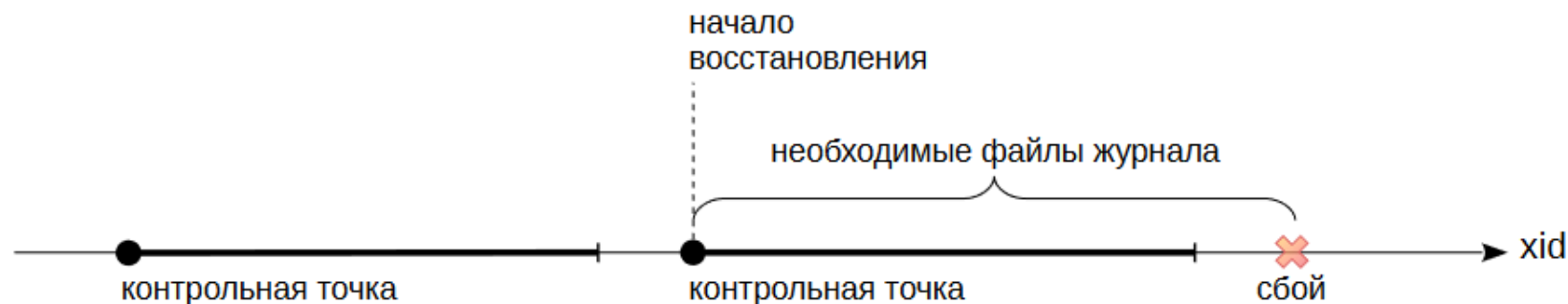
временные и нежурналируемые таблицы

Периодический сброс всех грязных буферов на диск

гарантирует попадание на диск всех изменений до контрольной точки
ограничивает размер журнала, необходимого для восстановления

Восстановление при сбое

начинается с последней контрольной точки
последовательно проигрываются записи, если изменений нет на диске

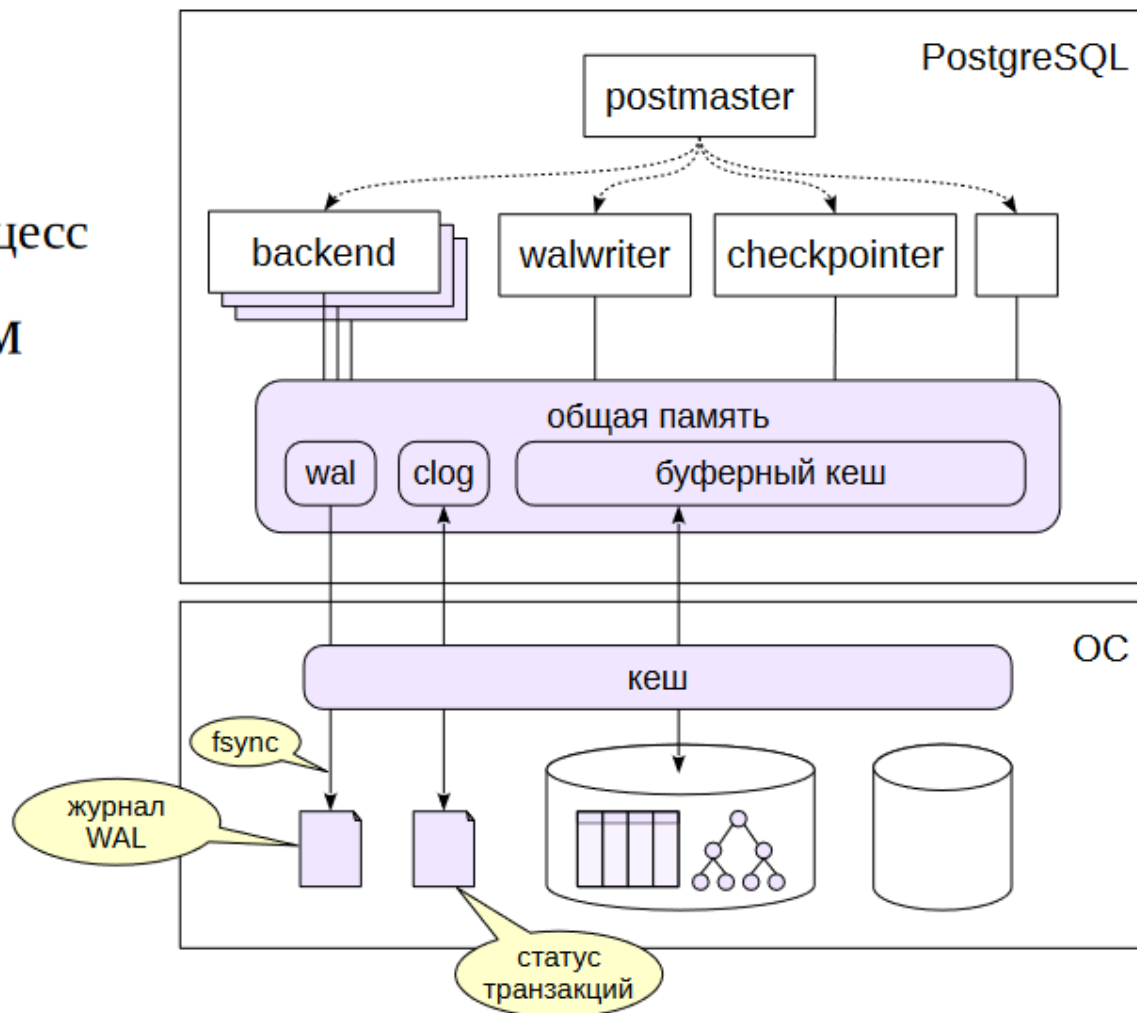


Синхронный режим

запись при фиксации
обслуживающий процесс

Асинхронный режим

фоновая запись
walwriter



Основные процессы

Запись журнала

Контрольная точка

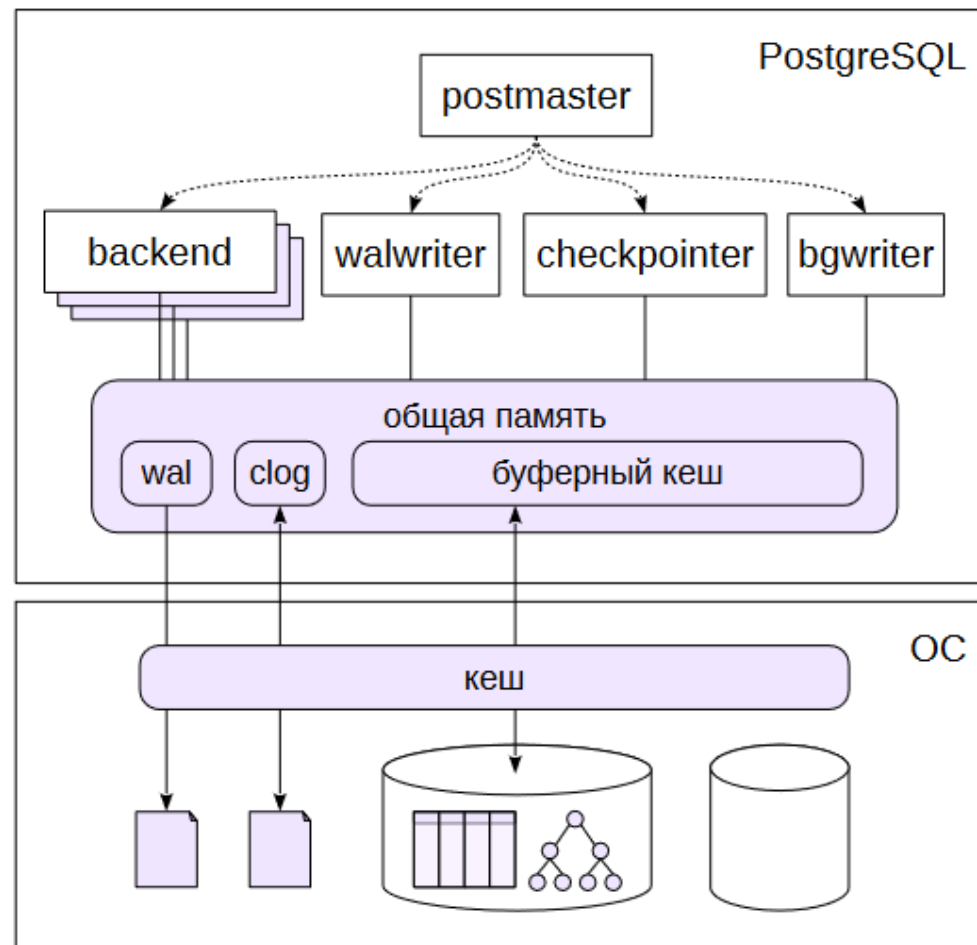
сброс всех
грязных буферов

Фоновая запись

сброс части
грязных буферов

Обслуживающие процессы

сброс вытесняемого
грязного буфера



Minimal

гарантия восстановления после сбоя

Replica *(по умолчанию)*

резервное копирование

репликация: передача и проигрывание журнала на другом сервере

Logical

логическая репликация: информация о добавлении, изменении и удалении табличных строк

Буферный кеш существенно ускоряет работу,
уменьшая число дисковых операций

Надежность обеспечивается журналированием

Размер журнала ограничен благодаря контрольным точкам

Журнал удобен и используется во многих случаях

- для восстановления после сбоя

- при резервном копировании

- для репликации между серверами

1. Средствами операционной системы найдите процессы, отвечающие за работу буферного кеша и журнала WAL.
2. Остановите PostgreSQL в режиме fast; снова запустите его. Просмотрите журнал сообщений сервера.
3. Теперь остановите в режиме immediate и снова запустите. Просмотрите журнал сообщений сервера и сравните с предыдущим разом.

2. Для останова в режиме fast используйте команду

```
pg_ctlcluster 16 main stop
```

При этом сервер обрывает все открытые соединения и перед выключением выполняет контрольную точку, чтобы на диск записались согласованные данные. Таким образом, выключение может выполняться относительно долго, но при запуске сервер сразу же будет готов к работе.

3. Для останова в режиме immediate используйте команду

```
pg_ctlcluster 16 main stop -m immediate --skip-systemctl-redirect
```

При этом сервер также обрывает открытые соединения, но не выполняет контрольную точку. На диске остаются несогласованные данные, как после сбоя. Таким образом, выключение происходит быстро, но при запуске сервер должен будет восстановить согласованность данных с помощью журнала.

Для PostgreSQL, собранного из исходных кодов, останов в режиме fast выполняется командой

```
pg_ctl stop
```

а останов в режиме immediate — командой

```
pg_ctl stop -m immediate
```