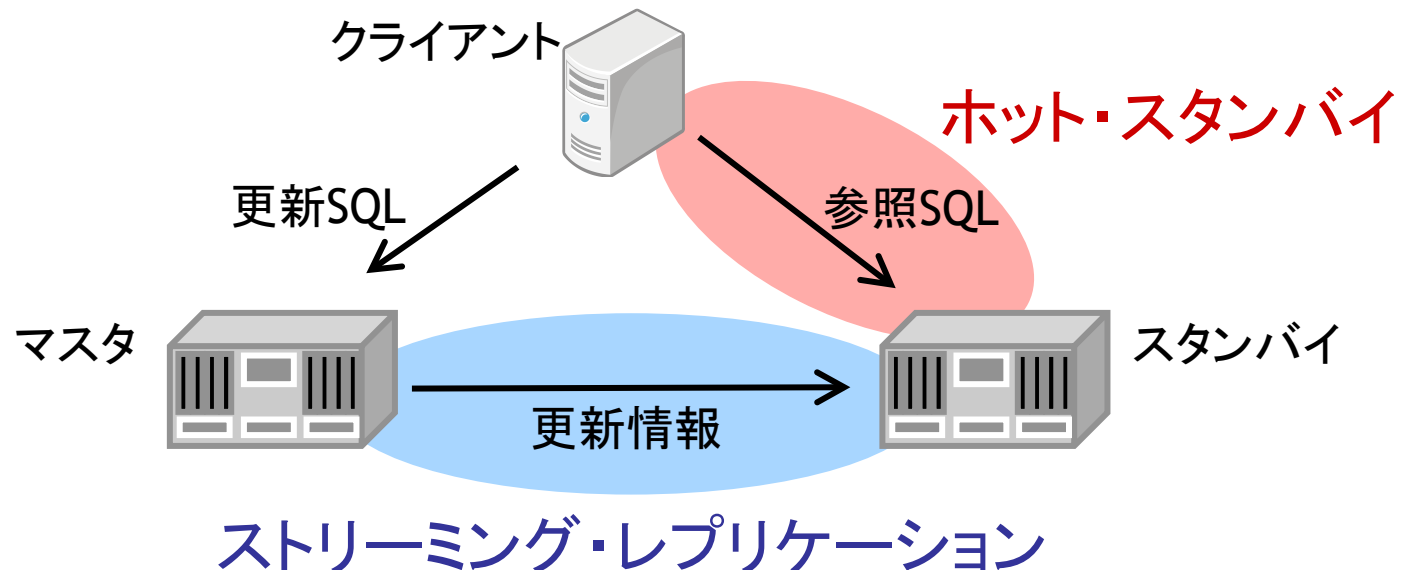


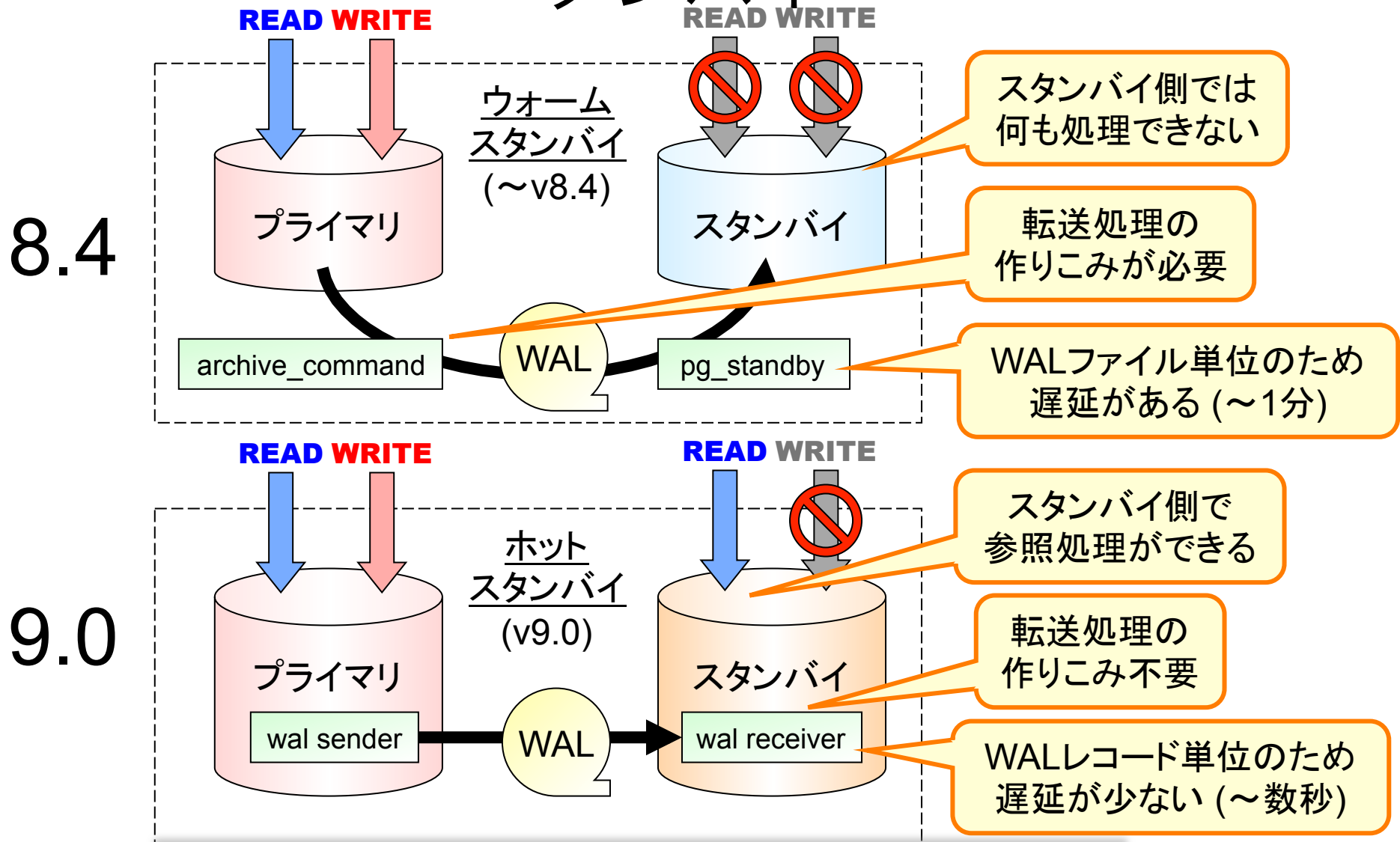
ハンズオン補足資料

PostgreSQL9.0のレプリケーションとは？

- ストリーミング・レプリケーション
 - マスタの更新内容を自動的にスタンバイに複製する機能
- ホット・スタンバイ
 - スタンバイで参照SQLを実行可能にする機能



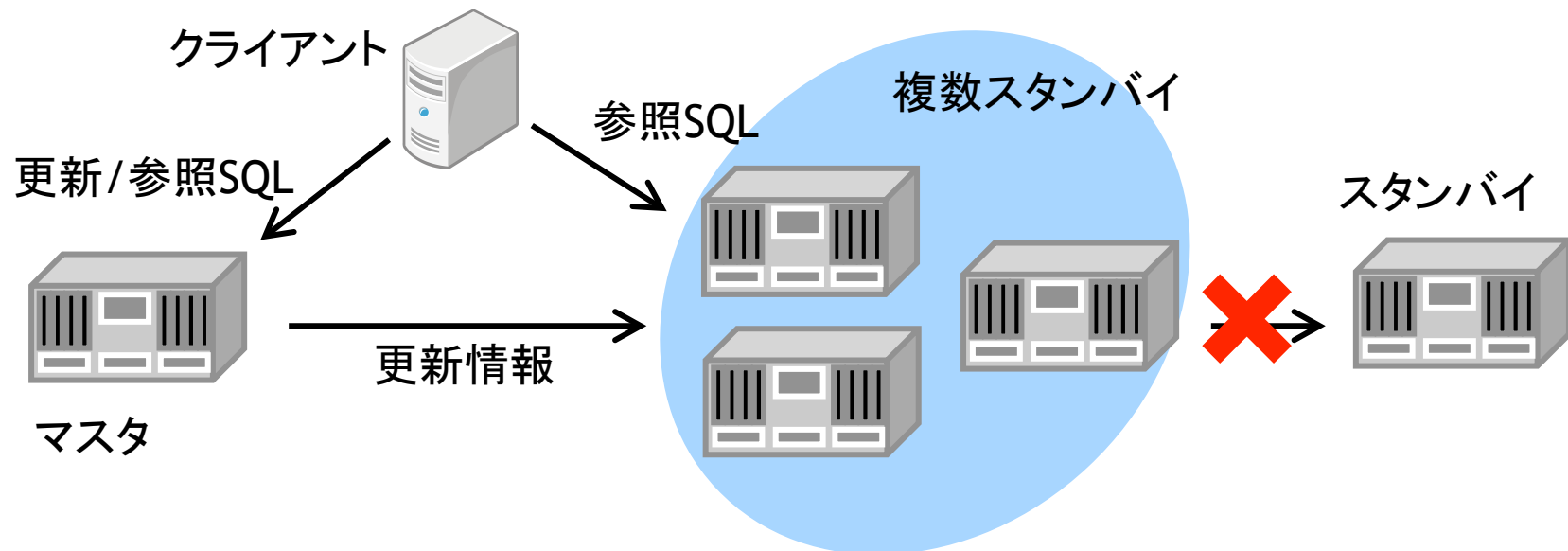
ウォーム・スタンバイ vs. ホット・スタンバイ



資源を有効活用 & 構築の手間が小さい

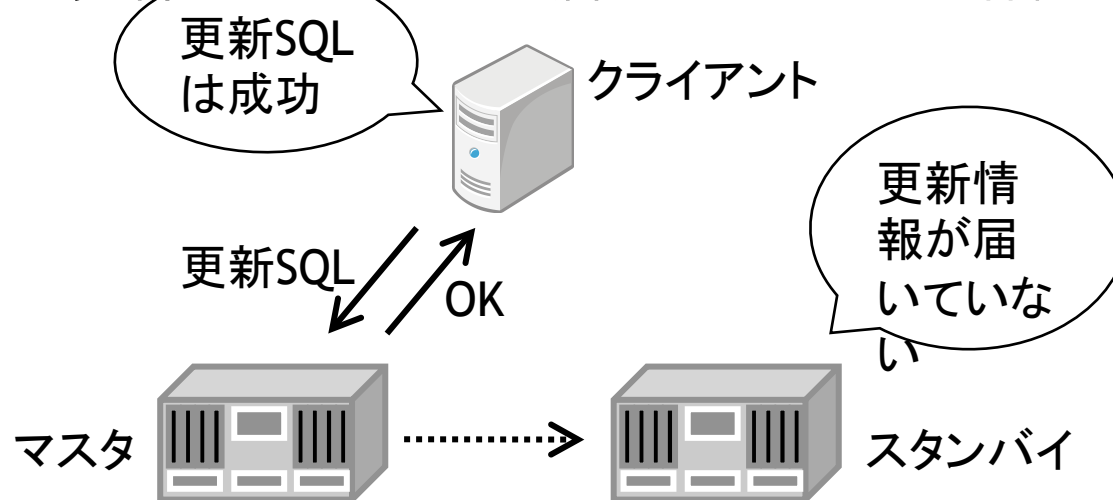
マスタ/スタンバイ型

- マスタ1台
 - 更新/参照SQLを実行できる
- スタンバイ複数台
 - 参照SQLだけ実行できる
 - VACUUM等のメンテナンスもマスタでの実行結果がスタンバイに伝播
 - カスケード接続(スタンバイにスタンバイを接続)はNG
- マスタ/スタンバイ間はNW接続
 - 高価な共有ディスクは不要
 - 遠隔地にマスタとスタンバイを配置できる



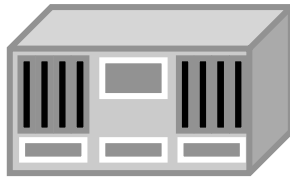
非同期レプリケーション

- マスタで完了した更新がスタンバイに届いていない可能性あり
 - フェイルオーバー時に直前(数ミリ秒前)の更新結果は失われるかも
 - 直前の更新結果をスタンバイで参照できない
- レプリケーションのオーバーヘッドは小さい
 - スタンバイに更新情報が届くのを待たずに更新SQLを完了できる
- 同期レプリケーションは9.1に提案中
 - 完了した更新がスタンバイに届いていることを保証

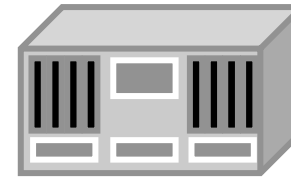


セットアップの3ステップ

マスタ

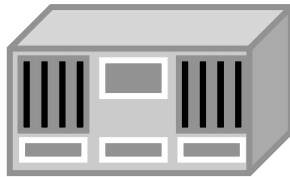


スタンバイ



セットアップの3ステップ

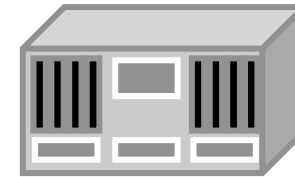
マスタ



postgresql.conf

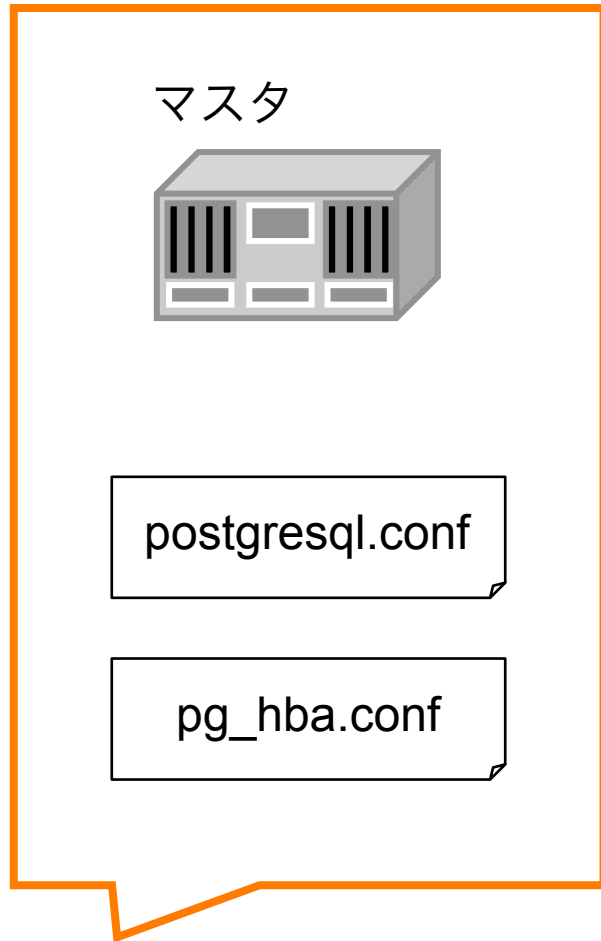
pg_hba.conf

スタンバイ

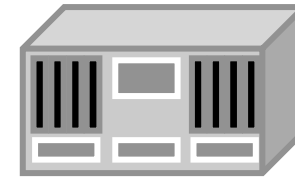


1. マスタの設定・起動

セットアップの3ステップ



スタンバイ

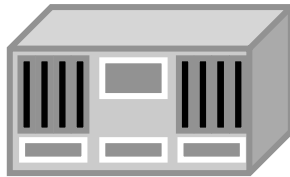


2. バックアップ
→

1. マスタの設定・起動

セットアップの3ステップ

マスタ



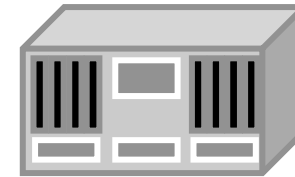
postgresql.conf

pg_hba.conf

1. マスタの設定・起動

2. バックアップ
→

スタンバイ



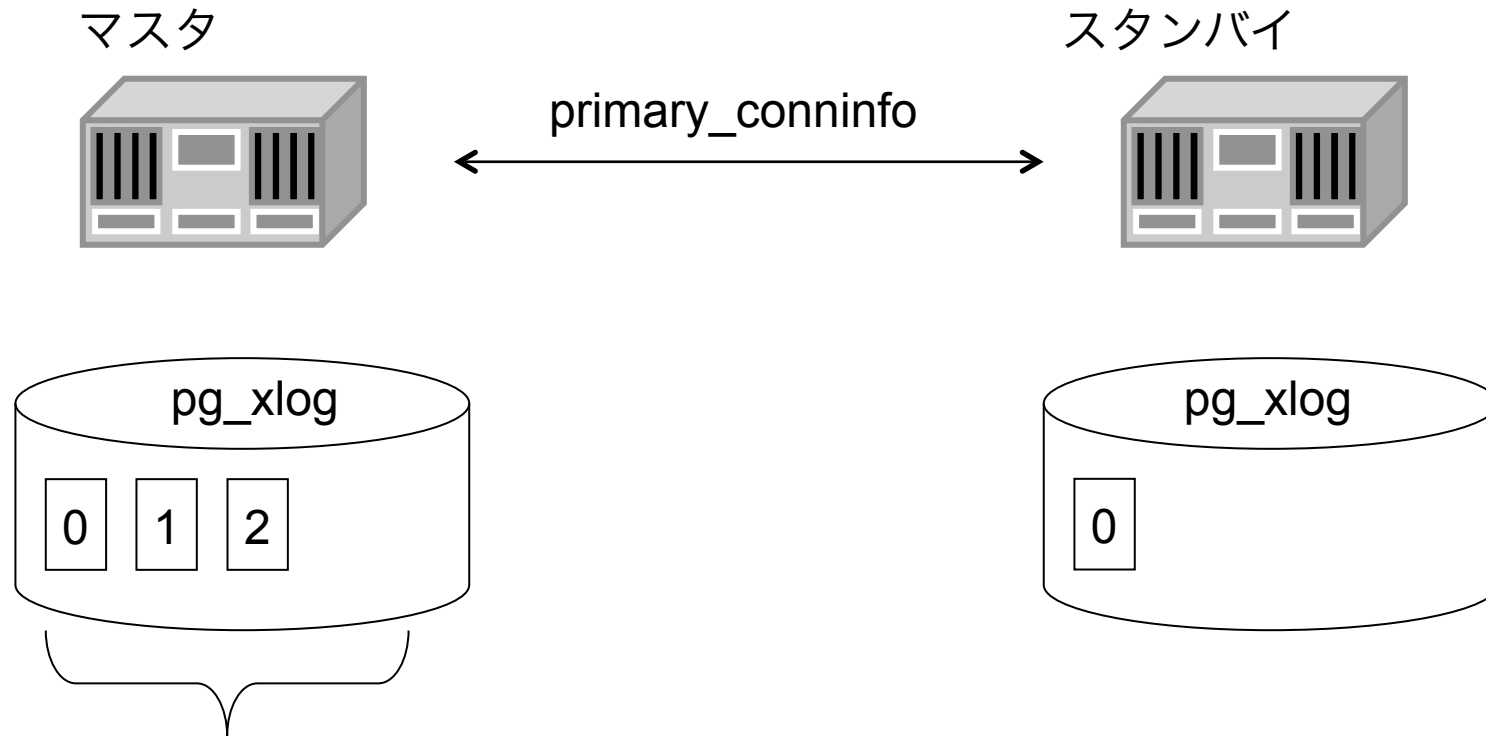
postgresql.conf

pg_hba.conf

recovery.conf

3. スタンバイの設定・起動

構成1: wal_keep_segments

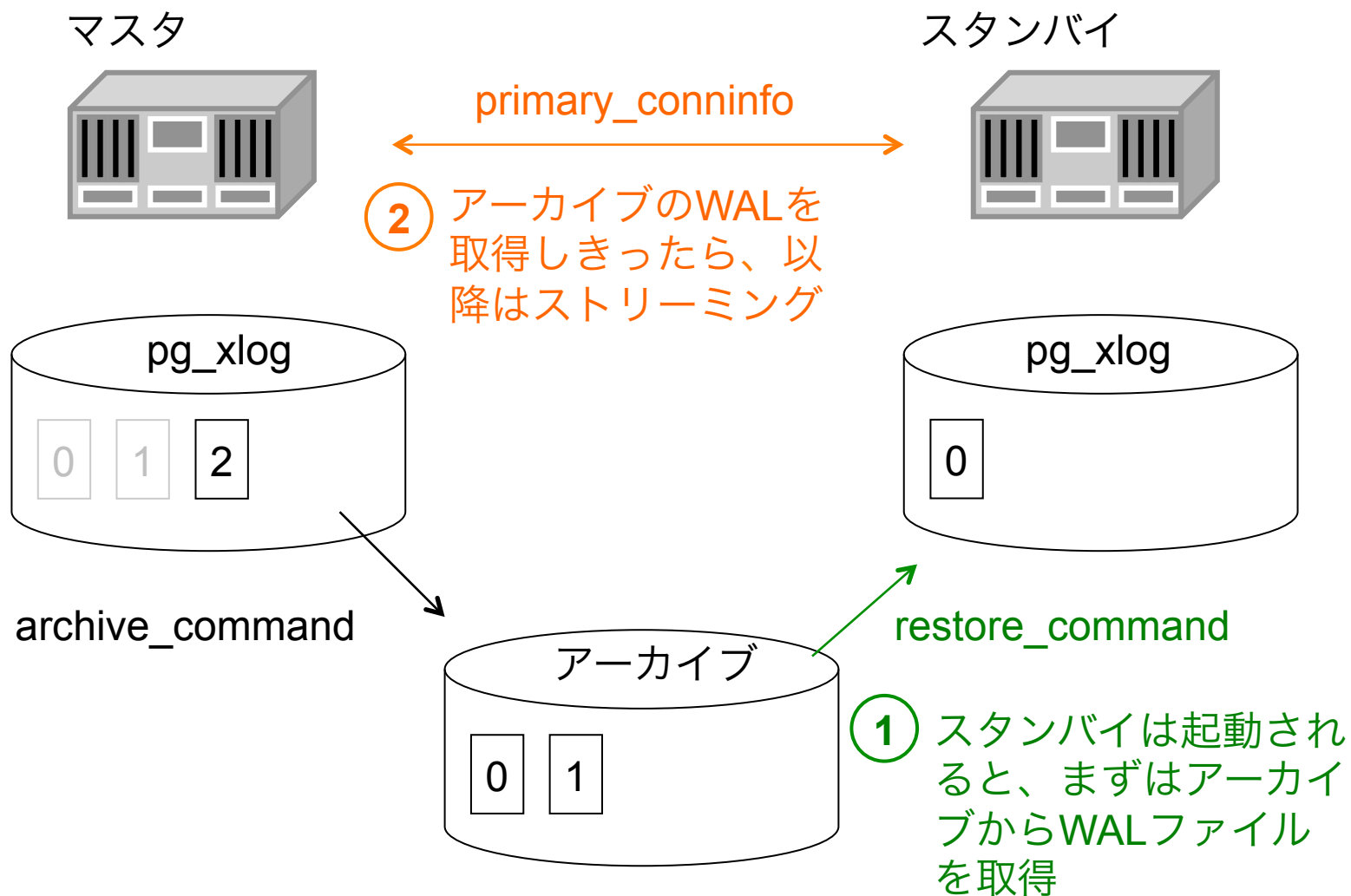


wal_keep_segments

レプリケーション用に残すWALファイルの数を指定。

転送前にWALファイルがマスタから削除されるのを回避

構成2: アーカイブの利用



ホットスタンバイって何？

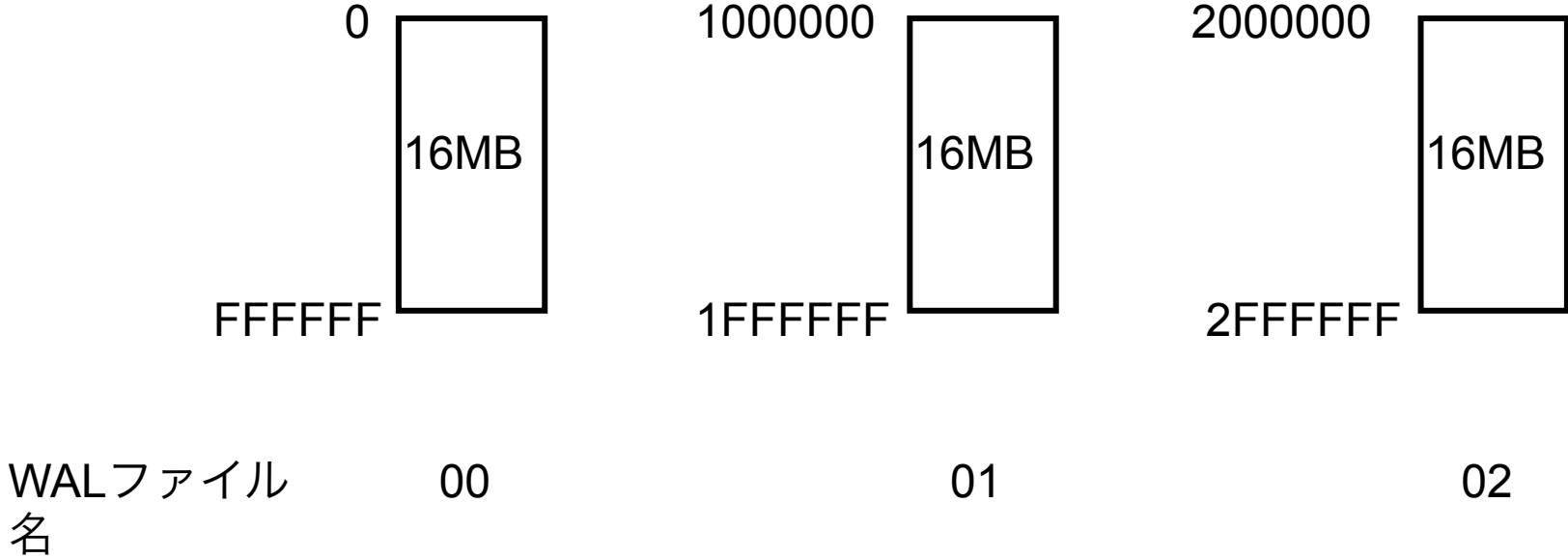
実行OK

- クエリ・アクセス
 - SELECT、COPY TO
- カーソル
 - DECLARE、FETCH、CLOSE
- パラメータ設定
 - SHOW、SET、RESET
- トランザクション
 - BEGIN、END、ABORT
- チェックポイント
 - CHECKPOINT
- 論理ホットバックアップ
 - pg_dump

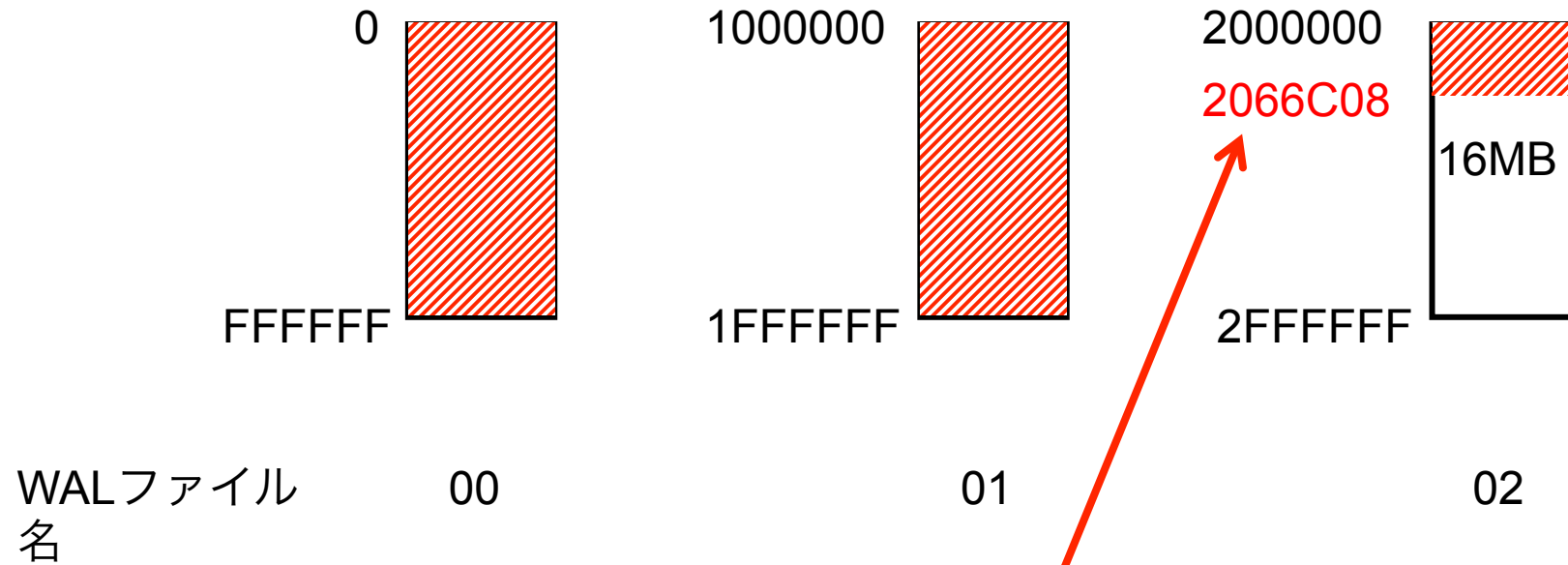
実行NG

- DML
 - INSERT、UPDATE、DELETE
 - SELECT FOR UPDATE
- DDL
 - CREATE、DROP、ALTER
- 参照以外のトランザクション
 - BEGIN READ WRITE
- メンテナンス
 - VACUUM、ANALYZE
- 物理ホットバックアップ
 - pg_start/stop_backup

LSN



LSN

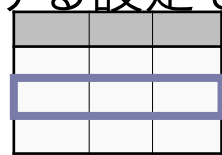


```
$ pgrep -fl postgres | grep streaming
```

```
12506 postgres: wal sender process postgres 127.0.0.1(42152) streaming 0/2066C08
```

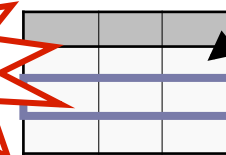
スタンバイで発生する競合とは？

- スタンバイの参照SQLとリカバリが競合する
 - マスタがVACUUMやHOTでゴミ掃除したデータを、スタンバイの参照SQLがまだ見ていた
 - HOTによるゴミ掃除の頻度は高いため、競合は高頻度で起こる可能性あり
 - スタンバイでアクセス中のデータベースを、マスタがDROPした
- 競合中はリカバリが止まる
 - 何秒競合したら参照SQLをキャンセルして、リカバリを進ませるか設定可能
 - 長時間かかる参照SQLをスタンバイで実行するときは、この値を大きく設定しておく
 - リカバリを優先したい場合は、この値を0に設定しておく
 - 競合を抑制するために、VACUUMやHOTによるゴミ掃除の頻度を下げる設定も可能



↑ 掃除

VACUUM



↑ 掃除

リカバリ