



postgresql

오픈소스dbms이며 벡터 csv등 다양한 데이터 형태를 지원한다.

420 systems in ranking, May 2024

Rank			DBMS	Database Model	Score		
May 2024	Apr 2024	May 2023			May 2024	Apr 2024	May 2023
1.	1.	1.	Oracle +	Relational, Multi-model ⓘ	1236.29	+2.02	+3.66
2.	2.	2.	MySQL +	Relational, Multi-model ⓘ	1083.74	-3.99	-88.72
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model ⓘ	824.29	-5.50	-95.80
4.	4.	4.	PostgreSQL +	Relational, Multi-model ⓘ	645.54	+0.49	+27.64
5.	5.	5.	MongoDB +	Document, Multi-model ⓘ	421.65	-2.31	-14.96
6.	6.	6.	Redis +	Key-value, Multi-model ⓘ	157.80	+1.36	-10.33
7.	7.	↑ 8.	Elasticsearch	Search engine, Multi-model ⓘ	135.35	+0.57	-6.28
8.	8.	↓ 7.	IBM Db2	Relational, Multi-model ⓘ	128.46	+0.97	-14.56
9.	9.	↑ 11.	Snowflake +	Relational	121.33	-1.87	+9.61
10.	10.	↓ 9.	SQLite +	Relational	114.32	-1.69	-19.54
11.	11.	↓ 10.	Microsoft Access	Relational	104.92	-0.49	-26.26
12.	12.	12.	Cassandra +	Wide column, Multi-model ⓘ	101.89	-1.97	-9.25
13.	13.	13.	MariaDB +	Relational, Multi-model ⓘ	93.21	-0.60	-3.66
14.	14.	14.	Splunk	Search engine	86.45	-2.26	-0.19
15.	↑ 17.	↑ 18.	Databricks +	Multi-model ⓘ	78.61	+2.28	+14.66
16.	↓ 15.	16.	Microsoft Azure SQL Database	Relational, Multi-model ⓘ	77.99	-0.41	-1.21
17.	↓ 16.	↓ 15.	Amazon DynamoDB +	Multi-model ⓘ	74.07	-3.50	-7.04
18.	18.	↓ 17.	Hive	Relational	61.17	-1.41	-12.44

2024년 자료의 기준으로 oracle - mysql - mssql - pgsql 순이다.

OpenSource DBMS

OpenSource 사용에 대한 우려 사항들

01
구축·운영·관리 측면

- 난립하는 수많은 오픈소스의 통합 및 최적화 어려움
- IT 운영 복잡성 증대
- 안정적 운영을 위한 전문 인력 확보 및 기술 내재화 어려움
- 신기술 및 지속적인 변화에 대한 신속 대응 어려움

최적화가 너무 어렵고 전문성 부족으로 불편함

02
안정성 및 보안 측면

- 패치, 업그레이드 등의 안정화 및 개선에 대한 불안
- 소스 공개로 인한 태생적 보안 취약성
- 오픈소스 품질에 대한 신뢰성 의문

새로운 버전의 사이드 이펙트 및 제품 취약점에 대한 걱정

03
기술 지원 측면

- 장애 발생 시 모호한 책임 이슈
- 이슈에 대한 불확실한 분석 및 불투명한 조치
- 신속하고 안정적인 전문 기술 지원 서비스 부족
- 오픈 소스 전환을 위한 Migration의 절차적 복잡성 및 어려움

기술 지원에 대한 신뢰성 및 전문성에 대한 만족도 부족

기술지원 비용 대비 낮은 서비스 만족도

6'00

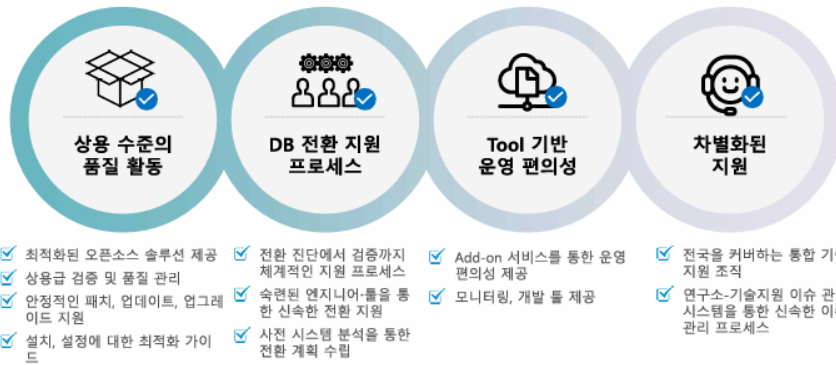
ImaxTibero

이미지의 2의 경우 소스코드가 공개되 있기에 취약점을 발견하여
 악용도리 가능성도 있으나 또한 공개되 있기에 문제를 해결하기도 쉬운 경우도 있다.

라이선스 이름	복제, 배포, 수정 권한	배포 시 소스코드 제공의무	배포시 라이선스 사본 첨부	조합저작물 작성 및 타 라이선스 배포 허용	보증의 부인	책임의 제한
GNU General Public License 2.0 (GPL 2.0)	O	전체 코드	O	조건부	O	O
GNU General Public License 2.0 (GPL 3.0)	O	전체 코드	O	X	O	O
Apache License 1.1	O	X	X	조건부	O	O
Apache License 2.0	O	X	O	O	O	O
MIT License	O	X	O	조건부	O	O
BSD-New License	O	X	O	조건부	O	O
PostgreSQL License	O	X	X	O	O	O

OpenSource DBMS

OpenSQL은 기존 Open Source의 기술 서비스 불안요소를 해결하고
티맥스티베로만의 독보적인 기술 경험과 Know-how로 최고의 서비스를 제공합니다.



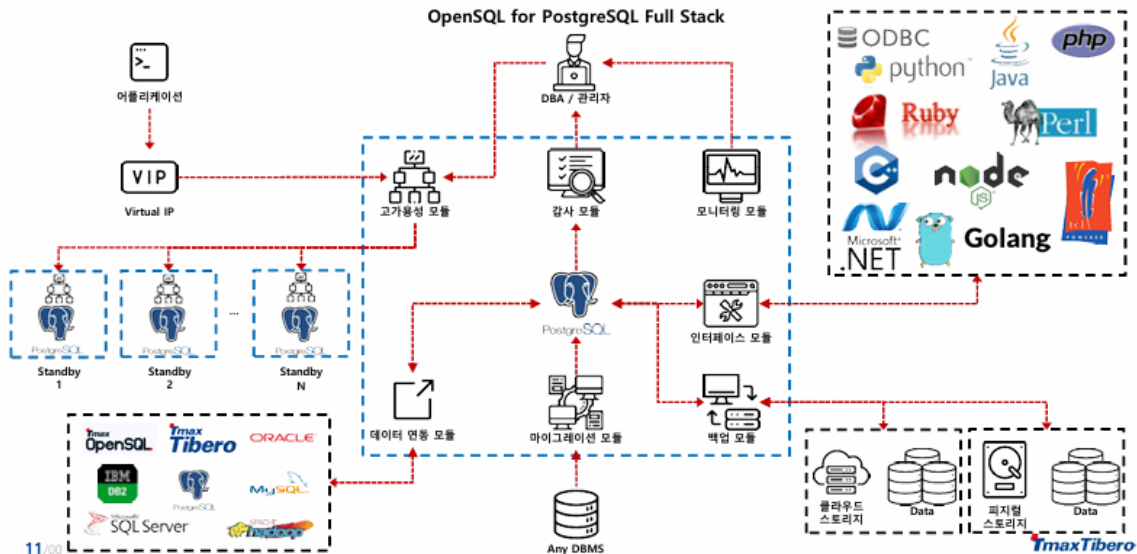
7/00

TmaxTibero

postgresql 같은 경우는 릴리즈 버전을 자주 출시하는 편인듯.

라이선스 이름	복제, 배포, 수정 권한	배포 시 소스코드 제공의무	배포시 라이선스 사본 첨부	조합저작물 작성 및 타 라이선스 배포 허용	보증의 부인	책임의 제한
GNU General Public License 2.0 (GPL 2.0)	O	전체 코드	O	조건부	O	O
GNU General Public License 2.0 (GPL 3.0)	O	전체 코드	O	X	O	O
Apache License 1.1	O	X	X	조건부	O	O
Apache License 2.0	O	X	O	O	O	O
MIT License	O	X	O	조건부	O	O
BSD-New License	O	X	O	조건부	O	O
PostgreSQL License	O	X	X	O	O	O

OpenSQL for PostgreSQL 전체 구조

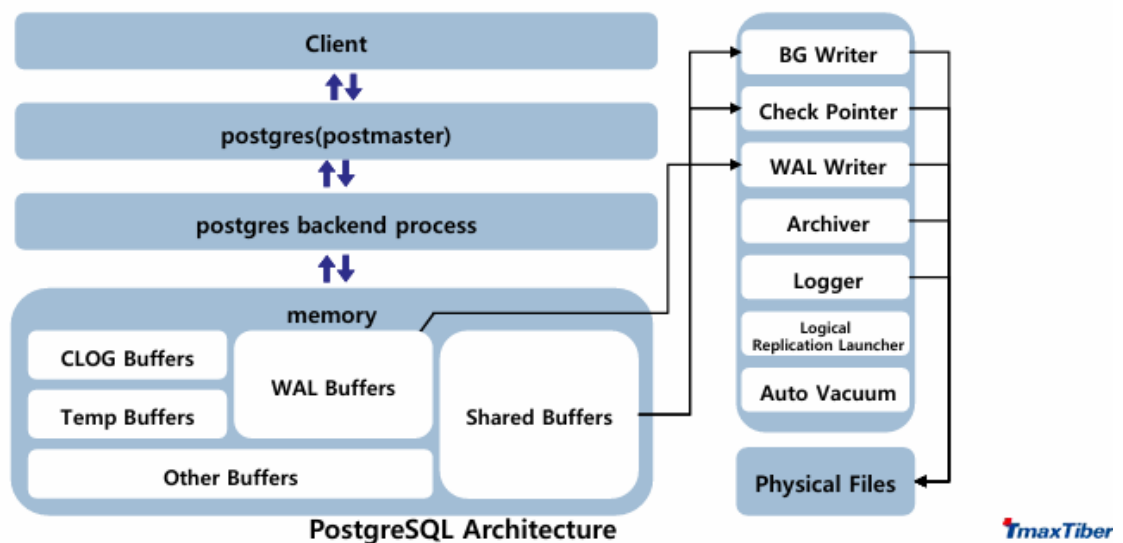


감사모듈에서 db인스턴스에 대한 제어를 수행한다.

모니터링 모듈은 node 1 or node 2 에서 분산시켜 사용한다.

데이터 연동 모듈: t_up과같은 걸 db링크라고 함.

postgres(Postmaster)의 메인 프로세스와 파생된 프로세스들로 구성



멀티 프로세스 방식의 구성이므로 커넥션풀은 지원하지 않는다.

클라이언트 - 서버 모델

리눅스에서의 프로세스를 확인가능하다.

postmaster = oracle listner

postmaster 에서 오라클의 리스너와 같이 세션의 연결할당하는 프로세스를 부여하는 데몬 프로세스이다.

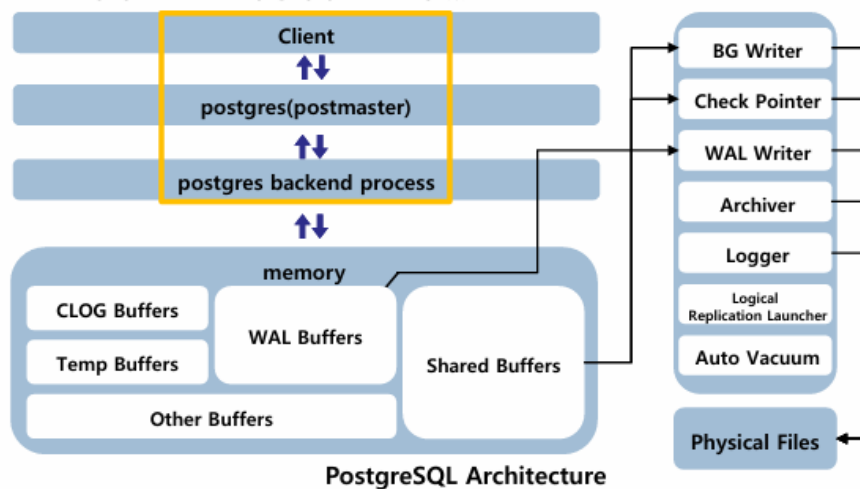
비정상 작동 체크도하며 하위프로세스가 장애시 이를 재기동 시키는 역할도 수행한다.

```
ps -ef | grep postgres
```

```
[opensql@pgpool-1:~]$ ps -ef | grep postgres
opensql 26592      1  0 Mar19 ?        00:03:07 /postgres/bin/postgres -D /home/opensql/pg/15/data
opensql 26593    26592  0 Mar19 ?        00:00:00 postgres: logger
opensql 456250   26592  0 10:08 ?        00:00:00 postgres: checkpointer
opensql 456251   26592  0 10:08 ?        00:00:00 postgres: background writer
opensql 456252   26592  0 10:08 ?        00:00:00 postgres: walwriter
opensql 456253   26592  0 10:08 ?        00:00:00 postgres: autovacuum launcher
opensql 456254   26592  0 10:08 ?        00:00:00 postgres: archiver
opensql 456255   26592  0 10:08 ?        00:00:00 postgres: logical replication launcher
opensql 456483  456409  0 10:11 pts/9    00:00:00 grep --color=auto postgres
```

• Postgres Backend Process

- Client가 요청한 SQL 및 Command를 처리하는 프로세스로 Client와의 연결이 끊어지면 종료
- Postmaster 프로세스에 의해 시작되며 Client 와는 1:1 관계
- max_connections 수치 만큼 Client가 동시에 연결 될 수 있음

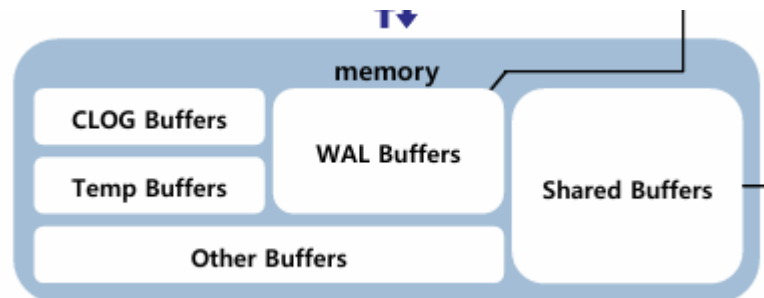


i00

TmaxTibero

백엔드 프로세스 기동 종료시 새로운 백그라운드 프로세스가 생성된다.

오라클 sga



- bg writer

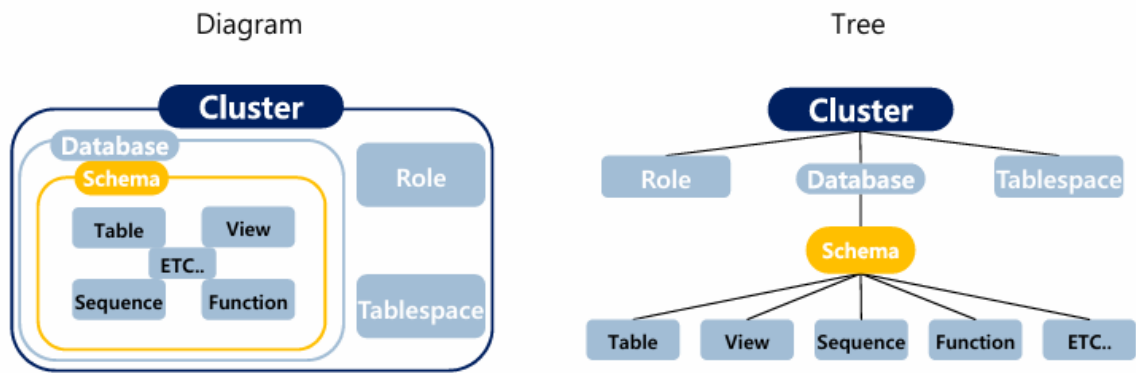
쉐어드 버퍼 디스크 블록을 기록하는 프로세스

dirty blcok: db 버퍼캐쉬에서 변경된 블록.

파라미터 이름	설명
bgwriter_delay	bgwriter 기록 간 sleep 시간 기본 값 : 200 ms
bgwriter_lru_maxpages	disk로 내릴 수 있는 최대 buffer page 개수 기본 값 : 100 buffers

대부분의 dbms는 8kb를 고수한다.

check point : sga 공간부족시 데이터 파일과 메모리 영역을 동기화 시켜준다.



3

TmaxTibero

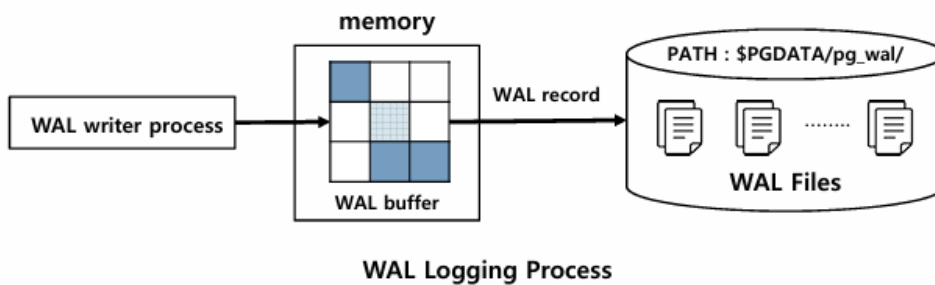
위와 같은 형태의 구조를 가지게 된다.

Checkpoint

checkpoint를 수행하는 프로세스

파라미터 이름	설명
checkpoint_timeout	Checkpoint가 수행되는 주기입니다.
max_wal_size	WAL 세그먼트가 저장되는 pg_wal 디렉토리의 용량을 제한합니다.

변경을 기록하는 체크포인트가 기록된 시점부터 장애 복구 조치를 취하게 된다.



TmaxTibero

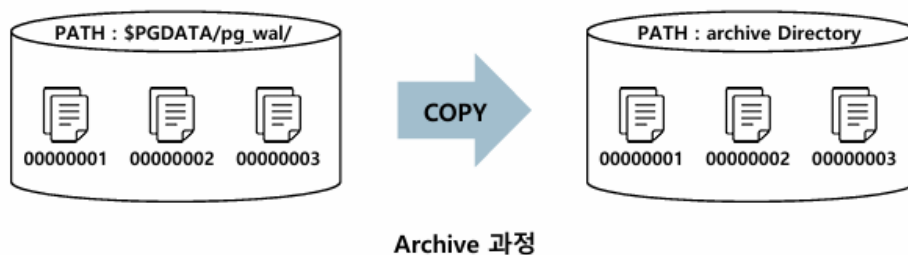
트랜잭션이 끝나기전 wql 파일로 저장된다.

변경사항은 wal buffer 버퍼에 저장되며 writer process가 조건에 따라 buffer wal file에 저장합니다.

wal writer

파라미터 이름	설명
bgwriter_delay	bgwriter 기록 간 sleep 시간 기본 값 : 200 ms
bgwriter_lru_maxpages	disk로 내릴 수 있는 최대 buffer page 개수 기본 값 : 100 buffers

archive



TmaxTibero

wal file을 백업하거나 이동함으로써 데이터베이스 복구에 일관성을 주는 작업

archive 모드 활성화시 설정된 archive 디렉토리에 wal file을 복사합니다.

일반적으로 physical 백업을 한다.

Process

archiver

아카이빙을 실행하는 프로세스

아카이빙은 wql 세그먼트가 전환될때 wal파일을 아카이브영역으로 복사하는 기능

- 관련 파라미터

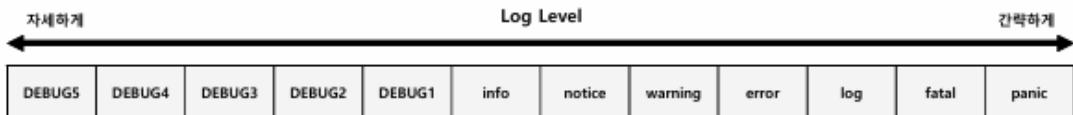
파라미터 이름	설명
archive_mode	on : archiving 사용 off : archiving 미사용 always : Standby도 archiving 사용
archive_command	WAL file을 archive 할 때 사용할 command 대상 파일명인 %f 예약어와 절대경로를 포함한 파일명 %p 예약어를 사용 할 수 있습니다.
archive_timeout	WAL file이 스위칭 할 정도로 데이터베이스가 사용량이 적은 경우를 대비하 여, 강제로 WAL file을 스위칭 하도록 주기를 설정할 수 있습니다.

database 특정 시점에 복구할때 사용된다.

wal file이 데이터량이 많을때 실행 속도가 느리므로 archive_timeout 파라미터에서 스위칭 주기를 선택가능하다.

Looger 오류발생시 로그파일에 기록하는 프로세스

파라미터 이름	설명
log_min_messages	로그 파일에 저장되는 로그의 수준을 설정합니다. 기본 값은 warning 입니다.
client_min_messages	클라이언트에게 메시지 형태로 전달되는 로그의 수준을 설정합니다. 기본 값은 notice 입니다.
log_min_duration_statement	이 파라미터 값 이상의 수행 시간이 걸린 모든 명령을 로그에 기록합니다. 단위는 miliseconds 입니다. 기본 값은 -1로 disable 입니다.



log_min_messages : 로그파일에 저장되는 로그의 수준을 설정함.

기본값: wating

client_min_messages

클라이언트에게 메시지 형태로 전달되는 로그의 수준을 설정

기본값: notice

log_min_duration_statement

이 파라미터 값 이상의 수행시간이 걸린 모든 명령을 로그에 기록하는 단위는 milliseconds 입니다.

기본값은 -1 disable입니다.

MVCC

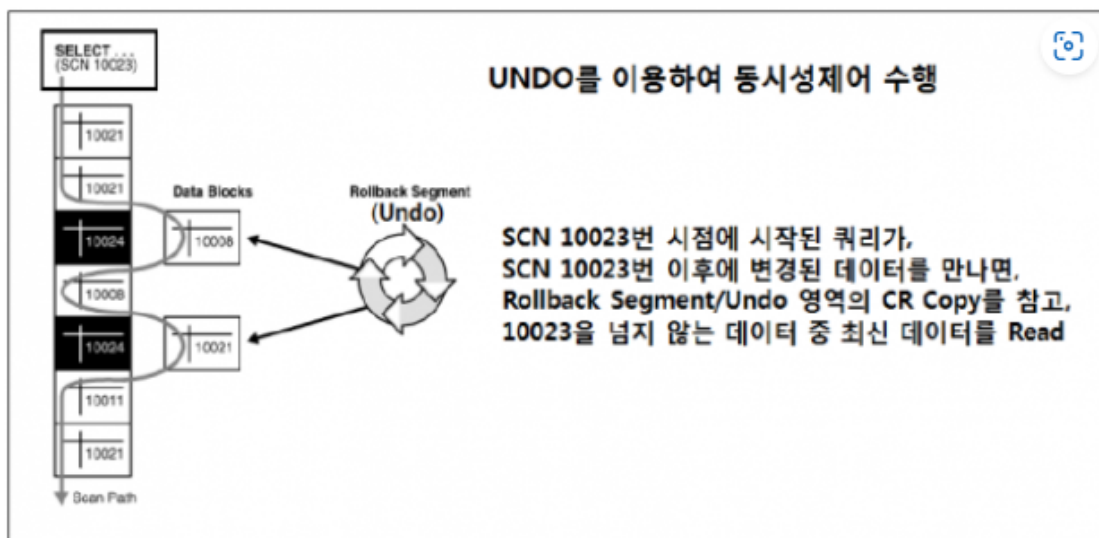
B 데이터를 C로 업데이트 했을 때

no	t_xmin	t_xmax	t_cid	t_ctid	data
1	100	105	0	(0, 2)	'A'
2	105	0	0	(0, 2)	'B'

no	t_xmin	t_xmax	t_cid	t_ctid	data
1	100	105	0	(0, 2)	'A'
2	105	105	0	(1, 1)	'B'
3	105	0	1	(1, 1)	'C'

24/00

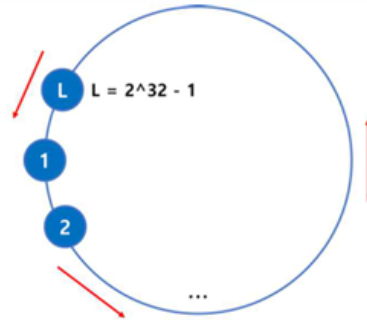
다중 버전 동시성 제어



- 데이터를 변경할 때마다 그 변경사항을 UNDO 영역에 저장

XID(Transaction ID)란?

- PostgreSQL은 트랜잭션을 관리하기 위해 트랜잭션 아이디라는 개념을 사용하고 있습니다.
- PostgreSQL은 32-bit unsigned integer($1 \sim 2^{32} - 1$)로 표현되어 약 42억 개의 트랜잭션 아이디를 가질 수 있고 1부터 시작하여 새로운 트랜잭션이 시작될 때 증가합니다.
 - (0은 invalid tx_id로 예약되어 있습니다).
 - 1과 2는 특수한 용도로 사용되므로, 실제 오브젝트에 대한 트랜잭션 아이디는 3부터 시작합니다.
 - PostgreSQL은 이를 원형(Circular) 형태로 관리하고 있습니다.



25/00

TmaxTibero

실제 운영은 3부터 시작한다.

VACUUM

- 오래되거나 사용되지 않은 데이터나 dead tuple 를 삭제 또는 갱신된데이터로 정리해주는 기술
- 디스크 공간이 확보되고 시간이 지남에따라 속도가 느려진느 것을 방지한다.

chat

1. MVCC (Multi-Version Concurrency Control):

- MVCC는 데이터베이스 관리 시스템(DBMS)에서 동시성 제어를 위한 기술입니다.
- 여러 사용자가 동시에 데이터베이스에 접근할 때, 일관성과 동시성을 보장하기 위해 사용됩니다.
- MVCC는 각 트랜잭션에 대해 데이터의 여러 버전을 유지하고, 트랜잭션 간의 간섭 없이 병행 처리를 가능하게 합니다.
- 보통 MVCC는 트랜잭션의 시작 시점에 트랜잭션에서 읽을 수 있는 데이터의 스냅샷을 만들어 사용합니다.

2. 베큘 (Bekkun):

- 베큘은 데이터베이스에서의 데이터 바인딩을 의미합니다.

- 일반적으로 베akup은 데이터베이스의 테이블 간에 데이터의 일관성을 유지하기 위해 사용됩니다.
- 예를 들어, 하나의 테이블에서 다른 테이블로의 참조 무결성 제약 조건을 설정하여 데이터베이스의 일관성을 보장할 수 있습니다.
- 이러한 베akup은 데이터베이스 관리 시스템의 성능과 데이터 무결성을 유지하는데 중요한 역할을 합니다.

Autovacuum Launcher

- Autovacuum 을 수행관리하는 프로세스

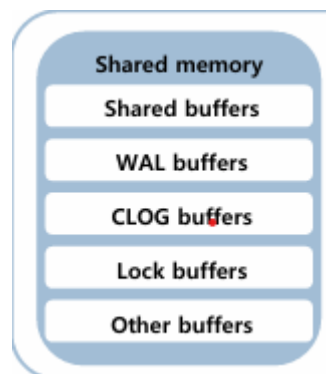
*Autovacuum_max_workers*의 Autovacuum Workers 프로세스 호출합니다.

- 관련 파라미터

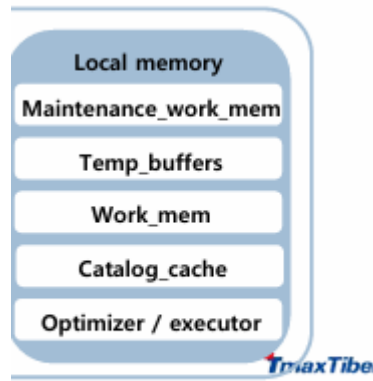
파라미터 이름	설명
<code>autovacuum</code>	default가 on으로 autovacuum process를 enable 함
<code>autovacuum_max_workers</code>	worker process 최대 수
<code>autovacuum_naptime</code>	autovacuum 수행 간의 sleep 시간



shared memory area



program global area



maininstance 베류 스키마 자원들 제어

temp_buffers: 각 데이터베이스에 임시 버퍼를 설정한다.

work_mem: 정렬 작업 개선을 위해 사용된다.

catalog_cahce

optimizer / executor: 실행계획을 구해 최적의 실행계획을 짜는 옵티마이저가 최선의 선택을 하는 작업공간

executor:사용자에게 값을 반환할 수 있게 도와주는 공간

Shared Memory

datablock 및 트랜잭션 로그와 같은 정보들을 캐싱하는 공간.

shared buffer

data와 data의 변경사항을 기록하고 자주 사용하는 데이터블록을 저장함에 따라 디스크 까지 경로를 안잡아도되 바른 i/o를 유도한다.

일단 shared buffer로 가고 없다 → os 버퍼에 가고 → 마지막으로 디스크에서 데이터 블록을 가지고 온다.

wal buffer

session 들이 수행한느 트랜잭션에 대한 변경로그를 캐싱하는 공간

복구 작업시 data를 재구성할 수 있도록하는 영역이다.

clog buffer

각 트랜잭션의 정보를 캐싱하는 공간

따로 파라미터를 설정 불가하며 데이터베이스 엔진에 의해 관리된다.

lock buffer

shared memory 영역중 lock 과 관련된 내용을 보관하는 영역

Localmemory

- 개별 백그라운드 프로세스가 할당 받아 사용되는 공간
- local memory의 수치는 개별공간의 크기를 의미하기 때문에 전체 Connection 수치를 감안하여 설정.

maintenance work memory

- 유지관리 작업에 사용되는 메모리
- vacuum 관련작업 인덱스 생성 테이블 변경 foreign key 추가 등의 작업에 사용되는공간
- 관련작업의 성능을 향상시킴려면 해당영역의 사이를 변경해야됨.

temp buffer

temporary 테이블에 사용되는 공간

temp파일과는 상관없다.

temp tablespace를 사용할때만 사용되는 영역입니다.

work memory

과도한 sort/hash 작업이 발생하여 temp 파일을 사용전에 사용되는 공간

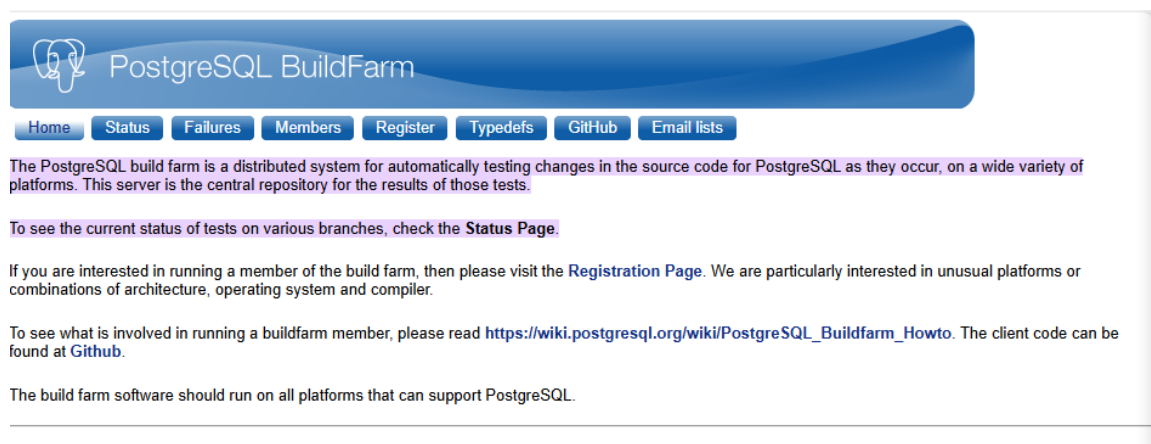
catalog cache

systemcatalog 메타데이터를 이용시 사용하는 공간

세션이 메타데이터 조회시 디스크에 들어오는 속도 저하가 발생되기에 개별 메모리로서 존재한다.

optimizer / excutor

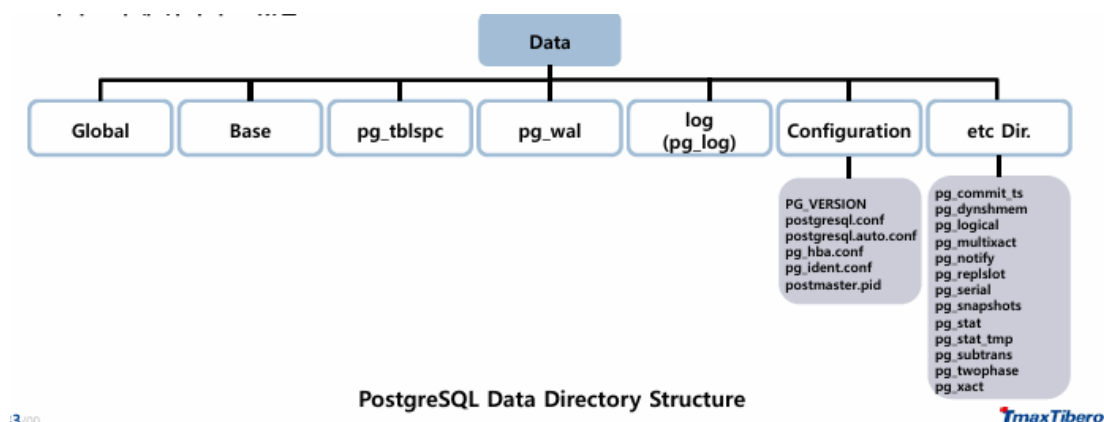
수행한 쿼리들에 대한 최적의 실행계획을 수립하고 실행계획의 실행을 전담하는 공간



PostgreSQL BuildFarm

컴파일러 버전 확인 및 아키텍처들 os의 조합확인이 가능하다.

Data 디렉토리구조(PGDATA)



cluster 초기화시 data 디렉토리가 생성되며, 해당 디렉토리 밑으로 구성 요소인 다양한 파일 및 디렉토리 위치

디렉토리 구조

• Data 영역

- 실제 사용자 데이터가 저장되는 공간이며, Global, base, pg_tblspc 디렉토리로 구분
- Global Directory
 - 초기 설치 시 생성되는 Default Global Tablespace 디렉토리로, pg_global Tablespace라고도 명칭
 - 해당 공간은 PostgreSQL Cluster 하위의 Database들이 서로 공유하는 정보가 저장
 - pg_internal.init, pg_filenode.map : Database 내에 Object의 속성 정보 및 실제 Object 들의 정보, Data 파일과의 Mapping 정보
 - pg_control : 메타 정보를 가지고 있는 파일, Oracle의 Control 파일과 유사
 - 기타 숫자 파일들 : Tablespace, Database, Role 등 Database가 공통으로 사용하는 Data 정보를 포함
- Base Directory
 - 초기 설치 시 생성되는 Default Tablespace 디렉토리 중 하나이며, pg_default Tablespace 라고 명칭
 - 테이블, 인덱스, 함수와 같은 Object가 실제로 저장되는 공간
 - Base 디렉토리에는 Database oid를 이름으로 갖는 디렉토리들이 존재

```
[opensql@pgpool-1:base]$ ls
1 14485 14486

[opensql@pgpool-1:base]$ oid2name -U postgres
All databases:
Oid Database Name Tablespace
-----
5 postgres pg_default
4 template0 pg_default
1 template1 pg_default
```

34/00

maxTibero

Global Directory

초기 설치 생성되는 디폴트 글로벌 테이블 스페이스 디렉토리로 pg_global tablespace라고 명칭

해당 공간은 postgresql 데이터베이스 내에 object의 속성 정보 및 실제 object들의 정보 data file과의 mapping 정보

pg_control : 메타정보를 가지고 있는 파일 oracle control과 유사한 구조

```
pg_control_system();
```

숫자파일들: tablespace,database,roid 등 데이터베이스가 공통으로 사용하는 데이터들을 저장.

base directory

초기 설정시 생성되는 default tablespace 디렉토리중 하나이며 pg_default tablespace라고 명칭