

Table des matières

Avant-propos	xi
1. Instance	1
1.1. Concept général	1
1.2. Au niveau du système d'exploitation	1
1.3. Au niveau du système de base de données	3
1.4. Groupes d'instances	4
2. Fichiers	5
2.1. Répertoire de données principal	5
2.2. Bases de données, relations et métadonnées des relations	8
Répertoire <i>base</i>	8
Répertoire <i>global</i>	21
2.3. Tablespace	23
2.4. Transactions	26
Répertoire des journaux de transactions	26
Répertoire d'état des transactions	31
Répertoire d'horodatages des transactions	32
Répertoire des points de retour	33
Répertoire des verrous partagés	33
Répertoire des données 2PC	34
2.5. Statistiques d'activité	34
2.6. Traces	36
2.7. Sous-systèmes	36
Système de notifications	36
Snapshots de base	37
SSI	38
Mémoire partagée dynamique	39
Réplication logique	39
Slots de réplication	39
2.8. Autres fichiers	40
2.9. Répartition du stockage	43
3. Contenu physique des fichiers	45
3.1. Structure générale d'un bloc d'une relation	45
3.2. Tables	51
Fichier HEAP	51
Fichier TOAST	56

3.3. Index	58
Index B-tree	58
Index Hash	66
Index GIN	70
Index GiST	74
Index SP-GiST	75
Index BRIN	76
Outil amcheck	80
3.4. Journaux de transactions	81
4. Architecture des processus	85
4.1. Démarrage et gestion des sous-processus	87
postmaster	90
startup	98
4.2. Écriture dans les fichiers de données	105
writer	106
checkpointer	108
4.3. Écriture dans les journaux de transactions : wal writer.....	118
4.4. Gestion des statistiques d'activité : stats collector.....	120
4.5. Gestion des traces : logger.....	124
4.6. Gestion automatique des VACUUM et ANALYZE	127
autovacuum launcher	128
autovacuum worker	130
4.7. Archivage des journaux de transactions : archiver.....	134
4.8. Gestion de la réplication	136
wal sender	137
wal receiver	138
4.9. Gestion des processus d'arrière-plan : background worker.....	142
4.10. Gestion des connexions : postgres.....	142
4.11. Liens entre répertoires et processus	145
5. Architecture mémoire	147
5.1. Mémoire partagée	148
Vue générale	148
Cache disque des relations	149
Cache disque des journaux de transactions et mémoire partagée	
du système transactionnel	154
Verrous	155
Sessions	158
Divers	160

Implémentation au niveau système	160
5.2. Mémoire par processus	161
Mémoire de travail	161
Mémoire pour les opérations de maintenance	163
Mémoire cache pour les objets temporaires	164
5.3. Manque de mémoire et traces	164
5.4. Liens entre répertoires, processus et mémoire	166
6. Protocole de communication	169
6.1. Protocole standard	169
Connexion	170
Exécution de requêtes	173
Cas particuliers	177
Déconnexion	178
6.2. Protocole de réplication	178
Messages pour la réplication	179
Messages pour la sauvegarde des fichiers	180
6.3. Outil d'étude	182
7. Gestion des connexions	187
7.1. Aperçu du processus de connexion	187
7.2. Établissement d'une connexion	189
Via un socket de domaine Unix	189
Via un socket réseau	190
Gestion des processus	192
7.3. Authentification	193
7.4. Impact des connexions	195
8. Gestion des transactions	197
8.1. Aspects théoriques	197
Définition d'une transaction	197
Identifiants de transaction	198
Transaction implicite et transaction explicite	200
Propriétés ACID	200
Niveaux d'isolation	204
Transaction imbriquée et transaction autonome	206
Savepoint	206
Durée de vie d'une transaction	207
Two-Phase Commit (2PC)	208
Cas particulier des procédures stockées	208
8.2. Implémentation interne	211

- MVCC 211
- Vie et mort d'une ligne 213
- Horodatage d'une ligne 215
- Emplacement physique d'un enregistrement 215
- Limite des identifiants de transaction 216
- Commit log et hint bits 218
- Accès concurrents et verrous 218
- 8.3. Contournement des inconvénients 221
 - Maintenance 221
 - Optimisations 223
- 9. Gestion des objets 225**
 - 9.1. Bases 225
 - 9.2. Tablespaces 228
 - 9.3. Schémas 230
 - 9.4. Tables 232
 - 9.5. Index 237
 - 9.6. Vues 240
 - 9.7. Séquences 241
 - 9.8. Langages de procédures stockées 241
 - 9.9. Procédures stockées 242
 - 9.10. Triggers 245
 - 9.11. Types 246
 - 9.12. Opérateurs 247
 - 9.13. Recherche plein texte 248
 - 9.14. SQL/MED 249
 - 9.15. Méthodes d'accès 250
 - 9.16. Extensions 250
 - 9.17. Rôles 251
- 10. Planification des requêtes 253**
 - 10.1. Introduction à l'optimiseur de requêtes 253
 - Fonctionnement de l'optimiseur 253
 - Plans d'exécution 255
 - 10.2. Nœuds d'exécution d'un plan 256
 - Accès aux données 256
 - Jointures 271
 - Agrégats 274
 - Autres 278
 - 10.3. Utilisation d'EXPLAIN 282

Informations de base	282
Options	284
10.4. Statistiques et coûts	287
Mise à jour des statistiques	287
Statistiques	288
Statistiques étendues	291
Paramétrage	291
Estimation de coûts	293
10.5. Problèmes et optimisations	296
Index inutilisés	296
Index manquants	297
Parcours séquentiels synchronisés	298
Suppression de jointures inutiles	298
Suppression d'éléments inutiles dans une clause GROUP BY	299
10.6. Outils	300
Affichage graphique d'un plan d'exécution	300
Aide à la recherche du nœud lent	301
Trace automatique des plans d'exécution	303
10.7. Exécuteur	305
11. Sauvegarde et restauration	309
11.1. Sauvegarde physique des fichiers à froid	310
11.2. Export logique	313
Fonctionnement interne	314
Format de sauvegarde	316
Sauvegarde complète	319
Sauvegarde partielle	320
Cohérence de la sauvegarde et verrous	321
Sauvegarde parallélisée	321
Sauvegarde des Large Objects	323
Compatibilité entre versions majeures	323
11.3. Export global	324
Format de sauvegarde	324
Sauvegarde complète	325
Sauvegarde partielle	325
Cohérence de la sauvegarde	325
Exemple d'utilisation de pg_dump et pg_dumpall	326
11.4. Import logique	326
Restauration d'une sauvegarde texte	327
Restauration d'une sauvegarde logique au format binaire	327

Opérations post-import	331
11.5. Sauvegarde physique des fichiers à chaud, en continu	332
Avantages et inconvénients	332
Archivage des journaux de transactions	333
Sauvegarde des fichiers	336
Restauration d'une sauvegarde PITR	338
11.6. Gestionnaires de sauvegardes	341
11.7. Remarques sur les fichiers de configuration	342
11.8. Considérations sur les sauvegardes	342
12. Réplication	343
12.1. Terminologie et vue d'ensemble	343
Cluster de réplication	343
Réplication physique	343
Réplication logique	344
Log Shipping et Streaming Replication	344
Warm Standby et Hot Standby	345
Secondaire asynchrone/synchrone	345
Réplication en cascade	347
12.2. Préparation du serveur primaire	349
Répertoire d'archivage	350
Configuration du serveur primaire	350
12.3. Mise en place d'un serveur secondaire	351
Copie des fichiers	352
Configuration	353
Configuration des différents types de serveurs secondaires	354
12.4. Utilisation d'un serveur secondaire	358
Type de requêtes	358
Durée d'exécution	359
Retard d'un serveur secondaire	360
Nettoyage des journaux archivés	361
Configuration d'un serveur secondaire	362
Bascule	362
Supervision	364
12.5. Réplication logique	364
Particularités des publications	367
Particularités des instances	368
Limitations	368
12.6. Autres solutions	369

13. Statistiques d'activité	371
13.1. Fonctionnement	371
Processus	371
Mémoire	372
Fichiers	372
Aperçu du fonctionnement global	372
Âge des statistiques	374
13.2. Statistiques et calculs intéressants	375
Sessions	375
Processus	380
Bases de données	381
Tables	385
Index	386
Séquences	387
Procédures stockées	388
Requêtes	389
13.3. Historisation et alertes	390
Analyse en direct : pgstat	390
Sondes Nagios	391
Analyse graphique en direct des requêtes : PoWA.....	393
Autres outils	394
Pour la suite	394
14. Maintenance	395
14.1. Opérations de maintenance	395
Lutter contre la fragmentation des tables	395
Mettre à jour les statistiques sur les données	403
Lutter contre la fragmentation des index	404
Réorganiser les données d'une table	405
14.2. Fréquence et automatisation	406
Automatisation de <i>VACUUM</i> et <i>ANALYZE</i>	406
Automatisation de <i>REINDEX</i>	407
Automatisation de <i>CLUSTER</i>	408
15. Sécurité	409
15.1. Au niveau du système d'exploitation	409
Utilisateur postgres	409
Accès aux fichiers	410
Accès aux données	411
Espionnage de la communication	413

- Pare-feu 414
- Injection SQL 414
- SELinux 414
- 15.2. Au niveau de la base 415
 - Rôle 415
 - Mot de passe 417
 - Droits par défaut d'un rôle 417
 - Attributs 418
 - Droits sur la définition des objets 420
 - Droits sur l'interaction avec les objets 420
 - Droits sur des objets à créer 421
 - Droits sur les lignes d'une table 422
 - Options de sécurité des objets 422
 - Suppression d'un rôle 423
 - Labels de sécurité 424
- Lexique 425**
- Index 431
- À propos des auteurs 447