

Table des matières

À propos des auteurs	xii
Avant-propos	xiii
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	9
Répertoire <i>base</i>	9
Répertoire <i>global</i>	22
2.3. Tablespaces	25
2.4. Transactions	27
Répertoire des journaux de transactions	27
Répertoire d'état des transactions	33
Répertoire d'horodatages des transactions	34
Répertoire des points de retour	34
Répertoire des verrous partagés	35
Répertoire des données 2PC	35
2.5. Statistiques d'activité	36
2.6. Traces	37
2.7. Sous-systèmes	38
Système de notifications	38
Snapshots de base	38
SSI	40
Mémoire partagée dynamique	40
Réplication logique	40
Slots de réplication	41
2.8. Autres fichiers	41
2.9. Répartition du stockage	44
3. Contenu physique des fichiers	45
3.1. Méthode d'accès à une relation	45
3.2. Structure générale d'un bloc d'une relation	46
3.3. Tables	51

Fichier HEAP	52
Fichier TOAST	56
3.4. Index	58
Index B-tree	59
Index Hash	67
Index GIN	71
Index GiST	75
Index SP-GiST	76
Index BRIN	77
Outil amcheck	81
3.5. Journaux de transactions	82
4. Architecture des processus	86
4.1. Démarrage et gestion des sous-processus	88
postmaster	91
startup	100
4.2. Écriture dans les fichiers de données	107
writer	108
checkpointer	110
4.3. Écriture dans les journaux de transactions : wal writer	120
4.4. Gestion des statistiques d'activité : stats collector	122
4.5. Gestion des traces : logger	127
4.6. Gestion automatique des <i>VACUUM</i> et <i>ANALYZE</i>	130
autovacuum launcher	131
autovacuum worker	133
4.7. Archivage des journaux de transactions : archiver	137
4.8. Gestion de la réplication physique	139
wal sender	140
wal receiver	141
4.9. Gestion de la réplication logique	145
logical replication launcher	145
logical replication worker	145
4.10. Gestion des processus d'arrière-plan : background worker	146
4.11. Gestion des connexions : postgres	147
4.12. Liens entre répertoires et processus	149
5. Architecture mémoire	150
5.1. Mémoire partagée	151
Vue générale	151
Cache disque des relations	152

Cache disque des journaux de transactions et mémoire partagée du système transactionnel	158
Verrous	158
Sessions	161
Divers	164
Implémentation au niveau système	164
5.2. Mémoire par processus	166
Mémoire de travail	166
Mémoire pour les opérations de maintenance	167
Mémoire cache pour les objets temporaires	168
5.3. Manque de mémoire et traces	168
5.4. Liens entre répertoires, processus et mémoire	170
6. Protocole de communication	172
6.1. Protocole standard	172
Connexion	173
Exécution de requêtes	176
Cas particuliers	180
Déconnexion	181
6.2. Protocole de réplication	181
Messages pour la réplication	182
Messages pour la sauvegarde des fichiers	183
6.3. Outil d'étude	185
7. Gestion des connexions	189
7.1. Aperçu du processus de connexion	189
7.2. Établissement d'une connexion	191
Via un socket de domaine Unix	191
Via un socket réseau	192
Gestion des processus	194
7.3. Authentification	195
7.4. Impact des connexions	197
8. Gestion des transactions	199
8.1. Aspects théoriques	199
Définition d'une transaction	199
Identifiants de transaction	200
Transaction implicite et transaction explicite	202
Propriétés ACID	202
Niveaux d'isolation	206
Transaction imbriquée et transaction autonome	208

Savepoint	208
Durée de vie d'une transaction	209
Two-Phase Commit (2PC)	210
Cas particulier des routines	210
8.2. Implémentation interne	213
MVCC	213
Vie et mort d'une ligne	215
Horodatage d'une ligne	217
Emplacement physique d'un enregistrement	217
Limite des identifiants de transaction	219
Commit log et hint bits	220
Accès concurrents et verrous	221
8.3. Contournement des inconvénients	224
Maintenance	224
Optimisations	226
9. Gestion des objets	228
9.1. Bases	228
9.2. Tablespaces	231
9.3. Schémas	233
9.4. Tables	235
9.5. Index	241
9.6. Vues	244
9.7. Séquences	245
9.8. Langages de routines	246
9.9. Routines	247
9.10. Triggers	249
9.11. Types	251
9.12. Opérateurs	252
9.13. Recherche plein texte	253
9.14. SQL/MED	253
9.15. Méthodes d'accès	255
9.16. Extensions	255
9.17. Rôles	256
10. Planification des requêtes	258
10.1. Introduction à l'optimiseur de requêtes	258
Fonctionnement de l'optimiseur	258
Plans d'exécution	260
10.2. Nœuds d'exécution d'un plan	261

Accès aux données	261
Jointures	276
Agrégats	279
Autres	283
10.3. Utilisation d'EXPLAIN	288
Informations de base	288
Options	290
10.4. Statistiques et coûts	293
Mise à jour des statistiques	294
Statistiques	295
Statistiques étendues	297
Paramétrage	298
Estimation de coûts	300
10.5. Problèmes et optimisations	303
Index inutilisés	303
Index manquants	304
Parcours séquentiels synchronisés	304
Suppression de jointures inutiles	305
Suppression d'éléments inutiles dans une clause GROUP BY	306
10.6. Outils	307
Affichage graphique d'un plan d'exécution	307
Aide à la recherche du nœud lent	308
Trace automatique des plans d'exécution	310
10.7. Exécuteur	312
11. Sauvegarde et restauration	315
11.1. Sauvegarde physique des fichiers à froid	316
11.2. Export logique	319
Fonctionnement interne	320
Format de sauvegarde	321
Sauvegarde complète	325
Sauvegarde partielle	326
Cohérence de la sauvegarde et verrous	327
Sauvegarde parallélisée	327
Sauvegarde des Large Objects	329
Compatibilité entre versions majeures	329
11.3. Export global	330
Format de sauvegarde	330
Sauvegarde complète	330
Sauvegarde partielle	331

Cohérence de la sauvegarde	331
Exemple d'utilisation de pg_dump et pg_dumpall	332
11.4. Import logique	332
Restauration d'une sauvegarde texte	333
Restauration d'une sauvegarde logique au format binaire	333
Opérations post-import	337
11.5. Sauvegarde physique des fichiers à chaud, en continu	338
Avantages et inconvénients	338
Archivage des journaux de transactions	339
Sauvegarde des fichiers	342
Restauration d'une sauvegarde PITR	344
11.6. Gestionnaires de sauvegardes	347
11.7. Remarques sur les fichiers de configuration	348
11.8. Considérations sur les sauvegardes	348
12. Réplication	349
12.1. Terminologie et vue d'ensemble	349
Cluster de réplication	349
Réplication physique	349
Réplication logique	350
Log Shipping et Streaming Replication	350
Warm Standby et Hot Standby	351
Secondaire asynchrone/synchrone	351
Réplication en cascade	353
12.2. Préparation du serveur primaire	355
Répertoire d'archivage	356
Configuration du serveur primaire	356
12.3. Mise en place d'un serveur secondaire	357
Copie des fichiers	358
Configuration	359
Configuration des différents types de serveurs secondaires	361
12.4. Utilisation d'un serveur secondaire	365
Type de requêtes	365
Durée d'exécution	365
Retard d'un serveur secondaire	366
Nettoyage des journaux archivés	367
Configuration d'un serveur secondaire	368
Bascule	368
Supervision	370
12.5. Réplication logique	371

Particularités des publications	374
Particularités des instances	374
Limitations	375
12.6. Autres solutions	375
13. Statistiques d'activité	377
13.1. Fonctionnement	377
Processus	377
Mémoire	378
Fichiers	378
Aperçu du fonctionnement global	378
Âge des statistiques	380
13.2. Statistiques et calculs intéressants	381
Sessions	381
Processus	386
Bases de données	387
Tables	390
Index	392
Séquences	393
Routines	393
Requêtes	394
13.3. Historisation et alertes	395
Analyse en direct : pgstat	395
Sondes Nagios	397
Analyse graphique en direct des requêtes : PoWA	398
Autres outils	398
14. Maintenance	399
14.1. Opérations de maintenance	399
Lutter contre la fragmentation des tables	399
Mettre à jour les statistiques sur les données	408
Lutter contre la fragmentation des index	409
Réorganiser les données d'une table	410
14.2. Fréquence et automatisation	411
Automatisation de <i>VACUUM</i> et <i>ANALYZE</i>	411
Automatisation de <i>REINDEX</i>	412
Automatisation de <i>CLUSTER</i>	413
15. Sécurité	414
15.1. Au niveau du système d'exploitation	414
Utilisateur postgres	414

Accès aux fichiers	415
Accès aux données	416
Espionnage de la communication	418
Pare-feu	419
Injection SQL	419
SELinux	419
15.2. Au niveau de la base	420
Rôle	420
Mot de passe	422
Droits par défaut d'un rôle	423
Attributs	424
Droits sur la définition des objets	425
Droits sur l'interaction avec les objets	426
Droits sur des objets à créer	427
Droits sur les lignes d'une table	427
Options de sécurité des objets	428
Suppression d'un rôle	428
Labels de sécurité	429
Lexique	430
Liste des illustrations	436
Index	439