

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	8
Répertoire <i>base</i>	8
Répertoire <i>global</i>	21
2.3. Tablespaces	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	32
Répertoire des verrous partagés	33
Répertoire des données 2PC	33
2.5. Statistiques d'activité	34
2.6. Traces	35
2.7. Sous-systèmes	36
Système de notifications	36
Snapshots de base	37
SSI	38
Mémoire partagée dynamique	38
Réplication logique	39
Slots de réplication	39
2.8. Autres fichiers	39
2.9. Répartition du stockage	42
3. Contenu physique des fichiers	43
3.1. Structure générale d'un bloc d'une relation	43
3.2. Tables	49
Fichier HEAP	49

Fichier TOAST	54
3.3. Index	56
Index B-tree	56
Index Hash	64
Index GIN	68
Index GiST	72
Index SP-GiST	73
Index BRIN	74
Outil amcheck	78
3.4. Journaux de transactions	79
4. Architecture des processus	83
4.1. Démarrage et gestion des sous-processus	85
postmaster	88
startup	96
4.2. Écriture dans les fichiers de données	103
writer	104
checkpointer	106
4.3. Écriture dans les journaux de transactions : wal writer.....	116
4.4. Gestion des statistiques d'activité : stats collector.....	118
4.5. Gestion des traces : logger.....	122
4.6. Gestion automatique des <i>VACUUM</i> et <i>ANALYZE</i>	125
autovacuum launcher	126
autovacuum worker	128
4.7. Archivage des journaux de transactions : archiver.....	132
4.8. Gestion de la réplication	134
wal sender	135
wal receiver	136
4.9. Gestion des processus d'arrière-plan : background worker.....	140
4.10. Gestion des connexions : postgres.....	140
4.11. Liens entre répertoires et processus	143
5. Architecture mémoire	144
5.1. Mémoire partagée	145
Vue générale	145
Cache disque des relations	146
Cache disque des journaux de transactions et mémoire partagée du système transactionnel	151
Verrous	152
Sessions	155

Divers	157
Implémentation au niveau système	157
5.2. Mémoire par processus	158
Mémoire de travail	158
Mémoire pour les opérations de maintenance	160
Mémoire cache pour les objets temporaires	161
5.3. Manque de mémoire et traces	161
5.4. Liens entre répertoires, processus et mémoire	163
6. Protocole de communication	165
6.1. Protocole standard	165
Connexion	166
Exécution de requêtes	169
Cas particuliers	173
Déconnexion	174
6.2. Protocole de réplication	174
Messages pour la réplication	175
Messages pour la sauvegarde des fichiers	176
6.3. Outil d'étude	178
7. Gestion des connexions	182
7.1. Aperçu du processus de connexion	182
7.2. Établissement d'une connexion	184
Via un socket de domaine Unix	184
Via un socket réseau	185
Gestion des processus	187
7.3. Authentification	188
7.4. Impact des connexions	190
8. Gestion des transactions	192
8.1. Aspects théoriques	192
Définition d'une transaction	192
Identifiants de transaction	193
Transaction implicite et transaction explicite	195
Propriétés ACID	195
Niveaux d'isolation	199
Transaction imbriquée et transaction autonome	201
Savepoint	201
Durée de vie d'une transaction	202
Two-Phase Commit (2PC)	203
Cas particulier des procédures stockées	203

8.2. Implémentation interne	206
MVCC	206
Vie et mort d'une ligne	208
Horodatage d'une ligne	210
Emplacement physique d'un enregistrement	210
Limite des identifiants de transaction	211
Commit log et hint bits	213
Accès concurrents et verrous	213
8.3. Contournement des inconvénients	216
Maintenance	216
Optimisations	218
9. Gestion des objets	220
9.1. Bases	220
9.2. Tablespaces	223
9.3. Schémas	225
9.4. Tables	227
9.5. Index	232
9.6. Vues	235
9.7. Séquences	236
9.8. Langages de procédures stockées	236
9.9. Procédures stockées	237
9.10. Triggers	240
9.11. Types	241
9.12. Opérateurs	242
9.13. Recherche plein texte	243
9.14. SQL/MED	244
9.15. Méthodes d'accès	245
9.16. Extensions	245
9.17. Rôles	246
10. Planification des requêtes	248
10.1. Introduction à l'optimiseur de requêtes	248
Fonctionnement de l'optimiseur	248
Plans d'exécution	250
10.2. Nœuds d'exécution d'un plan	251
Accès aux données	251
Jointures	266
Agrégats	269
Autres	273

10.3. Utilisation d'EXPLAIN	277
Informations de base	277
Options	279
10.4. Statistiques et coûts	282
Mise à jour des statistiques	282
Statistiques	283
Statistiques étendues	286
Paramétrage	286
Estimation de coûts	288
10.5. Problèmes et optimisations	291
Index inutilisés	291
Index manquants	292
Parcours séquentiels synchronisés	293
Suppression de jointures inutiles	293
Suppression d'éléments inutiles dans une clause GROUP BY	294
10.6. Outils	295
Affichage graphique d'un plan d'exécution	295
Aide à la recherche du nœud lent	296
Trace automatique des plans d'exécution	298
10.7. Exécuteur	300
11. Sauvegarde et restauration	303
11.1. Sauvegarde physique des fichiers à froid	304
11.2. Export logique	307
Fonctionnement interne	308
Format de sauvegarde	310
Sauvegarde complète	313
Sauvegarde partielle	314
Cohérence de la sauvegarde et verrous	315
Sauvegarde parallélisée	315
Sauvegarde des Large Objects	317
Compatibilité entre versions majeures	317
11.3. Export global	318
Format de sauvegarde	318
Sauvegarde complète	319
Sauvegarde partielle	319
Cohérence de la sauvegarde	319
Exemple d'utilisation de pg_dump et pg_dumpall	320
11.4. Import logique	320
Restauration d'une sauvegarde texte	321

Restauration d'une sauvegarde logique au format binaire	321
Opérations post-import	325
11.5. Sauvegarde physique des fichiers à chaud, en continu	326
Avantages et inconvénients	326
Archivage des journaux de transactions	327
Sauvegarde des fichiers	330
Restauration d'une sauvegarde PITR	332
11.6. Gestionnaires de sauvegardes	335
11.7. Remarques sur les fichiers de configuration	335
11.8. Considérations sur les sauvegardes	336
12. Réplication	337
12.1. Terminologie et vue d'ensemble	337
Cluster de réplication	337
Réplication physique	337
Réplication logique	338
Log Shipping et Streaming Replication	338
Warm Standby et Hot Standby	339
Secondaire asynchrone/synchrone	339
Réplication en cascade	341
12.2. Préparation du serveur primaire	343
Répertoire d'archivage	344
Configuration du serveur primaire	344
12.3. Mise en place d'un serveur secondaire	345
Copie des fichiers	346
Configuration	347
Configuration des différents types de serveurs secondaires	348
12.4. Utilisation d'un serveur secondaire	352
Type de requêtes	352
Durée d'exécution	353
Retard d'un serveur secondaire	354
Nettoyage des journaux archivés	355
Configuration d'un serveur secondaire	356
Bascule	356
Supervision	358
12.5. Réplication logique	358
Particularités des publications	361
Particularités des instances	362
Limitations	362
12.6. Autres solutions	363

13. Statistiques d'activité	364
13.1. Fonctionnement	364
Processus	364
Mémoire	365
Fichiers	365
Aperçu du fonctionnement global	365
Âge des statistiques	367
13.2. Statistiques et calculs intéressants	368
Sessions	368
Processus	373
Bases de données	374
Tables	378
Index	379
Séquences	380
Procédures stockées	381
Requêtes	382
13.3. Historisation et alertes	383
Analyse en direct : pgstat	383
Sondes Nagios	384
Analyse graphique en direct des requêtes : PoWA	386
Autres outils	387
Pour la suite	387
14. Maintenance	388
14.1. Opérations de maintenance	388
Lutter contre la fragmentation des tables	388
Mettre à jour les statistiques sur les données	396
Lutter contre la fragmentation des index	397
Réorganiser les données d'une table	398
14.2. Fréquence et automatisation	399
Automatisation de <i>VACUUM</i> et <i>ANALYZE</i>	399
Automatisation de <i>REINDEX</i>	400
Automatisation de <i>CLUSTER</i>	401
15. Sécurité	402
15.1. Au niveau du système d'exploitation	402
Utilisateur postgres	402
Accès aux fichiers	403
Accès aux données	404
Espionnage de la communication	406

Pare-feu	407
Injection SQL	407
SELinux	407
15.2. Au niveau de la base	408
Rôle	408
Mot de passe	410
Droits par défaut d'un rôle	410
Attributs	411
Droits sur la définition des objets	413
Droits sur l'interaction avec les objets	413
Droits sur des objets à créer	414
Droits sur les lignes d'une table	415
Options de sécurité des objets	415
Suppression d'un rôle	416
Labels de sécurité	417
Lexique	418
Liste des illustrations	424
Index	427