

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>	20
2.3. Tablespaces	22
2.4. Transactions	24
Répertoire des journaux de transactions	24
Répertoire d'état des transactions	29
Répertoire d'horodatages des transactions	30
Répertoire des points de retour	31
Répertoire des verrous partagés	31
Répertoire des données 2PC	32
2.5. Statistiques d'activité	32
2.6. Traces	34
2.7. Sous-systèmes	34
Système de notifications	34
Snapshots de base	35
SSI	36
Mémoire partagée dynamique	36
Réplication logique	37
Slots de réplication	37
2.8. Autres fichiers	37
2.9. Répartition du stockage	40
3. Contenu physique des fichiers	41
3.1. Structure générale d'un bloc d'une relation	41
3.2. Tables	46
Fichier HEAP	46

Fichier TOAST	51
3.3. Index	52
Index B-tree	53
Index Hash	60
Index GIN	62
Index GiST	66
Index SP-GiST	66
Index BRIN	67
3.4. Journaux de transactions	71
4. Architecture des processus	75
4.1. Démarrage et gestion des sous-processus	77
postmaster	80
startup	88
4.2. Écriture dans les fichiers de données	94
writer	95
checkpointer	97
4.3. Écriture dans les journaux de transactions : wal writer.....	107
4.4. Gestion des statistiques d'activité : stats collector.....	109
4.5. Gestion des traces : logger.....	113
4.6. Gestion automatique des <i>VACUUM</i> et <i>ANALYZE</i>	116
autovacuum lancer	117
autovacuum worker	119
4.7. Archivage des journaux de transactions : archiver.....	123
4.8. Gestion de la réplication	125
wal sender	126
wal receiver	127
4.9. Gestion des processus d'arrière-plan : background worker.....	130
4.10. Gestion des connexions : postgres.....	131
4.11. Liens entre répertoires et processus	133
5. Architecture mémoire	134
5.1. Mémoire partagée	135
Vue générale	135
Cache disque des relations	136
Cache disque des transactions et des multi-transactions	139
Verrous	140
Sessions	143
Divers	145
Implémentation au niveau système	145

5.2. Mémoire par processus	145
Mémoire de travail	145
Mémoire pour les opérations de maintenance	147
Mémoire cache pour les objets temporaires	148
5.3. Liens entre répertoires, processus et mémoire	148
6. Protocole de communication	150
6.1. Protocole standard	150
Connexion	151
Exécution de requêtes	154
Cas particuliers	158
Déconnexion	159
6.2. Protocole de réplication	159
Messages pour la réplication	160
Messages pour la sauvegarde des fichiers	161
6.3. Outil d'étude	163
7. Gestion des connexions	167
7.1. Aperçu du processus de connexion	167
7.2. Établissement d'une connexion	168
Via un socket de domaine Unix	169
Via un socket réseau	170
Gestion des processus	172
7.3. Authentification	173
7.4. Impact des connexions	175
8. Gestion des transactions	177
8.1. Aspects théoriques	177
Définition d'une transaction	177
Identifiants de transaction	178
Transaction implicite et transaction explicite	180
Propriétés ACID	180
Niveaux d'isolation	184
Transaction imbriquée et transaction autonome	185
Savepoint	186
Two-Phase Commit (2PC)	187
Cas particulier des procédures stockées	188
8.2. Implémentation interne	190
MVCC	190
Vie et mort d'une ligne	192
Horodatage d'une ligne	194

Emplacement physique d'un enregistrement	194
Limite des identifiants de transaction	195
Commit log et hint bits	197
Accès concurrents et verrous	197
8.3. Contournement des inconvénients	198
Maintenance	198
Optimisations	199
9. Gestion des objets	201
9.1. Bases	201
9.2. Tablespaces	204
9.3. Schémas	205
9.4. Tables	207
9.5. Index	211
9.6. Vues	213
9.7. Séquences	215
9.8. Langages de procédures stockées	215
9.9. Procédures stockées	216
9.10. Triggers	218
9.11. Types	219
9.12. Opérateurs	221
9.13. Recherche plein texte	221
9.14. SQL/MED	222
9.15. Extensions	223
9.16. Rôles	223
10. Planification des requêtes	225
10.1. Introduction à l'optimiseur de requêtes	225
Fonctionnement de l'optimiseur	225
Plans d'exécution	226
10.2. Nœuds d'exécution d'un plan	227
Accès aux données	227
Jointures	236
Agrégats	240
Autres	243
10.3. Utilisation d'EXPLAIN	247
Informations de base	247
Options	249
10.4. Statistiques et coûts	252
Mise à jour des statistiques	252

Statistiques	253
Paramétrage	255
Estimation de coûts	257
10.5. Problèmes et optimisations	259
Index inutilisés	259
Index manquants	260
Parcours séquentiels synchronisés	261
Suppression de jointures inutiles	261
10.6. Outils	262
Affichage graphique d'un plan d'exécution	262
Aide à la recherche du nœud lent	263
Trace automatique des plans d'exécution	265
11. Sauvegarde et restauration	267
11.1. Sauvegarde physique des fichiers à froid	268
11.2. Export logique	271
Fonctionnement interne	272
Format de sauvegarde	274
Sauvegarde complète	277
Sauvegarde partielle	278
Cohérence de la sauvegarde et verrous	279
Sauvegarde parallélisée	279
Sauvegarde des Large Objects	281
Compatibilité entre versions majeures	281
11.3. Export global	282
Format de sauvegarde	282
Sauvegarde complète	282
Sauvegarde partielle	283
Cohérence de la sauvegarde	283
Exemple d'utilisation de pg_dump et pg_dumpall	283
11.4. Import logique	284
Restauration d'une sauvegarde texte	284
Restauration d'une sauvegarde logique au format binaire	285
11.5. Sauvegarde physique des fichiers à chaud, en continu	288
Avantages et inconvénients	289
Archivage des journaux de transactions	290
Sauvegarde des fichiers	293
Restauration d'une sauvegarde PITR	294
11.6. Gestionnaires de sauvegardes	297
11.7. Remarques sur les fichiers de configuration	297

11.8. Considérations sur les sauvegardes	298
12. Réplication	299
12.1. Terminologie et vue d'ensemble	299
Cluster de réplication	299
Réplication physique	299
Log Shipping et Streaming Replication	300
Warm Standby et Hot Standby	300
Secondaire asynchrone/synchrone	301
Réplication en cascade	303
12.2. Préparation du serveur primaire	305
Répertoire d'archivage	305
Configuration du serveur primaire	306
12.3. Mise en place d'un serveur secondaire	307
Copie des fichiers	307
Configuration	308
Configuration des différents types de serveurs secondaires	310
12.4. Utilisation d'un serveur secondaire	313
Type de requêtes	313
Durée d'exécution	313
Retard d'un serveur secondaire	314
Nettoyage des journaux archivés	315
Configuration d'un serveur secondaire	316
Bascule	316
Supervision	318
12.5. Autres solutions	318
13. Statistiques d'activité	319
13.1. Fonctionnement	319
Processus	319
Mémoire	320
Fichiers	320
Aperçu du fonctionnement global	320
Âge des statistiques	322
13.2. Statistiques et calculs intéressants	323
Sessions	323
Processus	327
Bases de données	329
Tables	331
Index	334

Séquences	335
Procédures stockées	335
Requêtes	336
13.3. Historisation et alertes	337
Analyse en direct pgstat	337
Sondes Nagios	339
Analyse graphique en direct des requêtes : PoWA	340
Autres outils	341
Pour la suite	341
14. Maintenance	342
14.1. Opérations de maintenance	342
Lutter contre la fragmentation des tables	342
Mettre à jour les statistiques sur les données	350
Lutter contre la fragmentation des index	351
Réorganiser les données d'une table	352
14.2. Fréquence et automatisation	353
Automatisation de <i>VACUUM</i> et <i>ANALYZE</i>	353
Automatisation de <i>REINDEX</i>	354
Automatisation de <i>CLUSTER</i>	354
15. Sécurité	356
15.1. Au niveau du système d'exploitation	356
Utilisateur postgres	356
Accès aux fichiers	357
Accès aux données	358
Espionnage de la communication	360
Pare-feu	361
Injection SQL	361
SELinux	361
15.2. Au niveau de la base	362
Rôle	362
Mot de passe	363
Droits par défaut d'un rôle	363
Attributs	364
Droits sur la définition des objets	366
Droits sur l'interaction avec les objets	366
Droits sur des objets à créer	367
Droits sur les lignes d'une table	368
Options de sécurité des objets	368

Suppression d'un rôle	369
Labels de sécurité	370
Lexique	371
Liste des illustrations	377
Index	380