

Table des matières

À propos des auteurs	xiii
Avant-propos	xiv
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>	23
2.3. Tablespaces	27
2.4. Transactions	30
Répertoire des journaux de transactions	30
Répertoire d'état des transactions	36
Répertoire d'horodatages des transactions	37
Répertoire des points de retour	37
Répertoire des verrous partagés	38
Répertoire des données 2PC	38
2.5. Statistiques d'activité	39
2.6. Traces	40
2.7. Sous-systèmes	41
Système de notifications	41
Snapshots de base	41
SSI	43
Mémoire partagée dynamique	43
Réplication logique	43
Slots de réplication	44
2.8. Autres fichiers	44
2.9. Répartition du stockage	48
3. Contenu physique des fichiers	49
3.1. Méthode d'accès à une relation	49
3.2. Structure générale d'un bloc d'une relation	50
3.3. Tables	56

Fichier HEAP	56
Fichier TOAST	61
3.4. Index	64
Index B-tree	65
Index Hash	74
Index GIN	79
Index GiST	83
Index SP-GiST	84
Index BRIN	85
Extension amcheck	89
3.5. Journaux de transactions	90
4. Architecture des processus	94
4.1. Démarrage et gestion des sous-processus	96
postmaster	99
startup	108
4.2. Écriture dans les fichiers de données	116
writer	116
checkpointer	119
4.3. Écriture dans les journaux de transactions : wal writer.....	129
4.4. Gestion des statistiques d'activité : stats collector.....	131
4.5. Gestion des traces : logger.....	136
4.6. Gestion automatique des <i>VACUUM</i> et <i>ANALYZE</i>	140
autovacuum launcher	140
autovacuum worker	142
4.7. Archivage des journaux de transactions : archiver.....	147
4.8. Gestion de la réplication physique	149
wal sender	150
wal receiver	151
4.9. Gestion de la réplication logique	155
logical replication launcher	155
logical replication worker	155
4.10. Gestion des processus d'arrière-plan : background worker.....	156
4.11. Gestion des connexions : postgres.....	156
4.12. Liens entre répertoires et processus	159
5. Architecture mémoire	160
5.1. Mémoire partagée	161
Vue générale	161
Cache disque des relations	163

Cache disque des journaux de transactions et mémoire partagée du système transactionnel	168
Verrous	169
Sessions	172
Divers	174
Implémentation au niveau système	175
5.2. Mémoire par processus	176
Mémoire de travail	176
Mémoire pour le décodage logique	177
Mémoire pour les opérations de maintenance	177
Mémoire cache pour les objets temporaires	178
Quantifier la mémoire utilisée par un processus	179
5.3. Manque de mémoire et traces	180
5.4. Liens entre répertoires, processus et mémoire	182
6. Protocole de communication	184
6.1. Protocole standard	184
Connexion	185
Exécution de requêtes	189
Cas particuliers	193
Déconnexion	195
6.2. Protocole de réplication	195
Messages pour la réplication	195
Messages pour la sauvegarde des fichiers	197
6.3. Outil d'étude	199
7. Gestion des connexions	203
7.1. Aperçu du processus de connexion	203
7.2. Établissement d'une connexion	205
Via un socket de domaine Unix	205
Via un socket réseau	206
Gestion des processus	208
7.3. Authentification	209
7.4. Impact des connexions	211
8. Gestion des transactions	213
8.1. Aspects théoriques	213
Définition d'une transaction	213
Identifiants de transaction	214
Transaction implicite et transaction explicite	216
Propriétés ACID	217

Niveaux d'isolation	220
Snapshot	222
Transaction imbriquée et transaction autonome	224
Savepoint	225
Durée de vie d'une transaction	226
Two-Phase Commit (2PC)	226
Cas particulier des routines	227
8.2. Implémentation interne	229
MVCC	229
Vie et mort d'une ligne	231
Horodatage d'une ligne	234
Emplacement physique d'un enregistrement	234
Conséquences pour les index	235
Limite des identifiants de transaction	237
Commit log et hint bits	240
Accès concurrents et verrous	241
8.3. Contournement des inconvénients	244
Maintenance	244
Optimisations	247
9. Gestion des objets	252
9.1. Bases	252
9.2. Tablespace	255
9.3. Schémas	257
9.4. Tables	259
9.5. Index	265
9.6. Vues	268
9.7. Séquences	269
9.8. Langages de routines	270
9.9. Routines	271
9.10. Triggers	274
9.11. Types	275
9.12. Opérateurs	276
9.13. Recherche plein texte	277
9.14. SQL/MED	278
9.15. Méthodes d'accès	279
9.16. Extensions	279
9.17. Rôles	280
10. Planification des requêtes	282

10.1. Introduction à l'optimiseur de requêtes	282
Fonctionnement de l'optimiseur	282
Plans d'exécution	284
10.2. Nœuds d'exécution d'un plan	285
Accès aux données	285
Jointures	302
Agrégats	305
Autres	309
10.3. Utilisation d'EXPLAIN	318
Informations de base	318
Options	320
10.4. Statistiques et coûts	324
Mise à jour des statistiques	324
Statistiques	325
Statistiques étendues	327
Paramétrage	328
Estimation de coûts	330
10.5. Problèmes et optimisations	334
Index inutilisés	334
Index manquants	335
Parcours séquentiels synchronisés	336
Suppression de jointures inutiles	336
Suppression d'éléments inutiles dans une clause GROUP BY	337
10.6. Outils	338
Affichage graphique d'un plan d'exécution	338
Aide à la recherche du nœud lent	340
Trace automatique des plans d'exécution	343
10.7. Exécuteur	345
11. Sauvegarde et restauration	348
11.1. Sauvegarde physique des fichiers à froid	349
11.2. Export logique	352
Fonctionnement interne	353
Format de sauvegarde	354
Sauvegarde complète	358
Sauvegarde partielle	359
Cohérence de la sauvegarde et verrous	360
Sauvegarde parallélisée	360
Sauvegarde des Large Objects	362
Compatibilité entre versions majeures	362

11.3. Export global	363
Format de sauvegarde	363
Sauvegarde complète	364
Sauvegarde partielle	364
Cohérence de la sauvegarde	364
Exemple d'utilisation de <code>pg_dump</code> et <code>pg_dumpall</code>	365
11.4. Import logique	365
Restauration d'une sauvegarde texte	366
Restauration d'une sauvegarde logique au format binaire	366
Opérations post-import	371
11.5. Sauvegarde physique des fichiers à chaud, en continu	371
Avantages et inconvénients	372
Archivage des journaux de transactions	373
Sauvegarde des fichiers	376
Restauration d'une sauvegarde PITR	378
11.6. Gestionnaires de sauvegardes	381
11.7. Remarques sur les fichiers de configuration	381
11.8. Considérations sur les sauvegardes	382
12. Réplication	383
12.1. Terminologie et vue d'ensemble	383
Cluster de réplication	383
Réplication physique	383
Réplication logique	384
Log Shipping et Streaming Replication	384
Warm Standby et Hot Standby	385
Secondaire asynchrone/synchrone	385
Réplication en cascade	387
12.2. Préparation du serveur primaire	389
Répertoire d'archivage	390
Configuration du serveur primaire	390
12.3. Mise en place d'un serveur secondaire	391
Copie des fichiers	392
Configuration	393
Configuration des différents types de serveurs secondaires	395
12.4. Utilisation d'un serveur secondaire	399
Type de serveurs	399
Type de requêtes	399
Durée d'exécution	399
Retard d'un serveur secondaire	400

Nettoyage des journaux archivés	402
Configuration d'un serveur secondaire	402
Bascule	403
Supervision	405
12.5. Réplication logique	405
Particularités des publications	408
Gestion des transactions	409
Particularités des instances	409
Limitations	409
12.6. Autres solutions	410
13. Statistiques d'activité	411
13.1. Fonctionnement	411
Processus	411
Mémoire	412
Fichiers	412
Aperçu du fonctionnement global	412
Âge des statistiques	414
13.2. Statistiques et calculs intéressants	415
Sessions	415
Processus	420
Bases de données	421
Tables	424
Index	426
Séquences	427
Routines	427
Requêtes	428
Journaux de transactions	429
Slots de réplication	430
13.3. Historisation et alertes	430
Analyse en direct : pgstat	430
Sondes Nagios	432
Analyse graphique en direct des requêtes : PoWA.....	432
Autres outils	433
14. Traces	434
14.1. Configuration	434
Format des traces	434
Destination	434
Préfixe, fuseau horaire et langue utilisés	437

Niveau des traces	438
Activités tracées	439
Requêtes	442
14.2. Exploitation	443
Outils d'analyse en direct	443
Outils de rétro-analyse	444
15. Maintenance	445
15.1. Opérations de maintenance	445
Lutter contre la fragmentation des tables	445
Mettre à jour les statistiques sur les données	455
Lutter contre la fragmentation des index	456
Réorganiser les données d'une table	457
15.2. Fréquence et automatisation	458
Automatisation de <i>VACUUM</i> et <i>ANALYZE</i>	458
Automatisation de <i>REINDEX</i>	459
Automatisation de <i>CLUSTER</i>	460
16. Sécurité	461
16.1. Au niveau du système d'exploitation	461
Utilisateur postgres	461
Accès aux fichiers	462
Accès aux données	463
Espionnage de la communication	465
Pare-feu	466
Injection SQL	466
SELinux	467
16.2. Au niveau de la base	467
Rôle	467
Mot de passe	469
Droits par défaut d'un rôle	470
Attributs	471
Corruption du catalogue système	473
Droits sur la définition des objets	474
Droits sur l'interaction avec les objets	474
Droits sur les objets à créer	475
Droits sur les lignes d'une table	476
Options de sécurité des objets	476
Suppression d'un rôle	477
Labels de sécurité	478

Lexique	479
Index	485