

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>	22
2.3. Tablespaces	26
2.4. Transactions	29
Répertoire des journaux de transactions	29
Répertoire d'état des transactions	35
Répertoire d'horodatages des transactions	36
Répertoire des points de retour	36
Répertoire des verrous partagés	37
Répertoire des données 2PC	37
2.5. Statistiques d'activité	38
Avant la version 15	38
À partir de la version 15	39
2.6. Traces	39
2.7. Sous-systèmes	40
Système de notifications	40
Snapshots de base	41
SSI	42
Mémoire partagée dynamique	42
Réplication logique	43
Slots de réplication	43
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	55
Fichier HEAP	55
Fichier TOAST	60
3.4. Index	63
Index B-tree	64
Index Hash	73
Index GIN	78
Index GiST	82
Index SP-GiST	83
Index BRIN	84
Extension amcheck	88
3.5. Journaux de transactions	89
4. Architecture des processus	93
4.1. Démarrage et gestion des sous-processus	95
postmaster	98
startup	106
4.2. Lecture dans les fichiers de données	113
Signaux	114
Boucle principale	114
4.3. Écriture dans les fichiers de données	114
writer	115
checkpointer	117
Pour résumer	127
4.4. Écriture dans les journaux de transactions : wal writer.....	127
Vie et mort du processus	128
Signaux	128
Boucle principale	128
4.5. Gestion des statistiques d'activité : stats collector.....	129
4.6. Gestion des traces : logger.....	132
Fichiers	132
Signaux	133
Configuration	133
Boucle principale	134
4.7. Gestion automatique des VACUUM et ANALYZE	135
autovacuum launcher	136
autovacuum worker	138
4.8. Archivage des journaux de transactions : archiver.....	142
Signaux	142

Boucle principale	143
Boucle de traitement des journaux	143
Supervision et traces	145
4.9. Résumé des journaux de transactions : wal summarize	145
Signaux	145
Boucle principale	145
Supervision et traces	145
4.10. Gestion de la réplication physique	146
wal sender	146
wal receiver	148
4.11. Gestion de la réplication logique	151
logical replication launcher	151
logical replication worker	152
4.12. Gestion des processus d'arrière-plan : background worker	152
4.13. Gestion des connexions : postgres	153
Signaux	153
Initialisation	153
Boucle principale	154
4.14. Liens entre répertoires et processus	156
5. Architecture mémoire	157
5.1. Mémoire partagée	158
Vue générale	158
Cache disque des relations	160
Cache disque des journaux de transactions et mémoire partagée du système transactionnel	165
Verrous	166
Sessions	169
Divers	172
Implémentation au niveau système	172
5.2. Mémoire par processus	173
Mémoire de travail	173
Mémoire pour le décodage logique	174
Mémoire pour les opérations de maintenance	175
Mémoire cache pour les objets temporaires	176
Quantifier la mémoire utilisée par un processus	176
5.3. Manque de mémoire et traces	178
5.4. Liens entre répertoires, processus et mémoire	179
6. Protocole de communication	181

6.1. Protocole standard	181
Connexion	182
Exécution de requêtes	186
Cas particuliers	190
Déconnexion	192
6.2. Protocole de réplication	192
Messages pour la réplication	192
Messages pour la sauvegarde des fichiers	194
6.3. Outil d'étude	196
7. Gestion des connexions	200
7.1. Aperçu du processus de connexion	200
7.2. Établissement d'une connexion	202
Via un socket de domaine Unix	202
Via un socket réseau	203
Gestion des processus	205
7.3. Authentification	206
7.4. Impact des connexions	208
8. Gestion des transactions	211
8.1. Aspects théoriques	211
Définition d'une transaction	211
Identifiants de transaction	212
Transaction implicite et transaction explicite	214
Propriétés ACID	215
Niveaux d'isolation	218
Snapshot	220
Transaction imbriquée et transaction autonome	222
Savepoint	223
Durée de vie d'une transaction	224
Two-Phase Commit (2PC)	224
Cas particulier des routines	225
8.2. Implémentation interne	227
MVCC	227
Vie et mort d'une ligne	229
Horodatage d'une ligne	231
Emplacement physique d'un enregistrement	232
Conséquences pour les index	233
Limite des identifiants de transaction	235
Commit log et hint bits	238

Accès concurrents et verrous	239
8.3. Contournement des inconvénients	242
Maintenance	242
Optimisations	245
9. Gestion des objets	250
9.1. Bases	250
9.2. Tablespaces	253
9.3. Schémas	255
9.4. Tables	257
9.5. Index	262
9.6. Vues	265
9.7. Séquences	266
9.8. Langages de routines	267
9.9. Routines	268
9.10. Triggers	271
9.11. Types	272
9.12. Opérateurs	274
9.13. Recherche plein texte	274
9.14. SQL/MED	275
9.15. Méthodes d'accès	276
9.16. Extensions	276
9.17. Rôles	277
10. Planification des requêtes	279
10.1. Introduction à l'optimiseur de requêtes	279
Fonctionnement de l'optimiseur	279
Plans d'exécution	281
10.2. Nœuds d'exécution d'un plan	282
Accès aux données	282
Jointures	299
Agrégats	302
Autres	306
10.3. Utilisation d'EXPLAIN	315
Informations de base	316
Options	317
10.4. Statistiques et coûts	322
Mise à jour des statistiques	323
Statistiques	323
Statistiques étendues	326

Paramétrage	327
Estimation de coûts	329
10.5. Problèmes et optimisations	333
Index inutilisés	333
Index manquants	334
Parcours séquentiels synchronisés	335
Suppression de jointures inutiles	335
Suppression d'éléments inutiles dans une clause GROUP BY	336
10.6. Outils	337
Affichage graphique d'un plan d'exécution	337
Aide à la recherche du nœud lent	339
Trace automatique des plans d'exécution	341
10.7. Exécuteur	344
11. Sauvegarde et restauration	347
11.1. Sauvegarde physique des fichiers à froid	348
11.2. Export logique	351
Fonctionnement interne	352
Format de sauvegarde	353
Sauvegarde complète	357
Sauvegarde partielle	358
Cohérence de la sauvegarde et verrous	359
Sauvegarde parallélisée	360
Sauvegarde des Large Objects	361
Compatibilité entre versions majeures	362
11.3. Export global	362
Format de sauvegarde	362
Sauvegarde complète	363
Sauvegarde partielle	363
Cohérence de la sauvegarde	364
Exemple d'utilisation de pg_dump et pg_dumpall	364
11.4. Import logique	364
Restauration d'une sauvegarde texte	365
Restauration d'une sauvegarde logique au format binaire	365
Opérations post-import	370
11.5. Sauvegarde physique des fichiers à chaud, en continu	370
Avantages et inconvénients	371
Archivage des journaux de transactions	372
Sauvegarde des fichiers	375
Restauration d'une sauvegarde PITR	377

11.6. Gestionnaires de sauvegardes	381
11.7. Remarques sur les fichiers de configuration	381
11.8. Considérations sur les sauvegardes	381
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	388
12.2. Préparation du serveur primaire	390
Répertoire d'archivage	390
Configuration du serveur primaire	391
12.3. Mise en place d'un serveur secondaire	392
Copie des fichiers	392
Configuration	393
Configuration des différents types de serveurs secondaires	394
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
13. Statistiques d'activité	411
13.1. Fonctionnement pré-v15	411
Processus	411
Mémoire	412
Fichiers	412

Aperçu du fonctionnement global	412
13.2. Fonctionnement à partir de la v15	414
13.3. Âge des statistiques	415
13.4. Statistiques et calculs intéressants	416
Sessions	416
Processus	421
Bases de données	423
Tables	426
Index	427
Séquences	428
Routines	429
Requêtes	430
Journaux de transactions	431
Slots de réplication	431
13.5. Historisation et alertes	431
Analyse en direct : pgstat	431
Sondes Nagios	433
Analyse graphique en direct des requêtes : PoWA	434
14. Traces	435
14.1. Configuration	435
Format des traces	435
Destination	435
Préfixe, fuseau horaire et langue utilisés	439
Niveau des traces	440
Activités tracées	440
Requêtes	444
14.2. Exploitation	445
Outils d'analyse en direct	445
Outils de rétro-analyse	446
15. Maintenance	447
15.1. Opérations de maintenance	447
Lutter contre la fragmentation des tables	447
Mettre à jour les statistiques sur les données	458
Lutter contre la fragmentation des index	459
Réorganiser les données d'une table	460
15.2. Fréquence et automatisation	461
Automatisation de VACUUM et ANALYZE	461
Automatisation de REINDEX	462

Automatisation de <i>CLUSTER</i>	462
16. Sécurité	464
16.1. Au niveau du système d'exploitation	464
Utilisateur postgres	464
Accès aux fichiers	465
Accès aux données	466
Espionnage de la communication	468
Pare-feu	469
Injection SQL	470
SELinux	470
16.2. Au niveau de la base	470
Rôle	471
Mot de passe	473
Droits par défaut d'un rôle	474
Attributs	475
Corruption du catalogue système	476
Droits sur la définition des objets	477
Droits sur l'interaction avec les objets	478
Droits sur les objets à créer	479
Droits sur les lignes d'une table	479
Options de sécurité des objets	480
Suppression d'un rôle	480
Labels de sécurité	481
Lexique	483
Index	489