



Doctolib

Comment identifier des index corrompus

Flavio Gurgel - Meetup été 2017 - PeopleDoc - Paris

- Fichier de l'index est corrompu en disque. Causes :
 - ✓ fsync ou full_page_writes = off + crash
 - ✓ bug de fsync du système d'exploitation
 - ✓ barrettes de RAM sans ECC et défectueuses

```
flavio=> \d tbl_demo
          Table "public.tbl_demo"
  Column   | Type   | Modifiers
-----+-----+-----
 id        | bigint | not null default nextval('tbl_demo_id_seq'::regclass)
 name_hash | text   |
Indexes:
    "tbl_demo_pkey" PRIMARY KEY, btree (id)

flavio=> EXPLAIN (analyze) SELECT * FROM tbl_demo ORDER BY id;
ERROR:  could not read block 1280 in file "base/16385/16578": read only 0 of 8192 bytes
```

```
flavio=> SELECT relname, relkind FROM pg_class WHERE relfilenode = 16578;
 relname    | relkind
-----+-----
 tbl_demo_pkey | i
(1 row)
```

- Identifier la cause
 - corriger la configuration, mettre à jour tout le système, tester RAM et autre hardware.
- REINDEX
 - Avec un peu de chance, tester avant et basculer vers un serveur stand-by.

```
flavio=> reindex index tbl_demo_pkey ;
REINDEX
flavio=> EXPLAIN (analyze) SELECT * FROM tbl_demo ORDER BY id;
```

QUERY PLAN

```
Index Scan using tbl_demo_pkey on tbl_demo  (cost=0.42..35329.43 rows=1000000 width=41) (actual time=0.040..131.839 rows=1000000 loops=1)
Planning time: 0.228 ms
Execution time: 155.092 ms
(3 rows)
```

- Causé par un bug PostgreSQL jusqu'aux versions :
 - 9.2.19
 - 9.3.15
 - 9.4.10
 - 9.5.5
 - 9.6.1
- CREATE INDEX CONCURRENTLY
- Avec des UPDATE concurrents
- L'index peut rater quelques entrées de la table

- Des requêtes peuvent afficher des résultats différents si elles utilisent l'index corrompu.
- Des clés uniques ou primaires peuvent ne pas être respectées.
- Des clés étrangères peuvent ne pas être respectées.
- Une solution rapide est nécessaire.

- Utiliser une transaction en mode REPEATABLE READ ou SERIALIZABLE
 - Alternativement, la base sans écriture en mode READ COMMITTED
- Forcer un SELECT avec seqscan
- Forcer un autre SELECT avec l'index
- Comparer les résultats

- Fatigué ? Des centaines d'index à tester ? Des tables énormes ? Pas sûr ?
 - **pg_check_index.rb** fait le boulot fastidieux !
 - https://github.com/doctolib/pg_check_indexes
- Écrit en Ruby
 - Ruby 2.4.0 ou supérieur
 - Gem pg
 - Gem optparse
- Peut tester toute la base, seulement une table ou des tables jusqu'à une taille indiquée.
 - Testé sur PostgreSQL 9.4, 9.5 et 9.6

- Teste tous les index de type btree
 - intégrité
 - unicité
 - avec fonctions
 - multiples colonnes
 - partiels (avec clause WHERE)
- Ne teste pas les autres types d'index
- Ne modifie rien
- Ne corrige pas les éventuelles corruptions
- Peut être exécuté sur un serveur stand-by (réplica physique, streaming ou log-shipping)

```
$ ruby pg_check_indexes.rb
*** START OF INDEX/TABLE CONSISTENCY TEST ***

Table: public.tbl_demo - Index: tbl_demo_pkey - Columns: id, comparing data... index and table match
- OK. Index is unique or a PK, testing uniqueness... no duplicate entries - OK.
Table: public.unicorn_timings - Index: unicorn_timings_server_idx - Columns: server, comparing
data... index and table match - OK.
Table: public.unicorn_timings - Index: unicorn_timings_tmst_idx - Columns: tmst, comparing data...
index and table match - OK.

*** END OF INDEX/TABLE CONSISTENCY TEST ***
*** START OF FOREIGN KEYS CONSISTENCY TEST ***

Table: public.tbl_fk_demo - Constraint: tbl_fk_demo_tbl_demo_id_fkey - Columns: tbl_demo_id -
Referenced table: tbl_demo - Columns: id. Checking... no missing entries in referenced table - OK.
*** END OF FOREIGN KEYS CONSISTENCY TEST ***
```

- **Mettre à jour PostgreSQL !!!**
- REINDEX immédiat si possible
 - Pas d'écriture pendant le REINDEX
 - Les requêtes peuvent avoir des résultats différents après
- Rejouer pg_check_index.rb
 - Vérifier spécialement les clés étrangères qui peuvent se corriger avec le REINDEX

■ Alternative

- CREATE (UNIQUE) INDEX CONCURRENTLY ...;
- DROP INDEX ...;
- ALTER INDEX ... RENAME TO ...;

■ Si PK

- CREATE UNIQUE INDEX CONCURRENTLY ...;
- ALTER TABLE ... DROP CONSTRAINT ...;
- ALTER TABLE ... ADD CONSTRAINT ... PRIMARY KEY USING INDEX ...;

Merci !

- www.doctolib.fr

On recrute

- www.doctolib.fr/jobs