

Passerelle - Development #66694

base_adresse: ajouter des index sur unaccent_name et zipcode et tous les champs qui servent aux recherches

28 juin 2022 12:14 - Benjamin Dauvergne

Statut:	Fermé	Début:	28 juin 2022
Priorité:	Normal	Echéance:	
Assigné à:	Benjamin Dauvergne	% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposed:	Oui		
Description			
Pour éviter de charger la base avec des scans linéaires de la liste des 36000 communes de France.			

Révisions associées

Révision ee1baf8b - 15 janvier 2024 16:22 - Benjamin Dauvergne

base_adresse: add indexes on on geographic models names (#66694)

Révision bbf7aabb - 15 janvier 2024 18:34 - Benjamin Dauvergne

Revert base_adresse: add indexes on on geographic models names (#66694)

This reverts commit ee1baf8b50c7b629e082b83482dfc6c181d07de9.

Révision 33b39c52 - 14 février 2024 15:29 - Benjamin Dauvergne

base_adresse: add indexes on on geographic models names (#66694)

Historique

#1 - 28 juin 2022 12:14 - Benjamin Dauvergne

- Assigné à mis à Benjamin Dauvergne

#2 - 28 juin 2022 12:36 - Benjamin Dauvergne

- Fichier 0001-wip.patch ajouté

- Patch proposed changé de Non à Oui

C'est ce qu'il faudrait faire si on était en Django 3.2 pour avoir dex indexs sur les expressions, pour améliorer la rechercher sur les rues il faudrait un index trigramme sur une version .lower() du nom parce qu'on fait un icontains qui gagnerait à se transformer en simple __contains (pareil pour le __startswith sur unaccent_name, on aurait pu éviter un index custom si unaccent_name était déjà en casse fixe).

#3 - 28 juin 2022 12:38 - Benjamin Dauvergne

- Assigné à Benjamin Dauvergne supprimé

En attendant il faudrait passer par des migrations SQL :/

#4 - 17 mars 2023 12:10 - Robot Gitea

- Statut changé de Nouveau à Solution proposée

- Assigné à mis à Benjamin Dauvergne

Benjamin Dauvergne (bdauvergne) a ouvert une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/passerelle/pulls/147>
- Titre : Ajout d'index sur les colonnes unaccent_name des modèles du connecteur base_adresse (#66694)
- Modifications : <https://git.entrouvert.org/entrouvert/passerelle/pulls/147/files>

#5 - 17 mars 2023 12:11 - Benjamin Dauvergne

On est enfin en 3.2, j'ai donc pu finir ce patch. Ça demande à ce que l'extension trigram soit facilement installable, je ne pense pas que ce soit encore

un problème.

#6 - 20 avril 2023 14:31 - Benjamin Dauvergne

- Statut changé de Solution proposée à En cours

#8 - 21 avril 2023 13:23 - Benjamin Dauvergne

- Statut changé de En cours à Solution proposée

#9 - 20 septembre 2023 19:17 - Benjamin Dauvergne

Ping.

#10 - 28 novembre 2023 16:50 - Benjamin Dauvergne

Validé qu'il n'y a aucun souci avec l'utilisation de pg_trgm sur nos déploiements.

#11 - 08 décembre 2023 16:46 - Robot Gitea

- Statut changé de Solution proposée à Solution validée

Nicolas Roche (nroche) a approuvé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/passerelle/pulls/147>

#12 - 15 janvier 2024 18:18 - Robot Gitea

- Statut changé de Solution validée à Résolu (à déployer)

Benjamin Dauvergne (bdauvergne) a mergé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/passerelle/pulls/147>
- Titre : base_adresse: ajouter des index sur unaccent_name et zipcode et tous les champs qui servent aux recherches (#66694)
- Modifications : <https://git.entrouvert.org/entrouvert/passerelle/pulls/147/files>

#13 - 15 janvier 2024 18:35 - Benjamin Dauvergne

- Statut changé de Résolu (à déployer) à En cours

Comme j'ai le nez creux je n'ai pas taggé et forcément un truc ne passe pas dans la migration en dev :

Running migrations:

```
Applying base_adresse.0031_add_indexes...Traceback (most recent call last):
  File "/usr/lib/python3/dist-packages/django/db/backends/utils.py", line 82, in _execute
    return self.cursor.execute(sql)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
```

```
psycopg2.errors.UndefinedObject: operator class "gin_trgm_ops" does not exist for access method "gin"
```

The above exception was the direct cause of the following exception:

Traceback (most recent call last):

```
File "/usr/lib/passerelle/manage.py", line 10, in <module>
  execute_from_command_line(sys.argv)
File "/usr/lib/python3/dist-packages/django/core/management/__init__.py", line 419, in execute_from_command_line
  utility.execute()
File "/usr/lib/python3/dist-packages/django/core/management/__init__.py", line 413, in execute
  self.fetch_command(subcommand).run_from_argv(self.argv)
File "/usr/lib/python3/dist-packages/django/core/management/base.py", line 354, in run_from_argv
  self.execute(*args, **cmd_options)
File "/usr/lib/python3/dist-packages/passerelle/utils/defer.py", line 93, in BaseCommand_execute
  return old_BaseCommand_execute(self, *args, **kwargs)
  ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/core/management/base.py", line 398, in execute
  output = self.handle(*args, **options)
  ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/hobo/multitenant/management/commands/migrate_schemas.py", line 77, in handle
  self.run_migrations(tenant, settings.TENANT_APPS, step, len_tenants)
File "/usr/lib/python3/dist-packages/hobo/multitenant/management/commands/migrate_schemas.py", line 98, in run_migrations
  command.execute(*self.args, **self.options)
```

```

File "/usr/lib/python3/dist-packages/passerelle/utis/defer.py", line 93, in BaseCommand_execute
    return old_BaseCommand_execute(self, *args, **kwargs)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/core/management/base.py", line 398, in execute
    output = self.handle(*args, **options)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/core/management/base.py", line 89, in wrapped
    res = handle_func(*args, **kwargs)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/core/management/commands/migrate.py", line 244, in handle
    post_migrate_state = executor.migrate(
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/db/migrations/executor.py", line 117, in migrate
    state = self._migrate_all_forwards(state, plan, full_plan, fake=fake, fake_initial=fake_initial)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/db/migrations/executor.py", line 147, in _migrate_all_forwards
    state = self.apply_migration(state, migration, fake=fake, fake_initial=fake_initial)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/db/migrations/executor.py", line 227, in apply_migration
    state = migration.apply(state, schema_editor)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/db/migrations/migration.py", line 126, in apply
    operation.database_forwards(self.app_label, schema_editor, old_state, project_state)
File "/usr/lib/python3/dist-packages/django/db/migrations/operations/models.py", line 761, in database_forwards
    schema_editor.add_index(model, self.index)
File "/usr/lib/python3/dist-packages/django/db/backends/postgresql/schema.py", line 218, in add_index
    self.execute(index.create_sql(model, self, concurrently=concurrently), params=None)
File "/usr/lib/python3/dist-packages/django/db/backends/base/schema.py", line 145, in execute
    cursor.execute(sql, params)
File "/usr/lib/python3/dist-packages/sentry_sdk/integrations/django/__init__.py", line 562, in execute
    return real_execute(self, sql, params)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/db/backends/utils.py", line 66, in execute
    return self._execute_with_wrappers(sql, params, many=False, executor=self._execute)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/db/backends/utils.py", line 75, in _execute_with_wrappers
    return executor(sql, params, many, context)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3/dist-packages/django/db/backends/utils.py", line 79, in _execute
    with self.db.wrap_database_errors:
File "/usr/lib/python3/dist-packages/django/db/utils.py", line 90, in __exit__
    raise dj_exc_value.with_traceback(traceback) from exc_value
File "/usr/lib/python3/dist-packages/django/db/backends/utils.py", line 82, in _execute
    return self.cursor.execute(sql)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
django.db.utils.ProgrammingError: operator class "gin_trgm_ops" does not exist for access method "gin"

```

#14 - 15 janvier 2024 18:39 - Benjamin Dauvergne

C'est le bazar visiblement la classe d'opérateur a été créée dans un schéma spécifique :

```

gist      | text                                | demo_passerelle_dev_entrouvert_org.gtrgm | demo_passerelle_dev_entrouv
ert_org.gist_trgm_ops | no

```

#15 - 15 janvier 2024 18:55 - Benjamin Dauvergne

Et encore merci stackoverflow <https://stackoverflow.com/questions/12986368/installing-postgresql-extension-to-all-schemas>

Je n'avais pas eu le problème sur authentic car je créais l'extension dans le schéma public qui est toujours inclus dans le search_path par défaut mais donc la bonne solution c'est de créer les extension dans pg_catalog et Django ne sait pas faire cela, donc ce sera du RunSQL :

```
+ migrations.RunSQL(
+     [
+         'CREATE EXTENSION IF NOT EXISTS pg_trgm WITH SCHEMA pg_catalog',
+     ],
+     reverse_sql=[
+         'DROP EXTENSION IF EXISTS pg_trgm'
+     ]
+ ),
```

#16 - 15 janvier 2024 18:56 - Robot Gitea

Benjamin Dauvergne (bdauvergne) a ouvert une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/passerelle/pulls/440>
- Titre : WIP: base_adresse: add indexes on on geographic models names (#66694)
- Modifications : <https://git.entrouvert.org/entrouvert/passerelle/pulls/440/files>

#17 - 15 janvier 2024 19:51 - Robot Gitea

- Statut changé de En cours à Solution proposée

#18 - 22 janvier 2024 10:47 - Benjamin Dauvergne

- Statut changé de Solution proposée à En cours

Déplacer l'extension du schéma pg_catalog à public sur suggestion de Pierre D.

#19 - 23 janvier 2024 15:21 - Benjamin Dauvergne

- Statut changé de En cours à Solution proposée

#20 - 24 janvier 2024 13:42 - Robot Gitea

- Statut changé de Solution proposée à Solution validée

Nicolas Roche (nroche) a approuvé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/passerelle/pulls/440>

#21 - 14 février 2024 15:52 - Robot Gitea

- Statut changé de Solution validée à Résolu (à déployer)

Benjamin Dauvergne (bdauvergne) a mergé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/passerelle/pulls/440>
- Titre : base_adresse: add indexes on on geographic models names (#66694)
- Modifications : <https://git.entrouvert.org/entrouvert/passerelle/pulls/440/files>

#22 - 14 février 2024 16:18 - Transition automatique

- Statut changé de Résolu (à déployer) à Solution déployée

#23 - 21 avril 2024 04:42 - Transition automatique

Automatic expiration

Fichiers

0001-wip.patch	3,01 ko	28 juin 2022	Benjamin Dauvergne
----------------	---------	--------------	--------------------