

Migrer la base données en utf8mb4 pour éviter les crash en cas d'utilisation d'emoji

17/10/2022 06:48 - Victor G/tuxayo

Statut:	Fermé	Début:	17/10/2022
Priorité:	Normale	Echéance:	08/02/2023
Assigné à:	Obitanz .	% réalisé:	100%
Catégorie:		Temps estimé:	1.00 heure
Version cible:	Sprint 2023 février		

Description

Et avoir un support de langues complet. Je crois que c'est pas juste émojis et autres symboles la différence entre utf8mb3(l'ancien défaut) et utf8mb4.

test: mettre un emoji comme 🍌 dans la description d'un sondage, résultat: page blanche a la dernière étape au lieu d'avoir la création qui marche.

Ticket chez framadate:
<https://framagit.org/framasoft/framadate/framadate/-/issues/536>

Historique

#1 - 31/12/2022 10:14 - Obitanz .

- Assigné à mis à Obitanz .

Procédure inspirée de deux autres :

- une partagée sur le ticket Framagit cité

- et une réalisée sur l'instance de LaToileScoute

1. Préparation

```
mysqldump date_chapril_org_db > AAAAMMJJ-dump-date-db.sql
```

2. Procédure

```
alter database `date_chapril_org_db` character set = utf8mb4 collate = utf8mb4_unicode_ci;
use date_chapril_org_db;
alter table fd_comment convert to character set utf8mb4 collate utf8mb4_unicode_ci;
alter table fd_framadate_migration convert to character set utf8mb4 collate utf8mb4_unicode_ci;
alter table fd_poll convert to character set utf8mb4 collate utf8mb4_unicode_ci;
alter table fd_slot convert to character set utf8mb4 collate utf8mb4_unicode_ci;
alter table fd_vote convert to character set utf8mb4 collate utf8mb4_unicode_ci;
```

3. Tests

Créer des sondages en testant de mettre des emoji partout (titre, nom, commentaire, etc.)

4. Retour arrière (au cas où)

Première action : remettre la base en utf8

```
alter database `date_chapril_org_db` character set = utf8 collate = utf8_general_ci;
use date_chapril_org_db;
alter table fd_comment convert to character set utf8 collate utf8_general_ci;
alter table fd_framadate_migration convert to character set utf8 collate utf8_general_ci;
alter table fd_poll convert to character set utf8 collate utf8_general_ci;
alter table fd_slot convert to character set utf8 collate utf8_general_ci;
alter table fd_vote convert to character set utf8 collate utf8_general_ci;
```

Si ça ne marche, on restaure la base à l'aide de la sauvegarde

```
source AAAAMMJJ-dump-date-db.sql
```

5. Détail technique*

En changeant l'encodage de la bdd, on impacte le nombre de caractères que l'on peut stocker dans les variables. En passant en utf8mb4, on passe de 3 octets à 4 octets par caractère. On a deux types de variable stockant du texte long, ces types sont varchar et text. Pour ce premier, la nouvelle

limite est à 191 caractères un seul champ dépasse cette valeur, donc pas de soucis à se faire pour les autres.
Le champ en question est un varchar(256) mais la longueur présente en base ne dépasse pas la nouvelle limite.
Pour les champs en text, la limite est 65535 octets (donc en divisant par 4 on a le nouveaux maximum de caractères : $2^{16} / 2^2 = 2^{14} = 16384$.
Le texte le plus long contient 2206 caractères donc pas de soucis non plus.

#2 - 08/02/2023 03:56 - Obitanz .

- *Echéance mis à 08/02/2023*
- *Statut changé de Nouveau à Résolu*
- *% réalisé changé de 0 à 100*
- *Temps estimé mis à 1.00 h*

Procédure déroulée sans accro :

```
mysqldump date_chapril_org_db > 20230208-dump-date-db.sql
mariadb
```

```
MariaDB [(none)]> alter database `date_chapril_org_db` character set = utf8mb4 collate = utf8mb4_unicode_ci;
Query OK, 1 row affected (0,170 sec)
```

```
MariaDB [(none)]> use date_chapril_org_db;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A
```

Database changed

```
MariaDB [date_chapril_org_db]> alter table fd_comment convert to character set utf8mb4 collate
utf8mb4_unicode_ci;
Query OK, 639 rows affected (14,093 sec)
Records: 639 Duplicates: 0 Warnings: 0
```

```
MariaDB [date_chapril_org_db]> alter table fd_framadate_migration convert to character set utf8mb4 collate
utf8mb4_unicode_ci;
Query OK, 13 rows affected (1,117 sec)
Records: 13 Duplicates: 0 Warnings: 0
```

```
MariaDB [date_chapril_org_db]> alter table fd_poll convert to character set utf8mb4 collate utf8mb4_unicode_ci
;
Query OK, 749 rows affected (17,790 sec)
Records: 749 Duplicates: 0 Warnings: 0
```

```
MariaDB [date_chapril_org_db]> alter table fd_slot convert to character set utf8mb4 collate utf8mb4_unicode_ci
;
Query OK, 4164 rows affected (21,192 sec)
Records: 4164 Duplicates: 0 Warnings: 0
```

```
MariaDB [date_chapril_org_db]> alter table fd_vote convert to character set utf8mb4 collate utf8mb4_unicode_ci
;
Query OK, 9501 rows affected (26,905 sec)
Records: 9501 Duplicates: 0 Warnings: 0
```

Tests fonctionnels et possibles encore pendant quelques jours sur : <https://date.chapril.org/wKm7qc9gaReVSSej>

#3 - 07/03/2023 22:14 - Quentin Gibeaux

- *Statut changé de Résolu à Fermé*

#4 - 08/03/2023 19:17 - Pierre-Louis Bonicoli

- *Version cible changé de Backlog à Sprint 2023 février*