

MYSQL使用UTF8中文乱码终极解决办法

作者：刘明灵 liumingling@hotmail.com 2009.8

MYSQL乱码的问题主要有几个环节决定，一个是前端网页；二是字符传输；三是连接MYSQL；四是MYSQL的存贮方式。

MYSQL的缺省字符集为latin1；如果不做任何处理，则存贮方式为latin；此时虽然在PHP程序中设置set names 'utf8'，可以正确存取中文；但其实utf8的字符按照latin1来存贮，占用的字节数差不多翻倍。这样的方式，本质上是按照乱码存放，读取的时候再还原，是不太可取的。

我们要做的是实现前端UTF-8，传输UTF-8，接口UTF8，存贮UTF8。

1、前端问题：网页和PHP程序存贮为UTF-8格式；网页中指定编码格式：

```
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
```

2、传输问题：apache的配置

```
vim /etc/apache2/conf.d/charset
AddDefaultCharset UTF-8
```

3、MYSQL的连接和存贮问题：修改MYSQL配置

```
vim /etc/mysql/my.cnf
[client]
default-character-set = utf8
[mysqld]
character-set-server=utf8
collation-server=utf8_general_ci
skip-character-set-client-handshake
[mysql]
default-character-set = utf8
```

注1: MySQL官方建议 `[mysqld]` 的部份应该要使用 `character-set-server` 与 `collation-server` 等选项,

而不是使用 `default-character-set`, 因为此选项已经是 deprecated。

注2: `skip-character-set-client-handshake`是忽略客户端的字符集, 使用服务器的设置; 这样就使得接口也是UTF8; 很多人建议在PHP程序中,

在连接数据库时设置`mysql_query("set names 'utf8'");` 或者在`my.cnf`中配置`init-connect='SET NAMES utf8'` (仅对非超级用户有效);

其实这里配置`skip-character-set-client-handshake`是最简洁有效的, 客户端可以不做任何处理。

注3: `[client]`配置客户端部分, 用于服务器本地维护; `[mysql]`配置MySQL启动条件。

重启mysql

```
/etc/init.d/mysql restart
```

进入MySQL, 查看配置是否正确:

```
mysql> show variables like 'character_set_%';
```

+-----+-----+	
Variable_name	Value
+-----+-----+	
character_set_client	utf8
character_set_connection	utf8
character_set_database	utf8
character_set_filesystem	binary
character_set_results	utf8
character_set_server	utf8
character_set_system	utf8
character_sets_dir	/usr/share/mysql/charsets/

```

+-----+-----+
8 rows in set (0.00 sec)

mysql> show variables like 'collation_%';
+-----+-----+
| Variable_name      | Value          |
+-----+-----+
| collation_connection | utf8_general_ci |
| collation_database  | utf8_general_ci |
| collation_server    | utf8_general_ci |
+-----+-----+
3 rows in set (0.00 sec)

```

MYSQL创建库和表的时候，都不要指定编码格式。此时MYSQL数据库的编码格式为
 utf8_general_ci ,表格式也是utf8_general_ci,中文字段也是用
 utf8_general_ci

实现以上步骤，则网页、phpmyadmin都不是乱码。而且phpmyadmin导出中文均为
 正常可见的字符。

如果在此之前使用了不正确的配置，那么转换到新的正确配置，之前的内容会出现乱
 码，所以需要注意做旧数据保存和重新导入。

有两篇文章论述比较清晰，对于理解这些有帮助：

<http://www.adsenseor.com/mysql/256.html>

<http://php.tmacn.cn/tmacn/20090516112.html>

参考：

<http://www.adsenseor.com/mysql/256.html>

<http://php.tmacn.cn/tmacn/20090516112.html>

需要做到：前台UTF-8,接口UTF8,存贮UTF8

1、MYSQL配置

```
vim /etc/mysql/my.cnf
```

```
[client]
default-character-set = utf8
[mysqld]
default-character-set = utf8
[mysql]
default-character-set = utf8
```

重启mysql

```
/etc/init.d/mysql restart
```

```
mysql> show variables like 'character_set_%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| character_set_client | utf8 |
| character_set_connection | utf8 |
| character_set_database | utf8 |
| character_set_filesystem | binary |
| character_set_results | utf8 |
| character_set_server | utf8 |
| character_set_system | utf8 |
| character_sets_dir | /usr/share/mysql/charsets/ |
+-----+-----+
8 rows in set (0.00 sec)
```

```
mysql> show variables like 'collation_%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| collation_connection | utf8_general_ci |
| collation_database | utf8_general_ci |
| collation_server | utf8_general_ci |
+-----+-----+
3 rows in set (0.00 sec)
```

2、MYSQL数据库的编码格式为utf8_general_ci ,表格式也是utf8_general_ci, 中文字段也是用utf8_general_ci

3、网页和PHP程序存为UTF-8格式

4、网页中指定<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />编码格式

5、apache的配置

```
vim /etc/apache2/conf.d/charset
AddDefaultCharset UTF-8
```

6、PHP程序在数据库连接后加上

```
mysql_query("set names 'utf8'");//注意是utf8,不是utf-8
```

实现以上步骤，则网页、PHPMYADMIN都不是乱码。

修改其中的某个ini配置文件，添加utf8字符集处理

[mysql]下添加:

```
default-character-set=utf8
```

[mysqld]下添加:

```
default-character-set=utf8
```

```
collation_server=utf8_unicode_ci
```

```
//server端编码为utf8
```

```
character_set_server=utf8
```

//强制客户端编码为utf8

skip-character-set-client-handshake

```
-----  
[mysqld]  
character-set-server=utf8  
collation-server=utf8_general_ci  
init-connect='SET NAMES utf8'  
skip-character-set-client-handshake  
  
[client]  
default-character-set=utf8
```

註1:

MySQL官方建議 [mysqld] 的部份應該要使用 `character-set-server` 與 `collation-server` 等選項,

而不是使用 `default-character-set`, 因為此選項已經是 deprecated。

註2:

對於擁有 `SUPER` 權限的使用者(`root`), `init-connect` 選項無效。因此最好不要用具有 `SUPER` 權限的

帳號來存取資料庫。會有這個現象的原因是 MySQL 怕你要是不小心設錯編碼造成無法登入MySQL Server

時, 至少還有 `root` 帳號可以登入 MySQL Server 去進行故障排除。
