字符集
钱：５１０３１６６６０

１．字符集的基本知识
字符集最早的编码方案来自于ASCII（７位编码，后扩展为８位编码）,此字符集中定义了１２８个字符，其中９５个是可显示字符，剩下的是３３个无法显示的控制字符.
oracle最早支持的编码方案也就是US7ASCII.

后来为了解决多语言的支持，出现了Unicode编码，此编码的口号是：给每个字符提供唯一的数字，无论是什么平台，不论是什么程序，无论是什么语言.

Unicode编码方案主要有３个实施标准：UTF-8，USC-2，UTF-16．Oracle从７．２开始支持ＵＴＦ－８编码.

Unicode编码方案可以表示更多的字符，但是由于多位的存储，需要额外的存储空间和网络传输，所以选择最适合的数据库字符集仍然需要慎重考虑.

２．数据库的字符集

在创建数据库时，可以指定字符集(CHARACTER SET)和国际字符集(NATIONAL CHARACTER SET).

字符集的主要作用:
    用于存储CHAR ,VARCHAR2, CLOB, LONG等类型数据;
    用来标示诸如表名，列名以及PL/SQL变量等.
    用于存储SQL和PL/SQL代码等.

国家字符集用以存储ＮＣＨＡＲ，　ＮＶＡＲＣＨＡＲ２，ＮＣＬＯＢ等类型数据.

对于简体中文平台，一般缺省的字符集是ZHS16GBK.一旦字符集选定了，数据库中能够存储的字符就受到了限制，所以选择字符集应该尽可能容纳所有用到的字符.
常见的中文字符集有:
ZHS16CGB231280　　CGB2312-80 Simplified Chinese MB,ASCII
GB2312码是中华人民共和国国家汉字信息交换用编码，全称<信息交换用汉字编码字符集－基本集>．通行于中国内地，新加坡等地也使用此编码.

ZH16GBK　　GBK　　Simplified Chinese MB,ASCII, UDC
ＧＢＫ编码是１９９５年１２月颁布的指导性规范，ＧＢＫ与国家标准ＧＢ２３１２－８０信息处理交换码所对应的，事实上的内码标准兼容；同时，在字汇一级支持ISO/IEC 10646-1和GB 3000-1的全部中日韩(CJK)汉字(20902字)，包含了更多的编码.

注意：ZHS16GBK并非是ZHS16CGB231280的严格超集.

oracle的字符集命名通常遵循以下命名规则：<Language><bit size><encoding> .即<语言><比特位数><编码>．例如ＺＨＳ．１６．ＧＢＫ

３．NLS_LANG的设置与影响
3.1NLS_LANG环境变量的格式
```
# export NLS_LANG=<Language>_<Territory>.<Clients Characterset>
    #export NLS_LANG=SIMPLIFIED CHINESE_CHINA.ZHS16GBK
    #export NLS_LANG=AMERICAN_AMERICA.ZHS16GBK
```