[**Base64编码**](http://www.cnblogs.com/BeyondTechnology/archive/2011/03/27/1997165.html)

下面，详细介绍Base64的编码转换方式。  
  
所谓Base64，就是说选出64个字符——小写字母a-z、大写字母A-Z、数字0-9、符号“+”、“/”（再加上作为垫字的“=”，实际上是65个字符）——作为一个基本字符集。然后，其他所有符号都转换成这个字符集中的字符。  
  
具体来说，转换方式可以分为四步。  
  
第一步，将每三个字节作为一组，一共是24个二进制位。  
  
第二步，将这24个二进制位分为四组，每个组有6个二进制位。  
  
第三步，在每组前面加两个00，扩展成32个二进制位，即四个字节。  
  
第四步，根据下表，得到扩展后的每个字节的对应符号，这就是Base64的编码值。  
  
　　0　A　　17　R　　　34　i　　　51　z  
  
　　1　B　　18　S　　　35　j　　  52　0  
  
　　2　C　　19　T　　　36　k　　  53　1  
  
　　3　D　　20　U　　　37　l　　　54　2  
  
　　4　E　　21　V　　　38　m　　 55　3  
  
　　5　F　　22　W　　  39　n　　  56　4  
  
　　6　G　　23　X　　　40　o　　  57　5  
  
　　7　H　　24　Y　　　41　p　　　58　6  
  
　　8　I　　 25　Z　　　42　q　　　59　7  
  
　　9　J　　 26　a　　　43　r　　　 60　8  
  
　　10　K　 27　b　　　44　s　　　 61　9  
  
　　11　L　  28　c　　　45　t　　　 62　+  
  
　　12　M　  29　d　　  46　u　　　63　/  
  
　　13　N　  30　e　　　47　v  
  
　　14　O　  31　f　　　48　w　　　  
  
　　15　P　　32　g　　  49　x  
  
　　16　Q　  33　h　　　50　y  
  
因为，Base64将三个字节转化成四个字节，因此Base64编码后的文本，会比原文本大出三分之一左右。  
  
4.  
  
举一个具体的实例，演示英语单词Man如何转成Base64编码。

Text content　　 　      M 　　　　                a 　　　　               n　　  
ASCII 　　　　　　　     77 　　                   97 　　                  110　　  
Bit pattern          0 1 0 0 1 1 0 1       0 1 1 0 0 0 0 1       0 1 1 0 1 1 1 0   
Index                         19             22               5                      46  
Base64-Encoded          T              W                F                       u  
  
  
第一步，“M”、“a”、“n”的ASCII值分别是77、97、110，对应的二进制值是01001101、01100001、01101110，将它们连成一个24位的二进制字符串010011010110000101101110。  
  
第二步，将这个24位的二进制字符串分成4组，每组6个二进制位：010011、010110、000101、101110。  
  
第三步，在每组前面加两个00，扩展成32个二进制位，即四个字节：00010011、00010110、00000101、00101110。它们的十进制值分别是19、22、5、46。  
  
第四步，根据上表，得到每个值对应Base64编码，即T、W、F、u。  
  
因此，Man的Base64编码就是TWFu。  
  
5.  
  
如果字节数不足三，则这样处理：  
  
a）二个字节的情况：将这二个字节的一共16个二进制位，按照上面的规则，转成三组，最后一组除了前面加两个0以外，后面也要加两个0。这样得到一个三位的Base64编码，再在末尾补上一个“=”号。  
  
比如，“Ma”这个字符串是两个字节，可以转化成三组00010011、00010110、00010000以后，对应Base64值分别为T、W、E，再补上一个“=”号，因此“Ma”的Base64编码就是TWE=。  
  
b）一个字节的情况：将这一个字节的8个二进制位，按照上面的规则转成二组，最后一组除了前面加二个0以外，后面再加4个0。这样得到一个二位的Base64编码，再在末尾补上两个“=”号。  
  
比如，“M”这个字母是一个字节，可以转化为二组00010011、00010000，对应的Base64值分别为T、Q，再补上二个“=”号，因此“M”的Base64编码就是TQ==。  
  
6.  
  
再举一个中文的例子，汉字“严”如何转化成Base64编码？  
  
这里需要注意，汉字本身可以有多种编码，比如gb2312、utf-8、gbk等等，每一种编码的Base64对应值都不一样。下面的例子以utf-8为例。  
  
首先，“严”的utf-8编码为E4B8A5，写成二进制就是三字节的“11100100 10111000 10100101”。将这个24位的二进制字符串，按照第3节中的规则，转换成四组一共32位的二进制值“00111001 00001011 00100010 00100101”，相应的十进制数为57、11、34、37，它们对应的Base64值就为5、L、i、l。  
  
所以，汉字“严”（utf-8编码）的Base64值就是5Lil