

Re: ??Difference Between utf8encoder.GetBytes and Encoding.ASCII.GetBytes

Source:

<http://www.derkeiler.com/Newsgroups/microsoft.public.dotnet.framework.aspnet.security/2005-02/0262.html>

From: Phil C. (*charlestek_at_rcn.com*)

Date: 02/25/05

Date: Thu, 24 Feb 2005 18:17:52 -0500

Thanks Joe,

It's the usual "too many ways of doing something" make things more complex.

"Joe Kaplan (MVP – ADSI)" <joseph.e.kaplan@removethis.accenture.com> wrote in message news:%23AgR7NsGFHA.4084@TK2MSFTNGP14.phx.gbl...

> *The easiest thing to do is to write some code to test it and see, but I'll try to explain too.*

>

> *UTF8 and Unicode (which is really UTF16 as an encoding) are just two different ways to create a binary encoding of a unicode string. UTF8 uses a variable number of bytes for each character (depending on the character) and UTF16 will use 2 bytes for each character. Since you test string, "test", is all ASCII characters, the UTF8 version will be 4 bytes and the same as the ASCII version. The Unicode version will be 8 bytes. To see differences between ASCII and UTF8, you need to use non-ASCII characters in your test.*

>

> *The various static/shared properties on the Encoding classes are just shortcuts to keep you from having to build a new instance of the encoding class. Generally, it will always be a little faster to just use them:*

>

> *Encoding.UTF8*
> *Encoding.Unicode*

>

> *HTH,*

>

> *Joe K.*

>

> *"Phil C." <charlestek@rcn.com> wrote in message news:O1\$VxNrGFHA.3628@TK2MSFTNGP15.phx.gbl...*

>> *Thank you Joe, you saved me a lot of grief.*

>> *However, then, what is the difference between*

>> *UTF8Encoding.GetBytes("text")*

>> *and Encoding.Unicode.GetBytes("text")*

>> *or the converse*

>> *UTF8Encoding.GetString(Byte())*

>> *Encoding.Unicode.GetString(Byte())*
>> *??*
>>
>>

>> *"Joe Kaplan (MVP – ADSI)" <joseph.e.kaplan@removethis.accenture.com>*
>> *wrote in message news:uyWgMXqGFHA.2616@tk2msftngp13.phx.gbl...*
>>> *Generally speaking, the different encoding classes will give you an*
>>> *array of bytes from a string corresponding to how that encoding actually*
>>> *represents a string. Unicode (UTF16) represents each character as 2*
>>> *bytes. UTF8 will use a variable number of bytes for each character, but*
>>> *uses only one for ASCII characters, so it generally uses much less space*
>>> *to store the same Unicode data.*
>>>
>>> *ASCII converts characters into a single byte using only 7 bits of each*
>>> *byte. Since it only supports ASCII characters, it can result in data*
>>> *loss if the string in question contains non-ASCII characters. It rarely*
>>> *has a use in .NET crypto since strings are unicode in .NET.*
>>>
>>> *If your encryption key is stored as text, it is probably stored in*
>>> *Base64. In that case, you probably want to use Convert.FromBase64String*
>>> *to convert the string key into a byte array.*
>>>
>>> *Joe K.*

>>> *"Phil C." <charlestek@rcn.com> wrote in message*
>>> *news:u6LpKgpGFHA.2616@tk2msftngp13.phx.gbl...*
>>>> *Hi. (Using VB.Net) I have a symmetric encryption key stored as text,*
>>>> *encrypted by DPAPI in my web config that I use a handler*
>>>> *class to decrypt by the DPAPI and pass to the class that does the*
>>>> *encryption/decryption.*
>>>> *The decrypted DPAPI key is a string and needs to be converted to a byte*
>>>> *array for use by the encryption/decryption class. I'm confused as to*
>>>> *the difference using utf8encoder.GetBytes() or*
>>>> *Encoding.ASCII.GetBytes() to do this.*

>>>> *Thanks,*
>>>>
>>>> *Phil*
>>>> *Boston, MA*

>>>>
>>>
>>>
>>
>>
>
>