
Subject: URL Parsing for wget in IDL
Posted by [Ben Panter](#) on Tue, 17 Jul 2007 15:54:24 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Folks,

Does anyone know of a handy way of parsing URLs in IDL? Or even a nice perl script to do it? The situation is that we have a server that can accept SQL queries via URL encoding of a get command, which works in the main part but falls over with '&' and '+' type syntax.

I'll write a regexp version myself if necessary, but it's not trivial
- you have to ensure the ordering is correct as there are several encodings which use characters which would otherwise be replaced.

cheers,

Ben

--
Ben Panter, Edinburgh, UK.
Email false, <http://www.benpanter.co.uk>
or you could try ben at ^^^^^^^^^^^^^

Subject: Re: URL Parsing for wget in IDL
Posted by [Michael Galloy](#) on Wed, 18 Jul 2007 18:48:11 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Jul 17, 9:54 am, Ben Panter <m...@privacy.net> wrote:

> Hi Folks,
>
> Does anyone know of a handy way of parsing URLs in IDL? Or even a
> nice perl script to do it? The situation is that we have a server that
> can accept SQL queries via URL encoding of a get command, which works in
> the main part but falls over with '&' and '+' type syntax.
>
> I'll write a regexp version myself if necessary, but it's not trivial
> - you have to ensure the ordering is correct as there are several
> encodings which use characters which would otherwise be replaced.

There is a PARSE_URL routine in IDL 6.4 that would probably get you started.

Mike
--
www.michaelgalloy.com

Subject: Re: URL Parsing for wget in IDL
Posted by [Conor](#) on Wed, 18 Jul 2007 19:11:29 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Jul 17, 11:54 am, Ben Panter <m...@privacy.net> wrote:

> Hi Folks,
>
> Does anyone know of a handy way of parsing URLs in IDL? Or even a
> nice perl script to do it? The situation is that we have a server that
> can accept SQL queries via URL encoding of a get command, which works in
> the main part but falls over with '&' and '+' type syntax.
>
> I'll write a regexp version myself if necessary, but it's not trivial
> - you have to ensure the ordering is correct as there are several
> encodings which use characters which would other wise be replaced.
>
> cheers,
>
> Ben
>
> --
> Ben Panter, Edinburgh, UK.
> Email false, <http://www.benpanter.co.uk>
> or you could try ben at ^^^^^^^^^^^^^^^

Googling "perl parse url" brought up plenty of promising looking candidates. Here's one:

<http://textsnippets.com/posts/show/523>

IDL does have regular expressions, although I haven't tried to see if they can do as much as perl (they probably can). So you can always take one of those regular expressions and just convert it to an IDL regular expression. IDL does use different syntax than perl, so it would take some work. You would also have to know a lot about the IDL regular expression system, which I'm afraid I can't help you with. There's a website somewhere that explains it all, but I can't find it at the moment. Someone here knows though...

Subject: Re: URL Parsing for wget in IDL
Posted by [Michael Galloy](#) on Thu, 19 Jul 2007 15:19:25 GMT
[View Forum Message](#) <> [Reply to Message](#)

Conor wrote:

> On Jul 17, 11:54 am, Ben Panter <m...@privacy.net> wrote:
>> Hi Folks,
>>

```

>> Does anyone know of a handy way of parsing URLs in IDL? Or even a
>> nice perl script to do it? The situation is that we have a server that
>> can accept SQL queries via URL encoding of a get command, which works in
>> the main part but falls over with '&' and '+' type syntax.
>>
>> I'll write a regexp version myself if necessary, but it's not trivial
>> - you have to ensure the ordering is correct as there are several
>> encodings which use characters which would other wise be replaced.
>>
>> cheers,
>>
>> Ben
>>
>> --
>> Ben Panter, Edinburgh, UK.
>> Email false,http://www.benpanter.co.uk
>> or you could try ben at ^^^^^^^^^^^^^
>
> Googling "perl parse url" brought up plenty of promising looking
> candidates. Here's one:
>
> http://textsnippets.com/posts/show/523
>
> IDL does have regular expressions, although I haven't tried to see if
> they can do as much as perl (they probably can). So you can always
> take one of those regular expressions and just convert it to an IDL
> regular expression. IDL does use different syntax than perl, so it
> would take some work. You would also have to know a lot about the IDL
> regular expression system, which I'm afraid I can't help you with.
> There's a website somewhere that explains it all, but I can't find it
> at the moment. Someone here knows though...
>

```

The regular expression they suggest is:

```
^((http[s]?|ftp):V)?V?([^\s]+)((\w+)*V)([\w\-\.\.]+[^\# ?\s]+)(.*)?(#[\w\-\.\.])$
```

which I think will be OK for IDL except \s (whitespace) needs to be
 [[:blank:]] and \w (word character) needs to be [[:alnum:]]

Check out:

<http://michaelgalloy.com/2006/06/11/regular-expressions.html>

which links to a paper with more details at:

<http://www.ittvis.com/codebank/search.asp?FID=311>

Mike

--

www.michaelgalloy.com

Subject: Re: URL Parsing for wget in IDL

Posted by [Ben Panter](#) on Thu, 19 Jul 2007 15:45:50 GMT

[View Forum Message](#) <> [Reply to Message](#)

Michael Galloy wrote:

> The regular expression they suggest is:

>

> ^((http[s]?|ftp):V)?V?([^\s]+)((\w+)*V)([\w\-\.\.]+\s+)(.*)?(#[\w\-\.\.]+)\$

>

>

> which I think will be OK for IDL except \s (whitespace) needs to be

> [[:blank:]] and \w (word character) needs to be [[:alnum:]]

>

> Check out:

>

> <http://michaelgalloy.com/2006/06/11/regular-expressions.html>

>

> which links to a paper with more details at:

>

> <http://www.itvis.com/codebank/search.asp?FID=311>

>

Thanks guys - will do.

Ben

--

Ben Panter, Edinburgh, UK.

Email false, <http://www.benpanter.co.uk>

or you could try ben at ^^^^^^^^^^^^^
