

Re: Sed advice needed

Source: <http://linux.derkeiler.com/Mailing-Lists/Debian/2007-05/msg04570.html>

- *From:* Bob McGowan <bob_mcgowan@xxxxxxxxxxxxxx>
 - *Date:* Thu, 31 May 2007 09:42:05 -0700
-

Piers Kittel wrote:

Hello all,

I need to delete some words out of a large file containing information about packets I'm analysing. I know I can use sed to do this, but haven't really used it before, so am a bit unsure of how to do it. Two example lines are as of below:

```
"181","1324.014027","111.111.111.111","111.111.111.111","RTP","Payload type=ITU-T
H.261, SSRC=2008229573, Seq=54520, Time=1725612773, Mark"
"185","1324.078941","111.111.111.111","111.111.111.111","RTP","Payload type=ITU-T
H.261, SSRC=2008229573, Seq=54521, Time=1725616276"
```

I need to convert the above to the below:

```
"181","1324.014027","111.111.111.111","111.111.111.111","RTP","54520"
"185","1324.078941","111.111.111.111","111.111.111.111","RTP","54521"
```

What's the best way to do this? I've been reading the man pages of sed, cut and awk but I can't quite figure out how to do this. Any ideas?

Thanks very much for your time in advance!

Regards – Piers

—To UNSUBSCRIBE, email to debian-user-REQUEST@xxxxxxxxxxxxxxxxxxxxxx a subject of "unsubscribe". Trouble? Contact listmaster@xxxxxxxxxxxxxxxxxxxxxx

To rephrase your question and data, so I'm sure I understand and you know what I'm doing: You want to print items 1, 2, 3, 4, and 5, below, as is, with only the Seq value from item 6, enclosed in double quotes.

- 1: "181",
- 2: "1324.014027",
- 3: "111.111.111.111",
- 4: "111.111.111.111",
- 5: "RTP",

Re: Sed advice needed

6: "Payload type=ITU-T H.261, SSRC=2008229573, Seq=54520,
Time=1725612773, Mark"

If my description is correct, this will do the trick, assuming that all the lines in the file have the same spacing and relative positioning:

```
sed 's/"Pay.*Seq=\(.*\), Time.*$/"\1"/'
```

What I mean by the above assumption would be, for example, that a Seg number of 386 would look like "Seq=386" rather than "Seg= 386", or other formatting differences.

The above sed substitution finds the part of the string starting with a double quote followed by Pay, then any characters repeated up to the portion with Seq=, followed by anything at all up to a comma space followed by the word Time followed by anything up to the end of the line. The portion between the \(...\) is remembered and is accessed in the substitute by the \1. So, this part of the line:

```
"Pay....Seq=#####....Mark"
```

is replaced with:

```
"#####"
```

Where ##### is the sequence number for each line.

--

Bob McGowan

Attachment: smime.p7s

Description: S/MIME Cryptographic Signature

Re: Sed advice needed