
Subject: Re: How to read first and last line of an ascii file FAST
Posted by [Heinz Stege](#) on Sun, 29 Aug 2010 16:32:35 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Sun, 29 Aug 2010 08:32:15 -0700 (PDT), ahs wrote:

> Hello group,
>
> My data is stored in a bunch of ascii files, each 5-10MB big. Each
> line represent a sample with a timestamp. I would like to analyse all
> my files by reading the timestamp from the first and last line of
> every file. Getting the timestamp from the last line is the issue
> here. Right now I'm doing it like
>
> readf, fid, firstline
> nlines = FILE_LINES(fnames[i])
> skip_lun, fid, nlines-2, /LINES
> readf, fid, lastline
>
> skip_lun is really slow when the files are big and numerous. Does
> anyone have a tip of I could do this faster? Is there a way to
> directly point to the end of the file minus one line?
>
> regards,
> Agnar

Hi Agnar,

I would expect, that the file_lines function is taking some time too.
You can try the following:

Get the file size in bytes by the fstat function. Then use point_lun
to set the file pointer to n bytes before the file end. n should be
greater than the length of the last line. Read the n bytes into a byte
array and search for the last occurrence of CR and/or LF in that array.
If you don't find a CR/LF increase n and repeat the steps from
point_lun to here. Finally convert the bytes after CR/LF to string
type.

HTH, Heinz
