
Subject: Re: Lots of files

Posted by [lasse](#) on Mon, 19 Mar 2007 11:32:18 GMT

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

On 18 Mar, 03:59, David Fanning <n...@dfanning.com> wrote:

> Lasse Clausen writes:

>> Well thanks, that works, however it did not bring the speed boost I
>> had hoped for. So I had another thought: Actually, all data is one
>> line, not in one line per station as I said earlier. But I know that
>> each data set is 1440 characters long, so here is the outline of my
>> code, after I opened all the files:

```
>
>> info = file_info(input_filename)
>> lines = info.size/1440L
>
>> for i=0L, lines-1L do begin
>>     point_lun, fin, i*1440L
>>     readf, fin, line, format='(A1440)'
>>     ; extracting station name
>>     hstat = strlowercase(strmid(line, 12, 3))
>>     ; find correct file unit
>>     tmp = where(stats eq hstat)
>>     printf, tmp[0]+1, line
>> endfor
```

>
> Well lots of string processing and WHERE's going
> on here, which I think is what is slowing things
> down. How about something like this:

```
>
> theLines = Assoc(lun, BytArr(1440))
> maxYear = Max(stats)
> for l=0L, lines-1L do begin
>     aLine = theLines[l]
>     ; extracting station name
>     hstat = String(aLine[12:15])
>     ; find correct file unit
>     printf, (maxYear-hstat)+1, String(aLine)
> endfor
```

>
> Cheers,

>
> David

> --

> David Fanning, Ph.D.

> Fanning Software Consulting, Inc.

> Coyote's Guide to IDL Programming:<http://www.dfanning.com/>

> Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Yeeeha! This solution (thanks Lajos/David)

```
openr, fin, filename, /get_lun
theLines = Assoc(fin, BytArr(1440))
for i=0L, lines-1L do begin
  aLine = theLines[i]
  hstat = String(aLine[12:14])
  tmp = where(stats eq hstat)
  printf, tmp[0]+1, string(aLine)
endfor
```

goes like sh** off a shovel! I will use that. However, I still do not know why

```
openr, fin, filename, /get_lun
for i=0L, lines-1L do begin
  point_lun, fin, i*1440L
  readf, fin, line, format='(A1440)'
endwhile
```

(note that all the string processing and where is taken out) takes about 30 seconds. Well, I suspect the format code, there isn't much else to suspect. Can anybody explain to me why the above solution is so slow?

Cheers
Lasse
