Subject: Re: Lots of files
Posted by lasse on Sat, 17 Mar 2007 17:32:14 GMT
View Forum Message <> Reply to Message

On 16 Mar, 21:23, David Fanning <n...@dfanning.com> wrote:
> Paul van Delst writes:
>> I know you didn't intend to suggest hardwiring 99 different fileid's :o)
>
> With Cut and Paste it's not so bad. Of course, you
> spend the next five hours fixing typos, but... :-)
>
> 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.")

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 = strlowcase(strmid(line, 12, 3))
   ; find correct file unit
   tmp = where(stats eq hstat)
   printf, tmp[0]+1, line
endfor
```

I chose the above solution because my favoured one:

```
while not(eof(fin)) do begin
   readf, fin, line, format='(A1440)'
   hstat = strlowcase(strmid(line, 12, 3))
   tmp = where(stats eq hstat)
   printf, tmp[0]+1, line
endwhile
```

does not seem to work in the way I expected, i.e. read 1440 byte, parse station, write data, read next 1440 byte... until end of file. Rather, it reads the first 1440 bytes and then hits the end of the file (while loop is executed once). So that is why I wondered what the readf command with the above format code actually does. Since it hits the end of file border after the first read command, I suspect it actually reads in all data, and then extracts the first 1440 bytes from that. Which would explain why the solution I am running now (with the for loop) is so slow: about 20 seconds for 3000 lines (4MB file). On some chunky Sun server, mind you. Any more ideas?

Cheers
Lasse
