Subject: Re: Handle big data files
Posted by lucesmm on Mon, 02 Nov 2015 15:02:10 GMT
View Forum Message <> Reply to Message

Hello Helder
Thank you for your help, I am implementing the first option you suggested. I am having trouble now because one of the columns is calendar format and I am getting this error
PRINTF: Value of Julian date is out of allowed range
:( is there an easy way to solve this, or shoudl I just keep date info in different colums?

HEre is what the data look like and my format:

2014-12-01T00:00:12.905
C(CYI, X ,CMOI02, X ,CDI02,X, CHI02,X, CMI02, X, CSF0)
I am using the same fomat in both reading and writing the data
is this correct?

Now I wanted to ask something else. I have a bunch of columns that I don't really need, is there a way to create a save file just with the date column and the one that I care? (I am guessing this is the easiest version of files to save big data because they are binary)

Thanks again

-Luz Maria
On Monday, November 2, 2015 at 1:13:43 AM UTC-8, Helder wrote:
> Hi,
> I think that this line is responsible for making things slow:
>
> data=[[data],[line]]
>
> If the array data gets to be lon, then it will take a long time to copy the previous data to a new variable and add one element...
>
> You have two options:
> 1) only valid for IDL version >8.0. Use a list(). before the for use:
> data = list()
> then instead of data=[[data],[line]] use:
> data->add, line
> Then at the end:
> PrintF, outLun, data->toArray(), FORMAT= '...'
>
> 2) it's more complicated, but general. Create the data array loooong, then fill it up. You could also actually guess it's length:
> nData = 0l
> FOR i=0,nfiles-1 DO BEGIN
> nlines = FILE_LINES(files[i])
> nData += nlines-1 ;one line you always disregard
> ENDFOR

```
>
>  Now create data so that it is long enough:
>  myDataStructure = make_array(17,1, type=5)
>  data = replicate(myDataStructure, nData)
>
>  and in the cycle you fill up. You will also need a "fill-up" counter:
>
>  fillCounter = 0l
>  FOR...
>      ...
>      while...
>          ...
>          data[fillCounter] = line
>          fillCounter++
>      endwhile
>      ...
>  endfor
>
>  I hope it helps...
>
>  Cheers,
>  Helder
```

---