Subject: Re: Handle big data files
Posted by Helder Marchetto on Mon, 02 Nov 2015 09:13:39 GMT
View Forum Message <> Reply to Message

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,