
Subject: Re: error reading a large number of binary files
Posted by [Paul Van Delst\[1\]](#) on Mon, 23 Apr 2007 17:47:24 GMT
[View Forum Message](#) <> [Reply to Message](#)

mark@atmos.colostate.edu wrote:

>
> Hi Paul. Thanks for taking an interest in this. On Friday I actually
> tried something similar to your idea of just reading the files and not
> doing any computations, and that also failed. And just now I took the
> program you created above, saved it as readudaily2.pro and tried it
> three times. Here are the results:

Well, to be honest I didn't think it would work, but it eliminates one of the facile answers.

What a bizarre problem!

The only thing that occurs to me right now (and insert a bunch of hand waving here), is that maybe the IDL procedures are not as sophisticated as the RTLs used by the Fortran codes in reading the files. By that I mean if the record markers in the files do not *exactly* correspond to the length of the data you want to read in, maybe IDL chucks a wobbly and stops. For Fortran code, as long as the input data list is less than or equal to what is in the record, it doesn't matter if you read *less* than what the record contains - the next read will skip the remainder of the last record. If IDL, however, just stops reading the current record and simply starts up again for the next read, then it may treat the remaining data of the previous record as the record marker for the next one and, thus, things go pear shaped. (I told you I was waving my hands about :o)

It may be worthwhile writing a test/check IDL script that opens the files as direct access (i.e. without the /f77_unformatted) and explicitly reads the record markers that bookend each record to check their size and compare that to what you expect. If they're different, then maybe that's the problem. However, it's a not insignificant amount of work to do it.

cheers,

paulv

>
> attempt #1:
> daily_4/u196112.bin
> daily_4/u196201.bin
> daily_4/u196202.bin
> % READUDAILY2: READU: Corrupted f77 unformatted file detected. Unit:
> 100, File: daily_4/u196202.bin
>
> attempt #2:
> daily_4/u198409.bin
> daily_4/u198410.bin
> daily_4/u198411.bin

> % READUDAILY2: READU: Corrupted f77 unformatted file detected. Unit:
> 100, File: daily_4/u198411.bin
>
> attempt #3:
> daily_4/u199610.bin
> daily_4/u199611.bin
> % READUDAILY2: READU: Corrupted f77 unformatted file detected. Unit:
> 100, File: daily_4/u199611.bin
>
> As you can see, it fails on a different file each time. Although
> since it seemed to get further along on each successive trial, maybe I
> should try one or two more times and perhaps I'll get lucky??? :)
> Actually, I think this is just random and not a trend.
>
> I wonder if the same error would occur if I generated a large number
> of binary files of a smaller size (say a 5x5 array of random numbers),
> because maybe it's not how large the files are but how many of them
> there are. ?????
>
> Thanks,
> Mark
>

--

Paul van Delst Ride lots.
CIMSS @ NOAA/NCEP/EMC Eddy Merckx
