Subject: Re: openr or openw with COMPRESS flag broken for large data files
Posted by KRDean on Thu, 09 Jun 2011 14:37:17 GMT
View Forum Message <> Reply to Message

On Jun 8, 2:06 pm, JJ <j...@cornell.edu> wrote:
> I've been using openr, /compress as a way to read gzipped files
> without first gunzipping them in the OS. I have also created gzipped
> files in idl using openw, /compress. This has worked fine until now -
> when I tried to do this with a really big file. I'm guessing this has
> something to do with IDL using a long integer to do the indexing,
> because I find that this will work fine if I try to write/read an
> array of size 2^31-1 bytes, but will fail for an array of size 2^31
> bytes.
>
> Example:
>
> IDL> a = bytarr(2ul^31-1)
> IDL> openw, 1, 'test.dat', /compress
> IDL> writeu, 1, a
> IDL> close, 1
> IDL> openr, 1, 'test.dat', /compress
> IDL> readu, 1, a
> IDL> close, 1
>
> This works fine, but if you try with a = bytarr(2ul^31), it fails. It
> seems to fail in different ways depending on circumstances. With an
> array of 2^29 long integers, it writes the file, but gives an end-of-
> file error when trying to read it. With an array of 2^31 bytes, it is
> not able to write the file in the first place. With the particular
> files I was dealing with, IDL gave no complaints when writing or
> reading, but the data was corrupted (all zeros after a certain point).
>
> The save routine in IDL seems to work fine with such large arrays and
> the /compress keyword set. Reading/writing this data without /
> compress works fine. Writing an uncompressed file, then gzipping it
> in the OS, then attempting to read it using /compress does not work.
> Likewise, creating a file using openw, /compress, then gunzipping it,
> then reading it in without /compress also does not work. So the bug
> appears to be in both read and write - probably calling the same
> routine somewhere.
>
> Using the IDL save routine is not a workaround for me because I am
> reading/writing PDS (Planetary Data System) files. Doing gzip/gunzip
> at the OS level works fine and is a workaround for now - though
> annoying overhead.
>
> This seems like a bug in IDL. I upgraded to IDL 8.1 to test this, and
> the same error is there (at least for Solars x86 64-bit). I first

> noticed the problem in IDL 7.1 linux x86 64-bit version.
>
> Thanks.
>
> -JJ

From my dealing with gzip, inside and outside of IDL, there is a 2Gb
file size limit.

Kelly Dean
Milliken, CO
