Subject: h5_parse without READ_DATA
Posted by simulana on Mon, 03 Feb 2014 18:39:25 GMT
View Forum Message <> Reply to Message

My problem is probably a little esoteric, but I hope someone has tried this before and can help me. I'm trying to read in and use a VERY large HDF5 file. Previously when my datasets were smaller, I was using h5_parse with the /READ_DATA tag. SO convenient and helpful. But now I need to read in less than the full file. Not only is h5_parse taking over 7 minutes to complete, but having all that data in memory is forcing my subsequent rem_dups, sorts, and mask_locates to go into kswapd. Ugh.

So here's the question - can I use h5_parse without /READ_DATA to get the structure and then use the other h5 read commands on those structure IDs to extract the data? How would this look with an example h5d_read?

Subject: Re: h5_parse without READ_DATA
Posted by Michael Galloy on Mon, 03 Feb 2014 20:36:45 GMT
View Forum Message <> Reply to Message

On 2/3/14, 11:39 AM, simulana@gmail.com wrote:
> My problem is probably a little esoteric, but I hope someone has
> tried this before and can help me. I'm trying to read in and use a
> VERY large HDF5 file. Previously when my datasets were smaller, I
> was using h5_parse with the /READ_DATA tag. SO convenient and
> helpful. But now I need to read in less than the full file. Not
> only is h5_parse taking over 7 minutes to complete, but having all
> that data in memory is forcing my subsequent rem_dups, sorts, and
> mask_locates to go into kswapd. Ugh.
>
> So here's the question - can I use h5_parse without /READ_DATA to get
> the structure and then use the other h5 read commands on those
> structure IDs to extract the data? How would this look with an
> example h5d_read?
>

I have some routines to deal with HDF5 files that allow you to pull out
a single variable (or a part of a variable):

https://github.com/mgalloy/mglib/tree/master/src/hdf5

MG_H5_GETDATA should do what you want:

IDL> data = mg_h5_getdata(filename, 'path/to/variable')

Mike
--
Michael Galloy

www.michaelgalloy.com
Modern IDL: A Guide to IDL Programming (http://modernidl.idldev.com)
Research Mathematician
Tech-X Corporation