Subject: Reading part of a netcdf file?
Posted by rjp23 on Tue, 24 Sep 2013 15:26:51 GMT

View Forum Message <> Reply to Message

I have a very large (22GB) netcdf file and I was hoping there was some way to read part of it at a time into IDL but I can't figure out how.

For example, I have this variable:

Var   FLOAT   = Array[144, 72, 17, 32161]

The 4th dimension is time, i.e. the are 32161 timesteps.

Is it possible to read only the first X in one go, then read the next X, etc meaning that less is stored in memory at once?

Cheers

Rob

Subject: Re: Reading part of a netcdf file?
Posted by David Fanning on Tue, 24 Sep 2013 16:06:53 GMT

View Forum Message <> Reply to Message

rjp23@le.ac.uk writes:

> I have a very large (22GB) netcdf file and I was hoping there was some way to read part of it at
> a time into IDL but I can't figure out how.
>
> For example, I have this variable:
>
> Var   FLOAT   = Array[144, 72, 17, 32161]
>
> The 4th dimension is time, i.e. the are 32161 timesteps.
>
> Is it possible to read only the first X in one go, then read the next X, etc meaning that less is
stored in memory at once?

You might find the tools I have developed to browse, read, and write
netCDF files helpful:

   http://www.idlcoyote.com/fileio_tips/ncdf_browser.html

Cheers,

David
--

David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.idlcoyote.com/
Sepore ma de ni thue. ("Perhaps thou speakest truth.")

On Tuesday, September 24, 2013 5:06:53 PM UTC+1, David Fanning wrote:
> rjp23@le.ac.uk writes:
>
>
>
>> I have a very large (22GB) netcdf file and I was hoping there was some way to read part of it
at a time into IDL but I can't figure out how.
>
>>
>
>> For example, I have this variable:
>
>>
>
>> Var   FLOAT   = Array[144, 72, 17, 32161]
>
>>
>
>> The 4th dimension is time, i.e. the are 32161 timesteps.
>
>>
>
>> Is it possible to read only the first X in one go, then read the next X, etc meaning that less is
stored in memory at once?
>
>
>
> You might find the tools I have developed to browse, read, and write
>
> netCDF files helpful:
>
>
>
>   http://www.idlcoyote.com/fileio_tips/ncdf_browser.html
>
>
>
> Cheers,

>
>
>
> David
>
> --
>
> David Fanning, Ph.D.
>
> Fanning Software Consulting, Inc.
>
> Coyote's Guide to IDL Programming: http://www.idlcoyote.com/
>
> Sepore ma de ni thue. ("Perhaps thou speakest truth.")

Hi David,

Do you mean in reference to copying part of the file into another file or for just reading part of the file?

It looks like both might be possible once I get my head around your code...?

Cheers

Rob

Subject: Re: Reading part of a netcdf file?
Posted by David Fanning on Tue, 24 Sep 2013 16:31:28 GMT

rjp23@le.ac.uk writes:

> Do you mean in reference to copying part of the file into another file or for just reading part of the file?
>
> It looks like both might be possible once I get my head around your code...?

Well, I have to confess, I've never looked at a 22.5 GByte netCDF file before, but I suspect both are possible.

It *is* possible to read just a portion of a large data set from a netCDF file. See the GetVarData method in the NDCF_File object. See the COUNT, OFFSET, and STRIDE keywords. I can't advise sensibly without knowing more about the data, but I think it is possible to do what you want to do. You certainly don't want to read that data file all at once into the variable "a". ;-)

Cheers,

David

--
David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.idlcoyote.com/
Sepore ma de ni thue. ("Perhaps thou speakest truth.")

Subject: Re: Reading part of a netcdf file?
Posted by David Fanning on Tue, 24 Sep 2013 16:40:22 GMT

David Fanning writes:

> It *is* possible to read just a portion of a large data set from a
> netCDF file. See the GetVarData method in the NDCF_File object. See the
> COUNT, OFFSET, and STRIDE keywords. I can't advise sensibly without
> knowing more about the data, but I think it is possible to do what you
> want to do. You certainly don't want to read that data file all at once
> into the variable "a". ;-)

This article might be of interest to you, too:

   http://www.idlcoyote.com/fileio_tips/readslab.html

Cheers,

David
--
David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.idlcoyote.com/
Sepore ma de ni thue. ("Perhaps thou speakest truth.")

Subject: Re: Reading part of a netcdf file?
Posted by Paul Van Delst[1] on Tue, 24 Sep 2013 16:46:58 GMT

See:
   http://www.exelisvis.com/docs/NCDF_VARGET.html

Similar to what David said, the COUNT and OFFSET keywords are the ones

---

that should do the job for you (don't think you need STRIDE).

Just of the top of my head you'd need something like
  COUNT=[144, 72, 17, X]
and for the first read
  OFFSET=[0,0,0,0]
For the second read of X timesteps,
  OFFSET=[0,0,0,X]
third,
  OFFSET=[0,0,0,X*2]
etc..

I think I got the general idea right (anyone?). The X-related numbers
may be off by one since count is 1-based and offset is 0-based and I
always get 'em mixed up.

cheers,

paulv

On 09/24/13 11:26, rjp23@le.ac.uk wrote:
> I have a very large (22GB) netcdf file and I was hoping there was
> some way to read part of it at a time into IDL but I can't figure out
> how.
>
> For example, I have this variable:
>
> Var   FLOAT   = Array[144, 72, 17, 32161]
>
> The 4th dimension is time, i.e. the are 32161 timesteps.
>
> Is it possible to read only the first X in one go, then read the next
> X, etc meaning that less is stored in memory at once?
>
> Cheers
>
> Rob
>

Subject: Re: Reading part of a netcdf file?
Posted by Michael Galloy on Tue, 24 Sep 2013 20:01:42 GMT
View Forum Message <> Reply to Message

On 9/24/13 9:26 AM, rjp23@le.ac.uk wrote:
> I have a very large (22GB) netcdf file and I was hoping there was
> some way to read part of it at a time into IDL but I can't figure out

> how.
>
> For example, I have this variable:
>
> Var   FLOAT   = Array[144, 72, 17, 32161]
>
> The 4th dimension is time, i.e. the are 32161 timesteps.
>
> Is it possible to read only the first X in one go, then read the next
> X, etc meaning that less is stored in memory at once?
>
> Cheers
>
> Rob
>

You might want to try MG_NC_GETDATA. You can do stuff like:

    IDL> data = mg_nc_getdata(filename, 'var[*, *, *, 0:9]')

Get it from the mglib GitHub repo:

    https://github.com/mgalloy/mglib/blob/master/src/netcdf/mg_n c_getdata.pro

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