Subject: Re: Large data sets
Posted by Liam E. Gumley on Fri, 04 May 2001 13:47:27 GMT
View Forum Message <> Reply to Message

Neil Talsania wrote:
> I am new to IDL, and everything that I have done so far works great on
> small data sets. But, I dont know how to use it for a large (>2 gig)
> dataset, as I obviously just cant load that into an array. The data is on
> disk and can either be tiled so that it is stored in 1024x1024 segments, or
> it is stripped which is one complete row at a time. Even a complete row is
> large (>80000 samples).
>
> Any thoughts as to how I would be able to use this data?
>
> Another question: is there thread support in IDL? I have a machine with 18
> processors, so it sure would be nice to process this large dataset in 18
> parallel threads. Any suggestions?

The following may be of interest:

http://sag-www.ssl.berkeley.edu/~korpela/mmap/

Cheers,
Liam.
http://cimss.ssec.wisc.edu/~gumley/

Subject: Re: Large data sets
Posted by Paul van Delst on Fri, 04 May 2001 14:00:16 GMT
View Forum Message <> Reply to Message

Neil Talsania wrote:
>
> Hi,
> I am new to IDL, and everything that I have done so far works great on
> small data sets. But, I dont know how to use it for a large (>2 gig)
> dataset, as I obviously just cant load that into an array. The data is on
> disk and can either be tiled so that it is stored in 1024x1024 segments, or
> it is stripped which is one complete row at a time. Even a complete row is
> large (>80000 samples).
>
> Any thoughts as to how I would be able to use this data?
>
> Another question: is there thread support in IDL? I have a machine with 18
> processors, so it sure would be nice to process this large dataset in 18
> parallel threads. Any suggestions?

You work for kodak - can't you walk down the hall and ask the RSI/IDL engineers.....? :o)

Seriously, though. What errors (if any) are you getting? Some OS'es don't support files > 2Gig but if you have a machine with 18 processors I, uh, don't think that would be a problem. More info on what you wanted to do would be helpful. 80000 datapoints doesn't sound very big to me (I read in millions of double values daily on a regular little old linux box). Reading in cunks or strips at a time and processing it sequential sounds like a reasonable method - if only for a first cut (umless the data is interdependent somehow.)

paulv

--
Paul van Delst A little learning is a dangerous thing;
CIMSS @ NOAA/NCEP Drink deep, or taste not the Pierian spring;
Ph: (301)763-8000 x7274 There shallow draughts intoxicate the brain,
Fax:(301)763-8545 And drinking largely sobers us again.
Alexander Pope.

Subject: Re: Large data sets
Posted by Alex Schuster on Fri, 04 May 2001 14:12:33 GMT
View Forum Message <> Reply to Message

Neil Talsania wrote:

> I am new to IDL, and everything that I have done so far works great on
> small data sets. But, I dont know how to use it for a large (>2 gig)
> dataset, as I obviously just cant load that into an array. The data is on
> disk and can either be tiled so that it is stored in 1024x1024 segments, or
> it is stripped which is one complete row at a time. Even a complete row is
> large (>80000 samples).
>
> Any thoughts as to how I would be able to use this data?

Have a look at the ASSOC() function.

Alex
--
Alex Schuster Wonko@weird.cologne.de PGP Key available
alex@pet.mpin-koeln.mpg.de
