
Subject: Re: Huge datasets under Win2000

Posted by [Craig Markwardt](#) on Wed, 04 Jul 2001 17:37:28 GMT

[View Forum Message](#) <> [Reply to Message](#)

Joe Means <joe.means@orst.edu> writes:

> I have to read in over 13,000,000 observations [eventually over
> 33,000,000] with the structure
> dataShort = {dataShort, returnNum:0L, x:0D,y:0D,z:0D}
>
> At first, either when reading in or processing, I ran out of memory with
> a message about IDL not being able to allocate memory to make the
> array. So I read this into an Assoc variable which works fine. Now ,
> however, when I try to access parts of this data [dataS is the
> Associated structure array]:
> datSvX = (dataS.y-yl)/cellsize
> I get the error:
> % File expression not allowed in this context: DATAS.
>
> Any suggestions would be appreciated. It appears to me that Eric
> Korpela's VARRAY referred to in a thread earlier this week will not work
> on Win2000 Intel machines. I run Win2000 on a PC with 1Gb RAM and lots
> of hard disk space.

It looks like ultimately your 33,000,000 will just barely fit into
core. Thus, any strategy that involves reading the entire set of data
into memory at once seems doomed to failure. Also, recall that IDL is
fond of creating temporary arrays for expressions, so you will always
need a factor of 2-4 more memory in practice.

I have no direct experience with ASSOC or VARRAY.

However, I routinely process over 500 million events of a similar
type. The secret is to break things up into chunks, and process them
that way. Somehow I doubt that you really need all the data in memory
all the time (I hope). But if you do, then you probably need a huge
memory upgrade right now! :-)

Craig

--

Craig B. Markwardt, Ph.D. EMAIL: craigmnet@cow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response
