Subject: Speed penalty using START and COUNT with HDF_SD_GETDATA
Posted by Bob Fugate on Wed, 05 Sep 2001 03:20:27 GMT

I have a large number of 128x128 pixel arrays stored as SDS's in HDF files. Since I am only interested in a 32x32 subset of each array, I tried using the START and COUNT keywords to read only that part of the array I need --- thinking this would be faster and less taxing on memory. However, I learned today that it is much faster to read in the entire array. Here are the numbers:

8000 frames of 32x32 pixels x2 bytes/pixel in 85 seconds using START and COUNT, ~193KB/sec

8000 frames of 128x128 pixels x2 bytes/pixel in 10.5 seconds (not using START and COUNT), ~25 MB/sec, or 145 times faster.

By the way, if I read the whole array but use START and COUNT, there is no speed penalty, so the routine seems to know to get all the data.

I realise that reading data in larger chunks may be more efficient and I guess I can read in the whole array and discard the parts I don't need, but it is taxing my available memory, I am going to have to use some loops, and most importantly, why is reading the 32x32 subset so slow?

This is a so-so Windows NT machine; IDL 5.4. The data is on a server. I have a good connection to the server.

Anyone had any similar experiences or suggestions on how to speed up reading only the part of the array I need?

Thanks,
Bob Fugate