Subject: reading/writing large files
Posted by Russell Ryan on Fri, 01 Feb 2013 15:15:25 GMT

Okay gang I've been working on this for a few days and have given up.

I've got this simulation that outputs an array of floating point numbers (roughly 5000 or so), which I want to put into a file. If the file exists, I want to append to it; if not, I want to create it. I want to do this of order a million times (at least append of order a million times). When the simulation finishes, I want to read these numbers and do some post-processing. I don't want to read the entire file at once because I'm afraid I'll run into memory problems (especially since I can envision doing the appending 10^7 or even 10^8 times). So, instead I'd like to read say all 10^6 (or 10^7 or 10^8) trials of the k-th element of the array and get a single floating-point array of 10^6 elements (or what have you). Basically, I'm envisioning a table with say 5000ish columns but the number of rows is variable, and I want to read the k-th column.

Any ideas on the most efficient way of doing this? Obviously, my idea of a table is just illustrative and I don't actually care the format of the data in the file --- or even the file type. Currently, I'm opening a file with open and writing unformatted data to it with writeu. Then to read it, I create an assoc and loop over the rows. I've estimated it takes about 1.7e-4 s (on my 4 year-old laptop) so I'm guessing of order 3 minutes for 10^6. I can live with 3 min (if I have to), it just seems that given all the file types that IDL can read/write there's a way to do just this.

Any ideas?
Russell