Subject: Efficient reading of scattered data
Posted by eharold on Tue, 14 Jun 1994 18:38:15 GMT
View Forum Message <> Reply to Message

It seems that most of the time in my program is going into
reading individual complex numbers spaced widely across a file.
What I have is an unformatted file that contains a three-dimensional
array of complex numbers. x is along the fast axis, y along the slower
axis, and time along the slowest axis. I need 3-D FFTs but don't have enough
memory to just do it. So first I do 2-D space FFTs at each
time point. Then I need to do Fourier transforms
in time at each spatial position so I use point_lun to position
myself at each time_point I need and read the complex number there
into an array. I transform this array and write it back out to the file.
As I said it seems that most of the time in this program is in the reading
and writing needed for the time transforms.
Here's the code:

```
xform=complexarr(tdim)
csize = 8L ; size in bytes of a complex number
plane_size = xdim*ydim*csize
temp=complex(0,1)
printf,progress,"doing 320 (by 256) forward time transforms"
for x=xstart,xend do begin
 printf,progress,x
 for y=ystart,yend do begin
  rowsize = (y*xdim+x)*csize
  offset=(lindgen(tdim))*plane_size+rowsize
  for t=tstart,tend do begin
   point_lun, output, offset(t)
   readu, output, temp
   xform(t)=temp
  endfor
  xform=fft(xform,+1,/overwrite)
  for t=tstart,tend do begin
   point_lun, output, offset(t)
   writeu, output, xform(t)
  endfor
  offset=0 ; free offset
 endfor
 flush,output,progress
endfor
```

It has been suggested that I should read more than one complex number
at each point, and collect several time sequences before Fourier
transforming. Would this be substantially more efficient? The original
motivation was to read a full record at each step, but I think that suggestion

was motivated by the suggester's experience with Fortran. Perusals of the IDL Reference Manual indicate that the concept of record length has no meaning in UNIX or IDL.

The real questions are:

1. When I "readu, output, temp" does UNIX read more than the eight bytes required for the complex number?
2. Is substantial time taken by the point_lun procedure?
3. Any other suggestions?

You may wish to keep in mind that memory is a concern, and I can't fit the whole dataset into memory at once.

This runs under Solaris 2.3 on a four processor Sun Sparc. However I do not have exclusive use of the machine.

--

Elliotte Rusty Harold   National Solar Observatory
eharold@sunspot.noao.edu   Sunspot NM 88349