Subject: Re: Where vs Histogram vs ??
Posted by Stein Vidar Hagfors H[1] on Tue, 22 Oct 2002 15:58:41 GMT
View Forum Message <> Reply to Message

Andrew Cool <andrew.cool@dsto.defence.gov.au> writes:

> Hello All,
[...snip...]
> Roughly speaking, a third of the data is for any given WRF (waveform
> repetition frequency),
> a quarter is at any given frequency, and a quarter is in each of the 4
> possible beams.
> Or, for any given day, the data is split over 4 beams, and cycled over
> 4 frequencies and
> 3 WRF's.
[...snip...]
> start_year = 2000
> end_year = 2002
> start_day = 120
> end_day = 133
> start_half_hr = 0
> end_half_hr = 47
> WRF = 1
> FREQ = 2
> start_beam = 0
> end_beam = 3
> nominated_parameter = 2
>
> index = Where(!database.year GE start_year AND $
> !database.year LE end_year AND $
> !database.day GE start_day AND $
> !database.day LE end_day AND $
> !database.beam GE start_beam AND $
> !database.beam LE end_beam AND $
> !database.half_hr GE start_half_hr AND $
> !database.half_hr LE end_half_hr AND $
> !database.WRF EQ WRF AND $
> !database.FREQ EQ FREQ AND $
> !database.parameter(nominated_parameter) NE
> bad_data_value)
[...]

Given the above, could you perhaps try a multi-stage selection, e.g.,

```
wrf_ok = !database.WRF EQ WRF
wrf_freq_ok = !database.FREQ EQ FREQ AND temporary(wrf_ok)
;; By now you should have 1/12th of the data left!
;; Don't know how many bad_data_values you expect, the next one might
```

```
   ;; not gain much:
   wrf_freq_good =!database.parameter(nominated_parameter) NE bad_data_value $
                AND temporary(wrf_freq_ok)

   index1 = where(wrf_freq_good)

   ;; Build a new database on this subset (smaller than 1/12th),
   ;; continue with the rest of your searches...
```

Otherwise, I'd say that going from year/day/half_hr to Julian Day (modified to fit into a smaller data type, perhaps, by multiplying JD with 48 half-hours & subtracting earliest possible epoch?) is good advice, as is the multiple-array (instead of structure) approach.

However, as with many other problems of this type, the "killer" approach would be staying with a structure, using a DLM that goes through the data once, producing a single byte array with 0B/1B given input start/end times, beams, WRF, FREQ and Nominated-parameter!

There's no way IDL can optimize these statements the way a C programmer would do. Depending of course on the number of times you expect to do these selections over your project lifetime, I'd say writing a DLM may be a good investment of time!

--

 -- --------------
Stein Vidar Hagfors Haugan
ESA SOHO SOC/European Space Agency Science Operations Coordinator for SOHO

NASA Goddard Space Flight Center, Email: shaugan@esa.nascom.nasa.gov
Mail Code 682.3, Bld. 26, Room G-1, Tel.: 1-301-286-9028/240-354-6066
Greenbelt, Maryland 20771, USA. Fax: 1-301-286-0264
 -- --------------