Subject: Re: point_lun is slow
Posted by George McCabe on Wed, 27 Oct 1999 07:00:00 GMT
View Forum Message <> Reply to Message

thanks for the response, david and craig.

chunking is certainly much faster, and my algorythm is 'chunking' away
nicely.

in instances where a small number of the total data elements from the
file are required, the 'chicken pecking' approach is much faster. but
when in doubt, chunk.

do you have a rule of thumb for optimizing chunk size?

george

Subject: Re: point_lun is slow
Posted by Craig Markwardt on Wed, 27 Oct 1999 07:00:00 GMT
View Forum Message <> Reply to Message

George McCabe <george.mccabe@gsfc.nasa.gov> writes:
>
> reading from a data file at regularly spaced byte locations, 2 bytes at
> a time using point_lun - my program is abnormally slow. i don't have
> enough experience to guess whether the poor performance is inherent to
> point_lun & readu approach or if there are options which are affecting
> execution adversely.
>

As David mentions, you don't want to be doing lots of POINT_LUN calls
inside a large loop. I have found that one of best ways to do sparse
reads is to read a large chunk of data into memory, and then operate
on it from there. Who cares if you read a little too much data.
That's really what the underlying operating system has to do anyway,
but you can avoid doing lots of small READU calls.

buflen = 65536L
buffer = bytarr(buflen)
readu, unit, buffer, transfer_count=cc

You will have to do error checking here if cc is less than buflen.
You need to decide whether the entire file can fit in memory at once,
or if you need to do it in chunks.

At this point, you have a big chunk of memory and can operate on it
without doing any more reads. For example, if you want every other

byte, you could do something like this:

```
buffer = reform(buffer, 2, cc/2, /overwrite)  ;; This is fast!
result = buffer(0, *)                         ;; Gets the first of two bytes
```

By the way, this also answers the question of the fastest way to get ever other element of an array.

If you need to walk some more complicated data structure, that's harder. You probably won't be able to do that without a FOR loop. Hopefully your data file structure has a natural block size, and you can read a whole number of blocks at once.

I do chunking like this all the time, and get excellent boosts in performance.

Craig

P.S. I'm in Building 2 at Goddard. Stop by if you want. :-)

--

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

Subject: Re: point_lun is slow
Posted by George McCabe on Wed, 27 Oct 1999 07:00:00 GMT
View Forum Message <> Reply to Message

hi david,

i am reading individual detector array elements from frame sequential image data. the file is too large to be read completely into memory. i wrote three different algorythms to compare: one 'pages' through frames using assoc() and keeps groups of pixels, one chunks through groups of frames with point_lun & readu, and lastly, the algorythm in question, reads each 2 byte piece indiviually. the number of program steps in the last approach is much higher but, the amount of data read from disk is much smaller for select sets of elements. i did not have a way of estimating the speed of each part of these procedures a priori. if there are no subtleties to using point_lun in a unix file system that can change performance significantly (buffer sizes, ...?) the answer to my original question, what is faster, is known. what i was fishing for was information about idl and details related to accessing files which might be relevant to this task.

george

data David Fanning wrote:
>
> George McCabe (george.mccabe@gsfc.nasa.gov) writes:
>
>> reading from a data file at regularly spaced byte locations, 2 bytes at
>> a time using point_lun - my program is abnormally slow. i don't have
>> enough experience to guess whether the poor performance is inherent to
>> point_lun & readu approach or if there are options which are affecting
>> execution adversely.
>>
>> i'ld appreciate any thoughts on the topic which might lead to a
>> solution.
>
> Can you give us some idea about why in the world you
> are doing this!? :-)
>
> There is probably an easier (and MUCH faster) way
> if I had some idea what you were trying to do.
>
> Cheers,
>
> David
>
> P.S. You aren't doing this in a loop are you? :-)
>
> --
> David Fanning, Ph.D.
> Fanning Software Consulting
> Phone: 970-221-0438 E-Mail: davidf@dfanning.com
> Coyote's Guide to IDL Programming: http://www.dfanning.com/
> Toll-Free IDL Book Orders: 1-888-461-0155

Subject: Re: point_lun is slow
Posted by davidf on Wed, 27 Oct 1999 07:00:00 GMT
View Forum Message <> Reply to Message

George McCabe (george.mccabe@gsfc.nasa.gov) writes:

> reading from a data file at regularly spaced byte locations, 2 bytes at
> a time using point_lun - my program is abnormally slow. i don't have
> enough experience to guess whether the poor performance is inherent to
> point_lun & readu approach or if there are options which are affecting
> execution adversely.
>
> i'ld appreciate any thoughts on the topic which might lead to a

> solution.

Can you give us some idea about why in the world you
are doing this!? :-)

There is probably an easier (and MUCH faster) way
if I had some idea what you were trying to do.

Cheers,

David

P.S. You aren't doing this in a loop are you? :-)

--
David Fanning, Ph.D.
Fanning Software Consulting
Phone: 970-221-0438 E-Mail: davidf@dfanning.com
Coyote's Guide to IDL Programming: http://www.dfanning.com/
Toll-Free IDL Book Orders: 1-888-461-0155
