
Subject: Re: point_lun is slow

Posted by [Liam Gumley](#) on Thu, 28 Oct 1999 07:00:00 GMT

[View Forum Message](#) <> [Reply to Message](#)

George McCabe wrote:

> chunking is certainly much faster, and my algorithm 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.

I believe you can use the 'chunking' method in both cases with high efficiency. The key here is to access the disk in sequential order, with as few disk accesses as possible. I'm assuming that the goal is to read small (say 2 byte) sections of data from the disk at random locations.

The following pseudo-algorithm reads records (chunks) of data from the disk in sequential order. Only records that cover the specified read locations are actually read from disk. Each record is only read once.

```
Sort the array of read locations from lowest to highest
Set the record size to 512 bytes (you can experiment with record sizes)
Set the old record number to -1
Start a loop over the read locations
  For this read location, compute the record number in the file
  If the record number is different than the old record number
    Read the current record
    Set the old record number to the current record number
  End If
  For this read location, compute the byte offset within the record
  Extract data from the record at the byte offset
End Loop
```

This method should be just as efficient for small or large numbers of read locations.

Cheers,
Liam.

--

Liam E. Gumley
Space Science and Engineering Center, UW-Madison
<http://cimss.ssec.wisc.edu/~gumley>
