
Subject: Re: Read Total lines in an ASCII file
Posted by [Rick Towler](#) on Tue, 17 Dec 2002 22:32:27 GMT
[View Forum Message](#) <> [Reply to Message](#)

<wmc@bas.ac.uk> wrote
> Mark Hadfield <m.hadfield@niwa.co.nz> wrote:
>> Method 1, the two-pass method, is
>> surprisingly quick here, but suffers if the file is on a slow network
>> drive.
>
> If the read and re-read are close in time (as they should be) the file
> will still be in whatever cache the o/s uses, and the second read (as
> far as the o/s is concerned) will be much faster. On the ex-dec alphas
> I use, faster by a factor of 10 than the first. Whether putting IDL in
> between harms this gain I don't know. This may well survive being on a
> network drive (I'm assuming what you say above is a guess, rather than
> experience? Forgive me if not)

You are assuming that whatever file you are reading will fit into the cache which wouldn't always be the case. To complicate things you have the hard disk drive cache, possibly a cache in the controller, and a cache at the OS level. Each hardware/OS combination will yield different results. So it would be safe to say that YMWV with method 1.

Even if Mark was guessing about the performance of method 1 over the network it is a very safe assumption. Network latency, bandwidth limitations, and server load will always impact read/write performance. Again, the extent depends largely on the hardware/OS/network combination but since most of us aren't lucky enough to have a rack of 15,000 rpm ultra-160 SCSI drives running in RAID 0+1 networked via myriant we generally have to wait a little longer for files from the server than files stored locally ;)

-Rick
