Subject: Re: reading and writing very slow
Posted by oxfordenergyservices on Fri, 11 Feb 2011 17:06:39 GMT
View Forum Message <> Reply to Message

On Feb 11, 3:27 pm, Ben Tupper <ben.bigh...@gmail.com> wrote:
> On 2/11/11 10:16 AM, geoff wrote:
>
>
>
>
>
>
>
>
>
>
>> On Feb 11, 2:09 pm, Reimar Bauer<R.Ba...@fz-juelich.de> wrote:
>>> Am 11.02.2011 14:12, schrieb geoff:
>
>>>> Hi
>
>>>> I have some 1-2 GB text files (lots of them!), each containing weather
>>>> data for many thousands of stations for 1 year (per file). I need to
>>>> get the data out of the year files and into files which have all the
>>>> data for 1 weather station. It's easy but slow.
>
>>>> I am reading each year file line by line and appending that line to
>>>> the filename of the station (which is one of the fields on the line).
>>>> (openw.../append, close). Does opening and closing files so many
>>>> times have such an overhead? Is there a quicker way?
>
>>> no
>
>>> but reading line by line has.
>
>> only way i know how. variable length ascii unfortunately :(
>
> Hi,
>
> I can't tell from your description if exactly how you are managing the
> output process so this might not be all that helpful. If I had to do
> this in IDL, I would use something like Mike Galloy's resizeable array
> list. See...
>
> http://michaelgalloy.com/2006/04/24/collection-package-mgarr aylist.html
>
> I might use that array list to aggregate all of the data. When the
> aggregation is complete I would then dump it all to file at once.
>

> For each input file you can read in all the data at a swipe and then
> parse as needed within IDL. That would might be a lot faster than
> trying to read in formatted lines one-at-a-time.

Thanks for this I'll take a look. Turns out is was my fault.
Processing on one linux machine, filesystem on another.
