Subject: Re: Writing a very large file
Posted by David Fanning on Fri, 07 Dec 2007 20:23:35 GMT
View Forum Message <> Reply to Message

wlandsman writes:

> So the first four images take ~0.3s each to write, while subsequent
> images require more than 6 seconds each. I suspect that the slowing
> down is due to IDL (or the OS) needing to extend the file size. (I
> checked that it is not a memory usage problem.) So I think
> things would speed up if I could specify the final file size at the
> beginning -- perhaps there is a way to do this in Unix? I have
> experimented with the BUFSIZE and RAWIO keywords to OPENW but so far
> without any improvement.

I don't have any suggestions, but I wonder if disk fragmentation
might have something to do with it? Perhaps it is faster if
it can grab contiguous memory space.

Cheers,

David
--
David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.dfanning.com/
Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: Writing a very large file
Posted by pgrigis on Fri, 07 Dec 2007 20:39:03 GMT
View Forum Message <> Reply to Message

Hi Wayne, I don't see this effect on my Mac...

Ciao,
Paolo

```
0      0.47319698
1      0.86947393
2       1.3010280
3      0.88083601
4      0.90536904
5      0.87050986
6       1.1563799
7       1.3074460
8      0.86915302
9      0.90732503
```

10 0.87124896
11 1.3134921
12 0.87872696
13 0.88537407
14 1.2747021
15 0.88525605
16 0.87100291
17 0.89240909
18 1.3015611
19 0.89991498
20 0.88492012

wlandsman wrote:
> I am writing a sequence of images to a single very large file on my
> Linux system. I find that the processing dramatically slows down
> after the first few images. The simplified code looks like the
> following:
>
> pro test
> ; Display the time required to write a series of image to a single
> large file
> im = intarr(4096,4096)
> t = systime(1)
>
> close,1 & openw,1,'test.dat'
> for i=0,20 do begin
> writeu,1,im
> print,i,systime(1)-t & t = systime(1)
> endfor
>
> close,1
> return
>
>
> IDL> test
> 0 0.22054195
> 1 0.26708603
> 2 0.35127902
> 3 0.37285185
> 4 3.3877730
> 5 6.1666460
> 6 6.1697872
> 7 6.2481630
>
>
> So the first four images take ~0.3s each to write, while subsequent
> images require more than 6 seconds each. I suspect that the slowing

> down is due to IDL (or the OS) needing to extend the file size. (I
> checked that it is not a memory usage problem.) So I think
> things would speed up if I could specify the final file size at the
> beginning -- perhaps there is a way to do this in Unix? I have
> experimented with the BUFSIZE and RAWIO keywords to OPENW but so far
> without any improvement.
>
> Thanks for any suggestions, --Wayne

Subject: Re: Writing a very large file
Posted by Jean H. on Fri, 07 Dec 2007 21:23:37 GMT
View Forum Message <> Reply to Message

No problem on windows either...
I tried with a loop of 100, so the file would be 3.6Gb, and it was a
constant speed...
Jean

```
IDL> print,!version
{ x86 Win32 Windows Microsoft Windows 6.3 Mar 23 2006      32      64}

IDL> test
      0      0.64000010
      1      0.68700004
      2      0.67199993
      3      0.67199993
      4      0.68700004
      5      0.71900010
      6      0.70299983
      7      0.65700006
      8      0.65599990
      9      0.65600014
     10      0.75000000
     11      0.65599990
     12      0.89100003
     13      0.64100003
     14      0.67100000
     15      0.65700006
     16      0.68700004
     17      0.70299983
     18      0.86000013
     19      0.60899997
     20      0.64100003
```

Subject: Re: Writing a very large file

Posted by Jim Pendleton, ITT Vi on Fri, 07 Dec 2007 21:47:14 GMT
View Forum Message <> Reply to Message

"wlandsman" <wlandsman@gmail.com> wrote in message
news:f3c3ac4e-a5c2-4d57-a1f2-d0356d7a0d00@i12g2000prf.google groups.com...
> I am writing a sequence of images to a single very large file on my
> Linux system. I find that the processing dramatically slows down
> after the first few images. The simplified code looks like the
> following:
>
> pro test
> ; Display the time required to write a series of image to a single
> large file
> im = intarr(4096,4096)
> t = systime(1)
>
> close,1 & openw,1,'test.dat'
> for i=0,20 do begin
> writeu,1,im
> print,i,systime(1)-t & t = systime(1)
> endfor
>
> close,1
> return
>
>
> IDL> test
> 0 0.22054195
> 1 0.26708603
> 2 0.35127902
> 3 0.37285185
> 4 3.3877730
> 5 6.1666460
> 6 6.1697872
> 7 6.2481630
>
>
> So the first four images take ~0.3s each to write, while subsequent
> images require more than 6 seconds each. I suspect that the slowing
> down is due to IDL (or the OS) needing to extend the file size. (I
> checked that it is not a memory usage problem.) So I think
> things would speed up if I could specify the final file size at the
> beginning -- perhaps there is a way to do this in Unix? I have
> experimented with the BUFSIZE and RAWIO keywords to OPENW but so far
> without any improvement.
>
> Thanks for any suggestions, --Wayne

Wayne,

The POINT_LUN procedure can be used to define a file's
size at the outset. That is, you can OPENW a file then POINT_LUN
to define the maximum file size without having to write any data to
the file first. POINT_LUN to an offset of 0 to begin writing.
See also TRUNCATE_LUN, and the "Reading and Writing Very
Large Files" section of the on-line help.

Jim P.

Jim P.

Subject: Re: Writing a very large file
Posted by Kenneth Bowman on Sat, 08 Dec 2007 04:50:35 GMT
View Forum Message <> Reply to Message

On Dec 7, 2:17 pm, wlandsman <wlands...@gmail.com> wrote:
> I am writing a sequence of images to a single very large file on my
> Linux system. I find that the processing dramatically slows down
> after the first few images. The simplified code looks like the
> following:
>
> So the first four images take ~0.3s each to write, while subsequent
> images require more than 6 seconds each. I suspect that the slowing
> down is due to IDL (or the OS) needing to extend the file size. (I
> checked that it is not a memory usage problem.) So I think
> things would speed up if I could specify the final file size at the
> beginning -- perhaps there is a way to do this in Unix? I have
> experimented with the BUFSIZE and RAWIO keywords to OPENW but so far
> without any improvement.
>
> Thanks for any suggestions, --Wayne

When writing a sequential file like this there should be no need to
pre-allocate the file space.
What I suspect is happening is the following. The first few images go
into cache (memory).
Once the cache space is filled, it starts writing the files to disk,
which is a much slower process.
6 seconds for a 32 MB file isn't great (~5 MB/s), but maybe you have
older hardware or you are writing to
a network volume?

Ken Bowman

Subject: Re: Writing a very large file

Posted by Bringfried Stecklum on Sat, 08 Dec 2007 16:19:33 GMT
View Forum Message <> Reply to Message

Hi Wayne,

these are the results I get on my old laptop

0	0.19904590
1	0.20663810
2	0.20261288
3	0.20232391
4	0.20302701
5	0.20535898
6	0.21716094
7	0.20427299
8	0.20487905
9	0.21450901
10	1.1617432
11	0.22542715
12	0.21031618
13	0.21396208
14	0.21563792
15	0.21073008
16	0.22037601
17	0.66201711
18	0.50106001
19	3.6788890
20	1.0018129

So every now and then it takes longer. I noticed that even after the IDL procedure finished the disk is still busy for while which indicates that the buffer is being flushed to disk. Another thing is whether DMA/UDMA is enabled which you can check/set with hdparm. But I think you would have noticed that already since without DMA the system is awfully slow.

regards,

Bringfried

Subject: Re: Writing a very large file
Posted by wlandsman on Mon, 10 Dec 2007 15:24:24 GMT
View Forum Message <> Reply to Message

On Dec 7, 11:50 pm, KenBowman <k-bow...@tamu.edu> wrote:

> What I suspect is happening is the following. The first few images go
> into cache (memory).

> Once the cache space is filled, it starts writing the files to disk,
> which is a much slower process.
> 6 seconds for a 32 MB file isn't great (~5 MB/s), but maybe you have
> older hardware or you are writing to
> a network volume?

Thanks for this and to the others who responded. I do have old hardware, so my question should have been "how can the first few images be written so quickly?". And it does appear that the answer is that the images are being cached, so that the writing to the disk continues after the call to WRITEU is completed. For example, if I add a WAIT,6 between each call to WRITEU (so that the cache has time to be emptied) then the call to WRITEU always completes quickly (<0.3 s) no matter how big the file. --Wayne

---

Subject: Re: Writing a very large file
Posted by Kenneth Bowman on Mon, 10 Dec 2007 15:43:13 GMT
View Forum Message <> Reply to Message

In article <f169a807-d4a4-4e89-accf-7b07329b12ca@l16g2000hsf.googlegroups.com>,
wlandsman <wlandsman@gmail.com> wrote:

> On Dec 7, 11:50 pm, KenBowman <k-bow...@tamu.edu> wrote:
>
>> What I suspect is happening is the following. The first few images go
>> into cache (memory).
>> Once the cache space is filled, it starts writing the files to disk,
>> which is a much slower process.
>> 6 seconds for a 32 MB file isn't great (~5 MB/s), but maybe you have
>> older hardware or you are writing to
>> a network volume?
>
> Thanks for this and to the others who responded. I do have old
> hardware, so my question should have been "how can the first few
> images be written so quickly?".

Once the data has been written to the cache, which is managed by the OS, not by IDL, IDL thinks the write is complete and moves on to the next write.

Memory access is one to two orders of magnitude faster than disk access. (You have to wait for the platters to rotate, which is a relatively slow process, electronically speaking.)

Ken Bowman
