Subject: Re: lost data?
Posted by Paul van Delst on Thu, 10 May 2001 13:12:57 GMT
View Forum Message <> Reply to Message

src wrote:
>
> Is there a bug in IDL's Save/Restore command? I've just spent the last 18
> days running a Monte Carlo simulation to seemingly lose all my data. The
> problem occured when our license manager stopped responding (network
> problem) hence the IDL session running the simulation crashed. The MC
> code is designed to save results periodically as it runs (just in case
> this sort of thing happens). I've just tried:
>
> "restore, 'mc_file.sav', /Verbose"
>
> only to get:
>
> % RESTORE: IDL version 5.3 (linux, x86).
> % RESTORE: Truncated save file, restored as much as possible:
>
> That "resored as much as possible:" is in fact 0 (zero). Despite the file
> itself being 17 Mb! Some of my .sav files are a lot bigger than this, yet
> don't seem to have any problems. Is there anyway to recover this file, or
> prevent this happening again in the future?

I'm not trying to be facetious when I suggest using something other than an IDL save file for storing output. Say, netCDF? Or maybe Liam Gumley's binary data I/O tools for IDL. See

http://cimss.ssec.wisc.edu/~gumley/binarytools.html)

While I can't help you with your current dilemma, Craig Markwardt has posted about his tools to interrogate IDL save files. See

http://astrog.physics.wisc.edu/~craigm/idl/cmsave.html

good luck

paulv

--
Paul van Delst A little learning is a dangerous thing;
CIMSS @ NOAA/NCEP Drink deep, or taste not the Pierian spring;
Ph: (301)763-8000 x7274 There shallow draughts intoxicate the brain,
Fax:(301)763-8545 And drinking largely sobers us again.
Alexander Pope.

Subject: Re: lost data?
Posted by George N. White III on Thu, 10 May 2001 13:24:13 GMT
View Forum Message <> Reply to Message

On Thu, 10 May 2001, src wrote:

> Is there a bug in IDL's Save/Restore command? I've just spent the last 18
> days running a Monte Carlo simulation to seemingly lose all my data. The
> problem occured when our license manager stopped responding (network
> problem) hence the IDL session running the simulation crashed. The MC
> code is designed to save results periodically as it runs (just in case
> this sort of thing happens). I've just tried:
>
> "restore, 'mc_file.sav', /Verbose"
>
> only to get:
>
> % RESTORE: IDL version 5.3 (linux, x86).
> % RESTORE: Truncated save file, restored as much as possible:
>
> That "resored as much as possible:" is in fact 0 (zero). Despite the file
> itself being 17 Mb! Some of my .sav files are a lot bigger than this, yet
> don't seem to have any problems. Is there anyway to recover this file, or
> prevent this happening again in the future? I'm going to very upset to
> lose 18 days work...
>
> cheers,
> S

Some OS's (SGI Irix) have a checkpoint facility that works at the level of processes, and doesn't require support built in to the application. I know there has been some work on checkpointing for linux, as the same capability is required to migrate a process to a new node in some distributed processing systems. I don't know if there is anything you can use with IDL, but it is certainly worth a look.

We run IDL batch jobs on a compute server that almost never goes down (big UPS and generator), but some jobs want an X-server, so the users have been setting the DISPLAY variable to an X-server on a workstation that doesn't have generator power. The jobs die if the power is out too long for the UPS's that run the network and workstations.

The trouble with the things you do to try to improve reliability for long batch runs is that it is almost impossible to test all the things that can go wrong -- power failures, disks getting full or failing, network failures, etc. Do other people have similar cautionary tales? What changes were needed to make batch processing more robust?

--

George N. White III <gnw3@acm.org> Bedford Institute of Oceanography

Subject: Re: lost data?
Posted by Jaco van Gorkom on Thu, 10 May 2001 14:52:28 GMT
View Forum Message <> Reply to Message

George N. White III wrote in message ...
...
> Some OS's (SGI Irix) have a checkpoint facility that works at the level of
> processes, and doesn't require support built in to the application. I
> know there has been some work on checkpointing for linux, as the same
> capability is required to migrate a process to a new node in some
> distributed processing systems. I don't know if there is anything you can
> use with IDL, but it is certainly worth a look.
...

One question here, just for my understanding:
what kind of thing is checkpointing?

Jaco

Subject: Re: lost data?
Posted by Liam E. Gumley on Thu, 10 May 2001 14:55:01 GMT
View Forum Message <> Reply to Message

src wrote:
> Is there a bug in IDL's Save/Restore command? I've just spent the last 18
> days running a Monte Carlo simulation to seemingly lose all my data. The
> problem occured when our license manager stopped responding (network
> problem) hence the IDL session running the simulation crashed. The MC
> code is designed to save results periodically as it runs (just in case
> this sort of thing happens). I've just tried:
>
> "restore, 'mc_file.sav', /Verbose"
>
> only to get:
>
> % RESTORE: IDL version 5.3 (linux, x86).
> % RESTORE: Truncated save file, restored as much as possible:
>
> That "resored as much as possible:" is in fact 0 (zero). Despite the file
> itself being 17 Mb! Some of my .sav files are a lot bigger than this, yet
> don't seem to have any problems. Is there anyway to recover this file, or
> prevent this happening again in the future? I'm going to very upset to
> lose 18 days work...

I ran into the same problem a long time ago when I used to run long
FORTRAN jobs on a VAX/VMS system. My batch job wrote data to a netCDF
file as it ran, but if the batch job died, the file would be incomplete
and therefore unusable.

Fortunately, netCDF offers a mechanism that allows you to synchronize an
open file to disk. In IDL, it's done as follows:

ncdf_control, cdfid, /sync

where cdfid is the identifier of an open netCDF file. You will need to
create a netCDF file which has an unlimited dimension; see the example
for ncdf_vardef in the online help for an example.

Another option, which I have not tested, is to write the results from
your simulation to a binary output file (not a SAVE file), and then
periodically execute the flush procedure, e.g.,

flush, lun

where lun is the logical unit number of the file to be synchronized to
disk.

Cheers,
Liam.
http://cimss.ssec.wisc.edu/~gumley/

Subject: Re: lost data?
Posted by George N. White III on Thu, 10 May 2001 16:15:32 GMT
View Forum Message <> Reply to Message

On Thu, 10 May 2001, Jaco van Gorkom wrote:

>
> George N. White III wrote in message ...
> ...
>> Some OS's (SGI Irix) have a checkpoint facility that works at the level of
>> processes, and doesn't require support built in to the application. I
>> know there has been some work on checkpointing for linux, as the same
>> capability is required to migrate a process to a new node in some
>> distributed processing systems. I don't know if there is anything you can
>> use with IDL, but it is certainly worth a look.
> ...
>
> One question here, just for my understanding:
> what kind of thing is checkpointing?

>
> Jaco

Checkpointing saves an "image" of a running program in such a way
that you can restart the program at that point in the event of a
crash. For Linux, see:

http://www.cs.rutgers.edu/~edpin/epckpt/

You can write applications that save their state, but if you have
OS-level support you can checkpoint applications. Checkpointing
does impose some constraints on the application, and you may need
lots of disk space to hold the checkpoint files.

--
George N. White III <gnw3@acm.org> Bedford Institute of Oceanography

---

Subject: Re: lost data?
Posted by nmw on Mon, 14 May 2001 09:19:12 GMT
View Forum Message <> Reply to Message

In article <Pine.SGI.4.33.0105101003100.24771-100000@wendigo.bio.dfo.ca>, "George N. White III" <WhiteG@dfo-mpo.gc.ca> writes:

>
> Some OS's (SGI Irix) have a checkpoint facility that works at the level of
> processes, and doesn't require support built in to the application. I
> know there has been some work on checkpointing for linux, as the same
> capability is required to migrate a process to a new node in some
> distributed processing systems. I don't know if there is anything you can
> use with IDL, but it is certainly worth a look.
>

Do you know if this will actually work with IRIX?

I was under the impression that jobs which required a license manager
could not be checkpointed (or more correctly, they could not re-aquire
the connection to the license manager upon restart). The sgi license
manager has been modified to take account of this, but other products
which don't use the sgi license manager (such as IDL) cannot be
successfully checkpointed.

--

Nigel Wade, System Administrator, Space Plasma Physics Group,
          University of Leicester, Leicester, LE1 7RH, UK
E-mail :   nmw@ion.le.ac.uk

Phone : +44 (0)116 2523568, Fax : +44 (0)116 2523555