

Re: using mmap on large (> 2 Gig) files

Source: <http://coding.derkeiler.com/Archive/Python/comp.lang.python/2006-10/msg04419.html>

- *From:* "Travis E. Oliphant" <oliphant.travis@xxxxxxxx>
 - *Date:* Mon, 23 Oct 2006 17:57:31 -0600
-

Martin v. Löwis wrote:

myeates@xxxxxxxxxxxxx schrieb:

Anyone ever done this? It looks like Python2.4 won't take a length arg

2 Gig since its not seen as an int.

What architecture are you on? On a 32-bit architecture, it's likely impossible to map in 2GiB, anyway (since it likely won't fit into the available address space).

On a 64-bit architecture, this is a known limitation of Python 2.4: you can't have containers with more than 2Gi items. This limitation was removed in Python 2.5, so I recommend to upgrade. Notice that the code has seen little testing, due to lack of proper hardware,

NumPy uses the mmap object and I saw a paper at SciPy 2006 that used Python 2.5 + mmap + numpy to do some pretty nice and relatively fast manipulations of very large data sets.

So, the very useful changes by Martin have seen more testing than he is probably aware of.

-Travis

.