
Subject: Re: comparing and concatenating arrays...please help!!
Posted by [Pepijn Kenter](#) on Thu, 08 Jan 2004 22:57:03 GMT
[View Forum Message](#) <> [Reply to Message](#)

>
> My files are _huge_. File 1 is about 250000 lines and file2 about half
> that. which is why my code was taking about 4 days!

You also might want to consider using another programming language. IDL is designed for doing (scientific) calculations, not for a task like this. Loop constructs are particularly slow in IDL.

I can imagine that if you're in a hurry, you start programming in whatever language you're most familiar with (I'm also lazy in that respect ;-). But if you've got the time, a project like this is a good opportunity to learn a new language. IMHO, the only reason for using IDL for a task like this would be when the input is generated by another IDL program (or the output is used by an IDL program). In that case the programs could be integrated tightly if necessary.

Languages designed for processing text files are e.g. perl and awk. If speed is really important you can use C for everything; but I guess it would take a lot of effort to beat a specialized language, so better try these first.

And like I said before, the sort program is very fast in sorting text files. I've just generated a 480000 line test file. It took 11 seconds to sort on an atlon 1200.

>
> I'll give the sorting a try, but if anyone else has any suggestions,
> they'll be gratefully received. I'm not a newbie, but am daunted by
> some of the operations you suggested below... I certainly take
> inspiration from all you guys though!
>

If you want to read some more about computational complexity, check out:

http://en2.wikipedia.org/wiki/Computational_complexity_theor_y
<http://users.forthnet.gr/ath/kimon/CC/CCC1b.htm>

Pepijn.
