Subject: Re: The fastest way to extract some points in million poits data set
Posted by David Fanning on Tue, 20 Dec 2005 13:49:01 GMT
View Forum Message <> Reply to Message

Park writes:

> When I use a where function,
> in my computer, if data have one million points it takes 0.01
> sec(order).
> It makes me depressed because I have to do that step so many times.

I'd try Zoloft. It won't speed the calculation up, but
you will certainly feel better about it. :-)

Cheers,

David

--
David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.dfanning.com/

Subject: Re: The fastest way to extract some points in million poits data set
Posted by Craig Markwardt on Wed, 21 Dec 2005 11:45:22 GMT
View Forum Message <> Reply to Message

"Park" <snfinder@naver.com> writes:

> HI~ ALL^^,
>
>
> I usually use where function on a lot of situations.
> It gives me easily the indices of array that I want among data set.
>
> BUT, it is a little slow I think.
>
> When I use a where function,
> in my computer, if data have one million points it takes 0.01
> sec(order).
> It makes me depressed because I have to do that step so many times.

"So many" is such a vague concept. Even if you have to do that
operation a million times, it takes just a few hours. That is not too
depressing!

> Is there a faster method to extract the points which satisfy some

> condition
> in points data set(3D)?
> I need the indices of points that satisfy specified conditions.
> (For example, specified conditions: inside the particular cube, inside
> the sphere or etc.)

I think you might want to take a different approach. Rather than looking for a faster version of WHERE, you should be looking for better ways to precondition your data.

If you can pre-arrange your array, you may be able to speed up the individual searches by excluding some parts of phase space. For example, perhaps you can sort the array by one coordinate. Then you can restrict your search to only those points that have the right initial coordinate value. This can be made even faster if you pre-bin the data with HISTOGRAM and use the REVERSE_INDICES keyword. If that is not enough, the same approach can be expanded with some difficulty to multiple dimensions.

Good luck,
Craig

--

Craig B. Markwardt, Ph.D. EMAIL: craigmnet@REMOVEcow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response