Subject: Re: h5_parse() in the profiler
Posted by Jim Pendleton on Sat, 09 Jul 2016 14:05:11 GMT
View Forum Message <> Reply to Message

On Friday, July 8, 2016 at 10:38:22 PM UTC-6, Edward Hyer wrote:
> On Friday, July 8, 2016 at 10:47:02 AM UTC-7, Markus Schmassmann wrote:
>> for i=1,n-1 do temp=create_struct(temp,tagname[i],tagvalue[i])
>> terribly inefficient, better to create a string and then use
>> execute(string)
>
> Hmmm... Yes! EXECUTE() is a non-starter, this needs to be fully usable in compiled code. But I'm sure there is some clever way to do this with fewer calls to CREATE_STRUCT().
> If I come up with something that actually is faster, I'll post to this thread.

Have you tried returning an ordered hash instead?

Each structure array in IDL represents a chunk of contiguous memory. That is, each of the consecutive tags in the structure is consecutive in memory, with some redirection for items such as strings. The nested calls to CREATE_STRUCT will be much like an array append operation, a = [a, newstuff], which can become quite inefficient for large arrays due to the need to make a new copy of the data at each iteration.

By using the /ORDEREDHASH keyword to H5_READ (added in 2014), the storage of the individual values is not restricted to contiguous memory and the overhead of recursive copying is no longer present.

Jim P.