Subject: Re: HDF SDS array access in IDL
Posted by William Clodius on Wed, 28 Oct 1998 08:00:00 GMT
View Forum Message <> Reply to Message

Dr. G. Scott Lett wrote:
> <snip>
> IDL uses a storage scheme like C, not Fortran. The question of
> column-majority and row-majority was a source of confusion at RSI for some
> time, and this confusion was evident in some of their documentation. Recent
> editions should be better.
> <snip>

Technically what we are talking about is an access scheme not a storeage scheme. Storeage per se is typically defined by the OS/processor, and the language only defines a mapping to that storeage. For multidimensional arrays, Fortran promises that access is most efficient if the left most index varies most rapidly. C promises that access is most efficient if the right most index varies most rapidly. In two dimensional array the first index is commonly termed the column index, the second index the row index. In Fortran, and IDL, varying the column index and keeping the row constant is most efficient, in C varying the row index and keeping the column constant is most efficient. Fortran's access mode is commonly refered to as column major, I have never hear the term row order.

IDL's (5.1) Help is confusing on the issue. For Arrays and Matrices it states

"In a computer, a multidimensional data set ... can be indexed in column-major format, which means that the linear order of the data elements proceeds from the first element in the first column through the last element in the first column before beginning on the second column, and so on. This is the format used by FORTRAN to index data."

so far correct

"Alternatively, data can be indexed in row-major format, meaning that the linear order of the data elements proceeds from the first element of the first row through the last element of the first row before beginning on the second row, and so on. This is the format used by the C and Pascal computer languages, and is traditionally associated with image processing, because it keeps all the elements of a single image scan line together. Because IDL's origins are in image processing, it indexes data in row-major format."

Almost all languages keep scan lines together and this naming convention has no impact on image processing. They just access them together in different indices, C groups lines by the leftmost index Fortran by the

rightmost. C may be more commonly used for image processing for a variety of reasons (more often taught at universities, easier access to hardware, structs have more flexibility than what is available in F77, wide availability of cheap compilers), but this is not one of them.

"NoteMany computer languagesï¿½e.g. C, Pascal, Visual Basic, and Fortanï¿½index 2-dimensional arrays in (row, column) order. IDL indexes arrays in (column, row) order."

Note the above is Humpty Dumpty. The C and Fortran standards do not define which index is associated with the row and which with the column so the definition should follow other conventions. Historically, the convention in column major versus row major is that in column major the first dimension elements are contiguous and in row major the last index elements are contiguous. Under that convention, Fortran and IDL are both column major, C and Pascal are row major. IDL is trying to make up a confusing convention to make itself sound as though it is more like C than Fortran. See for example

http://member.aol.com/CORLISS100/chapter5.html
http://webster.ucr.edu/Page_asm/ArtofAssembly/CH05/CH05-2.ht ml
http://www.cs.uregina.ca/dept/manuals/Manuals/3_0NumRep/3_Nu m_Rep.html

(Note because of the current limitations of C arrays I suspect that IDL's arrays are actually implemented as C pointers with indexing similar to F2C's.)

--

William B. Clodius Phone: (505)-665-9370
Los Alamos Nat. Lab., NIS-2 FAX: (505)-667-3815
PO Box 1663, MS-C323 Group office: (505)-667-5776
Los Alamos, NM 87545 Email: wclodius@lanl.gov