Factors (tradeoffs) to consider for SAS Program Efficiency: Need to Prioritize Factors

Computer Processing Unit (CPU)	Main Memory	Input/Output (I/O)	Data Storage Device	Programming Time
Measured in Seconds: Program time	**The maximum number of bytes of actual main memory the compilation and execution of the program used**	**Number of times a program requests a read/write operation between main memory and a data storage device**	**Physical amount of space required to store and process data sets**	**Time to develop, test, document and maintain programs**
- One vs. Multiple Pass? - Sorting/Indexing? Proc sort – cpu intense, physical order, requires 2.5 times size of input file Index – faster, logical order, need to be maintained, use on few variables, storage is about 10% of input file, best if data does not change often and expected to select no more than 25% of all records - Logic and priority conditioning to process fewer records? - Libname engines to directly access databases?	- Keep/Drop variables? - Summary/Partitioned/ Snapshot data sets? - System option settings (BUFNO=, BUFSIZE=, MEMSIZE=, CATCACHE=, SORTSIZE 8m, NOMACRO, SASFILE) - SAS version 9 (THREAD option) - Assign large space to WORK library - Customize config or autoexec file.	- Keep/Drop variables? - Subset records using where/if? - Data set options/Data set statements? - Metadata info only to get variables and record counts – no need to read data set?	- Compress data sets to take less space? Most useful for short and wide data sets with character variables instead of long and narrow data sets. Can not be used with POINT= option. Requires SAS to compress & uncompress. - Variable length reduction, for example, to 3, for numeric variables? - Track stopping point when writing to large data sets? - Delete temporary data sets? - Use data view?	- Proc SQL to do multiple steps in one? - Check syntax only?- Min Proc Sort? - Class statement? - Format catalog? - Documentation? - Use random sample when possible? - Code for unknown and missing data values to avoid, for example, division by zero? - Assign descriptive and meaningful variable names? - Use the right SAS code for the task, for example to copy only the data set structure?

Copyright © 2011 by Gupta Programming, See more SAS papers and references at: http://www.SASSavvy.com. Become a member today.

Copyright © 2011 by Gupta Programming, See more SAS papers and references at: http://www.SASSavvy.com. Become a member today.