NEW TECHNOLOGY - TRANSFORMING A MAINFRAME SYSTEM TO THE PC USING SAS/AF® SOFTWARE

Henry F. DeLong Jr., Valerie J. Martin, Beth A. Schaline,
Pinkerton Computer Consultants, Inc.

ABSTRACT

The PC has become an almost indispensable tool in the workplace today. It has only been recently, however, that PCs have become large and/or powerful enough to take over processing tasks previously slated for the less convenient mainframe environments. This paper will outline the steps necessary to successfully convert a mainframe processing system to a PC environment capable of handling data from a variety of sources including ASCII, EBCDIC, dBASE®, and Lotus 1-2-3®.

Certain questions need to be answered whenever changes are made to an existing production system.

- Will the changes save valuable resources?
- Will the changes enhance the reliability of the data?
- How will the user adapt to the changes?

These questions become even more important since our available technology is changing so rapidly. Examples of converted code will be given as well as suggestions for integrating the existing procedures into a menu-driven system using SAS/AF®.

INTRODUCTION

In making the decision to transform an existing mainframe based system to a PC based system, one must consider cost-effectiveness, data reliability, and user friendliness. Efficient use of resources is paramount to the successful operation of any data processing environment and must be incorporated into this decision. It must means the PC can save valuable resources for an operation in a short period of time, thus making it a viable alternative to the mainframe.

Will the changes save valuable resources?

First, it is important to assess the cost-effectiveness of converting the system from the mainframe environment to the PC.

In many instances there will be significant savings in cost and processing time that will be realized after the conversion is complete. The mainframe, though fast and powerful, is an expensive data processing tool. The company is faced with either supporting the expense of a mainframe (and the staff and maintenance that goes along with it) or leasing CPU time from someone else with a mainframe. In the case of the former, some of the expense of supporting the mainframe (e.g., tapes, disk space, personnel) can be reduced if the transition to the PC is made. In the case of the latter, substantial savings can be seen over a period of time if monies used to lease CPU time and data storage were to be invested in PC equipment.

Recently the PC has become a manageable investment for most organizations. Due to their decreasing cost and improved performance, PCs have been found in increasing numbers in the workplace. Although it is still more feasible to run some applications on the mainframe due to size and processing time, a PC should be able to handle most applications. In many instances PCs are used as mainframe terminals thus eliminating the need to buy all new equipment.

Data storage is also a consideration in determining whether it is cost effective to convert an existing production system from the mainframe to the PC. With the introduction of networks and powerful, inexpensive file servers, large capacity data storage devices can be delegated to one machine thereby eliminating the expense of high capacity drives on individual machines and the needless duplication of data files.

The SAS® system for the PC is similar to the mainframe version and can be utilized with minimal instruction, so existing personnel can easily make the transition to PC with little training. Also, with the incorporation of the SAS/AF menu system, even a novice computer user can be quickly trained to use the data processing system. The user will be able to handle all of the data on their own PC thereby eliminating time spent waiting for tapes to be logged-in, for jobs to be run, and for files to be uploaded and downloaded. This ease of training results in a quick transition and tangible savings in time and money.

Will the changes enhance the reliability of the data?

Data reliability is a primary consideration for any processing system. The transition from mainframe to PC will enhance data reliability because of the increase of control that the user can exercise over it. Data residing on the disk can be easily edited and maintained with little danger of loss because it cuts out the processes of uploading and downloading. Also SAS/AF provides a menu system that will allow the user to go on line and edit existing data without having to go into another PC software tool.

How will the user adapt to the changes?

A PC data processing system with the front-end consisting of a series of menus programmed using SAS/AF will make the conversion much easier for users. These menus will allow the user to choose the data processing function they require. By taking full advantage of the SAS PC system and the PC-DOS operating system, tasks can be simplified by integrating peripheral devices and software packages to input data, process it, and produce reports with a minimal amount of training.

Also, as mentioned above, the SAS system for the PC is similar to the mainframe version so it can be used by existing personnel with limited instruction. Users will easily
adjust to the changes because the processing is still done in the SAS system. This ease of training results in tangible savings in time and money.

**Components of the System**

Our SAS/AF menu-based PC system consists of three main parts:

- Data Input
- Data Update
- Data Output/End of Processing (EOP) Reports

In most cases it will be desirable to keep as much of the existing mainframe code as possible. Separate the mainframe programs into these three groups. This systematic approach will ensure that existing code is used whenever possible.

Review the programs in the data input group. It will be necessary to give some thought to the differences in data input for the PC as compared to the mainframe. Data submission on diskette will be much easier to read into the system, obviously, as it will not be necessary to upload the data to the mainframe. Tape data submission may be handled in different ways. We have decided to utilize a tape drive device to copy the data from the tape directly to the hard disk of the PC.

Most of the programs should be able to remain intact. We used PROC DOWNLOAD to transfer the programs from the mainframe to the PC. Delete all of the JCL statements from the beginning of the programs and replace them with LIBNAME or FILENAME statements. It is advisable to add the "RUN;" statement at the end of each data step and proc step before attempting to use the programs with PC SAS. At this point we suggest running the programs in a test environment on the PC where you plan to do your processing. You will be able to document all of the occurrences where changes may be necessary to successfully complete the conversion.

Some of the errors, notes, or warnings you may see include notes for uninitialized variables, errors for formats that have been defined on the mainframe that will have to be redefined for the PC environment, insufficient memory, as well as errors for specific procedures that are written for the mainframe that are not executable on the PC. It seems obvious that the variable names in the input datasets must have the same naming conventions on the PC as they do on the mainframe in order for the programs to get the same results. However, you should scan the log for notes indicating that there are uninitialized variables as they may be a symptom of a larger problem.

The use of formats makes report generation an easier task, but they must be redefined and stored in a catalog to execute on the PC.

An interesting problem that you may encounter if your program contains PROC SORT on a large dataset is an error for insufficient memory.

**PROC SORT DATA=NEW;**
**BY STATE;**
**RUN;**
**ERROR: Out of Memory.**

**NOTE:** The data set WORK.NEW has 0 observations and 0 variables.

**NOTE:** Data set WORK.NEW was not written because of NOREPLACE option.

This does not mean that converting your system to the PC is impossible. Sorting a dataset, especially a large one with many variables, takes up considerable space and memory because there must be enough work space to accommodate the entire file and sufficient space to output the dataset. This problem may be solved by using the TAGSORT option on PROC SORT. Enough workspace is needed for key variables only, not the whole record. Employing this feature may cause an increase in CPU time, since it is necessary to make an additional pass over the data at the end of the sort to retrieve the rest of the record before outputting it to the new dataset.

**PROC SORT DATA=NEW TAGSORT;**
**BY STATE;**
**RUN;**

You may also encounter an error for invalid length specification if you have length statements in your program.

**DATA NOCURR NOPREV PASS1;**
**LENGTH ID1 2;**
**ERROR: Invalid length specification 2.**

Remember that on some operating systems the smallest length specification allowed for numeric variables using PC SAS is 3, therefore a length specification of 2 may be invalid.

If a specific procedure used on the mainframe does not execute on the PC you will need to do some research to determine if a similar procedure has been written for the PC. Try the SUGI Proceedings to see if anyone has already done the conversion. If not, it will be necessary to determine the feasibility of writing the code to execute the procedure.

**PROC DATACHK DATA=NEW;**
**ERROR: PROCEDURE DATACHK not found.**
**VAR VARDI-NUMERIC-VAR5;**
**TITLE 'STATISTICAL RESULTS FOR NEW DATA';**
**RUN;**

In this example, we were not able to locate a PC version of PROC DATACHK, which is documented in the SUGI Supplemental Library User's Guide.

The printed output from DATACHK consists of a count of the non-missing observations, the five largest values, the five smallest values, and the number of observations with six special missing values. A through .E. The Guide directs you to refer to PROC UNIVARIATE. UNIVARIATE produces simple descriptive statistics, but does not print a summary of many variables on one page. In our specific case, we reviewed the output generated by PROC DATACHK and decided that a PROC UNIVARIATE would suffice. Therefore, we replaced our mainframe code with the following:

```sas
PROC UNIVARIATE DATA=WORK.NEW;  
BY STATE;  
RUN;  
```

ERROR: PROCEDURE UNIVARIATE not found.

In this example, we were not able to locate a PC version of PROC UNIVARIATE, which is documented in the SUGI Supplemental Library User's Guide.

The printed output from PROC UNIVARIATE consists of a count of the non-missing observations, the five largest values, the five smallest values, and the number of observations with six special missing values. A through .E. The Guide directs you to refer to PROC UNIVARIATE. UNIVARIATE produces simple descriptive statistics, but does not print a summary of many variables on one page. In our specific case, we reviewed the output generated by PROC DATACHK and decided that a PROC UNIVARIATE would suffice. Therefore, we replaced our mainframe code with the following:
Although this will produce significantly more output, we decided that we needed to take a closer look at the data. Other alternatives are available such as PROC FREQ.

Building the Menus with SAS/AF

People often judge PC software according to how user friendly it is. SAS/AF enhances this component of any PC based system. The capabilities of the menus that are built into a processing system that has been converted from mainframe SAS to PC SAS will have a substantial impact on the overall acceptance of the final product. If you are able to show that converting the system from mainframe to PC environment not only is a cost effective endeavor, but will also enhance the interactive capabilities of the end user, then you will have indeed targeted a likely candidate for conversion. Therefore we will highlight some of the SAS/AF menu capabilities that will help you promote mainframe to PC system conversions.

**SYSTEM FLOW**

**MAINFRAME SYSTEM FLOW**

- **INPUT**
  - **REPORTS**
- **UPDATE**
  - **REPORTS**
- **OUTPUT**
  - **REPORTS**

**P.C. SYSTEM FLOW**

- **MAIN MENU**
  - **INPUT**
  - **UPDATE**
  - **OUTPUT**
  - **REPORTS**

**DATA INPUT**

By selecting option "1" from the Main Menu, the user is presented with the Data Input Menu as can be seen in figure 2. Depending on the option selected under this menu, the data is introduced into the system using a SAS/AF SCL (Screen Control Language) program linking the PC’s peripherals to the SAS system. The utilization of SCL under this menu system allows automation of data input by diskette (figure 3), tape (figure 4), or manual entry in formats which include SAS, dBASE, and Lotus. It may also be beneficial to place help screens in the application to help the user as can be seen in figure 5. After the data is input into the system, it is automatically converted into a SAS PC dataset for editing purposes. An example of some code to do this can be seen in figure 6. Using predetermined criteria for data reliability, the SAS dataset is checked for errors within the specified range. Error-free observations are output into a "clean" dataset; records containing errors are placed into a separate dataset and an error report pinpointing the observations and variables involved is produced.
DATA UPDATE

By selecting option "2" from the main menu, the user is presented with the Data Update menu as can be seen in figure 7. From this screen, the user can take advantage of the observation functions available in a SCL program to make necessary corrections to any errors that were found in the editing process. This allows data correction at the observation level, thereby speeding up the editing process.
DATA OUTPUT

By selecting option "3" from the main menu, the user is presented with 3 options as can be seen in figure 8. Data Output option "1" allows the edited and error-free data to be merged into the master dataset. With the master dataset created the data can then be released in a format and on a media of the users choosing, using the options available under Data Output menu "2". Again, using the SCL language allows the automation of these processes in a fashion similar to the Data Input option available in the main menu. Data Output option "3" allows final customized reports to be processed using the SAS/AF print capabilities.

CONCLUSION

Converting existing mainframe code so that it will run successfully in a PC environment may be an attractive option. Minimal expense, increased data control, the ease of conversion, and the possibility of more user-friendly systems using SAS/AF make using new technology a more manageable task. We have discussed the major components of a data processing system as it exists on the mainframe and how it could be converted to a PC environment. Examples of code that needed modification for the conversion were provided as well as sample menu screens and SAS/AF code that read various input files into the PC system and converted them to compatible SAS datasets. We have only touched on the major areas one would need to consider when transforming a mainframe system to the PC, however our objective was to provide a general overview of the steps necessary. In utilizing the versatility of the SAS/AF product, the user is presented with many options to transform numerous mainframe applications into a customized, integrated, user friendly application on the PC.

ACKNOWLEDGMENTS

SAS and SAS/AF are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.

Other brand and product names are registered trademarks or trademarks of their respective companies.

FOOTNOTES:

1. DATACHK is supported by the author, not by SAS Institute Inc. Author: Dr. Frank E. Harrell, Jr. Duke University.