

SAS® Software on the IBM RISC System/6000® Workstation – The Wave of the Future for Information Delivery

Darius S. Baer, SAS Institute Inc.

Abstract

After years of using the **SAS® System** for a wide variety of applications on both MVS and VM mainframe platforms, many SAS users may feel somewhat intimidated by moving to the **IBM RISC System/6000®** (RS/6000) workstation environment. Actually, once the users become somewhat familiar with the **AIX®** operating system (IBM's UNIX offering), they will feel rather comfortable. We discovered that running the SAS System on the workstation was not only similar in many ways to running on a mainframe, but it was actually easier. The availability of the graphical interface and locally attached terminal made the SAS experience both more interesting and enjoyable as well as faster in most cases.

The AIX environment offers unique opportunities and challenges for an organization heavily dependent upon the mainframe environment. Historically, the Burlington Northern Railroad has developed SAS applications solely on the mainframe. PC's were used only by specialized users, and had few allies in the world of corporate data processing. The arrival of the AIX based work station forced the company to rethink much of its data processing practices and generated a variety of possibilities for creative minds.

This paper will outline the migration path for a mainframe enduser who is moving to a workstation. Required AIX knowledge will be specified and the user will be advised of any quirks and pitfalls that may arise when moving the code as well as the environmental differences. Detailed benchmark comparisons of a few applications that have been migrated will be presented as well as some first-hand explanation of the Burlington Northern Railroad experience. Finally, the use of the mainframe and workstation together in a coupled environment will be discussed. The multi-platform ease-of-use of the SAS product is the message to be gotten from this paper. The strengths and weaknesses of the SAS System on the mainframe and workstation will be presented so that the SAS user can make a better informed decision concerning which platform to run their application on and how to use the SAS System effectively on both platforms.

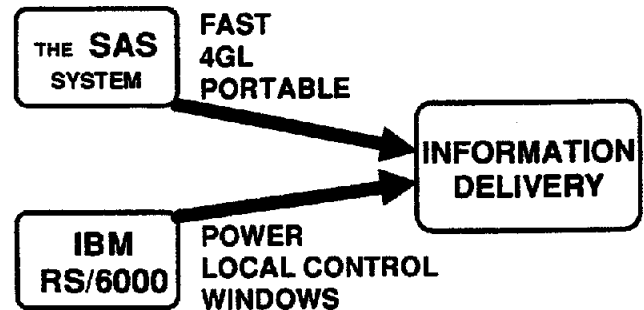


Figure 1. New Era of Computing

Introduction

The advent of having the computing power of a mainframe computer on our desk is upon us. Not only do we have the power, but we have the sophisticated application development languages that facilitate our development of end-user, menu-based application systems in relatively short amounts of time. No longer do we have to wait many months to a year for a sophisticated report generator or even a system to provide interactive, application-specific editing and what-if capabilities. With the power of the Fourth Generation **SAS System** combined with the hardware power of the **IBM RS/6000** and its associated AIX operating system and X-Windows environment, both the developer and enduser are provided with a unique environment in which to extract the **maximum** amount of **information** available from the data at hand (See Figure 1). Further, the environment lends itself well to rapid collection of data in a myriad of fashions. In fact, it is not unusual for sophisticated enduser systems to be prototyped within one to two months. We accomplished just that for the Burlington Northern Railroad.

In order to optimize capability, the Burlington Northern is using the mainframe and workstations in conjunction with each other; taking advantage of the natural abilities of each. Burlington Northern personnel have discovered the new levels of power that an enduser can experience by porting mainframe SAS applications to the AIX environment. Application developers are impressed with the ease with which applications already existing on the mainframe in SAS format, can be migrated to the workstation. The benefits of running applications in the AIX environment are evident. The

Burlington Northern Railroad has used the SAS System on an RS/6000 to quickly prototype and develop applications that would be beyond the capabilities of both a PC and mainframe. By using the mainframe and RS/6000 in conjunction with each other, Burlington Northern has optimized their data processing and decision support applications using the SAS System.

Nonetheless, this is a different environment, and the issues associated with those differences should be studied carefully and well understood prior to making the commitment to moving applications from the mainframe to the AIX environment and/or developing new applications. Finally, the workstation with its UNIX operating system is foretelling the main business computing system by the mid 1990s. Although it is not necessary at this time to make the migration to or include the use of the workstation in a business environment, those who can and do will have a head start on the rest of their competition.

Experiences on the Mainframe

SAS originated as a batch-oriented language. Until the late 1980s, in fact, that was the method of choice to build and run applications. Most end users were provided with reports, and interactive data analysis was something that might have been desirable, but was too time consuming. With the advent of Display Manager and Application Facility in Version 5 of the SAS System, the change was beginning. Even on the mainframe, there were environments where the use of interactive menus were desirable. Now, on the RS/6000 in the AIX X-Windows environment, interactive processing is fast and easy, ensuring that it is the processing platform of choice for all but the largest problems.

As a result of cycle time competition, the mainframe environment was often plagued with slow interactive response time. Although this was not relevant to batch processing, it did affect the development of the batch programs. In addition to cost/benefit justification, we have two major user-oriented reasons for the use of the workstation environment. First, application development is speeded up due to faster interactive response time. Secondly, because testing and debugging of sections of code can be done on the fly, almost instantaneously, no longer does one need wait until large segments of code have been written before testing can be done. This continual feedback is a very positive motivation for programmers as well as ensuring that the programmer does not stray from the desired direction and requirements of the application.

The Need for a Different Platform

More and more, endusers are becoming accustomed to getting the kind of information when they want it without having to wait for an analyst or programmer to generate

a report. This need requires an array of enduser tools, all built in the same fourth generation language. As was spelled out in a SUGI 14 paper, "Expectations for a Fourth Generation Language,"¹ application developers should be able to quickly build dynamic, syntax-free, point and click menu based systems from which different users can obtain information for their different needs. Although this has been done on the mainframe, the RS/6000 workstation provides a much less costly and more efficient platform.

In a follow-on paper, "DB2 and the SAS System — Information Delivery for the 1990s,"² Baer emphasized the need for enduser access to data that could be turned into information virtually instantaneously. Application development speed is enhanced significantly on the RS/6000. But, more importantly, application enduser learning time to become proficient on the system is reduced, and best of all, the amount of time to use the application and acquire needed information is reduced. It is this last item that provides the biggest cost benefit back to a company.

The major problem users had on the mainframe, and the reason for running jobs in batch was that each user had to compete with other users for the limited CPU and DASD resource. This competition necessitated running jobs in batch. To try to run jobs interactively caused greater expenditures of development and testing time. Furthermore, interactive applications were frowned upon, particularly if any amount of data manipulation were required, because the users would eat up inordinate amounts of CPU. As a result, the mainframe has remained the domain of batch processing. Conversely, the workstation with its lower cost CPU cycles and lower overhead can easily be justified for the interactive SAS user as well as many of the batch jobs.

Although the UNIX operating systems have been around for some time now, UNIX originated as a programmer's operating system, with a wealth of tools, but a lack of discipline. This problem is rapidly being corrected through adherence to international standards for both UNIX operating systems and the X-Windows environments. As such, the business community, in particular, can expect AIX and its siblings to become as robust as the mainframe operating systems, and be able to support the most critical and demanding of application requirements.

Our move from the mainframe to the RS/6000 was interesting in the sense that it required a firm commitment and a leap of faith. It was not possible to try to work in both environments. In order to force the issue, a period of four to six weeks of working only on the RS/6000 was necessary. Otherwise, it became obvious very early on that the programmers would quickly move back to the environment that they were used to. This commitment to change was very successful, and in a short period of time, new applications were developed

that, if they could have been developed on the mainframe, would have taken much longer.

Achieving Comfort in the AIX Environment

The seasoned SAS programmer and/or enduser should be able to migrate comfortably to the SAS Display Manager system within a week or two. However, we estimate that between one to three months may be necessary before a user will become comfortable in the AIX environment with the use of the AIX commands and the ability to manage AIX with relative ease. The major differences between the SAS System on the mainframe and on the RS/6000 is the much greater use of the mouse and the ability to have and manipulate a greater number of windows. Because a user can have many windows open within an X-Window session, more than one SAS Display Manager session could be running allowing the user to work on different programs simultaneously. In order to facilitate the migration of the SAS user community to the RS/6000 in the shortest amount of time, we recommend that there be one experienced AIX user who also knows SAS for every five to ten SAS users. This support process will ensure that any issues that are AIX related will be resolved by the expert without the AIX neophytes needing to spend or waste their time. This methodology is currently being implemented successfully at Burlington Northern.

One of the most powerful and enjoyable advantages of working on the RS/6000 is the graphics speed and capacity. Within the windowing environment, complex graphics can be created, edited, and resized. In the Burlington applications, maps of the U.S. with designated and named train stations are constantly being produced. These provide very important visual tools for planning purposes. Although the same graphs and maps can be created using mainframe resources, the RS/6000 is so much faster that enduser applications using the graphs and maps are being created on the RS/6000 that would have been rejected on the mainframe simply due to the time required to draw and redraw the graphics on the screen.

One of the major advantages of AIX over the other UNIX offerings is the system administration capability through tools such as SMIT and INFO. AIX was developed with a focus on Usability, RAS (Reliability, Availability, & Serviceability), and Security. For these reasons, we believe that the SAS user is best served by working in the AIX environment.

AIX is a case sensitive operating system. As such, file names specified in a SAS program under AIX must match exactly casewise with the named file. Other than that, case sensitivity is ignored within the SAS program. The file reference conventions on the RS/6000 are similar to that of the PC in using directory notation versus what is used under MVS (datasets) or VM (files

within disks). Other than the above, once small modifications are made, SAS programs can be run either in batch or display manager mode. We found that users were surprised and impressed that SAS was as similar in look and feel on such different systems.

Mainframe & RS/6000 Strengths & Weaknesses

This section contrasts the differences for developers and endusers on the three major platforms on which software are run: mainframes, workstations, and PCs. Figure 2 provides a visual of the authors' perceptions of the strengths and weaknesses of each platform. The feel of the PC and mainframe environments are provided on the RS/6000 by the mouse/windows capability and power, respectively.

Figure 2. Usability Comparisons

	Main-frame	Work-station	PC
Performance	+	+	-
Problem Capacity	+	+	-
Price/Performance	-	+	•
Storage & data access	+	•	-
Data backup & Protection	+	•	•
Batch Processing	+	•	-
Editing/Coding	•	+	+
Testing/Debugging	•	+	-
Windowing Environment	-	+	•
Graphical Enduser Interface	-	+	•

+, •, - stand for relative advantage, neutral, and disadvantage with respect to a neutral platform

When the SAS programmer is writing a program, the interactive capabilities and speed of the RS/6000 facilitate rapid development because of the continual ability to run the pattern of code, test, debug, code, test, debug, etc. We have yet to find a platform that works faster for SAS application development. This aspect of rapid development indicates that the RS/6000 is an appropriate platform for development of all SAS applications no matter where the eventual production environment will reside.

For those users who are still writing programs in their favorite editor and submitting them through batch, we strongly recommend becoming comfortable working in the display manager environment. On the RS/6000, the display manager environment offers total access through pull-down menus to the wealth of SAS tools and functions. This minimizes the need to remember SAS commands and/or write short programs to view or manage data during testing of programs.

For CPU performance, the RS/6000 has capabilities that come close to matching those of lower end mainframes. This means that the problem solving capacity is also fairly close. However, when

price/performance is compared, the workstation will win virtually every time as long as the SAS application or problem will fit within the confines of the workstation environment. Those very large problems which often use multidimensional arrays frequently will need the horsepower of the largest of mainframes such as the ES/9000, usually with a vector facility. Unfortunately, it is difficult to specify at what point in the size of a SAS application where a mainframe would be most appropriate without first benchmarking the application on both systems. This approach is used quite often with a great deal of success in many customer shops.

One must always consider storage and data access requirements of users and/or applications. It has been our experience that many SAS applications, particularly batch jobs, are data intensive and require large amounts of reading and writing to disk. The RS/6000 can be used as a backend machine, and a number of SAS users have chosen to run their batch jobs that way because of the lower cost of the CPU cycle. Using SAS/Connect or some other way of moving the jobs from the mainframe to the RS/6000 and back, the users can save the expensive mainframe cycles for jobs that require the horsepower and/or disk I/O of that platform.

Because of the local power of the RS/6000, the graphics possibilities are incredible. For Graphical User Interface (GUI) based SAS applications, the workstation is the platform of choice. The nature of Motif and X-Windows makes presentation of graphics and interaction with the user powerful, enjoyable, and very functional. Although graphical applications are built and operate adequately on the mainframe, they often suffer from inferior displays and slower response times. Even when these problems are corrected, the cost exceeds that of the RS/6000 solution.

A very important area to consider when migrating SAS applications to the RS/6000 workstation is data backup and protection. As the RS/6000 is often a multi-user platform, an individual should be assigned to be responsible for regular data and program backups. There should be system programmers assigned to manage workstation environments just like mainframes. It is the value of the application and associated data that dictate appropriate personnel be assigned. This is NOT like the PC world where each person ends up being their own system administrator.

Performance Comparisons Between the Mainframe and the RS/6000

As previously stated, and as the benchmarks bear out, the RS/6000 workstation environment offers processing power roughly equivalent to low-end mainframes (see Figure 3). It is only in the area of I/O that the mainframe has significantly greater capability and speed. But even in the I/O area, as shown in Figure 4, the RS/6000 can

hold its own, particularly with medium sized datasets. It is our contention that the RS/6000 can satisfactorily handle SAS datasets or flat files under 100 megabytes in size (one million 100 byte records). However, applications with considerably larger datasets have been successfully run on the workstation. As long as time is not a factor, and the workstation can be dedicated to the job running, the processing will be done. The considerably lower cost of the RS/6000 platform allows for running single applications for days, if necessary, to save the more expensive cycles on the mainframe.

Figure 3. Performance benchmark comparisons

Sample batch pricing job - CPU intensive

3084 Mainframe

70.0 minutes real time
10.0 minutes user CPU

Model 540 RS/6000 with 48 MB Memory

50.0 minutes real time
42.5 minutes user CPU
10.2 MB memory used

Sample batch data processing job - I/O intensive

50,000 observation subset of data

3090 mainframe

2:33 minutes real time
1:23 minutes user CPU
1.4 MB memory used

Model 540 RS/6000 with 48 MB Memory

2:40 minutes real time
1:26 minutes user CPU
0.6 MB memory used

We found that the Fulltimer Option is a useful SAS option that will measure the real and system time as well as the memory required to execute each step of the job and the whole job. By running carefully controlled benchmarks on each platform, organizations can best decide on how to partition their SAS workloads between the mainframe and the workstation.

Networking with SAS

Another major advantage of using the SAS System on the RS/6000 is the communication capabilities among various platforms. IBM promotes the use of TCP/IP as the network software, but others are supported as well by the SAS System. The communication vehicle in the SAS System to communicate with other platforms is SAS/CONNECT™ which allows the user to access data from remote systems as if the data were local to the user. In addition, the user can submit jobs on remote sessions while being signed on to the RS/6000 or can sign onto the RS/6000 remotely from the mainframe. If SAS/CONNECT is not being used, the CPORT and CIMPORT procedures may be used to move SAS datasets from one platform to the next. CPORT places

the data into ASCII file format that can be input through CIMPORT on any other platform. Through the use of Local Area Networks (LANs) and Wide Area Networks (WANs), the users and their data can be on any systems and still have ready access. As IBM supports open systems architecture, the user can be assured that a networking solution exists for any customer communications requirement.

Code and Data Migration

Code that runs on a mainframe is transferable to the RS/6000 with minimal changes. In fact, flat files can be moved to the workstation via 3.5" floppy disks. For direct transmission, there ought to be a hardware and software communications path between the mainframe and workstation. TCP/IP running on a Local Area Network works quite well. Code and data can then be moved through the communications pathway. If needed data is already stored in SAS datasets, the data can be moved by the use of PROC CPORT on the mainframe and PROC CIMPORT on the workstation. Alternatively, SAS/CONNECT can be installed which will make use of the existing hardware and software network to allow the SAS user to access SAS datasets on the mainframe as if those datasets were on the workstation. If data are stored on the mainframe in EBCDIC format, the SAS System provides input format capability which will allow the user to read those data on the workstation (which uses ASCII) without having to modify the data first. The ONLY code modification that need be done is to change the LIBNAME and FILENAME statements so that they accurately reflect where the datasets reside and to ensure that referenced filenames match the case sensitivity of the filenames under AIX.

Migration Procedures

In order for an organization to successfully migrate SAS usage from the mainframe to the workstation, the following requirements apply:

- At least one individual should obtain adequate training in the AIX Operating System including Motif and X-Windows. This person would be the AIX technical support.
- Familiarity with TCP/IP and other communications between the mainframe and workstation is desirable.
- One person should take responsibility for performing system administrator duties on a regular basis and as required. This person would also monitor system usage to determine at what point the system might need to be increased in size.
- The differences between SAS on the mainframe and the RS/6000 should be outlined in a document to ease the transition.

- SAS users should be provided procedures to execute SAS both in batch and Display Manager and should be provided with a standard environment in which to work. This should include a common set of routines and key definitions.

Depending on the size of the organization, one person might be sufficient to act in the capacity of AIX technical support, communications specialist, system administrator, and support person for the SAS System under AIX.

Tips and Tricks

- Customize your X Interface to the SAS System as described in Chapter 4 of the *SAS Companion for the UNIX Environment and Derivatives*³. Although you can customize for each session, we recommend that you specify the desired resources in a resource file. We specified color and window size. The key customizations turned out to be the most useful and the greatest time savers. You can define hard-key bindings. On the RS/6000, we had to define the insert key to work as a toggle insert key. The SAS System did not default to that. Following is the SAS.hardkeys from the defaults file we currently use:

```
SAS.hardKeys: <key>Insert: sas-toggle-insert()\n<Key>Delete: sas-delete-char()\nCtrl<Key>Right: sas-next-word()\nCtrl<Key>Left: sas-prev-word()\nCtrl<Key>End: sas-delete-end()\n<Key>End: sas-move-end()\n<Key>Home: sas-move-begin()
```

- In the SAS session itself, the user has access to a user profile in which the actions of specific keys can be assigned. We prefer to assign the numeric keypad to assorted functions. For example, the '+' key is assigned the 'clear' function and the '*' key, the '?' (retrieve previous command) function. Also, we liked to set up the 'O' on the keypad to swap between the pull-down menu and the command line. The command line is necessary for things like changing the current directory or invoking the AF command for testing. You will be able to decide how to customize your SAS KEYS window after using the SAS System for a short amount of time.
- We recommend that for ease of text searching, the program is written in all caps. Otherwise, the pull-down menu will need to be used for text searching.
- The directory in which the desired SAS programs are stored should be made the current working directory prior to initiating the SAS System or through a 'x cd' system command before executing the pull-down menu open command.

The Burlington Northern Railroad Testimonial

The Burlington Northern Railroad has had a long history of developing and maintaining mainframe applications. For most of the railroad the mainframe is the proper environment. However, there are a few isolated users within the company whose applications need the processing power of a mainframe, and the functionality of a PC. It is for these reasons that the Burlington Northern Railroad has decided to migrate its compute intensive applications to the UNIX environment. The primary use of the workstations will be to run operations research models. Not surprisingly, the railroad found many applications for its new environment within a very short time.

One of its most exciting discoveries was the SAS System running under the AIX environment. After a brief experience with AIX SAS, it was determined that many of the mainframe SAS applications could be moved to the AIX platform. The new platform offers a number of advantages over the mainframe. For instance, the user interface feels much like a PC. The user can use mouse driven menus to navigate the SAS System. It may feel like a PC but it processes like a personal mainframe. Many applications that took minutes on the mainframe are now functioning on the workstation with little or no dead-time. The railroad has taken full advantage of the high resolution graphics that can be achieved using the SAS System on the workstation. It is currently producing system maps that would be almost impossible to view using mainframe graphics, and more importantly, extremely slow to generate.

The migration to the new operating system was made easy by the portability of the SAS System. Using PROC CPORT the railroad was able to move its SAS applications and data to the workstation with a minimum amount of effort. The railroad even experimented with porting SAS/AF® applications from the mainframe with great success. The SAS/AF applications that were brought down took on an entirely different look and feel.

The porting was so successful that Burlington Northern, in conjunction with IBM, is currently developing an interactive schedule maintenance tool in the AIX-based SAS System. The tool will mimic a PASCAL application that is currently being used on the mainframe. The mainframe tool was extremely difficult to use--it had no graphics functionality, it was very slow, and had a typical mainframe look and feel. Because it was such an "ugly" tool it was used, unwillingly, only by those who had to. With the introduction of the AIX-based SAS System, the railroad was able to quickly develop and implement a compute and data intensive tool that felt like a windows application. Development was quick and inexpensive. It is mouse driven. It is extremely easy to use, and supports a wide variety of graphics.

Currently, several organizations within the railroad have begun using AIX as the primary SAS platform. The mainframe is used only when very large amounts of data need to be processed. The railroad has realized such a dramatic growth in its SAS capabilities that non-technical users, who were previously kept from the functionality of the SAS System by the mainframe environment, are now eagerly using the SAS System under AIX.

Conclusion

The workstation environment now provides a platform that can be used by a multitude of organizations requiring computing power. Those small groups which could not afford a mainframe can couple their PCs into a LAN and have the power of the RS/6000 when the need arises. Those large organizations with groups distributed remotely from the mainframe will find that much of the work can be accomplished by providing one or more workstations in each remote location while maintaining the mainframe to process company-wide data and act as a data repository from which the remote sites can obtain data. This, in fact, is the approach being taken by Burlington Northern. The time has arrived that computing power, speed, and visualization is on our desks, at our fingertips. The limiting factor is how we collect and organize our data, and then analyze and present it. The information revolution continues to march along. The SAS System on the RS/6000 platform is a glimpse into the future.

Bibliography

1. Baer, Darius (1989). Expectations for a Fourth Generation Language. In SAS® Users Group International Proceedings of the Fourteenth Annual Conference, SAS Institute, Inc., Cary, NC.
2. Baer, Darius (1990). DB2™ and the SAS® System — Information Delivery for the 1990s. In SAS® Users Group International Proceedings of the Fifteenth Annual Conference, SAS Institute, Inc., Cary, NC.
3. SAS Institute Inc. (1990). SAS® Companion for the UNIX® Environment and Derivatives, Version 6, First Edition, Cary, NC: SAS Institute Inc, 273 pp.

For more information, contact:

Darius S. Baer, Ph.D.
SAS Institute Inc.
1950 Stemmons Freeway
Suite 5052
Dallas, TX 75207-3109
(214) 746-3916

*SAS, SAS/AF, and SAS/CONNECT are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA

*IBM RISC System/6000 and AIX are registered trademarks of the IBM Corporation, Armonk, NY, USA