

A Historical Perspective of SAS's Future

Rick Nicola

SPS Software Services, Inc.

Abstract: The history of SAS with its roots in 370 architecture is similar to many other products. The jump to the PC and the offering of like facilities on many other platforms was paralleled by a much smaller subset of other products. Companies looking for a cohesive approach to make the next jump to object oriented programming and beyond should study history before choosing their "strategic" tools. This paper will touch on the history of SAS and its use at many data centers in the seventies and early eighties. It will examine present capabilities and how they follow the same basic philosophy as the original facilities. Finally, the paper will discuss why SAS will successfully make its next jump, by drawing from its base philosophy and its technical roots.

Introduction: In the Beginning:

SAS, like SPSS and BMDP, were initially aimed at statistical analysis and at the student and professional statistician. However, SAS usage quickly became broader because of its unique ability to aid the user in handling data. At its inception, SAS ran only on IBM's "370 architecture."

Coincident with the availability of SAS, growth in computer applications and the attendant growth in computer hardware requirements and costs, highlighted the urgency for methods to better manage the data processing utility. Because of its ability to handle and store data and perform analytical tasks, SAS quickly became a favorite of those charged with analyzing and presenting computer usage information. SAS became a focal point for "Computer Performance Evaluation" or CPE, and was used for the development of a computer "Performance Data Base" (PDB), in many IBM-370 shops. Early "SAS readers" of computer usage information were available through CMG (Computer Measurement Group) and the IBM User's group, SHARE. These were popularized and discussed in various papers by Barry Merrill and others who pioneered this discipline.

The concept of the "PDB" ultimately resulted in the availability of Merrill's MXG and other similar SAS based products. These products formalized

the handling of computer usage data and made the use of SAS for this purpose even more cost effective (and thus attractive). This resulted in two phenomena having little to do with CPE.

A. SAS was acquired for CPE tasks, but was quickly picked up by the rest of the user community. Engineers and others in technical and analytical areas who had been exposed to computers in college, were hungry for tools that were easy to use, but had strong analytic capabilities. SAS became one of the first tools to put the power of the computer directly in the hands of the user and to reduce the dependence on DP Programming groups to obtain basic information.

B. Several methods and concepts were introduced to handle computer usage data with the inception of the "Performance Data Base" (PDB). It was been recognized that these concepts could be applied to handling almost any kind of data describing a broad range of disciplines. While these concepts have been expressed in a number of ways, they might be stated as the following "rules" for building usable data reporting systems:

1. Store desired information in a form that is easily available.
2. Provide tools that can easily analyze and relate data.
3. Provide tools that can easily and clearly present base information and results.

The SAS people divide these elements into four SAS characteristics or cornerstones:

SAS Cornerstones

Data access.

Data management.

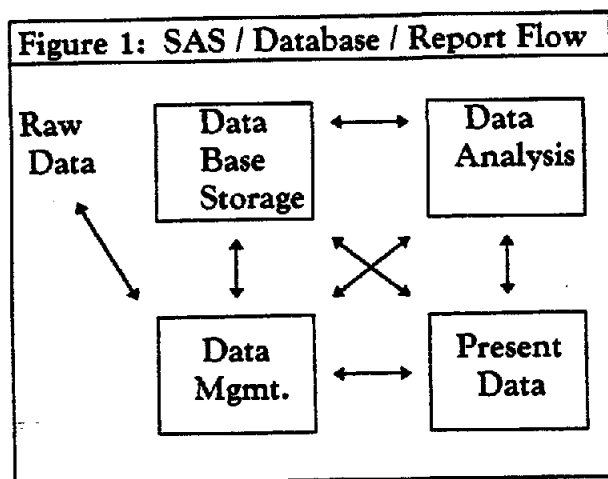
Analysis.

Data presentation.

In the twenty years since these concepts surfaced, our expectations for storing, analyzing and presenting data have changed and grown. However, the **base cornerstones remain constant**. And, while many other products have come (and many have gone), SAS has continued to provide enhanced facilities to meet growing and changing expectations.

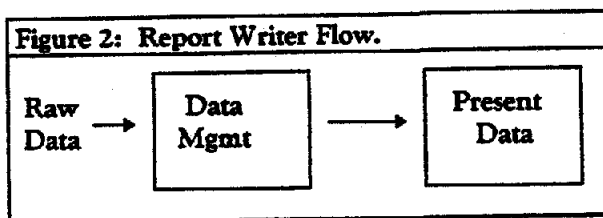
History: (Note that Figure 3 at the end of this paper gives a history of some SAS facilities. All dates are approximate and the list is by no means all inclusive.)

The SAS people often show the cornerstones graphically. They might be presented as follows:



History: 1975

In the mid seventies, several "report writers" emerged. These products were aimed at presenting printed reports directly from raw data. The basic method used was to manipulate the raw data and create a variety of detailed or summarized reports. A report writer chart is shown below.



These report writers were fairly easy to use. The user coded some base file definitions for the raw data and defined his output reports. As the SAS and Report Writer flow figures indicate, the report writers lacked an intermediate information repository (database) and generally had limited

analytic capabilities. SAS stored its information in its own unique format, and already offered substantial statistical procedures and analytic tools

The report writers were powerful and flexible. SAS had PROC PRINT, but to truly duplicate some of the report writer facilities one needed direct report coding and PROC SUMMARY. SAS lagged the report writers in this respect.

History: 1980

Going into the **eighties**, report writers continued to flourish. SAS continued to lag the report writers, but offered some basic "printer graphs" via PROC CHART and PROC PLOT.

At this time, there was also more of an interest in creating true business type graphs and other graphical representations. Throughout the 1970's, plotters and graphics terminals were available. However, much of the related software was aimed at a scientific audience that included those requiring design drawings and other engineering type applications. Building graphs, often required programming via FORTRAN or other language, and was driven by a series of subroutine calls.

During this period, SAS Institute introduced SAS / GRAPH® with SAS79. This product offered those on the 370 architecture a reasonably priced alternative to other products and a more easily used interface for creating charts and plots through generalized software.

In addition, to graphics, enhancements were also made to data management, analytic and presentation elements of SAS in SAS79 and SAS82. (See Figure 3 at the end of this paper.)

History: 1985

The release of SAS V5 in 1985 introduced numerous new facilities, such as the presentation manager, SAS/AF and some early entries into a portable version of SAS. Upgrades were made on each of the four "cornerstones." with this version, while sub products continued to be added to SAS

Competitive products began to become available at this time. Products such as FOCUS presented elements of the four "cornerstones" (although the analytic tools were not as strong) and made a stronger initial entry into the PC market.

The Institute recognized the need for portability and made the decision to rewrite SAS using the "C" compiler. Since "C" was available on most platforms, it held the promise of minimizing development costs while crossing platforms.

The decision to completely rewrite SAS using the "C" compiler, is the most pivotal in the product's history. This positioned SAS to cross platforms and laid the groundwork for SAS products into the twenty first century. The first "C" based SAS introduced was SAS 6.04 for PC-DOS in 1987.

History: 1990

The SAS "C" based system was made available on mainframe, UNIX and other systems in 1990. Version 6.06 of SAS had some shortcomings and some unnecessary incompatibilities with version 5. The first truly viable Version 6 for the mainframe was 6.07.

Version 6 provided a major design improvement to the "data management" and "data access" cornerstones by introducing the concept of an access "engine." It also provided major improvements to the presentation element within improved SAS / AF® and other interactive products.

While these changes provided a great improvement for the user familiar with mainframe SAS, it was inadequate for the user who was used to the responsiveness and "glitz" of a PC. This pushed the Institute to further enhancements and additional function on all four cornerstones of the system.

History / the Present: 1995

The products and facilities made available by SAS from 1990 through the present outstrips all other development prior to this period. SAS has enhanced all elements of it of its data storage, data management, analysis, and data presentation cornerstones.

1. Access of Data: The handling of "raw" data and even the storage of information in a data base has changed drastically. SAS / ACCESS® and views into numerous non-SAS database files makes it possible to retrieve and operate on data that is built and shared with other database systems. The 1990 change to a multiple engine architecture does not preclude the use of SAS formatted data bases

as well. In fact, the use of SAS or other database facilities to store data for reporting became even more popular. Somewhere along the line, we started to call this Data Warehousing.

2. Data Management: The availability of SAS on numerous platforms is presently complemented by products like SAS / CONNECT® and the remote library facility. These facilities address changes in where data is stored, rather than how. The ability to selectively retrieve and download information "on the fly" as a part of an interactive application helps position SAS as a fully flexible tool that can adapt to any definition or "model" of a client server application. This characteristic, together with display methods provide functionality available in no other product.

3. Analysis: The analytic aspect of SAS continues to be enhanced. The addition of even more industry or discipline based products complements the existing set of products and procedures. Interestingly enough, SAS / CPE® has been introduced to simplify some of the computer performance evaluation tasks that introduced SAS to many data centers in the first place. This product works with Barry Merrill's MXG and extends CPE tasks beyond the 370 architecture onto other platforms.

4. Data Presentation: Perhaps the most dramatic recent changes in SAS have occurred in the way that data may be presented and displayed. SAS / AF®, which was available in a limited away on version 5, is in a constant state of development in version 6. The use of the FRAMES elements of SAS / AF® extends the facility. SAS / EIS® provides an easier to use entry into building "point and click" applications. Continued enhancement and the availability of interfaces to "industry standard" facilities (such as object embedding, ODBC) has kept SAS in the forefront as "object programming" grows.

Why SAS been so Successful:

SAS has been successful in providing products that have remained competitive for nearly twenty years. This success has been due to "good management" and continuing product development. While many companies can claim similar characteristics, SAS has an advantage because of two processes.

1. Recognizing needs.

The use of the SASware ballot has kept the Institute in tune with the requirements of its users. Their involvement in SAS User Group International (SUGI) and the use of other innovative methods for recognizing change and needs has played an increasing role in reacting to changing requirements.

2. Costs and charges.

When SAS was simply a mainframe product, most mainframe software vendors charged a large, initial fee for their product, plus a yearly maintenance fee that ran about ten to fifteen percent of the initial fee. SAS on the other hand charged a relatively low initial fee, but charged a fifty percent "renewal fee." This enabled them to provide a steadier income flow that was less dependent on "new business." With this steadier cash flow, funds were available for continuing development. Since SAS is privately held, they are able to direct money to development rather than show short term gains to investors. While precise numbers are not available, sources state that over a third of SAS's revenue was directed to R&D in 1993, and it is expected that this investment level is standard.

Presently, SAS maintains a similar costing philosophy for products on all platforms, though PC versions of SAS have presented some challenges in pricing and packaging. The complaint is often heard that "SAS is too expensive" when compared to some of the "PC" packages. The "other" packages have no renewal fee. While these complaints have some validity, the fact that their prices are "initial cost only" should be a reason for concern.

Some companies are dependent on new customers or major enhancements to the next release of their product to fuel sales. Only revenue from these new sales are available to underwrite continued software enhancements. This costing structure is tenuous at best. In the PC software industry we have seen companies with packages that were the "kings" of their particular areas (such as database, spreadsheet, or word processing) on the brink of disaster. Much of this was because they were unable to continue market domination and produce the revenue necessary to meet new competition.

If I am a "home computer" user of a word processor, I can survive the demise of the vendor of

my package. If, however, I am a company dependent upon a vendor for program development products, I want to be sure that this vendor will be around in ten years. Furthermore, I would like for that vendor to continue to enhance his product with those hardware and software facilities that are "state of the art."

The failure of a software vendor to survive or to continue to enhance his product, would ultimately result in the instability or loss of user built information systems (no matter how informal these systems are.) **The cost of SAS software renewal amounts to very little when compared with the cost of redeveloping production and user systems.**

Summary and Conclusions:

SAS has a history of providing state of the art software and continuing to improve it on a regular basis. We would expect this improvement to proceed while other products and vendors might fail because of methods for "reading" new needs and because of the SAS revenue structure.

It is difficult to predict what computing will be like in the year 2000. The selection of a tool for developing client / server systems, ad-hoc reporting, or data analysis is a difficult decision for the casual user or the data processing manager. But, if your company is in it for the "long haul" and you wish to provide a stable use and development environment, SAS is by far the best product and the **best strategic choice**

References for: A Historical Perspective of SAS's Future

MWSUG October, 1995 Rick Nicola
SPS Software Services Inc. Canton, Ohio

Mike Ricciuti, "Can SAS Serve the Enterprise",
DATAMATION, May 1, 1992.

Mike Bucken, "SAS Emphasizes R&D to Gain MIS
Edge", Software Magazine, May, 1994.

Kevin Strehlo, "Programming for Business",
DATAMATION, May 1, 1995.

Figure 3 SAS History / Cornerstones.

Interesting Raw data	Data Management	Information Database	Analysis	Presentation
1975 (Actual company founding: 1976)				
Statistical CPE	Handle data SUMMARY	SAS's own (360 Arch)	Statistical Modeling	PROC PRINT
1980.....SAS79, SAS82.				
As above	ARRAYS SAS / ACCESS® MACROS APPEND TABULATE	As above VM/CMS	As Above SAS/ETS®	Graphics SAS/FSP®
1985 SAS V5, PC SAS (V6).				
As above VSAM Other	SAS / ACCESS® (extraction). Better macros	As above First portable (PRIME)	As Above SAS/IML® others	Display manager V5 AF
1990 SAS V6 (C-based)				
Growing	VIEWS SAS/ASSIST®	As above Multi-engine New databases More platforms SAS / ACCESS® (ACCESS now direct, not extract)	As Above Numerous new products.	V6 SAS/AF® SAS/ASSIST®
Various improvements, new products, etc. through present:				
1995				
Growing	SQL DB Views Flat file views SQL QUERY	New databases SAS databases Relational DB Non-relational SAS / Connect DATA Warehousing More platforms ODBC	Numerous new products SAS / CPE® etc. etc.	Graphics PROC REPORT SAS / AF® and FRAMES SAS / EIS® Better graphics Objects SAS/INSIGHT® etc.
2000				
	?????	New databases	New	PC Graphics