

## Paper 156-28

**How to Access PC File Data Objects Directly from UNIX**

Howard Plemmons, SAS® Institute Inc., Cary, NC

**ABSTRACT**

The ability to directly access PC data stores from SAS on UNIX has become a reality in SAS V9. This paper highlights the behind the scenes implementation of this enhancement to the SAS/ACCESS® Interface to PC Files product line.

**INTRODUCTION**

With SAS V9 SAS PC File Format customers on UNIX will have direct access to PC File data that is stored on the PC. The culmination of this development project resulted in enhancements to both the PC and UNIX versions of the SAS/ACCESS to PC File format products. The focus of this paper is to provide information on how this interface works and how you can put it to use. As of the submission date of this paper the PC File enhancements are still under development; therefore, some content and description may be different in the final product.

**SAS VERSION 8 PC FILE ACCESS**

The PC File product for SAS V8 differs based on OS. The following details the differences on PC and UNIX platforms:

PC File ACCESS	PC Product	UNIX Product
Microsoft® ACCESS	Yes	No
Microsoft Excel	Yes	No
WK1, WK3 and WK4 files	Yes	No
DIF files	Yes	Yes
DBF files (version III, III Plus, IV or 5)	Yes	Yes
Delimited files	Yes	Yes

Table 1

**On the PC platforms the product contains:**

Provides access to:

Several different formats see Table 1 above and:

- Proc ACCESS – creates descriptor files that describe data in a PC file to SAS, enabling you to directly read, update or extract PC files data into a SAS data file.
- Proc DBLOAD – creates and loads PC files.
- Import/Export Wizard – provides a wizard interface that provides a transfer method between SAS and PC file formats (export) and PC file formats and SAS (import).
- Proc Import and Proc Export – provides a procedure interface to allow import and export to PC file formats.

**On the UNIX platform the product contains:**

Provides access to:

Several different formats see Table 1 and:

- Import/Export Wizard – provides a wizard interface that provides a transfer method between SAS and PC file formats (export) and PC file formats and SAS (import).
- Proc Import and Proc Export – provides a procedure interface to allow import and export to PC file formats.

Note these descriptions represent what is available when using the SAS V8 release. The additional content provided in this paper details components added to SAS V9.

**SAS VERSION 9 PC FILE ACCESS**

Includes all the features mentioned in the SAS V8 section plus the following:

PC File ACCESS	PC Product	UNIX Product
Microsoft® ACCESS	Yes	Yes
Microsoft Excel	Yes	Yes
WK1, WK3 and WK4 files	Yes	Yes
DIF files	Yes	Yes
DBF files (version III, III Plus, IV or 5)	Yes	Yes
Delimited files	Yes	Yes
SAS JMP®	Yes	Yes
FoxPro®	Yes	Yes
Paradox®	Yes	Yes
PRODAS®	Yes	Yes
SPSS®	Yes	Yes
MATLAB®	Yes	Yes
Others to be named	Yes	Yes

Table 2

Note: Table 2 contains a list of new file formats we will be able to support in SAS Version 9. The actual support list when SAS Version 9.1 is released may be different.

**On the PC platform the product contains:**

Provides access to:

Support was added for several new data formats, see Table 2 above. The following software enhancements were added to the PC platform support in SAS Version 9 as well:

- A SAS Libname engine to allow for direct communication with Microsoft excel and Microsoft Access.
- Enhancements to the import/export procedures and wizard to include new file formats.

### On the UNIX platform the product contains:

Provides access to:

Support was added for several new data formats, see figure 2 above. The following software enhancements were added to the UNIX platform support with the release of SAS Version 9.1.

- A SAS Libname engine for direct access of PC files. Specifically Microsoft Excel, Microsoft ACCESS, other file supported via ODBC connectivity to the PC.
- Enhancements to the import/export procedures and wizard to include new file formats. These include PC accessible and UNIX accessible file formats as shown in Table 2 above.

## MAJOR VERSION 9 ENHANCEMENTS

The extension of the PC File product for SAS V9 comprises several key components. The diagram below depicts the flow of these new SAS Version 9 UNIX features:

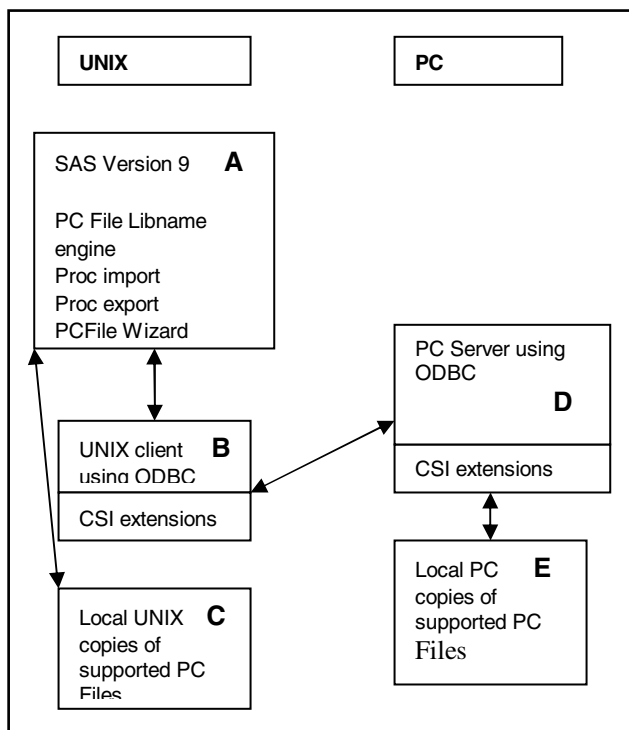


Figure 1 – Architectural Overview of Direct PC File Access

The functionality of Conceptual Software Inc. (CSI) extensions resulted from the acquisition of CSI by SAS in 2002. The ability to access PC file data flows from these component parts is described below:

UNIX portion of the product include a new libname engine, enhancements to PROC IMPORT/EXPORT functionality and the PC File Wizard. These new features allow you to access locally stored/accessible PC File data (C). You can also use the client interface (B) to obtain direct access to PC file data stored on the PC using the server (D) to access the local PC data (E). Enhancements to the SAS code include CSI obtained from an acquisition of CSI in 2002.

**B** The UNIX client and CSI extensions are used to provide communication between UNIX and the PC Server (D). There are two components that make up the client software. The first is an ODBC-based client. The second is an extension of the CSI interface that is used to interface with PC side CSI code.

**C** Local file access refers to the PC File data that is accessible from UNIX. This data can be FTP'd over or network accessible, mounted PC drives. The SAS CSI components available from PROC IMPORT, PROC EXPORT and the PC File Wizard can then directly read and manipulate these files.

**D** The PC server component is used to directly access the PC data via an ODBC interface to PC ODBC drivers or via CSI extensions from PROC IMPORT/EXPORT or the PC File Wizard on UNIX. The PC server task is described in greater detail in the PC SERVER TASKS section below.

**E** Local PC data can be accessed from UNIX if an ODBC driver is available for use by the PC server or it is a data source that can be processed by CSI extensions from PROC IMPORT/EXPORT or the PC File Wizard.

Note: The PC File Product on the PC has been enhanced to provide with the ability to interface with CSI extensions and the local data stores they support in SAS Version 9.1. The PC File Product also contains libname engines that allow direct access to Microsoft Excel and Microsoft ACCESS. This engine functionality was released in SAS Version 9.0.

### PC SERVER TASKS

As shown in Diagram 1 the PC server is required for you to access PC File data from UNIX. When viewing the diagram you should note that SAS Version 9 software is not required on the PC. The requirements are that the server must be installed and administered on the PC that you will connect to from UNIX. The process below details how that works:

- Install the PC Files product on UNIX and installing the server component on the PC using directions in the PC File install documentation. This will place the server components on your PC.
- Once the PC server is installed you can invoke it by running it from the install location. For example:

```
C:\sas install location\pcfserver
```

After invoking the dll you will see the server control screen.

**A** As mentioned earlier some of the new features added to the

SAS PC Server Control	
Server Name: <i>myserver</i>	
<u>User Name</u>	<u>DSN/File access</u>
<i>User1</i>	<i>excel</i>
<i>User2</i>	<i>access</i>
<i>User3</i>	<i>wks</i>
Port Number: 524	
Maximum Connections: 5	
<b>RESTART</b>	
<b>SHUTDOWN</b>	

Figure 2 – PC Server Administration Screen

This server control screen contains information that must be used by the UNIX client to access the PC server. Definition of these components:

*Server Name* – Identifies the name of the PC server. This information is required for libname access from UNIX. This is the node name of your PC.

*User Name* – Identifies who is connected to the server, for example, UNIX userid.

*DSN/File access* – What the user is connected to, either the name of the ODBC DSN file or the file type being accessed. The DSN would need to be set up with connection information on the PC.

*Port Number* – Identifies the port that the server is listening on. This information is required for libname access from UNIX.

*Maximum Connections* – How many UNIX users can connect to this PC server. The default is five connections; however, you can adjust the setting and restart the server.

**RESTART** – Takes the server down and restarts it. You can use this option when you change Maximum Connections, Server Name or Port Number.

**SHUTDOWN** – Takes the server down.

Note: the server is a multi-threaded server in that each user connection will use a different thread on the PC. The RESTART and SHUTDOWN commands will terminate all threads/user connections.

### DIRECT ACCESS TO PC DATA USING THE NEW UNIX LIBNAME ENGINE

Once the PC server is up and running you are ready to start accessing data directly from the SAS/ACCESS to PC Files on UNIX. As explained previously the PC server and associated UNIX client take advantage of ODBC. Once you have specified the required components, Server Name and Port you will be given access to PC File data on the PC. The DSN parameter is optional and would be used if you had configured ODBC access on the PC by creating an ODBC initialization file. If you do not specify a DSN SAS will determine the access method based on the libname options that you enter.

The new libname statement syntax specific to the PC Files libname statement is shown below:

```
libname <library name> pcfiles
      path='c:\supported datafile
location.extension'
      type=<access|excel|...>
      user="userid"
      password="password"
      dbpassword="db password"
      dbsysfile="workgroup information file"
      dsn="PC DSN location"
      version=2002|2000|97|95|5"
      header=yes|no
      <other common libname options>
```

Note user, password, dbpassword and dbsysfile may be needed for connection to an access database.

Additional information on the generic libname options may be obtained from the SAS online documentation or on the SAS web site <http://www.sas.com/service/library/onlinedoc>

With the PC server up and running and the SERVER, PORT and DSN if any, in hand you are ready to access PC data from UNIX. If you do not have a DSN setup on the PC SAS will figure out the type of access you are requesting and formulate the connection dynamically. Here are several code examples on how that access can occur:

#### Example 1 – accessing Excel data on the PC

```
/*-access pc server and use the Excel data ---*/
/*--- source set up on the PC server -----*/
libname x pcfiles type=excel port=524
server=myserver
path="c:\my excel file.xls" version=2002
header=yes;
```

```
/*-view the sheets within the excel file -----*/
proc datasets lib=x; run;
```

#### Example 2 – accessing multiple sources from the PC

```
/*--- several libnames to different PC data
stores -----*/
libname x pcfiles type=excel port=524
server=myserver
path="c:\my excel file.xls";

libname y pcfiles type=access port=524
server=myserver
path="c:\my access database.mdb";

/*-use with SAS -----*/
proc contents data=y.acctable; run;
proc contents data=x.excel; run;

proc sql;
  create table xx as select * from y.acctable;
quit;
```

Some highlights of Example 2 include –

- Multiple connections to the PC File server – you should see these on the server control panel (Figure 2).
- PROC CONTENTS – identifies the attributes of the data files that you are accessing. From the excel example you can access individual sheet names. From the access example you can access individual table names within the Microsoft Access database.
- PROC SQL – using this procedure with the pcfile libname engine enables you to create a SAS data set called xx from an

extraction of all the columns and rows that exist in the acctable in your access database.

### DIRECT PC FILE ACCESS USING THE IMPORT/EXPORT PROCEDURES AND WIZARD

The import/export procedures and wizard on UNIX have been extended to take advantage of this new direct access method. There are two access methods used to enable you to directly read PC data from UNIX. The first is the client/server model described earlier. The import/export procedures and wizard have been enhanced to use this model to read PC File data directly from the PC. The second method is to read these files on UNIX. To accomplish this task the file must be copied over from the PC or the PC disks mounted to UNIX. The sample code below shows these access methods using the import/export procedures:

#### Example 3 – The procedure interface getting data from the PC using the client/server model:

```
/*-import Excel data as a SAS dataset on UNIX */
proc import datafile="c:\mydatafile.xls"
out=work.a dbms=excelcs port=524
server=myserver;
```

This code would connect to the server on the PC and request data stored in the data file identified by the datafile parameter. The data would be transferred directly from the PC to the SAS dataset identified in the out= parameter.

#### Example 4 – The procedure interface getting data from a local UNIX file:

```
/*-import Excel data into SAS dataset on UNIX */
proc import
datafile="/u/yourdir/subdir/excelfile.xls"
out=work.a dbms=excel;
```

This code reads the local file. By not specifying the server and port a local file and or accessible disk is expected.

#### Example 5 – Writing SAS datasets on UNIX to data files on the PC:

```
/*-export SAS data to Excel data on the PC ---*/
proc export data=work.a
outfile="c:\pc data file.xls"
dbms=accesscs port=524 server=myserver;
```

The client server model supports both input and output processing. This code would connect to the server on the PC and then send data to be stored on the PC in the form of an access file. The data in the SAS dataset would be formatted and sent directly to the PC data file.

#### Example 6 – Writing SAS datasets on UNIX to data files on UNIX or UNIX accessible disks:

```
/*-export SAS data to excel data on UNIX -----*/
proc export data=work.a
outfile="/u/yourdir/excelfile.xls"
dbms=excel;
```

The SAS data is converted to the file format and written to a local directory or accessible disk.

## CONCLUSION

Starting in Version 9 of SAS software, there is a major change in the

ability to directly access PC data from UNIX using the SAS/ACCESS Interface to PC Files product. We have provided tools that will allow you to read and write a variety of PC file data directly from UNIX. You will now be able to access Microsoft Excel, Microsoft ACCESS and any other ODBC supported PC data source. With the incorporation of the CSI software into the PCFiles product you will be able to access additional data stores (see Table 2). These components and enhancements provide a reduction in process steps in order to get PC data to UNIX and provide new data sources that you can analyze with SAS.

## ACKNOWLEDGMENTS

SAS/ACCESS Interface to PC Files product implementation team:

Chris Dehart  
Henry Feldman  
Michael Ho

SAS/ACCESS Interface to PC Files product testing:

Robin Boyles

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Howard Plemmons  
Senior Software Manager  
Database Interface Development Department  
SAS Circle  
SAS Institute, Inc.  
Cary, NC 27513

E-mail: Howard.Plemmons@sas.com  
Phone: 919-531-7779

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

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