The Top Ten Questions to Ask Before Your Data Migration

by Diane Olson

Since SUGI 28, David Wiehle and I have traveled around the country giving "The MIGRATE presentation." We thought the presentation contained all the information our customers needed to know in order to migrate their data libraries and start using the new features of SAS®9. However, as more and more people began to use the MIGRATE information, they identified several migration-related issues, like what operating system they were on and how it was set up, finding their SAS libraries, how permissions were set up, and so forth. Customers said they needed to find out what version their libraries were in and who was using them and for what. As these customers began calling Tech Support with their questions and sending mail to Migration.Feedback@sas.com, they generously provided us with material for this paper - the tricky parts of data migration. This paper asks the top ten questions you need to answer in order to make your migration the most successful yet.

The goal of this paper is not to teach you how or why to use PROC MIGRATE. We're assuming you have that knowledge already. If you don't have that knowledge base, this paper should still be helpful; however, it may not make quite as much sense. All the information in the previous PROC MIGRATE talk, "How to Migrate Your Data and Know You've Done It Right," can be found on the support.sas.com Migration Community at http://support.sas.com/rnd/migration/resources/procmigrate/index.html . The PowerPoint slides from the presentation along with the Speaker Notes for each slide are also located there.

And now, in the time-honored style of David Letterman, I present to you the top ten questions to ask before your data migration.

10. It's 10 p.m. Do you know where your data libraries are?
If you are only responsible for migrating your own libraries, perhaps you know exactly where they are. However, what if you are responsible for the data migration of your entire company? Some organizations have hundreds or perhaps thousands of SAS libraries, making the task rather daunting.

In order to find SAS bound libraries on z/OS, use the LISTCAT command in conjunction with the LISTDS command to create a list of potential SAS libraries. SAS bound libraries after Version 5 have the attributes of DSORG=PS and

RECFM=FS or FBS; Version 5 libraries have the DSORG=DA attribute. Additionally, they usually have the LRECL and BLKSIZE attributes set to 6144. However, the only way to determine whether the bound library is truly a SAS library is to examine the first eight bytes of that z/OS dataset for the characters "#SASLAAS". Often sites have specific programs that use the VTOC (Volume Table of Contents) which can be used to automate the location of OS datasets that have the attributes of SAS bound libraries. There are a number of shareware programs to be found on the web that can help with this task. This topic is discussed in much greater detail on the Migration Community.

For directory-based hosts, looking for expected filename extensions is the most expedient way to find your SAS libraries, but you need a comprehensive list of extensions. Previous to Version 7, the filename extensions were different for each platform; you need to look for those as well as the common extensions used on all platforms for Version 7 and afterwards. If you are on a PC and have ever used a fat file system, you must search for special three-character extensions as well. You may need to craft some Perl scripts or use another method to make a list of locations from which to migrate. There is a complete list of SAS library member name extensions on the Migration Community. After you know where your data libraries are located, you're ready to move on to the next step.

9. Has your library been abducted by aliens?

Or is your data simply hiding? Even when you know where your data libraries are, you still must take care when you assign your LIBNAMEs for migration. For directory-based hosts, PROC MIGRATE migrates a SAS library, not a directory. A directory is a physical location that may contain one or more logical libraries. This means that you must be aware of how many libraries are located in a single directory.

- Directories that contain more than one engine's members are known as *mixed-engine directories.* For example, you can have Version 6 data stored in the same location as Version 8 data. In that case, you need two separate migrations: one with V6 as the source library's LIBNAME engine and one with V8 as the source library's LIBNAME engine.
- Directories containing non-sequential members and sequential members also create a mixed-engine directory. For example, data created with the BASE engine and data created with the TAPE engine that are stored in the same directory create two distinct logical libraries which require separate migrations.
- A similar kind of mixed-use directory occurs on the PC platforms when a directory contains members with both short filename extensions and regular filename extensions. A libname assignment with the SHORTFILEEXT option

is necessary to migrate the members with short file extensions, separate from the migration with a libname without the option.

- Yet another concern is determining whether you still have Version 5 data. Many people do, especially with archived data stored on tape. If that data is not migrated before you remove Version 6 SAS from your system, it will become inaccessible. Because Version 5 data was supported through Version 6 and PROC MIGRATE came on the scene in SAS®9.1, PROC MIGRATE will not migrate from Version 5. The solution is to use Version 6 SAS' PROC V5TOV6 to convert to Version 6 and then migrate to SAS®9.

8. Blessed are the cheese makers?

With a nod to Monty Python, blessed are the peace makers indeed; but what about the trouble makers? Most of the members in your SAS libraries will not be a problem during migration. As explained in the previous PROC MIGRATE talk, your migrated data will have exactly the same data and attributes, including passwords, compression, indexes, integrity constraints, and so on. Here are some trouble makers to look out for:

- PROGRAM member types, DATA STEP views without the stored source code, and some catalog entry types will not migrate. See the "PROC MIGRATE: How to Migrate Your Data And Know You've Done it Right" paper for more details on the solutions for those members.
- If you are storing data sets with non-native data representations, after the migration they will have the data representation of your target operating system. If you did not want that result, you need to migrate that library using a means other than PROC MIGRATE.
- Another concern is operating system file permissions. Your user ID is set up to create files with certain default permissions. When PROC MIGRATE creates members in the target library, it must use the default permissions of your user ID and not the ones that were on the source library member. You may want to log the permissions during the original search for the data libraries and use that information after the migration to reset the permissions. This is something that PROC MIGRATE cannot do for you.

7. Should you have taken that left turn at Albuquerque?

Ah, if you had a map, you might know the answer to that one. And so, you also need a map for migration. After answering the question of where your source libraries are located, you need to map those libraries to their destination. You cannot migrate in place; you must map out locations for the new SAS®9 libraries. You also need to determine whether you have enough disk space to allow you to migrate and validate all your libraries at once, before you eliminate the source libraries. Or perhaps you want to process one user's libraries at a time. After validating that your new data library is intact, you may want to delete the source library and

replace it with the target library. Regardless, having a detailed map of your migration is a must.

6. If the bolt cutter won't work…what else is in your toolbox?

If you attended one of our previous PROC MIGRATE presentations, you know that "the bolt cutter" analogy was used to describe having the right tool for the job, with PROC MIGRATE being the right tool for migration. During that presentation, you also learned that PROC MIGRATE wasn't meant to migrate across platforms. If that is your situation, you have several choices:

- You might be most comfortable using the tried-and-true conversion routines, like PROC COPY, PROC CPORT/CIMPORT, or PROC UPLOAD/DOWNLOAD with the right options set for your situation.
- SAS/SHARE or SAS/CONNECT would also be good tools to use in this case.
- You can migrate across platforms in two steps, which can be done in one of two ways. Some customers have chosen to move their data to the new operating system and then use PROC MIGRATE at a future date. Others have chosen to first use PROC MIGRATE and then move their data to the new platform.
- You can simply FTP your data to the new platform and use CEDA (Cross-Environment Data Access) to access the data. Using CEDA to access a non-native member adds a slight overhead to translate the data in memory to the current platform's encoding. There are also some tasks you can't perform with a non-native file, such as updating. However, for some locations it might be the best fit.

5. Have you hugged your system administrator today?

It is not only important to know how SAS and PROC MIGRATE work, but also imperative to understand your own operating system. For example, it is extremely simple to use PROC MIGRATE on a UNIX-based system to migrate from one machine to another (in the same operating system family) using NFS, when the source and target libraries are NFS-mounted. One customer we talked with this year could not determine how to use NFS, despite numerous contacts with us. What they needed was help from their system administrator. In the end, the lack of communication between the customer and their system administrator led to a migration using PROC CPORT, FTP to the target machine and then PROC CIMPORT. It was an unnecessarily unwieldy migration. Get to know your system administrator; bring cigars, a box of chocolates, or even a box of chocolate cigars…whatever makes them blissfully happy.

4. SLIBREF…Is that a new official in the NFL this year?

With all the rule changes in the NFL these days, it might be; however in the SAS lexicon, SLIBREF is short for "server libref." This option specifies a libref that is

assigned through a SAS/SHARE or SAS/CONNECT server. If you are migrating across machines, this is especially helpful so that you can migrate in one step rather than two steps or more. As was explained in the previous MIGRATE presentation, in the 32-bit to 64-bit UNIX or OpenVMS Alpha migration, using the SLIBREF option is the only way for PROC MIGRATE to migrate your catalogs—even on the same machine. However, be aware that the migration validation tools do not use the SLIBREF libref to access the source library when you migrate from 32-bit SAS to 64-bit SAS. Therefore, the validation tools are subject to CEDA limitations. The %AFTER macro will produce meaningful validation output, but the %CHECKEM macro will produce lots of false errors because it won't have access to certain members and attributes. Note that the migration validation tools can be used even if you are not using PROC MIGRATE to migrate your libraries. In those cases, the same limitations hold true if the migration is cross-host. Be sure to check your log for ERRORs or WARNINGs in combination with using the migration validation tools for complete information.

3. Are we there yet? Are we there yet? Are we there yet?
Sometimes timing is everything. When can you start your data migration? Everyone's situation is different. First estimate how much time you need to complete the migration and validation. Now consider the fact that you must have access to unchanging data. You can't migrate a library that is being constantly updated. When can you take your users' data off-line in order to migrate static libraries? If you are planning to migrate in stages, inherent in this decision will be what libraries interact with other libraries. Only knowing what SAS processes are run on the data can provide you with this information. Checking the data itself is crucial as well. Imagine if two libraries are connected by referential integrity constraints. Those libraries need to be in the same stage of a migration. You need to plan which libraries to migrate as well as when to migrate unchanging data.

2. Who's on first?
"I don't know" isn't the right answer to that question - he's on third base. But seriously, who is in charge of your data migration? Is one person planning and performing the migration, or will each user at your company perform their own migration? Will you migrate some users' data before others? Is any data shared? Keep in mind that if only some users migrate to SAS®9, they can change group data such that Version 8 users cannot access the data. For example, a SAS®9 user could associate a long format name with a data set such that Version 8 cannot read that data set. On the other hand, migrating everyone at once may not be feasible. This leads us to the most important question of all…

*1. What is your **plan**?*

Data migration is an extremely important step in order to take advantage of all that SAS®9 has to offer. However, it requires careful assessment of your particular situation that only you can make, because your situation is unique. Answering all the important questions—who, what, where, when, and how—is imperative before you begin your data migration. The more users and the more data at your company, the more extensive your planning must be.

We at SAS are always here to help answer your data migration questions. Please don't forget the compendium of migration information to be found in the Migration Community under support.sas.com. Technical support is happy to answer your migration questions as well. If you have other questions or suggestions for PROC MIGRATE or data migration in general, please email me at Diane.Olson@sas.com. If you have data validation questions or suggestions, please send them to David.Wiehle@sas.com.