Tips and Tricks for Organizing and Administering Metadata

Michael G. Sadof, Bedford NH

ABSTRACT

The SAS® Management Console was designed to control and monitor virtually all of the parts and features of the SAS Intelligence Platform. However, administering even a small SAS Business Intelligence system can be a daunting task. This paper will present a few techniques that will help you simplify your administrative tasks and enable you and your user community to get the most out of your system. The SAS Metadata server stores most of the information required to maintain and run the SAS Intelligence Platform which is obviously the heart of SAS BI.   It stores information about libraries, users, database logons, passwords, stored processes, reports, OLAP cubes and a myriad of other information.   Organization of this metadata is an essential part of an optimally performing system. This paper will discuss ways of organizing the metadata to serve you organization well.   It will also discuss some of the key features of the SMC and best practices that will assist the administrator in defining roles, promoting, archiving, backing up, securing, and simply just organizing the data so it can be found and accessed easily by administrators and users alike.

BI ARCHITECTURE

At the core of the SAS Intelligence Platform is the SAS Metadata Server.   Metadata is data about your data like who can access the data and where the data is stored. The metadata server basically controls all access to the metadata including users, libraries, reports and Stored Processes. Most of the metadata is stored in a set of SAS files known as the 'Foundation Repository'. The repository contains library definitions, report definitions, user logons, database passwords and much more.

The SAS Management console is a client tool that enables you to monitor and maintain the metadata. It acts as a gate-keeper and catalogue of the information that is stored in the SAS BI system.   It can perform backups, control access to data, organize and promote objects within the repository.   Figure 1 shows a familiar diagram of the key parts of the SAS BI Architecture.
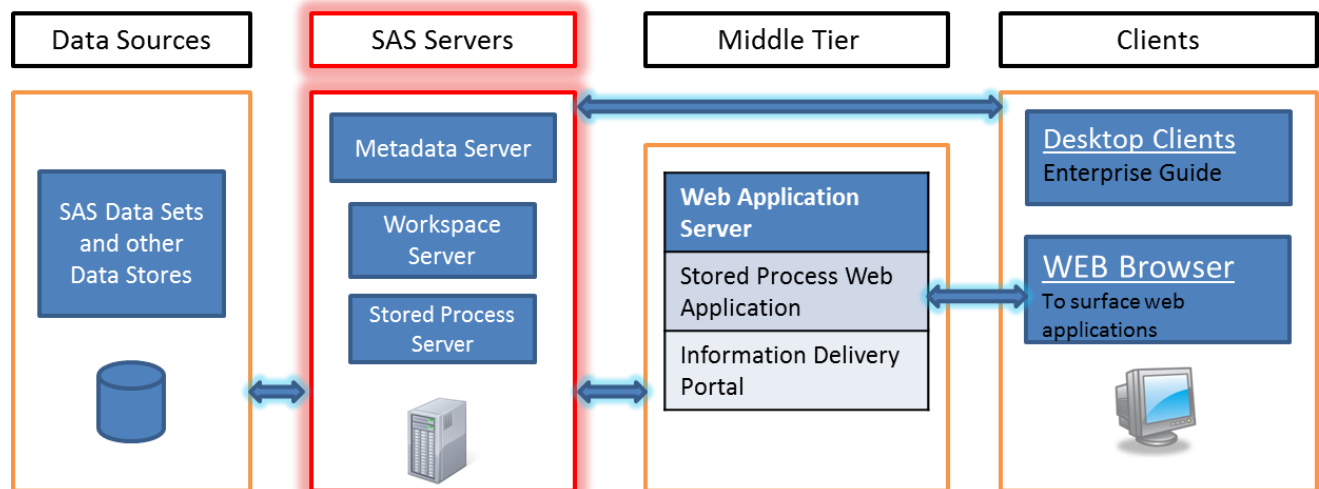

Figure 1.   SAS BI Platform Architecture

THE SAS MANAGEMENT CONSOLE (SMC)

SAS Training teaches an excellent 5 day course on administration of the SAS BI platform.   Presented in this paper are some important concepts on how to administer without delving into the specifics which you can find elsewhere. The key items that you can manage in with the Management Console are basically where things are stored and who can access them. Here are some examples:

          Users Access
          Definition of servers
          Definition of databases and external files
          SAS Libraries
          Reports
          EG Projects
          Stored Processes
          SAS programs
          Data Integration Studio Jobs
          Cubes

While security and access are always a big part of administering a system another important factor is the organization of data within that system.

FOLDER STRUCTURE

As an administrator it is important to create a framework that users can interact with and build upon in an organized manner.   Best practices dictate that information and programs should be stored in a central location that has controlled access, rather than on a particular user's laptop or thumb drive.   A basic folder structure is provided at installation but certainly you will want to modify and enhance that structure.   While the folder structure (as shown in Figure 2) is similar to a directory structure that you might find on Windows or Unix it is not the same. It exists only in the Foundation Repository.   You may store a program or EG project either in the physical directory of your operating system or within the folder structure of the metadata.   It is important to emphasize this concept to your user community and develop some guidelines as to where the information will be stored.   When the data is stored within the SAS Folder structure access to the data can be controlled via the management console. As an administrator you should define for your organization where data and reports and programs should be stored.   It can certainly be a combination of the two but whichever framework is decided upon it should be enforced throughout the user community.   Just as books can be arranged on a library bookshelf by author, title, or topic you must create a structure for your metadata by department, functionality, or user.   Figure 2 displays a folder structure for a fictitious XYZ company with a 3 departments (Data Analysis, Marketing, and Sales).   Note that within each department the objects will be further organized by function.   You can create folders and subfolders by simply right clicking on a level much in the same way as is done in windows.   By setting the appropriate permissions on each folder you can control who can view and who can change.   In this manner each of your user groups can have their own unique set of folders. I find this is a very simple and effective method to control access and organize.   SAS provides several other methods of segregating data which may ease the burden of promoting and migrating code.   You can set up multiple metadata servers each with its own foundation repository or you could set up multiple repositories within the same metadata server.   The repository is basically one big file that contains all the information.

As an administrator you are responsible for the maintenance, functioning and protection of the data within the system. Data access can be controlled at a folder level so good folder design with ease the burden of maintenance as well as provide for accountability.   It will also provide for better collaboration of data, programs, and reports.

Typically a SAS BI system is installed with segregated if not separate environments such as Development, Testing, and Production.   If separate physical servers are not possible then the appropriate implementation of folder security can isolate each system and provide for safety and accountability.

Folder structure becomes very important if you are maintaining multiple environments (dev, test and prod). Usually a program, cube, or other object is created in the test environment and then gets promoted through the other environments.   The first promotion to test can be simply to move everything in a particular folder from the dev environment to test.   But as a system grows there will be some new development, some fixes, and perhaps some deletion of objects.   I find the easiest way to determine which members are to be moved is to place them in multiple folders.   So you might have folders labeled by release number so you can determine which objects are to be moved. Another method might be to demote or checkout everything from the prod environment and promote the whole package again.   You really need to see how your developers and users will be operating before you can make a decision.   Another idea is to name the member objects appropriately with a date and select objects based on that name in the export wizard.   You can select based upon name, date created, and type of object.
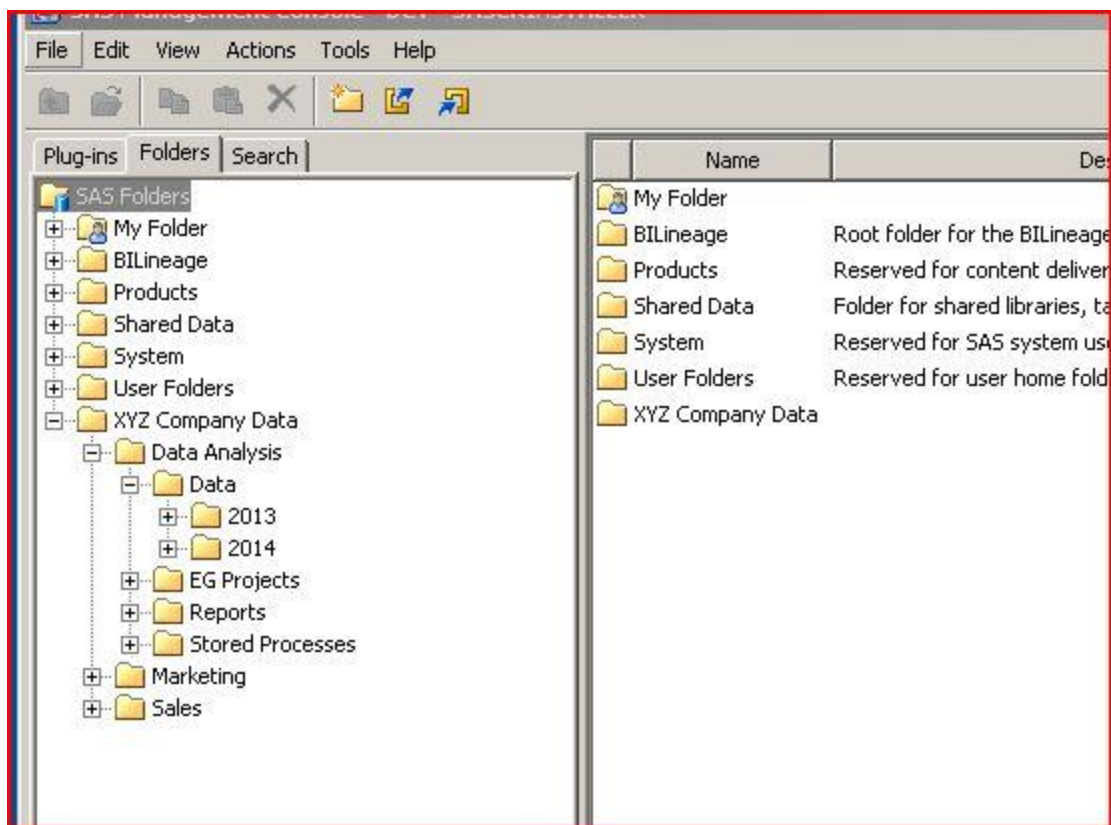
Figure 2. Example Folder Structure

PROMOTION AND MOVING DATA BETWEEN ENVIRONMENTS

The concept of moving or promoting data from one environment to another is easily accomplished in the management console but pre-planning can assist in keeping track of the various pieces. Basically the idea of promoting code or reports is accomplished through creating export packages and then importing those exported packages to the new system or location. Typically a large SAS implementation will have development, test, and production servers. The code, projects, and reports can be created and unit tested in the development environment and then moved to the test and subsequently the production environment. The management console provides a method of copying or 'exporting' metadata from one environment to another. In the export process you can create a system independent package of information that can be imported to the next environment. You can create these packages easily in the management console by right clicking on a folder or object and selecting "Export SAS Package". The wizard will then give you the option of selecting all the objects or allows you to filter the list of objects to be exported. Here is where a good folder structure can help. If all of your marketing programs are stored in the same folder you will have to select the whole folder or hunt for the particular objects you need. By creating sensible sub-folders (let's say by project or by date) you can make this job a bit easier and repeatable. Figure 3 depicts the filtering dialog of the export process.

Once the package is created you must physically move that package file to the next environment if the environment is not on the same server. The import process requires that you place your cursor over the appropriate folder before right clicking and choosing "Import SAS Package". During the import process you will be prompted to assign servers, libraries, and physical locations since they may be different in different environments.

The BI system also provides the capability of creating multiple repositories or multiple metadata servers. It is beyond the scope of this paper to explore these methods but the administrator should be aware of them. I have found that appropriate folder creation and security is generally sufficient for most systems.
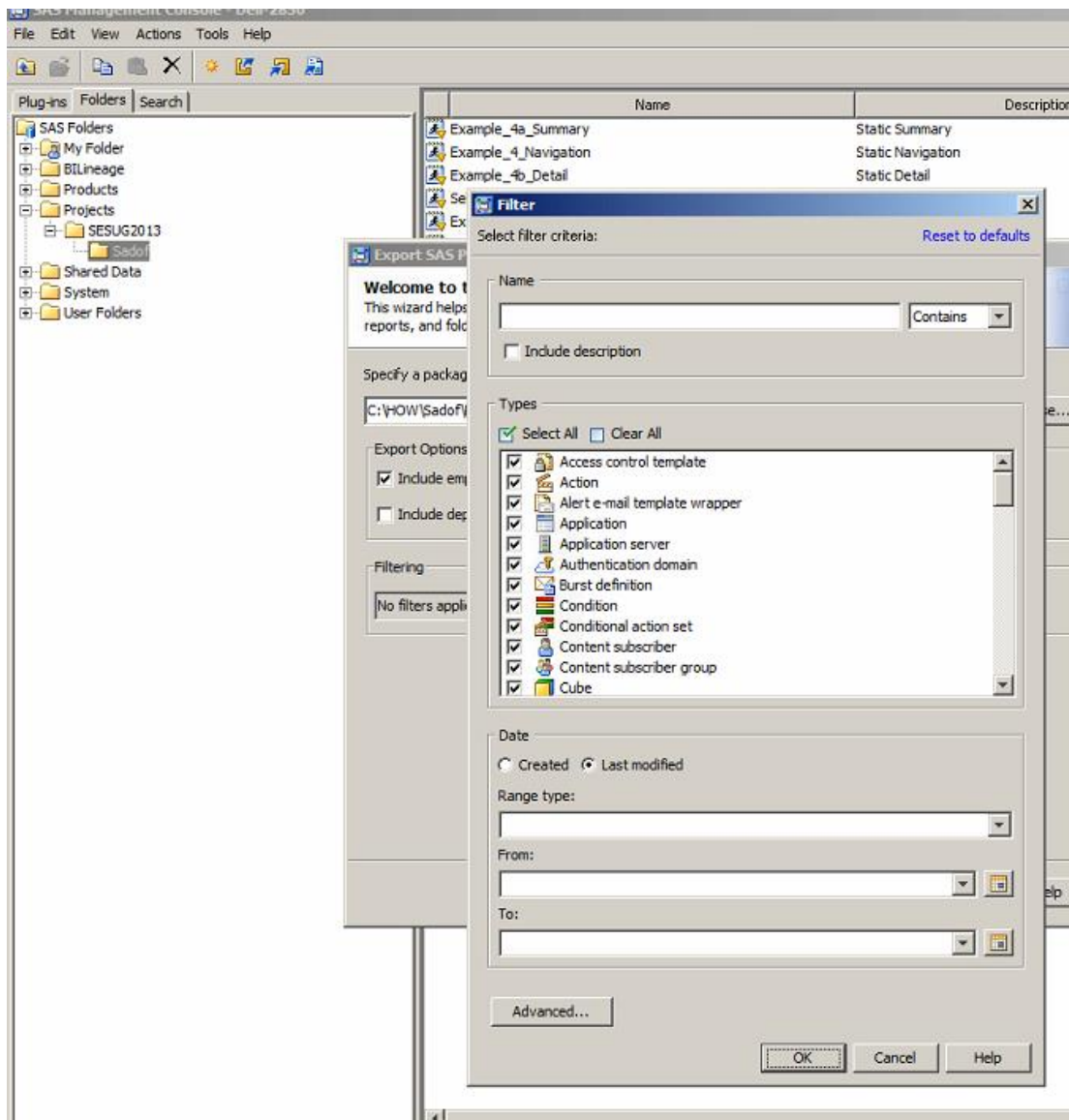
Figure 3. Export Filter Dialog

USER AND FOLDER SECURITY

One of the most important aspects of the metadata server and management console is the ability to restrict access to certain folders and objects.   This can be through role based or group based security.   The basic BI installation provides most needed roles which can be used to restrict or grant access by function like the ability to drill down into a cube.   Group based security is more for controlling the data like the access to a particular set of tables/data files that may contain sensitive information. I have found that the SAS supplied roles are sufficient for most needs and can easily be altered and customized for your particular situation.   Groups on the other hand are very specific to your organization and you will want to build your own.   Best practices indicate that you should not alter or delete the groups supplied by SAS like SAS Servers or SAS Administrators, but rather add your own as necessary like "programmers", "Market Analysts" etc. SAS provides the ability to control whether a member of a particular group can see the folder, read the objects within that folder, read the data associated with those objects or modify or delete those objects or data.   You can set the security on a particular object by right clicking and choosing properties and then clicking on the authorization tab

BACKUPS

The SAS 9.3 BI platform now provides a way of automatically scheduling your metadata backups. It also provides a method of restoring your repositories from a previously saved backup. I cannot stress enough the necessity of backups. You should allow the system to backup your repositories on a regular basis but I would also recommend that on a weekly and monthly basis you save backups in an alternate location. The automatic backup system will only keep a few generations of backups. I would also recommend backups of the data directories on the production server before any promotions or migrations. Export packages can be used for backup or archival purposes as well. You can export the whole folder structure to a package file and label it with a date. The package files are fairly compact and worth their weight in gold if there is a problem or metadata becomes corrupted. It is also possible to copy the entire physical repository directory and restore from that but you will not be able to do a selective restore. For a selective restore you can create a backup using the export function and then selectively import the objects you need. I find it very helpful to keep several full export packages at a folder level. Invariably a user will lose an object and need a restoration.
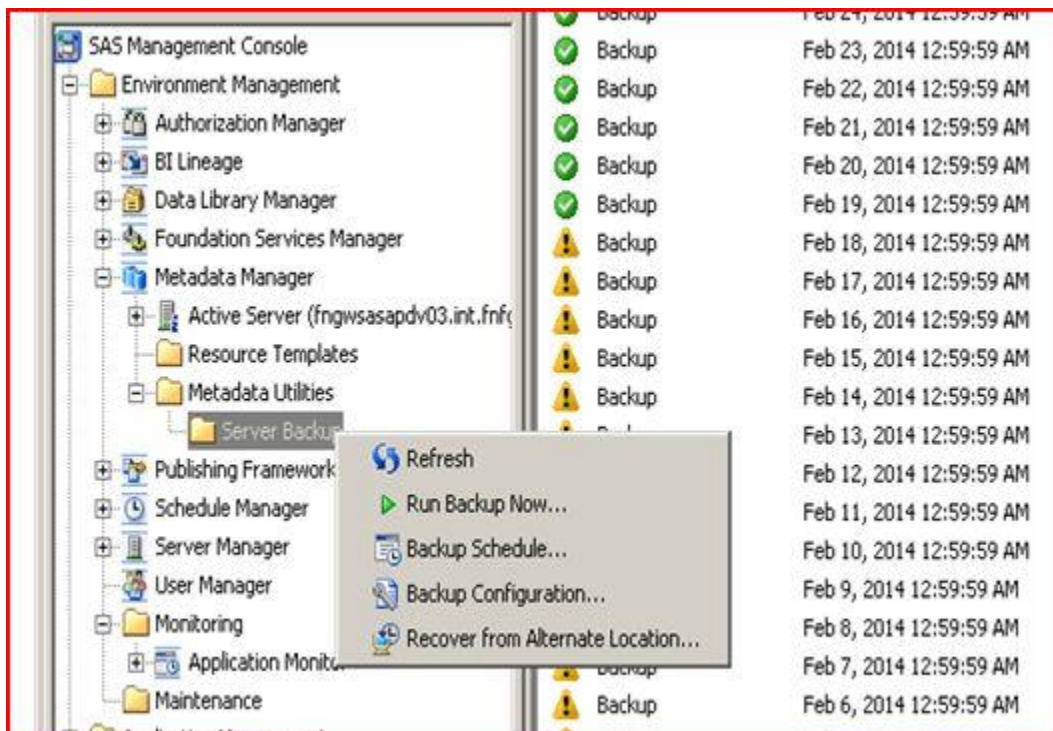

Figure 4. Backup Dialog

CONCLUSION

In conclusion I would say that the most important factors in creating you environment are user access, folder structure and solid backup strategy. If the folders are created properly then access to the folders and subsequent promotion of the content of those folders is easily accomplished. With a solid and sensible folder structure you can selectively give access, promote members, and restore objects easily.

RECOMMENDED READING

- *SAS(R) 9.3 Intelligence Platform: System Administration Guide*

CONTACT INFORMATION

Mr. Sadof has over 25 years of experience as an Information Technology professional and currently maintains four levels of SAS Certifications. Most recently has been involved in design and development with the SAS Business Intelligence products, optimization and data warehousing projects, information visualization and the presentation of results on the web. He has been involved with installation and maintenance of SAS BI systems including 9.13-9.4

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

Michael G. Sadof
MGS Associates, Inc.
Bedford, NH
301.613.3987
mgs@mgsnet.net