So You Want to Install SAS®?

Rafi Sheikh, Analytiks International Inc., Minneapolis, MN

ABSTRACT

Demonstrating the importance of preparing for a SAS install using SAS Office Analytics® bundle as an example. There are many nuances, as well as requirements, that need to be addressed before an install can occur. The majority of these requirements and preparations are similar yet differ slightly according to the target install Operating System, whether it be for Windows or *Nix . Our discussion will focus on these three topics:

1. Pre-install considerations such as sizing, storage, and proper credentials
2. Third party requirements
3. Install steps and requirements, and post- install configuration

In addition to discussing preparations for an install this paper also addresses potential issues and pitfalls to watch out for as well as best practices.

INTRODUCTION

Any well executed plan is based on carrying out the most fundamental and basic components flawlessly. A SAS install is no exception. Here the fundamental building blocks of a successful SAS install; whether it is a single machine or multiple, serving one end user or scores of them, all depends on how well the basic components are carried out.

We will use loosely the SAS Office Analytics bundle as an example to talk about necessary preparations for the install. SAS Office Analytics (OA) generally has SAS/Foundation®, SAS/MetadataServer®, SAS/ACCESS®, ODBC, and SAS/STAT® modules bundled together. The purpose for using a bundle is not to specifically prepare addressing install of this particular bundle but to provide some sort of a preparation reference or framework.

OA supports single or multiple machine configurations i.e. in multiple you may install SAS Metadata on one node and SAS compute on a different node. This bundle does not include any SAS web applications however, a short discussion and link is provided for further information on Web application servers and SAS requirements in one of the following sections. SAS Installs are divided into the following phases:

1. Pre-Install
2. Install
3. Post Install

PRE-INSTALL

Installing SAS is like cooking a meal, just as it is important that before we start cooking, we have all the ingredients ready, so is the same with an Install. We must have all the pre-install requisites gathered, assembled, and acted upon for the particular SAS install to be completed successfully.

PROFILE YOUR INSTALL

It is important to know the system which will host the SAS environment, the users who will use the platform, and how the platform itself will be used. The main focal points for the profiling exercise are:

a. **End User**:
   i. **Work Habits** Profiling end users provides a base line and an insight to an average user's psyche and work habits. It is important to understand the work whether it's good, bad or inefficient since that is the same behavior that will be practiced in the new SAS environment and provides opportunities for improvements. The user profile will provide insight from a number of end users that the install is expected to support, what kind of programming and processing that these end users require? The answers in turn will give insight to type of performance the new SAS install is expected to sustain.
   ii. **Pain Points** Awareness of any pain points that these users face such as performance, access and programming in-efficiencies provide opportunities to address issues that often are not surfaced and are typically unknown to the management.
   iii. **Process Awareness** Another outcome is process awareness and an understanding of how things run

within a certain entity. This type of information becomes an input for the security model that will be needed post SAS install. Profiling also provides an insight to the overall end user technical skill level. This allows for grouping and understanding of roles and responsibilities which are an integral part of any security model.

b. **Data Sources**: Profiling data sources provides an opportunity to develop a comprehensive strategy for accessing data.

 i. **Data Access Strategy** This addresses how data libraries will be created; global or local? Will it be efficient to use an explicit pass through, implicit pass through or a blended approach? Which data sources should be provided on a global level versus group and individual user level. How will the data libraries be assigned, via Metadata, auto-assign or autoexec. Based on these decisions, this will determine if there will be any performance considerations. In the case of third party databases, how will access be provided and managed? Will it be via creating login profiles; groups or individual credential basis.

 ii. **Third Party Databases** Listing of all databases allows for gathering of logistics per database such as access policy; will users access the database using a functional or common set of credential or individual user credentials. Logistics include information regarding database schema(s), system DSNs, client installation, verification of connection to target database independent of SAS, if any drivers are needed, and if non-traditional drivers are used; such as EasySoft® additional configuration tasks.

c. **Usage**: Profiling usage is done to gain understanding of resource and performance requirements. Typically, profiling includes these questions: number of data steps in a typical code, how many sorts and merges, what type of statistical and SAS programming features are being used, how many SAS session concurrently are used, number of programs typically run by an end user. Answers to these questions provide insight to the performance requirements that the SAS environment is expected to provide. Information gathered becomes an input in the performance design; Dynamic allocation of temporary space (if needed), type of storage to use, minimum number of CPUs and memory needed to sustain such performance. Often this information becomes an input for OS and hardware configuration decisions.

A successful profiling exercise provides input to performance requirements, security needs and data access strategies. If explored further, profiles may also be helpful in pinpointing any process flaws, training gaps and opportunities for introducing efficiencies.

DESIGN

At this juncture all profiled information manifests into different design elements, such as:

1. System Resources Design: The SAS Install system resources such as hardware, operating system (OS), number of CPU and amount of memory (RAM) are the focal points. The information from end users profiling provides usage estimates; these numbers may change. Information per SAS version and OS are available at http://support.sas.com/resources/sysreq. When designing system resources, it is important to note that each OS brings with it some nuances that range from particular OS patch level requirements to building proper servers. This also includes CPU tuning, IO, storage, and memory considerations. For example, let us consider Windows 2008 R2 64 bit server; The CPU tuning aspect includes consideration for the power level set on the server (C to C0 – C06). C0 and C1 settings mean all processors are to be available at all times while each stage after C1 are at various level of idle and sleep. The decision needs to be made collaboratively as to which C state to keep the server's processors. In the same scenario, the storage array configuration settings matter; for example in case of multiple LUNS it is better to not stripe at the OS level. The type of file system chosen has its own impact or lack thereof; using NTFS is a recommended type of storage however the choice of storage array has a significant impact on performance.

 o In short, while designing it is important to work with your IT administrator and to research and gather information relevant to that target operating environment. The following are recommended categories: CPUs: Based on number of users, concurrent SAS sessions and active SAS sessions at one given time. CPU Clock speed and number of cores are important considerations, a general rule of thumb is 5-8 users per CPU; based on usage.
 o RAM: A general rule of thumb here is 4 GB/CPU.
 o IO Subsystem: This will be based on usage pattern. Work with the storage personnel to determine the best configuration and sustained IO. Sustained IO rate is important.
 o SAS Configuration: Options selected and used are based on all design decisions to tie varied components like IO to CPU together and leverage OS and hardware level components.

2. Security Model Design:. It is advised that the SAS security model works in tandem with the security implemented at the OS level. In case it does not, restrictions applied on SAS metadata will be in-effective if the security permission on the OS level is wide open and vice versa.

3. Infrastructure Design: The profiled information provides great insight as to the SAS environment infrastructure. Infrastructure includes directory structure, end user groups, security model implementation on the OS as well as

surfacing it in SAS Metadata.

4. <u>Performance:</u> A critical item for any environment and SAS is not an exception. Decisions such as CPU/RAM, network, and IO are critical elements. IO considerations vary depending on storage system used. The main concern when setting up a file system is to ensure that SAS gets the sustained I/O needed to complete the SAS jobs in the timeframe required by the end-users. The I/O rate for SAS data varies by process

   o With the large, block IO that SAS handles the IO throughput rate is very important. The I/O throughput rate for SAS depends on what SAS tasks are being done: For most Extract, Transform, and Load (ETL) processes, 50-75MB/second of IO throughput for each file system SAS is using is required per SAS session.

   o For end-user exploitation, 15-25MB/second per user per file system is required. However, for heavy SAS users, 50-75MB/second per user per file system is required.

   o Disks allocated to SAS working files - including "SASWORK" require 15-25MB/second per user for most SAS processes with heavy SAS users requiring 50-75MB/second per user. Disks allocated for SAS utility files for SAS threaded procedures and OLAP processing - "UTILLOC" - require 50-75MB/second per user

Type	Sustained IO
ETL	50-75 MB/sec/Session
WORK (Temporary area) and Utility Files	50-75 MB/sec/Session
End users	15-75 MB/sec/Session

Table 1: General IO throughput requirements

SAS PRE-INSTALL REQUIREMENTS

Once design and hardware resources are agreed upon the next phase begins. At this juncture we know what type of server and OS will host the SAS install. The next steps are to gather the ingredients needed for the install. These are:

- SAS Depot
  - o Depending on the SAS install, a SAS Depot would require anywhere from 4 GB to 30 GB of space. Typically the storage requirement is provided by the SAS Software Order communication. It is preferred to have the depot local or as close to their server as possible. Having the SAS depot on a commonly accessible host is also an option for enterprise access; however, it is imperative that the traffic and network bandwidth do not become a hurdle. The SAS Software Order email carries all the necessary information to download and create a SAS depot. Typically this email will be sent to the SAS representative for the company.
- Environment

Categories	Detail
SAS System Required IDs	SAS install related IDs are an integral part of the SAS environment. Based on the SAS bundle that is installed, the minimum IDs are: o sasadm@saspw = SAS unrestricted user (admin)-internal account o sassrv = Used with Stored procedure o sastrust@saspw = Used with authentication (internal sas account) o sasdemo = First user, after an install normally end users are not added yet and this ID helps in validating the install. **Note:** *apart from the internal sas accounts (@saspw) all IDs in the case of a WINDOWS environment may have a domain such as domain\sasdemo to the IDs. Internal accounts do not need creation at OS level.*
Group	o In the case of a Windows environment, it is suggested that a SAS Server User group is created with a 'Log on as Batch' right assigned. All end users who wish to access the SAS environment become members of this group. Members could be individual users or (domain) groups. *On a system where Windows Integrated Authentication is used, this requirement can be ignored.* o On the *Nix environment, the information gathered becomes an input to the OS security design in terms of groups and permissions. Primary and secondary group access will need to be addressed. It is good practice to keep the installer ID groups separate from end users' and careful consideration should be given to secondary group membership.
SAS System Required Ports	Depending on which SAS bundle you are installing ranging from SAS VDD to SAS (E)BI, particular ports are required by SAS. SAS provides a range of ports that is recommend per component. It is necessary

Categories	Detail
	that ports are available and registered to be used with SAS. In case a port is not available, you may assign an available port and ensure that the correct value is provided during install, and that the port 3 is documented. See appendix A for a detailed listing.
Firewall	It is important that all firewalls on the server to the depot (if not local) are turned off and in case of multiple SAS servers in between the firewalls will need to be turned off as well. If that is not possible, ensure that the network administrator creates an exception to allow traffic between target servers.
Antivirus Software	If possible have antivirus software turned off at time of install if this is not possible then have an exception rule created.
Installer ID Rights	It is preferred that the install occur under a different set of credentials than of a person or an administrator, even in the case of a single user. Example: SAS
Mail	Get mail related information from the administrator. The information needed is: o Mail hosting server o Port o Does mail require authentication (yes or no) o Who the (alert) email will be sent to (admin or a mail box)
JAVA/JRE/Junit	SAS provides its own JRE. JDK is used when the install contains mid-tier or SAS Web application components that can be downloaded from the Oracle site. Note, it is advised that you have a profile created on the Oracle site as you will be asked to provide that information at the time of the download. All third party information and appropriate download links are available at http://support.sas.com/resources/thirdpartysupport/ You will also find Junit download information at the URL above. Junit is used with Deployment Tester that is used to validate an install.
Install Location: • WORK • SASHome • Config • Data	Work with the IT infrastructure team to ensure that relevant storage shares are accessible accordingly. SASHome is where all the executables and binaries are installed to and is also host to "SASFoundation". WORK, of course, is the SAS temporary space and "config" is where all the Lev-N configuration, documentation and logs for the SAS logical servers reside. Consideration should also be given to which storage share will host the data structure.
Data Sources, Drivers, and Clients	o From data source profile information, ensure that all the drives and clients are installed for any databases such as Oracle, DB2, Sybase, etc. o Ensure that all the database connections are independently verified.
Plan File	Your SAS representative would have received emails from SAS containing order information and SAS Depot download instructions. For multiple machine, complex SAS installs it is often necessary that custom plan files are created and provided by SAS. However, SAS does provide a standard set of plan files for different software bundles and machine combinations which could be utilized during an install. The standard plan files are at: http://support.sas.com/demosdownloads/setupcat.jsp?cat=Standard+Deployment+Plans
OS Patches	Based on the operating system that will host the SAS install, it is important to pay attention to the OS patch level requirements. There are always nuances that are particular to an OS. For example, on WIN 2008 R2 there are particular patches that address Low Memory Foot Print issues. However this issue has been fixed going forward with SAS 9.2M3. In other operating systems such as HP Itanium, AIX, Solaris, and Linux Red Hat there are different requirements for OS level patched that SAS requires and it is important that the System Requirement guides are consulted. Read more at http://support.sas.com/resources/sysreq/
Passwords	Having a password policy is beneficial. It is recommended that some of the passwords that are used with SAS install related IDs are non-expiring and non-changing. During installation there are mainly two usernames that are created and used behind the scenes. These usernames are SASTRUST and SASSRV. SASTRUST is used for backend trusted communication between the different servers, (i.e. MetadataServer to Workspace Server). SASSRV is an account that is used for pooled servers such as the Pooled Workspace Server and Stored Process Server. If a password automatically expires, then elements of the SAS implementation will no longer work. If passwords need to be changed, then it is suggested to do so during a planned outage and use the Deployment Manager to update passwords each time a password expires. You can read more at http://support.sas.com/documentation/cdl/en/bisecag/63082/HTML/default/viewer.htm#p0kb2gtnuy

Categories	Detail

	jnrnn1hduu5by88u4f.htm
	
	Note
	As part of doing an install of an external utility, sasauth, manages authentication and runs setuid as root to access user database. As a result, during installs on any of the *Nix hosts, a step prompts for the root password to change permission to 4755 as a root (super user ID) with the option of completing the steps after the install. Normally that root password is not available to the installer ID. Some methods of addressing this requirement are:
	
	• Have an admin present who has the appropriate privileges to input the password for you, or
	• Post install have the installer ID granted a sudo su - <installer-id> access by the admin to do the following steps:
	
	      su root
	      # cd !SASROOT/utilities/bin
	      # mv setuid/* .
	      # chown root elssrv sasauth sasperm
	      # chmod 4755 elssrv sasauth sasperm
	      # exit
	
	• Sudo su as root and execute the script at: !SASROOT/utilities/bin/setuid.sh

Note: Although SAS Office Analytics does not include a mid-tier component, any other SAS install that does have SAS web applications ought to consult SAS guidelines for Web and HTTP servers. There are currently three Web application server types; Jboss AS, IBM Websphere, and Oracle's Weblogic. Note that Jboss is also available in GA (general availability, freeware) SAS will require the AS (supported version by Red Hat). Read about specific requirements and bench line patch levels at: http://support.sas.com/resources/thirdpartysupport/v93/appservers/index.html

INSTALL

Once all the ingredients are gathered and combined they will amount to a flawless Install. Keep all the gathered information in one place. Once the execution starts make a provision to record the install. There are many different ways to record an install, some methods include; manual screen shots and recording the entries to creating a screen recording of the install. Camtasia Studio is a useful tool to record installs and some use SnagIt® as a tool for capturing install steps. In addition, a response file SDWResponse properties is also created where all the values input during deployment are recorded. Typically this file will be in the install ID's home directory.   This is a simple hidden text file.

Tip
Under the installer ID's home directory there are many SAS Deployment Wizard related files that are helpful in investigating install failures or issues. These files are normally hidden on the *Nix environment doing a ls –la and on a Windows environment set options to view hidden files. Typically on Windows the path is C:\Users\admin\AppData\Local\SAS\SASDeploymentWizard\...

Components	Details
SDW.log	Log contains information about the SAS Deployment Wizard itself.
Bootstrap.log	Carries information on the SAS logical servers especially on Windows environment where the servers are often run as services.
SDW_<date>/log	Contains information pertaining to all install activities
sdwprefs.txt	This file contains information on SAS Deployment wizard configuration and system information. Note: On environment's that host multiple SAS installs under the same installer ID ensure that this file is saved and renamed. Otherwise after the first install all information will be re-used taking input from the previous install.

VALIDATION

Once the installation is completed, SAS deployment wizard generates an html file called Instructions.html that has all the generated install information as well as validation steps which will lead an installer to verify a successful install.

Typically, Instructions.html will include all the listings of the SAS Application Servers including Metadata server that are installed, their paths, paths to the logs and validation steps. Instructions.html will automatically start as soon your install is finished, however if it fails to do so you may have to manually open the html file. Normally the file location is <sasroot>/config/Lev1/Documents   Follow the steps in the Instructions.html to validate your install. SAS Deployment Tester also plays part in validating the install by running a suite of tests that looks at the SAS Foundation to IQ and OQ which is particularly relevant for Clinical/Pharmacy clients.

Instructions.html also provided information on the default backup setup that was created during the install. Following the Instructons.html document provides information about the backup location, setup and schedule which are all changeable from the default settings.

Before moving on to the post install phase another important step is configuration. From profiling end users and having discussions on the system, resource estimation and the design phase SAS configuration settings can be determined. Some of the important options are:
   A. MEMSIZE: Default allocation is 80% of available memory however, having large memory settings does not translate into better performance. Recommended values are 512M to 2 GB
   B. SORTSIZE: Default value is 64M, suggested values between 256M to 512M unless you have large physical memory available, than you may go higher than 512M
   C. BUFFSIZE, IBUFFSIZE, UBUFFSIZE: The setting needs to be in conjunction with storage setup and to be set different from default when doing large sequential reads.
   D. SGIO: Most relevant in WIN environments and if working with SAS9.2M3 or higher.
   E. WORKTERM: Specifies whether SAS erases WORK files at the termination of a SAS session.
   F. WORKINIT: Initializes a new work session library
   G. WORK: Important if using dynamic storage allocation as well as when default location is not acceptable.
   H. UTILLOC: Initializes a new work session library

Best Practice Use of WORK and UTILLOC
Starting with SAS 9.2, SAS has provided options to allocate WORK sessions dynamically to multiple separate file systems that are hosted on separate logical or disk volumes. The goal is to minimize footprint collisions. Divide and conquer holds true. This setup works best for a large number of end users. There are several techniques that could be utilized to spread the work, however, SAS provides two algorithms RANDOM or SPACE (least used). You may read more at: http://bit.ly/dynamicwork.

The same concept that may be used with UTILLOC specifies where the SAS utility files are written. The default is to use the WORK location. In high usage environments it is better to spread reads and writes. The file system considerations will be the same as WORK.   The best places to specify the settings is either at the individual SAS application server level, such as <*BIConfigDir*> /Lev1/SASMeta/MetadataServer/sasv9_usermods.cfg or <*BIConfigDir*>/Lev1/SASApp/StoredProcessServer/sasv9_usermods.cfg or to apply globally at the <*BIConfigDir*>/Lev1/level_env.sh as a TKOPT environment variable (TKOPT_ENV_UTILLOC). If you define the UTILLOC at <BIConfig>/Lev1/SASApp/sasv9_usermods.cfg level for global application, it will be ignored.

Multiple locations may be used where SAS THREADED application could access the space in. However, a single utility file will not span multiple locations.

The latest trend in managing storage is virtual storage pools. Work closely with your storage administrator to get the sustained IO which is still king in performance considerations. The benefit of virtual pools is that they expand dynamically and if done correctly may not require any dynamic WORK allocation.

Third Party Database Access Configuration
   A. Verify all target database drivers and clients are installed and the connection to target databases is verified independent of SAS.
   B. Using SAS/ACCESS modules that were installed configure the access to the target databases according to each database flavor. Configuration steps vary per database. Configuration requirements and steps are provided by SAS at http://bit.ly/OjtxP5 (Select relevant OS and then in next page select System Requirement for SAS 9.x Foundation for <OS>

Information about each of these SAS options may be obtained at: http://bit.ly/N4Btkj

POST INSTALL

Design Implementation At this juncture the remaining part of the profiling that was done comes into action. At this

stage, implementation of all data access setup, SAS security models and directory structure starts. This stage ends with surfacing all appropriate information in SAS Metadata.

On-Boarding Once all data structure, end user additions, data access, and security designs are implemented, the next and important phase is on-boarding of end users. On-boarding is preparing your end users for the new environment. This includes providing basic information such as new server names, login information, data structure, and new access policies; if any, permissions, any group designation or roles changes, data access methods as well as a listing of what they ought to bring along from the old environment to new environment (code, data, report, etc.) along with support contact. End users get a review of the documentation and possible "migration" dates. In a nut shell, all relevant information that will provide a road map for the end user to navigate without much trouble in their new environment.

Migration This is the step where end users move their code, any relevant historic data and any identified item that they would like to move to their new environment. Normally, support is provided during this transition period with navigating in the new environment, moving objects, and any technical issue(s). If planned correctly, this migration can be painless.

CONSIDERATIONS

In summary, the following represents a list of considerations the customer should account for prior to the SAS installation.

1. Create and mount disk volumes. Within these volumes create necessary folder structures to account for user and group folders if desired, folders for SASCode, Stored Processes, SASOutput, etc.
2. Apply necessary operating system security to these folders. SAS has metadata security, but that does not stop someone from remote connecting to the server and navigating to these folders.
3. Install and configure whatever database client you are going to use. Make sure that the appropriate databases are defined to the client and that they can be accessed with the client or a tool like SQL+.
4. If using ODBC as the RDBMS access method, make sure that appropriate data sources have been defined to ODBC. Install any necessary drivers. Make sure that the databases can be accessed by other third party ODBC compliant applications as well.
5. Make a note as to the ODBC definition names and database names are if using actual RDBMS clients as well as any username and passwords.
6. Determine where SAS output will be written, if to a web server; where is that web server installed and does SAS have access to it?
7. Determine User Name and Group Name naming conventions and what resource they will have access to.
8. Ensure that the software depot has been successfully downloaded and that it is accessible from the server in which the installation will occur. Do a test transfer of files between the software depot and the SAS server to ensure adequate throughput for the installation. If not, then the depot will have to be copied locally to the server.
9. Ensure that the appropriate users and groups are created on the OS with the appropriate permissions as defined in the pre installation checklist.
10. Make sure the OS has been patched with the most up- to-date patches.
11. Having SAS Depot downloaded locally to the install server is ideal, otherwise ensure it is downloaded to a common/accessible storage area

Installing SAS is like cooking a gourmet meal, the more thorough in gathering and organizing your ingredients the better the cooking experience will be. Happy Cooking!

RECOMMENDED READING
- The indices at the end of this document provides more information on ports and general considerations http://support.sas.com/documentation/installcenter/index.html
- Windows 2008 Server Low Memory Footprint Issue: http://support.microsoft.com/default.aspx?scid=kb;EN-US;974609

To get more information and or templates for profiling please contact the author.

CONTACT INFORMATION

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

Name:	Rafi Sheikh
Enterprise:	Analytiks International, Inc
Address:	10 S 5th Street, Suite 720
City, State ZIP:	Minneapolis, MN 55402
Work Phone: 612.305.4312	
Fax:	866-347-2021
E-mail:	rsheikh@aii-3.com
Web:	www.aii-3.com
Twitter:	@analytiks

INDEX – A

Port Name	Used For	Port No.
sasconnect	SAS/CONNECT Server and Spawner	7550 - 7559
ossched	Operating System Services scheduler	8450 - 8459
sasmeta	SAS Metadata Server Note: Include machine name	8560 - 8569
sasobjspawno	SAS Object Spawner - Operator Port	8580 - 8589
saspool1	SAS Object Spawner: pooled workspace server port bank 1	8800 - 8809
saspool2	SAS Object Spawner: pooled workspace server port bank 2	8810 - 8819
saspool3	SAS Object Spawner: pooled workspace server port bank 3	8820 - 8829
saswork	SAS Workspace Server Note: In SAS 9.2, two or more workspace servers can share the same port even if they are running at the same time.	8590 - 8599
sassp	SAS Stored Process Server: Bridge connection	8600 - 8609
sasspmb1	SAS Stored Process Server: load balancing connection 1 (MultiBridge)	8610 - 8619
sasspmb2	SAS Stored Process Server: load balancing connection 2 (MultiBridge)	8620 - 8629
sasspmb3	SAS Stored Process Server: load balancing connection 3 (MultiBridge)	8630 - 8639
saspworksp	SAS Pooled Workspace Server	8700 - 8709
sasdepltsrv	SAS Deployment Tester - Server	10020 – 10029