SAS® INSTALLATIONS: SO YOU WANT TO INSTALL SAS®?

Rafi M. Sheikh, Analytiks International, Inc.

ABSTRACT

The paper demonstrates the importance of preparation for a SAS installation using the SAS Office Analytics® bundle as an example. There are many nuances, as well as requirements, that need to be addressed before an installation can occur. The majority of these requirements and preparations are similar yet differ slightly, according to the target operating system, whether it be Windows or *NIX (Unix and Linux) flavors. Our discussion will focus on these three topics:

1. Pre-installation 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 installation, 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 installation is no exception. Here, the fundamental building blocks of a successful SAS installation, whether on a single or multiple machines serving one end user or scores of them, all depends on how well the basic components are assembled.

We will use the SAS Office Analytics bundle as an example to talk about the necessary preparations for the installation. SAS Office Analytics® includes SAS/Foundation®, SAS/Metadata Server®, SAS/ACCESS®, ODBC, and SAS/STAT® modules. The SAS Office Analytics example is used to provide a preparation reference or framework.

SAS Office Analytics supports single or multiple machine configurations; i.e., you may install SAS Metadata, SAS Compute, and, in the case of SAS 9.4, bundled mid-tier components on separate nodes. Extending the concept further, you may decide to provide for failover nodes and, with SAS 9.4, vertical or horizontal clustering of mid-tier components. SAS has enhanced its mid-tier configuration process for ease of use. SAS 9.4 includes mid-tier components by default. Some of the new components include SAS Web Infrastructure Platform Data Server, SAS Web Infrastructure Platform Scheduling Services, SAS Environment Manager, SAS Web Server, Web Application Server, and Web Applications. SAS 9.4 mid-tier is not in the scope of this paper; however, a short discussion and link are provided for further information on Web application servers and SAS requirements in the "SAS Pre-install" section below. SAS installs are divided into the following phases:

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

PRE-INSTALL

PROFILE YOUR INSTALLATION

It is important to know the system that 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 the following:

End User:

Work Habits

Profiling end users provide a baseline and an insight into the average user's 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 it also provides opportunities for improvements. The user profile provides end-user

insight regarding what the installation is expected to support and what kind of programming and processing these end users require. The answers provide the type of performance the new SAS installation is expected to sustain

Pain Points

Awareness of any pain points that these users face such as performance, access, and programming inefficiencies provides opportunities to address issues that often do not surface and are typically unknown to the management.

Process Awareness

Another outcome is process awareness and an understanding of how things run within a certain entity. This type of information becomes input for the security model that will be needed post-installation. Profiling also provides insight into 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.

Data Sources:

Profiling data sources help develop a comprehensive strategy for accessing data.

Data Access Strategy

This addresses how data libraries will be created: either global or local. Will it be efficient to use an explicit pass-through, an implicit pass-through or a blended approach? Which data sources should be provided on a global level, as opposed to the group and individual user levels? How will the data libraries be assigned, via metadata, auto-assignment, or autoexec? These decisions will determine the need for performance considerations. In the case of third-party databases, how is access provided and managed? Will it be using a service or individual credentials?

Third Party Databases

Listing all target databases helps in the gathering of information that will help make a successful connection. The information is connection-specific such as on Windows components needed to construct System DSNs or odbc.ini in case of *NIX systems used by SAS ACCESS –ODBC for successful connection.   Verify connection to target database independent of SAS. Where non-traditional drivers are used, such as EasySoft®, additional configuration tasks will be required. *NIX OS will require additional steps such as installing and configuring a data manager. SAS Office Analytiks bundles ACCESS/ODBC module. Target databases require appropriate CLI and drivers installed and configured on the server hosting SAS since ACCESS-ODBC does not come with the drivers. Some third parties such as EasySoft and Data Direct® provide drivers.

Usage:

Profiling usage helps gauge resource and performance requirements. Typically, profiling includes these items: the number of data steps in a typical code, how many sorts and merges, what type of statistical and SAS programming features are being used, the number of concurrent SAS sessions, and the number of programs typically run by an end user. Answers provide insight to the performance requirements that the SAS environment is expected to provide. Information becomes input for performance design, the dynamic allocation of temporary space (if needed), the type of storage to use, the minimum number of CPUs, and the memory needed to sustain such performance. Often, this information becomes input for OS and hardware configuration and file system considerations.

A successful profiling exercise provides input for performance requirements, security needs, and data access strategies. 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 the following:

System Resources Design:

The SAS installation system resources such as hardware, operating system (OS), the number of CPUs, and the amount of memory (RAM) are the focal points. The information from end-user 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 has OS patch level requirements that the target server OS needs to meet or exceed. Other considerations include CPU tuning, IO, storage, and memory considerations. For example, let us consider the 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 levels of idle and sleep. The decision needs to be made collaboratively as to which C state should keep the server's processors. In the same scenario, the storage array configuration matter, for example, in the case of multiple LUNS, it is better to

not stripe at the OS level. The type of file system chosen has its own impact. The choice of storage array has a significant impact on performance.

- 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 the recommended categories: CPUs: This is based on the 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.
- RAM: The general rule of thumb is 6-8 GB per CPU.
- IO Subsystem: This will be based on the usage pattern. Work with the storage personnel to determine the best configuration and sustained IO. A sustained IO rate is important.

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.

Infrastructure Design:

The profiled information provides great insight as to the SAS environment infrastructure. Infrastructure includes directory structure, end user groups, and security model implementation on the OS as well as surfacing it in SAS Metadata.

Performance:

Decisions such as CPU/RAM, network, and IO are critical elements. IO considerations vary depending on the 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

- With the large block I/O that SAS handles, the I/O throughput rate is very important. SAS tasks determine I/O throughput required rates.

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
UTILLOC (utility files) & OLAP	50-75 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 installation. The next steps are to gather the ingredients needed for the installation. These are the following:

SAS Depot

- 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 the 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 Components for Installation

Categories	Detail

Categories	Detail
SAS System Required IDs	SAS install related IDs are an integral part of the SAS environment. Based on the SAS Office Analytics bundle , the minimum IDs are: ○ sasadm@saspw = SAS unrestricted user (admin)-(internal sas account) ○ sassrv = Used with stored procedure ○ sastrust@saspw = Used with authentication (internal sas account) ○ sasdemo = First user, after an install normally end users are not added yet and this ID helps in validating the install. ○ dbmsowner: Used with the SAS Web Infrastructure Platform Data Server (Internal sas account) ○ adminowner: Used with SAS Environment Manager: Administration Database Configuration (internal sas account) ○ sasevs@saspw: SAS Environment Manager Service Account ○ EVManager: SAS Environment Manager: Database Configuration **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	○ 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. ○ 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 a good practice to keep the installer ID groups separate from end users' and careful consideration should be given to secondary group membership. For example ID sassrv used with stored processes requires the same access and privileges that the developer has.
SAS System Required Ports	Based on the SAS solution being installed SAS maintains a particular list of ports for its use. Starting with SAS 9.4 default midtier components are shipped with SAS products. SAS Office Analytics is included. The port list therefore, has expanded. Based on SAS Office Analytics It is necessary 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 installation and that the port number is documented.
Firewall	It is recommended 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	Turn off antivirus software during installation. If this is not possible then work with your security and network people to have an exception rule created for .lck, .sd2, .sc2, .sas*, .utl and WORK location to prevent any file locking and access issues.
Installer ID Rights	It is preferred that the installation occur under a different set of credentials than from a person or an administrator, even in the case of a single user. Example ID: SAS. Note that for *NIX flavor it is strongly recommended that the installation is not done as a root or by an individual user.
Mail	Get mail related information from the administrator. The information needed is as follows: ○ Mail hosting server ○ Port ○ Whether mail requires authentication (yes or no) ○ Who the (alert) email will be sent to (admin or a mail box)
JAVA/JRE/Junit	With SAS 9.4 all the required JRE and JDK components are bundled and shipped with the product with the exception of Junit which is required for the Deployment Tester SMC plugin to function, for IQ/OQ testing post installation. For SAS versions earlier than 9.4 you may refer to the requirements and download links here: http://support.sas.com/resources/thirdpartysupport/ *JRE/JDK downloads require an Oracle profile on the Oracle website.*
Install Location: • WORK • SASHome • Config	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 and it 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.

Categories	Detail
• Data	
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 planned SAS installations 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 installation. The standard plan files are at: http://support.sas.com/demosdownloads/setupcat.jsp?cat=Standard+Deployment+Plans It is advisable to perform a four point check when dealing with orders and plan files before downloading depot as well as when applying the license file. These four points to match your plan file, SAS Install data and SAS Order Email (SOE) are the following: 1. Order number 2. SAS Tech Site Number 3. Target machine OS, and 4. Category of machine (prod, failover, etc.)
OS Patches	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 patches that SAS requires and it is important that the System Requirement guides are consulted. Read more at http://support.sas.com/resources/sysreq/
Passwords	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. There are two types of accounts required for a typical SAS install. External; based on OS & maybe domain and Internal; non OS based. : - External - based on OS & maybe domain; and - Internal - non-OS-based. External IDs are the installer ID, sasdemo, and sassrv used for installation, validation, and stored procedures, respectively. Internal IDs are SASTRUST@SASPW and SASADM@SASPW. SASTRUST is used for backend trusted communication between the different servers (i.e., MetadataServer to Workspace 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#p0kb2gtnuyjnrn n1hduu5by88u4f.htm **Note** For *NIX installations , 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

Table 2: List of Environment Components

SAS 9.4 and Midtier

Staring with SAS 9.4 SAS products are bundled with midtier components. These minimum midtier components are as follows:

- SAS Web Server®
- SAS Web Application Server ®
- SAS Web Infrastructure Platform®
- SAS BI Web Services®
- SAS Environment Manager®
- SAS Web Application Themes®
- SAS Flex Application Themes ®
- SAS Content Server®
- SAS Logon Manager®

Note: SAS 9.3-based Analytics did not include midtier components.

Since SAS9.4 Midtier is not in within this paper's scope, SAS 9.4 Midtier is discussed in brevity. SAS 9.4 Midtier has significant changes, including the bundling of the SAS Web and Application servers by default. The SAS Web Server® is an HTTP server, and the SAS Web Applications Server is a VMware vfabric lightweight server. Many enhancements have been introduced, including but not limited to a new PostGres 9.1.9 transactional data store, the SAS Web infrastructure Platform Data server, replacing the SAS Framework Server in SAS 9.3, and the addition of the SAS Environment Manager, a web-based SAS server management tool, which in the future, may replace the SAS Management Console®. All the components listed above have gone through enhancements such as a more robust custom content support, themes, and content server repository management. The vertical or horizontal cluster configuration is also simplified, and the SAS Web application can be deployed automatically as a deployment option. Read more in-depth information for SAS 9.4 Midtier here:
http://support.sas.com/documentation/cdl/en/bimtag/66823/HTML/default/viewer.htm#bimtagwhatsnew94.htm

INSTALL

Once all the ingredients are gathered and combined, they will amount to a flawless installation. Keep all the gathered information in one place. Once the execution starts, make a provision to record the installation. There are many different ways to record an installation. Some methods include manual screen shots and recording the entries or creating a screen recording of the installation. Camtasia® Studio is a useful tool for recording installations, and some use SnagIt® as a tool for capturing install steps. In addition, a response file, sdwresponse.properties, may also be created, which captures deployment input values. This file resides in the installer ID's home directory. The typical command for recording the installation via the sdwresponse.properties file is <sas-depot>\setup.exe –record –deploy. Omitting the –deploy option will only record the responses but will not actually deploy the software. For *NIX, the file resides in /home/<installer-id>.

Tip

Under the installer ID's home directory, there are many SAS Deployment Wizard-related files that are helpful in investigating installation failures or issues. These files are normally hidden. To display them on a *NIX environment, use ls –la, and on a Windows environment, set the option that will allow these hidden files to be displayed.

Typically on Windows the path is C:\Users\admin\AppData\Local\SAS\SASDeploymentWizard\...

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

Table 3: List of SAS Deployment Wizard Related Logs and Files

VALIDATION

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

Instructions.html includes all the listings for SAS Application Servers, including the Metadata server, installation, log

paths, and validation steps. Instructions.html will automatically start as soon your installation 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. The SAS Deployment Tester also plays a vital part in validating the install by running a suite of tests that look at the SAS Foundation to IQ and OQ, which is particularly relevant for clinical/pharmacy clients. Instructions.html contains other relevant information regarding backups, SAS web application URLs, port listings, and alert email set during install time. This is an important file along with DeploymentSummary.html to keep in your records. DeploymentSummary.html is also generated during install time and lists all installed components and status.

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 are between 256M and 512M. If you have a large physical memory available, then you may go higher than 512M.
C. BUFFSIZE, IBUFFSIZE, and UBUFFSIZE: The setting needs to be in conjunction with storage setup and be set different from default when doing large sequential reads.
D. SGIO: This is most relevant in WIN environments and if working with SAS 9.2M3 or higher.
E. WORKTERM: This specifies whether SAS erases WORK files at the termination of a SAS session and is enabled by default.
F. WORKINIT: This initializes a new work session library and is enabled by default.

Best Practice Use of WORK and UTILLOC

Starting with SAS 9.2, SAS has provided options to allocate WORK sessions dynamically to multiple and separate file systems that are hosted on separate logical or disk volumes. The goal is to minimize footprint collisions. The saying "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://www.aii-3.com/kb/Allocating%20SAS%20WORK%20Libraries%20Dynamically%20in%20SAS%209.2.pdf

The same concept is applicable to the UTILLOC option which specifies where the SAS utility files are written. The default option 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 either at the individual SAS application server level, such as *<SAS ConfigDir>* /Lev1/*Server/sasv9_usermods.cfg. Global application requires the use of the TKOPT environment variable (TKOPT_ENV_UTILLOC) in the *<SAS ConfigDir>*/Lev1/env.sh. If the UTILLOC option is defined at any other level for example *<SAS ConfigDir>*/Lev1/SASApp/sasv9_usermods.cfg, the option would be ignored. A limitation to note with utility files is multiple locations are most suited to SAS THREADED applications, which utilize the space; 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 I/O, which is still king in terms of 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 that 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 requirements vary per database. Configuration requirements and steps are provided by SAS at http://support.sas.com/documentation/installcenter/94/index.html (Select the relevant OS and then on the next page select System Requirement for SAS 9.x Foundation for <OS>
C. SAS 9.4 bundles all third party components except Junit which is used by the Deployment Tester component. This include SAS midtier, all Java and JRE components.

POST INSTALL

DESIGN IMPLEMENTATION

At this juncture the remaining part of the profiling that was done comes into action. At this stage, the 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, changes in data access methods, as well as a listing of what users may migrate from the old environment to the new environment (code, data, report, etc.). Providing end users with support contact during this transition keeps the exercise with minimal hurdle

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 for navigating in the new environment, moving objects, and any technical issue(s). If planned correctly, this migration can be painless.

CONSLUSION

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, and folders for SAS Code, Stored Processes, SAS Output, etc.
2. Apply necessary operating system security to these folders. SAS has metadata security, but that does not stop someone from connecting to the server and navigating to these folders remotely.
3. Make sure that the appropriate database clients are defined on the SAS Server and that the target database is accessible via client tool like SQL+ or Toad
4. For ACCESS-ODBC module define System DSNs or odbc.ini files. Ensure that the database is accessible via another third party ODBC compliant application; independent of SAS.
5. Ensure that the software depot has been successfully downloaded and 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. It is recommended that the SAS Depot is downloaded locally to the install server, otherwise download to a common storage area
6. Ensure that the appropriate users and groups are created on the OS with the appropriate permissions.
7. Make sure the OS follows SAS minimum system patch level guidelines.
8. Follow pre and post install steps for midtier components if installing on a *NIX environment which includes raising the open file descriptor and number of processes as well creating specific directories for the SAS web server.

Installing SAS is like cooking a gourmet meal, and preparation is the proverbial magic wand. The more thorough the gathering and organization of ingredients, the better the cooking experience will be. Happy Cooking!

RECOMMENDED READING

- 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
- Windows 2008 R2 Performance Degradation: http://support.microsoft.com/kb/2207548

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