

Answer Frequently Asked SAS® Usage Questions with the Help of RTRACE

Airaha Chelvakkanthan Manickam, Cognizant Technology Solutions, Bridgewater, NJ, USA

ABSTRACT:

SAS® license of any organization consists of a variety of SAS components such as SAS/STAT®, SAS/GRAPH®, SAS/OR®, etc. SAS Administrator do not have any automated tool supplied with Base SAS software to find how many of licensed copies are being actively used, how many SAS users are actively utilizing the SAS server and how many SAS datasets are being referenced. These questions help a SAS administrator to take important decisions such as controlling SAS licenses, removing inactive SAS users, purging long-time non-referenced SAS datasets, etc. SAS provides a system parameter called RTRACE to answer these questions. The goal of this paper is to explain the set-up of RTRACE parameter and to explain the usage of it in making the SAS administrator's life easy. This paper is based on SAS 9.2 running on AIX operating system.

INTRODUCTION:

SAS contains several components such as SAS/STAT, SAS/GRAPH, SAS/OR, etc. and every organization gets licenses to use a variety of SAS components for their SAS environment. These SAS components are usually installed on a server and utilized by hundreds of SAS users. There is no direct way or tool to identify the usage of the installed SAS components against the list of users.

The SAS users produce thousands of SAS datasets every month. Several of these SAS datasets are continuously refreshed, used and updated. However in every SAS server, a portion of SAS datasets never gets referenced or updated. So they are lying on the server occupying huge amount of disk space.

This paper is trying to help the SAS administrator to answer the following important questions:

1. How to identify the most active SAS users and the least active SAS users?
2. What are the most used SAS components and what are least used?
3. How to get the list of inactive SAS datasets for clean-up and space recovery?
4. How much disk space is consumed by inactive datasets?

RTRACE is a system option provided by SAS which, if used at the right place, can provide the solutions for the above mentioned questions for SAS Administrator and can record activities of SAS users accessing the server.

RTRACE DEFINITION:

Definition of RTRACE option as per support.sas.com is as follows:

```
To trace the loading of SAS modules and files, you can use the SAS system option
-RTRACE. The syntax is as follows:
    -rtrace ALL -rtraceloc <file_name>
The -RTRACELOC option is *required* when setting RTRACE ALL.
```

RTRACE captures a lot of details about the activities carried out in a particular SAS session. Sometimes it produces very large files with long time running sessions. It is the responsibility of the SAS Admin to know what is useful and filter out what is required to produce usage metrics.

HOW TO SET-UP RTRACE SYSTEM OPTION AND START GENERATING RTRACE LOGS?

SAMPLE SAS SYSTEM:

The set-up of the SAS environment decides the set-up of RTRACE option. For the purpose of this paper, let's consider a SAS environment as mentioned below.

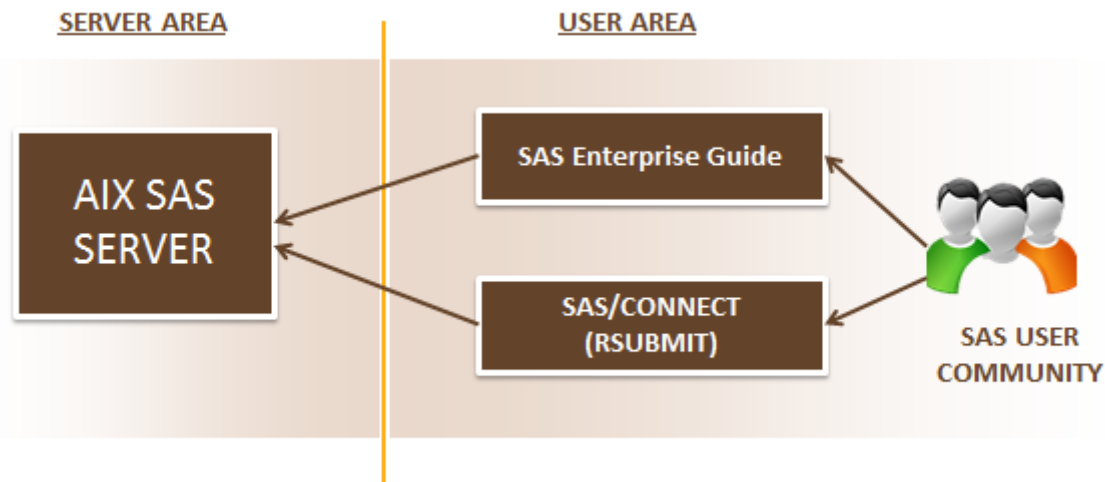


Figure 1 Sample SAS System

It is important to note that deploying administration mechanism such as RTRACE should possibly be invisible to the users. This way, users are not disturbed about the tracking mechanism running behind the scene and RTRACE can produce true utilization metrics.

As shown in the picture above (figure: 1), SAS users connect to SAS server in two ways 1) by invoking SAS/CONNECT and Remote-Submit their jobs on the server 2) connecting to the server via SAS Enterprise Guide. RTRACE must be deployed such that it can track the traffic on both the ways.

The better place to invoke RTRACE system option is to include it on the server in places which get invoked whenever a SAS users connects to the server, either using SAS/CONNECT or SAS Enterprise Guide.

SAMPLE SAS SYSTEM FOLDERS CONFIGURATION:

For the purpose of this paper, following are the important SAS system folders, in a typical AIX environment, that a SAS Admin should be aware of in order to place RTRACE system option.

- (1) SAS software is installed under location : /ROOT/SAS_9.2
- (2) SAS Configuration files are located at : /ROOT/SAS_9.2/SASFoundation/9.2
- (3) SAS executable link is located at : /ROOT/SAS_9.2/SASFoundation/9.2/sas
- (4) SAS executable link calls for wrapper script : /ROOT/SAS_9.2/SASFoundation/9.2/bin/sas_en

(3) is a link, not an actual file or folder, that calls for the SAS wrapper script (4) – as shown below in figure (2).

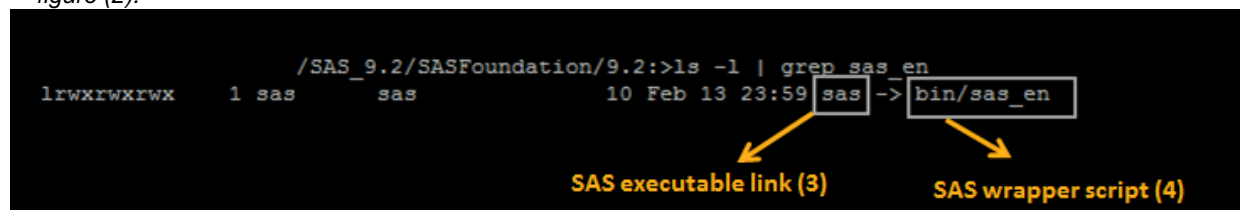


Figure 2 SAS wrapper script link

- (5) SAS Meta-data server configuration is located at : /ROOT/SAS9.2/CONFIG/Lev1
- (6) SAS variables for meta-data connection located at : /ROOT/SAS9.2/CONFIG/Lev1/level_env.sh

HOW TO SET-UP RTRACE FOR CAPTURING STATISTICS OF USERS CONNECTING USING SAS/CONNECT?

When user invokes SAS using SAS/CONNECT, SAS executable link mentioned at (3) is invoked, which in turn calls SAS wrapper script mentioned at (4). It is where the RTRACE system option should be included to capture the statistics of every user connecting using SAS/CONNECT.

The idea that the author uses is to intercept between SAS executable link (3) and SAS wrapper script (4) such as make SAS executable link (3) connect to a new script with RTRACE system option which will then connect to SAS wrapper script (4) – as shown in figure (3).

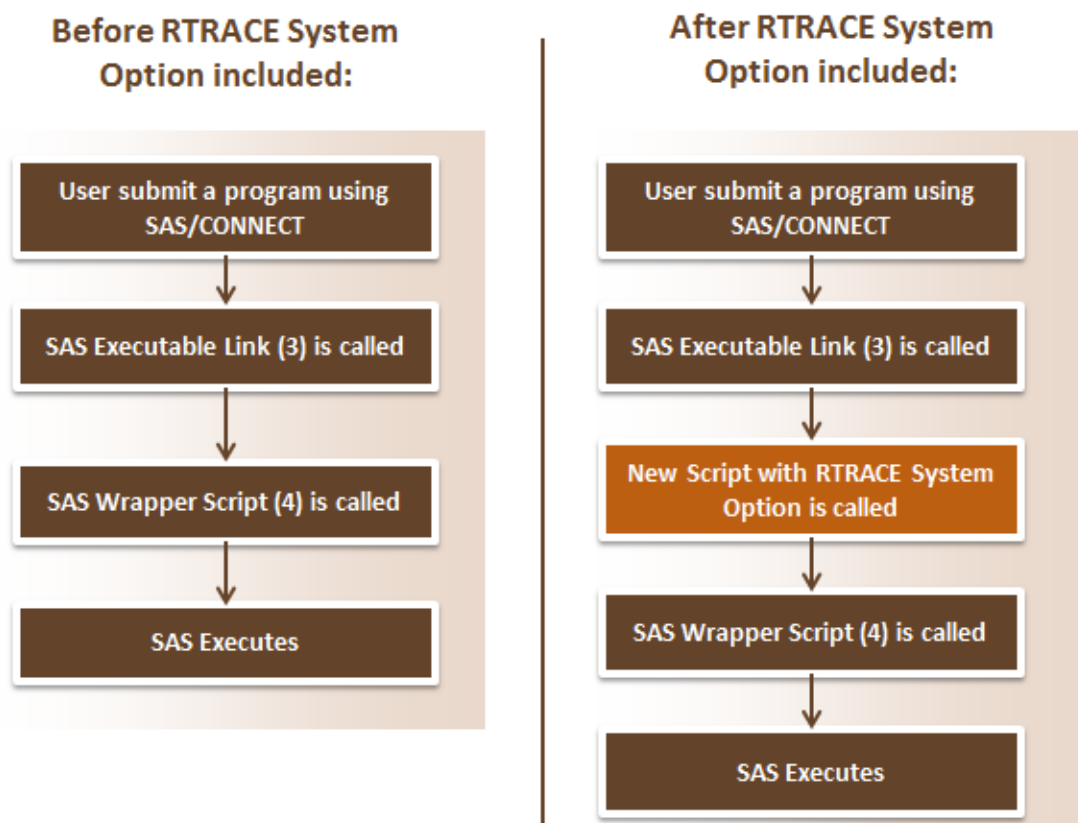


Figure 3 RTRACE set-up for SAS/CONNECT

The content of the new script with RTRACE system option is given below – in figure (4).

```

/appl/sas/SAS_9.2/SASFoundation/9.2:>ls -l | grep sas_en
lrwxrwxrwx  1 sas  sas      10 Feb 13 23:59 sas -> bin/sas_en_alias
/appl/sas/SAS_9.2/SASFoundation/9.2:>cat bin/sas_en_alias
OUTPUTDir=/Info-One/adhoc/manickal/RTRACE
PID=$$
Host=$(hostname)
Pgm=$@
/appl/sas/SAS_9.2/SASFoundation/9.2/bin/sas_en $Pgm -RTRACE ALL -RTRACELOC $OUTPUTDir/$USER.$Host.$PID.RTRACE.LOG

```

Annotations in the image:

- An arrow points from the `bin/sas_en_alias` file to the text: **New wrapper script that invokes RTRACE option**.
- An arrow points from the `bin/sas_en` file to the text: **SAS executable link (3)**.
- An arrow points from the `bin/sas_en` file to the text: **Existing SAS wrapper script (4) called from inside new script**.

Figure 4 Script with RTRACE system call – for SAS/CONNECT

The new wrapper script has 3 variables and these variables help to name the RTRACE log file unique to every user/session.

1. "USER" captures the user id of the person connecting to the user
2. "PID" captures the process ID of the user / this helps to name the RTRACE log file unique across different sessions of the same user
3. "Host" captures the server name in which user was logged in

Along with the above variables, "OUTPUTDir" variable specifies the location where RTRACE logs should be sent to.

At the end of the script, the control is sent back to the original SAS wrapper script – but along with the invocation of RTRACE option and a location to store the log file.

HOW TO SET-UP RTRACE FOR CAPTURING STATISTICS OF USERS CONNECTING USING SAS ENTERPRISE GUIDE?

When user invokes SAS using SAS Enterprise Guide, SAS invokes the environment variables mentioned at (6). It is an optimal place to include RTRACE system option so it can be automatically invoked for every user connecting using SAS Enterprise Guide.

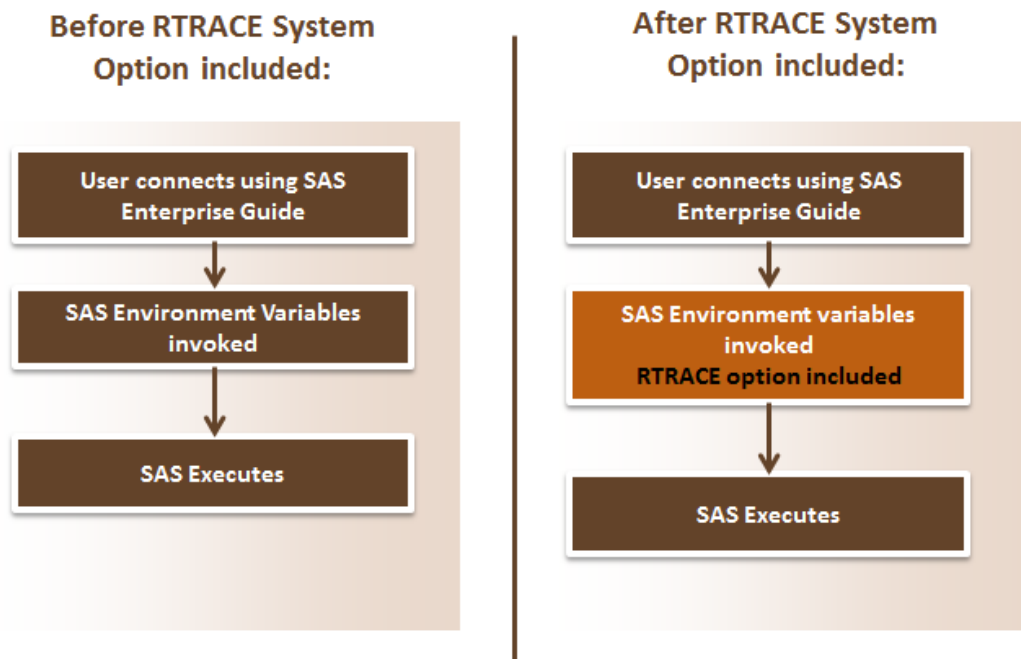


Figure 5 RTRACE set-up for SAS Enterprise Guide

The content of the modified environment variables file (6) with RTRACE system option is given below – in figure (6).

```

#!/bin/sh
#
# level_env.sh
#
# Set environment variables relative to the configuration level structure
#

LEVEL_ROOT=/appl/sas/SAS_9.2/CONFIG/Lev1
UTILITIES=/appl/sas/SAS_9.2/SASDeploymentManager/9.2/products/cfgwizard_92380_prt_xx_sp0_1/Utilities
DEPLOYWIZ=/appl/sas/SAS_9.2/SASDeploymentManager/9.2/products/deploywiz_92380_prt_xx_sp0_1/deploywiz
SAS_HOME=/appl/sas/SAS_9.2
JAVA_JRE_COMMAND=/appl/sas/SAS_9.2/jre1.5_SR5/jre/bin/java

SASVJR_HOME=/appl/sas/SAS_9.2/SASVersionedJarRepository/9.2
SASVJR_REPOSITORYPATH=/appl/sas/SAS_9.2/SASVersionedJarRepository/9.2/eclipse
SASROOT=/appl/sas/SAS_9.2/SASFoundation/9.2

OUTPUTDir=/Info-One/adhoc/manicka1/RTRACE
PID=$$
Host=$(hostname)

SAS_COMMAND="$SASROOT/sas -RTRACE ALL -RTRACELOC $OUTPUTDir/$USER.$Host.$PID.RTRACE.LOG"

```

Figure 6 Script with RTRACE system call – for SAS Enterprise Guide

Since the environment variable file is invoked by every SAS Enterprise Guide user, RTRACE system option is included in every session of the SAS Enterprise Guide user.

HOW TO UNDERSTAND RTRACE LOG FILES? WHAT IT CONTAINS?

As explained in the previous sections, RTRACE log files are created with unique name consisting of user id, session id and server name. They are also kept at specific folders passed as parameter to the RTRACE option. Figure (7) shows sample RTRACE log files.

	User ID	Server Name	Session ID	Last modified time of the log file / provides when user sessions ends
-rw-r----- 1 ramkish1 staff 5804	Ramkish1	sasservr1.na.company1.net	3762810	Feb 15 15:35
-rw-r----- 1 kanthma1 systmgr1 36025	kanthma1	sasservr1.na.company1.net	3762234	Feb 16 12:23
-rw-r----- 1 ramkish1 staff 18059	ramkish1	sasservr1.na.company1.net	3819829	Feb 17 04:12
-rw-r----- 1 Suchisa1 prodct34 466243	Suchisa1	sasservr1.na.company1.net	9398102	Feb 17 06:06
-rw-r----- 1 Jananbh1 unixgrp1 48229	Jananbh1	sasservr1.na.company1.net	6386198	Feb 17 06:07
-rw-r----- 1 Siddhak1 allgrps 104452	Siddhak1	sasservr1.na.company1.net	1092932	Feb 17 10:09
-rw-r----- 1 jananbh1 unixgrp1 963605	jananbh1	sasservr1.na.company1.net	8376512	Feb 18 23:22

Figure 7 Sample RTRACE log files

RTRACE log contains every action performed by the SAS session in terms of files opened/referenced/renamed/closed.

- Every SAS component such as SAS/GRAPH, SAS/ETS, etc. called by the user is recorded by the opening of respective SAS component's executable files
- Every SAS dataset read by the user is recorded by the opening/referencing/closing of respective SAS dataset
- Every SAS dataset modified/newly created is recorded by the opening/referencing/renaming/closing of respective SAS dataset
- Every temporary SAS dataset created by the user is also recorded

Figure (8) shows sample content of RTRACE log files and brief explanation of what each line refers to.



Figure 8 Sample content of RTRACE log file

It is a challenge to clearly decode/link the SAS executable file to its respective SAS component such as SAS/GRAP, SAS/ETS, SAS/IML, etc. The following sections provide an approach that the author has come across and used in his experience.

OKAY. LET'S PUT RTRACE TO BETTER USE:

Till now we have explored how to set-up of RTRACE option and how to understand the log files that it produces. Now, it is time to put the logs captured to better use.

Following sections answer many SAS usage questions that normally arise in an active SAS environment. The intent of this paper is to provide a logic/algorithm to the audience to find answers to the usage questions. The author can be contacted privately to collect the actual programs which implement the algorithms in AIX 6.1 environment.

HOW RTRACE HELPS TO OBSERVE SAS USER CONNECTION STATISTICS?

In a typical SAS environment, SAS users connect to the server 24x7. It is costly and resource consuming to have an active monitoring mechanism for user connections. RTRACE comes in handy to answer frequent questions in this section, such as:

- How many users connect in a month?
- How many users did not connect to the SAS server in a month?
- How many sessions created by a user in a day?
- What are the users connecting trends (density) over a period of time?

The answer lies with the name of RTRACE log files. As shown in figure (7), RTRACE log file names contain what is needed to know about user statistics.

- The RTRACE log file names contain user ids which can be mapped to an internal database to identify the actual user names.
- The session ids on the RTRACE file names help to identify how many times a user connects to the SAS server in a day, because each separate connection creates a separate session id.
- The timestamp of when RTRACE log files were last modified provide details of when the user last connected to the SAS server.
- By looking at the list of log files created for a month, the number of users connected to the SAS server for that month can be obtained. By matching it against a pre-defined list of SAS licensed users, the list of users who did not make any connection can also be derived.
- Along with this, the user ids on the log file names and their modified time can help to identify which days of the month most number of users connected and which days of the month least number of users connected to the SAS server – called User Density report.

The statistics obtained from these answers help to plot a lot of interesting graphs that are useful for SAS Admins and to the management to make important decisions about the server usage. Some sample graphs are given below. The graphs can be created using Microsoft® Excel or SAS Enterprise Guide.

Figure (9) depicts the number of users connected to the SAS server Vs number of users did not connect to the server. As per the chart, 45 users who are allotted with a SAS license, but never made a connection in Jan-2014. Such continuous monitoring will help SAS Admins to decide if SAS license can be revoked from non-connecting users.

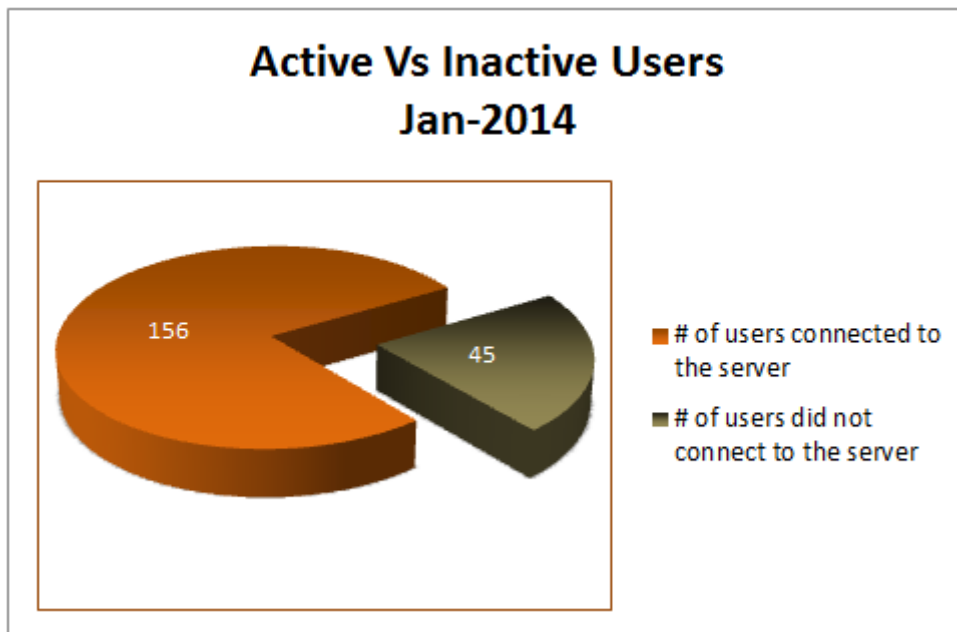


Figure 9 Chart to display active Vs inactive users

Figure (10) depicts the most active users and least active users. Users submitted programs on the server more than 15 days of the month are most active users and uses connected less than 3 days are lease active users.

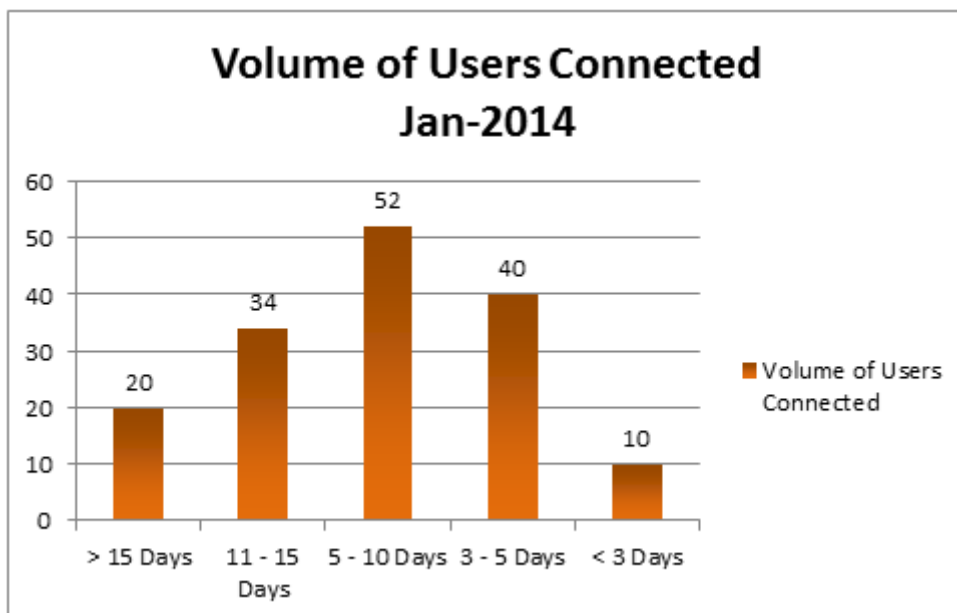


Figure 10 Chart to display most active and least active users in a month

Figure (11) shows a deeper look at the number of users connected in a particular day by hour. As per the statistics, 2PM to 4PM is the most active time of the day. It is also observed that there are 15 users who connected to the SAS server after office hours (between 6PM to Midnight).

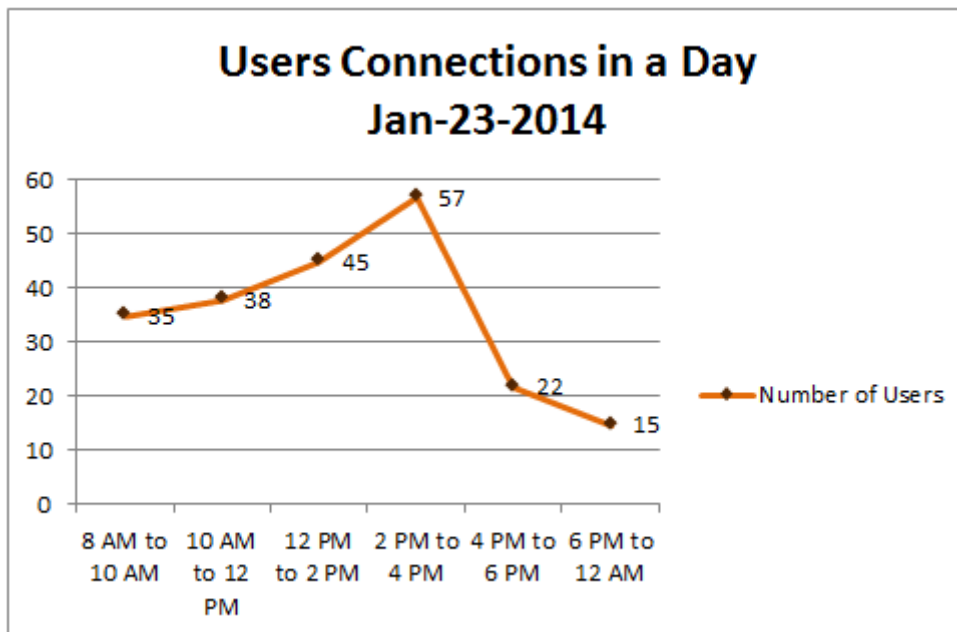


Figure 11 Chart to display user connection trend in a day

HOW RTRACE HELPS TO OBSERVE SAS COMPONENT USAGE STATISTICS?

The challenge question any SAS admin need to face is how many users use a particular SAS component? There is no direct way of answering this question. RTRACE logs contain information that can help answering SAS component usage questions.

Frequently asked questions about SAS component usage are:

- What are the most used and least used SAS components?
- How many users used a particular SAS component?
- Usage of SAS components over a period of time? Or which days of the month certain SAS components are widely used?

The content of RTRACE log files has answers to these questions. As shown in figure (8), RTRACE log contains what are the SAS executable files opened and closed. These executable files refer to the respective SAS components such as SAS/STAT, SAS/OR, SAS/IML, SAS/ETS, SAS/GRAPH, etc.

The biggest challenge here is to map the SAS executable names to the respective SAS component name. For instance, following snippet from a RTRACE log shows that "saswzx", "saswobs" and "sasxug" executable files used within a SAS session. What is the name of the SAS component that consists of these executable files?

```
File opened: /appl/sas/SAS_9.2/SASFoundation/9.2/sasexe/saswzx
File opened: /appl/sas/SAS_9.2/SASFoundation/9.2/sasexe/saswobs
File opened: /appl/sas/SAS_9.2/SASFoundation/9.2/sasexe/sasxug
```

On an AIX system, there is no direct mapping of SAS executable files to SAS components on the SAS installation directories. When SAS installed in AIX environment, all executable files are kept in the directory /appl/sas/SAS_9.2/SASFoundation/9.2/sasexe and has no distinction by SAS component.

This is different in Microsoft® Windows 7 SAS installation. On a Windows 7 environment, SAS executable files are properly organized into respective SAS component named directories. For instance, the directory called C:\Program Files\SAS\SASFoundation\9.2\graph\sasexe contain all executable files (as .dll files on Windows 7) for SAS/GRAPH. Similarly, a separate folder contains executable files for each SAS component such as SAS/STAT, SAS/OR, SAS/ETS, SAS/IML, etc.

It is a matter of arranging SAS executable file names by SAS component using Windows 7 installed folders and use this organized structure for analysis in AIX environment.

Following chart Figure (12) shows sample executable file names for SAS/GRAPH and SAS/ETS components taken from Windows 7 environment. By omitting the .dll extension, this mapping can be used directly on an AIX environment.

COMPONENT	Corresponding .DLL	COMPONENT	Corresponding .DLL
GRAPH	gbarline.dll	ETS	sasarima.dll
GRAPH	geocode.dll	ETS	sasautor.dll
GRAPH	sasamgis.dll	ETS	sascitib.dll
GRAPH	sasammap.dll	ETS	sascompu.dll
GRAPH	sascht7.dll	ETS	sascount.dll
GRAPH	sascht8.dll	ETS	sasdataf.dll
GRAPH	sasg3d.dll	ETS	sasdiag.dll
GRAPH	sasg3gri.dll	ETS	sasdtsrc.dll
GRAPH	sasgaedt.dll	ETS	sasecrsp.dll
GRAPH	sasgafix.dll	ETS	sasefame.dll
GRAPH	sasgalog.dll	ETS	sasehavr.dll
GRAPH	sasganno.dll	ETS	sasentro.dll
GRAPH	sasgarea.dll	ETS	sasesm.dll

Figure 12 Table with sample SAS executable file names from Windows 7

When we know how to map a SAS executable file name to a SAS component, it is an easy job to mine into the RTRACE log to understand the SAS component usage statistics.

- By analyzing a RTRACE log file and mapping it against the SAS executable table, the list of SAS components used within a user's session can be found. By extending this analysis on all RTRACE log files produced over a period of time, the usage of all SAS components of that period can be obtained.
- The user ids on the RTRACE log files help to map the usage of SAS components to a particular user and to the user volume.

Following are some sample charts that depict the SAS component usage produced with the help of the statistics obtained in this section.

Figure (13) shows the most used and least used SAS components of Jan-2014. SAS/ACCESS has the most number of users – 66 and SAS/IML & SAS/OR has the least number of users – 3 and 2 respectively.

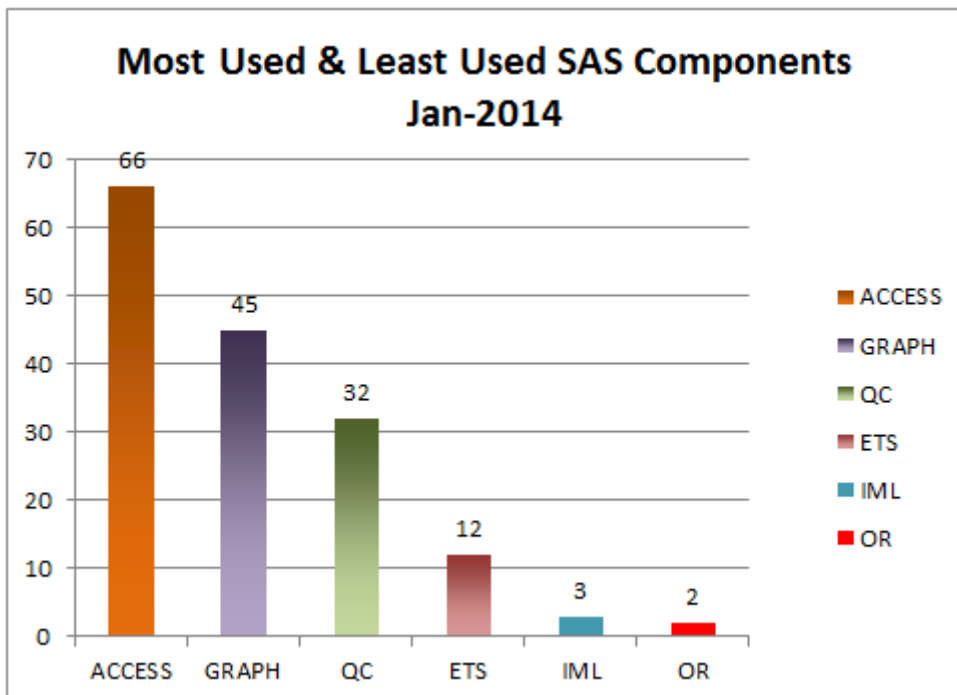


Figure 13 Chart to display most used and least used SAS components in a month

Figure (14) shows the usage statistics of SAS/ACCESS, SAS/IML and SAS/GRAPH components for the entire month of Feb-2014. The statistics has some key observations.

- Demand for SAS/ACCESS is more in the 1st week of the month compared to rest of the month
- Need for SAS/GRAPH is lesser in the middle of the month compared to rest of the month
- SAS/IML, lowest utilized SAS Component, is not used by anyone in the last week of the month

Observations of this nature will help SAS Admin to prepare the environment specific to SAS components, cater the need of respective business units and shuffle SAS licenses between users, if needed.

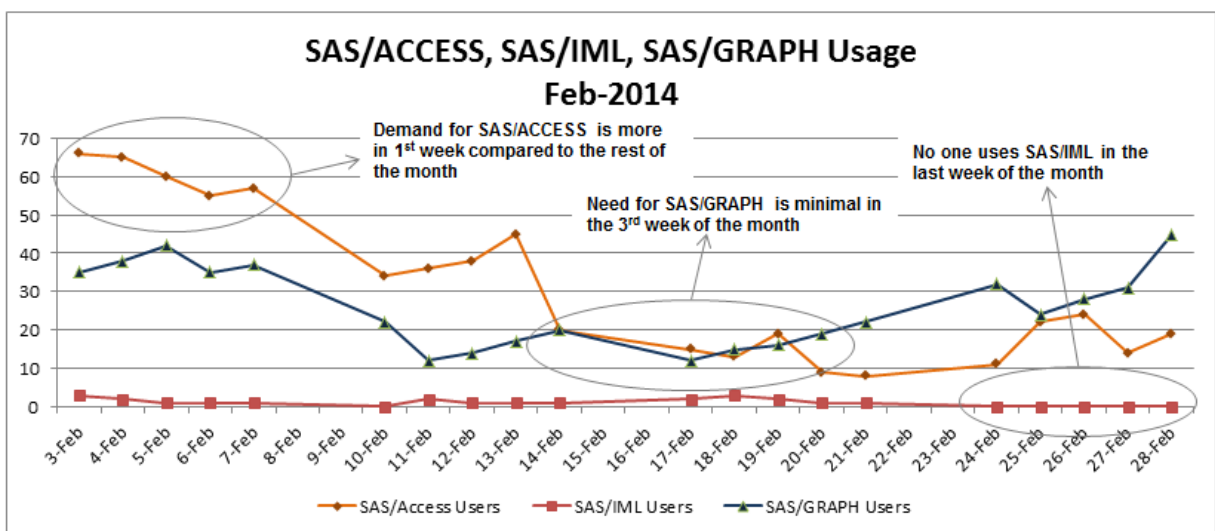


Figure 14 Chart to display usage statistics of various SAS/components in a month

There can be lot of SAS Components purchased and installed in the SAS server for end users. Few of them never get referenced. By monitoring the usage of SAS components by end users using RTRACE logs, organizations can decide whether they really need to pay for these licenses for unused components existing on the server during next SAS license renewal.

HOW RTRACE HELPS TO IDENTIFY INACTIVE SAS DATASETS?

In a SAS environment where SAS users are allowed to create their own SAS datasets for performing statistical analysis, it is a challenging work to automatically identify and purge inactive permanent SAS datasets. It may look simple, but in an environment with hundreds or thousands of SAS users are activity working, it is not an easy job.

Frequently asked questions about inactive SAS datasets are:

- How much of overall space used by inactive SAS datasets?
- Provide a list of all inactive SAS datasets that were never used in last 6 months and of size greater than 100 MB

RTRACE logs have a solution for this. Inside the RTRACE logs, there is information about the datasets that user had accessed during the session – as shown in the snippet below.

```
File opened: /Project1/Library1/assessment.sas7bdat
File closed: /Folder1/Testing/user1/align/y2008z2/zip_2008_trd.sas7bdat
File renamed: /Folder2/Testing/SAS_work3F8D01DB005E_server/afftest.sas7bdat
```

- Using the information captured in the RTACE logs, collect the names of all the datasets being opened, closed, referenced, and renamed by users.
- Compare this information with the list of all datasets present on the server.
- Check for the datasets that are older than a threshold period and that were never used by any user for more than 6 months.
- Publish the identified inactive SAS datasets and plan to purge them.

There are several charts that can be generated using the statistics collected.

Figure (15) below provides a summary of inactive SAS dataset analysis performed for all files of size > 100MB and using last 6 months of RTRACE logs. As shown in the chart, 323 GB + 450 GB of space can be recovered by purging the inactive 123 + 324 SAS datasets that are older than 2 and 5 years respectively.

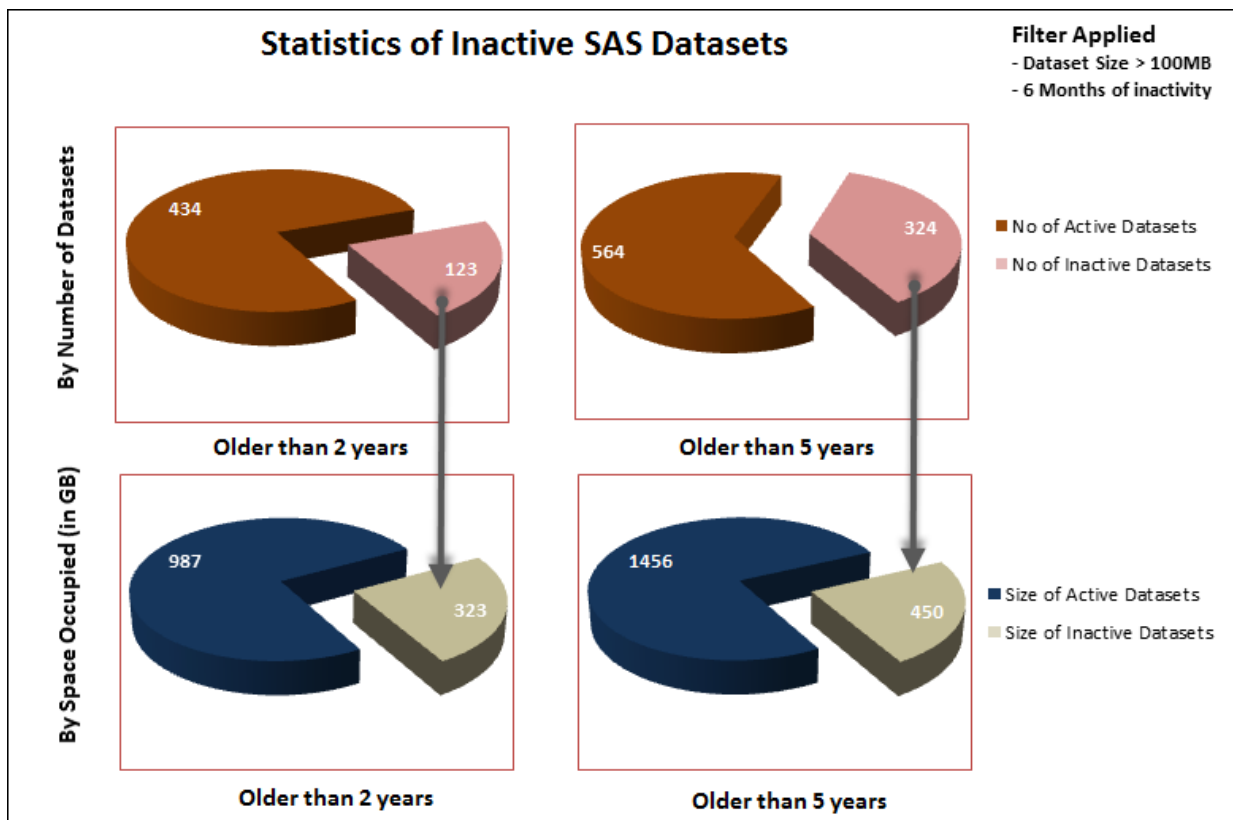


Figure 15 Chart to display inactive dataset statistics

CONCLUSION

In a typical environment, SAS business users tend to create datasets and programs on-the-fly and in ad-hoc manner. RTRACE is inevitable in such environments to better understand the SAS system usage and make proper recommendations to the management.

The drawback of RTRACE facility is its own space consumption to store the log files and the need for extensive programming skills to decode the information captured. However the level of in-depth details that RTRACE log files provide, will overcome this drawback in a large organization. Overall, RTRACE is a tool that any SAS admin must know and be familiar with.

REFERENCES

Michael A. Raithel, "Measuring SAS® Software Usage On Shared Servers With the RTRACE Facility", SUGI 29. Available at: www2.sas.com/proceedings/sugi29/215-29.pdf

David Bosak, "Reporting SAS®9 License Usage", SUGI30. Available at: www2.sas.com/proceedings/sugi30/216-30.pdf

SAS Support Documents from <https://support.sas.com>

ACKNOWLEDGMENTS

I like to acknowledge Shrikanth, Thota who initiated this paper in SAS Global Forum 2013 but couldn't complete. A part of his effort has been reused in this paper. I like to thank my Cognizant managers Sunil Varanasy, Ajit Kumar Sahu and Dhananjay Kelkar for encouraging me to submit this paper. Last but not the least, I like to thank my wife, Janani Kanthan, for giving me worry-free nights from my 2 year old to compile the paper from pieces of information that I had kept at various places.

CONTACT INFORMATION

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

Name: Airaha Chelvakkanthan Manickam
Organization: Cognizant Technology Solutions
City, State ZIP: Bridgewater, NJ
Work Phone: 610 316 5780
Email: airahachelvakkanthan.manickam@cognizant.com / mairaha@gmail.com
Web: <http://www.linkedin.com/pub/airaha-chelvakkantha-manickam/23/b65/761>

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 brand and product names are trademarks of their respective companies.