

A Quick Guide for your Data Load Problems in SAS Visual Analytics

Swetha Vuppalanchi

ABSTRACT

SAS® Visual Analytics is a user-friendly yet powerful and intuitive visualization tool that one can use for data exploration and reporting. However, before one can work with any data, that data has to be made available to the SAS Visual Analytics environment. This paper introduces to the high-level architecture of SAS Visual Analytics, the connection setup between the SAS Visual Analytics servers, the kinds of data sources that can be accessed, and the different ways of extracting data into SAS Visual Analytics. Most importantly, this paper addresses the most common errors encountered when loading data into SAS Visual Analytics. This paper also discusses useful options for automating the start of SAS® LASR™ Analytic servers, loading and refreshing SAS Visual Analytics tables, and directing signature files. Finally, the paper includes a checkpoint list that can be used to ensure that the appropriate permissions are in place at the metadata and operating system levels to perform data load operations effortlessly. The techniques and troubleshooting methods discussed in this paper are based on the distributed SAS Visual Analytics installation on a UNIX framework, but methodologies discussed here can be useful for troubleshooting common SAS Visual Analytics issues on any environment.

INTRODUCTION

SAS Visual Analytics offers two types of deployments: distributed and non-distributed. In distributed deployment there is one root node and one or more worker nodes connected through network whereas in non-distributed deployment there is only one server which acts as both the root and worker node. A typical deployment for distributed SAS® LASR™ Analytic Server is to use a series of blades in a cluster. The same operating system is used throughout and the same SAS software is installed on each blade that is used for the server. To enable the communication among the blades, MPI implementation is used. The workload is distributed among different machines to perform massively parallel processing. The main component of SAS Visual Analytics is SAS® LASR™ Analytics server which is an in-memory engine that provides secure and concurrent access to data loaded in memory. The server handles both big data and smaller sets of data, and it is designed for high-performance, multi-threaded, analytic code. The server processes client requests at extraordinarily high speeds due to the combination of hardware and software that is designed for rapid access to tables in memory. For the data to be available in SAS Visual Analytics environment, it has to be loaded into SAS® LASR™ Analytics Server. The term loaded into memory is also used interchangeably.

LOADING DATA INTO LASR ANALYTIC SERVER

There are two types of LASR servers available with the default SAS VA installation where the data can be stored in memory. PUBLIC LASR and PRIVATE LASR. Apart from the default LASR servers, custom LASR servers can also be defined for each department, project etc. The PUBLIC LASR can be used by any user who is not defined in SAS metadata but would like to store data in LASR whereas to load data into PRIVATE LASR, a user has to be defined in metadata. A registered metadata user can also load tables into PUBLIC LASR along with PRIVATE LASR. There are different ways of loading data into PUBLIC and PRIVATE LASR servers. Data can be loaded from excel or csv files, database tables or SAS datasets. Data can also be combined from different data sources using VA Data Builder and the target dataset can also be loaded into LASR server for quick access. Data can also be loaded into LASR using other SAS clients such as SAS Enterprise Guide and SAS Data Integration Studio and the inbuilt interface SAS VA Administrator. For the data to be loaded into LASR from the above-mentioned interfaces, the data has to be first registered in metadata. Only when loading data through autoloading process, data need not be registered in metadata. The data loaded in LASR server (aka in memory) is available for other VA interfaces, Visual Analytics Explorer, Visual Analytics Designer. The reports developed from these interfaces can be consumed by Visual Analytics Viewer, SAS Mobile Viewer and also SAS Add-In for Microsoft Office as long as they connect to

the same metadata server. The following figures show different interfaces that can be used to load tables into memory

Figure 1. Data Load in VA

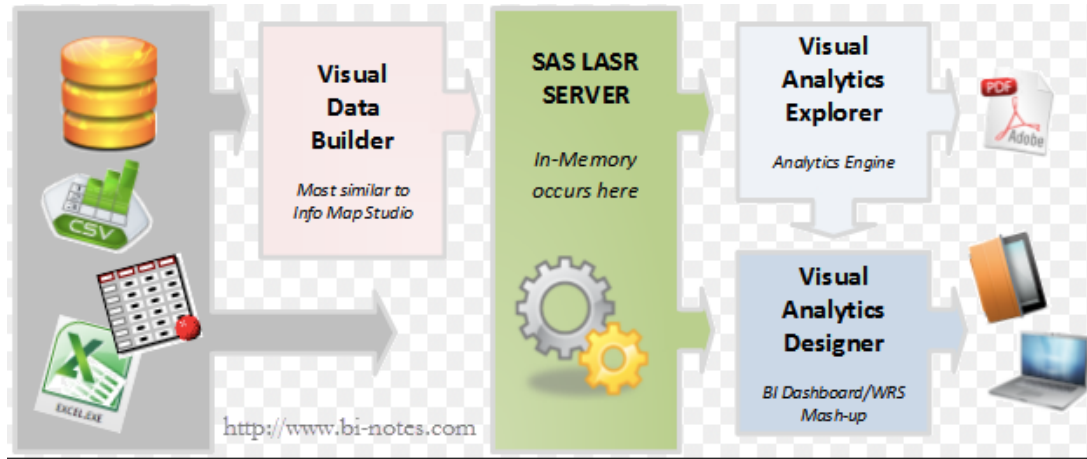


Figure 2. Loading from SAS Enterprise Guide

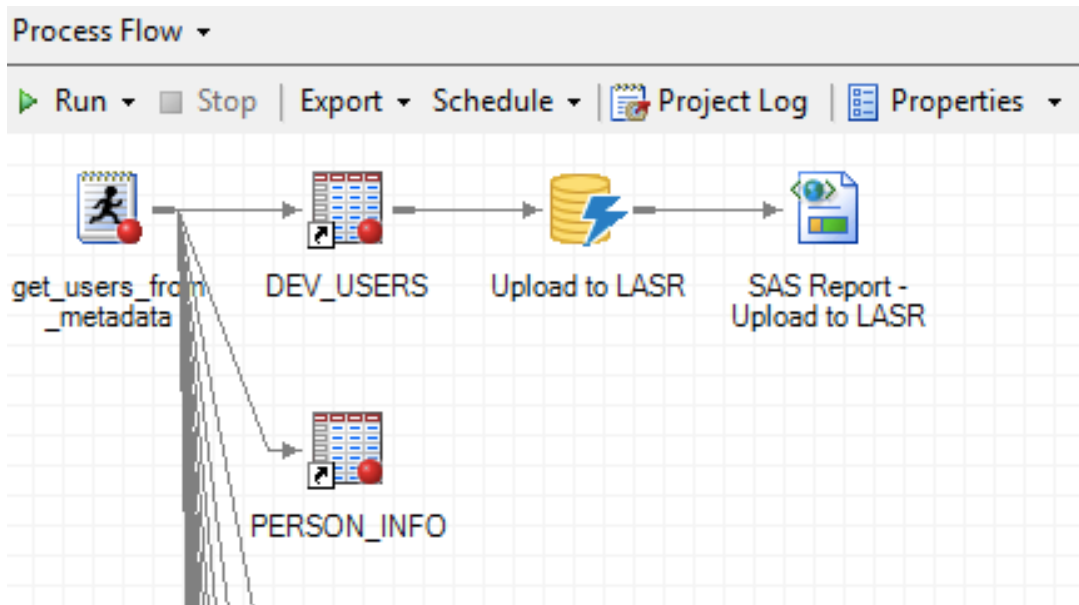
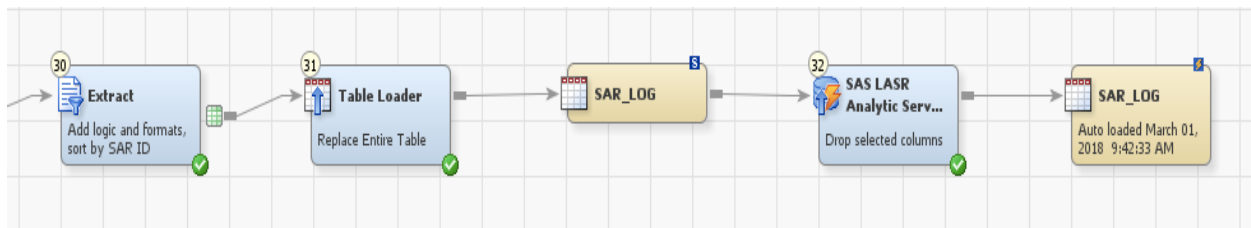


Figure 3. Loading from SAS Data Integration Studio



AUTOMATIC LOADING OF TABLES INTO LASR SERVER

The default set up of SAS Visual Analytics allows for the automatic load of tables into PUBLIC LASR Server. The autoloading directory for PUBLIC LASR is configured at the time of installation and by default points to **<SASCONFIG>/Lev1/AppData/SASVisualAnalytics/VisualAnalyticsAdministrator/AutoLoad**.

This path can be configured to point to different directory at the time of installation or by changing the settings in SAS Management Console described in configuration section. In order to load the tables into PUBLIC LASR server, just drop the excel files, csv files or SAS tables into this folder and the tables should be available in memory after 15 mins according to the default schedule set up for autoloading. This interval again can be changed according to the requirements. The scripts associated for the autoloading are located under **<SASCONFIG>/Lev1/Applications/SASVisualAnalytics/VisualAnalyticsAdministrator**.

In the standard configuration, a new run of autoloading is started every 15 minutes. The timing is controlled by a setting in schedule.sh (on UNIX) or schedule.bat (on Windows) in the autoloading scripts directory. The autoloading script can be scheduled through CRON in UNIX and Task scheduler in windows by giving appropriate permissions to the scheduler account. The shell script for autoloading can also be customized to be invoked by other schedulers utilizing the scripts set up for default autoloading

Below is an example of schedule.sh in UNIX. The time interval can be changed by changing the value for TIME_INTERVAL_MINUTES

```
#!/bin/bash
#
# schedule.sh
#
# syslog      Schedule AutoLoad.sas
#
RUNSAS_PATH="/opt/aml/SASCONFIG/Lev1/Applications/SASVisualAnalytics/VisualAnalyticsAdministrator/runsas.sh"
TIME_INTERVAL_MINUTES=15

#call to runsas.sh
cat <(fgrep -i -v $RUNSAS_PATH <(crontab -l)) <(echo "**$TIME_INTERVAL_MINUTES * * * $RUNSAS_PATH > /dev/null ") | crontab -
```

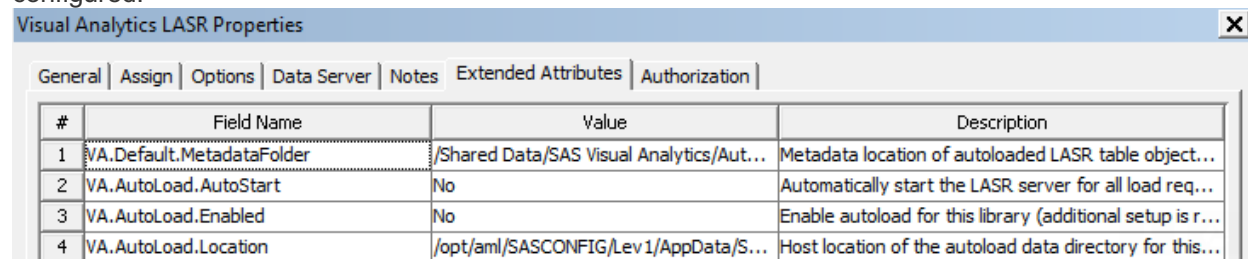
Once the autoloading schedule is setup, any tables that exist under *Autoloading* folder will be loaded into memory and any tables that exist under *Unload* sub directory in Autoloading will be unloaded from memory according to the schedule set in the schedule.sh script if using the standard autoloading. To troubleshoot any errors associated with autoloading process, check the logs under Logs sub directory under Autoloading folder

As mentioned additional LASR servers can be configured for various departments/projects and similar set up of autoloading can be done for the custom LASR servers. For autoloading database tables into PRIVATE LASR Analytic Server, some additional processes need to be setup and scheduled. Both of these processes are described in detail in the links pointed out in the recommended reading section.

CONFIGURATION OF LASR LIBRARY FOR AUTOLOAD

To Configure LASR Servers for autoloading, following options need to be enabled from SAS Management Console for the pertaining LASR Server.

Data Library Manager Wizard → Libraries and select the LASR server for which the options need to be configured.



#	Field Name	Value	Description
1	VA.Default.MetadataFolder	/Shared Data/SAS Visual Analytics/Aut...	Metadata location of autoloading LASR table object...
2	VA.AutoLoad.AutoStart	No	Automatically start the LASR server for all load req...
3	VA.AutoLoad.Enabled	No	Enable autoloading for this library (additional setup is r...
4	VA.AutoLoad.Location	/opt/aml/SASCONFIG/Lev1/AppData/S...	Host location of the autoloading data directory for this...

VA.AutoLoad.AutoStart – This option would start the LASR server automatically when tables are being loaded to this LASR server and when the server is switched off

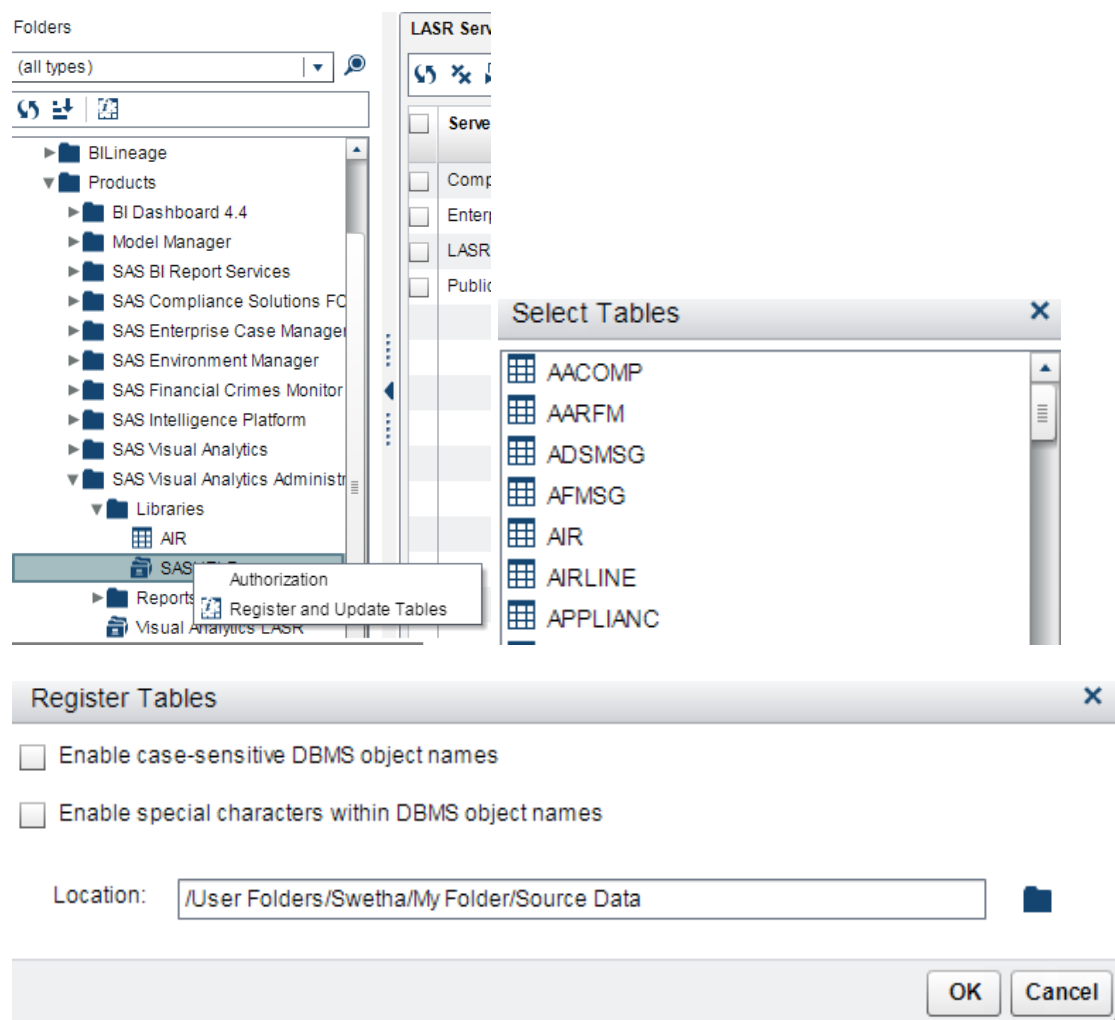
VA.AutoLoad.Enabled – This option would enable the autoloading process to this LASR server. Apart from this option, autoloading scripts should also be configured if it is a custom LASR server

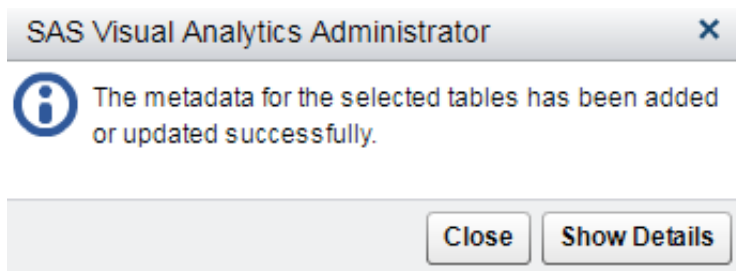
VA.AutoLoad.Location – This option is to set the directory for autoloading. This can be changed to any other directory on the server where the user has write access

LOADING TABLES FROM VISUAL ANALYTICS ADMINISTRATOR

In order to load tables manually, first register the data source in SAS metadata either from SAS Management Console or SAS Visual Analytics Administrator

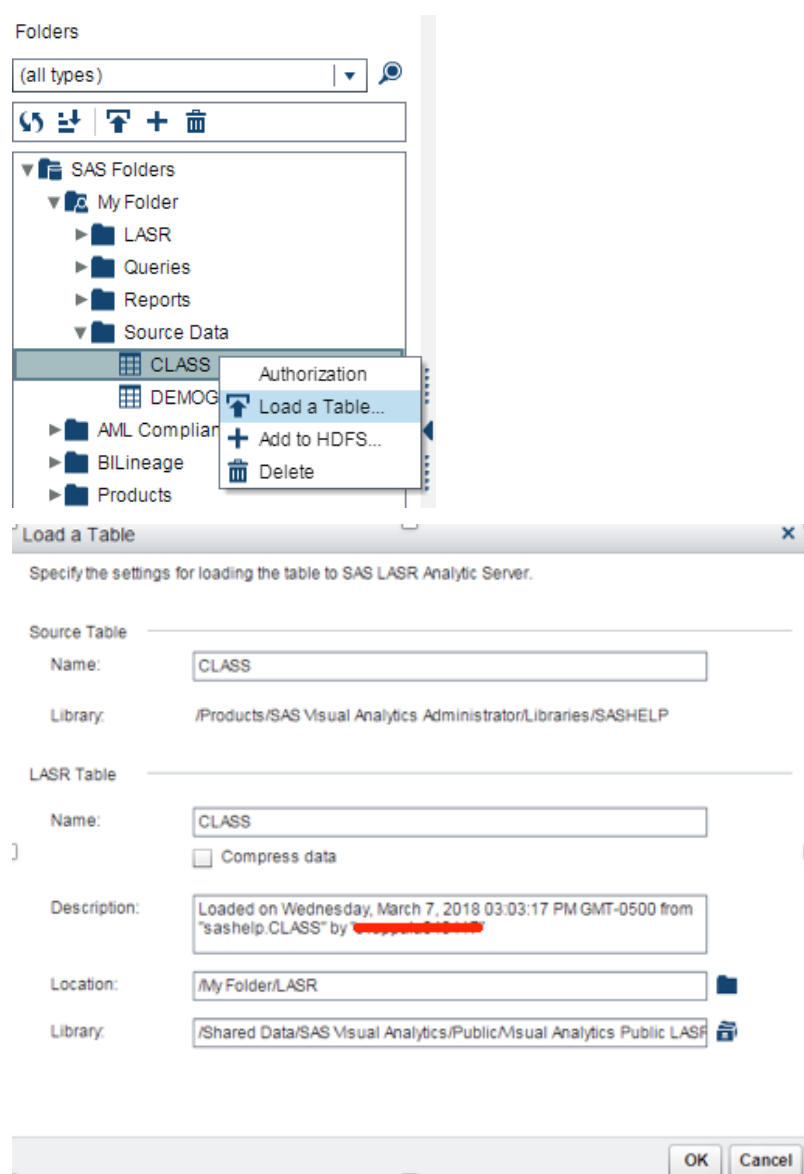
Log on to SAS Visual Analytics Administrator → Go to Folders → Select any metadata library and right click → Select Register and Update Tables → Select all the tables that need to be registered → Specify the metadata location where the tables will be located and click OK. A message will be displayed that the tables are registered in metadata

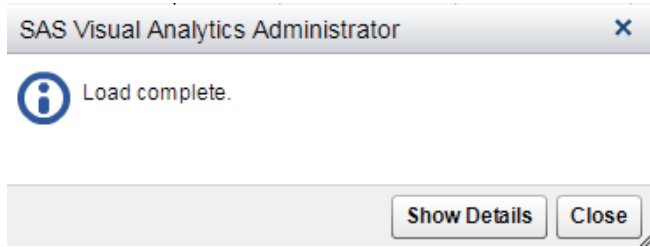




Once the tables are available in metadata, follow the below steps to load the table into memory

Select a registered table in metadata → Right Click and select Load a Table → Specify options for LASR library name, table name and metadata location and click OK. A message will be displayed that the table has been loaded to memory





TROUBLESHOOTING DATA LOAD ERRORS

For troubleshooting any errors related to loading data into memory, click on Show Details button when a message is displayed that the load failed. In this section, most of the conventional errors while loading tables into memory are discussed.

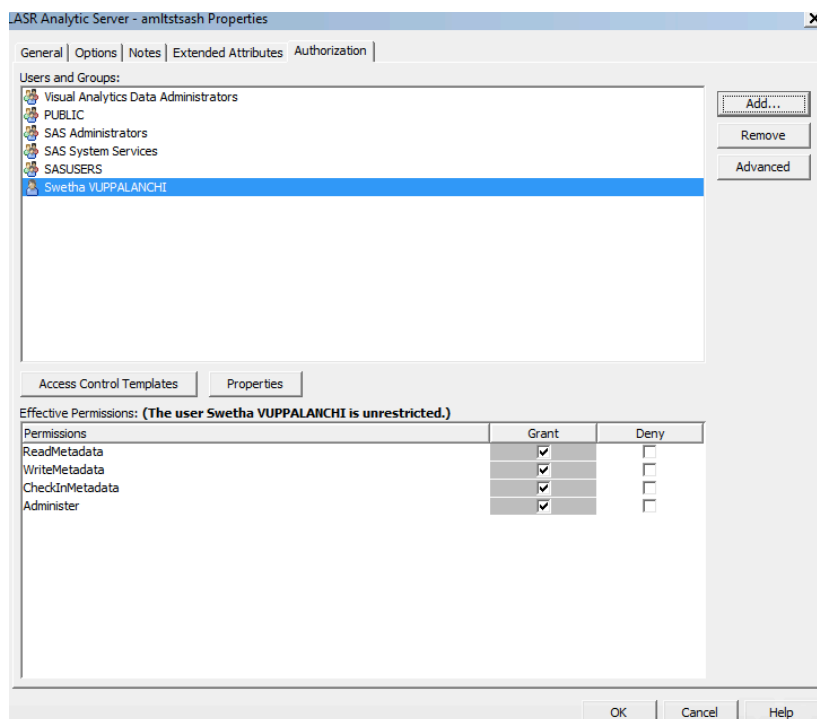
- **PERMISSION ISSUES**

The User ID that is being used to load the tables into VA should have read and write access to the physical dataset that is being loaded into memory and also should have write permissions to the SAS folders on SAS Metadata where the data is being written. Following error may be encountered in the event of a permission issue

```
ERROR: An I/O error has occurred on file CSLASR.CUSTOMER_EXTRACT_MAIN.DATA.
NOTE: The SAS System stopped processing this step because of errors.
```

Make sure, the LASR server where the tables are being written to has write access to the user loading the tables as well as the metadata table that is being loaded into memory has write access to the user

To provide write permissions on the LASR server, from the LASR server library properties, go to the authorization tab, add the user or group to be given permissions and grant all permissions to enable the user to load datasets to the associated LASR Server library. Administer role is optional.



To provide write permissions to the metadata table, access the table either from SAS Management Console or Visual Analytics Administrator and provide write permissions through Authorization tab

From SAS Visual Administrator, select the metadata table, right click and select Authorization to check whether the user loading the table has write permissions

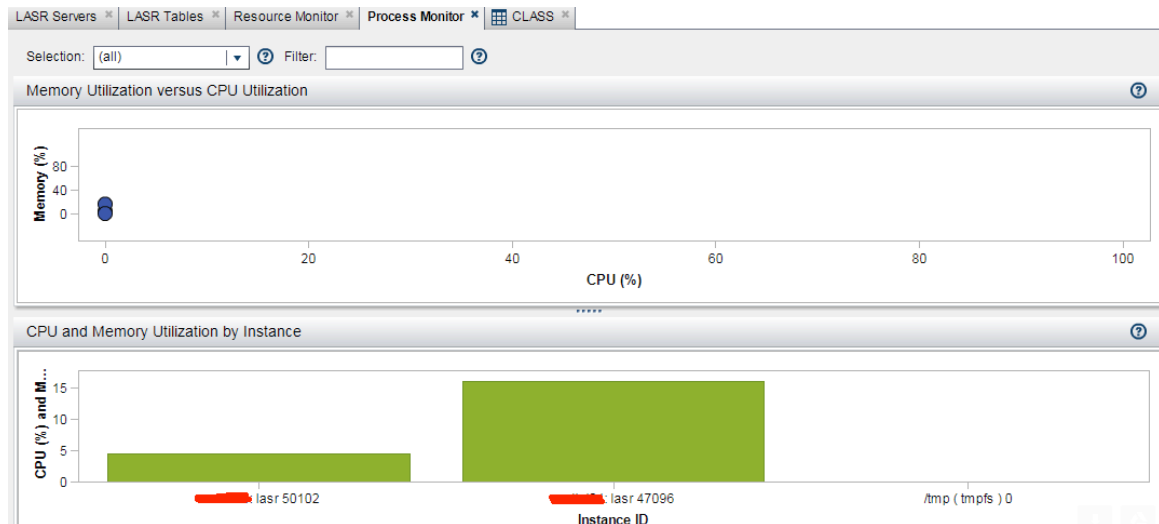
Identity	ReadMetadata	Read	WriteMetadata	Write	Administer	Create
PUBLIC	⊗	⊗	⊗	⊗	⊗	⊗
SAS Administrators	⊙	⊙	⊙	⊙	⊙	⊙
SAS General Servers	⊗	⊙	⊗	⊗	⊗	⊗
SAS System Services	⊙	⊗	⊗	⊗	⊗	⊗
SASUSERS	⊗	⊗	⊗	⊗	⊗	⊗
Swetha VUPPALANCHI	⊙	⊙	⊙	⊙	⊙	⊙

- **MEMORY ISSUES**

In some instances, the LASR may run out memory and give the following error when it is not able to load any more tables into memory. Following error is encountered if the memory consumption exceeds the specified limit.

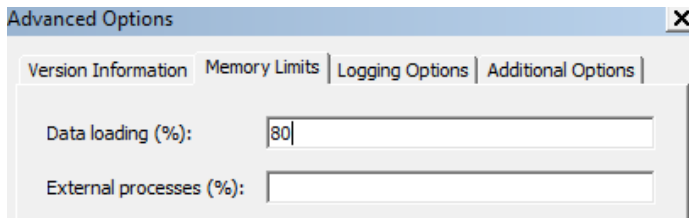
```
ERROR: A server-side limit on the consumption of memory resources has been reached. These limits can be adjusted by the owner or by the administrator of the LASR Analytic Server.
ERROR: 1016: A server-side limit on the consumption of memory resources has been reached. These limits can be adjusted by the owner or by the administrator of the LASR Analytic Server.
```

There may be multiple reasons for memory resources to be exhausted. It may be that the tables loaded in memory are large in number and very huge in size that may eat up all available memory or the actual in memory datasets may be occupying limited space but there are some hanging processes on the VA servers that are consuming a lot of memory. So before unloading any tables, check the Process Monitor tab to see if there are any rogue processes on the VA head node or worker nodes and kill the corresponding processes manually. Typically, if any processes show up under /tmp they are defunct processes and can be terminated. Restarting the web server may also kill some idle processes and free up some memory.



The memory limit for a LASR server is set to 80% by default and can be changed in settings under Advanced Options of LASR server in SAS Management Console. If the consumption of physical memory exceeds this limit, then no more tables can be loaded into memory until the memory is released or increased. The memory utilization of the VA servers can also be checked using a script

and the process is described in detail in LASR memory configuration guide referenced in the recommended reading section

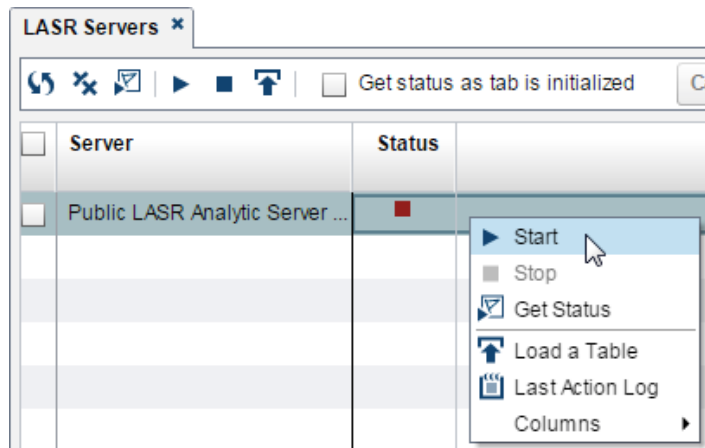


- **LASR SERVER NOT STARTED**

If the LASR server is not switched on, following error might be encountered when loading the tables. Turning on the LASR server manually from VA Administrator Console with administrative privileges would resolve the issue. Following error might be encountered when trying to load a table to a LASR server but it is not running

```
ERROR: A connection to the LASR Analytic Server on '██████████', port 10081, could not be made. Make sure that the host and port
are correctly specified, that you are attempting to connect to a LASR Analytic Server of the correct vintage, and that the
server is still running.
NOTE: The SAS System stopped processing this step because of errors.
NOTE: PROCEDURE LASR used (Total process time):
      real time          0.00 seconds
      cpu time           0.00 seconds
```

The LASR server can be started from the Visual Administrator by navigating to LASR Server → Check the box for the server that has to be started and click on **Start**

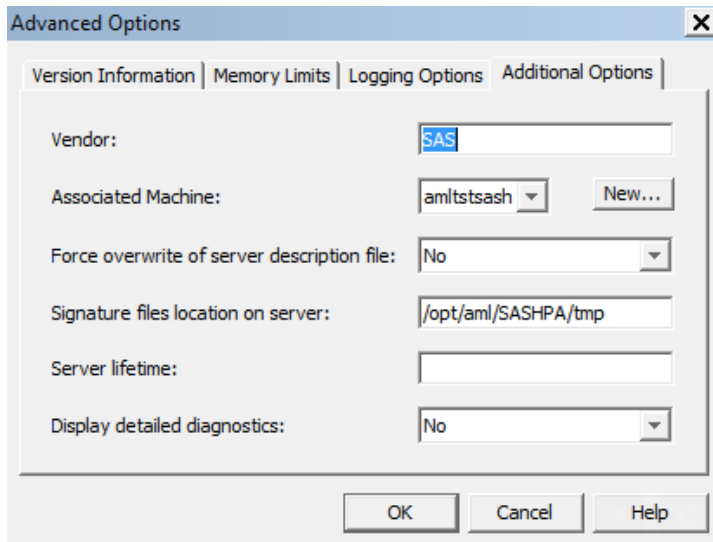


- **ACCESS TO LASR SIGNATURE FILES**

LASR Signature files are used as a security mechanism for server management and for access to data in a server. These signature files are created whenever LASR server is started, stopped or table is loaded or unloaded from LASR server so write access is needed to the signature file location for all these operations. If the user loading the tables does not have write access to this location the following error might be encountered

```
ERROR: Failed to create LASR Analytic Server Signature File. Make sure the path is specified correctly and that
you have write permission to the path
```

This signature file path can be configured from the SAS Management Console under LASR Server properties → Advanced Options → Additional Options → Signature file location on the server



- **LICENSE EXPIRATION**

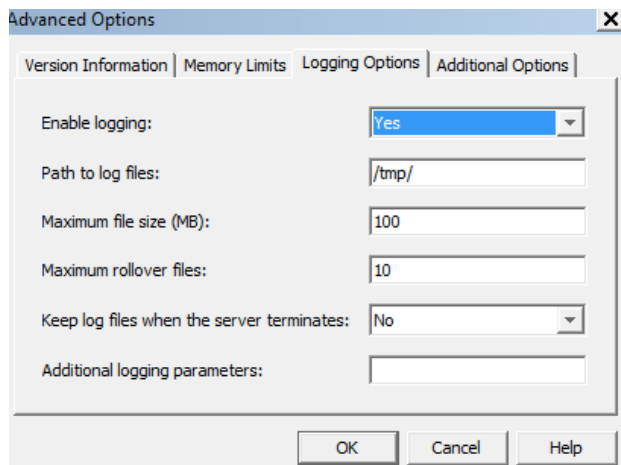
In the event of license expiration on SAS Metadata Server, the following error may be encountered

```
UNLOADING CSLASR.ALERT_INVESTIGATION_EXTRACT FROM LASR
-----
NOTE: REMOVE FROM LASR
ERROR: Metadata Server denied access to operation.
```

Updating the SID file in SAS Metadata Server with new license information using the SAS Deployment Manager would resolve the issue with the license

- **NONE OF THESE**

If the errors encountered does not match to any of the errors described above, then check the corresponding log where the failure has occurred by selecting the option 'Show most recent log' from the table or server that have issues from VA Administrator or click on Show Details button when message is displayed about the failure to get some direction on the error. The LASR logs are also stored under /tmp by default for UNIX if the option to store logs is set to 'YES'. These logs may also provide some insight into the underlying issue. The setting for these logs is under LASR Server Properties → Advanced Options → Logging Options



CONCLUSION

This paper addresses different ways of loading data into SAS Visual Analytics before it can be used for exploration and reporting, it also describes the autoloading procedure and the process of loading data manually into LASR, lays out the differences between the PUBLIC and PRIVATE LASR servers, talks about some important configuration options for LASR servers and libraries and most importantly helps diagnosing the most frequent errors encountered when loading data into SAS LASR servers. Therefore, this paper serves as a quick guide for dealing with data load problems in SAS Visual Analytics.

REFERENCES

- **SAS® LASR™ Analytic server Reference Guide**

<http://support.sas.com/documentation/cdl/en/inmsref/67213/HTML/default/viewer.htm#n0a090rw1gy20un1xvhhbe7deb1i.htm>

RECOMMENDED READING

- **Details on how LASR Server works**

<https://blogs.sas.com/content/sgf/2017/04/17/how-lasr-servers-are-started-from-sas-visual-analytics-administrator>

- **Memory Management in LASR Server**

<http://go.documentation.sas.com/?docsetId=inmsref&docsetVersion=2.81&docsetTarget=p050lknh5xepngn1s2kaa5nijrer.htm&locale=en>

- **New Implementations of Autoload for PRIVATE LASR**

<http://support.sas.com/training/tutorial/gel/gelva03.html>

- **Autoloading database tables in PRIVATE LASR Server**

<http://support.sas.com/resources/papers/proceedings16/3660-2016.pdf>

CONTACT INFORMATION

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

Swetha Vuppalanchi
732-397-1022
swethavup@gmail.com