Date: February 17, 2023

SAS® Visual Analytics - Data Preparation - Visual Statistics - Econometrics - Optimization - Visual Forecasting on VIYA 3.5 Hardware Estimate for: University of Hawai'i of Manoa

Sales Tracking Number: 16403010
ServiceNow Number: RITM0525268

This is not an architectural design or installation document.

A new Sizing document is recommended for requests more than 90 days old due to potential changes in SAS software or hardware/operating system technologies.

Critical Considerations:

The SAS Account team and University of Hawai'i of Manoa are responsible for ensuring the necessary licensing strategies.

The numbers presented here represent Usable RAM and DISK storage. For this exercise 2x was used as part of usable DISK storage space number.

A SAS Professional Services engagement is required for this installation.

This estimate is not a performance benchmark and does not provide any performance guarantee. Any use or reliance on this estimate is at customer's risk. SAS disclaims any liability with regards to (i) the hardware sizing estimate provided herein; or (ii) University of Hawai'i of Manoa reliance on the estimate.

University of Hawai'i of Manoa acknowledges that it is responsible for all costs associated with procuring any hardware

Note- This assume that the appropriate data management activities will happen outside of the SAS In Memory and resources for the data management activities are not included in this exercise.

As end-user concurrency increases, response times may be impacted.

Total amount of data to use in the entire Viya environment – this assumes basically the same data – although potentially different datasets – will be shared throughout the Viya environment.

Data Related Questions - Hint note the data increment (MB, GB or TB)	
What is the total size of the data that will be loaded onto Disk?	160 GB
What is the total amount of data that will be lifted into memory?	170 GB
Additional details that may assist in understanding the workload.	

**** All data sizes assumed to be uncompressed.**

SAS Visual Analytics

SAS Visual Analytics – User and Data details		
How many total users will have access to the environment?		7
Session Type	***Concurrent Sessions**	***Data set sizes (specify MB/GB/TB)**
Explorer/Admin (i.e. Heavy users)		
Exploiting largest data set	1	103 GB
Exploiting average data set	2	10 GB
Designer/Visual (i.e. Light users)		
Exploiting largest data set	1	10 GB
Exploiting average data set	2	2 GB
Mobile Users		
How many mobile device users	0	

**** All data sizes assumed to be uncompressed.**

SAS® Data Preparation and Data Quality

SAS Data Preparation/Quality *(During overnight BATCH Window)* – User and Data Details		
How many total jobs will be executed in this environment?		7
Session Type	***Concurrent Sessions**	***Data set sizes (specify MB/GB/TB)**
Data Preparation	2	1.2 GB
Data Quality	2	1.2 GB
Join Tables (joining cleansed and ordered data tables)	2	1.2 GB

**** All data sizes assumed to be uncompressed.**

SAS Visual Statistics

SAS Visual Statistics - User & Data Details		
Total number of users in environment	7	
Number of VStat User Sessions by Data Set Size	**Sessions**	**Size (Give Units: MB, GB or TB)**
Concurrent sessions using Linear Regression	2	10 GB
Concurrent sessions using GLM	1	10 GB
Concurrent sessions using Logistic Regression	1	10 GB
Concurrent sessions using Decision Tree	1	10 GB

**** All data sizes assumed to be uncompressed.**

SAS® Econometrics (On Viya)

SAS Econometrics Server User and Data details		
Do they plan to run this as batch or at prime-time hours?	Run as batch (separate from all other products)	
Total number of users in environment	7	
Number of Econometric Sessions by Data Set Size	**Sessions**	**Size (Give Units: MB, GB or TB)**
Concurrent sessions using CCOPULA	1	10 GB
Concurrent sessions using CNTSELECT	1	10 GB
Concurrent sessions using CPANEL	0	
Concurrent sessions using CQLIM	2	10 GB
Concurrent sessions using CSEVERITY	2	10 GB

**** All data sizes assumed to be uncompressed.**

SAS® Optimization (On Viya)

SAS® Optimization Server User and Data details	
Do they plan to run this as batch or at prime-time hours?	Run as batch
How many total users will have access to the environment?	4
How many concurrent sessions?	2
What is the average dataset size?	14 GB
What is the largest dataset size?	103 GB
Additional details that may assist in understanding the users, such as expected growth over the next 12 months.	

**** All data sizes assumed to be uncompressed.**

SAS® Visual Forecasting

SAS® Visual Forecasting (VF) Server User and Data Details			
How many total users will have access to the environment?			7
Maximum desired run time [in minutes] for VF projects?			30 minutes
How many maximum concurrent sessions?			2
What is the maximum number of time series to be forecasted in a Time Series table?			1000
What are the maximum attributes that are associated with each time series in the Time Series table?			20
What is the maximum number of time series in an External Forecast table?			0
Is hierarchical modeling is being used? (If yes, answer below question)			Yes
If hierarchical modeling is being used, what is the maximum number of levels in the forecasting hierarchy for the Time Series table?			3
Number of VF User Sessions by Data Set Size	**Concurrent Sessions**	**File Size**	**MB/GB/TB**
Number of concurrent sessions using a Time Series table	2	1.5	GB
Number of concurrent sessions using an External Forecast table			
Additional details that may assist in understanding the users, such as expected growth over the next 12 months.			

**** All data sizes assumed to be uncompressed.**

SPRE

Prime Time workloads

Session Type	Concurrent Sessions	Average Input Data Volume per Session (specify MB/GB)
Base SAS related to data prep / ETL	3	1 GB
Basic Statistics	3	1 GB
Advanced Analytics (STAT)	2	1 GB
IML Sessions	1	300 MB
ETS Sessions	1	1 GB
QC Sessions	3	1 GB
OR Sessions	1	1 GB

Non-Prime Time (Batch) Workloads

Session Type	Concurrent Sessions	Average Input Data Volume per Session (specify MB/GB)
Base SAS related to data prep / ETL	2	300 MB
Basic Statistics	3	1 GB
Advanced Analytics (STAT)	2	1 GB
IML Sessions		
ETS Sessions	1	1GB
QC Sessions		
OR Sessions	1	1GB

Hardware Estimate – Based on customer specific request and is a standard SAS Viya Server configuration from a hardware perspective. Any additional data will require increased hardware.

This is an estimate of the hardware resources required by the SAS solutions being sized. Implementation issues and architectural design for this environment should be addressed by the appropriate teams within Professional Services, Global Enablement Teams, or domain experts that can provide more specific information based on the needs of the University of Hawai'i of Manoa.

SAS Visual Analytics - Data Preparation - Visual Statistics - Econometrics - Optimization - Visual Forecasting on Viya 3.5 Server	
# Servers	5 (4 CAS Worker Nodes + 1 CAS Controller Node)
CPU per server	CAS Worker Node: 2 x 8 cores Intel Xeon Gold 6234 processors (3.3 GHz) CAS Controller Node: 1 x 8 cores Intel Xeon Gold 6234 processors (3.3 GHz)
Total cores	72
Memory Clock Speed	2933 MHz
RAM per node	CAS Worker Node: 192 GB CAS Controller Node: 96 GB
Operating System	Red Hat Enterprise Linux
NIC	10 GbE
SAS Version	VIYA 3.5
Local Disk per node	2 x 480 GB SSD

- **Server power settings need to be set to maximum, not the factory setting.**
- **Hyper-threading is recommended for all production CPU's.**
- **For better performance, it is recommended to use SSD's drives instead of HDD's.**
- **This assumes uncompressed data.**

For this configuration 1 server (8 cores) is dedicated for CAS Controller and Micro services**.**

The remaining 4 servers (64 cores) are CAS worker nodes.

Recommended CASCACHE minimum: 280 GB

SPRE: 1 server with 4 cores and 96 GB RAM and 2 x 480 GB SSD.

Dev / Test: 1 server with 16 cores and 192 GB RAM and 2 x 480 GB SSD.

This response is based on the **Dell** servers with Intel Xeon processors; other server models are also available provided they meet the minimum specifications listed above.

The Sizing team is responsible for providing the required number of cores and memory for the solutions as requested.

Our best practice is to provide the topology as developed by R&D and try to provide as unified a presentation of the requirements as possible. When questions on deployment arises, the sizing team defers to the account team to decide, based on the customer's desires and the best practices of the solutions involved, on the most practical installation.

Assumptions:

SAS is a very powerful tool and therefore can require significant resources. The following assumptions were used to arrive at the results and estimation of needed resources:

- SAS is a fourth-generation programming language and allows customers the flexibility and power to use the software in novel ways that exceed the original intended ordinary and typical use of the software. Such use may result in workloads in excess of those described in the estimate and may require additional resources.
- This sizing is based on SAS VIYA best practices, guidelines, and benchmarks.
- This sizing estimate is based on a combination of guidelines provided by SAS R&D, SAS Product Management and test data.
- Changes to the workload (in either number of sessions or data volumes), operating system, or preferred vendor or chipset may render this sizing as void. In the event of changes, the SAS Account Team should resubmit the questionnaire with the needed updates for reprocessing.

Additional Information – Recommended Reading:
Storage Assumptions:

University of Hawai'i of Manoa may wish to consider the peak I/O throughput requirements of their system and work with their storage provider to ensure that the storage environment can provide the level of I/O required. A significant percentage of "performance problems" reported to SAS Technical Support can be directly attributed to insufficient levels of I/O throughput.

Please refer to these papers for more information on how to configure your storage for SAS applications:

- **SAS® Viya™: What It Means for SAS® Administration -** http://support.sas.com/resources/papers/proceedings17/SAS0644-2017.pdf
- **Parallel Programming with the DATA Step: Next Steps** https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2184-2018.pdf

- **Top SAS® Architecture Pitfalls: Lessons Learned the Hard Way**
  https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3609-2019.pdf
- **Important Performance Considerations When Moving SAS® to a Public Cloud**
  https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2020/4312-2020.pdf
- **Moving SAS applications to MS Azure:**
  https://communities.sas.com/t5/Administration-and-Deployment/Moving-SAS-applications-to-MS-Azure/m-p/630295/highlight/false#M18426
- **Engineering CAS Performance Hardware Network, and Storage Considerations for CAS Servers -**
  https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3351-2019.pdf
- **2019 - Optimizing SAS on Red Hat Enterprise Linux (RHEL) 6 and 7 (v.1.3) -**
  https://access.redhat.com/articles/1593123
  https://access.redhat.com/sites/default/files/attachments/optimizingsasonrhel6and7_2.pdf
- **Migrating from SAS® 9.4 to SAS® Viya® 3.5**
  https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2020/4457-2020.pdf
- **Proper Planning Prevents Possible Problems: SAS® Viya® High-Availability Considerations -**
  https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3481-2019.pdf
- **Multi-tenancy in SAS® Viya®: Considerations and Implementation**
  https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3322-2019.pdf

Disclaimer

The hardware sizing recommendation herein is a no fee service. The recommendation is based on expected usage and other technical details provided by University of Hawai'i of Manoa and SAS' knowledge and expertise regarding their respective products. This hardware sizing recommendation is not a performance benchmark and does not provide any performance guarantee. University of Hawai'i of Manoa actual usage of the SAS software on the recommended hardware may not be consistent with the usage data provided by University of Hawai'i of Manoa resulting in differing performance. SAS disclaims any liability with regards to (i) the hardware sizing recommendation provided herein; or (ii) University of Hawai'i of Manoa reliance on the recommendation. University of Hawai'i of Manoa acknowledges that it is responsible for all costs associated with procuring any recommended hardware and acknowledges that changes in hardware may result in SAS software upgrade license fees being payable.