

Using the SAS® Intelligence Platform for Clinical Trials Programming

Donna Dutton, SAS Institute Inc., Cary, United States

ABSTRACT

In the past, much of the programming for Clinical Trials analysis has been performed using Windows or UNIX-based SAS installations running interactive SAS sessions or batch jobs. This makes the sharing of computer resources and consistent permission enforcement very difficult. However, by utilizing the SAS® Intelligence Platform architecture in a server-based configuration, computing power can be maximized and maintenance of PC-based software reduced. Security can be managed through SAS® Management Console and interaction with server-level OS security. The SAS® Enterprise Guide client is used to run jobs directly on the server, and the SAS® Add-in for MS Office is available for use within Microsoft desktop applications. Other SAS Intelligence Platform products can be used to surface results.

INTRODUCTION

Clinical Trials programmers have long been familiar with the use of the SAS Windowing Environment (formerly known as the Display Manager System or DMS) for the development of programs. However, with the overhead of maintaining local PC SAS software and security, or the consideration of running many interactive SAS sessions on a single UNIX server, is this really the best development model today? The SAS Intelligence Platform was introduced in 2002 with Base SAS® 9.1, and has evolved through SAS 9.2, 9.3, and the newest release, SAS 9.4. In this presentation, we examine the various components of the SAS Intelligence Platform Architecture and how they can be used to form a robust, enterprise-class programming environment for the Clinical Trials programmer.

THE BASIC INTELLIGENCE ARCHITECTURE

The Intelligence Architecture distinguishes itself by having four logical tiers: Data Sources, SAS Servers Tier, Middle Tier and the Client Tier. These logical tiers may reside on a single physical machine for very small applications, or on individual servers or server farms for larger implementations.

The Data Source tier is undoubtedly the most familiar tier to those of us who have used SAS Software for any length of time. It consists of any kind of data which can be accessed by any of a number of SAS products. In some cases, it could reside on its own physical server, for example, in the case of a large data base system. However, we are most familiar with this logical tier residing on one or more hard drives or file systems which are accessible to the computer on which we run our SAS programs.

In the SAS Server Tier the SAS® Metadata Server, first among the critical SAS Servers, controls users, permissions, access to libraries, and the SAS Application servers. The Workspace servers, Stored Process servers, and Batch Execution servers are dynamically executed (spawned) SAS sessions which run submitted SAS programs. Each one has its own role. SAS Workspace servers are used by SAS products such as SAS Enterprise Guide, SAS® Clinical Data Integration, SAS® Data Integration Studio, and SAS® Drug Development to give users a personal space in which to execute their SAS programs. Workspace servers are much like an individual SAS session on a Windows PC, except that they are created dynamically, usually on a larger computer, when a SAS Client application calls for them. Web-based clients such as Web Report Studio or the SAS Information Delivery Portal typically employ Stored Process servers to execute predefined SAS programs (stored processes) either with or without parameters passed to them. Batch Execution servers are used to run background non-interactive, predefined SAS programs, frequently on a scheduled basis, by various clients.

The Middle Tier presents and provides data to web-based clients such as Web Report Studio and the Information Delivery Portal. Prior to the SAS 9.4 Intelligence Platform, a third party Web application server such as JBoss Application Server, IBM WebSphere Application Server, IBM WebSphere Application Server Network Deployment or Oracle WebLogic Server was required. However, the SAS 9.4 Intelligence Platform includes the SAS® Web Application Server and the SAS Web Server, eliminating the need for third-party web application server software.

Making up the Client Tier, we have PC-based clients such as Enterprise Guide or Clinical Data Integration (running in Data Integration Studio) and web-based clients such as Web Report Studio and the SAS Information Delivery Portal. The PC-based clients can be installed on individual Windows PCs or on a Windows server accessible by means of Remote Desktop Protocol (RDP). These clients use the various servers described above to allow users to develop and run SAS programs. The SAS Management Console (SMC) client allows administrators to define user Roles and Groups to control access to data and servers. See Figure 1 below.

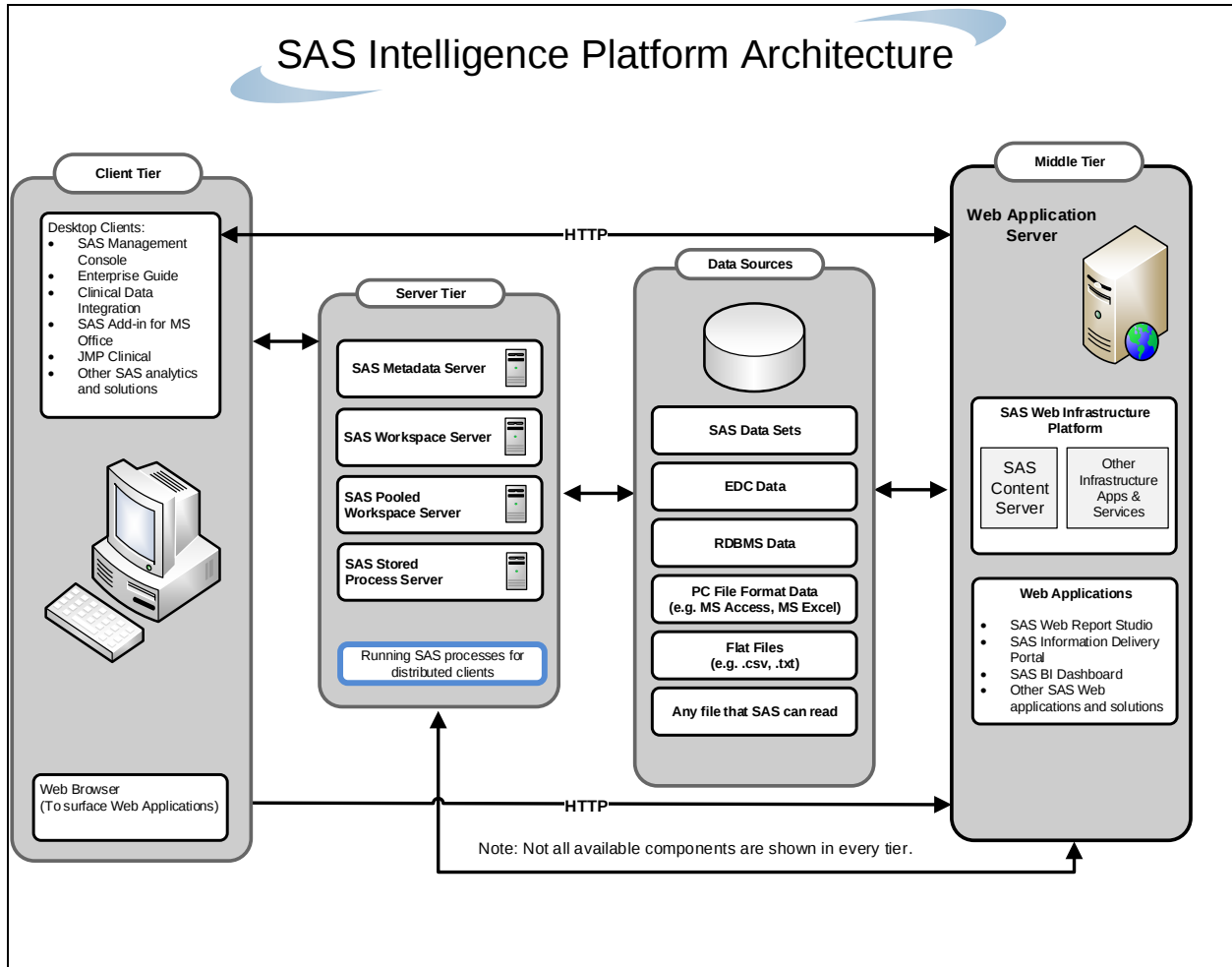


Figure 1: Logical Tiers of the SAS Intelligence Architecture

A DAY IN THE WORK-LIFE

From the programmer's local Windows computer, which is running the Enterprise Guide and Clinical Data Integration clients, our Clinical Trials programmer writes programs to work with small data sources stored locally, and large data sources stored on centralized storage devices or data base servers which are accessible from one or more SAS Application servers. Those application servers may be running on a large unix-based server in a data center, providing the computing power to handle large amounts of data.

Enterprise Guide provides a programming environment similar to that provided by SAS Display Manager System, but with enhancements such as autocomplete for keywords, dynamic syntax help, and the ability to group programs into process flows and projects. In addition, Enterprise Guide is able to execute the code locally, or to submit the code for execution on one or more SAS Application servers. The log and listing are returned to the Enterprise Guide session.

In addition, SAS Clinical Data Integration may be used to develop jobs which transform data into standardized CDISC data structures (SDTM, ADaM) using prebuilt data models and processes, then validate the structure and content of data for conformance to selected data standards (such as CDISC SDTM). These capabilities can be combined with custom user-written code and with other Data Integration Studio transformations to integrate various data feeds, translate legacy data into more modern formats, and to create complex job flows.

PhUSE 2013

Results can be published and made available over the web through pre-defined secured reports in Web Report Studio and stored processes and/or progress indicators surfaced through the Information Delivery Portal.

Using the SAS Metadata Console client, an administrator can define various Groups and Roles to implement membership in different projects. SAS Clinical Data Integration includes the Clinical Administrator role. An administrator can also create other roles such as Data Manager, Clinical Programmer, Biostatistician, Reviewer, and so on. When combined with operating system level data security, the ability to grant the right people the right level of access to the right data is achieved.

A SAMPLE VERSION 9.3 IMPLEMENTATION

As stated earlier, the four logical tiers can be distributed across multiple computers in order to achieve the scalability and performance needed to support the number of programmers and the amount of data expected on the system. While it is possible to place all tiers on a single Windows computer, it is not a typical or advantageous use of the capabilities. If there were only a single programmer working in isolation, a single Windows installation of SAS Software would suffice. But, in our work with clinical trials data, this is never the case. Instead, there is data that must be obtained, shared and secured. Programs must be developed, tested, and then moved to production. Perhaps a separate server is needed for batch execution of finalized analytical programs. The SAS Intelligence Platform can support all of these needs, and grow and change as a customer's needs change.

A typical configuration separates the SAS Server tier onto multiple computers, using one for the SAS Metadata Server and one (or more) for the SAS Application servers, such as the SAS Workspace Server, the SAS Pooled Workspace Server, the SAS Stored Process Server and the SAS Batch Execution Server. Or, one or more of these Application servers could be placed on different computers. See Figure 2 below:

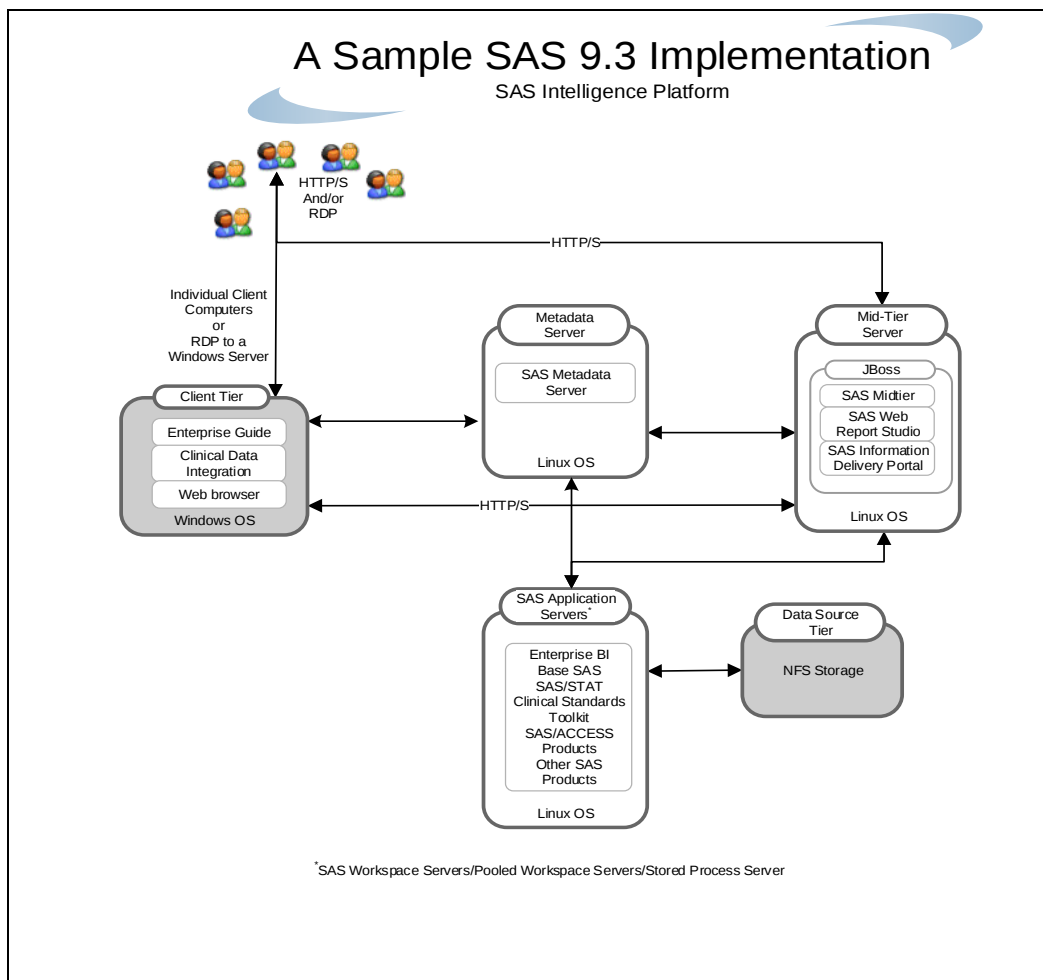


Figure 2: Sample SAS 9.3 Implementation

PhUSE 2013

CONCLUSION

In our state-of-the-art Clinical Programming environment, we observe a number of improvements over the older model. In addition to reduced software maintenance, data does not have to be duplicated and stored on local machines. Larger servers can be used to support large data and programs with high computational requirements. Security can be controlled from SAS Management Console and the physical servers. The ability to run a program on any of several different servers from a single interface allows for precise security definitions and appropriate resource scaling – administrators can choose to associate physical and logical servers with specific studies, groups, roles or resource requirements. The Enterprise Guide programming interface provides at-your-fingertips syntax assistance, autocomplete for keywords, the ability to group programs into process flows and projects, and a myriad of options to customize things like white space, indentation, fonts, and colors. You may just end up liking it even more than our beloved Program Editor window! And to make it easier on the IT Department, resources can be increased by adding to the centralized servers rather than by upgrading or adding resources to an individual programmer's local machine. Overall, a move to a SAS Intelligence Platform based Clinical Trials programming environment is a leap into SAS' most current technology, which yields enterprise-wide benefits.

REFERENCES

SAS V9.3 Intelligence Platform: Overview

<http://support.sas.com/documentation/cdl/en/biov/63143/HTML/default/viewer.htm#titlepage.htm>

SAS V9.4 Intelligence Platform: Overview

<http://support.sas.com/documentation/cdl/en/biov/64876/HTML/default/viewer.htm#titlepage.htm>

CONTACT INFORMATION

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

Donna Dutton
SAS Institute Inc.
SAS Campus Dr.
Cary, NC 27513
USA
Work Phone: +1 (325)622-4541
Fax: +1 (325)622-4541
E-mail: Donna.Dutton@sas.com

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.