Paper AD06

SAS 9.3 Migration — A Project Experience

Jules van der Zalm, OCS Consulting, 's-Hertogenbosch, the Netherlands
Raymond Ebben, OCS Consulting, 's-Hertogenbosch, the Netherlands

INTRODUCTION

As a top 20[1] pharmaceutical company, Astellas is an important player in this market. With major offices in Japan, the Netherlands and the USA the company is globally active. A vast team of statisticians and clinical trial programmers use SAS® to review and report clinical data.

OCS Consulting has been involved in the maintenance of Astellas' SAS environment for quite a few years. Recently that maintenance has been transferred to an offshoring partner with OCS Consulting still involved in various projects that go beyond the regular maintenance of the environment. The migration to SAS 9.3® is one of those projects.

This paper will elaborate on why Astellas and OCS Consulting started this project, the approach that has been taken and the challenges that were faced. A glimpse of the outcome will also be given but as the project is still ongoing at the time of writing, the ultimate result cannot yet be disclosed.

[1] Source: http://seekingalpha.com/article/287269-the-50-largest-pharmaceutical-companies-by-sales

BUSINESS DRIVERS

For Astellas there are two important reasons to start this project.

1) The primary objective is to increase the availability of the system by introducing a load balanced system with failover functionality in order to increase the system's uptime.

2) The secondary objective for this project is the migration from SAS 9.1.3 to the newest SAS release, 9.3. This release provides many desired improvements such as the new ODS graphics. Migrating to the newest SAS version also ensures that the Level A Support Status with SAS is maintained.

During the project a tertiary objective was added, which is to make Astellas' custom portal environment available on the development and validation servers instead of only on the production server.

SYSTEM UPTIME/DOWNTIME AS A BUSINESS DRIVER

Until 2010 the SAS environment of Astellas could handle extensive usage of approximately 50 concurrent users. At that time they were only users based in Europe. Given the single time zone users were in, there were flexible maintenance windows as there were more than 12 out-of-business hours. Also, due to the relatively small numbers of users, potential downtime of the servers would not cause too much financial damage, especially when compared to the cost of running redundant servers.

In 2010, Astellas started a project to make the existing SAS environment available to users in the United States. After the completion of this project, the SAS environment serviced a maximum of 250 users. Due to the time difference with Europe the maintenance windows became much smaller and on less convenient times; the Dutch and German support teams could perform maintenance only during the nightly hours in Europe in order to not interrupt the users' daily activities. The expanded user base also increased the financial risk in case of system downtime.

An important aspect of this project is the introduction of a redundant server environment. The idea is to have two servers sitting side by side sharing all the work. When one server fails the other server will take over all the activities, albeit with the performance of only one single machine and thus less responsiveness for the users. Introducing this failover mechanism should significantly reduce the system downtime. The target is to reach an uptime of 99,8% on a yearly basis; that means less than 18 hours of downtime per year. Performance gain is not an important aspect of this project; the current server suffices easily satisfies the performance requirements.

USER BENEFITS

A key goal of this project is to introduce SAS 9.3. Astellas has skipped SAS 9.2 but as they use the metadata server quite extensively, this new version provides many important improvements in terms of security and ease of maintenance. Besides the metadata server, SAS 9.3 and Enterprise Guide 4.3 provide significant enhancements from an end user perspective, such as improved graphical output, several new procedures and new options.

THE APPROACH

There are many steps to take before the new environment can be installed; an approach should be defined. This is an important activity as it defines the path to follow to get to the desired end result. It will describe the responsibilities and tasks of each member of the project team.

The sections in this chapter each describe one step in the chosen approach.

DOCUMENT THE EXISTING SAS ENVIRONMENT

The first milestone of the project should be to map the existing server environment. Even though the new environment will be set up from scratch, it should and will become very similar to the existing one, to avoid confusing the users unnecessarily with changes to their daily routines.

PRODUCTION, VALIDATION, DEVELOPMENT

In the current situation the SAS environment comprises three separate servers: development, validation and production. Each server is a stand-alone machine running the same installation of SAS 9.1.3 on the Solaris 5.10 operating system. End users only use the production server for their daily work, within which they have a development, validation and production environment for SAS development. The validation and development machines are mainly used for software updates, updates to bespoke applications and to develop and validate per-study customisations and settings for an interface from Astellas' raw study data to a Clinical Trial Management System. Furthermore the production server runs a Stored Process Web Application hosting a bespoke portal environment, which Astellas uses to host their own bespoke framework running a collection of web-based applications supporting the monitoring, cleaning and reporting of study data.

DATA STORAGE

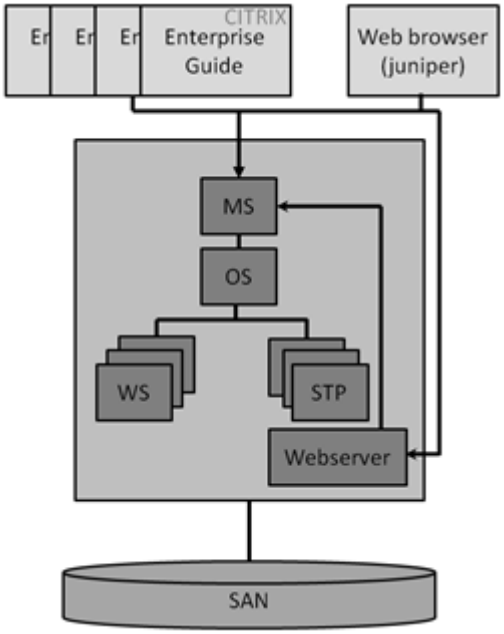
All the study related data is stored on an external Storage Area Network (SAN). More specific, there are two SAN's; one for the production environment and one shared by development and validation. These SAN's have their own failover systems implemented and are outside the scope of this project. The SAN's are mounted on the respective servers so the data on them is available to the users as if it were stored on the local server.

ADDITIONAL APPLICATIONS

On each server additional applications are installed and configured to support the use and maintenance of the systems. Among these applications are:

- Astellas' custom built web-based data management tools;
- clients to connect to external Oracle, Sybase and MS SQL databases;
- Java (JRE, JDK);
- Samba, to allow connections to the data on the server from within Windows;
- StatXact, to provide additional statistical procedures to the SAS users;
- EPSI, a custom built interface between Astellas' raw data sources and the CTMS;
- several other system tools.

SCHEMATICALLY



This is what the existing environment looks like schematically. Within the grey square is the SAS server environment with the SAS Metadata Server (MS) and the Object Spawner (OS) spawning multiple Workspace Servers (WS) and Stored Process Servers (STP). Users use either Enterprise Guide (through Citrix) or their web browser to connect to this environment. At the bottom is the SAN on which all the data is located. The server environment is connected to this SAN.

PREPARE FOR THE INSTALLATION

Before undertaking any action, the way the installation and configuration will be executed must be thought through very thoroughly.

PARTIES INVOLVED

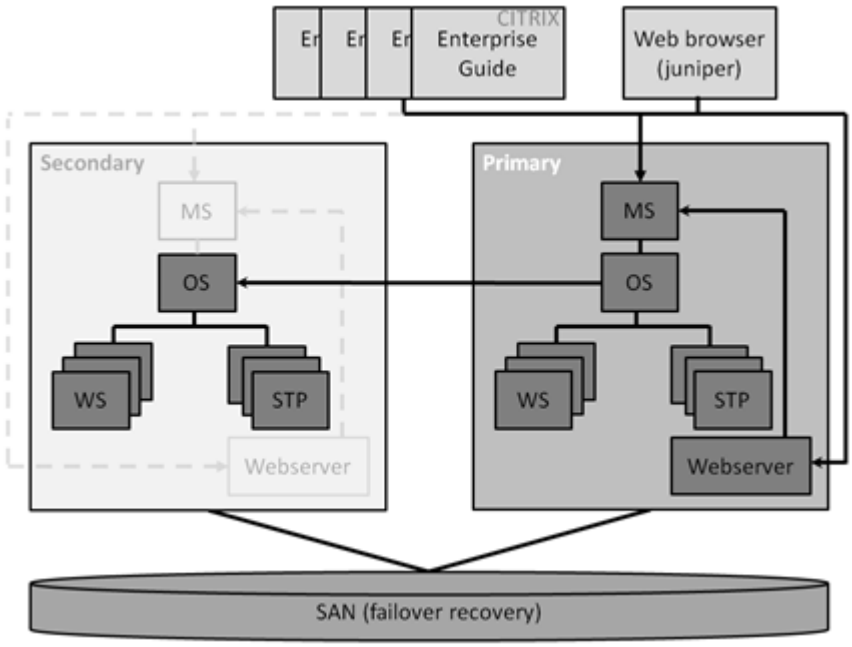
This project knows five significant involved parties.

- Astellas, the project sponsor. The project manager is an employee of the sponsor. He keeps close watch on the project timelines and ensures all parties involved are aligned.
- “Information management”, a third party company. Astellas has outsourced the maintenance of the SAS environment to this party;
- “Technical support”, a third party company. Astellas has outsourced the technical (infrastructure and Operating System) support and maintenance of its servers to this party;
- SAS, in an advising role for the system configuration;
- OCS Consulting, a third party company who was previously responsible for the functional and technical maintenance of the SAS environment and is now a preferred partner to Astellas for SAS related projects.

This paper is written mainly from the perspective of OCS Consulting.

DESIGN

Before any further preparations can be undertaken, the architecture of the target environment must be designed. The starting point for the new design is that to the users, as little as possible should change. Furthermore, the primary objective of the project must be met: the environment should have failover functionality to increase the availability, and the performance of the system must be improved by introducing load-balancing between servers. In consultation with all parties involved, the schematic below has been selected to be the design of the new environment architecture.



The right part of this schematic is similar to the existing environment: the MS and OS sitting on the server, the users connecting to it through EG and the web browser, and data being stored on the SAN. A similar environment will be set up on the primary server in the new environment. Alongside, there will be a secondary server, similarly installed and configured, with only the Object Spawner running and the Metadata Server inactive. Users don't connect to this server directly; it is set up to load balance with the primary server.

In order to achieve failover functionality, the secondary server is set up in such a way that it can fully replace the primary server. The Metadata Server on the secondary machine is disabled because within an single environment, there should be only one Metadata Server. As soon as the primary server would crash or shut down for whatever reason, the Metadata Server on the secondary machine is enabled. The server is then capable of taking over all duties. The users however would not be able to connect to it because their systems are configured to connect to a machine with the name of the primary servers. To resolve that issue, the DNS will be updated in order to route requests for the primary server to the secondary server. Enabling the Metadata Server and updating the DNS are manual processes for which detailed instructions will be documented and responsible parties will be assigned.

DOCUMENTATION

As required in a controlled (GCP) environment, the installation of the new environment will first be performed and tested on the development server before proceeding with validation and production. During this first deployment there will be a little back and forth on the installation and configuration in order to obtain the best sequence of events to get to the desired environment. The most important deliverable of this phase will be IQ and OQ documentation presenting how to install and configure the environment and all its components, and how to verify that the environment is according to the specifications. These IQ's and OQ's will be executed to install the environment on the validation and production servers.

LOAD BALANCING AND SERVER CONFIGURATION

An important component of this project is the addition of the load balancing server. It is also a challenging one, as at the end of the project only the production environment will consist of two servers while the validation and development environments consist of only one server. The approach taken is that the load balancing setup is introduced in development. Once the desired environment has been created and documented, one server will be taken out of the configuration and the physical device will be promoted to the validation environment. Once the validation environment is successfully installed (based on IQ's and OQ's), the secondary server will be promoted to production and installed there. Therefore there will be 6 installations in total, of which eventually only 4 will remain (2x production, 1x validation, 1x development).

The installation of one server consists of the following tasks in this order:

- Format/clean the server;
- Install the operating system + related applications (Java, database clients);
- Install and configure SAS + related software;
- Install and test other applications.

SOFTWARE

Besides the hardware configuration there is the software component of this project.

Operating system

The selected operating system is Solaris 5.10. This is the same as in the existing environment, but with a higher update level. The operating system is installed by the Technical Support team prior to each fresh installation. On the development server that may be required many times as the installation and configuration of the environment will be the result of 'trial and error' until the desired environment is obtained. Each fresh installation is restored from a backup instead of installing from scratch each time.

The installation of the operating system includes the installation of Unix applications such as Sudo, LFTP and Rsync.

SAS software

The most recent version of SAS 9.3 will be installed. The installation of SAS consists of the following three components:

- SAS Analytics (Base, Stat, Graph)
- SAS/ACCESS to MS SQL, Sybase and Oracle
- SAS Integration Technologies

Furthermore the following SAS services will be installed on each of the two servers:

- Metadata Server
- Object Spawner
- Stored Process Server
- Workspace Server
- Jboss Application Server

All these services will be enabled on the primary server. The secondary server will also have them, but only the Object Spawner will be active. This server will work side-by-side with the same service on the primary server in order to provide load-balancing functionality. The other services will remain disabled until the primary server would come to fail; then the services will be started to let the secondary server take over the duties of the failed primary server.

Bespoke portal application

Astellas' bespoke portal application, which provides several data management related applications, is currently only installed on the production server. There it has four instances: sandbox, development, validation and production. These four instances run within the same SAS services. Astellas' will use this project as an opportunity to copy the IDP to the development and validation environments. The production instance will be used as a basis for twelve new instances: sandbox, development, validation and production on each of the three servers.

Only the migration of existing functionality is within the scope of this project; if any issues are identified then these will be documented and resolved under standard change control.

EPSI

EPSI is the External Patient Status Interface. It is a bespoke application built in SAS and Java. It runs on a daily basis to push figures and patients, visits and withdrawals from EDC systems into a Clinical Trial Management System. It is developed and tested in SAS 9.1.3 and will need to run on SAS 9.3 in the new environment. After the installation of SAS on the development server EPSI will be installed and tested according to existing IQ and OQ documentation. Changes required to make the application run on the new Java and SAS version will be made and documented.

Non-standard changes

Since its initial installation, the existing production environment has undergone many changes. These are mainly related to the configuration, but may also be the installation of newer version of applications or creating a mountpoint to a new data source. They are non-standard changes. Such changes are not documented in the original IQ's or OQ's of the applications they apply to. To ensure that those changes are also applied to the new environments, each non-standard change is evaluated. If the change is still relevant then it will be considered for application in the new environments. In order to keep documentation as complete as possible the changes will no longer be applied after, but during the installation of the respective application or component.

METADATA

The upgrade to SAS 9.3 delivers enhanced security and configuration features, which gives Astellas the opportunity to review its existing SAS system and make enhancements to it. SAS distinguishes three levels on which security can be configured: Access Control Template (ACT's), Access Groups and Users. In order to keep things simple it has been decided to keep a one-to-one relationship between ACT's and Access Groups. Security is then set up by assigning users to groups, which have been assigned ACT's.

The new Metadata Server finally has a built-in metadata backup facility. This facility will replace the bespoke application that currently runs periodically. In contrast to this bespoke application, the built-in facility does not require the SAS services to be paused while the backup runs. The backup will be scheduled at the most likely time that no users are on-line and using the system.

STUDY CONVERSION

All of Astellas' study programming has been done in Enterprise Guide on the SAS 9.1.3 server. SAS 9.3 not only introduces new functionality, there are also several new defaults for certain functions and procedures which may cause existing programs to produce different results. The purpose of study conversion is to ensure that SAS programs generate the same output in SAS 9.3 as they did in SAS 9.1.3.

WHICH STUDIES TO CONVERT

Study programmers decide which SAS programs will be converted. In practice that means that if reports are still to be created for a study, these programs will be converted because otherwise the results may be unreliable.

PREPARING THE CONVERSION

In order to make the conversion process as smooth as possible, a Project Code Conversion & Verification Protocol is prepared. This is a document that specifies the proposed approach to converting a program, most common changes to programs, and additionally describes potential issues. Programmers can use this document as a guideline to converting their programs.

In order to prove that this document is sufficiently adequate, it will be verified using the Project Code Conversion Test Script. Two selected studies will be converted using the Conversion & Verification Protocol, and the result will be verified following the Conversion Test Script. If the outcome is positive then the Conversion & Verification Protocol is considered verified, and can be used by study programmers to convert their programs.

The validation of the programs converted using this script is the responsibility of the study team and is outside the scope of this project.

CHALLENGES

LOAD BALANCING *AND* FAILOVER MECHANISM

The setup with two servers that are supposed to take over from each other in case of a system failure, while also being each other's load balancing partner, is quite unique, even for SAS. As far as we know it has not been done this way before. Coming up with a method that would actually work was quite a challenge. Load balancing is a very common feature in SAS environments; one server usually can't handle the complete load, so multiple servers are connected to each other and share the requests. Failover, however less common, is also not too exotic to SAS. Combining these two principles however *is* rather exotic and requires careful planning and detailed instructions describing the steps to be taken in case one server should actually fail.

SHARED DATA STORAGE

All the study data is stored on a SAN. A SAN (Storage Area Network) is an architecture that connects servers and storage devices in such a way that they appear like locally attached devices. It only became apparent during a later phase of the project the type of SAN that Astellas uses can be accessed from only one device at a time, unless a very costly additional license would be obtained from the vendor. This would be an issue to the dual server setup, as both the servers need to be able to connect to the data; after all, the user's request might be passed to either server, so either server must be able to connect to the study data and that data must be only in one place, otherwise there can be discrepancies between both servers.

To resolve this issue, it has been decided to mount the SAN only on the primary server, and use NFS sharing to make the same data available on the secondary server. That way there's only one connection to the SAN, but the data is still identical and available on both servers.

DIFFERENT PARTIES, DIFFERENT RESPONSIBILITIES

The parties participating in this project each have their own responsibility and their own distinct competence: technical support can install and configure the Solaris operating system, but does not know their way around SAS. We, at OCS Consulting, know how to install and configure SAS, but are not Solaris administrators. Astellas, as the project sponsor, have their wishes and desires, but have no in-depth knowledge of the technical specifics of either SAS or Solaris. Finally SAS, who are only involved upon request for situations that we cannot resolve within the project team.

Furthermore the users, whose feedback is required at times, are in a different time zone and participating parties work from their own offices having responsibilities for other projects as well.

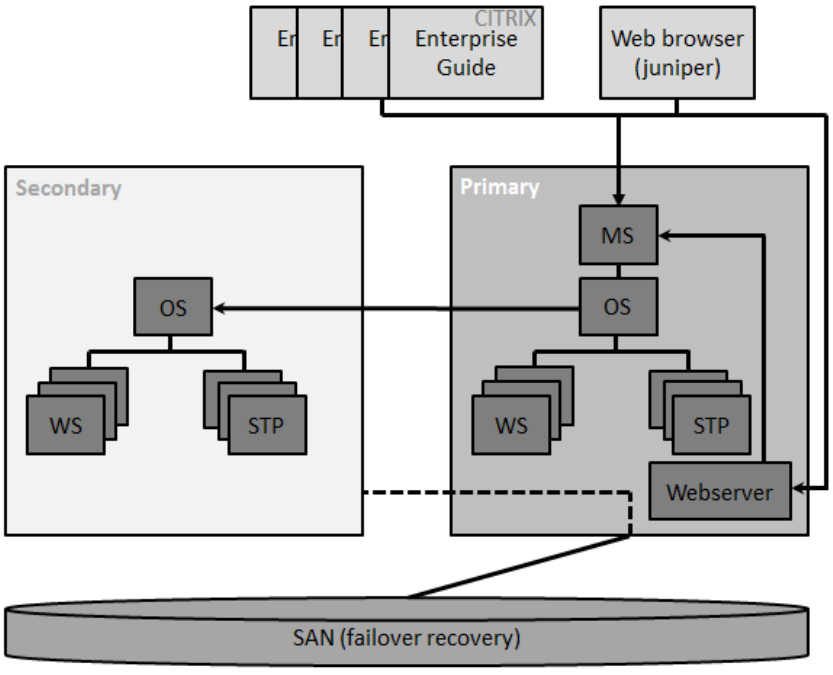
Because of all of this, there is the chance of slowing down the project. To ensure the project does progress as planned, frequent meetings with all parties are essential. During such meetings all current issues are discussed and tasks are distributed. It has also proven useful to meet in person every now and then to get a 100% focus on the tasks at hand. In hindsight we think that it might have been a good idea to organise workshops where all parties would sit together in the same room for multiple consecutive days thus making faster progress.

CHANGED DESIGN

When OCS discussed the proposed design of the dual server configuration with SAS, an alternative architecture was proposed. In this alternative there is only one Metadata Server and one web server. The secondary server functions only as a load balancing server running Stored Process Servers and Workspace Servers. Would this secondary server fail, the primary server keeps running without balancing the load to the secondary server. The difference between the proposed design and this alternative becomes visible in the scenario where the primary server fails: instead of starting up the Metadata Server and web server on the secondary machine (as in the proposed design) the *mount* of the primary server on which the SAS environment is installed is moved to the secondary server. The DNS is then updated and the name of the primary server is associated with the IP address of the secondary server. After that the primary server is reproduced on the hardware of the secondary server.

The most important advantage of this alternative is that it is easier to maintain. In the proposed design, the metadata is not always fully synchronised between the two environments so when the secondary server becomes the primary server, some changes in the metadata may be 'undone'. In this alternative the metadata is *moved* to the secondary server, so it doesn't change. Also having one web server makes maintenance easier. The one drawback is that if the installation of the primary server would become *corrupted*, a backup would have to be restored, whereas with the proposed design the secondary server could be activated. The advantages however outweigh this drawback as restoring a backup can be done relatively fast.

This alternative design was accepted by Astellas and selected as the definitive design.



CONCLUSION

Undertaking a project like this is a quite a challenge. It requires lots of concise planning and dedication, especially when the participating parties are spread over a number of geographical locations. At this point the project is still ongoing, but we're sure it will be a success.

ACKNOWLEDGEMENTS

Thanks to Harald van der Winden of Astellas, for granting us the permission to write and publish this paper, and to present it at PhUSE 2012. Also thanks to everyone in the project team for providing us with your continuous support.

CONTACT INFORMATION

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

Jules van der Zalm and Raymond Ebben
OCS Consulting
P.O. Box 3434
5203 DK 's-Hertogenbosch
+31 (0)73 523 6000
sasquestions@ocs-consulting.com
http://www.ocs-consulting.nl/