Paper: 173-2013

Modern SAS Programming: Using SAS Grid Manager and Enterprise Guide in a Global Pharmaceutical Environment

David Edwards, Director, Global Statistical Programming - Amgen, Inc.

Gregory S. Nelson, President and CEO - ThotWave Technologies, Cary, North Carolina

Susan Wang, Data Fabric, Program Manager - Amgen, Inc.

Abstract

Amgen, like most large biotechnology companies, uses SAS to support the drug discovery process. Used throughout the organization for data management, analytics and reporting activities, SAS use extends to research, operational and manufacturing units.   Like many organizations, Amgen has grown organically and their use of SAS is no exception. Equipped with a vision to fully leverage its global workforce and to maximize their IT investments, Amgen developed a SAS-based, research informatics infrastructure to deliver value around the globe.

This paper will highlight many of the aspects of the SAS Analytical Platform project including business justification, requirements, design, verification and validation (V&V) and production migration for over 1500 programmers and statisticians spread across three continents. We will highlight some of the challenges we faced and how these were overcome using improved processes, modern technologies such as SAS Grid Manager and Enterprise Guide, and the combined efforts of a global project team.

Having just finished the third and final wave of production migration, we will be able to talk in detail about what worked and what could have been improved – both from a project perspective as well as the technologies deployed.

Introduction

Business Context

Over the last decade Amgen's Clinical Development organization has changed significantly. For example, in 2003 Amgen had only two Clinical Development sites, one in Thousand Oaks, California, the other in Cambridge, UK. Although Amgen ran global Clinical Trials, the two Clinical Development sites had a large degree of operational independence, particularly with regard to the conduct of clinical data management and statistical programming activities. Initially, SAS Programmers, Data Managers and Statisticians were based almost exclusively onsite, and they analyzed/cleaned study data, stored locally, using UNIX SAS servers at the same location.

Fast forward to 2012 and we see a very different situation: Amgen now has five Clinical Development sites; has embraced remote working for permanent staff; and has augmented its workforce with a considerable number contingent workers and Functional Service Provider (FSP) staff, located primarily in North America, the UK and India. The operational independence previously enjoyed by Clinical Development sites no longer exists; Data Management and Statistical Programming activities are now performed by staff selected from a global pool of resources. Over last decade, Amgen has seen significant growth in the number and size of the clinical trials it conducts.
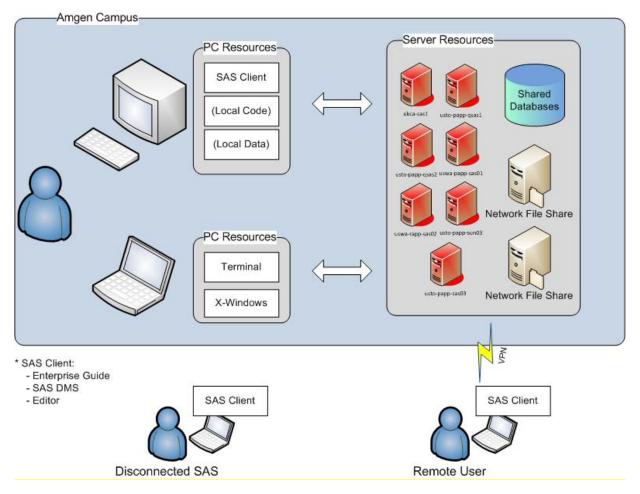
For a number of years, Amgen attempted to keep pace with these changes by expanding its existing SAS UNIX server-based infrastructure by adding new servers and network attached storage (NAS) at its Cambridge and Thousand Oaks locations. Over time it became evident that this strategy was not sustainable. It did not, for example, address the latency issues experienced by SAS users when using client applications such as Universal Viewer and Display Manager SAS (running on UNIX). Programmers were required to behave as *Grid Managers*, deciding which server to logon to each day, and as a result Amgen was unable to make effective use of its SAS resources, creating a situation where some servers were overloaded to the point of being unresponsive while others had idle capacity.

Like many other Biotechnology and Pharmaceutical companies Amgen has, in recent years, made a significant investment in the analysis of Real World data, adding the challenge of *big data* to its already overtaxed statistical computing environment.

This paper will describe how Amgen approached re-architecting its SAS environment using technologies such as SAS Grid and Citrix XenDesktop to address these issues and deliver a performant, globally accessible SAS Analytical Platform capable of adapting to the changing needs of the organization.

Current State Architecture (Legacy)

Prior to the deployment of Amgen's new SAS Analytical Platform, the *production* Clinical Development SAS environment at Amgen consisted of five Oracle Solaris servers - one in the UK and four in the US. The figure below depicts the current (legacy) state prior to the implementation of the modernized architecture.

Each server had Foundation installations of SAS v8.2, v9.1.3 and v9.2, with additional SAS BI Server components installed for versions 9.1.3 and 9.2 on all but one server. These servers were used primarily to clean and analyze clinical study data, stored as SAS datasets on a number of shared NetApp NAS devices located in the same geographic location as the SAS servers.

The majority of SAS jobs running on these SAS servers used locally attached SAN storage for their SASWORK library. However, jobs used to analyze large (up to 3TB) Real World data were forced to re-direct their SASWORK library to slower NAS storage, due to the limited amount SAN storage available on each host.

Inconsistencies in the build and configuration of these SAS servers caused a number of problems, for example:

- Some programs developed in the US would not run on the UK server, due to a more restrictive set of SAS modules being deployed in the UK.
- It was not possible to include all US servers when creating a SAS BI Load-Balance Workspace Server cluster, because one of the servers used Kerberos authentication while the others all authenticated users against UNIX LDAP.

This architecture supported a SAS user base of approximately 1500 and handled in excess of 45K SAS jobs each day, the vast majority of which executed in less than 3 minutes. While the majority of jobs were done quickly, the SAS environment still had to handle a much smaller number (less than 5%) of long running, resource-intensive

statistical analyses and data manipulation jobs involving datasets containing billions of observations.

Traditional Programming Methods at Amgen

SAS programming at Amgen makes heavy use of Foundation SAS - in particular BASE, STAT & Graph. All *production* SAS code was being executed, in batch, from the UNIX command prompt. While the approach to submitting *production* code was consistent between teams, the preferred approach to developing SAS code, running exploratory analysis and performing model development varied significantly. Some variation was based on personal or team preference and some was based on the availability of software such as PC SAS.

Many US-based biostatistical programmers and most statisticians used PC SAS as their preferred tool to develop code and to run exploratory analysis. In general, Amgen programmers in the UK and at FSP partners in India are not given access to PC SAS. As a result, those programmers who like to develop code using an interactive development environment (IDE), had to either run Display Manager SAS on UNIX or use Enterprise Guide (which prior to the roll out of the new SAS Analytical Platform was an unsupported SAS client at Amgen). At Amgen, old school programmers did not bother with an IDE at all, simply editing code with a text editor and submitting it in batch at the UNIX command line. Based on personal preference, programmers used a number of different text editors to support this style of code development including the Windows editor UltraEdit and UNIX editors such as Vi, Vim, Nano, Emacs and Gedit.

To connect to a UNIX SAS server, programmers and statisticians used a variety of different terminal emulation packages including Reflections X, PuTTY and NX Server. While Reflections X provides X11 support, technically allowing users to run Display Manger SAS on UNIX, network latency made this an impractical option for anyone not on the same local area network as the UNIX servers. NX Server was introduced to make Amgen's legacy federated SAS environment more accessible to users who accessed the environment over the wide area network. NX server represented a first step toward providing a virtual desktop infrastructure. For example, it made it practical for all users to run graphical UNIX applications such as Display Manager SAS. Amgen's NX implementation involved co-locating a cluster of Linux servers, running NX Server software in the same data center as the SAS servers. Using NX Client software, users were able to connect to a virtual Gnome desktop from which they could start an xterm session on one of the UNIX SAS servers, at the same location, avoiding the network latency experienced with other X11 capable terminal emulators such as Reflections X.

While the introduction of NX server improved the accessibility of Amgen's legacy SAS environment for remote workers, it did not provide users with performant access to other key desktop applications used by staff on a daily basis, such as Windows SAS clients like Universal Viewer and SAS Enterprise Guide, Microsoft Office and Microsoft SharePoint.

To use Amgen's legacy SAS environment, users first needed to know the geographic location where the data they were interested in was located. Data from trials conducted in the EU was typically stored at the Cambridge UK Clinical Development site, and data from trials conducted in the rest of the world was typically stored at the Thousand Oaks Clinical Development site in California. To submit *production* SAS code, users would then need to connect to a SAS server in the same location. In the UK, the choice of SAS server was straightforward, as there was only ever one; in the US, however, there were four servers to choose between. In an ideal world, users would behave like a Grid Manager selecting the machine with the lowest load each time they submitted a SAS program. But in reality most users tended to select the same server every time they logged on and choose another only if their first choice was under extreme load. As a result, over time the different servers had very different load characteristics, and Amgen did not make effective use of their SAS resources.

Challenges and Opportunities

In re-architecting the Clinical Development SAS environment, the project team had to overcome numerous challenges including:

1. The identification and remediation of impacted systems.
2. The migration of, over 60 TB, of data from shared storage locate on different continents.
3. Managing the transition of a user base of 1500 users in multiple business functions and geographies.
4. Managing the migration of data and users without impacting business continuity.
5. The development of a training plan and materials.
6. Mitigating the impact of SAS tools such as Enterprise Guide and SASGSUB on existing programming practices.

7. Designing a Grid, which balanced the number and size (number of CPUs, amount of memory etc.) of Grid nodes to optimally support the typical mix of jobs to be run on it.
8. Architecting and tuning persistent and scratch storage to meet SAS performance guidelines.

While a radical re-architecture of Amgen's legacy Clinical Development SAS environment posed numerous challenges, it also offered an opportunity to build a SAS environment that:

1. Was aligned with Amgen's global operating model.
2. Would provide sufficient capacity to meet current and projected workloads.
3. Would provide a platform that could be scaled to address changing business needs.

The use of Citrix XenDesktop technology to support virtual access to SAS desktop clients had the added advantage of providing remote staff with performant access to other key Amgen systems such as its Learning Management System, various document management systems, and internal Microsoft SharePoint sites.
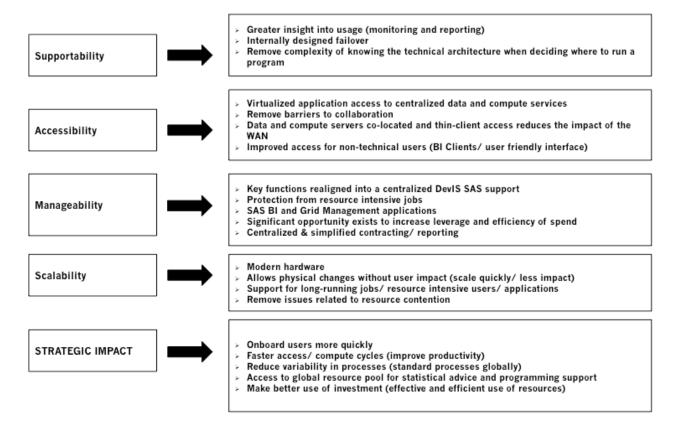

New SAS Analytical Platform

Motivation

BUSINESS AND TECHNICAL DRIVERS

The key business drivers for re-architecting Amgen's Clinical Development SAS environment included:

1. A need to replace end-of-life SAS servers and slow NAS.
2. An urgent need to address the inability of the SAS environment to support current and future projected peaks in workload without significant business disruption.
3. A desire to create a SAS platform that:
    a. Could be scaled to meet changing business demands in a manner that is transparent to end users.
    b. Is less complex from an end user perspective.
    c. Meets the needs of a globally distributed workforce.
    d. Provides a standard supported toolset for the development and execution of SAS code.
    e. Allows effective prioritization of critical work.
    f. Can effectively utilize all the available SAS capacity.
    g. Provides a truly homogenous SAS environment.
    h. Provides high-speed data storage for both persistent and scratch storage.
    i. Limits the ability of a user to overwhelm servers by submitting too many concurrent jobs.
    j. Provides scratch storage with sufficient capacity to support jobs that operate on Big Data.
    k. Is supported around the clock by a team of dedicate SAS Administrators.
    l. Reduces the organization's reliance on PC SAS.
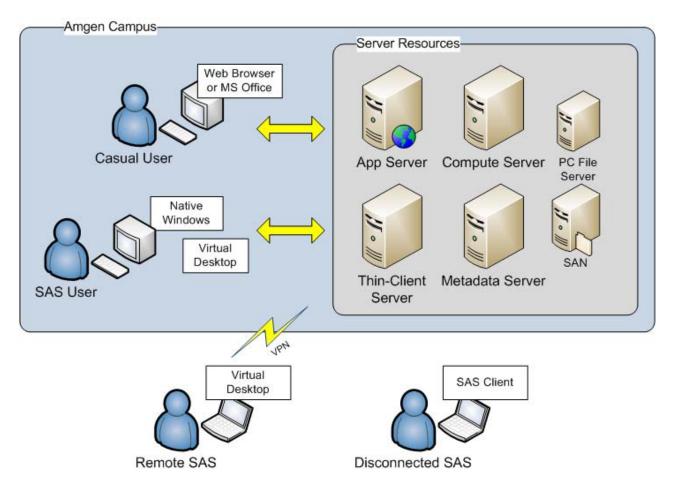
Amgen Benefits

Supportability	➢ Greater insight into usage (monitoring and reporting) ➢ Internally designed failover ➢ Remove complexity of knowing the technical architecture when deciding where to run a program
Accessibility	➢ Virtualized application access to centralized data and compute services ➢ Remove barriers to collaboration ➢ Data and compute servers co-located and thin-client access reduces the impact of the WAN ➢ Improved access for non-technical users (BI Clients/ user friendly interface)
Manageability	➢ Key functions realigned into a centralized DevIS SAS support ➢ Protection from resource intensive jobs ➢ SAS BI and Grid Management applications ➢ Significant opportunity exists to increase leverage and efficiency of spend ➢ Centralized & simplified contracting/ reporting
Scalability	➢ Modern hardware ➢ Allows physical changes without user impact (scale quickly/ less impact) ➢ Support for long-running jobs/ resource intensive users/ applications ➢ Remove issues related to resource contention
STRATEGIC IMPACT	➢ Onboard users more quickly ➢ Faster access/ compute cycles (improve productivity) ➢ Reduce variability in processes (standard processes globally) ➢ Access to global resource pool for statistical advice and programming support ➢ Make better use of investment (effective and efficient use of resources)

While the business drivers above sufficiently characterize the real desire for change, the way these goals translated in practice really proved challenging. As mentioned above, the access to SAS was delivered to users across the globe through UNIX servers located in two different data centers, which provided access to users depending on the country in which they originate (UK and US). Each time someone logged on to one of the servers, they had to decide which server to use for their programming that day.

The SAS products licensed across the various servers differ, which made running the same analyses on the one SAS server more challenging for users when sharing SAS code with their global counterparts. Furthermore, one of the SAS servers may have become over-loaded at the same time another server may have plenty of capacity. Users accessing SAS on the UNIX servers encountered code that just wouldn't run or took far longer than it should have.

In summary, what was required was a modern, scalable, and consistent environment in which to analyze data, that also supported a large, geographically disbursed workforce without location-based performance impact (or license restrictions). We really wanted an environment that utilized SAS resources more effectively and gave us the ability to scale as the business grew.

The diagram above highlights the logical architecture that was deployed. Note that we ended up not installing the SAS PC Files Server because we could achieve the required functionality through adaptations in SAS coding practices. Furthermore, this diagram does not depict the storage architecture, which was a significant part of this new architecture, or the implications for high performance storage (high throughput) when considering SAS Grid implementations.

Project Management Approach

SCOPE

The re-architecture of Amgen's Clinical Development production SAS environment was one aspect of a larger effort (the Data Fabric program) at Amgen, designed to address not only the issues with our legacy SAS environment discussed previously, but also the management (based on the principle of least privilege) of access to data stored on the various file shares accessed by the SAS environment. While the details of many other aspects of the program are not covered in any detail in this paper, you will see references to them later when we discuss the rationale for implementing a Phased Rollout strategy for the SAS Analytical Platform. Each of the major deliverables discussed below was managed by an IS Project Lead and the program as a whole was overseen by an IS Program Manager and a business Program Champion.

DELIVERABLES

The key deliverables of the Data Fabric Program were:

1. SAS Grid Manager.
2. Dedicated high speed Network Attached Persistent Storage (configured to use Windows NTFS access controls).
3. A high speed SAN storage array (providing SASWORK storage with sufficient capacity to support the analysis of large Real World datasets).
4. The File Access and Security Tool (FAST - a custom Web based application for the management of file

access controls in accordance with the principle of least privilege).
5. A Virtual Desktop Environment (implemented using Citrix XenDesktop).
6. Backup and Disaster recover solutions for the deliverables above.
7. Training and Support for the new environment.
8. A retirement and decommissioning plan for the legacy SAS environment and associated storage.

STAKEHOLDER ENGAGEMENT

As a core philosophy in delivering value to all stakeholders, we wanted to ensure that there was close collaboration between the information technology departments (infrastructure, storage, security, data center, business continuity, program management) and the business functions, which were as diverse as the IT functions. These included traditional clinical research departments (statistical programming), data management, outcomes and health economics (think big data) and early discovery research – each with very different usage patterns and functional requirements.

To that end, we focused heavily on stakeholder engagement through a number of key elements that proved successful. These included:

- Stakeholder representation at various levels
  - o Executive Steering Committee
  - o Program Management Review
  - o Change Management Board
  - o Business Unit representatives in testing and go-live planning
  - o Key Opinion Leaders in each business unit and site serving as a "go-to" person
- A comprehensive Readiness program in preparation for implementation, with
  - o Learning Management System training modules
  - o Interactive PDFs which navigated users through the various resources
  - o SharePoint site which contained FAQs, KnowledgeBase Articles and Late-breaking news
- Governance and Change Management Plan to facilitate large scale decision making and escalation processes

A key measure of success for the project was how well stakeholders were engaged in the transition of their business functions to the new platform. Further confirmation of a successful implementation was that we were able to shut down our hypercare (go-live support desk) a month early, and that there were no major defects and no downtime.

PILOTS

Because of the potential to change the daily routine of hundreds of statisticians and programmers around the world, and the pure magnitude of potential risk these changes had on daily operations, we decided to test the various components of the architecture through a series of pilots. Four pilots were planned with these goals:

- Prepare users for the upcoming migration and engage the user community in open dialogue.
- Minimize and mitigate risks that might be associated with the proposed architecture.
- Allow users to gain understanding of how the new SAS platform and storage architecture will function.
- Contribute to migration planning.
- Clarify what needs to be remediated and/or formally tested.
- Allow the business to make high-level process decisions so that the details of the final architecture can be determined and included in the migration planning activities.
- Demonstrate that critical high level business requirements can be met, as expected, from the proposed architecture.

We defined each of the pilots with very specific objectives:

- Pilot I: 'Hello World' messages to verify access for all personnel and initiate support for the Virtual Desktop Infrastructure (VDI).
- Pilot II: Introduction to the new environment (Using SAS in batch and Enterprise Guide).
- Pilot III: Identification of any impacts and implications of the changes.
- Pilot IV: Testing of final migration procedures to demonstrate that issues previously identified were

resolved.

The purpose of each pilot was to engage the business community in understanding how the new SAS platform and storage architecture would function so that we could minimize any risks and/or issues that were unforeseen. Additionally, this was also an opportunity for the SAS administrators to see how the system performed under some load so that we could evaluate the monitoring tools and begin development of the operations manual.

Each testing phase had specific components such as entry criteria, requirements, deliverables and exit criteria. The SAS Platform testing team met regularly throughout the testing phases to ensure that all concerns were being addressed, and we exercised the administrative and operational use cases that were critically important to getting the new SAS Administrative team up to speed.

PHASED ROLLOUTS

Because the SAS Platform reflected only a small change in the larger context of the Data Fabric program, we migrated users in three waves. There were two primary drivers of this decision: the volume of data that was to be migrated and the number of users to be impacted. The former was important insofar as we were moving from an environment that used multiple shared NetApp NAS devices for storage and we were upgrading the security infrastructure to a modernized version that necessitated security updates to the migrated files. That is, we moved data from a number of shared geographically disbursed NAS devices configured to secure files and directories using UNIX POSIX permissions to a single dedicated NAS device configured to secure data using Windows NTFS permissions. In terms of the number of users that could potentially be impacted, we wanted to ensure that there was sufficient support staff to provide hyper care for users impacted by each migration wave.

As it turned out, the data migration was the critical time-sensitive issue and the hyper care (support staff on call) was actually shut down early because there were no major issues after the first few days. The issues that did arise during the first full week of migration had to do with the virtualization environment, and there were no defects associated with the operation of SAS.
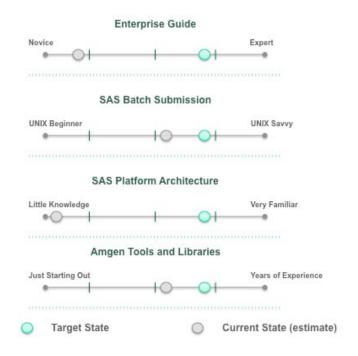
TRAINING

Early on, we made the determination that we needed a training strategy that would help individual users adapt to the new environment in a way that would meet their own personal learning styles. To that end, we developed and delivered training and instruction materials through the following modes:

- Instructor led training delivered at the major development sites (UK, US) and Web.
- Project team meetings to review major changes.
- Mandatory computer based training (through the Amgen Learning Management System).
- Instructional guides, How-to Demonstrations, Frequently Asked Questions (FAQs), and a Dynamic PDF accessible through the SAS Support site (SharePoint).

Philosophically, our intention was not to train users on SAS per se, but to ensure that they could adapt to the new Enterprise Guide and SAS Grid manager client (SASGSUB) tools, and to provide them with a basic understanding of the architecture. One of the key decisions early on was to avoid having them experience SAS in batch differently from what they were used to – that is, when they typed "sas" on the command line, it should behave in the way that they were used to. So we produced a "wrapper" for the SAS Grid Submission Utility Client (SASGSUB) and called it GSAS. This solution minimized much of the confusion about using this new tool, and ensured that legacy code and scheduled jobs could be supported with minimal change.

The training plan was highly effective, both in securing user buy-in, and also in building a foundation for new users who may be familiar to pharmaceutical programming but not SAS Grid or Enterprise Guide. The figure below highlights the skill updates required to use the new environment.

SUPPORT, MAINTENANCE AND OPERATIONS

As outlined above, we had what turned out to be a very strong transition strategy that was available to support the users, and a variety of support mechanisms were also put in place. The SharePoint site served as the primary self-service tool. The major components of this internal web site included:

- **Dynamic PDF** – this interactive PDF contained demonstration movies, animations, links to documentation and resources across the Data Fabric program; it was the one-stop-shop for all user information.
- **Frequently Asked Questions** – over the course of the project, we catalogued questions from Pilot users, Help Desk calls and emails; the answers were captured and made available as part of the user's knowledge base of information.
- **Amgen IT Help Desk** – all calls and support tickets acquired through the intranet were routed through a single Help Desk; since the project affected security, access, virtual desktop, storage and of course SAS, we needed a way to quickly route information to the right team. All second and third level support calls were managed by the appropriate team.
- **Documentation** – consisted of user manuals, updated standard operating procedures and even how-to video demonstrations; all of these documents were easy to find, accessed through the Dynamic PDF.

The SAS Administration Team was responsible for three primary roles that included:

Monitor:

- Provide break/ fix support
- Monitor server status and provide early detection of potential issues
- Monitor disk usage and I/O performance
- Monitor grid performance and participate in capacity planning
- Manage SAS audit and performance reporting framework
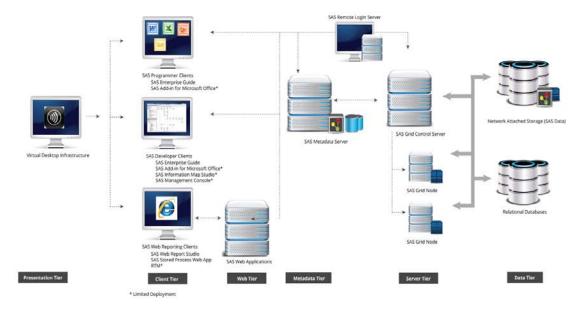- Manage Service Level Agreements (SLAs)

Maintain:

- Apply SAS and Platform LSF hotfixes
- Recommend operating system patches and security updates
- Maintain licenses
- Manage service requests

Support:

- Manage job queues and priorities
- Support client Applications
- Review/optimize code
- Manage roles

Technical Architecture

Within the global SAS environment at Amgen, a number of components make up the architecture. This is logically depicted in the figure below.

THE PRESENTATION TIER

The presentation tier is the environment used to access the SAS Platform. For most users, this is a Citrix XenDesktop Viewer launched from the web browser (or Citrix Receiver application) running on the local PC. We call this the presentation tier because the only thing that is running in the local PC is the Citrix Viewer application that accesses applications streamed from the VDI in the Washington data center.

VIRTUAL DESKTOP INFRASTRUCTURE

The deployment of a Citrix XenDesktop Infrastructure in the same data center as the SAS GRID was critical to the overall success of the project. Without it, users geographically removed from the SAS Grid would have experienced unacceptable network latency when using a number of Windows desktop clients to interact with the SAS GRID or directly with the data on which it operates, e.g. SAS clients such as Enterprise Guide Universal Viewer. The Virtual Desktop environment also allows users to access the new SAS Analytical Platform from a multitude of devices, including IOS and Android Tablets, and Windows, OS X and Linux PCs. In addition to being able to access the Virtual Desktop environment from the Amgen network, users can also access it over the Internet even when limited to a 3G wireless connection.

THE CLIENT TIER

Clients are the software applications that are used to access information. These are provided through the VDI environment as if the user is running a Windows desktop environment locally. Applications are installed (or streamed) to the virtual desktop. Users interact with an image on a local machine, and the client is actually running on the Virtual Desktop. The client only runs on the users local machine if it is installed on the local machine and the user chooses not to use VDI to access the SAS Analytical Platform. The clients available in VDI include SAS Enterprise Guide, Microsoft Office, and other tools such as UltraEdit. All the client applications users need to interact with the SAS Analytical Platform are part of the virtual desktop image or stream to it using Microsoft App-

V technology.

THE APPLICATION TIER

This logical tier includes SAS servers, a remote login server, the SAS metadata server and mid-tier web application servers. SAS can be accessed either through interactive clients such as Enterprise Guide, or in batch with a command line tool called GSAS. GSAS is accessed through the remote login server through PuTTY or Reflections X. (As discussed earlier, GSAS is an Amgen developed wrapper for the SASGSUB utility used to submit code in batch to the SAS GRID.)

THE DATA TIER

A user can read and write data to the "data tier" from SAS, and can also access both native SAS data sets (stored on filers) or those found in relational databases such as RAVE or Oracle Clinical.

SECURITY ACCESS AND CONTROL

Data stored on file shares accessed by the SAS GRID were secured using granular Windows NTFS permissions. Data ownership - the definition and application of consistent file system security models, the granting and revoking of user access, and the monitoring of *Locked* (read-only) areas of the file system for unexpected events was all managed by FAST. File security was configured in such a way that users could not modify the permissions applied to files and directories directly. All permission changes had to be effected by Business Data Owners using FAST.

User access to the SAS GRID itself was limited to users with a SAS Metadata Identity (define in the SAS Metadata Repository). Access to the Remote Login Server and the SAS GRID nodes was limited to users who were members of a UNIX Netgroup. Users were prevented from logging on interactively to the SAS GRID nodes (which would have allowed them to bypass the GRID manager) by setting their shell on these hosts to /bin/false.

To ensure that Amgen did not exceed the number of licenses purchased for XenDesktop and some of the client applications available on it (e.g. SAS EG and Universal Viewer), it was also necessary to limit user access to the Virtual Desktop environment. This was achieved by defining groups in Active Directory (AD) and using these AD groups to limit user access to the Virtual Desktop itself and to individual client applications when required.

As noted, a number of different techniques were employed to control access to the various components of the SAS Analytical platform. To avoid the complexity of manually creating Metadata Identities and having to manually manage the membership groups in both AD and UNIX LDAP, we established a set of business rules, based on user attributes (such as department) available in Amgen's Corporate Directory, to define who should have a Metadata Identity and who should be a member of the relevant AD and UNIX LDAP groups. We then created a simple web application that allowed us to manage the definition and modification of these business rules and to automate management of Metadata Identities and group membership.

SAS PLATFORM

SAS Grid Manager

In another paper, Ladd and Fenton describe the SAS Grid Manager and its relevance in modern SAS architectures. While we won't repeat that information here, we do want to describe some of the major elements of the SAS Grid Manager. One on end, we have Enterprise Guide, where the user submits his or her job. Once submitted, the SAS Grid Manager does all of the work to decide where and how that job gets distributed to the grid for processing. The Grid Manager assigns a job ID to the submitted program, and sends it to the Normal queue (by default.) The job is evaluated and scheduled based on job attributes, then dispatched to one of the SAS servers, and the Grid node processes the job and assigns a status of complete.

In a large, multi-user environment like Amgen, many jobs are submitted to the grid as more and more users are on the system. The Grid manages the work through Workload Balancing. Each new job is evaluated and dispatched to a SAS grid node based on the job characteristics and the availability of compute resources. If there is a point at which the servers can't process all the jobs, the Grid Manager holds them until resources are freed up.

Sometimes jobs will be prioritized based on criticality. For example, when Amgen needs to respond to a regulatory agency, those jobs will have a higher priority, and the user will specify the CRITICAL queue. Similarly, if a job is not critical, the user can request a night queue, which waits until a specified time (such as 10 pm) to dispatch the job

for execution on one of the SAS grid nodes.

The SAS Grid environment has a number of capabilities - it gives Amgen:

- Workload Balancing that enables us to manage the workload of hundreds of users.
- More control through job prioritization, and the opportunity to make the most effective use of the SAS resources we have.
- Enterprise scheduling capabilities.
- A high level of availability.
- The ability to run large and complex analytics.
- Ability to "scale out" by adding compute resources as needed with minimal disruption.
- Significant reduction in job execution times by Grid Enabling code that previously executed in a serial fashion on a single server.

SAS Business Intelligence

While not central to the discussion of the SAS architecture at Amgen, we do have SAS Business Intelligence that is used primarily for distributing reports via Web Report Studio and the SAS Add-in for Microsoft Office. Furthermore, there are a number of custom applications at Amgen that use the SAS Integration Technologies application programming interface (API) to call SAS from non-SAS clients such as Java. These clients all utilize the SAS Grid Manager to ensure that load is being managed by the SAS Grid Manager. We currently don't allow Grid sign-ons from EG or Java clients, so the Grid Manager cannot manage these sessions once they are started; but our SAS BI Workspace Server clusters use the Grid Load Balancing Algorithm to determine which node to start a SAS session on based on current load.

Interactive Programming (with Enterprise Guide)

Early on, we made the decision that we would support SAS Enterprise Guide as the primary SAS-based client for interactive use. That is, we explicitly deprecated support for SAS Display Manager via X-windows. It was not consistent with the overall strategy of our grid solution and by-passed the grid client submission framework. On the other hand, Enterprise Guide fit nicely within the grid strategy and enabled users to submit interactive jobs directly to the grid without additional work or intervention on their part.

Batch Programming (using SAS Grid Submission Client Utility)

The second client interface that users had at their disposal was essentially SAS in batch mode on UNIX. They were accustomed to using SAS via the command line/ shell in UNIX, so we created a wrapper in Java, which took their command and sent it off to SAS Grid Manager via the SAS Grid Submission Client Utility. This enabled us to completely mimic what they were used to and allowed for grid behavior such as checking the status of a job and terminating jobs, and for using this in UNIX shell programming or third party languages such as Java or Perl.

Lessons Learned

In any project, hindsight is 20/20. After the fact it is easy to identify things that should have been approached differently or aspects of the project that should have had greater focus earlier in the project life cycle. Two of the biggest lessons we learned during our Grid implementation were that is important to:

- Determine the optimal size and configuration of the Grid
- Understand the importance of storage performance

Accurately characterizing the different types and number of SAS jobs that your organization routinely runs is extremely important when designing the architecture of the SAS Grid environment for your organization. When Amgen first worked with SAS to design our SAS Grid architecture we designed a Grid with just four large compute nodes. When we piloted this architecture with a representative set of test jobs, we quickly discovered that the Grid Manager was unable to dispatch enough jobs in each dispatch window to keep the four large nodes busy. As a result of this early pilot / proof of concept, we redesigned our Grid to use nine smaller compute nodes allowing the Grid Manager to dispatch a greater number of SAS jobs in each dispatch Window and improve the utilization of the SAS compute capacity we had licensed.

The other lesson we learned during our implementation was the importance of designing and deploying a storage architecture that is able to deliver the <u>sustained</u> I/O throughput recommended by SAS. We did not run a robust set of IO tests that would have demonstrated whether our storage architecture could deliver the recommended sustained IO throughput under load until late in the project life cycle, and as a result had to make significant tweaks to our storage architecture late in the day.

Summary

 As we have shown, the overall architecture was designed to address a number of challenges, and it has given Amgen a number of key benefits, including:

- The ability to scale, so that we can easily to accommodate changing business needs.
- A guarantee of a highly available environment.
- Global support by dedicated IS SAS administrators, that allows us to make optimal use of available SAS resources.
- The opportunity to support workload prioritization.
- Resiliency – individual users are prevented from inadvertently impacting each other; and
- Reduced reliance on PC SAS.

References and Recommended Reading

Fenton, Ladd, and Spakes, *"A Practical Approach to Re-Architecting a SAS® Deployment"*, Invited paper presented at the SAS Global Forum, 2011.

Nelson, G. and Loether, J. 2007 *"Architecting SAS in a Modern World: Best Practices for Design, Configuration and Management of SAS® 9".*   Invited paper presented at the SAS Global Forum. Orlando, Fl. April, 2007.