

# Lessons Learned from our SAS Modernization

Craig Kasper

Winnipeg SAS User Group

October 2017

# A Little Background

- Our project was part modernization and part consolidation
  - Multiple user bases, multiple separate licenses
  - Our existing installations did not require significant investment at any one time, but resulted in a limited environment that was getting further and further from current technology.
  - Before the project, our SAS environments lagged significantly behind the capabilities of more modern installations.
- We could have put off the modernizing...
  - ... but it would only have made modernizing later even more difficult.

# A Little More Background

- The project was a modernization, not a transition from other analytics software.
  - SAS usage is well-established within our analyst base (many of our programmers have 5 - 10 years or more of experience programming in SAS).
  - SAS usage is well-established within the area of health analytics (MCHP, CancerCare, national organizations such as CIHI all have a significant existing SAS base, including shareable program code).
  - We have a built-up capacity using SAS for analytical and data processing which allows us to do things easily in SAS which other conventional off-the-shelf BI solutions would struggle with.
- Experiences (and takeaways) may be somewhat different for starting from scratch.

# Lesson #1 – Give yourself as much lead time as you can

- Even in a typical small-to-medium-scale deployment, there can be a large amount of data and programs to move.
- SAS may be being used to interact with really old technology (legacy software, DDE) or with really new technology. In both cases, extra lead time is helpful.
  - For really old technology, alternatives may need to be evaluated and tested.
  - For really new technology, additional configuration may be required for both SAS and the other technology.
- This is especially true if you need to maintain existing core service levels during the transition.

# Lesson #2 – You have more SAS code than you think...

- ... and it is likely less organized than you think...
- ... and the same likely goes for your datasets...
- ... but everyone who undergoes a major migration deals with this.
- Recommendation: classify your existing program code and your data holdings at a high level:
  - In current use versus retained for archival purposes only
  - By client (or client base) being served
  - By frequency of use or urgency level
  - By complexity
  - By analyst responsible
- Each of these classifications may indicate possible migration strategies.

## Lesson #3 – Getting organized is a must

- Modern SAS environments are highly structured. Operating them and administrating them well and wisely will require a foundation of solid organization.
- Failing to get your SAS environment organized and keep it organized will have operational risk associated with it.
  - If your modernization is already looking more difficult than you expected, this is one possible reason.
  - Conversely, if you are expecting to modernize, but not right away, improving the organization of code and data you hold is a step which will benefit you both now and later.

# Lesson #4 – Know your planning priorities

- Security Model
- Folder Structure (new environment) and Metadata Folder Structure.
- Inventory of programs and data mapped to folder structure.
  - Not all programs and data may need to be migrated immediately (or at all) but you will need to know which programs and data are most urgently needed and which ones will require the most time, expertise, or effort to move.

## ... Uhhh, “Metadata”?

- “Metadata” is “data about data”. It provides a software-friendly profile of important things about your data.
- The current standard client-server configurations of SAS receive requests to run programs and processes within the standard server environment.
  - These requests come from multiple sources (SAS Studio, Enterprise Guide, Office Add-in, Visual Analytics). Some of these are wizard-driven and will need contextual information about your data before they actually submit a request.
  - Additionally, these sources need to respect permissions and other policies you have regarding your data.
  - And all of this information needs to be findable and accessible by the programs making the requests.
- Metadata on a SAS server has a folder structure for the same reason that computer file systems do – because grouping and categorizing things is effective and efficient, but still human-friendly.

# Priorities: #1 – Security Model

- The security model for your deployment exists provides the answers to the following two questions:
  - Who gets to access which data and analytics components?
  - Who gets to access which software features?
- SAS's metadata server uses group membership to manage access to the data and analytics components inside its folder structure.
  - Groups will normally be specified based on functional business units, current processes, and/or approved access to data.
- Access to SAS software features within a metadata server uses roles.
  - For example, a person in a developer role may be given access to more features than a report viewer, but fewer features than an administrator.

# Priorities #2 – Folder Structure

- At a high level, folder structure should line up with how data is accessible to users, based on the security model.
  - In order to simplify data management, data with more restricted access should be sited deeper within the folder system. Data with less restricted access should be sited closer to the top-level folder.
    - (This is to avoid siting less-restricted data where its access will be unduly restricted due to a higher-level folder having more restrictive access control applied to it.)
- At a technical level, both metadata folder structure and data storage folder structures should be planned for.
  - For easiest maintenance, metadata folder structure and data storage folder structure should line up, even if some differences exist.

# Priorities #3 – Inventory of programs and data

- The larger your existing system and/or user base, and the more options they have, the more important this is...
  - ... and the more complicated it may be to collect this inventory.
  - We crawled our primary network drive to identify the location of all SAS files on the drive.
  - The inventory we have developed will allow us to migrate or archive the programs on the network drive in a controlled fashion as analyst capacity permits.

# Priorities #4 – IDing high-priority migrations

- Data and programs involved in mission-critical tasks should be migrated first.
- High-use data sets and programs should also be migrated early on if a continuous level of service is required.
- You will be making things much easier on yourself if you make major changes (other than migration-related changes) either well before or well after the migration period, if possible.

# Lesson #5 – Things don't always go according to plan

- In our case:
  - Mid-project role turnover
  - Multiple unavoidable delays
  - Timing issues
- You will want an aggressive plan to implement your modernization, to reduce the amount of time you are devoting extra time and resources to the transition.
  - ... just be ready to cope with the unexpected!

# Lesson #6: Implementation finishes before migration does

- Completed implementation of a modernized environment means that you can do everything you need to within the new environment.
  - This is not the same as your programs and processes (which do these things) being ready to run.
  - Data and data connections will be present, but programs will need to be adapted to access them.
- Most programs which do not use complex or obsolete techniques will be relatively straightforward to move over to the implemented new system, once users have become acquainted with the new system.
- The number of programs and amount of data to move over will likely extend the migration period beyond the implementation period (depending on how much there is to transition).

# Lessons #7: There is always more to learn

- The new system is up and running, and performing well.
- We still have a lot of work to do to fully adapt to using the system, and to use all of its capabilities (Add-in for Microsoft Office, Visual Analytics).
- As we use the new system, we will continue to learn more about how to best use the system to keep up with our growing needs and the growing needs of the stakeholders we serve.

Questions?