

Paper 215-31

## Banking on the Future: A Metadata Server Survival Guide

Peter Eberhardt, Fernwood Consulting Group Inc, Toronto, ON  
Andrew Farrer, Canadian Imperial Bank of Commerce, Toronto, ON  
Zulfiqar Habib, Habib Consultants Limited, Toronto, ON

### ABSTRACT

The strength of the SAS® Business Intelligence (BI) Architecture revolves around a well thought out metadata server. Unfortunately, this crucial piece needs to be designed and configured very early in the v9 learning curve; and in your haste to get your metadata server running you can make some poor choices. Sadly these poor choices early on can lead to long-term pain. This paper will discuss some of the trials and tribulations we experienced configuring our metadata server as part of a major compliance project - Basel II.

### INTRODUCTION

In the modern banking environment there is the business need for nimbleness to seize new opportunities competing with the regulatory requirements for security, accuracy and accountability. On a global level, the pendulum has shifted the emphasis towards the regulatory requirements as is evidenced in the Basel II compliance to which all major banks must adhere. However, a well thought out system that addresses and exceeds the regulatory quagmire will leave a bank not only a shining star in the eyes of the regulators, but more importantly it will leave the bank in a position to be aggressive in the new marketplace knowing it can act quickly on opportunities without worrying about the ramifications of non-compliance.

To meet the regulatory requirements a bank needs software that can efficiently and effectively manage and analyse mountains of data, but that is not enough. The bank needs software that can provide centralized management over access to data and processes, but that is still not enough. The bank needs software that will allow seamless access to disparate data sources through a common interface, but that is still not enough. All of this is important for success; however to be truly successful the bank needs software that not only provides all of the above, but also makes all of these data and processes available to the appropriate people using the appropriate tools; tools that range from advanced modeling and statistical methods to simple point-and-click environments. At the Canadian Imperial Bank of Commerce (CIBC) it was felt that SAS provided the software to do it all. In this paper we will discuss how the SAS system is being implemented, with an emphasis on the heart of the implementation - the metadata server. To set the stage for this discussion we will first review Basel II and the CIBC environment.

### BASEL II

The Swiss have given us chocolate, watches and the ubiquitous 'Swiss Army Knife'. Now, from the town of Basel we have the *International Convergence of Capital Measurement and Capital Standards - A Revised Framework*, or BASEL II as it is commonly called. The BASEL II accord defines international standards for measuring the adequacy of a bank's capital. It was created to promote greater consistency in the way banks and banking regulators approach risk management across national borders.

There are three main pillars to the BASEL II accord; these are:

#### PILLAR 1

The first pillar provides improved risk sensitivity in the way that capital requirements are calculated in three of the components of risk that a bank faces:

- credit risk,
- operational risk

- market risk

In turn, each of these components can be calculated in between two or three ways of varying sophistication.

Technical terms in the more sophisticated measures of credit risk include EL (Expected Loss) whose components are PD (Probability of Default), LGD (Loss Given Default), and EAD (Exposure At Default). Calculation of these components requires advanced data collection and sophisticated risk management techniques.

Pillar I sets out minimum regulatory capital requirements—the amount of capital banks must hold against risks. It retains Basel I's minimum requirement of 8 percent of capital-to-risk-weighted assets.

## **PILLAR 2**

The second pillar deals with the regulatory response to the first pillar, giving Bank regulators much-improved 'tools' over those available to them under Basel I. It also provides a framework for dealing with all the other risks that a bank faces, such as name risk, liquidity risk and legal risk, which the accord combines under the title of residual risk.

## **PILLAR 3**

The third pillar greatly increases the disclosures that the bank must make. This is designed to allow the market to have a better picture of the overall risk position of the bank and to allow the counterparties of the bank price and deals appropriately.

As can be seen from this brief overview, the BASEL II accord forces some stringent requirements upon the bank. Adding to these requirements is the Sarbanes-Oxley Act.

## **SARBANES-OXLEY ACT**

The Sarbanes-Oxley Act makes corporate executives explicitly responsible for establishing, evaluating and monitoring the effectiveness of internal control over financial reporting. For most organizations, the role of IT will be crucial to achieving these objectives. Some of the key areas of responsibility for IT include:

Understanding the organization's internal control program and its financial reporting process.

Mapping the IT systems that support internal control and the financial reporting process to the financial statements.

Identifying risks related to these IT systems.

Designing and implementing controls designed to mitigate the identified risks and monitoring them for continued effectiveness.

Documenting and testing IT controls.

Ensuring that IT controls are updated and changed, as necessary, to correspond with changes in internal control or financial reporting processes.

Monitoring IT controls for effective operation over time.

Participation by IT in the Sarbanes-Oxley project management office.

To comply with Sarbanes-Oxley, it is critical that organizations must understand how the financial reporting process works and must be able to identify the areas where technology plays a critical part. In considering which controls to include in the program, organizations should recognize that IT controls could have a direct or indirect impact on the financial reporting process. For instance, IT application controls that ensure completeness of transactions can be directly related to financial assertions. Access controls, on the other hand, exist within these applications or within their supporting systems, such as databases, networks and operating systems, are equally important, but do not directly align to a financial assertion. Application controls are generally aligned with a business process that gives rise to financial reports. While there are many IT systems operating within an organization, Sarbanes-Oxley compliance only focuses on those that are associated with a significant account or related business process.

The Basel II accord commits all financial institutions worldwide to adopt common standards for calculating and reporting capital reserves. All financial products need to undergo a regular risk calculation, which is aggregated and reported to banking regulators.

Sorbannes-Oxley requirements govern a minimum level of data protection and reporting. The standards of risk management generally supercede the minimum Sox standards.

## **CIBC**

Canadian Imperial Bank of Commerce (CIBC) is one of the 'Big Five' Canadian Banks. The bank has a heritage of sophisticated data analysis as well as financial and economic modeling, driven primarily through the Marketing and Finance departments; these business driven initiatives have leveraged the traditional SAS analytics. With a recognized need for a business analysis and modeling environment that is closely aligned with the strategic technical architecture, CIBC became the first Canadian financial installation to adopt the SAS 9 Business Intelligence (BI) architecture.

Decision Solutions is a department within the Retail Risk Management division that is responsible for delivering Basel II regulatory capital reports. These reports confirm that the bank has sufficient reserves (within consistent risk guidelines) to cover commitments. Although there had been internal efforts to address the complex project requirements, the SAS Risk Dimensions product was the closest to an 'off the shelf' solution.

The combination of Basel II compliance deadline, existing strong SAS modeling culture and SAS Risk Dimensions resulted in the fast track adoption of SAS 9.

There is also a less time-critical, regulatory requirement to protect customer and commercially sensitive data. Administering all data access through SAS Metadata servers ensures that appropriate controls are possible without significant re-engineering and loss of flexibility.

## **COMPUTING ENVIRONMENT**

Decision Solutions has a mixture of SAS applications and SQL Server databases on 6 Windows 2000/2003 servers of varying configuration. Storage is mixed between local disk and a 2Tb Storage Array. Non-SAS data occupies another 6Tb. All SAS Servers have SASv913 and Service Pack 3 installed with these features configured but not necessarily active or available:

- Metadata server
- Workspace server
- Connect server
- Stored Process server
- SAS/Share server
- OLAP server
- Data Step Batch server

By default, the SAS Management Console is also installed but only used via a remote desktop connection.

## **SAS BI ARCHITECTURE**

The SAS Intelligence Platform architecture is designed to efficiently access large amounts of data, while simultaneously providing timely intelligence to a large number of users. The platform uses an n-tier architecture that enables you to distribute functionality across computer resources, so that each type of work is performed by the resources that are best suited to the job.

You can easily scale the architecture to meet the demands of your workload. For a large company, the tiers can be installed across a multitude of machines with different operating systems. For prototyping, demonstrations, or very small enterprises, all of the tiers can be installed on a single machine.

The architecture consists of the following four tiers:

#### **DATA SOURCES**

Data sources store your enterprise data. All of your existing data assets can be used, whether your data is stored in relational database management systems, SAS tables, or ERP system tables.

#### **SAS SERVERS**

SAS servers perform SAS processing on your enterprise data. Several types of SAS servers are available to handle different workload types and processing intensities. The software distributes processing loads among server resources so that multiple client requests for information can be met without delay.

#### **MIDDLE TIER**

The middle tier enables intelligence data and functionality to be surfaced to users via a Web browser. This tier provides Web-based interfaces for report creation and information distribution, while passing analysis and processing requests to the SAS servers. Not currently in use by Decision Solutions.

#### **CLIENTS**

The client tier provides users with desktop access to intelligence data and functionality through easy-to-use interfaces. For most information consumers, reporting and analysis tasks can be performed with just a Web browser. For more advanced design and analysis tasks, SAS client software is installed on users' desktops.

#### **USER ENVIRONMENT**

All our users were confined to three main SAS clients:

- Enterprise Guide 3.0.2
- PC/SASv913 and SAS/Connect
- ETL Studio 3.2
- SAS Enterprise Miner 5.2 is also being considered but requires additional components (Analytical server) to be configured

#### **METADATA SERVER**

At the heart of the SAS BI framework is the Metadata server. The SAS Metadata server provides the central management and control of any organization's metadata. One of the more important facets in the SAS metadata architecture is inclusion of elements beyond the normal metadata. Some of the features include:

Data about data (Metadata), which could be characterized by the existing SAS dictionary.

Authentication interfaces to most protocols that we use include Unix, Active Directory, Z/OS, Oracle and SQL Server. Inherited Permissions that assigns access controls on all defined objects either down a hierarchy or across common layers.

Dynamic resource allocation. Like a dynamic version of sasv9.cfg and autoexec.sas, resources are assigned to users based on profiles.

Many other features are included for specific SAS products and solutions.

At the heart of the SAS Metadata server philosophy is the goal of an open architecture. First, it conforms to the Common Warehouse Metamodel standard, allowing interoperability with third party products. Beyond this, it has an

open interface which allows you to programme to the Metadata through a variety of environments such as Java, C/C++, COM/DCOM, and of course SAS.

A centralized metadata management facility coupled with an open architecture will help simplify and ultimately streamline your administration of the metadata; repetitive and recurring administrative tasks can be automated. Of course SAS also provides a specific tool to manage the Metadata server – the SAS Management Console, a Java based application. In addition, end user tools such as ETL Studio and Enterprise Guide provide mechanisms to not only manage data, but also to manage the related metadata.

#### METADATA SERVER CONSIDERATIONS

Since all SAS9 clients must first connect to a sas metadata server, robustness is critical. There is no redundancy or failover so the metadata server is a single point of failure that must be managed appropriately. Since it is a regular sas session under the covers, it requires all the usual nourishment on the physical server – disk space, memory, and cpu cycles. For this reason, it is recommended to isolate it from the hurly-burly of the workspace hosting machine.

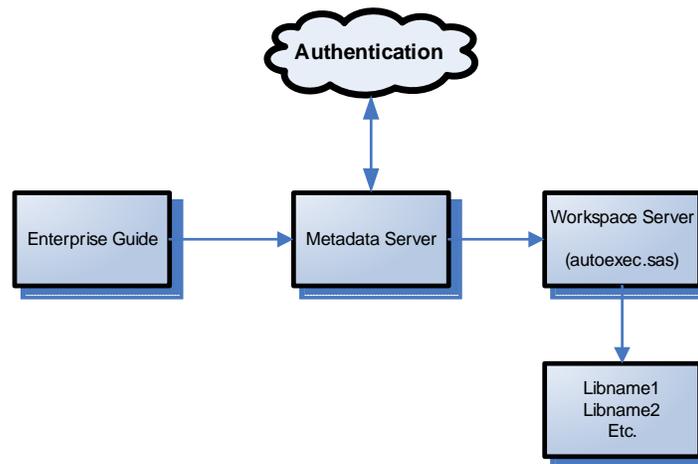
#### ENTERPRISE GUIDE LEGACY

Enterprise Guide has evolved separately from other SAS9 clients. Written in .Net for seamless compatibility with MS Office, many of the metadata interfaces behave differently from other SAS 9 clients written in Java such as ETL Studio, Enterprise Miner and Management Console. Since the Metadata server also uses Java, the minor differences in parsing of the underlying XML can give variable results.

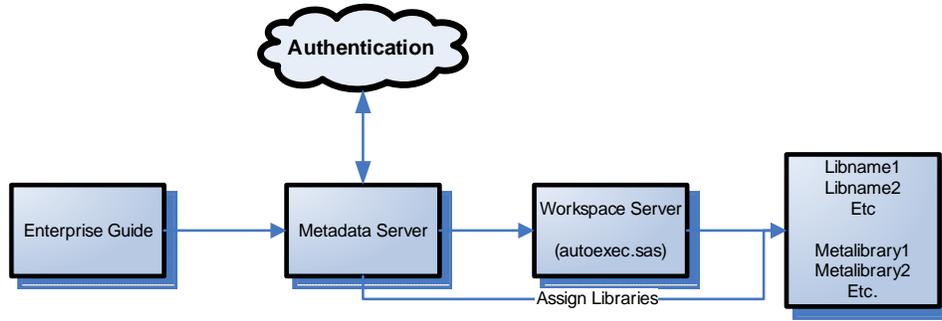
We created the multitude of environments to allow for isolated testing of different behaviour between SAS clients.

#### TWO ENTERPRISE GUIDE USE CASES

- 1 Enterprise Guide connecting straight to workspace server. Metadata server only used for authentication. Libraries assigned by autoexec or within code. No metaautoresources used to assign resources. Works well with Unix where shared configuration files (autoexec.sas and sasv9.cfg) are more commonplace.



- 2 Enterprise Guide connects to Metadata server for authentication, assignment of entitled workspace, libraries and resources. More flexible but subject to variations in assignment of permissions by the metadata server inheritance rules.



## LOGICAL ARCHITECTURE

Decision Solutions has three environments for each of the levels of activity:

### DATA PREPARATION (DATAPREP)

- Classic ETL of multiple data sources into a standardised SQL Server warehouse.

### DATA MODELING AND VALIDATION

- Risk Quantification, Risk Weighted Averages using Base/SAS and Enterprise Miner

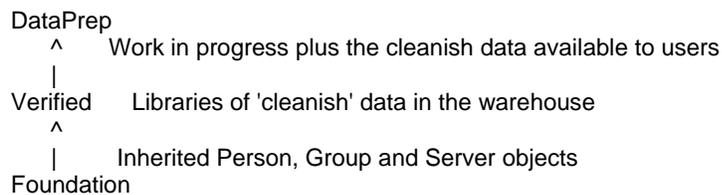
### SCHEDULED LIVE

- Capital calculations, external publishing, governance verification

Theoretically, each environment could have a separate metadata repository within the same metadata server. In practice, it is safer to isolate environments with separate metadata servers on the same machine or separate machines altogether. A minimum of 2 servers is recommended for any applications that depend on live metadata.

## METADATA REPOSITORIES

As a minimum, it is necessary to have at least a Foundation repository. This is created as a result of running the Configuration Wizard at the end of a SAS 9 installation. Keeping all metadata in a single Foundation repository is only suitable during construction or for simple environments. Custom Repositories can be created that inherit properties from the Foundation. A good split is to define all Person, Group and Server in the Foundation Repository and define SASlibrary and DatabaseSchema to dependant custom repositories. For example, using our 3 project environments above, there would be three repositories:



Each of these need to be controlled and isolated with separate failover, test and construction environments. This is achieved with multiple SAS metadata servers and repositories on multiple physical machines.

## PHYSICAL ARCHITECTURE

Physical Server	Lev0 - Devl/construction Logical servers on TCP ports ending in 0	Lev1 - Test/Live Logical servers on TCP ports ending in 1	Lev2 - Restricted/Scheduled Logical servers on TCP ports ending in 2	LevN - Other projects. Logical servers on TCP ports ending in N
SAS01 Primary Win2003 Server 8 processor	Metadata Workspace (Pooled) Connect (PC/SAS) StoredProcess	Metadata Workspace (Primary load balanced) Connect (Repository Promotion) StoredProcess	Metadata (Failover copy from sas02) Workspace (reserved) Connect (Repository Promotion) StoredProcess (inactive) OLAP	Nothing
SAS02 Secondary Win2003 Server 2 processor	Metadata (Backup/construction) Workspace (Functional Testing) Connect (Repository Replication) StoredProcess	Metadata (Testing copy from sas01) Workspace (light load only) Connect (Repository Promotion) StoredProcess (Test only) OLAP	Metadata (Primary for sas01) Workspace (copy of sas01) Connect (Repository Promotion) StoredProcess (inactive) OLAP	Metadata (Failover) Workspace defined on other SAS machines
CASTDEV Development Win2003 Server 2 processor	Metadata Workspace Connect StoredProcess OLAP DataStepBatch Share	Metadata (Construction of sas01) Workspace (inactive - use sas01 or sas02 instead) Connect (Repository Replication) StoredProcess (Test only)	Metadata (Failover copy from sas02) Workspace (Config testing) Connect (Repository Promotion) StoredProcess (inactive) OLAP	Metadata (Construction) Workspace (Config only) Whatever else is needed.

## CONFIGURATION MECHANISMS

If multiple repositories are overlaid on the grid above, a three-dimensional (Repository, PhysicalServer, Environment) configuration array can be visualized. Backup and restore add another dimension.

Unofficial Foundation repository tricks - not endorsed by SAS Institute

Metadata repositories are in-memory databases loaded from SAS datasets in a sub-directory. Pausing the Metadata server causes the physical library to be refreshed. At this point, the canny and cautious administrator can copy the library to somewhere else. Resume the server when done.

**Caveat:** Dependant custom repositories have all sorts of links defined. Only try this with a standalone Foundation repository between similar file systems i.e. NTFS to NTFS.

The target can be an equivalent location on another machine. Pause, copy and resume as on the source machine. If the embedded SAS system accounts (sasadm, sassrv, sastrust) are identical, then this seems to work fine.

## OFFICIAL SAS PROCEDURE

To back up interdependent repositories in a single environment, SAS provide %OMABKUP which is fine but does not have versioning. This minor modification can handle changes between weekly server backups:

```
%omabakup(DestinationPath="E:\SAS\EMiner\Lev1\SASBackup_%sysfunc(substr(&sysday, 1, 3))",
  ServerStartPath="D:\SAS\EMiner\Lev1\SASMain",
  RposmgrPath="MetadataServer\rposmgr",
  Reorg=Yes) ;
```

Although this can be scheduled daily, it is safer to run manually initially as Proc Metaserver cannot always resume the metadata server successfully.

There is no equivalent macro for restoring. Instructions are here: [Backing Up the SAS Metadata Server \(http://support.sas.com/rnd/eai/openmeta/v9/setup/backup.html\)](http://support.sas.com/rnd/eai/openmeta/v9/setup/backup.html)

## METADATA SERVER WORKAROUNDS

In a homogeneous environment where most users need the same access to base SAS libraries, a standard autoexec.sas works well. Pre-assignment can be added if needed.

When there are multiple groups with different access requirements to varying data sources on different platforms, then more thought is needed. For example, a group may require read access to data on SQL Server, write access to a shared Base SAS library and a semi-permanent library for persistence between sessions.

Enterprise Guide (EG), Workspace Server (WS), and Metadata server (MS) configurations

	Pros	Cons
EG + WS + pre-assigned libraries	Very reliable. Only need to add accounts to Foundation repository Shared autoexec.sas commonly used in Unix environments	Very little dynamic control
EG + MS + Foundation meta libraries	Easier to administer initially. Easily replicated on another machine	Can get clogged up with conflicting permissions Difficult to isolate stable definitions from other changes.
EG + MS + Foundation + Custom repositories	Better isolation of changes.	More complex to configure.



## CONCLUSION

For many organizations attaining a position of having a centralized metadata repository that surfaces disparate data sources to a variety of clients and processes, and at the same time maintaining stringent access controls would be nice. For the modern bank it is a necessity. At CIBC we found the SAS BI architecture, and particularly the Metadata server, essential tools to help us attain that goal. However, the road is not smooth. Careful planning and a willingness to stop, review, and periodically possibly abandoning your work to start over is required. Be prepared for problems on the road, but the journey will be worth it.

## ACKNOWLEDGEMENTS

Tina Berry and John Parsons, SAS Institute Canada  
For organizing an evaluation license of SAS ETL Server that would run on a laptop.

## ABOUT THE AUTHORS

**Peter Eberhardt** is SAS Certified Professional V8, SAS Certified Professional V6, and SAS Certified Professional - Data Management V6. In addition his company, Fernwood Consulting Group Inc. is a SAS Alliance Partner. Peter is a regular speaker at SUGI, SESUG and NESUG. If you have any questions or comments you can contact Peter at:

Fernwood Consulting Group Inc.,  
288 Laird Dr.,  
Toronto ON M4G 3X5  
Canada  
Voice: (416)429-5705  
e-mail: [peter@fernwood.ca](mailto:peter@fernwood.ca)  
[www.fernwood.ca](http://www.fernwood.ca)

**Andrew Farrer** is a SAS Specialist at CIBC currently engaged in Basel II activities. He specialises in the configuration of SAS products across multiple platforms with a bit of system administration thrown in. It has been 13 years since he discovered the joys of the data step. Since then, an aversion to data wrangling has led him to administration and support. If you have any questions or comments, you can contact Andrew at e-mail: [Andrew.Farrer@cibc.com](mailto:Andrew.Farrer@cibc.com)

**Zulfiqar (Zul) Habib** is an independent SAS Consultant with over 23 years of experience. He is a SAS Certified Professional (V6/V8) and a SAS Alliance Partner. His industry experience includes: Financial, Banking, Marketing, Insurance, Petrochemical, Public Utilities, Human Resources, Retail Distribution, Government, Pharmaceutical, and Credit Risk. If you have any questions or comments, you can contact Zul at:

Habib Consultants Limited  
10520 Yonge Street,  
Unit #35B, Suite 209  
Richmond Hill, ON L4C 3C7  
Canada  
(416) 721-8667  
Email: [zhabib@habibconsultants.com](mailto:zhabib@habibconsultants.com)  
[www.habibconsultants.com](http://www.habibconsultants.com)

SAS, and SAS Alliance Partner are registered trademarks of SAS Institute Inc. in the USA and other countries. Other brand and product names are registered trademarks or trademarks of their respective companies.