Integrating Your Analytics In Database with SAS, Hadoop and the Teradata EDW

Bob Matsey, Teradata Corporation

ABSTRACT

Submission Title: Integrating Your Analytics In Database with SAS, Hadoop and the Teradata EDW with cover how cutting edge companies are using Agile Analytics and Data Labs to combine and analyze many different data types in Teradata EDW.

INTRODUCTION

Many pharmaceutical companies are leveraging Agile analytics to increase
Turn-around time and performance in various aspects of the business. With Agile analytics and Data Labs, business analysts, Data Scientists and business users are empowered to test, Prototype and experiment with an hypothesis by having direct access to many types of corporate data without needing requirements definition, projects to define data, engaging IT to load their data, and other processes and procedures that bottleneck the 'quick' nature of Agile Analytics insights.

We will discuss technics on how to bring multiple data types together quickly in a Massively parallel environment to analyze your data quickly and get the results back in a fraction of the time that they use to.

AGILE ANALYTICS – DATA LABS

IT typically has a back-log of projects and cannot easily or quickly change
Priorities and help with loading of new data types quickly. By leveraging SAS analytic tools and Teradata's Data Labs, you can provide and environment that is flexible to upload disparate data types and merge it with corporate EDW data inside the (extremely powerful and quick) MPP Data Warehouse environment. Business Analyst's & Data Scientist groups can prescribe to an "analytic" model that allows groups to scale up quickly, allowing them to have a wide quantity of different data in one place while providing an increased quality of analytics while decreasing delivery cycle times. This co-developed technology solution provides an efficient alternative to the traditional Analytic development life cycles for analyzing and preparing data to go into building analytic models and increasing the velocity of analytic discovery.

In this presentation, we will share technology capabilities, best practices and success stories in the pharmaceutical industry. An example of a High Level Agile Analytic Architecture is in Figure 1.

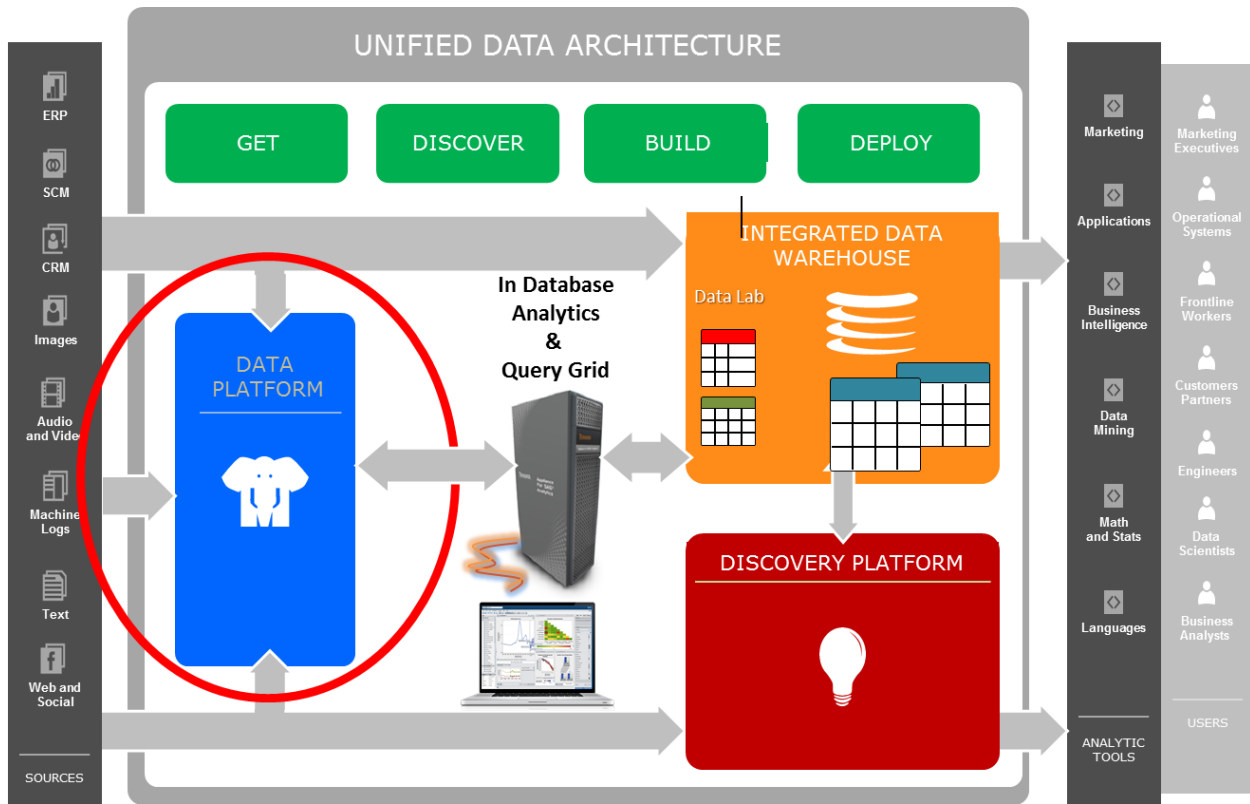Figure 1 Agile Analytic High Level Architecture

Figure 1. Agile Analytic Architecture

The above Agile Analytic Architecture is an example of how Pharma customers are providing and Agile Analytic environment with their Enterprise tools such as Teradata and SAS. Combining Hadoop and the ability to bring in other data types such as Excel spreadsheets, external data, Oracle & DB2 data along with Access Databases and other types provides an end user ample opportunities with multiple data types in one environment.

DATA LABS

Business Analysts, Data Analysts, Business Users & Data Scientists all have one thing in common: They Need Data. They have a wide range of data needs that they need loaded quickly can vary from Hadoop, excel spreadsheets, SQL Server data, Oracle, SAS data sets, external data, other DBMS's and non-structured data which all needs to be analyzed with their traditional Data Warehouse data. Being able to combine this in one place with the power of a Massively Parallel processing Data Warehouse provides many benefits to the uses. These benefits range from:

- Speed & Efficiency (Performance) – gets results back in seconds verses hours and days

- No data movement out of the warehouse
- No redundant data sitting in multiple places that need to be cleaned up
- Frees up IT to do other things verses having to load data for users
- Increases productivity of the Analysts (they get their results and can prove their theory quickly)
- One 'Version of the Truth'
- Reduces cost by eliminating many 'personal Data Marts'
- Increased Collaboration across business users and business lines
- Can be used for multiple purposes as a development platform, migration platform and others

Below is an example of a Data Lab Architecture that can be implemented for an Analytic organization. It encompasses any type of analytic tool along with the EDW and multiple data types depending on the business's needs.
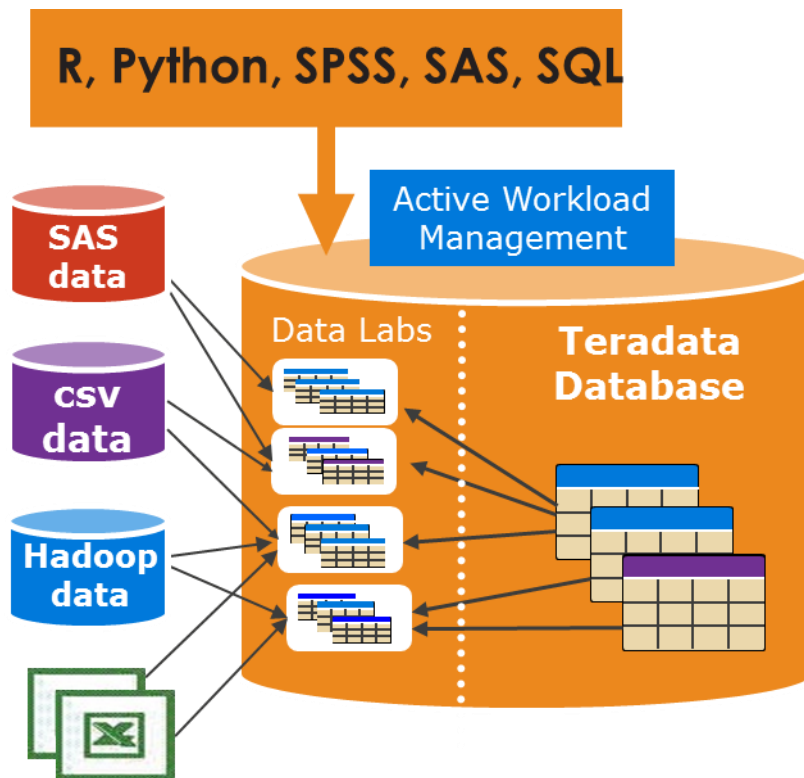
CONCLUSION

In closing, Bob will share how Pharma and all types of other organization are using Agile Analytics to their advantage and are able to change the way they are doing their data preparation and work to make business decisions quickly and easily in the Data Lab environment. By using this simple but collaborative environment, organizations are discovering new value in their data on an ongoing basis. He will provide

some case studies from existing customers using these capabilities. Integrating Teradata and SAS In Database technologies together opens the door to many capabilities and speeds not seen before. It provides analysts with the ability to find those valuable nuggets of new data attributes & information that weren't readily available before and are important for the organization to grow and develop new business opportunities

CONTACT INFORMATION

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

Bob Matsey
Teradata Corporation
704-258-6166
Bob.matsey@teradata.com