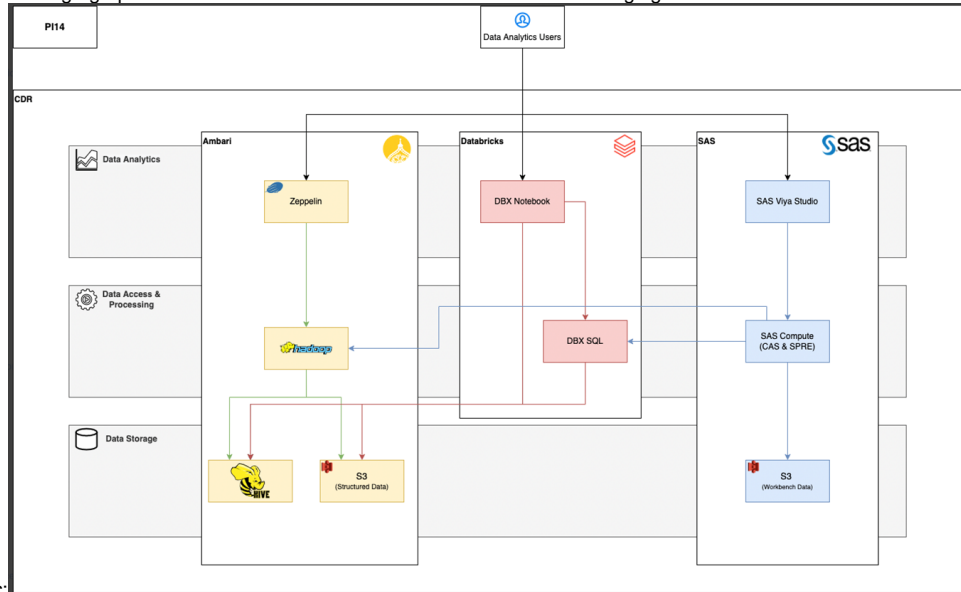# 6-28-2023 Townhall Q&A

Ambari Decommission

1. Is there a formal schedule for team migration regarding the pending Lustre file migration? (If possible, we would request a lead time of at least 30 days to ensure sufficient prep time on our part and any post-testing/verification.)
   A - The team is working on a schedule and will communicate details as soon as possible.
2. Is there any kind of architecture design graphic that can be shared so I can understand what is changing?

   A- Here is the current graphic:
3. Are all three phases coming in July 2023?
   A - No. The phases will take place over several months beginning in July 2023. Phase Two will begin on Friday, August 4, 2023. The schedule for Phase Three will be announced later.
4. How will this switchover impact the current SAS Viya code?
   A - No current SAS Viya code will be impacted. All the macros will execute as normal and connect to the Ambari cluster running them as is.
5. Will we be able to set this once or will we need to update all our existing hive_exec_sql statements for Databricks?
   A - Each macro call will need to be modified to use Databricks. However, your current code will still work but will be using Ambari until it is discontinued.
6. Is there an estimate of how long until Ambari is discontinued?
   A - We want to be off Ambari by Fall 2023.
7. Can you clarify what you mean by "Fall 2023" for the Ambari discontinuation? Is the aim September? November? A moving target?
   A - The aim is to have this done as quickly as possible, but I understand there are deliverables that are due. That is why the fall 2023 right now, but after we release SQL cluster, this is top priority.
8. After SQL integration, will R integration be on the agenda?
   A - Currently, we have an R only compute cluster, but it cannot access the Hive metastore. We have no way to access any data from an R node. Users will be aerated once this has changed.

Data Backup & Recovery

1. Will these storage tiers affect how we access these files? We often access end of year (EOY) files during the next EOY. Would we need to change the path of library on Workbench?
   A - For tiering, there will be no affect to your files. They will appear in the same spots regardless of the tier they reside in on the backend. For archiving, the files will not appear in your Workbench. This functionality will be implemented in a future state, and users will have the ability in the portal to provide self-service to manage un-archiving objects.
2. Is this applicable to data in Hive or Workbench or both?
   A - It is applicable to all user data.
3. Is it archived user-specific or only data that no one has accessed for 13 or 18 months respectively?
   A - It is not user-specific; it is governed by data usage patterns.

SAS Viya and Databricks (DBX)

1. Can you connect to both of your groups in one session by running the macro twice?
   A - Yes. Users can only connect to DBX using one group at a time but in the same SAS session you can connect to multiple groups.
2. This demo covered the HiveQL to Databricks transition but what are the implications for existing code and processes that operate in the CASLIBs?
   A - The CASLIBs will all have their underlying connections updated to the DBX cluster and all existing code and processes will run the exact same using the CAS engine within SAS Viya.
3. Will this example code be made available somewhere to us to review on our own?
   A - All the example demo code will be shared to all users as we migrate users over to use the DBX cluster. We will also have documentation available on Confluence for users to utilize.

4. When will the SAS Viya-Databricks connection be available?
   A - It is currently available but only to certain groups. We will be onboarding other groups in the near feature. Details of this will be made available as soon as possible.
5. Is there any documentation showing us how to access the SAS workbench using Databricks R cluster?
   A - We will be providing the sample notebooks. We are working on it to provide them to all users.
6. After SQL integration, will R integration be on the agenda?
   A - Currently, we have an R only compute cluster, but it cannot access the Hive metastore. We have no way to access any data from an R node. Users will be aerated once this has changed.