



Created : 01 December 2017

Last Updated :

Date Submitted : 01 December 2017

Version Number : 1.0

Author : Logan Data Team

Health Check Report – Boston University Cloud Data Integration Project

Contents

[1 INTRODUCTION 2](#_Toc500155946)

[2 OBJECTIVES 2](#_Toc500155947)

[3 HEALTH CHECK REPORT DETAILS 2](#_Toc500155948)

[3.1 BioRAFT 3](#_Toc500155949)

[3.1.1 Strengths 3](#_Toc500155950)

[3.1.1 Opportunities For Improvement 3](#_Toc500155951)

[3.2 Error Handling And Monitoring 3](#_Toc500155952)

[3.2.1 Strengths 3](#_Toc500155953)

[3.2.2 Opportunities For Improvement 4](#_Toc500155954)

[3.3 SnapLogic Infrastructure 4](#_Toc500155955)

[3.3.1 Strengths 4](#_Toc500155956)

[3.3.2 Opportunities For Improvement 5](#_Toc500155957)

[3.4 Integration Error Report 5](#_Toc500155958)

[3.4.1 Strengths 5](#_Toc500155959)

[3.4.2 Opportunities For Improvement 5](#_Toc500155960)

# **1 INTRODUCTION**

Logan Data Inc., located at 2 Lan Dr., Ste. 200, Westford MA, provides Data Consulting Services. Logan Data had worked previously on several Data Integrations Projects for Fresenius Medical Care, SnapLogic, Insulet and various other clients. Logan Data has a Master Service agreement with Boston University. The company has experienced employees with varied and diversified skillsets who are SME’s and Certified Professionals in various consulting spheres.

# **2 OBJECTIVES**

To create a Health Check Report for Cloud Data Integrations currently underway at Boston University, Logan Data team in consultation with BU SME’s and Technical Leads; picked up pipelines under 2 subject areas that are in advanced stages of development and/or are deployed to production. Purpose of this health check was to get an independent assessment on how well the pipelines were performing in accordance with their objectives and how well they adhered to the methodologies’ best practices. Existing SnapLogic Infrastructure was also looked at for a health check.

# **3 HEALTH CHECK REPORT DETAILS**

Logan Data Team looked at the following 2 project areas in BUTest Org – BioRAFT under Admin-Research-Systems/RIMS and Error Handling under BUIntegration/ErrorHandling. All objects pertaining to the subject areas were looked at. BU’s existing SnapLogic Production Infrastructure was also looked at for assessment.

## 3.1 BioRAFT

BioRAFT Integration is created as a triggered task that calls BioRAFT pipeline. This pipeline makes a REST GET call to BioRAFT API to get Training Information and returns the same to the caller.

### 3.1.1 Strengths

The task and the pipeline itself have been designed and developed as per standard Cloud Integration Guidelines. Naming conventions used on the Snaps in the pipeline are descriptive and convey what each Snap is expected to perform.

### 3.1.1 Opportunities For Improvement

Error Handling should be added to the pipeline. Already an effort is underway to use an Error Handler Pipeline to log and record errors and the same can be employed here OR as based on need of this pipeline, a different approach can be followed. A good practice is to include a brief description and comments in the Notes box under Info tab of each snap about what operation the snap is doing. This self-documents the code. Authentication fields are visible in the code – if these need to be kept secure then probably they can be stored in a file/table and read from it before being passed onwards in the pipeline.

## 3.2 Error Handling And Monitoring

BU is working on establishing a standard error handling and monitoring process for its newly adopted cloud integration platform SnapLogic. The objectives are:

1. Build a standard error handler pipeline that can be used by all other pipelines to send real-time error notifications for the non-fatal errors
2. Build error monitoring pipelines to run on an hourly basis to gather new failures and send notification emails as well as keep track of the errors in the disposition report tables for various reporting needs

### 3.2.1 Strengths

New feature ‘Error Pipeline’ in the recent SnapLogic release has been used for building an easy-to-use portable Error Handler pipeline that is ready to use for any pipeline.

Error monitoring pipeline provides, Admins pipeline health reports in the form of automated hourly error summaries and reduces Admin’s monitoring effort.

### 3.2.2 Opportunities For Improvement

**Error\_Handler**

Real-time error notification emails should be sent to respective developers if they are developing in their own project folders. Errors in the protected project areas will be sent to the admin distribution list, which can be done in following way:

Set receiver to pipe.user instead of the hardcoded distribution list in the Email Sender snap.

In Dev, when developers develop pipelines in their own project folders, error/failure emails will be sent to the pipeline owner only. Otherwise if we sent to the distribution list, all users will be flooded with error notification emails.

Only development leads or admins have the service accounts for moving the code to the protected project area using the service account. This makes sure the owner of everything under the protected project area is the service account. Any error in the protected project area will be sent to the service account email id. Email forwarding can be configured in outlook account settings for the service account email id to forward emails to a distribution list. Outlook will automatically forward the emails. It does not require the service account to stay connected to be able to forward the emails.

**Error\_Monitoring\_and\_Notification**

1. Include the parent ruuid and invoker in email notification.
2. Include the runtime label in the subject, for example:

Change from 'ERROR - Pipeline failure' to 'ERROR - Pipeline failure on Cloud - BUDev'

1. This will make it more flexible with configurable run interval which is currently using default

## 3.3 SnapLogic Infrastructure

Existing BUProd Org was looked at for this exercise. SnapLogic Environment at BU has been setup as a combination of a GroundPlex and CloudPlex which will allow it to smoothly create Integrations for Cloud Applications like AWS, Salesforce, etc. and allow it to maintain a on-premises connectivity to applications like Oracle, Mainframe, etc.

### 3.3.1 Strengths

BUProd CloudPlex is set up as a 2 node system, with each node having 4 cores, running Linux on amd64 with 16 GB of memory and a Java version of 1.8.0\_45. This is as per recommendation from SnapLogic.

BUProd GroundPlex is also set up as a 2 node system, with each node having 4 cores, running Windows Server 2012 R2 on amd64 with 16 GB of memory and a Java version of 1.8.0\_111. This is again as per recommendation from SnapLogic.

Above is very close to what BU envisions for its IPaaS Production Target Technology Architecture.

### 3.3.2 Opportunities For Improvement

Currently BU’s SnapLogic Integration Production Environment is at a nascent stage. Based on the roadmap drawn up by BU’s Technology Team, this is expected to see a rapid growth over the next few years. It is recommended to have a periodic check on the Groundplexes to ensure that they continue to perform optimally and serve business needs. Accordingly decisions to upgrade Hardware, OS, etc. can be taken.

## 3.4 Integration Error Report

### 3.4.1 Strengths

### 3.4.2 Opportunities For Improvement