# Nachiket Joshi

# SJSU ID: 011408956

# Extra Credit Assignment

TABLE OF CONTENTS

CHAPTER PAGE

CHAPTERS

CHAPTER 1 – **Introduction To Netflix Atlas** **3**

CHAPTER 2 – **DB2 On Cloud** **13**

CHAPTER 3 – **Graph Data store** **35**

CHAPTER 4 – **Summary** **49**

Introduction To Netflix Atlas

Every company nowadays is investing much into the realms of Big Data analysis. The sentiment analysis to know the near future analysis of the data is going to be useful for the business. Every company generated multiple terabytes of data every day. The real time analysis to give the operational insights of this data will generate positive results to decide the direction of the business.

In pursuit of above-mentioned phenomenon, Netflix developed Atlas to manage dimensional time series data for a near future real-time operational insight. The main feature of Atlas is that is provides in memory data storage. This allows gathering and reporting of very large number of metrics and that too quickly.

Atlas is a tool specifically designed to study operational intelligence. This study is done over the business data that is gathered over time. The main aim of operational system is to know what exactly is happening inside a system currently. This also helps establish trends among the customers.

The business that the Netflix does suddenly increased from 2011 to 2014 and the existing operational intelligence methods were not suitable to study this gigantic leap in data to study thus they designed this new Atlas based on the stack language which is able to handle big data easily. Netflx can also scale with the hardware that we use.

OVERVIEW OF NETFLIX ATLAS

Atlas is a system that the Netflix uses to manage highly complex to manage dimensional time -series data. The issues it addresses are scalability and high query attendance capability in the previous system.

The main goal of this system was to build a system that provided,

* A Common API
* Scale
* Dimensionality