Learning Storage Networks is enterprise software intended to be used by HPE and their customers to collect data and analyze networks automatically, without need for manual input or intervention until actual erroneous behavior is detected. This will reduce both time and money spent to deal with data loss and hardware damage.

Abstract

Learning Storage Networks is enterprise software intended to be used by HPE and their customers to collect data and analyze networks automatically, without need for manual input or intervention until actual erroneous behavior is detected. This will reduce both time and money spent to deal with data loss and hardware damage.

Approach

Data Collection/ Cleaning
Data Source Integration
Data Normalization
ML Algorithm Training
Next "State" Prediction

Overview

Hewlett Packard Enterprise (HPE) offers storage services to other corporations to fill today's massive demand of data storage spaces. Specifically, they offer storage area networks (SANs) to allow for large scale storage, tracking, and analysis of data. These networks are made up of interconnected Fibre Channel switches, that deteriorate over time and are sometimes manufactured defective. The magnitude of these networks prevents easy manual error detection, and thus, there is much time and money to be saved for companies like HPE with the use of machine learning to predict switch anomalies before they happen.

Storage Area Network

What are Switches?
Switches connect storage, servers, and other switches in a network called a fabric. Our architecture is comprised of a script running on a server that uses a single switch to collect data from fabrics of switches. This collected data is then analyzed via a machine learning model which predicts whether or not a switch will enter an error state.

Results

- Automated data collection of Brocade switch data.
- Trained an ML model that is able to predict 15 future errors with 55% accuracy.
- Can notify network operators of possible risks before they occur.
- Prevented errors could potentially allow for a significant amount of costs to be avoided.

Acknowledgments

We would like to extend a special thanks to Ayman Abouelwafa and Salil Gokhale for facilitating the process of communication between HPE and our team, as well as providing technical assistance whenever necessary. We deeply appreciate the care and effort which Professor Richard Jullig as well as our TAs Chandranil Chakraborti and Golam Muktaadir provided us to support the entire process of this project.