Outage Prediction

Predicting outage risk to assist resident preparation

Joanna Huang, Kalvin Kao, Justin Plumley

November 7, 2017

Background:

- “In terms of the total number of lost hours of electricity, Puerto Rico and the US Virgin Islands are in the midst of the largest blackout in US history...”
- Power outages due to downed power lines are a frequent occurrence in the US
- Business Problem: Extreme weather events result in economic, social, and physical disruptions and inconvenience and loss of critical lifeline systems for residents living in disaster areas

Project objective:

- Use weather and power datasets to predict outages and their potential magnitude to inform residents and assist in preparation

Datasets for Learning

- Learn whether an outage will occur given weather
- Learn the magnitude (number of affected customers) of a power outage if an outage occurred, given weather

NOAA

(National Oceanic and Atmospheric Administration)

Provides weather history over a long period of time and / or locations

Data fields: sky conditions, visibility, temperature, humidity, wind speed, press

Data size: up to 30 observations/day (per city)

https://www.ncdc.noaa.gov/

EIA Electric Power

Major Disturbances and Unusual Occurrences, Year-to-Date 2017

Data fields: Year, Month, Event Date and Time, Restoration Date and Time, Duration, Utility/Power Pool, NERC Region, Area Affected, Type of Disturbance, Loss (megawatts), Number of Customers Affected

Data size: 50-100 observations a year

https://www.eia.gov/electricity/monthly/epm_table_grapher.php?

Datasets for Prediction

- Weather and power forecasts
  - Predict future outages and their magnitudes given weather and power consumption forecasts
- Forecasts are updated regularly, and become more accurate for closer dates
  - Predictions may be updated as forecasts are updated

Weather Underground API

Real-time forecasts

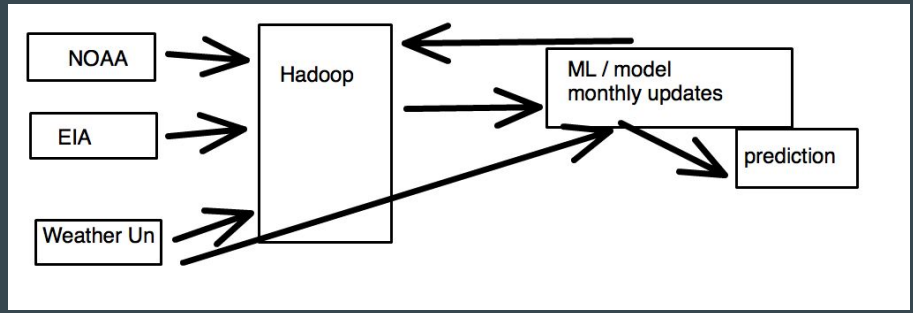
Data fields: location, observation time, weather description, temperature, humidity, wind (dir, mph, pressure), precipitation)

Data size: Depends on number of cities

http://www.wunderground.com/weather/api/

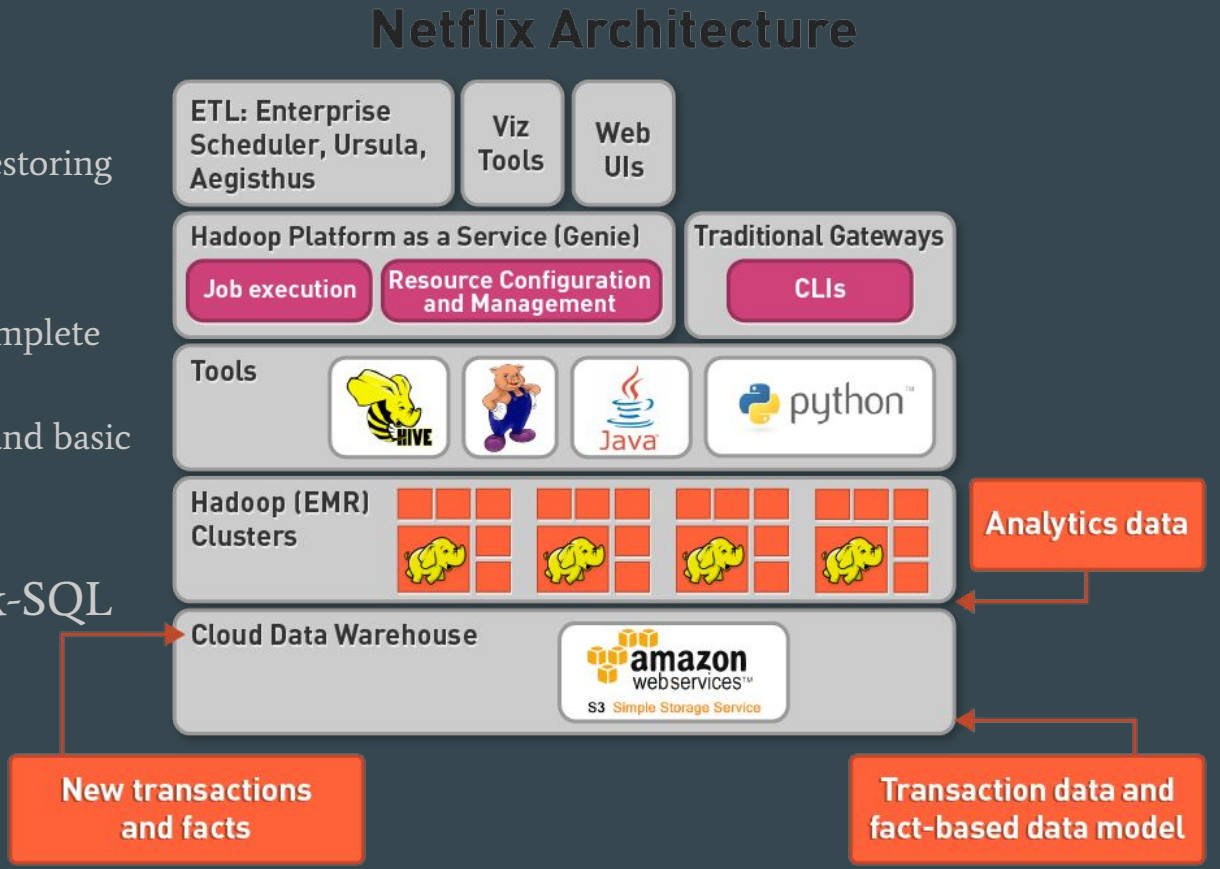
Design

- Steel Thread
  - Data: Weather and power outage data from single state (e.g. Maine)
    - 1 day of weather
  - Store: Hive/Hadoop
  - Model: Predict ‘no outage’
  - Output: csv
- Complexity overview:
  - Volume
    - Estimate maximum of ~100mb data per weather station
  - Variety
    - Excel, CSV
    - i.e. temperatures, sky conditions
  - Velocity
    - Update learning monthly
    - Update predictions daily



Architecture

- Architecture: Netflix
  - One data warehouse for restoring predictive analyses
- Storage: Hadoop
  - Small, structured, semi-complete data
  - Low latency, low quality, and basic query requirements
  - Already scaled out
- Processing: Apache Spark-SQL
  - MLlib, RDDs, pySpark



Questions/Concerns

- 3 Vs?
- A lot of previous work in this area
    - More accurate predictions with soil moisture, tree density

Next Steps

High-speed winds during a thunderstorm may cause trees around an electric grid to crash into the distribution system feeders causing an outage in that area. This model predicts potential vulnerability to ensure the trees in the most critical areas with the highest risk to be trimmed first.

"Power outages in the distribution system are primarily caused by **wind blowing trees and limbs onto power lines**," says Guikema. "Therefore, our model includes a number of different measures of **wind, the type of trees, and the soil moisture conditions**, which provides a measure of the stability of the soil and the likelihood of a tree being uprooted."

A group of researchers led by Seth Guikema at the University of Michigan has created a model that predicts the number of power outages "based on **wind speed estimates, population density, soil moisture levels, drought indices, and information about trees in each census tract.**"

Design/Architecture

Steel Thread

We will be building out steel thread starting with weather and power outage data from

Pipeline Overview:

Data ingestion → processing/transformations → analysis → serving layer

Complexity

Volume

Velocity

Variety

Architecture/Scale-Out

Background

- Power outages due to downed power lines are a frequent occurrence in the US
- Other developed nations, such as Germany and the UK, have replaced some overhead power lines with underground power lines
  - Aesthetic purposes
  - Less hazardous
  - Protected from weather or vandalism
- Why not the US?
  - High initial costs
  - Some geographical/geological limitations
- Business Problem:
  - Undergrounding may also decrease operational costs over the lifetime of the cables... but by how much?

Overview

- Historical weather and power outage data
  - Learn whether an outage will occur given previous weather
  - Learn the magnitude (number of affected customers) of a power outage if an outage occurred, given preceding weather
- Weather and power forecasts
  - Predict future outages and their magnitudes given weather and power consumption forecasts
- Forecasts are updated regularly, and become more accurate for closer dates
  - Predictions may be updated as forecasts are updated
- Business Use Case:
  - Determine which regions of the US will gain long-term savings by undergrounding power lines