# Metro Vancouver Bus Dashboard

JONATHAN BHASKAR

## Web demo:

This project has a live web demo which can be found at http://52.27.110.152:5000/

## Introduction

Bus services in the Metro Vancouver area is run by the Coast Mountain Bus Company. They have 210 bus routes and serve an 1800 $km^2$ area. On an average weekday, their buses collectively cover a distance equivalent to circumnavigating the earth five times. Even though they have such an impressive service network, personally, I experience bus delays and cancellations often. GPS devices are fitted in all buses, and Translink tracks their locations in real time. By visualizing this real time data with the bus timings, I try to present a tool which can be used to isolate and deal with bus delays.

## Methodology

The final tool takes the form of a dashboard. I'll talk about the different parts and how they work together.

### Data

The data comes from four sources. Translink provides information about all bus trips (Bus number, destination, etc.) and all stops (Stop name, location, etc.). This information is fixed and does not change unless Translink makes route changes. The other source is the Translink API.

The API has two end points. One end point provides real time data about all the buses operating at the moment and their current locations. The other end point provides the next stop of every bus, and if the bus is delayed, how late it is going to be to that stop. The dashboard uses a combination of these four sources. For instance, I get the real time location of a bus, use the bus trip information to find the route and destination, get the real time delay and next stop ID for that bus from a different source, and then use the stop information to get the stop name from that ID.

## Map

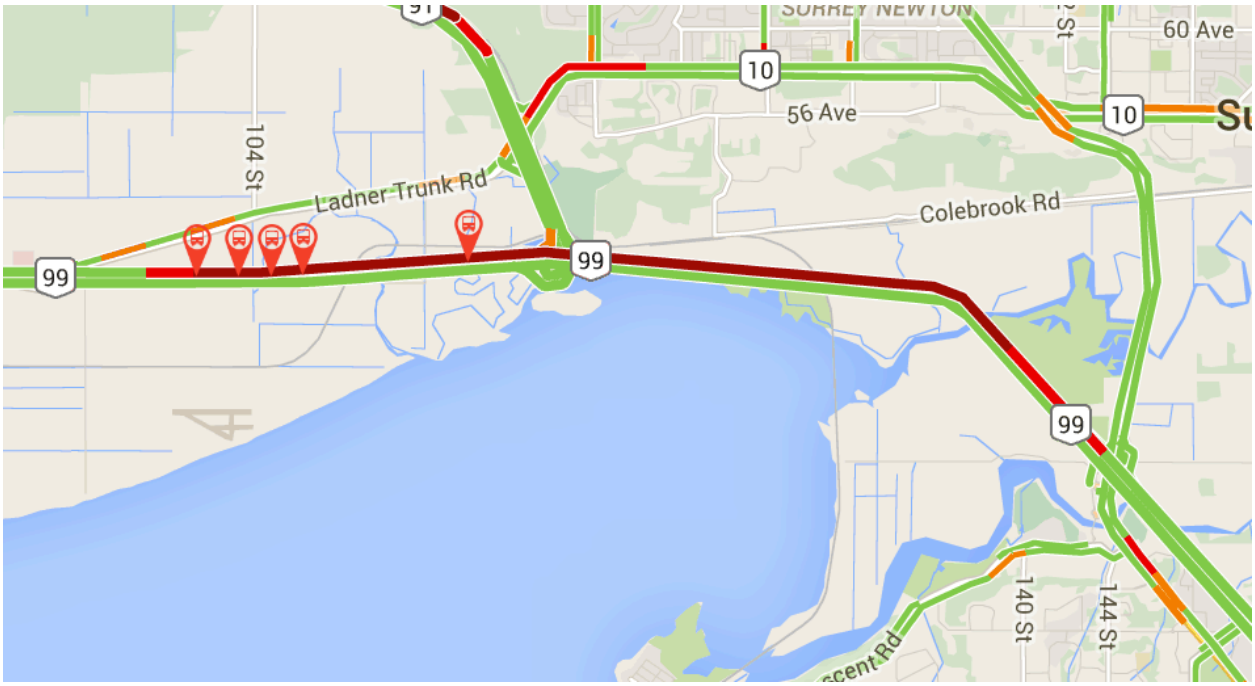
The most important part of the visualization is the map. It has markers which show the current location of every bus currently in service. It takes a short while to process the location information and plot it on the map. A loading indicator is present at the top of the page to make sure users know that the site is operational while the data processing takes place in the background.

I used Google Maps for the map service. Google plots the markers using the location information I send to them (more specifically, latitude and longitude co-ordinates).

Buses on time have blue colored markers and buses with a delay of over one minute have red colored markers. This gives the user a rough estimate of the proportion of the number of buses on time to the number of buses delayed. Clicking on a marker brings up an info window with the combined bus information mentioned earlier. Zoom controls are present in the bottom right of the map.

## Filters and Overlays

The dashboard has two overlays. The first one is the traffic overlay; it shows real time traffic information. This overlay is turned on by default. The traffic data is provided by Google. Roads with high traffic congestion are shown in red, and roads with no congestion are shown in green. The traffic overlay is useful to isolate buses that are late because of traffic problems. The screenshot shows five buses delayed because of a traffic jam in Delta on Highway 99.




The next one is the transit overlay. It highlights transit routes.

The refresh button updates the marker locations to the most up to date position of the buses.

A filter in the bottom right of the screen helps the user to filter buses which are delayed for more than a particular amount of time. This helps the user isolate the buses that need to be paid the most attention.

### Charts

There are two charts which show aggregated data. The first one shows the no. of buses that are on time and the no. of buses on time in a bar chart. This enables the user to get an estimate of the ratio of on time vs. delayed buses at any time. From observing the data for the past few days, this ratio can go up to 0.25 in peak hours, i.e. one in every four buses are delayed. Because of traffic and other reasons, a short bus delay is practically unavoidable. But longer delays can be. The second graph shows aggregated total buses delayed for less than 5 minutes, 5 minutes to 10 minutes and more than 10 minutes. It gives a better estimate of how many buses are running really late.

## Future Work

Some other ideas that could be incorporated:

1. Show bus routes visually: Clicking on a bus marker on the map could highlight the route the bus is currently on. This could tell the user if a bus was late because it was stuck in future, or if it is likely to run into traffic congestion later on.
2. Historical Data: Data could be stored to provide historical insights. For example, when a bus is selected, data like how many times the bus has been late in the past week could be retrieved and displayed.

## Conclusion

Even though visualization has traditionally been used to represent data that was not very volatile, with interactive visualizations and real time data becoming commonplace, real time visualizations could be a valuable tool. Good visual imagery makes perception and understanding easier, and so real time visualizations could be used to make sense of data quickly and make decisions on the fly.