NASCIO 2019 State IT Recognition Awards –
Category: Data Management, Analytics and Visualization

Commonwealth of Kentucky
Kentucky Transportation Cabinet

Intelligent Transportation System (ITS) Data

Project Initiation: April 2017
Project Completion: December 2018

Heather Stout, Executive Director
Office of Information Technology
Kentucky Transportation Cabinet

Dr. Charles Grindle, Chief Information Officer,
Commonwealth Office of Technology

Executive Summary

Our transportation cabinet is charged with "providing a safe, efficient, environmentally sound and fiscally responsible transportation system that delivers economic opportunity and enhances the quality of life in Kentucky." ITS uses innovative traffic management technologies as a cost-effective way to ease the problems. The Commonwealth's ITS project was a high impact, quick hitting effort that addressed some of the thorniest problems in data integration and analytics. Google has used this project to showcase Waze and the power of analytics.

Exemplar

The ITS Data system ingests and geo-enriches each data record so time stamps and road locations can be used to join data. The GoKy map is now is a single source for real-time status of the road network, and includes crowd sourced Waze reports, cameras, Dynamic Message Signs, speed and congestion data, lane closures and road maintenance corridors. This common operational real-time map is available to the public, local governments, and state police.

Concept

- The team created an on-premise production big data server environment using open source software and the following technologies: HBase, HDFS, Hive, Flume, Impala, Kafka, Spark, Sqoop, Zookeeper, Elastic Search, Kibana, and Tableau. Features and data steams were incrementally deployed using Agile project management and development methodologies.
- The ITS ingests and geo-enriches each data record so time stamps and road locations can be used to join data. The GoKy map is now is a single source for real-time status of the road network, and includes crowd sourced Waze reports, cameras,

Significance

ITS provides data to any citizen or entity that accesses the GoKY map making it difficult to assess just how far a reach this project has.

- Automatic Vehicle Location (AVL) 6 - 9 million records/week (reads every 10 seconds)
- Dynamic Messaging Signs (DMS): electronic roadway signs, 470K records/week
- HERE – traffic speed and volume, 49 million records/week
- iCones: speed and traffic volume through manually configured work zones, 180K records/week
- KyMesonet: 80 weather stations
- Roadway Weather Information System (RWIS): 39 location based weather points
- Traffic Operations Center (TOC): 1 million

Impact

While the financials for this project were excellent (completed in 21 months at a cost of $620,775 with a 6 months ROI (one winter), it is the **operational impact** that is tells the tale:

- Provides a common operational picture
- Incident Detection
- Workzone Monitoring
- Snow and Ice Monitoring
- Crowd-sourcing / citizen engagement
- Compliance with Federal (511)

Exemplar

The ITS Data system ingests and geo-enriches each data record so time stamps and road locations can be used to join data. The GoKy map is now is a single source for real-time status of the road network, and includes crowd sourced Waze reports, cameras, Dynamic Message Signs, speed and congestion data, lane closures and road maintenance corridors. This common operational real-time map is available to the public, local governments, and state police.

Imagine the public relations nightmare with 700-800 vehicles and semi-trucks stuck on a frozen snow-covered interstate for 10-12 hours overnight with temperatures at 25F degrees. Now imagine although they were stuck, they still had cell phone communications and access to social media like Twitter, Facebook and Instagram to tell friends, family, and the twitterverse. Finally, top it off by adding one famous celebrity/politician to the mix of stranded commuters. We do not have to imagine it; we saw it happen...twice! The scenario above happened in March of 2015 and halted traffic on I-24 and I-65. The second event was in January 2016 when a blizzard locked up a 30-mile stretch of I-75 and stranded thousands of commuters for over 11 hours.

Incident managers and highway engineers analyzed the available data from both incidents. In the March 2015 incident, data analysis showed weather systems reported significant rain. Since the beginning of the event contained very heavy rainfall, it made pretreatment ineffective. Each of our 120 county offices were on call, waiting for the rain to pass so they could begin treating and plowing. As the cold front moved through the state, temperatures dropped some 30 degrees within an hour, causing standing water to freeze, freezing rain, and eventually very heavy snowfall. Analysis showed the crowd sourcing information from Waze showing reports changing from “heavy rain” to “flooding” and eventually “ice on the road.” These crowd reports progressions were actually earlier than the weather data that was used by the maintenance crews.

In the 2016 incident, there was a lack of clear communication and data sharing between the Traffic Operations Center (TOC) and on-scene personnel. There was a blizzard and we needed to get people off the roadways to wait it out until crews could clear routes. Emergency Management personnel and local law enforcement rerouted traffic off the interstate to an unplowed road and caused additional traffic backups. The lack of a common operational picture and real-time status of the road network again created a problem.

Our problem was the data, or more accurately the massive amount of available but disparate data. Both events together demonstrated our need to use real-time processing of big data for the Traffic Operation Center to make time-sensitive decisions. We also needed to provide a common operational picture with maps, alerts, and data available to the public, first responders, and local law enforcement. The Traffic Operation Center needed to be able to see and visualize the data on a map, to see, validate, and respond to crowd source reported incidents and add them to the map, and to create and send alerts both to the map and to the crowd source community.

During extreme incidents data has to be current and readily available or it's useless. Our data warehouse allowed us to complete analysis after the fact, but it had to be pulled from multiple systems often with nothing effective to join data tables. The nightly ETLs for our data warehouse meant most data was 24 hours old. Our ITS had 10 available data streams that would amount to over 9 million records a day if set to real-time. The Transportation Cabinet's data warehouse could not effectively ingest data sets that size let alone process and operationalize it.

The ITS Data project built upon and became a production implementation of an earlier concept prototype, which delivered successful results using the open source Hadoop architecture stack. It was not part of a planned larger project, but was an incremental delivery of real-time big data capabilities augmenting our data warehouse and several of our applications. A staff of five (5) resources of varying allocation percentages expended 7,335 hours to complete. After implementation of the production hardware environment and initial software install, a service contract was let providing 24/7 monitoring and maintenance of the Cloudera Hadoop software suite. Total combined implementation cost was $620,775.

Our Office of Information Technology (OIT) followed a hybrid Agile Development methodology with three (3) week sprint development cycles for the ITS Data project. The completed project is 100% internally developed and maintained by transportation's personnel on the state's owned and operated hardware.

The Kentucky Transportation Cabinet's (KYTC) Public Affairs office created an ad campaign using videos, social media and press releases to inform the public of the new technologies and pointed them to the web URL: https://transportation.ky.gov/sites/GoKY/Pages/home.aspx . The page also advertised and encouraged the download and use of the Waze app from our partnership with Google.

The project was completed in 21 months at a cost of $620,775 with an estimated return on investment period of less than one winter season (6 months).

Significance

Because ITS provides data to any citizen or entity that accesses the GoKY map or uses the WAZE (Google/Alphabet app), it is difficult to assess just how far a reach this project has.

- Automatic Vehicle Location (AVL) 6 - 9 million records/week (reads every 10 seconds)
- Dynamic Messaging Signs (DMS): electronic roadway signs, 470K records/week
- HERE – traffic speed and volume, 49 million records/week
- iCones: speed and traffic volume through manually configured work zones, 180K records/week
- KyMesonet: 80 location based weather stations, 300K records/week
- Roadway Weather Information System (RWIS): 39 location based weather points, 200K records/week
- Traffic Operations Center (TOC): alerts, lane/road closures, etc. entered by TOC, 1 million records/week

OIT has greatly increased the reach and viewership of KYTC's data by developing strict data standards, which include everything from naming conventions to the timeliness of the aggregation, processing and publication of real-time data. Those standards eased the barriers for sharing data with partners such as Waze. Furthermore, it allowed KYTC's participation with the USDOT in Work Zone Harmonization data feed standard, which will be used by a vast array of private partners working on connected and autonomous vehicles.

The first phase of the project created a production open source Hadoop-based environment for real-time processing of big data. The system then established connections to 10 data feeds from ITS sources, including Waze and Twitter accounts (transportation operations, emergency management, local government, local news and political feeds). The standard data pull is every 2-minutes, except for AVL on snowplows, which is every 10 seconds during snow operations. The throughput is approximately 385,000 records per hour, totaling over 9.3 million/day.

Impact

While the financials for this project were excellent (completed in 21 months at a cost of $620,775 with a 6 months ROI (one winter), it is the **operational impact** that is tells the tale:

- Provides a common operational picture
- Incident Detection
- Workzone Monitoring
- Snow and Ice Monitoring
- Crowd-sourcing / citizen engagement
- Compliance with Federal (511)

Common Operational Picture - The ITS Data project provided situational awareness to everyone via the GoKy dashboard https://transportation.ky.gov/sites/GoKY/Pages/home.aspx). It allows users to explore all Kentucky's roadway network in real time with all of the data feeds provided. GoKy allows operators and the public to view congestion, construction, roadway alerts, snow events, traffic messages, and traffic cameras all from a single map.

Incident detection - Our automated algorithm combines commuter Waze reports and HERE traffic speed/volume data in real time to expected norms and indicates with a visual icon when the threshold suggests the likelihood of an incident. When the system detects that an incident has occurred, the KYTC Traffic Operation Center is notified to further investigate. The incident detection system also provides the ability to monitor the incident in real time, as well as performing after action reviews. After-action reviews are providing useful insights into the order of events that happen that involve crashes.

Work Zone monitoring - A combination of Waze reports, HERE, iCones data aids TOC operators in ensuring that work zones traffic remain in optimal condition. Also provides understanding of traffic control plans and work zone performance on traffic speed, congestion created, accidents, and secondary accidents.

Snow and Ice Operations - An additional 980 snowplows were equipped with AVL. Improved location snapping and geo-enrichment of data provided a fully implemented decision support system for real-time operations and fleet management.

Not only did the ITS Data project meet significant needs for highway traffic management; it was a colossal win to meet federal mandates for 511 (https://www.fhwa.dot.gov/trafficinfo/index.htm) and Map-21 requirements for congestion reduction, system reliability, and freight movement. Likewise, it directly aligns with several of KYTC's strategic goals. First, it improved mobility and access by improving traffic flow and freight movement. Second, it improved organizational performance by applying technology to improve effectiveness and efficiency. Lastly, it strengthens customer and stakeholder relationships by communicating accurate and timely traffic information.

OIT has fundamentally transformed the capabilities of KYTC's Intelligent Transportation Systems by the adoption of an open source, big data environment. This allowed ITS practitioners within the cabinet to develop new methods of incident detection by combining crowd sourcing and sensor data as well as new and innovative methods for developing real-time dashboards for a greater understanding of work zones, snow and ice operations, and overall network performance.

Although the 2018-2019 season was mild, KYTC was able to benefit commuters and commercial freight drivers by keeping the interstates open at least two separate times by using our snow and ice conditions dashboard. While other agencies were considering closing the interstates, the KYTC Director of Maintenance was able to use our data to show that the event didn't warrant any type of closure. That same dashboard is shared with KY Emergency Management, KY State Police, and a few local school systems. It allows KYTC to share real-time information with relevant parties and make more informed decisions during real-time snow operations.

The estimated potential savings in snow and ice operations could be enough to provide the RIO in approximately 3 months of use. The business unit analyzed some data during the initial prototyping. Using AVL data from the first 120 plows equipped, they determined 476.92 labor hours were spent distributing 1432.46 tons salt on non-designated routes. Based on these numbers, the business unit estimates that material applied off network was approximately $1.5 million during a period, which represents only 1/2 of winter season. Using the real-time map and fleet management to keep trucks on designated routes could potentially save $3 million per season. The implementation cost of the was $620,775 with annual support of costs of $345,000.

Congestion Mitigation – According to the latest TRIP report, Kentucky residents spend $1.6 billion due to congestion. The data provided and organized by this project will enable the business unit to gain a better understanding of that congestion and develop solutions for better network performance. [TRIP is a private nonprofit organization that researches, evaluates and distributes economic and technical data on surface transportation issues.]