THE EASTERN TRANSPORTATION COALITION

CONNECTING FOR SOLUTIONS

The ABC's of Conflation: TMC, LRS, OSM – What Happens When You Muck It Up

August 19, 2021

Welcome!

- We are using Zoom **Webinar**
- **AUDIO (Computer):** Use your computer speakers and microphone by clicking the “Join Audio” button at the bottom left of the screen. You will be muted.
- **Alternate Audio (Phone):** Call into the meeting by dialing the phone number based on your location (provided in the confirmation email) and enter the Meeting ID at the prompt. You will be muted.
- **This web meeting is being recorded.**
- **Questions** with the audio or web? Please contact Esther directly via the chat box or email (ekleit@kmjinc.com)

Using the Q&A box and Chatbox

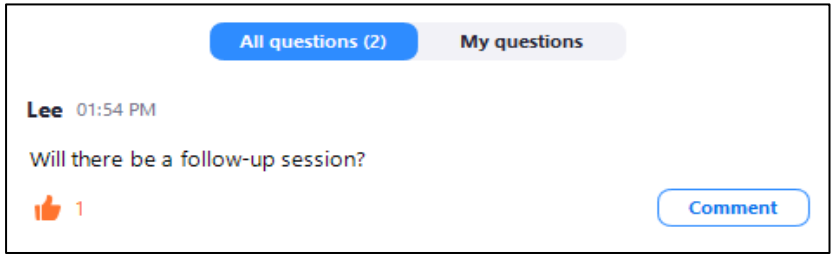
- Use the **Q&A box** to ask presenters questions
- Ex. “How accurate is the captured data?

- Use the **chatbox** for technical issues or to contact Coalition staff
- Ex. “I can’t hear the presenter”

Asking Questions in the Q&A Box

- Click on the Q&A icon at the bottom of your screen

- The questions in the Q&A box will be monitored and answered either between presentations or at the end of the meeting
- You can keep track of your questions in the “My Questions” tab in the Q&A box



Asking Questions Verbally

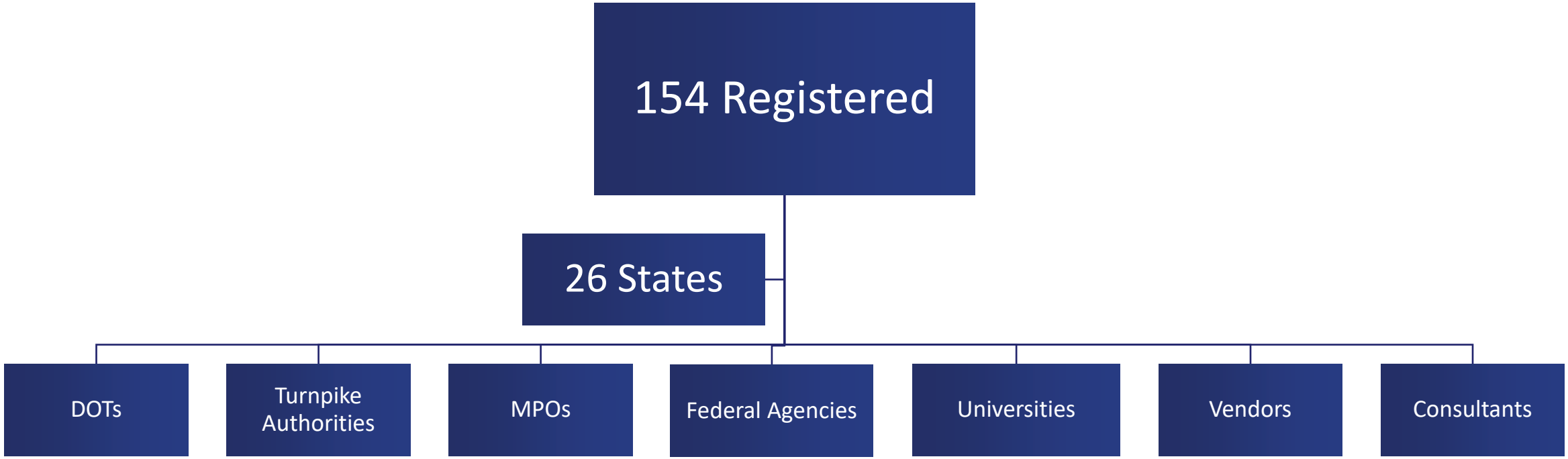
- Please raise your hand *(click on the hand icon at the bottom of the screen)*, and a host will unmute you.

- Please give your name and agency before asking your question
- **Please mute yourself when you are finished asking a question**

Welcome

Denise Markow, TSMO Program Director
The Eastern Transportation Coalition

The Eastern Transportation Coalition Sponsored Event



Coalition Update

RECENT

- ✓ **Coalition-wide Strategic Planning Web Workshop** - June 8, 2021
- ✓ **WAZE Technical Bi-Annual Working Group** - June 15, 2021
- ✓ **VPP-Traffic Data Marketplace State POC Meeting** - June 22, 2021
- ✓ **RITIS User Group Web Meeting** - July 15, 2021

UPCOMING

- ➢ **Traffic Data Marketplace - RFP Evaluation Meetings** - Aug, Sep, Oct 2021 *(member invite only)*
- ➢ **Meet the Work Zone Data Exchange Project Web Meeting** - September 9, 2021
- ➢ **RITIS User Group Web Meeting** - September 30, 2021
- ➢ **Travel Info Mapping Technical Meeting** - October 7, 2021 *(member invite only)*
- ➢ **TDADS National Webinar** - November 10, 2021
- ➢ **WAZE Technical Bi-Annual Working Group** - November 18, 2021 *(member invite only)*
- ➢ **VPP-Traffic Data Marketplace State POC Meeting** - December 7, 2021 *(member invite only)*

Just a reminder: Use the **Q&A box** for questions

- Use the **Q&A box** to ask presenters questions
- Ex. “How accurate is the captured data?

- Use the **chatbox** for technical issues or to contact Coalition staff
- Ex. “I can’t hear the presenter”

Introductions & Background

Stan Young, Chief Data Officer
The Eastern Transportation Coalition

Agenda

Topic	Speaker
Welcome from the Coalition	Denise Markow, The Eastern Transportation Coalition
Background	Stan Young, Chief Data Officer, The Eastern Transportation Coalition & Research Scientist, NREL
Connecting the Dots: Pennsylvania's Conflation Experience	Steve Gault, Chief of TSMO Arterials & Planning Section, Pennsylvania DOT
Probe Segments and LRS: A Tale of Two Giants and How Ohio Made Them Friends	Bill Welch, Program Administrator, Data Analytics and Automation, Ohio DOT
Conflation Software and Its Methodology in Real World Applications	Fred Hejazi, CEO, Citygate GIS
Wrap Up and Remaining Questions	Stan Young

Speaker Introductions

Steve Gault
Chief, TSMO Arterials & Planning Section
Pennsylvania DOT

Bill Welch,
Program Administrator
Data Analytics and Automation
Ohio DOT

Fred Hejazi
CEO
City Gate GIS


THE EASTERN
TRANSPORTATION
COALITION
CONNECTING FOR SOLUTIONS

Background

It is the best of times ….

- ❑ Digital Mapping Companies
- ❑ Traffic information disseminated through industry
- ❑ More and better info for DOT's and travelers than we have ever had

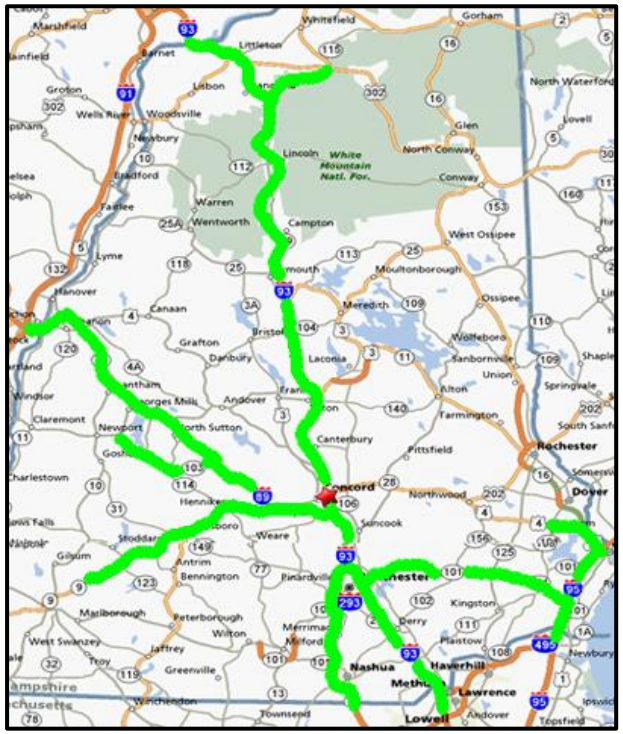
INRIX

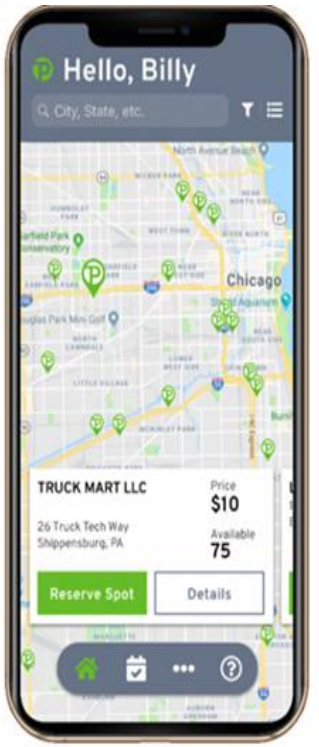
here

Google

WZDx
WORK ZONE DATA EXCHANGE

TOMTOM

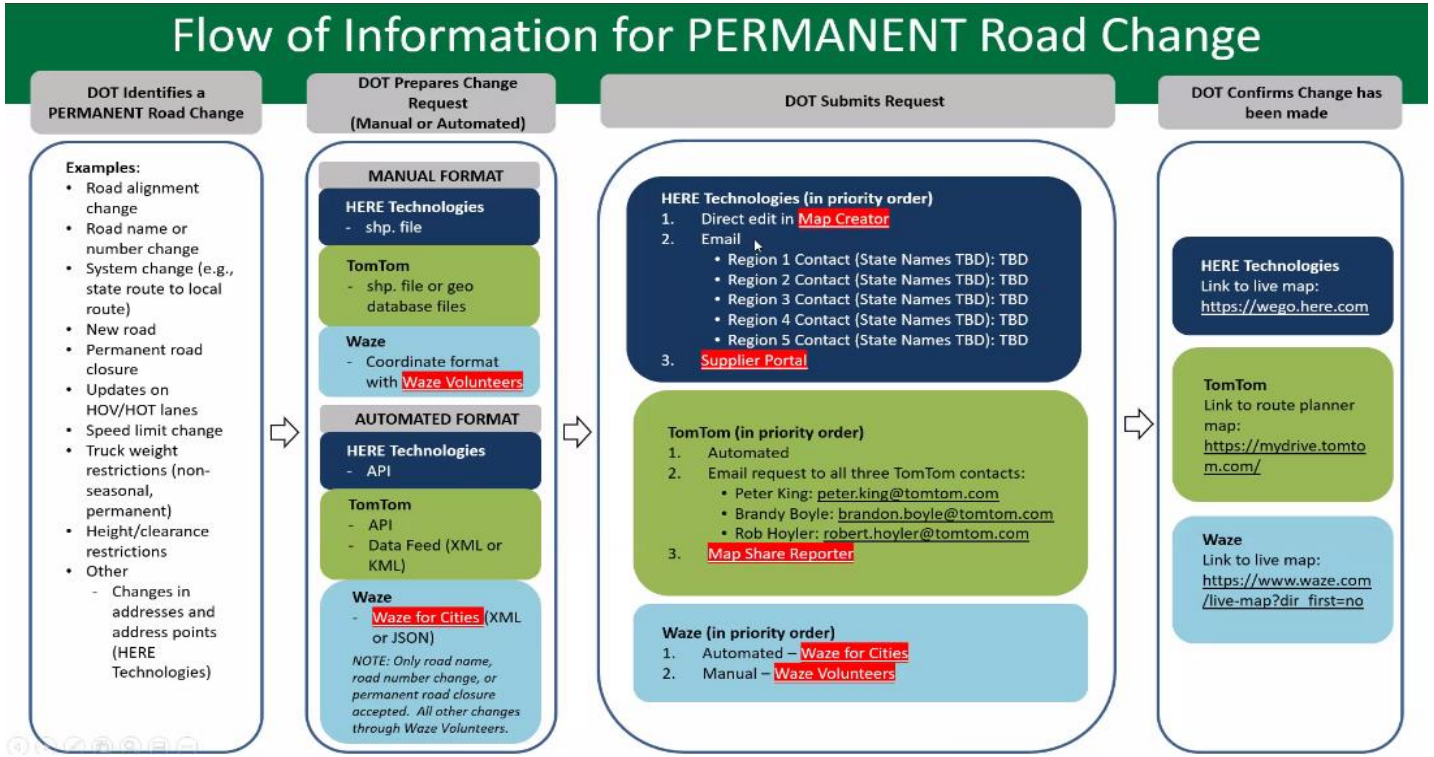




It is the ~~worst~~ (most challenging) of times ….

a. Just for permanent mapping changes – non-real time

Each maps change requires interaction with each Map Company tool to update changes



Source: Enterprise Pooled Fund Study

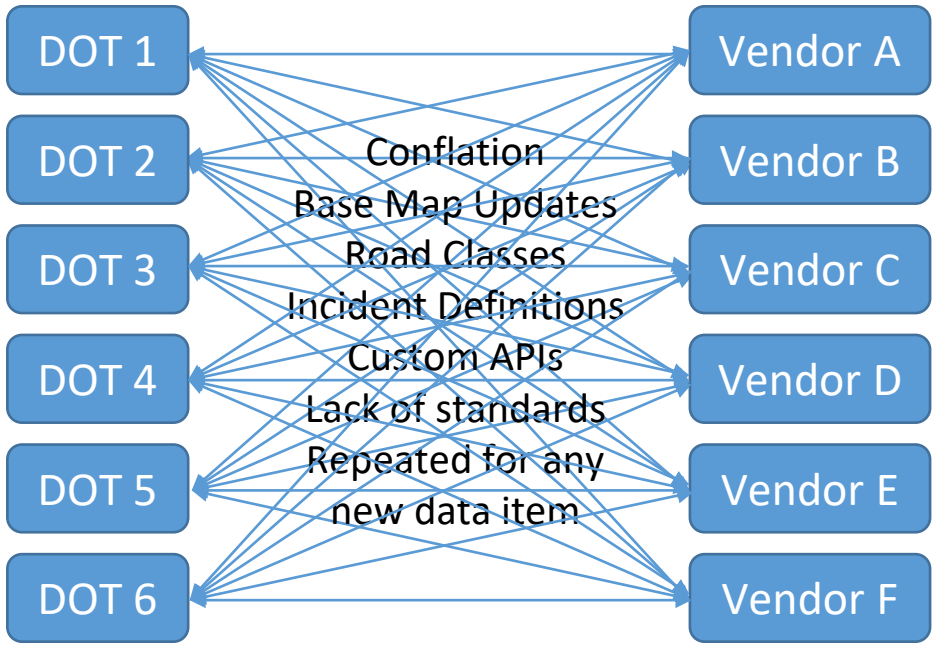
How about b. Work Zones (real-time) c. Incidents (real-time) d. Posted bridges e. ????

Each data set currently requires custom, vendor specific interaction

Common Digital Exchange Formats

- Our mission is to build effective common data exchange formats and sources so that all mapping companies have one location or portal to pick up the needed information- *“everyone drinking from the same fire hose”*

Today’s World

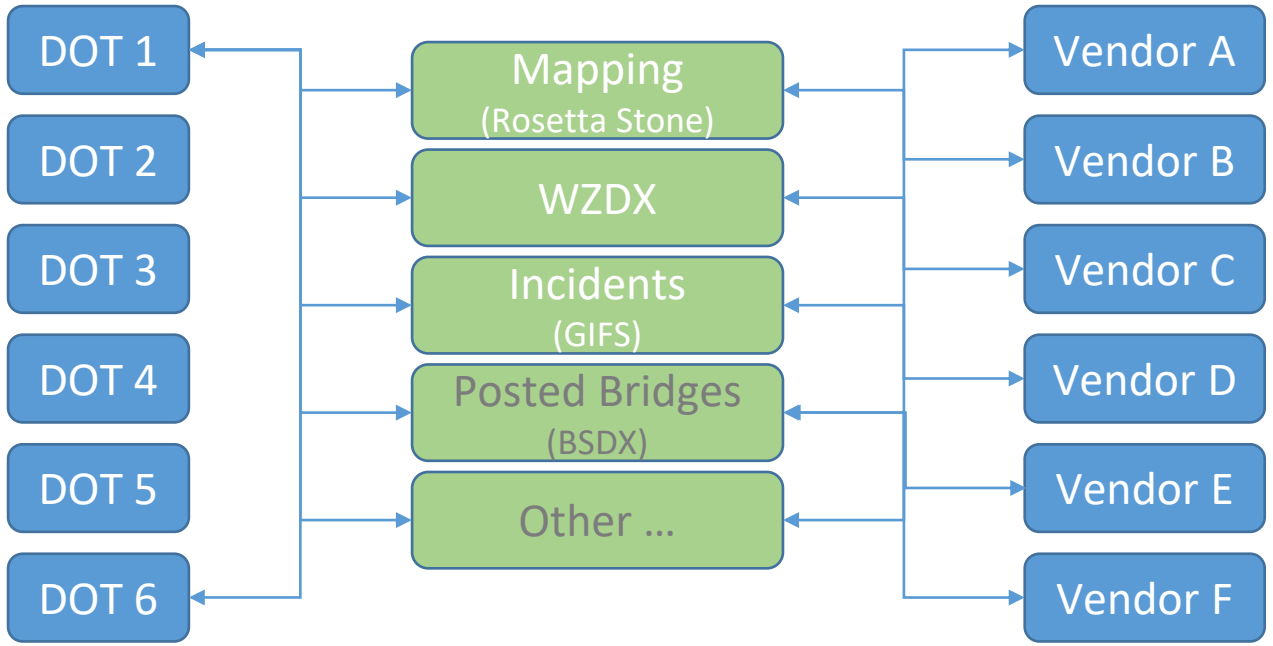


Complexity = Vendors X Data

Common Digital Exchange Formats

- Our mission is to build effective common data exchange formats and sources so that all mapping companies have one location or portal to pick up the needed information- *“everyone drinking from the same fire hose”*

Tomorrow’s World



Complexity = # of Defined Data

Spatial information – need for Conflation

State 1 LRS

State 2 LRS

:

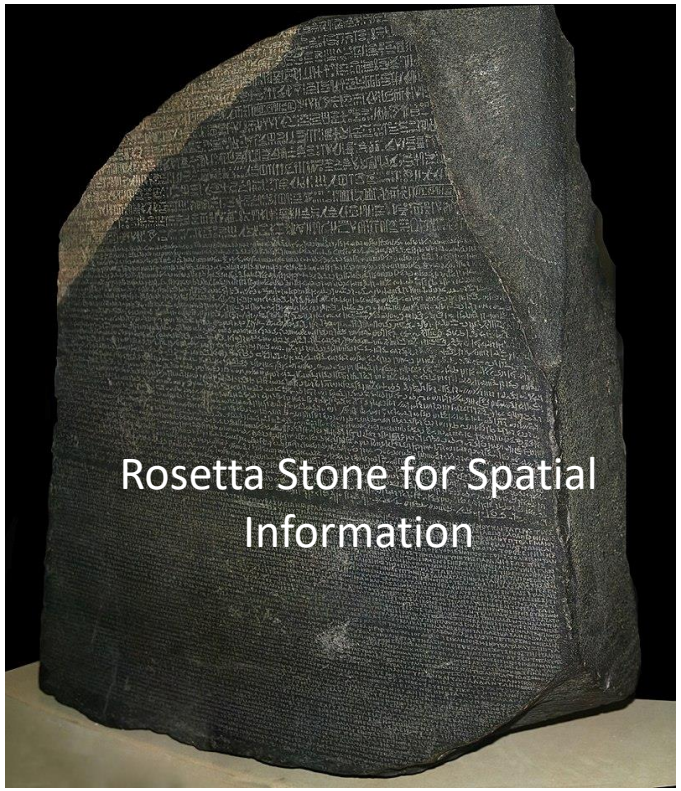
:

State 50 LRS

US Census Bureau

USGS

ARNOLD



HERE

TomTom

OSM

Google, Apple

TMC Flavor A

TMC Flavor B

Connecting the Dots: Pennsylvania's Conflation Experience

Steve Gault, Chief, TSMO Arterials & Planning Section
Pennsylvania DOT

CONNECTING THE DOTS:
PENNSYLVANIA'S CONFLATION EXPERIENCE

STEVE GAULT, P.E., PTOE • TSMO ARTERIALS & PLANNING • AUGUST 19, 2021

pennsylvania
DEPARTMENT OF TRANSPORTATION

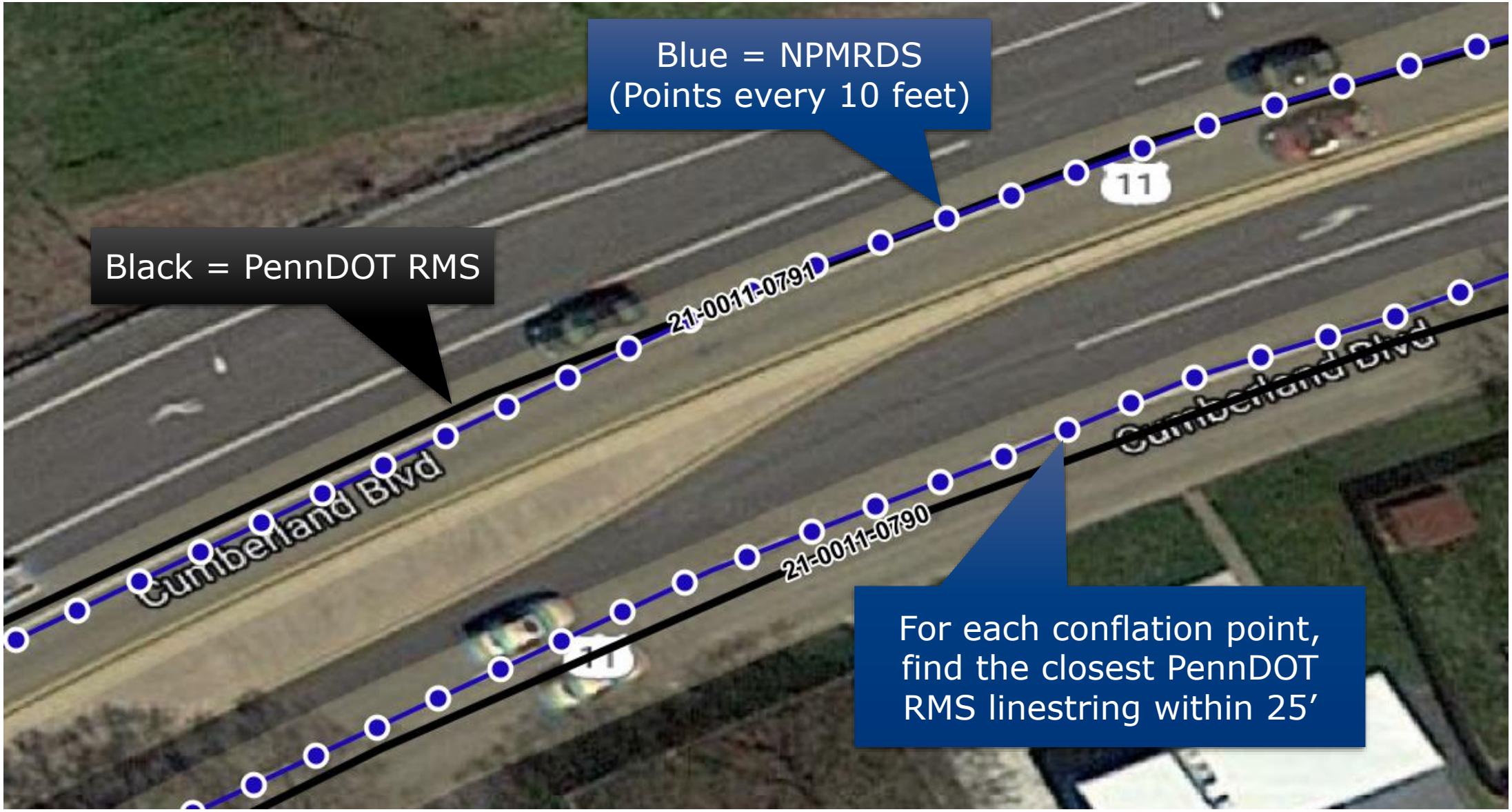
What's the Challenge?

- PHED measure requires speed limit to calculate delay
- Several measures use AADT; PennDOT discovered errors with the HPMS conflation process used for the NPMRDS data set
- PennDOT's most recent data for speed limits and AADT are in our Roadway Management System (1980s mainframe database also available in GIS shapefiles)
- Segment limits
  - NPMRDS TMCs: break at interchange ramps, major intersections
  - PennDOT LRS:
    - approximately ½ mile segments, break at major landmarks such as intersections or bridges
    - Interstate segments are always ½ mile and do not consider ramps
- Geometry doesn't match to exactly conflate points
- No common attributes to join datasets

Goal

- Provide one speed limit and one AADT which is representative of the entire NPMRDS TMC segment
- Produce CSV file with TMC code and speed limit/AADT
- Send CSV file to intake@nmprds.org to override defaults

Basic Conflation Process



Tally Conflation Points

TMC Segment	Conflation Point #	RMS County	RMS State Route	RMS Segment
101+0001	1	21	0011	0790
101+0001	2	21	0011	0790
101+0001	3	21	0011	0800

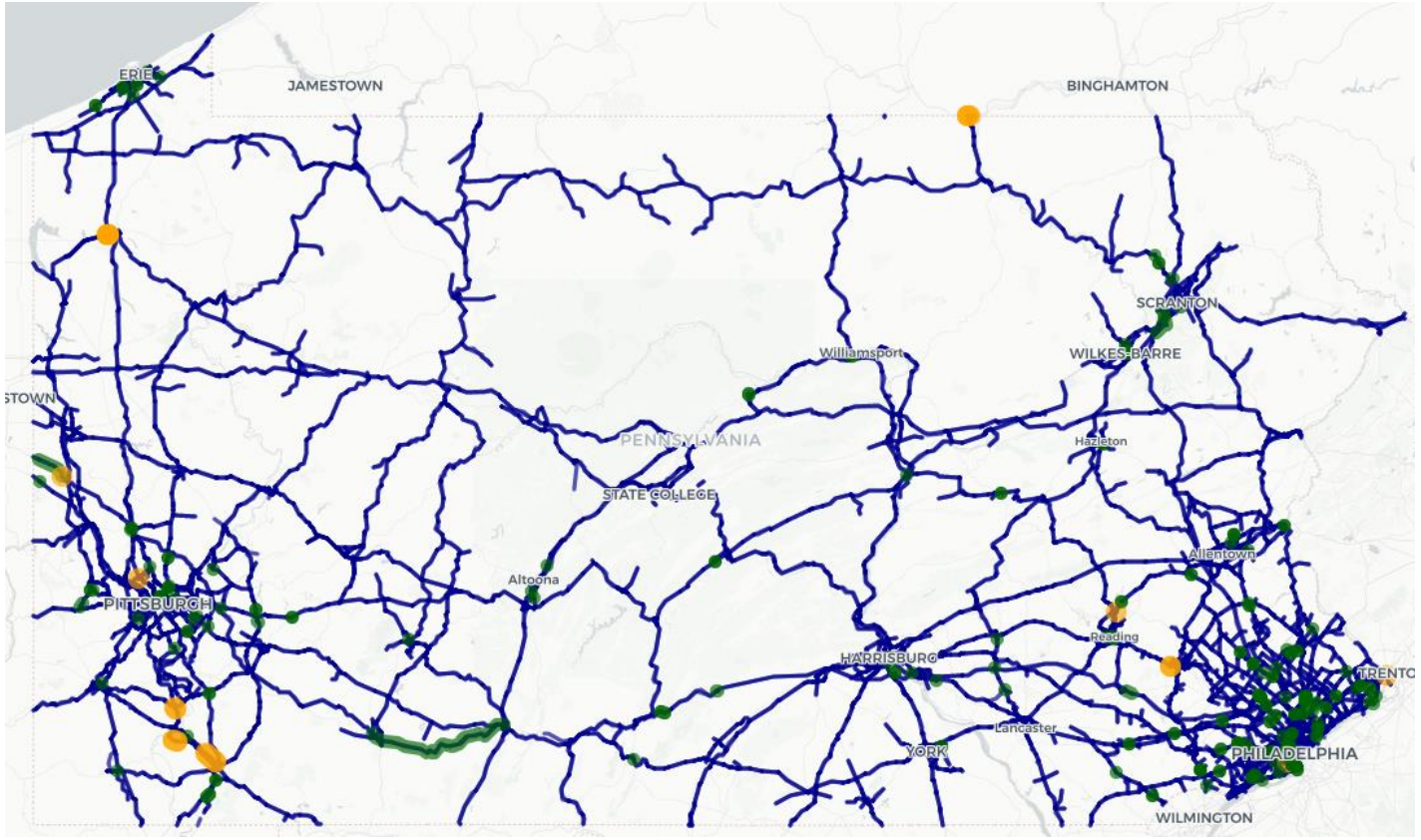
- Tally each RMS Key (County/State Route/Segment) – TMC combination
- Distinct selection of TMC id's, sorting by most occurring RMS key
- RMS key with the most matches for a TMC id has attributes applied (21/0011/790 in the example above)

Challenges

- Undivided highways
  - PennDOT RMS uses one linestring to represent BOTH directions
  - NPMRDS has separate lines for each direction
  - Need to divided PennDOT AADT in half to provide directional value
- Non-state routes on NHS in PennDOT RMS
  - Pennsylvania Turnpike (one linestring, bidirectional) → convert to parallel linestrings first
  - Locally-owned roads on NHS (may not have speed limits)
- Accuracy statistics
  - Number of conflation points per TMC id
  - Number of conflation points associated with 'winning' RMS seg
  - Percentage of 'winning' RMS coverage
- Goal is to minimize manual work – define **good enough**

QA/QC

- Created webmap with problem locations
- https://tmp-map.s3.amazonaws.com/tmc2020/npmrds-rms-aadt-mar2020-round2.html



Closing Thoughts

- Lessons Learned
  - More conflation points ➔ better results (10’ vs. 100’)
  - Weighted average attributes vs. “winning” segment data didn’t provide much benefit for the additional computation effort
- Further discussion
  - Formalize method to “override” assigned values with corrected values based on reviews
  - Using conflation to better link other data to help identify congestion causes or details
    - PennDOT Roadway Condition Reporting System (RCRS)
    - Weather
    - Congestion pie chart by road segment

Questions?

pennsylvania
DEPARTMENT OF TRANSPORTATION

Steve Gault, P.E. PTOE
Chief, TSMO Arterials & Planning
Pennsylvania Department of Transportation
sgault@pa.gov
717.787.6988

Probe Segments and LRS: A Tale of Two Giants and How Ohio Made Them Friends

Bill Welch, Program Administrator, Data Analytics and Automation
Ohio DOT

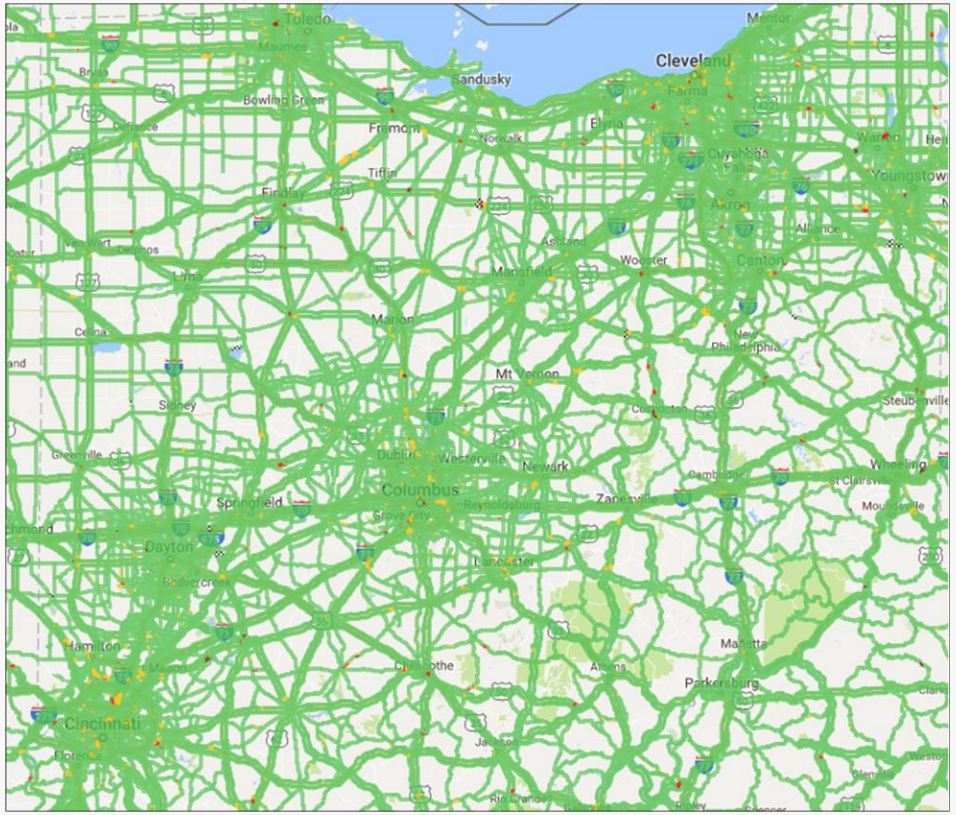
Probe Segments and LRS

A Tale of Two Giants and How Ohio Made Them Friends

odot

Probe Data in Ohio

- INRIX XD data for all of Ohio
- ~170,000 segments
- Supports TMC, operations analytics, maintenance operations



Ohio’s LRS System

- 121,000+ centerline miles
- 215,000+ segments

Ohio’s LRS System

121000

- 121,000+ centerline miles
- 215,000+ segments

- County-divided and statewide versions
- Primarily use county system

Ohio’s LRS System

121000

- 121,000+ centerline miles
- 215,000+ segments

- County-divided and statewide versions
- Primarily use county system

- Split identifiers by direction for divided highways
- Single identifiers otherwise

Ohio’s LRS System

121000

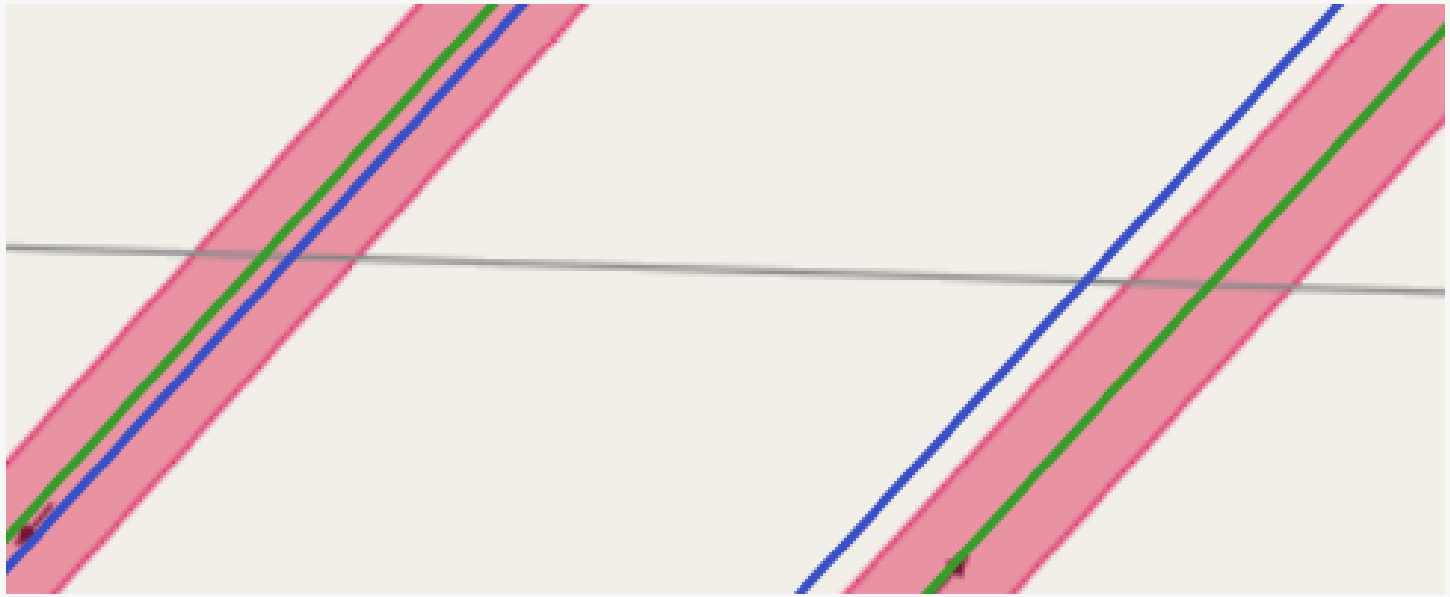
- 121,000+ centerline miles
- 215,000+ segments

- County-divided and statewide versions
- Primarily use county system

- Split identifiers by direction for divided highways
- Single identifiers otherwise

Annual updates

Challenge with Probe Data



Green: LRS
Blue: INRIX

- LRS segments and probe segments don’t always overlap
- Probe segments could cross multiple LRS segments
- LRS updates

Initial Use Cases

Snow and Ice Performance Evaluator (SNIPE)

Weekly

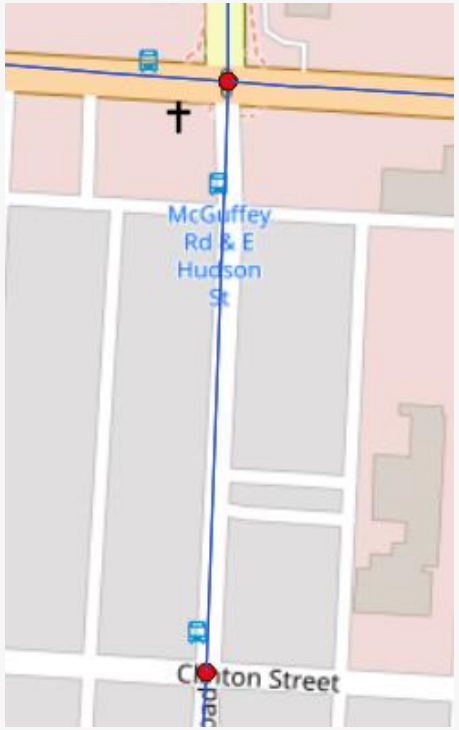
Traffic Operation Assessment Systems Tool (TOAST)

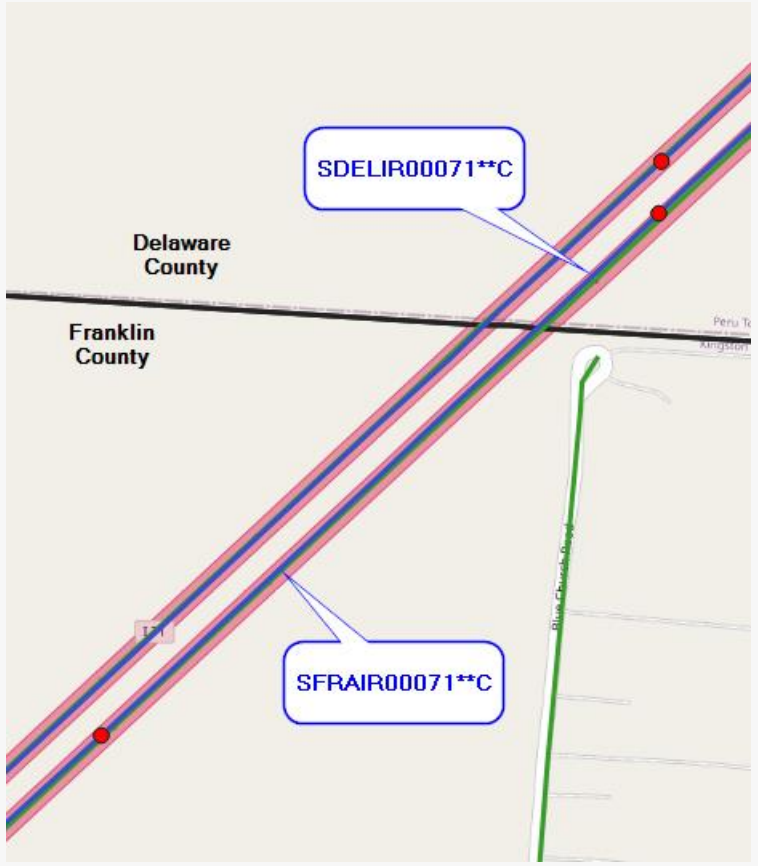
Yearly

Snow and Ice Priority Routes	Crash Reports	ATMS Incidents	Road Weather Information
Travel Time Performance	Bottlenecks	TOAST Corridors	Traffic Volumes

Our Solution: Three Versions

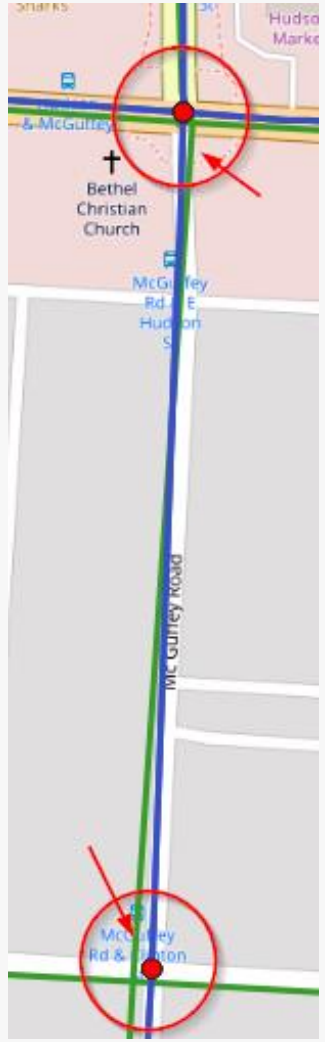
	Original Method (19.2 and earlier)
Matching Method	Start and End latitude/longitude
Handles start/end on intersection?	X
Handles cross-county segments?	X
Handles ramps?	X
Handles loop roads, roads start/end on same road?	X





Our Solution: Three Versions

	Original Method (19.2 and earlier)	Secondary Method (20.1) [addon to orig. mtd.]
Matching Method	Start and End latitude/longitude	Neighborhood around start and end latitude / longitude
Handles start/end on intersection?	X	Better but imperfect
Handles cross-county segments?	X	Better but imperfect
Handles ramps?	X	X
Handles loop roads, roads start/end on same road?	X	X

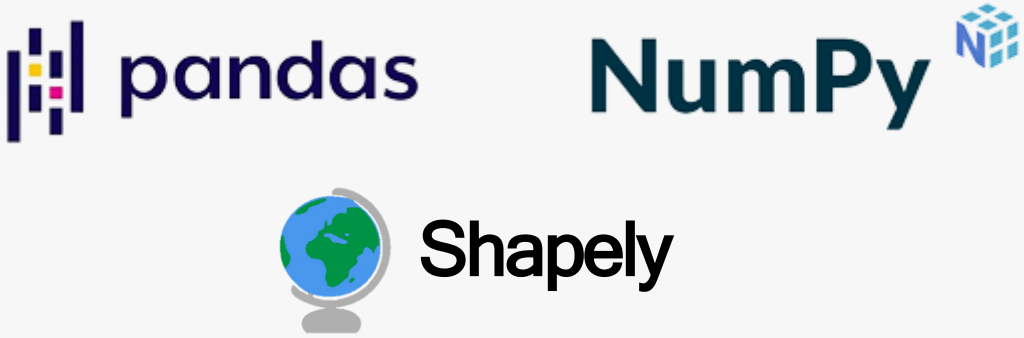


Our Solution: Three Versions

	Original Method (19.2 and earlier)	Secondary Method (20.1) [addon to orig. mtd.]	Sampling Method (20.2 and on)
Matching Method	Start and End latitude/longitude	Neighborhood around start and end latitude / longitude	10-point sample along INRIX geometry; identify NLFID with majority frequency
Handles start/end on intersection?	X	Better but imperfect	✓
Handles cross-county segments?	X	Better but imperfect	✓
Handles ramps?	X	X	✓
Handles loop roads, roads start/end on same road?	X	X	✓

Our Solution: Sampling Method

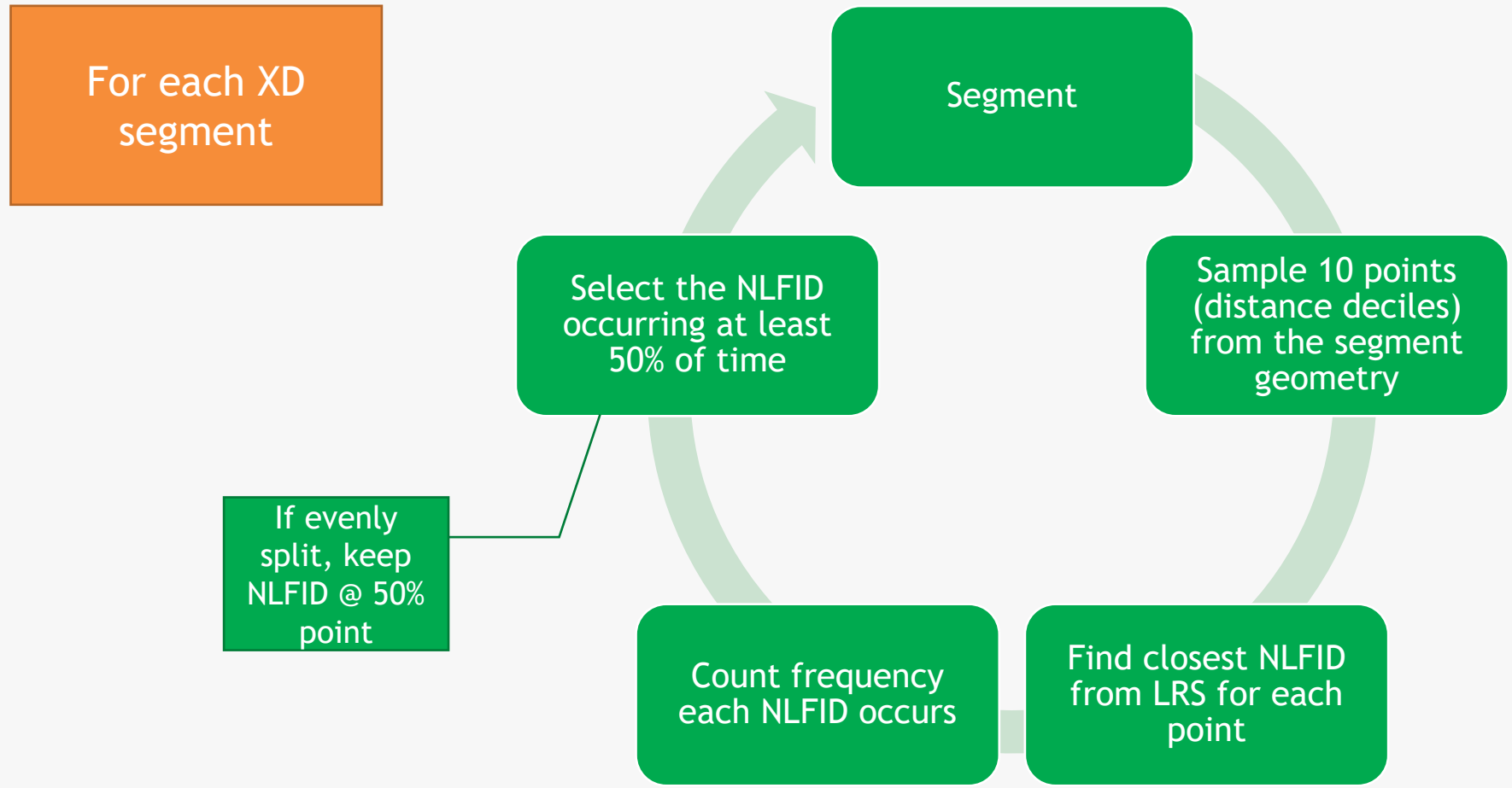
NLFID Identification → Logpoint Identification → Join and save



- Python script handles each of the three main steps
  - High-quality libraries for data analysis, numerical analysis, and geospatial analysis simplify development
- SQL Server is our database platform for most things
  - Geospatial queries for getting data
  - Output stored in a table in our TSMO Warehouse
  - Most modern database systems have the capabilities for this

Our Solution: NLFID Identification

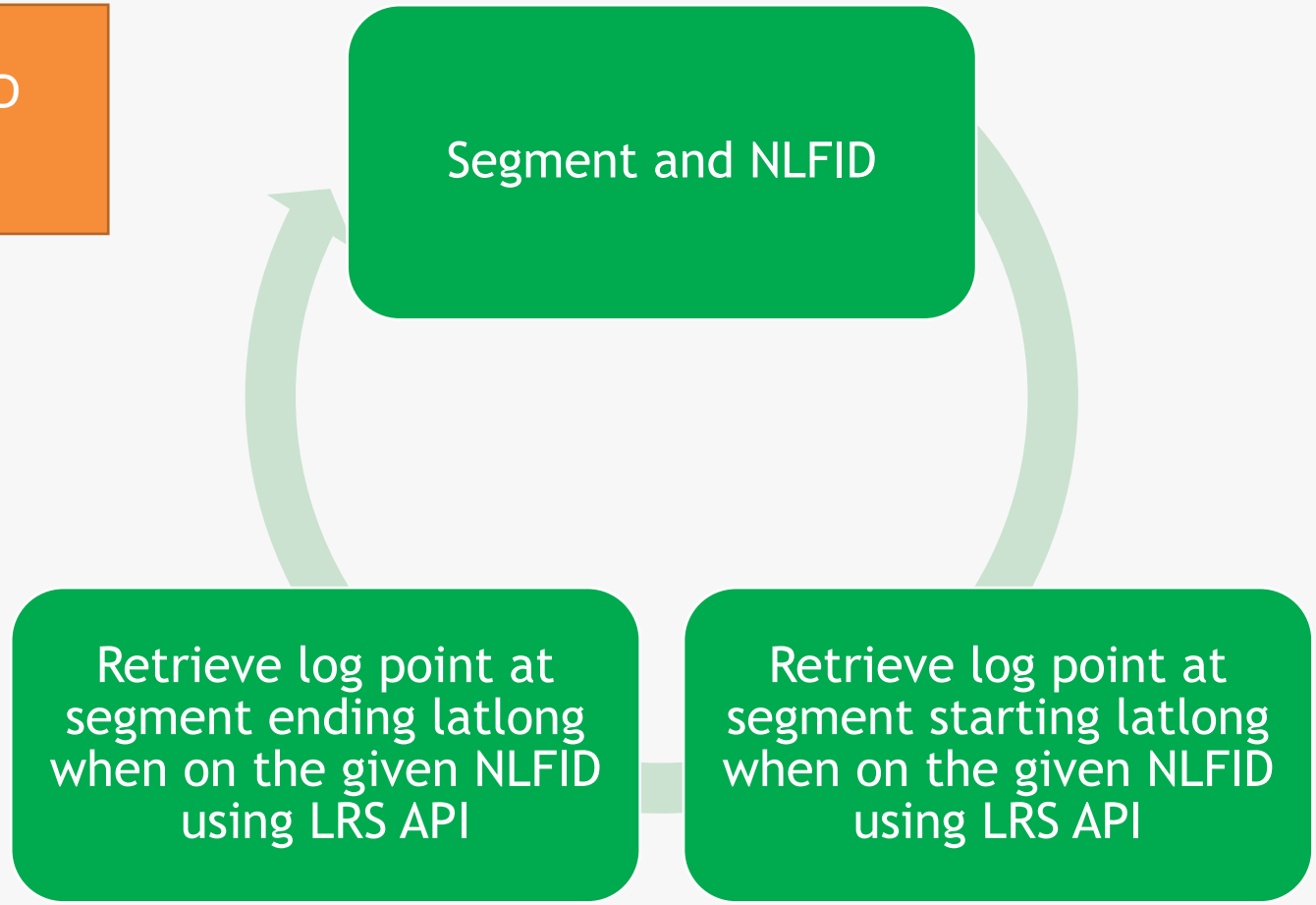
NLFID Identification | Logpoint Identification | Join and save



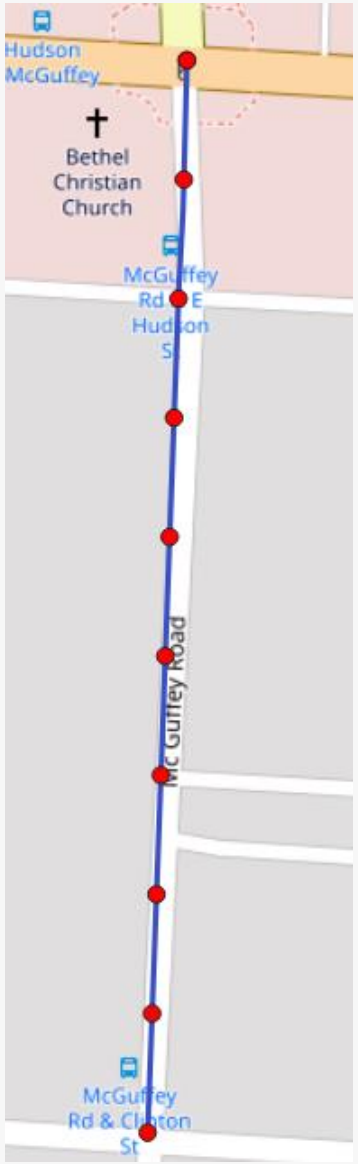
Our Solution: Logpoint Identification

NLFID Identification | Logpoint Identification | Join and save

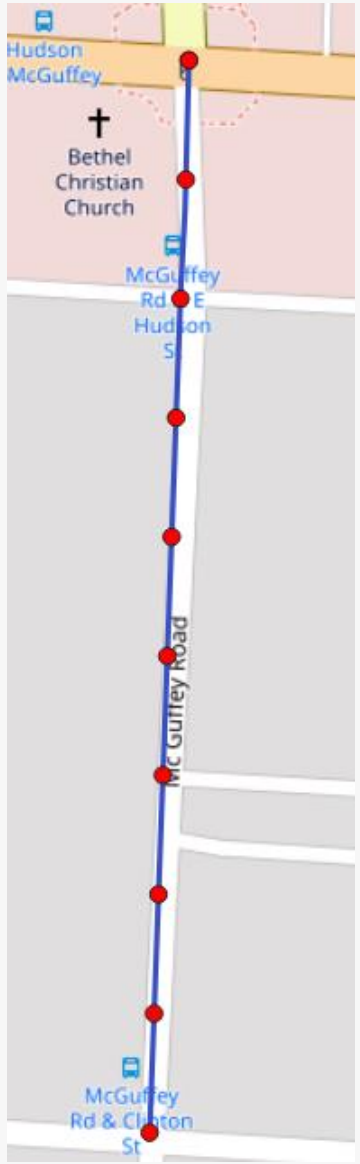
For each XD segment



Our Solution: Example

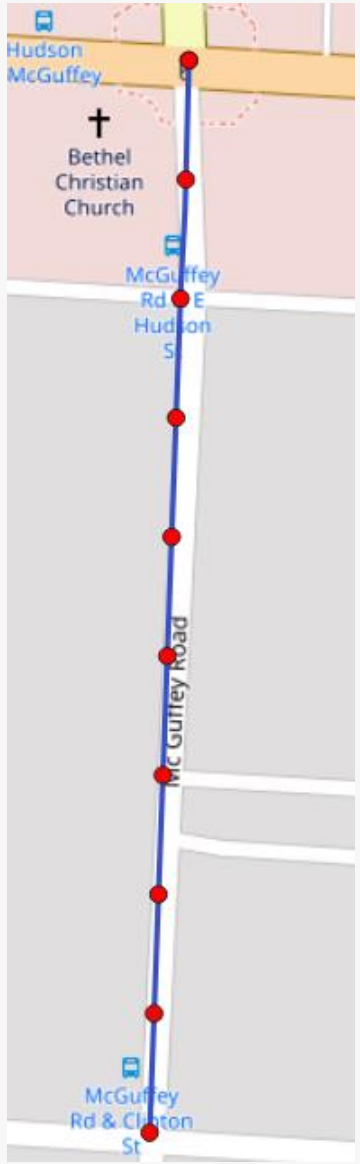


Our Solution: Example



NLFID	Count
MFRAMR04237**C	1
MFRAMR01342**C	8
MFRAMR00451**C	1

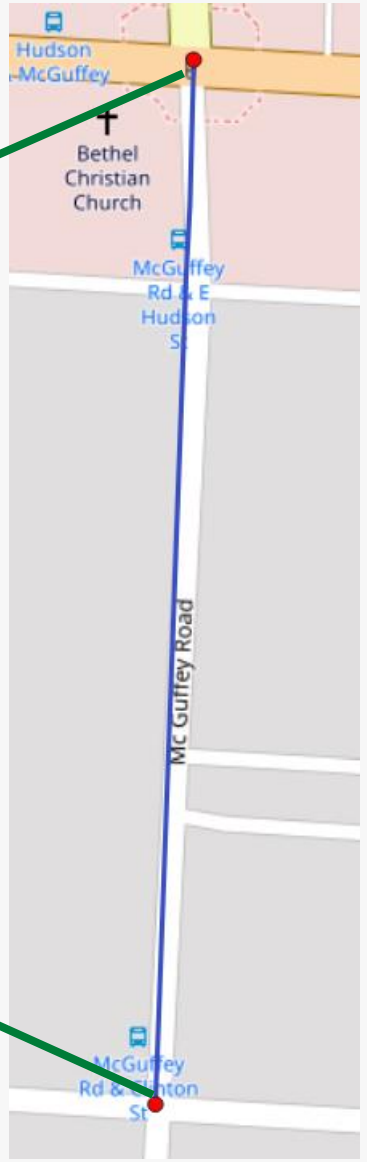
Our Solution: Example



NLFID	Count
MFRAMR04237**C	1
MFRAMR01342**C	8
MFRAMR00451**C	1

0.317 on MFRAMR01342**C

0.138 on MFRAMR01342**C



Initial Development and Implementation

~30 hours

Me

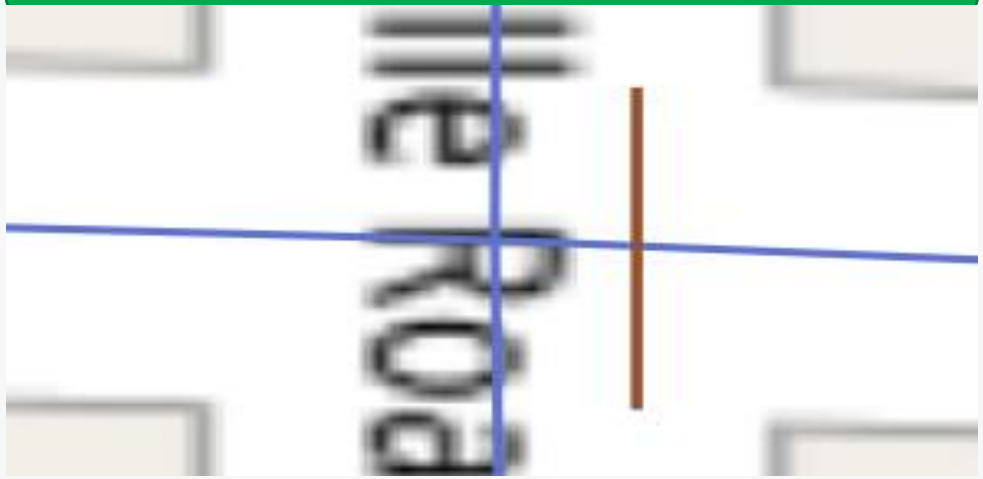
$0 additional outlay

Additional Enhancements

Performance

- Initial version: 12 hour runtime
- Addressing inefficiencies in coding reduced to 1.5 hours

Accuracy

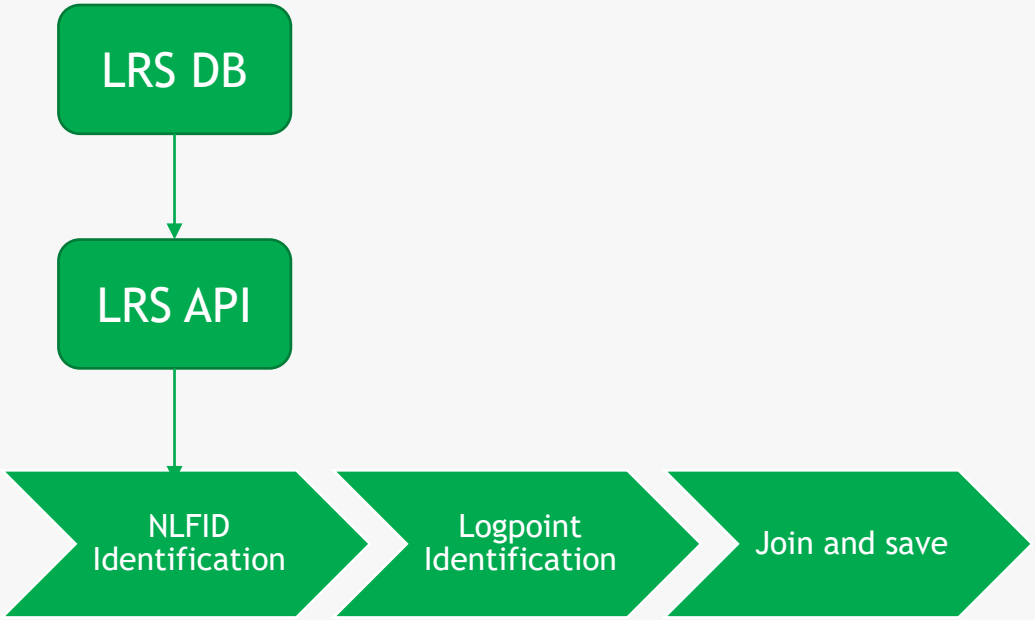


- For very short segments, sometimes the “closest” LRS segment was a cross-street
- Used direction of segments to ensure we matched to the right segment (vector analysis)

Additional Enhancements: Performance

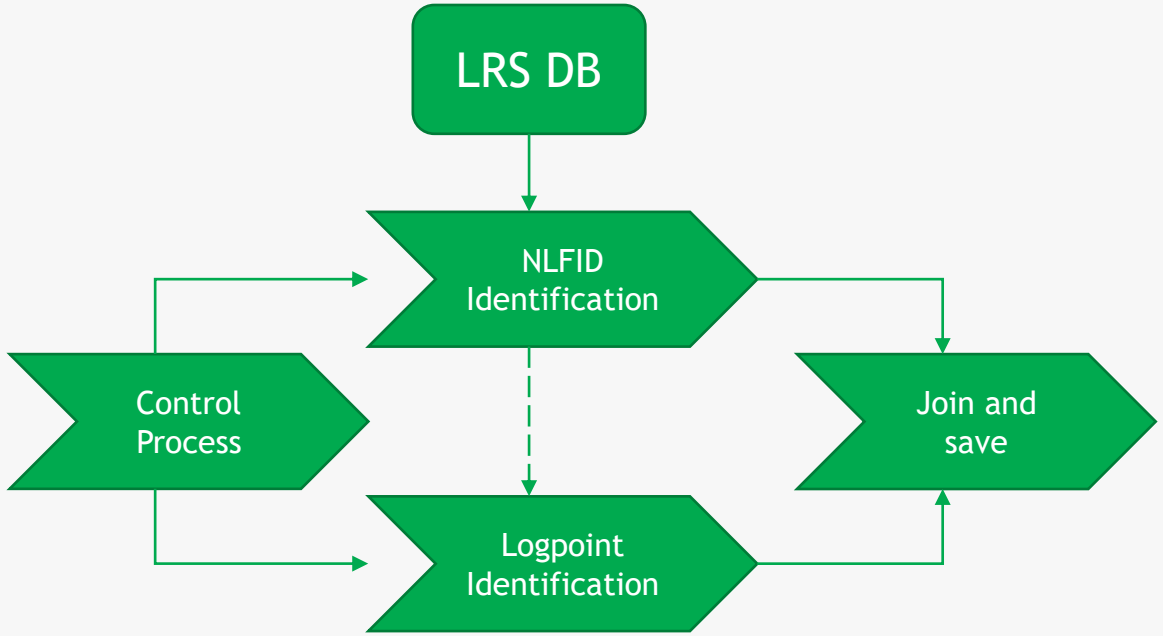
Before

- ✓ Sequential execution: must wait for *all* NLFIDs to be identified before identifying any logpoints
- ✓ Use LRS API to get NLFIDs



After

- ✓ Parallel execution: logpoints identified as NLFIDs identified
- ✓ Connect to LRS DB directly

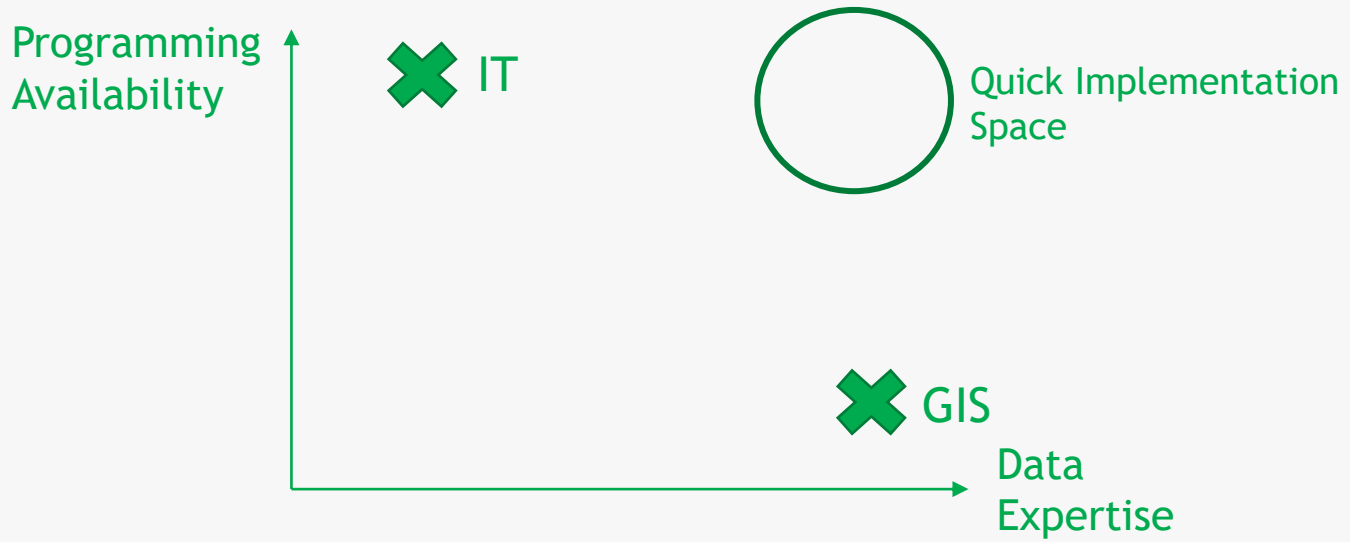


Next Steps

- Migrate away from LRS API for logpoint identification (do in script)
- Increase number of logpoints sampled (next target: 20)
- More QA processes
  - QA is our weakest area
  - Current: “Random” sample of ~500 segments, manual checks of known problem areas
  - Limitation: Staff time

Lessons Learned

- Conflating to LRS *can* be solved in-house
  - Speed of implementation is heavily dependent on internal environment
  - Time vs. cost tradeoff when deciding to make or buy



Lessons Learned

- We could have had a solution sooner if we had discussions with other states
- Never underestimate the power of 8:30 PM random thoughts ☺

Bill Welch, MPA
Program Administrator, Data Analytics and Automation
Ohio DOT
William.Welch@dot.ohio.gov

Conflation Software and Its Methodology in Real World Applications

Fred Hejazi, CEO
City Gate GIS

Conflation Technology

Development and deployment of automated conflation

By Fred Hejazi PLS

citygateGIS
Intelligence Mapped

Background

- Citygate GIS is a geospatial software development company
- The company was started in 1986 located in Annapolis, Maryland
- ESRI Business partner since 1999
- Primarily develop application on the desktop and the web using ESRI and Google Maps
- Fred Hejazi was the software architect for Conflex, Citygate's conflation technology
- Also the project manager for a 6 year conflation project for the US Army Geospatial Center

Background in Conflation

- Software was born from manual conflation of 1990 to 2000 Tiger files
- Project goal was to create a completely autonomous conflation technology
- Rather than basing conflation on proximity the approach was to create a statistical best match
- In 2004 Conflex was introduced as a product as the ESRI UC
- A file was submitted to us for processing during the Conference
- Several months later we were invited to meet with US Army and told they had been working with NGC for 6 years and the had never broken 65% match. Our approach exceeded 88%.

Technical Approach

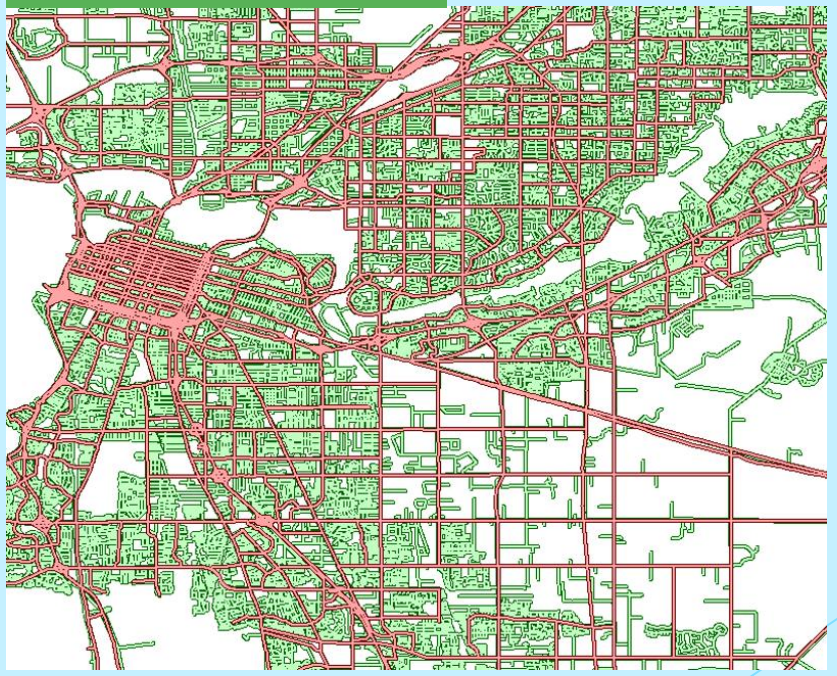
- At the time most conflation techniques relied on proximity
- You would buffer one set of lines then use polygon to line overlay to identify the matching lines
- Process did not work where geometric errors were inconsistent or
- If the scale between data layers varied greatly
- Our approach used a statistical best match, similar to how a person determines matches between two layers
- Matches were statistically matched based on:
  - Proximity
  - Network connectivity
  - Directionality
  - Length
  - Feature types

Technical Approach

- This approach allowed us to solve many types of Conflation problems

Scale Difference

Density Difference



Technical Approach

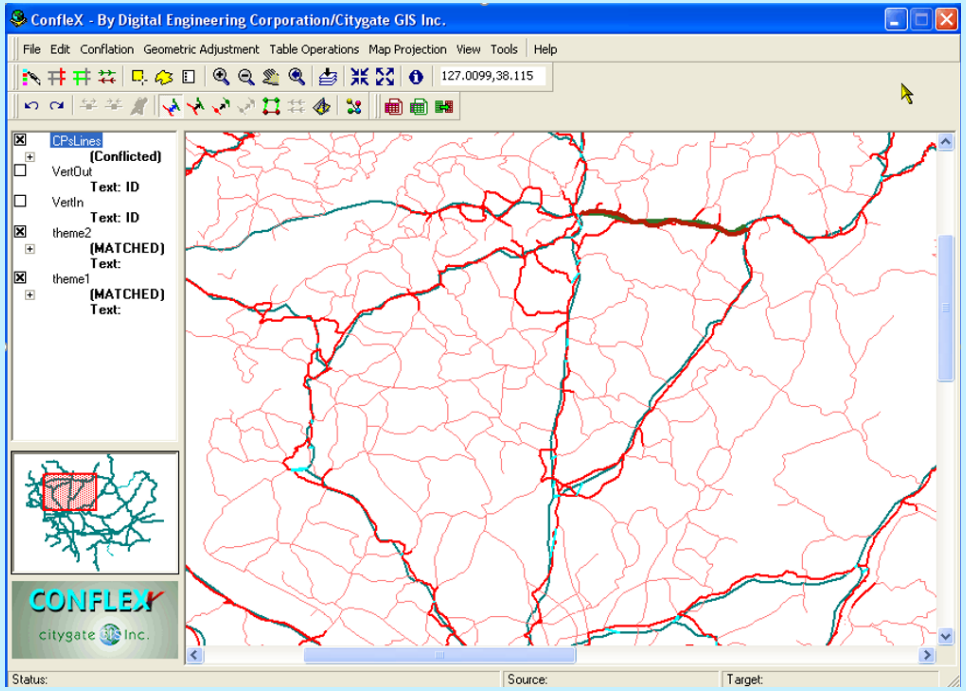
- This approach allowed us to solve many types of Conflation problems

Accuracy Difference

Source Difference

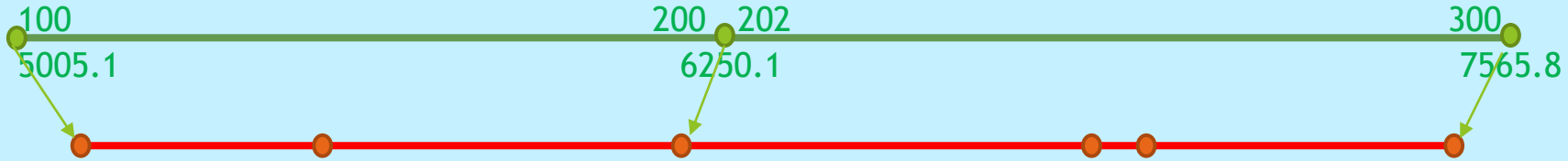
Technical Approach

- Our work with the US Army continued for 6 years
- The product that was developed was called ConfleX
- It included a standalone and ArcGIS addon versions
- Primary focus was on creating a “Best Map”



Limitations of a Commercial Solution

- Automation of linear data conflation is highly project specific
- The Best Map approach applies to a narrow group of projects
- It provides rapid results with good accuracy, matches names, and key intersections
- It does not automatically solve linear referencing, route and name combinations or address ranges



- In a DOT application, 100% of the roads have to be verified as correct. That can not be assured without manual review

Our approach

- For off the shelf purchases allow customers to extensively test usability and applicability to their project
- Built an internal set of software code that are customized based on the project requirements
- Integrate with ArcGIS for manual editing
- Productivity is increased through semi-automated processing
- Provide automated routines with key manual guidance
- Iterative approach where in each cycle the software uses previous matches to make new ones

Questions?

Fred Hejazi,CEO

CityGate GIS

fhejazi@citygategis.com

Wrap Up

Stan Young, Chief Data Officer
The Eastern Transportation Coalition

Additional Questions

Meeting information & presentations will be posted to The Eastern Transportation Coalition website.
Participants will receive a link to the presentations after they are posted.

THE EASTERN TRANSPORTATION COALITION

CONNECTING FOR SOLUTIONS

THANK YOU!

For Questions or Additional Information, please contact:

Denise Markow, TSMO Director, *301-789-9088,* *dmarkow@tetcoalition.org*