Transport for London, using data to keep London moving

Roland Major
Enterprise Architect
Breadth of Transport for London

- 30 million journeys daily
- In addition to all road and rail transport, we look after rivers, assisted travel, taxi and private hire regulation
- We do much more also from education to running the world’s largest out of home advertising estate of its type
- We’re 150 years old and chock full of heritage and design assets
Conceptual Architecture

Key Characteristics

- Bringing information and insight closer to decisions
  - Locale, map based UI
  - Timely, event driven not batch
  - Trusted, consistent and accurate
- Public cloud hosted
- Reusable commodity platforms
- Open Standards
The Data Hub

**Commodity Cloud**
- Scalable Object Store S3
- Persistence Services (SQL and No SQL) Aurora + Dynamo
- Elastic Search
- Spark
- Compute EC2

**Configured Platform Services**
- Spatial - ESRI
- Data Warehouse - Oracle RDBMS + Redshift
- SOA Platform – WS02
We currently provide a number of (digital) road information including:

- **Incidents** - Live traffic information and planned works/events disruption. Includes planned closures.

- **Road status tool** – Web interface

- **VMS** (including some journey times), **JamCams, Post Code** impact areas and **Other Camera** locations.

Many users, applications and value added providers
Analytics and Visualisation

- Everything we monitor is linked in time and place
- Improvement in identifying effects and finding cause
- Balance of models and sensors
- Where you have multiple data sources how do you bring these together to inform you of new insight
- How feasible is prediction?
Operational Systems

- TfL has an operations center which coordinates on the day activities across Surface.
- Many Events are planned, some events just occur.
- Plan $\rightarrow$ Do $\rightarrow$ Review.

- Situational Awareness is important.
- Silo systems just do not work.
Waze Alerts Operations View

Situational Awareness

#StrataHadoop
Traffic Control Sensors

- We have about 14,000 sensors measuring junctions approaching junctions
- This data is currently used by the Real Time control to manage optimisation
- We have been investigating how it can be processed
  - Scale 780 Million events per day
  - Latency to data center circa 1 Second
  - Resolution 250mS scans
Traffic Sensors and derived flow
Key points

- Sensor data is great for modern analytics platforms
- Meta data preparation is the hardest part
- It’s completely anonymous
- Granularity of data gives early identification of incidents
- ML predictions out to 45 mins, are quickly running at >80% accuracy
- Complex network, understanding its dynamics needs a lot of work
Operation Technology

London Works 2

- **Central Register** – a pan London system enabling visibility and management of works and related activities in London

- **Traffic Management Act Notifications (TMAN)** - A dedicated interface between London boroughs and TfL enabling the balanced delivery of major schemes and works on the TLRN and SRN

- **Forward Planning Tool** - An advance planning tool that allows promoters to provide early visibility of road and street works
Does this activity include a permanent change to the network or major development? No

More than 1 Phase? No

Traffic Management Type: 1 Lane Remains Running

Overnight Works Only (22:00 - 05:00) Yes
Review

- Use Technology as a commodity
  - Open Source is a key enabler
  - IoT, Big Data, Cloud don’t resonate
  - Skills are the biggest hurdle to overcome
- Bringing information sets together encourages new thinking
- Agile changes the organisation
- Composable systems delivers solutions quickly
Thanks You

- Questions

- rolandmajor@tfl.gov.uk