
SCTE
CABLE-TEC EXPO
FAST FORWARD 2021
ATLANTA, GA
OCTOBER 11-14
SCTE
a subsidiary of CableLabs
UNLEASH THE POWER OF LIMITLESS CONNECTIVITY
2021 Fall Technical Forum
SCTE • NCTA • CABLELABS

SCTE
a subsidiary of CableLabs®

Wireline Access Network

Solving the Mysteries of the Distributed Access Architecture

Matthew Stehman

Data Scientist
Comcast

2021 Fall Technical Forum
SCTE • NCTA • CABLELABS

SCTE CABLE-TEC EXPO FAST FORWARD 2021

VIRTUAL EXPERIENCE
OCTOBER 11-14

Overview

1. Introduction
2. Overview of DAA telemetry at Comcast
3. Implementation of Sherlock (big data analysis architecture)
4. Event classification and rankings
5. Event examples
6. Machine learning applications
7. Conclusions

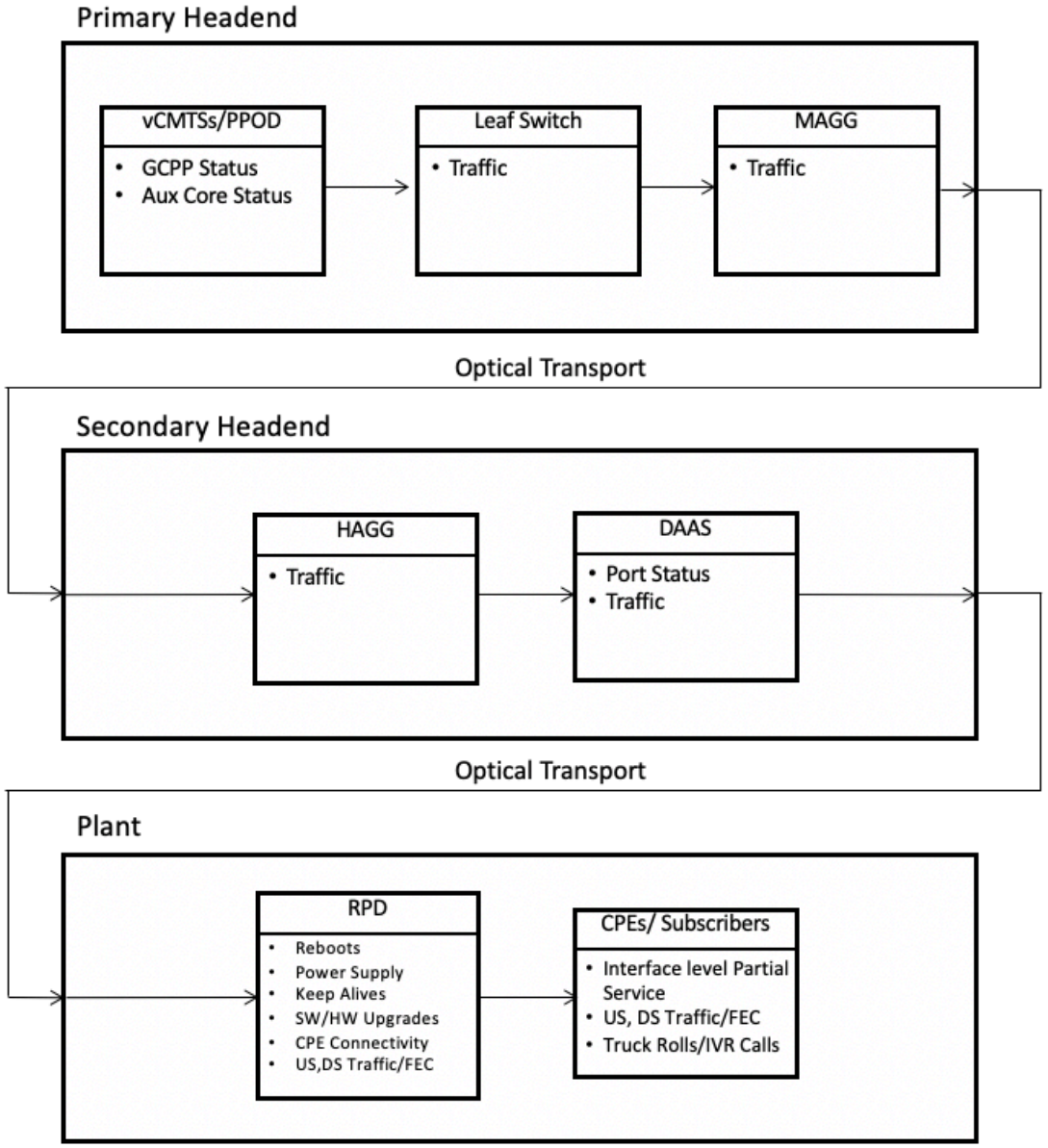
Problem Statement

- Comcast is continually deploying and scaling its distributed access architecture (DAA)
- DAA allows for extremely rich telemetry across all components in the network
- Advanced techniques are required to draw insights from the vast amount of telemetry data
  - Existing tools do not create a centralized view of the DAA network
  - No longer feasible to manually analyze all data

Develop an automated big data analysis platform that can analyze DAA data and provide actionable insights to the DAA deployment and operations teams at Comcast

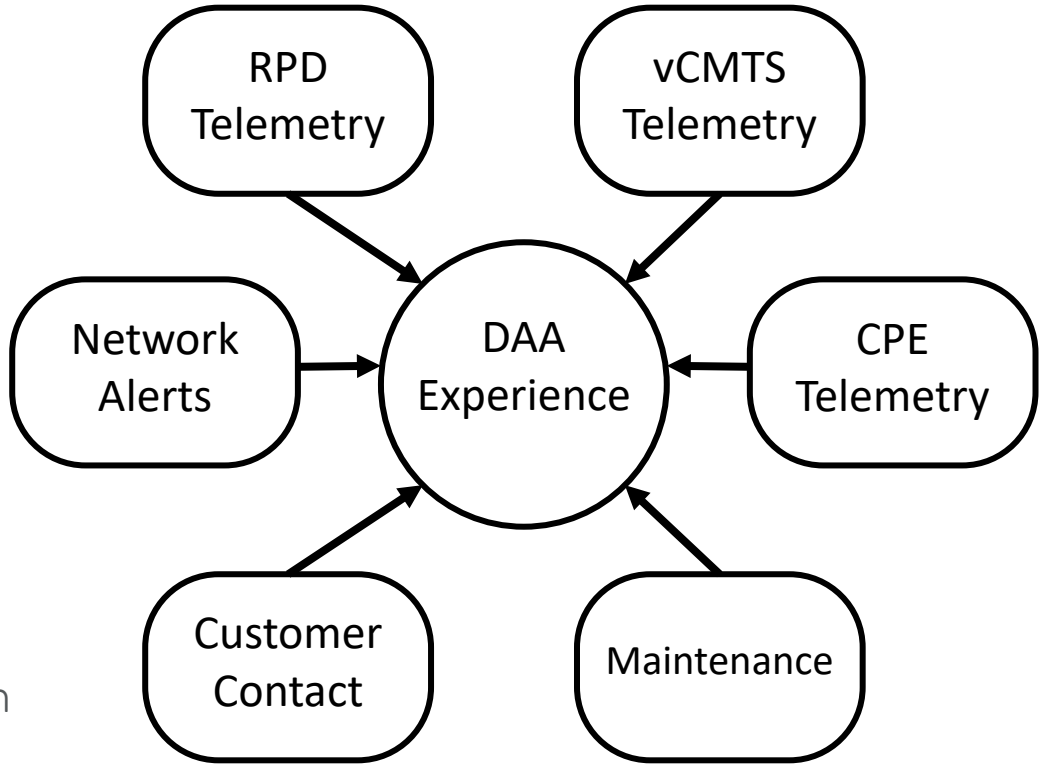
DAA Topology

- DAA's dedicated telemetry core captures real time measurements from the headend to the customer devices
  - Primary headend
  - Secondary headend
  - Plant
- The increased level of visibility over analog CTMS allows for constant monitoring of individual systems and root cause analyses for any issues



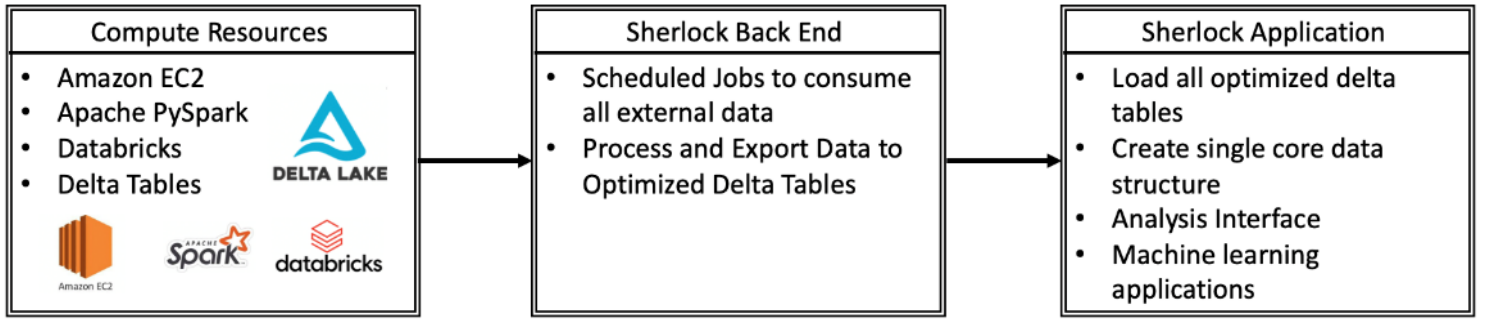
Telemetry Categories

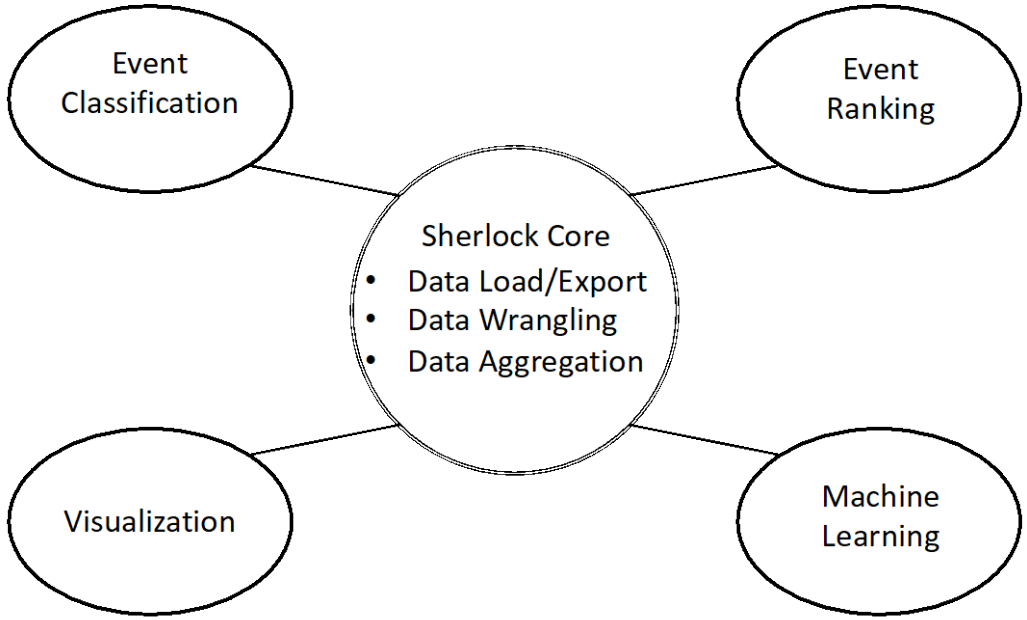
- The DAA core telemetry is combined with other Comcast data sources to provide a full view of the DAA customer experience
  - System statuses
  - RPD statuses and upgrades
  - Traffic
  - CPE connectivity/health
  - Signal quality
  - Customer contact
  - Existing Comcast network alerts
- Comcast DAA data generates over a billion data points per PPOD per day!
  - Raw data is captured every 15 seconds for most metrics
- Raw data needs to be aggregated and centralized in order to perform meaningful analyses
  - RPD-level aggregation
  - Five-minute frequency



Sherlock Overview

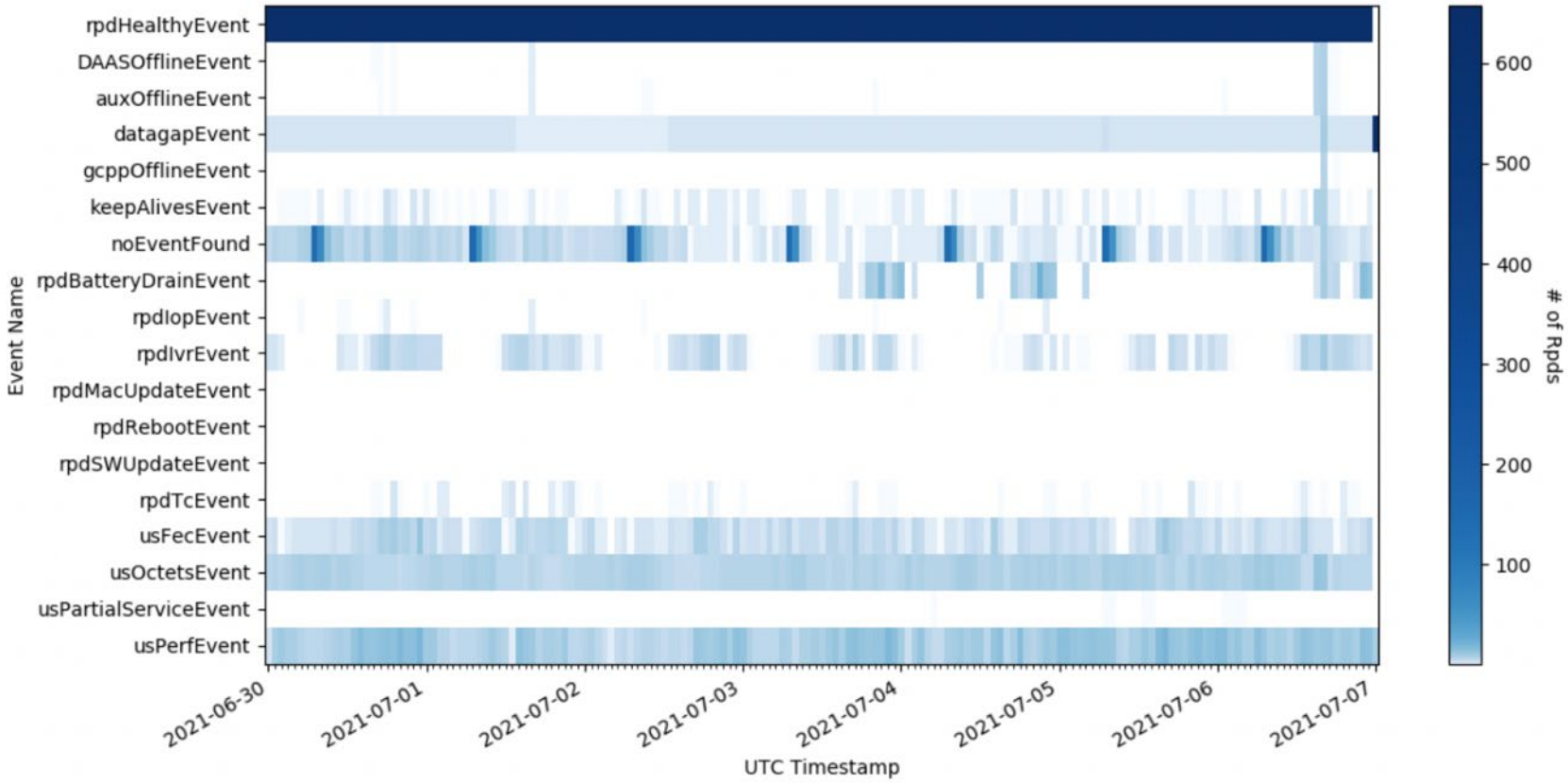
- Requirements
  - Full footprint coverage with future scalability
  - Scheduled and ad hoc analysis capabilities
  - Fast computation
    - Base libraries are not sufficient
  - Machine learning integration
- Features
  - Sherlock's core combines data from across Comcast into a centralized data structure
    - Single time-aligned table with all metrics at the RPD level
  - Auxiliary modules access the core data structure to perform dedicated analysis tasks





Event Detection and Ranking

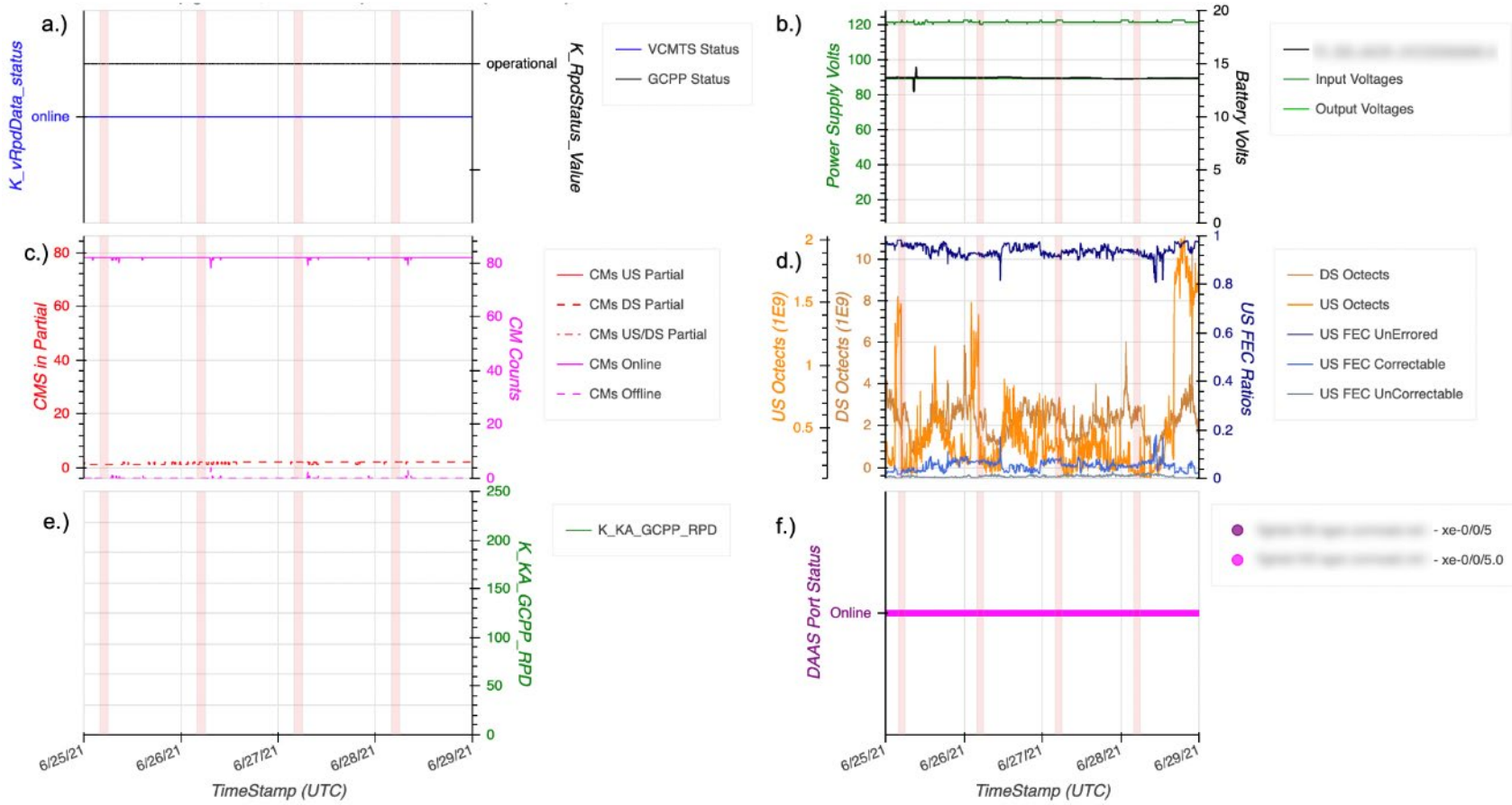
- Based on the available metrics Sherlock is currently able to detect 18 different events from the data
  - Event categories continue to grow with telemetry access
- Events are used to rank/prioritize elements of the DAA network to be reviewed and addressed by experts
  - Frequency
  - Duration
  - Scheduled maintenance
  - Customer impact



Example FEC Event

- DAA platform is pristine
- Plant is impaired
  - Persistent levels of correctable US codewords
  - Short spikes of uncorrectable codewords
  - US performance alerts indicate minor noise on all four US interfaces

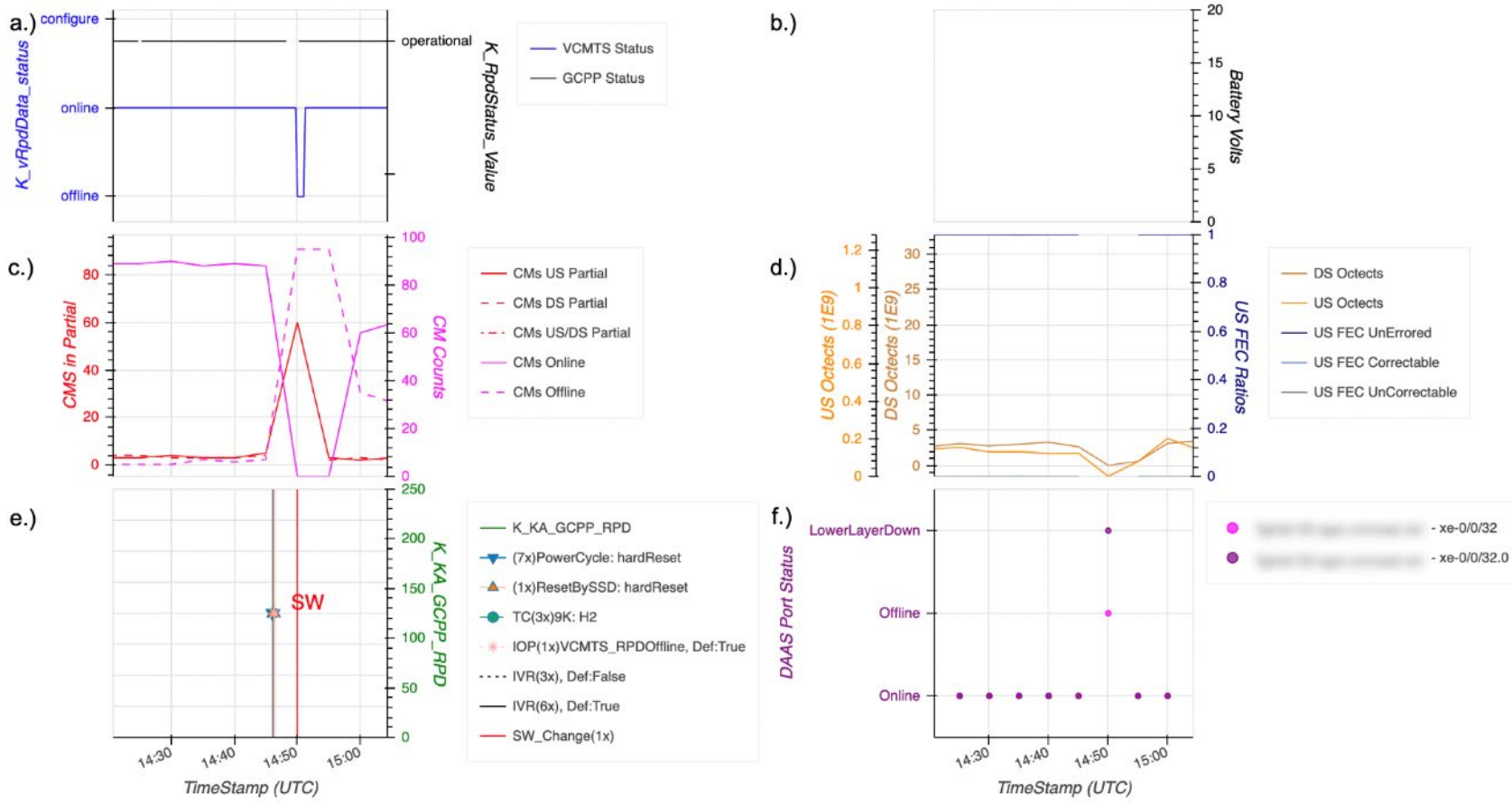
Event Type	Event Start Time (UTC)	Event End Time (UTC)
usFecEvent	6/25/21 9:35	6/25/21 9:35
usPerfEvent	6/25/21 16:35	6/25/21 22:35
usPerfEvent	6/25/21 22:50	6/26/21 2:15
usFecEvent	6/25/21 23:50	6/25/21 23:50
usFecEvent	6/26/21 1:05	6/26/21 2:00
usPerfEvent	6/26/21 2:35	6/26/21 3:10
usPerfEvent	6/26/21 4:40	6/27/21 0:35



Example SW Event

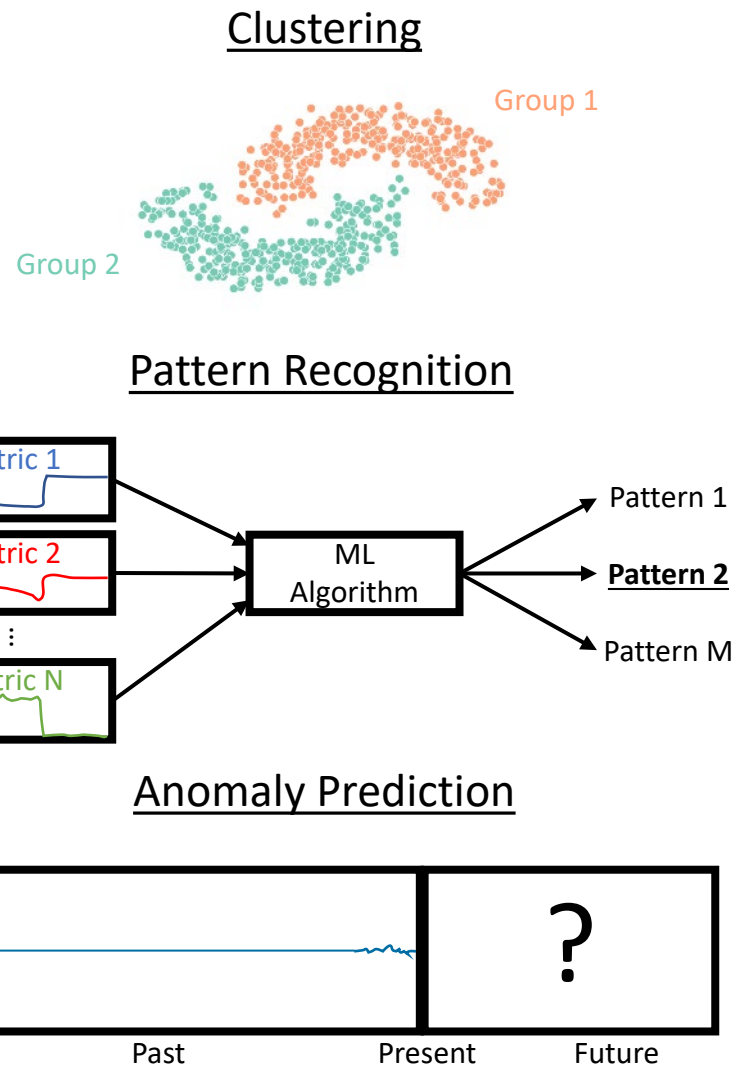
- DAA platform goes down for an RPD SW upgrade
  - Many events flagged
  - Short outages are expected
- Important to tie all platform events to planned outages

Event Type	Event Start Time (UTC)	Event End Time (UTC)
auxOfflineEvent	6/17/21 14:50	6/17/21 14:50
usOctetsEvent	6/17/21 14:50	6/17/21 14:50
rpdRebootEvent	6/17/21 14:50	6/17/21 15:45
rpdIopEvent	6/17/21 14:50	6/17/21 14:50
keepAlivesEvent	6/17/21 14:50	6/17/21 14:50
DAASOfflineEvent	6/17/21 14:50	6/17/21 14:50
datagapEvent	6/17/21 14:50	6/17/21 14:50
rpdSWUpdateEvent	6/17/21 14:55	6/17/21 14:55



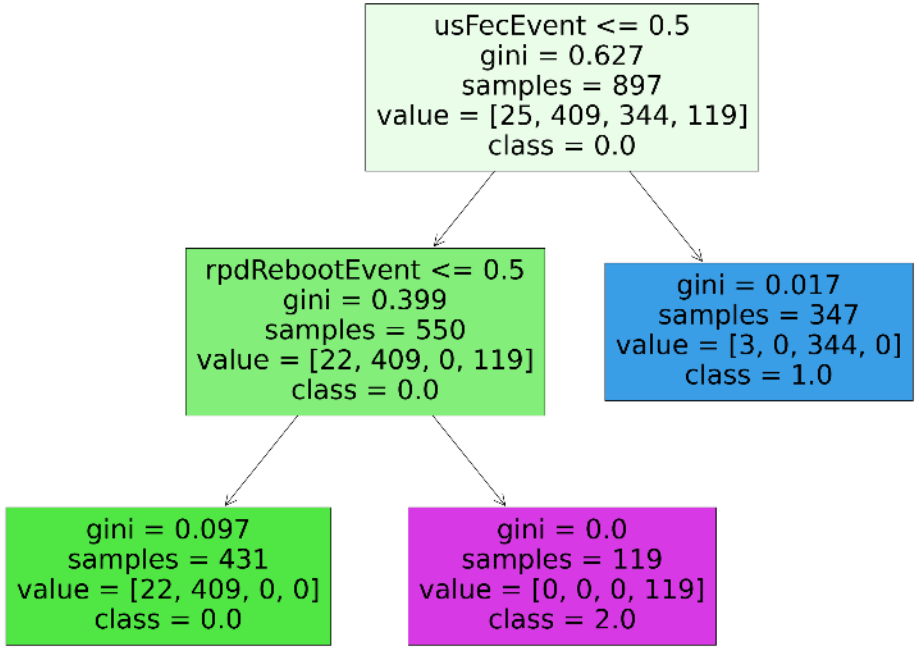
DAA Data is Ripe for ML

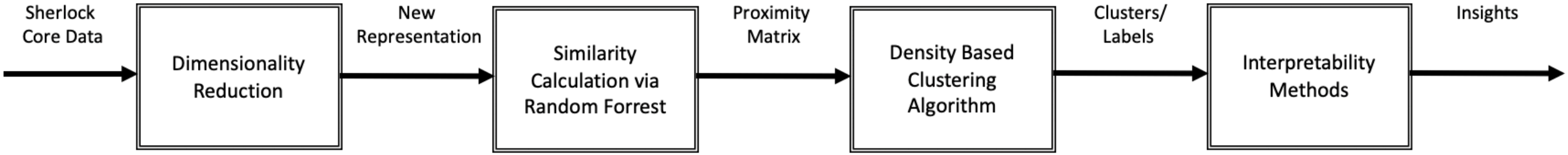
- Expansion in machine learning applications is relatively straightforward once the core data is combined into a single data structure
- Comcast sees three main applications of machine learning as part of Sherlock
  - Identifying similarities in the vast amounts of DAA data via clustering
  - Advanced pattern recognition across multiple dimensions of data
  - Future prediction of anomalies/events and proactive recommendations



Clustering Case Study

- Identified 897 RPD US partial service events in a three week period on the network
- Performed a clustering analysis to determine what other events could have caused the partial service
  - Majority correlated with FEC/noise and RPD reboots
- Only two events were included for simplicity of the study
  - Additional events would need to be added to address the uncorrelated events





Findings

- Big data analysis requires big data specific tools
- Combining all relevant data into a centralized data structure is key for ease/speed of analysis
- When building an analysis framework ensure it can scale with project needs
- Let the data drive the insights and always question the source/quality of data

Future Work

- Continue growing the list of metrics accessible to Sherlock
- Build out the machine learning architecture for more advanced use cases
- Integrate Sherlock and Comcast's network topology graph

Special thanks to the paper's coauthors:
- Ramya Narayanaswamy, Comcast
- Jude Ferreira, Comcast
- Robert Gaydos, Comcast

ATLANTA, GA
OCTOBER 11-14

SCTE
a subsidiary of CableLabs

Thank You!

Matthew Stehman

Data Scientist
Comcast
Matthew_Stehman@comcast.com

2021 Fall Technical Forum
SCTE • NCTA • CABLELABS

COMCAST