Predictive diagnostics for vehicle battery management

next generation vehicle telemetry

V 1.0

05/08/18

Abstract

Vantage Power designs and manufactures technologies that can connect and electrify powertrains in heavy-duty vehicles. Their technology appeals to fleet operators as a retrofit solution to existing vehicles, or to vehicle manufacturers as an OEM solution. By default, Vantage Power's electrification technology comes connected to AWS Cloud, extending the cloud platform on to the vehicle and providing a user interface where new insights and capabilities can be accessed. This solution had been in place and fully operational for more than 12 months by the time of writing this paper.

In the majority of cases, to electrify heavy-duty vehicles a large and expensive lithium-ion battery will need to be installed to provide energy storage. This whitepaper provides a description of a common problem experienced by automotive lithium-ion batteries and how utilising AWS S3, Athena and Sagemaker, a capable solution can be introduced that will provide an alert in advance of a problem capable of damaging the battery.

Introduction:

Battery powered vehicles are increasingly common and as we transition toward an electric low-carbon economy, that is only going to increase further. The design, operation and maintenance of its forebearer, the internal combustion engine, has been established and fine-tuned over more than 100 years to achieve its current worldwide adoption. But for electric vehicles, the transition and deployment is likely to be demanded in a much shorter timeframe – so the technology, operation and maintenance will have to mature much more quickly in comparison.

Automotive battery manufacturers are required to warrant their product for up to 10 years and sometimes beyond, and in many circumstances some have only tested their battery in the lab for a fraction of the time they will see in service. This leads to a couple of questions:

- Is lab testing representative of the conditions a battery will see in service?
- Is accelerated life testing over a few months representative of a lifetime in use?

Some would argue the answer to both questions is "no", and because of this every battery deployment includes a significant level of uncertainty.

Today, the cost of batteries for heavy duty vehicles, such as buses and trucks, is substantial. The weight, chemistry and high-voltage mean there are tough safety considerations when dealing with them. They are a mission critical component and their successful management is fundamental to the operation of an electric vehicle fleet. Batteries also have a limited shelf-life and, for heavy duty applications, can have a considerable lead time and logistical effort if they need to be replaced.

Putting this all together - the financial, operational and technical risks are considerable, but this has not stopped some organisations from fielding battery technology at low maturity – a decision frequently regretted. Using a specific battery failure mode known as cell self-discharge, this paper demonstrates that there are smart options and opportunities out there which organisations can take to manage these risks.

Automotive batteries are often constructed out of much smaller individual cells, assembled into a sub-assembly, monitored and then connected in series to create higher voltages. Self-discharge is a phenomenon that affects batteries at a cell level where undesired chemical reactions reduce the stored charge. Cells which suffer this issue become out of balance with each other with respect to how much charge they are holding.

State of the art on-board electronics, referred to as the Battery Management System (BMS), do enable a level of fault detection, however, detection levels vary and quantification of the problem is difficult. This means the effects can sometimes only be measured at a point where the problem is well developed and can be causing harm, deteriorating performance or presenting a safety risk. In addition to this, the size of some of the vehicle battery packs means there are thousands of cells that need their data evaluating over long periods of time, creating a big-data exercise that on-board embedded processors are not built for.

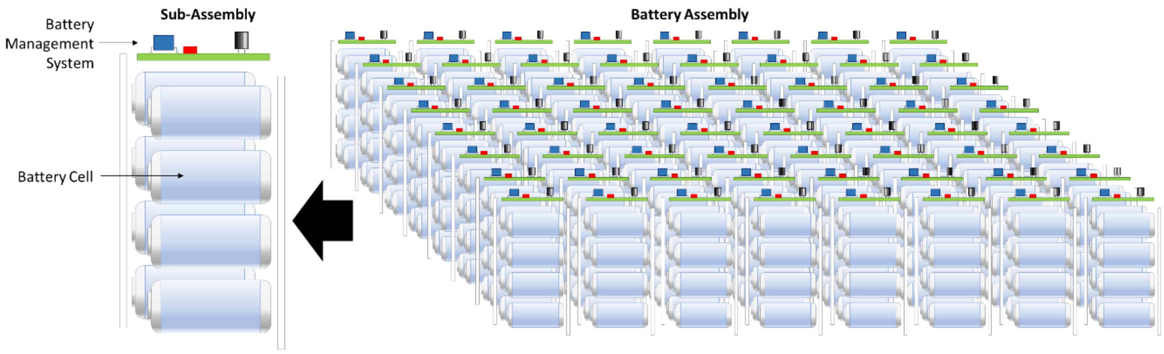

Figure 1: Example of a typical automotive battery. Multiple battery cells are assembled into a sub-assembly and monitored with a BMS. The BMS monitors the sub-assembly, not the individual cell, so small changes at a cell level are difficult to detect.

Cell self-discharge will manifest itself in different ways depending on the type of battery, state of charge, charging cycles, temperature and other factors, over short but also very long time periods. This can mean the volumes of data needed to detect and quantify the behaviour need the use of typical big-data tools.

To add to the difficulty, measurements by the BMS need to be considered in a wider context. What may appear as an obvious anomaly in the data may be easily explained when viewed in conjunction with past utilisation, temperature profile, past maintenance activities, or the age of the cell in question.

This whitepaper demonstrates how the data from multiple vehicle battery packs stored in AWS S3 can be analysed utilising Athena and Sagemaker, and used to create an advanced model that can detect cell self-discharge. The model takes into consideration the context of the measurements and can be used to identify the failure mode significantly earlier than current techniques. Using historical data and known instances of cell self discharge, this new model has been validated and is currently running in real-world operation on a daily basis to identify instances of self-discharge on fielded equipment.

This paper will demonstrate how the use of AWS's cloud-based tools with battery, powertrain and vehicle-specific domain expertise can significantly improve detection of in-field problems and provide the necessary scalability to monitor thousands of connected vehicles. This means battery OEM's can quickly learn how their product is performing and use these insights to innovate, create and mature new capabilities to address issues and edge-cases in near real-time, developing new revenue streams and improving aftermarket support in the process.

Solution:

A connectivity solution and AWS cloud platform had already been established prior to undertaking this project. The platform uses AWS IoT to stream critical data channels at 1Hz from individual vehicles, and compressed batched log files containing more comprehensive data are uploaded to an S3 bucket every 30 minutes. This platform had been in place and fully operational for 12 months storing data from double-decker London buses prior to this project starting. The data transmitted by individual vehicles was already being used to service a comprehensive browser-based user interface called VPVision.

To satisfy the requirements of this project, new AWS products were incorporated into the existing architecture to enable the S3 log data to be easily accessed and indexed, used to create a model and subsequently tested and refined on historical data and new data. The following describes how the main AWS services have been utilised by the engineers to develop and deliver a solution.

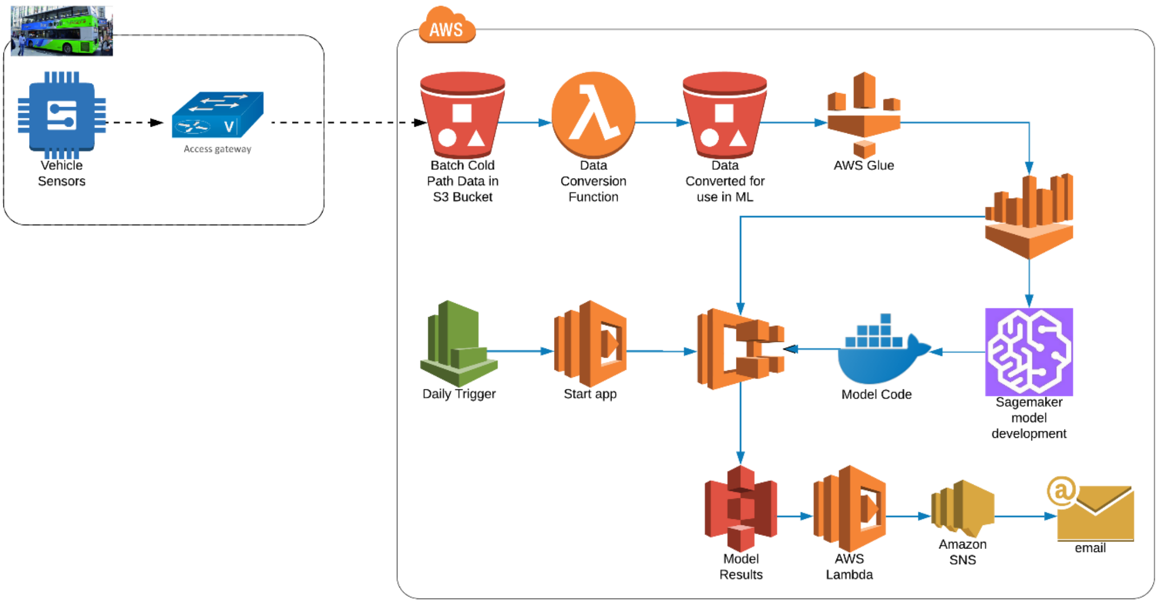

Figure 2: AWS Architecture used for Battery self-discharge model

AWS S3 – AWS S3 is object storage built to store and retrieve any amount of data from anywhere. It boasts impressive durability, provides comprehensive security, access control and compliance, and is highly flexible so that costs can be optimised.

Batch data from the buses is stored in an S3 bucket, with new data arriving every 30 minutes. This data makes the history of the buses, and hence the battery packs, available for engineers to view and understand the performance of the batteries. The data in the S3 was used to develop the self-discharge detection model by applying it to battery packs known to have a self-discharge issue.

AWS Glue – AWS Glue is a fully managed extract, transform, and load (ETL) service that prepares and loads data for analytics. It's easy to use - simply point it at the data storage and it will immediately create a searchable and queryable data catalogue. Glue can be highly customised and is reusable - once created it's operated in a fully managed scalable environment.

Glue meant that the S3 containing the batch data could be directly used as a data source for development. By providing a schema for Glue to read the data files, the data channels relevant to the battery and the self-discharge problem could be exposed for use in development of the model.

AWS Athena – AWS Athena is an interactive query service, fully integrated with AWS Glue, which can analyse large-scale datasets. Using AWS Glue, data repositories across various services can be united and modified as required, Athena then enables the AWS Glue Data Catalogue to be analysed using standard SQL.

Athena meant that the data exposed through Glue could be queried with a simple and familiar SQL interface. This meant that the data could be extracted from the files in a continuous, columnar format that made it easier to conduct analysis over the entire history of the vehicles. It also meant that data could be filtered against other contributing factors, such as temperature, vehicle state and battery state meaning only the relevant and valid data made it into the analysis. All of this was achieved with some simple SQL statements and completed in less than 3 minutes - a vast improvement on what could be achieved with local computing hardware.

AWS SageMaker – AWS SageMaker enables machine learning models to be built, trained, and deployed quickly and easily at scale. Seamlessly integrated with other parts of AWS, SageMaker provides the tools required to access the right data elements, select a model and framework, tune and refine the model, integrate it and deploy at scale. A lot of the complexities associated with creating these capabilities are removed meaning more effort is focussed on developing the best algorithm rather than the supporting infrastructure.

SageMaker allowed loading and exploration of the data in a familiar Jupyter Notebook environment. It was simple to set up an Athena query on the S3 data to pull in all relevant data for the self-discharge problem and get to work on developing the model. The model went through several stages of development in SageMaker and could quickly be tested on large amounts of historic data to prove out the algorithm.

The Model – The ideal conditions to perform the analysis need to be established. Until these conditions are met no analysis is conducted.

Once a suitable state has been achieved, the critical data from all the BMS devices within the battery pack are collected. The low-granularity digital readings produced by the BMS do not provide the accuracy or level of detail required and the readings will fluctuate around the real value.

The model will smooth out these readings and will generate a moving average mean value for each individual BMS. Given a sufficiently long period, the rate at which the moving average is changing can be calculated.

If the conditions required to identify these trends are interrupted, shorter periods can be stitched together in order to provide a sufficiently long duration. Once the analysis of individual BMS units is complete, an average for the whole pack is calculated and an acceptable deviation is created, this establishes a pass/fail criteria unique to each pack. On a daily basis, the model will recreate the acceptable limits for each pack and identify non-compliance. By re-evaluating the pass/fail limits each day, the model accounts for normal changes in the performance of cells as a result of pack utilisation and age. So as the battery pack matures and the performance changes, the assessment criteria automatically changes with it. Any anomalies identified during the daily analysis are as a result of an individual cell losing voltage faster than the rest, which indicates cell high self-discharge.

Using AWS Simple Notification Service (SNS) engineers are notified when an anomaly has been detected by the model, allowing preventive maintenance to be scheduled. With this model, existing monitoring techniques can be used to detect the self-discharge early, notify and have the battery looked at, which will increase its useable life.

The Process – SageMaker was used to explore the data to first understand the trends that were emerging. From this an initial model was developed that looked over 1 week of data to analyse the statistical trends and detect the outliers. This is achieved using both standard statistical deviation from the median and a Local Outlier Probability detection algorithm. The model was refined to use 1 day’s worth of data, increasing the resolution of the results and improving accuracy.

The model was adapted into a Docker container that could be run on AWS Elastic Container Service (ECS). Since this model runs only once per day and requires a few minutes of compute time, it was not required to use a long running ML instance. This model is triggered once per day using a CloudWatch-scheduled event to trigger a Lambda function. This then starts up one Fargate task per vehicle, exporting the results to an S3 bucket. This triggers a Lambda which will analyse the results and send email notifications if anomalies have been detected.

Conclusion:

Battery cell high self-discharge can now be automatically identified much earlier than existing techniques allow, using AWS S3, SageMaker and Athena. On the real-world applications tested, the problem can be identified 1 month earlier at a point where the fault has not caused more widespread damage.

The new technique was developed using 6 years of historical data (1 year / 6 vehicles) and 4 known examples of cell self-discharge to create a highly capable statistical model. The model ensures that the analysis is only conducted on suitable data and that the analysis will adapt to accommodate normal changes in the performance of the battery throughout its life.

After first developing the model, the analysis required a week of data in order to identify any trends and anomalies. By refining the model, this was reduced to 1 day and a further refinement may see this reduce further. Continuing to reduce the timeframe over which this analysis is conducted will allow the analysis to be pushed to the edge. Using AWS Greengrass the opportunity exists to delegate this analysis to operate on-vehicle, reducing the AWS running costs whilst massively increasing the on-vehicle software capability.

AWS provided a seamless environment for this capability to be created, tested and deployed.

AWS meant that it was simple to leverage the data uploaded by the vehicles and already stored in S3. By integrating Glue, Athena and SageMaker engineers were able to query the data and use this to develop the model. Getting from raw data in S3 to developing in SageMaker took a few hours of setup and running the analysis in SageMaker was fast, given the amount of data being crunched. Engineers also didn't have to worry about available hardware or computing power, as SageMaker handles all of this in the background. Engineers could get on with the important job of developing the model, without the need to build and manage the computing hardware and software required to perform the analysis.

The cloud provided the perfect environment in which to develop and refine the model. The 4 instances of cell self-discharge were spread across different vehicles, so data from each device was required during the development and testing phase. Having instant access to large data sets and to live data from multiple vehicles was invaluable in order to iterate and refine the design of the model quickly.

By developing and trialling the software in the cloud, as opposed to on-vehicle development and testing, the risk of disruption to vehicle operation is zero. This low risk environment is perfect to innovate and trial different models, whereas an on-vehicle development would likely have resulted in a more cautious and simplistic implementation.

<u>Future:</u> The capability developed can be used to monitor any battery powered application where safety, reliability, performance and maintenance cost are key considerations. This could extend across a range of hybrid / electric automotive applications and beyond in to other industries that utilise battery technology.

The approach taken can also be used to develop other battery monitoring and alert types, specifically those which require more complex analysis and benefit from a connected low risk environment that can provide almost instantaneous feedback. Given the uncertainty that accompanies any long term battery deployment, the approach taken in this project may continue throughout the operational life of the battery. Since this is based on real-world data and not lab data, developing new code in response to emerging problems may actually accelerate toward the end-of-life as problems become more frequent.

It does not need to stop at the batteries. Any connected and monitored device can benefit from failure and performance analysis created using AWS Sagemaker. Depending on the framework, this analysis does not need to reside in the cloud - when a sufficient maturity level and validation is achieved AWS Greengrass can be used to deploy the code on-vehicle, or on-device.

To achieve the rapid deployment of electric vehicles, this ability to learn and improve in-service is fundamental. By combining the industry capabilities of Vantage Power and the vast potential of VPVision’s architecture and platform, a valuable partnership is created that enables new possibilities and unlocks new revenue streams.