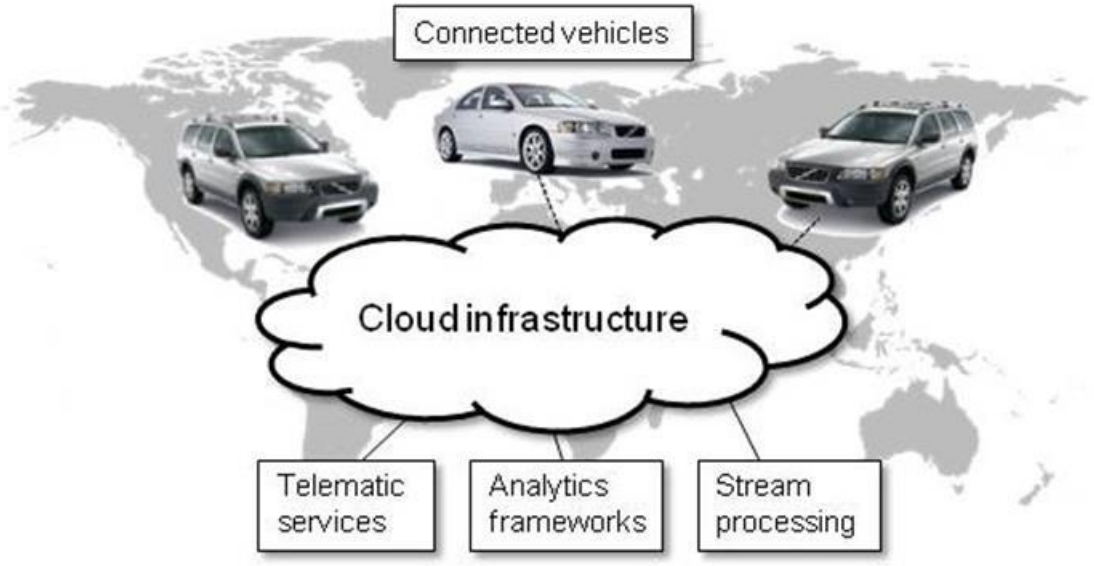
Final report: BAuD - Big Automotive Data



Project within FFI Fordonsutveckling

Mats Gjertz, Mathias Johanson

2015-03-03

Content

FFI in short

FFI is a partnership between the Swedish government and automotive industry for joint funding of research, innovation and development concentrating on Climate & Environment and Safety. FFI has R&D activities worth approx. €100 million per year, of which half is governmental funding. The background to the investment is that development within road transportation and Swedish automotive industry has big impact for growth. FFI will contribute to the following main goals: Reducing the environmental impact of transport, reducing the number killed and injured in traffic and Strengthening international competitiveness. Currently there are five collaboration programs: **Vehicle Development, Transport Efficiency, Vehicle and Traffic Safety, Energy & Environment and Sustainable Production Technology.**

For more information: www.vinnova.se/ffi

1. Executive summary

To be successful in the increasingly competitive consumer vehicle market, automotive manufacturers must be highly responsive to customer needs and market trends, while responding to the challenges of climate change. One key to achieving this is to promote knowledge-driven product development through large scale collection of data from vehicles in use, to capture customer needs and to gather performance data and statistics. In the testing and validation phases of automotive development, large volumes of measurement data and diagnostic data are being collected through powerful data loggers and diagnostic systems. With sophisticated telematics services enabling wireless collection of data from test vehicles from all over the world, the volume of data collected quickly grows beyond what is possible to manage by manual processing. This calls for automated data analysis mechanisms, that process data immediately as they are uploaded from connected vehicles.

In the scientific community, the scalability and manageability issues relating from the increasing demands of data is frequently referred to as the "Big Data" challenge. In this project we have studied the Big Automotive Data challenge, resulting from the need to collect large amounts of data about customer behaviour, performance metrics and statistics to support knowledge-driven development of next generation vehicles. The approach has been be to design a technological framework for scalable generic data capture and integrated analysis. Novel analysis methods have been developed, for integration in the framework. The system has been tested and evaluated in vehicle development projects at Volvo Car Corporation.

The project parners are Volvo Car Corporation, Alkit Communications AB, Luleå University of Technology and Uppsala University.

In summary, the results of BAuD project are as follows:

- Analysis and visualization mechanisms for time series data from connected vehicles have been developed, with a focus on needs related to battery development for hybrid cars
- Data capture and analysis mechanisms based on streaming data have been developed, focusing on the development of new features for active safety
- User studies have been conducted, with feedback to the development of the two case studies within the project (battery development and active safety)
- The project has contributed to an improved development process at Volvo Cars
- Scientific results have been produced and published
- A course in database technology with a focus on NoSQL database solutions has been organized within the project
- A Master's thesis work has been carried out at Luleå University of Technology
- A follow-up project, BAuD II, has been applied for and granted within the FFI program

2. Background

The globalization of markets, resources and knowledge require product development companies to be highly responsive to customer needs and to environmental changes. In the automotive industry, the major challenges of limiting CO_2 emissions while delivering high quality products to a growing number of customers in expanding and highly heterogeneous markets require very efficient and powerful tools and methods to capture customer needs and also to gather performance data and statistics to improve product development.

In the testing and validation phases of automotive product development, large volumes of measurement data are being gathered from fleets of test vehicles. With the advent of telematics systems and improved means of wireless vehicular communication more or less globally, the opportunities to collect data has improved tremendously over the past few years. This has an enormous potential of improving automotive product development, by making reliable performance data, statistics and customer behaviour information available as quickly and efficiently as possible in the development process. The ability to make good use of this valuable resource can be clearly identified as a key means to be competitive in the automotive industry. The big challenge is how to be able to efficiently collect, manage analyse and make good use of the large volumes of data collected.

In the BAuD project, we have studied the technological and usage-related mechanisms needed to scale up the systems currently in use for collection and analysis of data from vehicles at Volvo Cars. New technology and methods have been developed to meet the needs.

As an illustration of the need for Big Data mechanisms, Figure 1 shows the growth of the volume of data collected in the case studies during the BAuD project. As can be seen, the data volume has more than doubled, and the rate continues to increase.

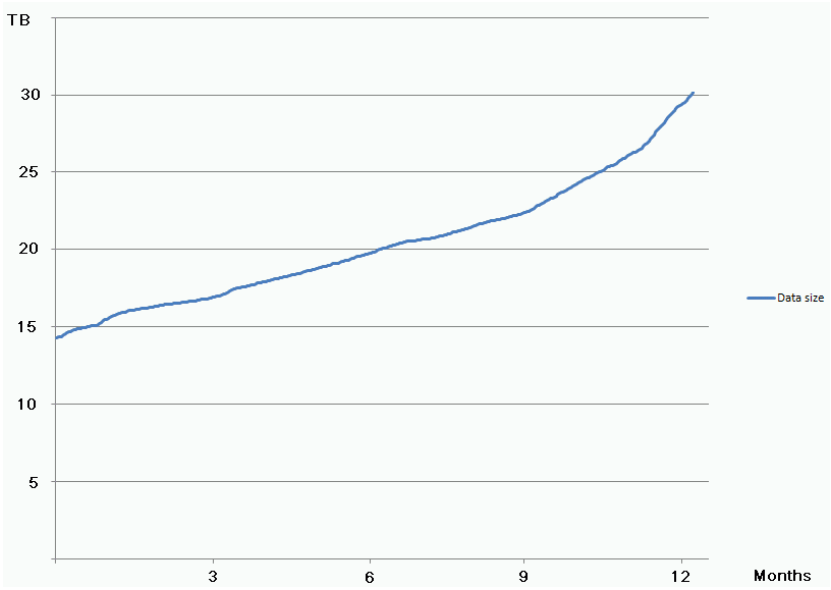


Figure 1: Data volume growth from connected vehicles in the BAuD project during the last 12 months

3. Objective

The project's main objective has been to develop technology and methods for large-scale capture and analysis of data from vehicles in order to improve the possibilities of knowledge-driven product development in the automotive segment. An important part of this has been to gain knowledge about the current state-of-the-art concerning Big Data technology and best practices, and to position the project partners within the subfield of Big Automotive Data.

A further aim has been to study the opportunities of developing new products and services enabled by the collection of large amounts of data for product development. We refer to this as *data-driven innovation*. Examples of new opportunities that arise are possibilities to exploit selected parts of the collected data together with third parties, or to develop new aftermarket services integrated into production vehicles that utilize the data collected from the vehicle fleet.

4. Project realization

The approach of the project has been to take a holistic view of automotive data capture, Big Data analytics and decision-making processes, to identify current weaknesses and opportunities for improvements beyond the current state-of-the-art. This includes the development of a scalable generic data capture and integrated Big Data analytics

framework, composed of both existing components (telematics services) and newly developed components (analytics framework).

The work has been carried out within six work packages (WP) as follows:

WP 1: Project management

Project management and coordination activities have been carried out by VCC with participation from all other parners.

WP 2: Use case definition and needs

The needs and requirements analysis was perfomed through a series of workshops arranged by VCC early on in the project, with participation from the main stakeholders of Big Data at VCC and the three other project participants. Around 30 Use Cases were identified and described, whereof two were selected as the pilot applications of the project. The two use cases selected were battery development for hybrid vehicles and active safety development.

WP 3: Scalable data capture and analysis framework

This work package was focused on design and development of the framework for integrated data capture and analytics, including visualization mechanisms for time series data and mechanisms for handling of streaming data. The work was mainly carried out by Alkit Communications.

The development was based on a system architecture which has also been developed within the project. The software components developed has been integrated with the telematics system in use by R&D at Volvo Cars for collecting data from test vehicles (WICE).

WP 4: Data analysis methods and tools

New analysis methods and tools have been studied and developed in the project, mainly by Luleå University of Technology in collaboration with Volvo Cars. The methods have also guided the development of the analytics framework performed in WP3.

WP 5: Development process optimization and case studies

Two case studies have been conducted related to ongoing product development projects at Volvo Cars. The first case study aims to improve the development of batteries for hybrid cars, through novel analytics capabilities integrated within the telematics system

WICE used for data collection. The second case study aims to enable handling of streaming data from test vehicles to support the design of a new active safety feature currently under development at Volvo Cars.

In parallel with the case studies, work on development process optimization has been carried out by Volvo Cars.

WP 6: Scientific supervision and academic dissemination

The work in this WP has been devoted to keeping the project up to date with cutting edge research and technology concerning Big Data. Academic dissemination of scientific results has also been performed and supervision of a Master's thesis student at Luleå University of Technology. A course in database technology for the project partners has been arranged by Uppsala University.

5. Results and deliverables

The project has resulted in new technology, developed and tested within the project in two case studies related to the ongoing product development projects at Volvo Cars. The project has also generated new analysis methods and approaches. A Master's thesis has been produced at Luleå University of Technology and a course in database technology, with a focus on NoSQL-mechanisms for Big Data applications has been developed and organized by the University of Uppsala.

The results are described in more detail below.

Scalable analysis of data from connected test vehicles

A technological framework has been developed in the project which enables large-scale analysis of data from connected vehicles. This framework has been implemented by Alkit and integrated with the telematics system WICE used by R&D at VCC. The framework is based on a system architecture also developed within the BAuD project, with the aim of meeting the demands of performance, scalability, integration in Volvo's IT infrastructure, extensibility and ease of use. Figure 2 shows a high level outline of the architecture.

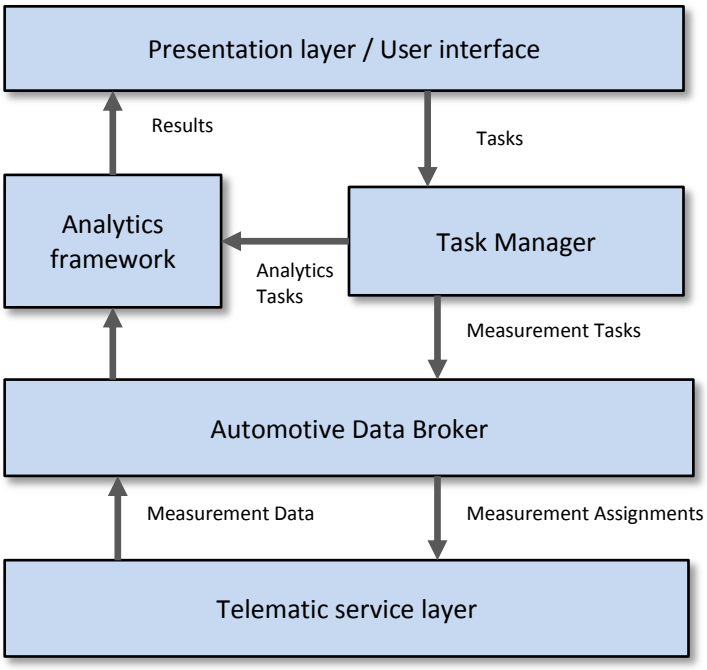


Figure 2: High level architecture of the BAuD system for capture and analysis of data from connected test vehicles

The approach enabling data to be captured from connected test vehicles via telematics services and directly analyzed and visualized, without requiring manual exporting to other systems, enables product development to be more efficient. To test this new approach, a use case related to battery development for hybrid vehicles was selected as a pilot application. A number of analysis functions was developed in collaboration between Luleå University of Technology (LTU) and VCC and implemented and integrated by Alkit in the WICE portal, which is the user interface of the WICE telematics system. This enables direct online analysis of data from the connected test vehicles.

Figure 3 below shows an example of what the web-based user interface for data analysis looks like, which has been developed in the BAuD project and integrated into the WICE telematics system.

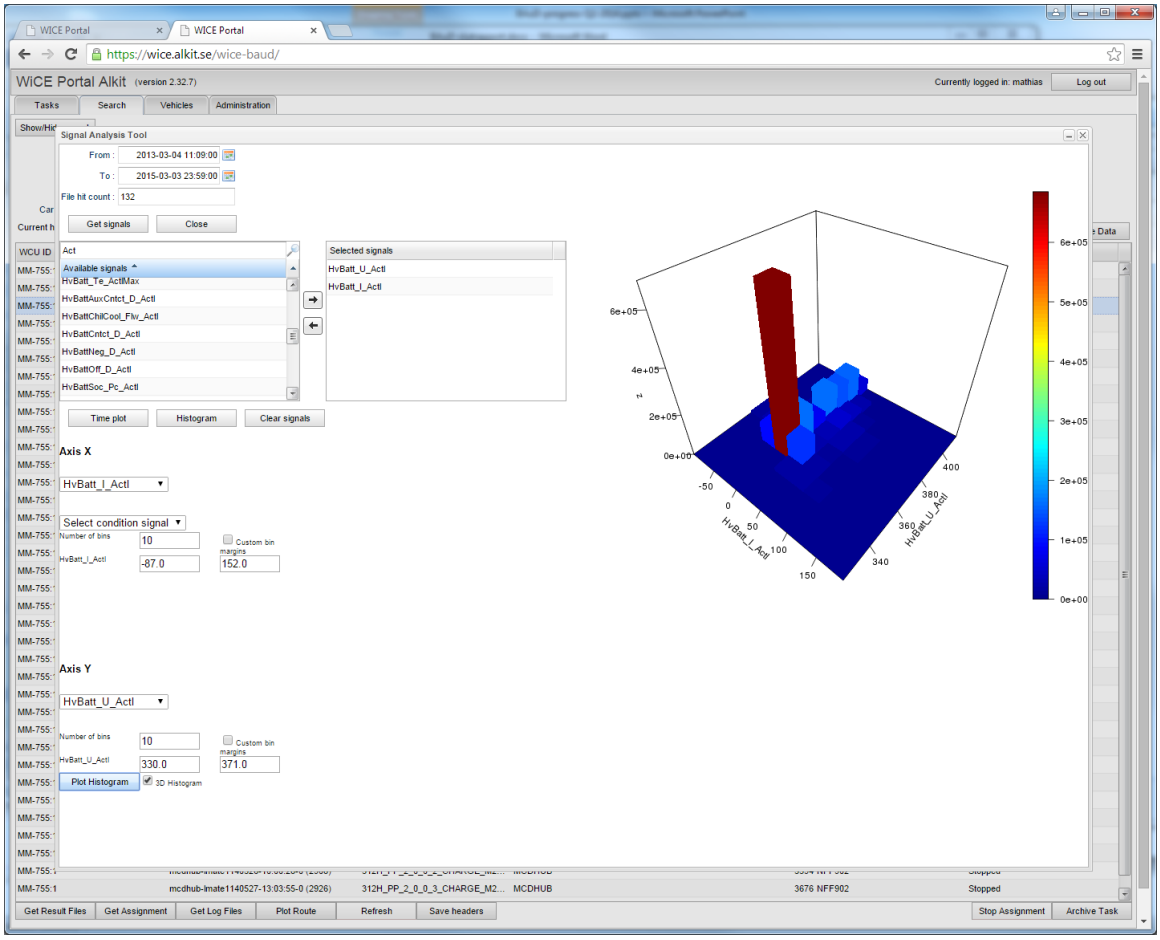


Figure 3: Screen shot of the user interface of the integrated analysis tool developed in the BAuD project

The example above shows a 3D histogram plot of signals. Several other possibilities of analysis and visualization mechanisms have been implemented by Alkit based on several different analyses (around 50) developed by Volvo and LTU to meet the battery group's needs.

Scalability mechanisms for data capture and analytics

In order for the data capture and analytics framework to be able to manage large data volumes uploaded from potentially very large fleets of vehicles, a system architecture has been designed based on Big Data principles.

The concept is based on having a so-called generic measurement task running in the data capture device in a large fleet of test vehicles. The generic measurement task captures a large number of parameters and signals for many different purposes. This results in measurement data files that are uploaded to a server infrastructure. When some analysis should be performed for a particular stakeholder, the signals that are relevant for the

analysis must be filtered out of from the uploaded data files and then fed to the selected analysis algorithm, whereupon the results are compiled and visualized. This can be formulated as a Map-Reduce operation where the filtering of signals and the analysis function constitutes the Map function, and aggregation of result and visualization constitutes the Reduce function. By designing the BAuD framework based on a general infrastructure for executing the Map-Reduce parallelized in a clustered environment (Spark), where data is stored in a distributed file system in a way that allows parallelization (HDFS), linear scalability can be achieved when the volume of data grows exponentially.

A plan for how the existing infrastructure of the WICE system can be migrated to the architecture developed in BAuD project has been produced and a project has been initiated at VCC for the industrialization of the results and integration into VCC's IT infrastructure.

Handling of streaming data from connected vehicles

One of the use cases of the project is linked to a VCC internal project called RFE (Road Friction Estimation), which aims to develop new active safety features which require analysis of data that are continuously streamed from connected vehicles. For this use case, the BAuD project has developed a monitoring function, whereby captured road friction data can streamed in real time to a cloud service, where the data is analyzed and used to send alerts to nearby cars on slippery roads. This "collaborative warning system" is very interesting from the BAuD project’s perspective because it not only requires the system to cope with many connected cars (hundreds of thousands or millions of term), but also poses requirements for real-time processing of data.

Various techniques for managing streaming data has been studied and tested for this application scenario. Among other things, Uppsala University has contributed expertise and solution principles regarding DSMS technology (Data Stream Management Systems).

Product development process improvement at Volvo Cars

Work is underway at Volvo Cars to introduce the approach based on large-scale capture and analysis of data from connected vehicles which has been developed in the BAuD project into Volvo’s BRAVe process. This includes both organizational changes and the introduction of new technology developed within the BAuD project and related projects.

Thesis work at Luleå University of Technology

A Masters thesis has been produced in which a measurement assignment has been defined for the Volvo V60 Hybrid vehicle. It primarily deals with balancing issues for the

conditioning of the car and charging of the battery. The work has generated the report "Measuring assignment for Volvo V60 Hybrid".

Course in NoSQL database technology at Uppsala University

A course was organized by the Database Laboratory of Uppsala University (UDBL) in the autumn of 2014, in which all project partners participated. The course presented various techniques for Big Data management according to NoSQL principles and also had a special focus on the analysis of streaming data.

The course was led by Professor Tore Risch and Dr. Kjell Orsborn.

5.1 Delivery to FFI-goals

The project is a demand-driven research and development project with a clear emphasis on improving industrial competitiveness and innovation. The project's main focus has been to foster knowledge-driven product development by means of large scale data collection and automated analysis, with an expressed goal of promoting fast and efficient development. This is in perfect harmony with the programme's focus area "Methods and tools for vehicle development" (Metoder och verktyg för fordonsutveckling).

The "Green, Safe and Connected" vision of the focus area "Vehicle electrics and electronics" (Fordonsel och elektronik) is also very well in agreement with the project's intentions, with connectivity through telematics services and data collection for optimization of fuel efficiency and security systems being key drivers. The project furthermore has strong elements relating to embedded systems and software (Inbyggda system och mjukvara); the telematics services, the data analysis software, and the database technologies.

Some of the key contributions of the project to the specific goals of the programme are listed below.

Contribution to targets within transportation, energy and environmental policy

Improved means to collect large volumes of data from connected vehicles improves the opportunities of studying and improving the environmental impact of vehicles. The emission of greenhouse gases can be measured on entire fleets of vehicles, and optimization of engine control software can be performed. This data can also help in the development of new fuel systems. Collected customer data can also be used to generate statistical background material for improving transport efficiency.

Contribution to the ability of industry to operate knowledge-based production in Sweden in a competitive way

Sweden is one of the world-leading R&D nations in vehicular communication and telematics technology. This project has strengthened this position. The ability to do large scale collection and analysis of data from vehicles all over the world supports knowledge-based production of vehicles in Sweden, while still having operational data available from all over the world.

Contribution towards a vehicle industry in Sweden that continues to be competitive

The key to competitiveness in the future is to be responsive to market needs and trends. This has been one of the main motivations of the project, and Sweden's leading role in telematics and IT makes Swedish industry competitive in this regard.

Undertake development initiatives of relevance to industry

The project is an industry-driven research and development project, originating from demands for improved data collection and analysis. The results will be implemented in industry (at Volvo Cars).

Ensure that new knowledge is developed and implemented, and that existing knowledge is implemented in industrial applications

See above. New knowledge is generated both with respect to the new technologies and tools developed, and also as a consequence of more and better data being made available for the development of next generation products. New business models and services will also be possible based on this data.

6. Dissemination and publications

6.1 Knowledge and results dissemination

Dissemination of knowledge and results has been conducted within the project both internally within the participating organizations, as well as externally through participation in public events, such as conferences and exhibitions, and through the publication of scientific results.

Big Data has emerged as one of the major technology trends in recent years. The BAuD project has helped to place the project participants on the map as innovative R&D

atkeholders in this domain. The project has established the concept of "Big Automotive Data" as an important focus area for vehicle manufacturers to stay competitive and innovative.

6.2 Publications

Selected publications of the project are listed below.

- M. Johanson, S. Belenki, J. Jalminger, M. Fant, M. Gjertz, "Big Automotive Data - Leveraging large volumes of data for knowledge-driven product development", IEEE International Conference on Big Data, Washington DC, October 27-30, 2014.

- O. Rabelius ”Mätuppdrag Volvo V60 hybrid med avseende på kundbeteende”, Master’s thesis, Luleå University of Technology, 2014.

- M. Johanson, M. Gjertz, M. Fant, J. Dahl, G. Vestlund, J. Jalminger. ”BAuD: Use Case Description and Requirements”, BAuD project deliverable 2.1.

7. Conclusions and future research

In the BAuD project we have explored the opportunities and challenges of leveraging Big Automotive Data for knowledge-driven product development. The BAuD framework, a scalable and efficient Big Automotive Data platform including integrated telematics and analytics services, has been developed and tested in two case studies conducted at Volvo Cars. The two case studies focused on active safety development and battery performance for hybrid vehicles respectively.

As part of our future work in the follow-up BAuD II project, we will explore how the BAuD framework can be extended to capture not only objective measurement data from connected vehicles, but also subjective usage information from customers. We intend to do this by designing a smartphone app, which will allow specialized questionnaires to be presented to selected customers, to capture the usage experience and feed the subjective data into the BAuD analytics framework. The BAuD II project will also continue the work on designing scalable and high performance data capture and analytics mechanism and develop a knowledge management system for handlig results of analyses.

8. Participating parties and contact person

The participating organizations of the BAuD project, with contact persons, are listed below.

	Volvo Personvagnar	Mats Gjertz	mats.gjertz@volvocars.com
	Alkit Communications AB	Mathias Johanson	mathias@alkit.se
UPPSALA UNIVERSITET	Uppsala universitet	Tore Risch	Tore.Risch@it.uu.se
LULEÅ TEKNISKA UNIVERSITET	Luleå tekniska universitet	Ove Isaksson	Ove.isaksson@ltu.se