PrEsto Cloud

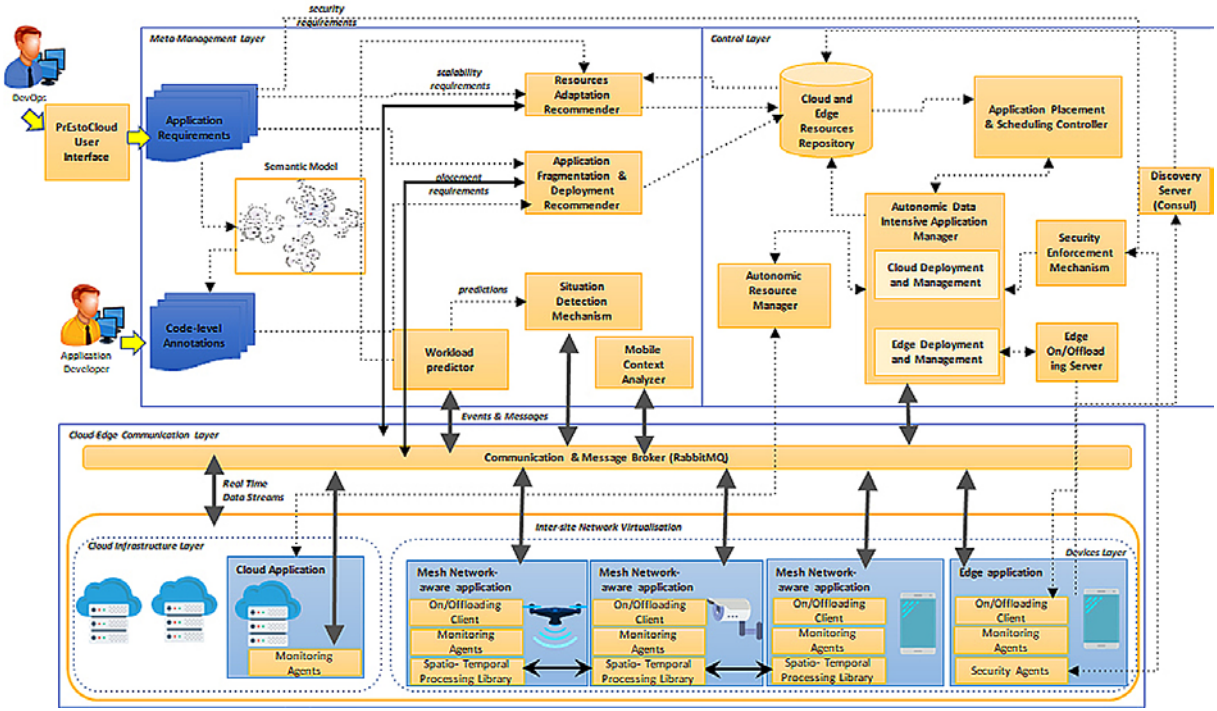
PROACTIVE CLOUD RESOURCES MANAGEMENT AT THE EDGE FOR EFFICIENT REAL-TIME BIG DATA PROCESSING

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732339".

PROJECT DESCRIPTION

Among the greatest challenges of cloud computing is to automatically and efficiently exploit infrastructural resources in a way that minimises cloud fees without compromising the performance of resource demanding cloud applications. In this aspect the consideration of using processing nodes at the edge of the network, increases considerably the complexity of these challenges. PrEstoCloud idea encapsulates a dynamic, distributed, self-adaptive and proactively configurable architecture for processing big data streams.

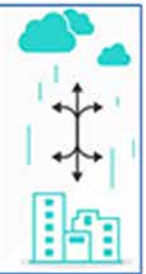
PROJECT OVERVIEW



For providing the aforementioned characteristics on the platform architecture is based on components, which are grouped in five layers as shown

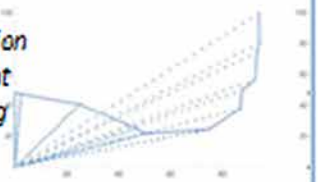
Meta-Management Layer

workload prediction
edge context analysis
situation detection
adaptation recommendation



Control Layer

deployment optimisation
application deployment
deployment scheduling and control



Cloud-Edge Communication Layer

Communication and message brokerage
Event queing

Cloud Infrastructure Layer

Monitoring probes

Devices Layer

Inter-site Network Virtualisation
Spatio-temporal processing
On/off loading agents
Monitoring Agents

PrEstoCloud basic layers

- The **Meta-management layer** which mainly consists of decision logic capabilities required for enhancing the Control layer. Components of this layer use as input the situation details, the variation of the data streams and the context of the mobile devices at the extreme edge of the network to recommend, at the appropriate time, the necessary adaptations, such as scaling of resources and on/off-loading of fragments.

- The **Control layer** which manages resources of the cloud and edge infrastructure layers and contains components which monitor and manage cloud resources capabilities that can be extended to the edge of the network. Moreover, this layer is responsible for the optimized scheduling of fragment execution over available resources. The control layer detects available edge resources and selects target resources for deployment and plan fragment scheduling according to the recommendation of the meta management layer. The selection of the individual cloud and edge infrastructural resources is followed by an optimization step, which involves the examination of a big variability space and find the most appropriate alternatives that satisfy certain business goals (e.g., reduce cloud costs while maintaining an adequate response time).

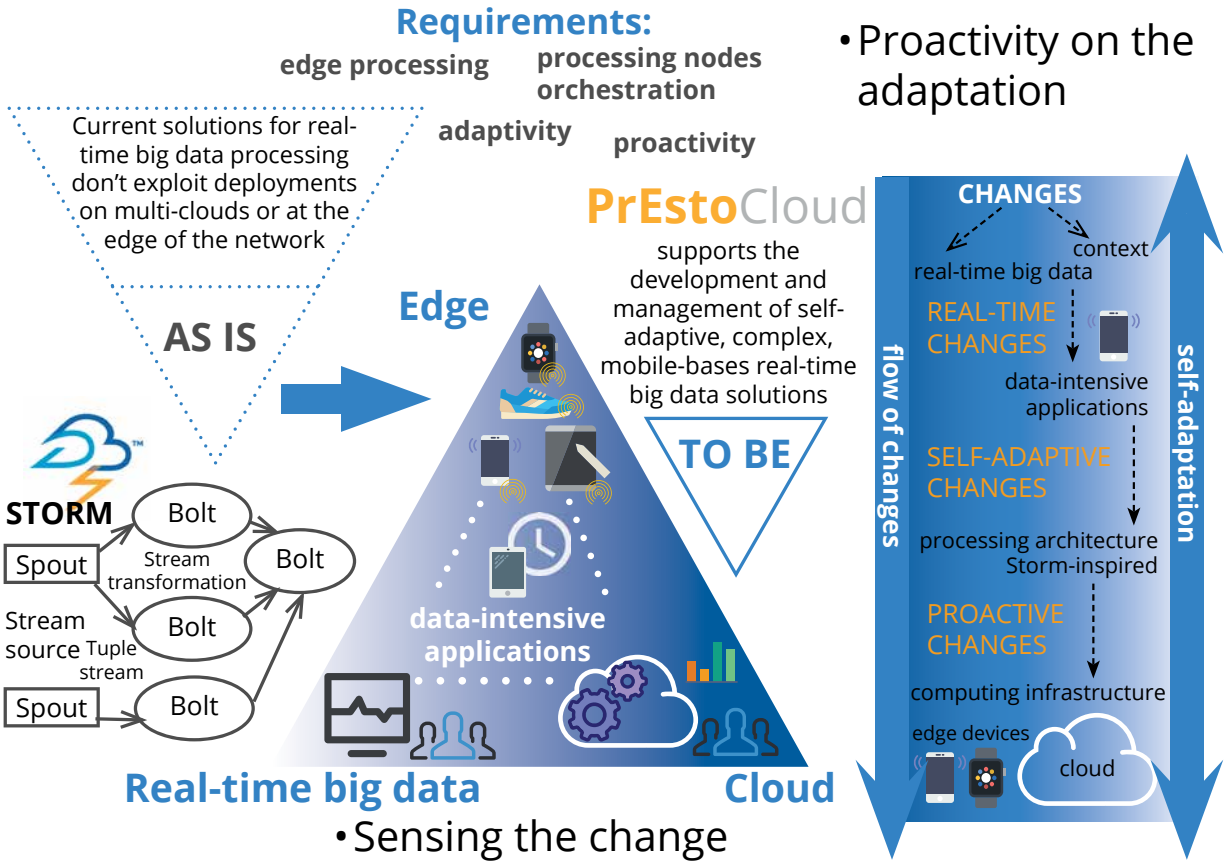
- The **Cloud infrastructure layer** which realizes the dynamic placement and scheduling capabilities, according to the instructions of the Control Layer, allows the utilization of the edge resources (i.e., Devices layer), private and public clouds. The deployment of fragments is handled based on deployment constraints elated to different properties like response time, security constraints or any other preferences of the DevOps.

- The **Cloud-Edge communication layer** contains the inter-site network virtualization technology for coping with the need for connecting resources situated in multi-cloud environments and managing their orchestration and provisioning across different and heterogeneous providers. This layer is also responsible for relaying data streams securely on and off the PrEstoCloud platform and providing publish/subscribe event brokering capabilities.

- The **Devices layer** consolidates any edge device that can be used as a processing node.

CHALLENGES

- Exploit multi-cloud environments for deploying big data processing frameworks extended to the extreme edge of the network
- Make intelligent cloud placements and configurations of applications based on the anticipated processing load with respect to data volume and velocity
- Elaborate on components that are capable to recommend and implement adaptations in real-time



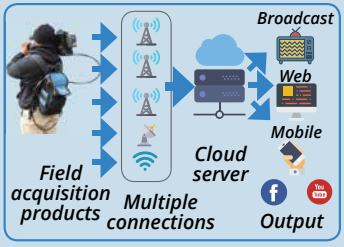
- Proactivity on the adaptation
- Sensing the change

USE CASES

Logistics (CVS mobile)

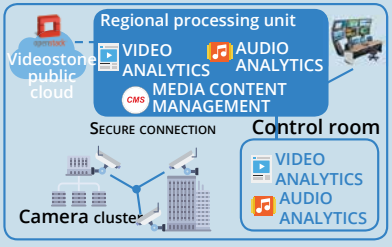
A vehicle/fleet management processes real-time information and alerts – based on data streams from GPS, on-board diagnostics, tire sensors and others.

Media (LiveU)



A media prosumer platform offers personalized and flexible consumption of real-time stories by combining freelance reporting, traditional broadcasting and social media streams.

Surveillance (ADITESS)



A surveillance solution combines real-time data streams from cameras and pre-processing results from groups of unmanned aerial vehicles.

OBJECTIVES

- Inter-site network virtualization and security management
- Multi-layer cloud resource management and monitoring
- Distribution management
- Adaptive scheduling of IoT big data processing tasks between devices and the cloud
- Proactive cloud adaptation
- Test and validate the proposed approach in complementary use cases

Software AG performed several roles in PrEstoCloud. As Project Coordinator, Software AG had the overall responsibility for intermediation between consortium and European Commission as well as the overall project controlling and supervision. Software AG managed the knowledge produced during the project lifecycle and made a significant contribution to the success of the project.

In its role as work package leader “Dissemination and Exploitation”, Software AG coordinated and managed the dissemination activities and thus gave the impetus for the high level of activity in dissemination at all levels. Software AG presented PrEstoCloud to a larger group of prospective customers and interested public on its booth at international fairs like Cebit and Hannover Messe. At Hannover Messe 2019 Software AG demonstrated PrEstoCloud not only on workstation but also on stage with interactive presentations and open discussions.

A detailed exploitation plan regarding products and partner exploitation with promising individual elements and applied to the pilot cases was elaborated. A comprehensive exploitation plan is based on a canvas model that identifies the role of each partner in disseminating and/or exploiting the project products.

Finally, Software AG contributed on technical level in supporting the architecture and integration as well as integrate components based on Software AG products. So PrEstoCloud platform also offers an alternative approach including Software AG commercial components e.g. APAMA as workload predictor and situation detection mechanism as well as MashZone for visualization and modeling purposes. Within PrEstoCloud’s architecture this allows to instantly get insights from big fast data streams from EDGE devices and to predict what workload is likely to happen next.

Nissatech is a ten-year old, innovation-driven SME with strong international cooperation and vision to become European leader in developing advanced IT solutions for advanced data processing to be used in various industries and businesses. The main objective is to develop own technological building blocks through an efficient implementation of the cutting-edge research and their usage for resolving very challenging real-world problems in different domains.

Nissatech developed two components: Communication and Message Broker and Workload Predictor.

Communication and Message Broker is made with opensource RabbitMQ broker and additional components (Hashicorp Vault and Bind9 DNS) for automatization. With specific configuration, we build high available cluster broker on the cloud that receives published messages over a secure layer and forwards them to subscribers. Messages are published from client-side directly to cluster of the nodes or through the edge (platform) deployed nodes. Edge nodes of the broker can be used for internal communication of components deployed in the same graph or/and for communication with the whole platform. Cluster workload is balanced with the usage of DNS and all nodes in our topology are dockerized for easier deployments. Broker supports AMQP/S, MQTT/S and IPv4/6 protocols.

Workload Predictor performs predictive analytics on time series. Conceptually, it connects real-time streaming and machine learning (ML). Due to nature of the learning algorithm (on-line prediction based on a limited time window), the computation can scale properly.

The solution is based on the data-analytics approach, which start with a deep understanding of the real-time dataset through an intensive data processing pipeline. The goal of the pipeline is to determine the main characteristics of the monitored data such as CPU and Memory usage and then to prepare data for robust predictions.

The main advantage is that in a rigorous data analytics pipeline, the dataset will be adapted to the conditions required by the prediction algorithm, ensuring the reliability of the whole prediction process.

Activeeon is an independant Software vendor company that assists their customer in workload automation. It has emerged in 2007 as spin'off from INRIA Rresearch institute, and addresses a field of expertise including High performance computing, cloud management and resource scheduling. The two edited softwares, ProActive Workflow automation (PWS) and ML Open Studio (MLOS), are designed to enact IT automation, big data processing, application orchestration, distributed computing and machines learning processing.

In the context of PrEstCloud, Activeeon has proposed its compentences in workload automation and compute resource management. The contributions are twofold:

- (i) The "Autonomic Data-Intensive Application Manager" components is responsible for orchestrating the placement and the execution of application fragments and stateless functions within PrEstoCloud in accordance with the constraints from components from the meta-management layer.

- (ii) The Autonomic Resource Manager is the component of the Control Layer which is responsible for the deployment and management of PrEstoCloud resources and applications. Initially based on the Resource Manager of ActiveEon, the Autonomic Resource Manager is designed to deploy and configure VMs in multiple clouds infrastructures (public and private) and make them available to all PrEstoCloud components, but was complemented to work tackle edge resource management.

Those two contribution are part of the constrol layer of the PrEstoCloud architecture, in charge of enforcing the determined business-logic policies on the lifecycles of cloud and edge resource supporting the operated PrestoCloud appliance.

The Artificial Intelligence Laboratory (ailab.ijs.si) is part od Jozef Stefan Institute, the leading research institution in Slovenia. AI lab, with approximately 60 researchers, is one of the largest European research groups involved in RTD projects and initiatives that introduce AI approaches and technologies to the field of big data analytics. The research topics include Artificial Intelligence, Machine Learning, Data-Mining, Text-Mining, Web-Mining, Multimedia Mining, Semantic Technologies, Social Network Analysis, Language Technologies, Natural Language Processing, Multi-lingual, Cross-lingual Technologies, Scalable, Real-time Data Analysis, Data Visualization, Knowledge Reasoning and recently Sensor Networks. Its key research direction is combining modern statistical data analytic techniques with more semantic/logic based knowledge representations and reasoning techniques with the purpose to make progress in solving complex problems such as text understanding, large scale probabilistic reasoning, building broad coverage knowledge bases, and dealing with scale. Its team consists of individuals with high competences in both research and technical development from data driven analytics to knowledge based systems. The department collaborates strongly with its spin-out companies that are providing real-life big data analytics solutions to companies such as Bloomberg, New York Times, British Telecom, Google and many others.

JSI as partner on the PrEstoCloud project has been responsible for "Task 3.4: Mobile offload processing". An efficient and reliable orchestration of distributed microservices in the PrEstoCloud framework is necessary that benefits from edge computing to upper layers of computing resources such as fog and cloud infrastructures. It means that the adoption of advanced IoT technologies has been enabled through the concept of applications' offloading from a computing node to another one, developed by JSI. To this end, JSI in the PrEstoCloud solution as Task 3.4 called "Mobile offload processing" offers two architectural components:

(i) On/Offloading Server and

(ii) On/Offloading Client.

This video is a training material to explain how to use the "Mobile offload processing" tool developed by JSI at the Artificial Intelligence Laboratory.

http://videolectures.net/prestocloud_task_mobile_offload/

ICCS is a non-profit private law body associated with the School of Electrical and Computer Engineering of the National Technical University of Athens (NTUA). ICCS was established in 1989 by the Ministry of Education of Greece in order to promote research and development activity in all diverse aspects of computer and telecommunications systems and their applications.

Fog computing environments are unstable and diverse. Edge devices with different hardware and software configuration have different processing capabilities and performance. Factors such as the available RAM, the type of memory chips, the number of CPU cores, frequency, type and generation, the current battery level (when the device lowers CPU clock frequency in order to lower energy consumption) or the speed of storage devices (HD, SSD, Flash) may affect the QoS (e.g. total execution time) of application and services in different ways. Considering the aforementioned characteristics of fog computing, with the term context we refer to any information (such as CPU, memory or utilization, network type and traffic, battery state, software and hardware configuration, etc.) that can help to infer conclusions about current and future state of fog devices.

A clear and accurate classification of context is helpful to uncover, understand, manipulate and sort out a variety of sensor-generated data to uncover useful information about the status of a computing resource and help DevOps make infrastructure management decisions. Many machine learning methods and mathematical representations of fog infrastructures have been used to convert low-level sensor data into higher-level context, which can facilitate decision making. Indicative methods used for context classification include Hidden Markov Models, Markov Chains, Bayesian Networks, Nearest Neighbour, Time Series, Threshold based Learning and Gaussian Mixture Models.

ICCS has analysed key methods of context classification and prediction and the application of these methods to facilitate fog infrastructure management. In the context of PrEstoCloud, ICCS has developed a Fog Context Analytics component, which can support infrastructure operators in case where there is an increased demand for management automation, robustness and overall management complexity reduction to achieve good quality of service and serve demands of different devices and applications. The software component is available at: https://gitlab.com/prestocloud-project/mobile-context-analyzer2

CNRS, the French National Centre for Scientific Research, is the largest fundamental science agency in Europe, with over 35,000 employees (researchers, engineers, administrative staff) across France.

CNRS has been involved in the high-level requirement analysis and design of the PrEstoCloud platform, as well as in the translation of the high-level (meta-management) constraints to low-level (e.g., device-based) constraints.

In PrEstoCloud, CNRS is represented by the associated I3S Laboratory located in Sophia-Antipolis, and, in particular, through the SigNet and SCALE teams.

The SigNet teams brings its extensive experience in computer networking and virtualization.

It is responsible for the conception, architecture, and implementation of a dynamic overlay network over hybrid clouds, spanning up to edge devices, through which deployed distributed use-case application components can exchange data securely.

The SCALE team brings its experience in application scheduling, in particular, its open-source scheduler BtrPlace (http://www.btrplace.org/).

BtrPlace has been extended as part of PrEstoCloud based on use-case application requirements for optimized placements, according to the idiosyncrasies of available hosting environments.

Ubitech is a leading, highly innovative software house, systems integrator and technology provider, established to provide leading edge intelligent technical solutions and consulting services to businesses, organizations and government in order to allow the efficient and effective secure access and communication with various heterogeneous information resources and services, anytime and anywhere. Ubitech enables real-time valid information processing and decision-making, the realization of intelligent business environments, and B2B and B2C transactions by providing high added-value business –oriented and –based solutions. Ubitech has been established in Athens, Greece back in 2005, concentrated initially in the Greek and Balkan market and acquiring several EC and national grants for novel R&D initiatives. Currently, Ubitech has extended its operations with targeted international activities through its subsidiaries, representation offices, business partners and affiliated companies in Limassol (Cyprus), Madrid (Spain), Buenos Aires (Argentina) and Guayaquil (Ecuador), concentrating mainly in the Spanish-speaking countries of Central and Latin America.

UBITECH is the integrator of PrEstoCloud and also took the initiative for the creation of an assistive web-based UI that ease users to achieve elastic deployments for processing data at both cloud and edge. In addition UBITECH developed and integrated a highly efficient security mechanism that can be installed in each node of the deployed service and is based on XPD/eBPF data filtering and Snort IDS. Also UBITECH provides a mechanism that allows connectivity with end-to-end encryption for cloud and edge devices, even on the go, based on latest mesh networking advancements.

Advanced Integrated Technology Solutions & Services Ltd is a Cyprus-based Small Medium Enterprise (SME) established in 2011. ADITESS is a scientific, consulting, and research company whose purpose is to conduct theoretical and applied research and to produce studies, at strategic and tactical level, on issues concerning Security policies, Critical Infrastructure Protection, Aftermath on crisis events, Transportation security and Border management (monitoring and surveillance), modeling solutions and to develop state of the art applied Security Solutions in the above-mentioned areas. ADITESS staff is composed of dedicated and experienced professionals who have outstanding professional and academic experience in the area of security, and have been involved in EU Research Programs as well as national and international projects for several years. Consisting and cooperating with consultants - researchers with professional and academic experience and a long-standing history in proposal preparation, project management and R&D activities, ADITESS LTD is at the forefront of our offering with clients stemming from a multitude of technological domains, such as Security, Defense, Information and Communication Technologies. Furthermore, a number of ADITESS researchers come with ex-military and ex-police forces background in IT Security, Electronic Warfare (EW) and Signal Intelligence (SIGINT) in large-scale National and International operations, over the last fifteen years. This field experience is enhanced by high academic knowledge especially in the IT and security (IT, UAV systems, Electro-optics, Communication, etc.) system's area.

Surveillance systems that capture video and audio in enterprise facilities and public places produce massive amounts of data while operating at a 24/7 mode. There is an increasing need to process, on the fly, such huge video and audio data streams to enable a quick summary of "interesting" events that are happening during a specified time frame in a particular location.

Through the PrEstoCloud Project, we enabled a novel and adaptive architecture that built on top of a distributed computing paradigm which is ideal for smart surveillance systems that uses resources at cloud, fog and edge.

LiveU Ltd. is an Israeli based startup company that has developed technology which enables live wireless high-quality multimedia transmissions from the field to any location. Our systems use any existing wired and wireless technology including cellular data streams, WiMAX and Wi-Fi hotspots, overcoming the inherent bandwidth limitations and fluctuations of these networks. Our main applications are live video transmission for cable and satellite TV (sport and news gathering) and IP TV/Internet portals streaming. Our products are currently being used by national and international TV News channels, Internet portals, Web TV and others. We deliver a lightweight, low-cost unit designed for every reporter or location. We currently transmit real-time video at an uplink bitrate of 1-2Mbps (SDTV) using MPEG-4 / H.264 for broadcast TV and WME (Windows Media Encoder) for Internet applications.

We are developing a solution for fast deployment, vehicles and man carried, that will integrate number of cameras and wireless links for Mobile Journalism applications where consumers can generate and exchange content, these live streams are likely to create large amount of data specifically at events. Within the context of PrEstoCloud project, we have developed the News Room Application which allows news companies, professionals and volunteers to work together on news making. The application is based on WebRTC framework and uses Kurento Media Servers. Besides, the News Room Application works with Video Analytics service developed by NAM.

NAM is an Israeli software development company providing expert development of custom software applications in the domain of video communications and telecommunication. We are involved in elastic cloud platform for media application, with mobile connection manager, with video in various formats such as HEVC, SVC, H264, and we focus on providing developments in the field of WebRTC and cloud services to various companies, in particular, for video conferencing and Mobile Journalism business cases.

Within the framework of PrEstoCloud project, we have developed the Video Analytics service integrated with LiveU's News Room Application. The Video Analytics service enables live streams analysis for detection of relevant news events such as appearance of famous politicians, actors etc. in video feeds. Such detection allows to minimize amount of transferred data by sending only relevant ones as a part of News Room Application work.

With the integration of the new Big Data system and by provisioning additional logistics services to customers, the data transfer and processing is expected to significantly rise in the future. Because of this, CVS Mobile exploits PrEstoCloud to control the data transfer to fog and cloud resources in a cost-efficient way.

CVS takes advantage of the following benefits:

- increasing fleet productivity which is an objective for the CVS Mobile. For example, telematics data is useful for evaluating and hence correcting driver behaviour that insurance companies hopefully begin to offer discounts for the CVS Mobile.

- telematics data allows the CVS Mobile to optimise maintenance schedules since developing custom maintenance schedules for each vehicle is necessary to identify when and where performance issues arise in a fleet vehicle. For example, the CVS Mobile can be able to determine when a fleet vehicle may need an oil change, new braking pads or a filter swap, and hence the CVS fleets will be capable of paying for maintenance more efficiently.

- decreasing the idle time which is a big problem for the CVS fleets. Comparing fuel consumption against miles driven is a way of identifying when idle engine activity is cutting into performance. For example, avoiding unnecessary stops can be reached by telematics technology.

- Minimising infrastructure costs through technologies by which services and data can be efficiently provisioned, monitored and migrated across federated cloud, fog and edge resources.

CONSORTIUM

Activeeon
SCALE BEYOND LIMITS

CVS
Advanced Telematics

software AG

ICCS

cnrs

aditess
Advanced Integrated Technology
Solutions & Services

UBITECH
ubiquitous solutions

Nissatech
innovation centre

IJS

N. Amram Technology

LiveU