Efficient Real-Time Big Data Processing at the Edge of the Network – The PrEstoCloud project

SALMAN TAHERIZADEH, Artificial Intelligence Laboratory, Jožef Stefan Institute, Ljubljana, Slovenia
BLAZ NOVAK, Artificial Intelligence Laboratory, Jožef Stefan Institute, Ljubljana, Slovenia
SEBASTJAN VAGAJA, CVS Mobile Inc., Ljubljana, Slovenia
MARIJA KOMATAR, CVS Mobile Inc., Ljubljana, Slovenia
MARKO GROBELNIK, Artificial Intelligence Laboratory, Jožef Stefan Institute, Ljubljana, Slovenia

The European Horizon 2020 PrEstoCloud [1] project creates substantial research contributions in real-time Big Data technologies in order to provide a dynamic, distributed architecture for proactive cloud resource management. In order to extract useful knowledge through real-time Big Data processing, the PrEstoCloud solution reaches the extreme edge of the network. The PrEstoCloud solution provides important benefits including (i) application development and deployment, (ii) integration and automation, (iii) data protection and privacy, (iv) cost optimization, and (v) big data management.

CCS Concepts: • **Computer systems organization → Cloud computing** • **Computing methodologies →** Distributed computing methodologies

KEYWORDS

Edge computing, Real-time analytics, Big Data

1 Brief description of the PrEstoCloud project

In recent years, a wide variety of software solutions, such as Internet of Things (IoT) applications, have emerged as cloud-based systems [1]. As a consequence, billions of users or devices get connected to applications on the Internet, which results in trillions of gigabytes of data being generated and processed in cloud datacenters. However, the burden of this large data volume, generated by end-users or devices, and transferred toward centralized cloud datacenters, leads to inefficient utilization of resources. To overcome above problem, edge computing framework is aimed at increasing capabilities of resources at the edge of the network compared to traditional centralized cloud architectures by not only placing services in the proximity of end-users or devices, but also using new data transfer protocols to improve the interaction with datacenter-based services. This also provides a low-latency response time for the application.

Edge computing, which is currently a hot research topic is a widespread method for providing many different types of Big Data technologies such as Internet of Things (IoT) systems [2]. Edge computing is a new computing paradigm optimizing applications to expand their data processing capabilities next to end-users rather than at the faraway centralized datacenters.

In this regard, the PrEstoCloud project creates a novel computing method and management solution for efficient deployment and execution of data-intensive applications orchestrated upon edge computing frameworks [3].

The PrEstoCloud solution provides important benefits including:

- Application development and deployment
- Integration and automation
- Data protection and privacy
- Cost optimization
- Big data management

2 Main innovations

The PrEstoCloud project aims to prototype the following innovations which emerge to meet requirements for real-time Big Data environments within edge computing scenarios:

- *Technologies to improve the execution of data-intensive applications:* The PrEstoCloud project provides a new solution based upon edge computing frameworks for different purposes: (i) to securely and reliably collect data from different types of sensors; (ii) to efficiently perform real-time Big Data processing and predictive analytics on IoT events and streams; (iii) to seamlessly upgrade enterprise applications and practices with IoT data; and (iv) to appropriately standardize the integration of sensors, objects or devices with enterprise applications.

- *Technologies to enhance the development productivity of data-intensive applications:* The PrEstoCloud project offers a new way of aiding data-intensive application providers to develop and customize their real-time Big Data processing systems based on edge computing frameworks. In the next decade, more than two billion sensors/objects/devices will get connected to the Internet. Therefore, this huge market obviously requires efficient solutions for deploying and customizing such business-oriented real-time data-intensive systems.

[1] PrEstoCloud, http://prestocloud-project.eu/

https://doi.org/10.0000/0000

- *Technologies to monitor data-intensive applications deployed upon edge computing frameworks:* The performance of data-intensive applications deployed upon edge computing frameworks varies depending on runtime conditions such as the workload density, availability and reliability of virtualized infrastructures, network connection quality between end-users and servers, and so on. Therefore, tracking dynamic changes of execution environments provided by the PrEstoCloud solution on the fly is necessary to identify any deterioration of system health.
- *Technologies to facilitate location-aware and context-driven adaptation recommender systems:* As an important advantage of modern cloud-edge solutions such as PrEstoCloud, tracking end-users' information for example their location, mobility and operational environment can be useful in order to offer fully customized services. In this way, various constraints for proper behavior of data-intensive applications (e.g. response time, security constraints, etc.) can be expressed during design-time, and further refined while running time, verified in real-time. This fact may appropriately support end-users' requirements and desires especially for data-intensive applications.

3 PRESTOCLOUD CONSORTIUM

The PrEstoCloud consortium, presented in Table 1, combines the resources of 11 partners categorized in four groups: (1) MP: Management Partner, (2) IP: Industrial Partners, (3) TP: Technology Providers, and (4) RI: Research Institutes.

Table 1. The PrEstoCloud consortium

Name	Category	Country	Website
Software AG	MP, TP	Germany	www.softwareag.com
CVS Mobile	IP	Slovenia	www.cvs-mobile.com
ADITESS	IP	Cyprus	www.aditess.com
LiveU	IP	Israel	www.liveu.tv
N.Amram Technologies	IP	Israel	amram-technologies
Nissatech	TP	Serbia	www.nissatech.com
ActiveEon	TP	France	www.activeeon.com
UBITECH	TP	Greece	www.ubitech.eu
IJS	RI	Slovenia	www.ijs.si
ICCS	RI	Greece	www.iccs.gr
CNRS	RI	France	http://www.cnrs.fr

REFERENCES

[1] Salman Taherizadeh and Marko Grobelnik, "A capillary computing architecture for dynamic internet of things: Orchestration of microservices from edge devices to fog and cloud providers", Journal Sensors 18(9), 2018. doi: 10.3390/s18092938.

[2] Salman Taherizadeh, Blaz Novak, Marija Komatar, and Marko Grobelnik, "Real-Time Data-Intensive Telematics Functionalities at the Extreme Edge of the Network: Experience with the PrEstoCloud Project", The 42nd Annual IEEE International Conference on Computers, Software and Applications (COMPSAC 2018), Tokyo, Japan, 2018. DOI: 10.1109/compsac.2018.10288.

[3] Salman Taherizadeh and Vlado Stankovski, "Dynamic multi-level auto-scaling rules for containerized applications", The Computer Journal, Oxford University Press, 2018. doi: 10.1093/comjnl/bxy043.