Cloud and Data Federation in MobiDataLab

Emanuele Carlini, Patrizio Dazzi, Francesco Lettich, Raffaele Perego, Chiara Renso
Information Science and Technologies Institute "A. Faedo"
National Research Council of Italy (ISTI-CNR)
Pisa, Italy
{name.surname}@isti.cnr.it

ABSTRACT

Today's innovative digital services dealing with the mobility of persons and goods produce huge amount of data. To propose advanced and efficient mobility services, the collection and aggregation of new sources of data from various producers are necessary. The overall objective of the MobiDataLab H2020 project is to propose to the mobility stakeholders (transport organising authorities, operators, industry, government and innovators) reproducible methodologies and sustainable tools that foster the development of a data-sharing culture in Europe and beyond. This short paper introduces the key concepts driving the design and definition of the Cloud and Data Federation that stands at the basis of MobiDataLab.

CCS CONCEPTS

• **Information systems** → **Data federation tools**; *Information systems applications*; • **Computer systems organization** → **Cloud computing**.

KEYWORDS

Mobility, Data Federation, Cloud Federation

1 INTRODUCTION

Mobility Data provides information on urban mobility patterns like trasportation networks, timetable information, car and traffic to name a few. The Europe vision for common mobility data space is for a *"single data space as a market for data [..] where personal as well as non-personal data [..] are secure and [..] boosting growth and creating value, while minimising the human carbon and environmental footprint."*[1] This will allow different stakeholders to share their data into a single platform to facilitate access, pooling and sharing of data.European Green Deal[2] and the European Data Space[3] also

---

[1]https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX %3A52020DC0066
[2]https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en
[3]https://digital-strategy.ec.europa.eu/en/policies/strategy-data

---

strictly depends on the capacity of organizations to digitalize and share mobility data as a pillar toward the decarbonization of the transportation sector, also improving of intermodal connections in transport hubs for optimal level of vehicle availability in car and bicycle sharing systems. A mobility data sharing ecosystem relies on fundamental data-related policies and regulatory measures including the establishing of data protection regulations and standardized processes. Such a data space will facilitate access, pooling and sharing of data from existing and future transport and mobility databases. A mobility data sharing infrastructure should consider several layers: (i) data collection and merging, (ii) data standards, (iii) data infrastructure, (iv) governance and accountability, and (v) use monitoring and analysis. Data collection and merging refers to the collection of data generated from diverse sources assembled within a data sharing initiative. Data infrastructure refers to policies for the development of the infrastructure for the management of data resources. Data standards include the adoption of metadata standards to ensure interoperability. Use monitoring and analysis refers to policies and tools to enable public, private, or other third parties to monitor data accesses to ensure the ethical usage of data to protect public interests.

2 CLOUD AND DATA FEDERATION

Almost all the existing IT behemoths offer their own cloud solutions, each operating according to specific models, e.g., communication protocols and virtualization formats. This leads to the well know issue of vendor lock-in and hinders the interoperability between different providers. These aspects are particularly relevant when, like in MobiDataLab, there is a need for conducting data analytics on data hosted in different clouds, possibly in different countries. To tackle these issues, several standards have been proposed to ease cooperation among different clouds [8]. Each of these standards cover, however, only specific aspects and no standard is universally adopted by cloud providers. A different approach is to federate heterogeneous clouds. Many approaches to Cloud Federations have been proposed so far [2, 3]. These proposals range from general approaches [4, 6], to rather specific ones [13]. According to the NIST [4], the characteristics that define a Cloud Federation are four:

- "a virtual security and collaboration context that is not necessarily owned by any one user or organization." – in MobiDataLab project distinct entities are willing to collaborate for mobility data sharing, without envisioning the need for building a single infrastructure.
- "participating entities have membership in the federation and identity credentials that are linked to each member." –

---

[4]https://www.nist.gov/publications/nist-cloud-federation-reference-architecture

the actors of MobiDataLab project will be granted with an access to the data provided by other parties.

- "Users, sites, and organizations can participate in a federation by choosing to share some of their resources and metadata." – Not necessarily all the partners involved in MobiDataLab are willing to share all the mobility data they own.
- "Participating members agree upon the common goals and governance of their federation – MobiDataLab is defined around a clear goal: incentive the data sharing culture in Europe for a improved management of transport systems.

The Federated approach is the one adopted in MobiDataLab to design its cloud architecture.

3 THE MOBIDATALAB VISION

The European Cloud Initiative [5], provides research funds in many research fields in a tentative that "aims to strengthen Europe's position in data-driven innovation, improve its competitiveness and cohesion, and help create a Digital Single Market in Europe." Indeed, several EU-connected initiatives have been dealing with data federations. Although not necessarily focusing specifically on transport and mobility data, these initiatives shape the technological landscape for data exchange and inter-operability. A major initiative in this direction is project GAIA-X [7] aimed at providing common requirements and a trusted infrastructure to allow secure and sovereign data operations, by means of certification and verifiability of the operations done by the actors in the system. In the context of the Next Generation Internet initiative of the European Commission, the Corteza project [12] is an open-source platform for data exchange and management, which mostly focuses on commercial enterprises. FENIX [10] aims to develop an European federated architecture for transportation and logistics data sharing, specifically designed for shippers, logistics service providers, mobility infrastructure providers, cities, and authorities. Several other H2020 EU funded projects have dealt with the building and maintenance of data federations. The SUNFISH project [11], focuses on the creation of a Federation-as-a-Service to enable the exchange of data and services in a secure and controlled fashion based on blockchain technology. The BEACON project [9] has dealt with the development of the federation of network resources and the secure exchange of data and deployment of federated cloud applications. The EU-Korea H2020 BASMATI Project [1] aimed at providing an ecosystem that integrates cloud federation with mobile devices. As such, it offers a reasonable reference and knowledge ground on top of which to develop the original contribution envisioned by MobiDataLab for federated mobility data sharing and processing.

The transport domain is witnessing a great increase in the use of digital technologies enabling new digital services in the public and private sector. The number of users grows at an exponential rate also thanks to the use of smartphones.The consequent sharing and aggregation of these different sources of data from various producers is necessary for enhancing the Intelligent Transport Systems (ITS) vision. It is essential to build on Open Data, promoting their sharing between the various stakeholders and developing tools for high added value reuse. The overall objective of MobiDataLab project[5] is to propose to the mobility stakeholders a methodology

and tools that foster the development of a data sharing culture in Europe and beyond. One of the pillars of the MobiDataLab approach is the *Transport Cloud*: a cloud-based prototype platform for sharing transport data, technically designed according to federated cloud principles to facilitate the access to mobility data, in an open, inter-operable and privacy-preserving way. The Transport Cloud is intended as the technological backbone of MobiDataLab and it is designed to support several core objectives. First, the deployment of a catalogue of reference Open Data compliant with current standards for data and service discovery; Second, a seamless and privacy-preserving access to mobility data sources, possibility in different formats and by using different connection technologies; Third, the execution of *data processors* that add value to the data, such as semantic enrichment, anonymization, translations from and to various data format.

In the near future we plan to define the architecture of the MobiDataLab cloud federation, that will be evaluated on the basis of the needs of the MobiDataLab stakeholders.

REFERENCES

[1] Jörn Altmann, Emanuele Carlini, Massimo Coppola, Patrizio Dazzi, Ana Juan Ferrer, Netsanet Haile, Young-Woo Jung, Dong-Jae Kang, Iain-James Marshall, Konstantinos Tserpes, et al. 2016. BASMATI-a brokerage architecture on federated clouds for mobile applications. In *CGW 2016 International Workshop, Krakow, Poland.*

[2] Gaetano F Anastasi, Emanuele Carlini, and Patrizio Dazzi. 2013. Smart cloud federation simulations with cloudsim. In *Proceedings of the first ACM workshop on Optimization techniques for resources management in clouds.* 9–16.

[3] Marcio RM Assis and Luiz Fernando Bittencourt. 2016. A survey on cloud federation architectures: Identifying functional and non-functional properties. *Journal of Network and Computer Applications* 72 (2016), 51–71.

[4] Emanuele Carlini, Massimo Coppola, Patrizio Dazzi, Laura Ricci, and Giacomo Righetti. 2011. Cloud federations in contrail. In *European Conference on Parallel Processing.* Springer, Berlin, Heidelberg, 159–168.

[5] European Commission. 2020. *The European Cloud Initiative.* Retrieved April 2021 from https://ec.europa.eu/digital-single-market/en/european-cloud-initiative

[6] Massimo Coppola, Patrizio Dazzi, Aliaksandr Lazouski, Fabio Martinelli, Paolo Mori, Jens Jensen, Ian Johnson, and Philip Kershaw. 2012. The Contrail approach to cloud federations. In *Proceedings of the International Symposium on Grids and Clouds (ISGC'12)*, Vol. 2. 1.

[7] GAIA-X. 2021. *GAIA-X: A Federated Data Infrastructure for Europe.* Retrieved April 2021 from https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html

[8] Ibrahim Mansour, Reza Sahandi, Kendra Cooper, and Adrian Warman. 2016. Interoperability in the heterogeneous cloud environment: a survey of recent user-centric approaches. In *Proceedings of the International Conference on Internet of things and Cloud Computing.* 1–7.

[9] Rafael Moreno-Vozmediano, Eduardo Huedo, Ignacio M Llorente, Rubén S Montero, Philippe Massonet, Massimo Villari, Giovanni Merlino, Antonio Celesti, Anna Levin, Liran Schour, et al. 2015. BEACON: A cloud network federation framework. In *European Conference on Service-Oriented and Cloud Computing.* Springer, 325–337.

[10] Fenix Network. 2019. *FENIX – A European Federated Network of Information eXchange in LogistiX.* Retrieved April 2021 from https://fenix-network.eu/fenix_european_architecture_for_logistics_pr

[11] Francesco Paolo Schiavo, Vladimiro Sassone, Luca Nicoletti, Andreas Reiter, and Bojan Suzic. 2016. Faas: Federation-as-a-service: The sunfish cloud federation solution. In *FaaS: Federation-as-a-Service: The SUNFISH Cloud Federation Solution.*

[12] Crust Technology. 2021. *The Corteza project.* Retrieved April 2021 from https://cortezaproject.org

[13] Xiaoyu Yang, Bassem Nasser, Mike Surridge, and Stuart Middleton. 2012. A business-oriented cloud federation model for real-time applications. *Future Generation Computer Systems* 28, 8 (2012), 1158–1167.

[5]http://www.mobidatalab.eu