1. Publishable Summary

The goal of the FERARI project is to provide the means for building large scale distributed systems that

- provide support for large scale services,
- provide Complex Event Processing technology for business users in Big Data architectures,
- integrates machine learning tasks in the architecture,
- provide support for flexible and adaptive analytics workflows and
- exemplifies the potential of the new architecture in the telecommunications and cloud domain.

In the FERARI project, we develop a reference architecture and framework that combines in-stream complex event processing (CEP) with in-situ methods (i.e., methods that process data right where it is generated), which are a first choice for scalable processing of large scale services. This way, the project brings stream processing much closer to the business world providing a seamless model that applies CEP as part of Big Data applications in a way easily consumable by business users. Additionally, distributed machine learning is integrated into the architecture to allow learning sophisticated models from the data streams. These functionalities are based on adaptive and event-driven methods, i.e., it is possible to adapt workflows within the architecture to changing data distributions. This architecture

- is implemented and made available as an open source framework and
- is the basis for the solutions FERARI provides for the telecommunications and cloud domain.

The beginning of the second reporting period was the 2st of February 2015.Overall, all project objectives have been met. All deliverables have been produced. Phase 2 – the reporting period - has as its major objective the release of an open-source architecture prototype as well as setting up first versions of the application scenarios. Further goals are setting up a test bed for the application scenarios as well as further development on the architecture, communication-efficient algorithms, and the flexible event processing application. During the second year, the major results have been:

1. the release of an open-source architecture prototype on bitbucket;
2. setting up a test bed and a first version of the application scenarios on a dedicated FERARI server at HT in Zagreb;
3. releasing a first open-source version of the distributed online learning framework as part of the architecture prototype;

4. and the development of “The Event Model” (TEM), i.e., a new way to model, develop, validate, maintain, and implement event-driven applications in a consumable fashion.

The definition of the architecture of the prototype as well as its further development and the distributed learning framework is provided in deliverable D2.2. The deliverable contains a detailed description of the reference architecture and shows how the building blocks, i.e., the application scenarios and the approaches developed in WP3-5, are embedded within this architecture. The architecture prototype can be reached from the FERARI open-source repository (https://bitbucket.org/sbothe-iais/ferari).

In order to implement solutions for the application scenarios and evaluate these solutions in terms of HT specific KPIs, HT set up a test environment on a dedicated server at HT in Zagreb. A first version of the FERARI solutions for the application scenarios is implemented on the server. For the mobile fraud scenario, a dashboard has been developed that helps HT personnel to track, monitor, and analyze information about detected frauds that help making decisions about future steps. A description about the application scenarios, the test bed and the dashboard can be found in deliverable D1.2.

A major goal of the FERARI project is to make in-stream and big data complex event processing accessible to business users. For that, it is essential to provide an abstract event processing model that allows to develop and maintain event-driven applications in a consumable fashion. To that extend, an event model has been developed that follows the model driven engineering approach and is based on the concept computing paradigm, i.e., all model artefacts are considered concepts. adsf

The major scientific contributions in the second year involve a novel technique for latent fault detection, in-situ anonymization methods, novel monitoring and sketching techniques for distributed stream monitoring and a novel way for adaptively optimizing queries in complex event processing. They can be found in deliverables D3.2, D4.2, and D5.2.

In summary, there has been great progress on implementing a FERARI framework – in particular the architecture prototype - which led to new and exciting ideas that resulted in high quality publications in top journals and conference. The open source release of the architecture prototyperaised interest from the press as well as from the industry and led to many invited talks at major companies such as Siemens and Deutsche Telekom. The consortium is collaborating intensively and is working in an atmosphere of partnership and open exchange of ideas.

Further information about FERARI can be found under www.ferari-project.eu.

Project funded by the European Community under the
Information and Communication Technologies Programme
Contract ICT-FP7-255951