PROJECT PERIODIC REPORT

Grant Agreement number: 619525

Project acronym: QUALIMASTER

Project title: A configurable real-time data processing infrastructure mastering autonomous quality adaptation

Funding Scheme: Small or Medium Scale focused Research Project (STREP)

Date of latest version of Annex I against which the assessment will be made: 2013-10-13

Periodic report: 1st **X**   2nd ☐   3rd ☐   4th ☐

Period covered: from **01st January 2014**   to **31st December 2014**

Name, title and organisation of the scientific representative of the project's coordinator:

Dr. Claudia Niederée

L3S Research Center

Appelstr. 9a

30167 Hannover Germany

Tel: +49 511 762 17796

Fax: +49 511 762 17779

E-mail: niederee@L3S.de


Project website address: www.qualimaster.eu

3.1 Publishable Summary

The growing number of fine-granular data streams opens up new opportunities for improved risk analysis, situation and evolution monitoring as well as event detection. However, there are still some major roadblocks for leveraging the full potential of data stream processing, as it would, for example, be needed the highly relevant systemic risk analysis in the financial domain.

The QualiMaster project will address those road blocks by developing novel approaches for autonomously dealing with load and need changes in processing pipelines for large-scale data streams, while opportunistically exploiting the available resources for increasing analysis depth whenever possible. For this purpose, the QualiMaster infrastructure will enable autonomous proactive, reflective and cross-pipeline adaptation, in addition to the more traditional reactive adaptation. Starting from configurable stream processing pipelines, adaptation will be based on quality-aware component description, pipelines optimization and the systematic exploitation of families of approximate algorithms with different quality/performance trade-offs. However, adaptation will not be restricted to the software level alone: We will go a level further by investigating the systematic translation of stream processing algorithms into code for reconfigurable hardware and the synergistic exploitation of such hardware-based processing in adaptive high performance large-scale data processing.

The project focuses on financial analysis based on combining financial data streams and social web data, especially for systemic risk analysis. Our user-driven approach involves two SMEs from the financial sector. Rigorous evaluation with real world data loads from the financial domain enriched with relevant social Web content will further stress the applicability of QualiMaster results.

The QualiMaster Infrastructure
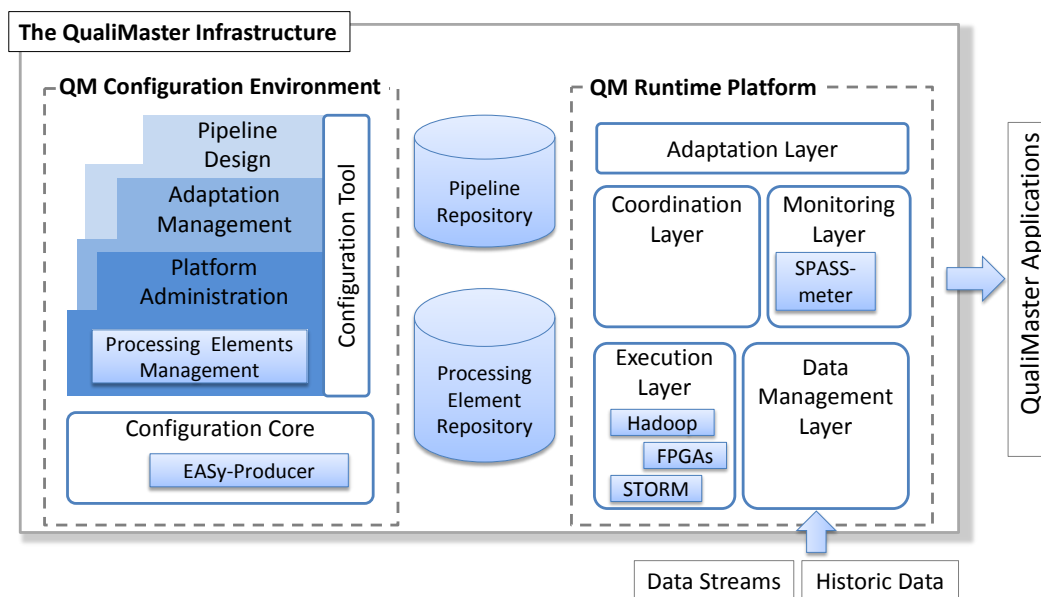
In the first year of the QualiMaster project the focus was on laying the foundations for the QualiMaster infrastructure. Starting from a systematic requirement analysis, which resulted in the identification of 10 key actors and 30 detailed use cases, the architecture for the QualiMaster infrastructure has been defined (see figure above for a high-level architecture overview). Work on the QualiMaster infrastructure and its architecture included the identification and grouping of the core components, e.g., into the various layers of the runtime environment (adaptation layer, coordination layer, monitoring layer, execution layer and data management layer), and the definition of the technology stack to be used (including Apache STORM and Apache Hadoop). Furthermore, the infrastructure has been set up and populated with components and

tools, in order to enable the construction and operation of first data stream processing pipelines. For focusing the activities in the project a QualiMaster priority pipeline has been define, on which all teams of the QualiMaster team work together.

Infrastructure and Pipeline Configuration

For setting up adaptive pipelines in a safe and flexible way, a domain specific configuration tool has been developed. This tool embraces the configuration of the QualiMaster infrastructure, the configuration of specific data stream processing pipelines based on algorithmic families as well as their verification, instantiation into source code, packaging and deployment as STORM topologies. This configuration tool builds upon a configuration core, which implements the configuration functionality based on concepts from software product-line development. Thereby, the QualiMaster configuration approach relies on a quality-aware configuration meta model (variability model) with validation constraints as well as an instantiation specification acting as a foundation for flexible evolution and adoption of the QualiMaster technology.

Algorithmic Families in Software and Re-configurable Hardware

Adaptation is enabled in QualiMaster by providing families of exchangeable algorithms with different performance/quality tradeoffs. In the first year of the project, algorithmic families have been developed for financial data analysis, analysis of social web content and streams as well as for data stream processing in more general. For financial analysis, work has focussed on the development of advanced correlation estimators, which provide core functionality for systemic risk analysis. In the area of social web analysis work has focussed on the analysis of relevant Twitter data streams and especially on sentiment analysis, which is expected to give additional signals in support of systemic risk analysis. Some of the algorithms -- including the correlation estimator and the SVM machine learning support for sentiment analysis -- have also already been implemented in re-configurable hardware adding a possible switch to hardware to the respective algorithmic families. These first implementations in hardware were also used to analyse, which algorithms are promising for translation to re-configurable hardware as well as for investigating a systematic translation approach.

Foundations for Adaptation

In addition to realizing the first algorithmic families, the QualMaster team has also worked on defining the adaptation approach, which will be followed in the QualiMaster project. In this context, the concepts for a flexible adaptation language enabling domain and application specific adaptations have been developed and enactment patterns have been investigated, which are crucial for implementing adaptation. Foundational work for the QualiMaster adaptation approach also included the definition of an initial quality taxonomy, which is the basis for the description of quality properties of members of algorithmic families as well as for the formulation of quality-based adaptation strategies. In addition, first mechanisms for integrating the adaptation support into the infrastructure have been implemented.

QualiMaster Applications

Systemic Risk analysis and other types of financial risk analysis have been selected as primary application domains for the QualiMaster infrastructure in the project. In close collaboration with the two application partners from the financial domain, Maxeler and Spring Techno, first application scenarios and financial analysis algorithms in support of those scenarios have been selected. Furthermore, work has been started on a design environment for stakeholder applications, which eases the creation of customized stakeholder application. Those applications make the results created by the data stream processing pipelines available in a user friendly way and enable the users further customize their application as well as to manipulate the underlying processing.

Expected impact

Intelligent processing of stream data clearly will gain in importance in the coming years. This is a logical consequence of the dramatic increase in the number of data streams that has become available in a variety

of areas and of the potential benefit that an adequate and timely processing of this data can bring. Actually, those sensor data and other data streams cover a much finer granularity and capture situations and events in many areas much better. This opens unpreceded opportunities for risk prediction, situation analysis and monitoring, early detection of events, monitoring and analysis of developments as well as triggering of reactions early. QualiMaster will have various types of impact in improving current data processing infrastructures and in helping to exploit the full potential of data stream processing. For example, the targeted systematic translation of parts of stream processing algorithms into hardware will further improve stream processing efficiency. Together, the contributions of QualiMaster will have an economic impact by helping European players in the Big Data market by increasing competitiveness (better performance through improved algorithms and intelligent exploitation of reconfigurable hardware), by enabling more sustainable data stream processing infrastructures (cf. better exploitation of available resources through adaptation) and by easing data processing service innovation (cf. easy configuration of new types of data processing pipelines)

The financial domain and especially the early recognition of risks in the financial domain has proven a crucial factor for the economic stability of Europe as well as for the overall stability of Europe, as has been seen in the recent crisis in Greece and Cyprus. The solution for automated systemic risk analysis, which is developed in QualiMaster, addresses the complex but highly relevant task of recognizing systemic risks, which due to their nature have a very high impact (effecting entire market sectors), but are difficult to detect. The QualiMaster support for systemic risk especially addresses regulatory bodies, which can intervene at an early point in case early warnings for systemic risk are detected. Moreover, institutional users in the small but entire societies in the large would profit from a more precise and longer prediction of financial risks. Furthermore, the stronger interweaving of financial analysis with observing and analyzing relevant social web developments, as it is addressed in the QualiMaster project, will result in more comprehensive financial analysis results, which also take into account the broader context of financial developments and decisions. In addition, a stronger consideration of societal aspects and grass-roots citizen interests can also be a first step towards better socio-economic systems in Europe, which, in the long run, is an important factor for the stability and well-being of the European society.

The QualiMaster Consortium:

- L3S Research Center, Gottfried Wilhelm Leibniz Universität Hannover
- Maxeler Technologies Limited
- Stiftung Universität Hildesheim
- Telecommunication Systems Institute
- Spring Techno GmbH & Co KG

Contact: Dr. Claudia Niederée (Coordinator), niederee@L3S.de

The results of the project work can be followed on the project's Web page: www.qualimaster.eu