Edge Computing for Complex Event Processing Systems

Abstract In complex event processing (CEP), large amounts of data have to be processed to generate business value. The most important dimension for value creation is latency between event and decision taken. The proposed approach reduces data latency and analysis latency by adding a CEP system where the data is generated to preaggregate events to complex events and detects state changes at the data source and reduce the amount of data sent over network, and thus, reduce the latency.

Connection to the lecture

During the talk on efficient complex event processing, Professor Weidlich introduces complex event processing systems as classical stream processing systems, like Apache Flink, that handle more complex tuples. The handled tuples are complex events, which consist of combinations of recorded events, and thus can have different types. Handling complex events to make business decisions is a time critical process. Thus, latency from event occurrence to the taken decision needs to be reduced to the minimum possible extent. The latency can be split in three parts. The first one is **data latency**, which describes the latency between event recording and starting to process the event. **Analysis latency** describes the time matching the given event to a pattern and **decision latency** describes the time for making the decision [1].

An adjacent lecture given by Professor Markl dealt with handling big data in general, where he focused on reducing two kinds of latency present in data processing systems. **Human latency**, the amount of time to solve the problem and write the required code, and **technical latency**, which describes the program runtime until obtaining the result. One approach to reduce the technical latency, presented in the lecture, was to build a system for data processing on the edge for internet of things devices [2]. The related research project is called **NebulaStream** and proposes strategies to handle data processing in IoT networks [3].

Goal

In the domain of complex event processing, efficiency plays a key role in decision making based on data. The goal of this research project is to reduce data latency and analysis latency to reduce the time between event occurrence and decision. We evaluate how composing and handling complex events at the data source, and thus reducing the amount of propagated data, influences the latency of CEP systems.

Problem

CEP systems handle huge amounts of data and derive business decisions from it. Individual raw events do not directly lead to decisions, but a composition of single events to complex events. Such complex events are matched to patterns to take actions. Most raw events can be discarded and do not contain relevant information. Even though, the event streams are sent to the CEP system to be handled.

Composing complex events at the source can decrease decrease the system latency by the factor 15, as proposed by several research groups [4,5,6]. The mentioned related work focuses on decreasing data latency. Analysis latency, the time to match patterns by processing complex events, is not reduced in their approach. With the increasing amount of data to be handled, e.g., in the domain of IoT sensors and more capable edge computing devices, more processing can be done where the data is generated. Thus, analysis latency can be reduced, as well.

Solution

The proposed approach to reduce data latency and analysis latency is to bring the computation to the edge by not only composing complex events where the data is generated to reduce the amount of information sent, but also start the analysis process when generating the data. Thus, state transitions can be detected at the source without transferring the data to the one CEP system. This extension is possible due to the increasing computational capabilities of IoT devices.

A lightweight CEP system is added where the data is produces, as shown in the graphic below. As a result, the system only propagates relevant state transitions to the main CEP system and reduces the information sent to the main CEP system.

We propose to build a lightweight CEP system , which can be run where data is produced. This system can use proposed methods by **NebulaStream** [3], such as in-network preprocessing of the data and hardware targeted code generation to decrease latency.

To evaluate the performance of such a system, we measure the data latency, analysis latency and decision latency and compare it with traditional CEP systems and the approaches mentioned above [4,5,6].

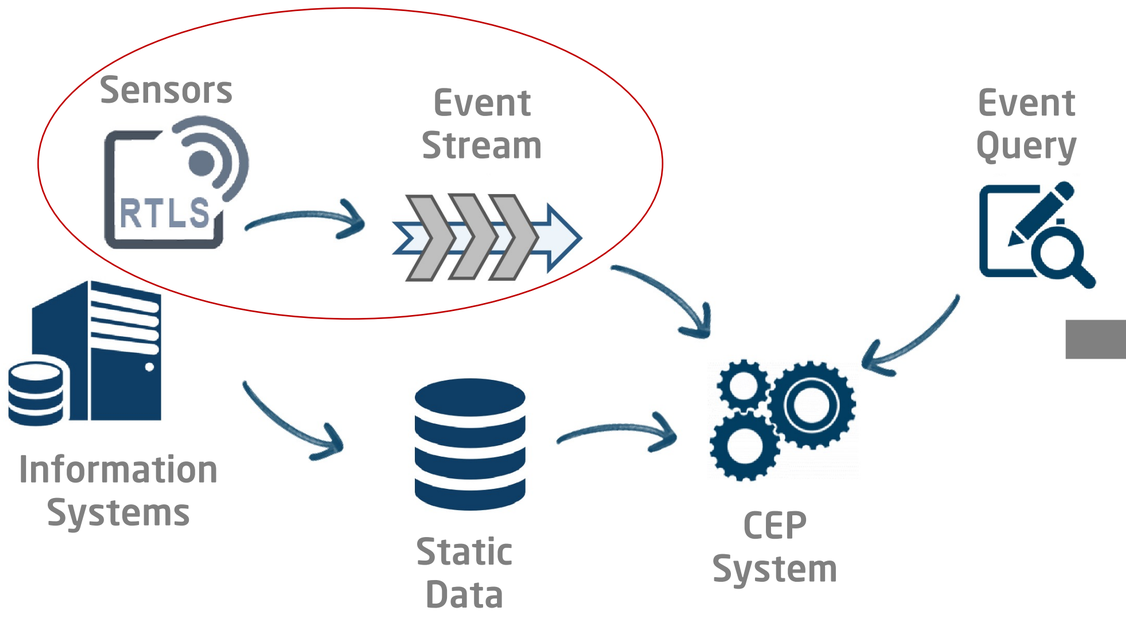

Fig. 1: Complex Event Processing System – Focus Area

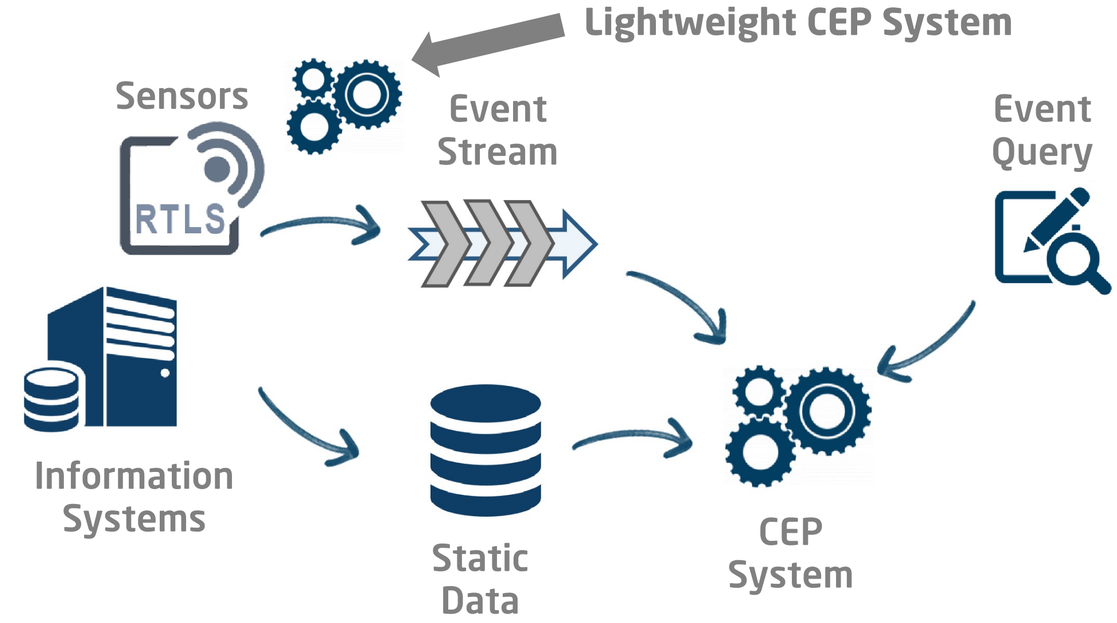

Fig. 2: Complex Event Processing System – Proposed Approach

Author
Hendrik Patzlaff
Master Student

Lecture Series on Database Research
Winter Term 2021/22
Hasso Plattner Institute

Sources:

[1] Weidlich, Matthias. "Efficient Complex Event Processing." Lecture Series on Database Research 2021/22, 18.01.2022.
[2] Markl, Volker. "Data Infrastructures." Lecture Series on Database Research 2021/22, 04.01.2022.
[3] Zeuch, Steffen, u. a. „The NebulaStream Platform: Data and Application Management for the Internet of Things". Conference on Innovative Data Systems Research (CIDR), 2020.
[4] L. Lan, R. Shi, B. Wang, L. Zhang and N. Jiang, "A Universal Complex Event Processing Mechanism Based on Edge Computing for Internet of Things Real-Time Monitoring," in IEEE Access, vol. 7, pp. 101865-101878, 2019, doi: 10.1109/ACCESS.2019.2930313.
[5] Luthra, Manisha, u. a. „Efficient Complex Event Processing in Information-centric Networking at the Edge". arXiv:2012.05070 [cs], Dezember 2020. arXiv.org, http://arxiv.org/abs/2012.05070.
[6] Dhillon A., Majumdar S., St-Hilaire M., El-Haraki A. (2021) Complex Event Processing in Sensor-Based Environments: Edge Computing Frameworks and Techniques. In: Mukherjee A., De D., Ghosh S.K., Buyya R. (eds) Mobile Edge Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-69893-5_20

Figures are based on [1] and modified by the author

Hasso Plattner Institut