



**CS5990 Masterarbeit Informatik**

Malte Schmitz, ISP ([schmitz@isp.uni-luebeck.de](mailto:schmitz@isp.uni-luebeck.de))

César Sánchez, IMDEA Software Institute ([cesar.sanchez@imdea.org](mailto:cesar.sanchez@imdea.org))

7. July 2017

# Parallel Analysis of Large Traces with Stream Based Specifications

*Analyzing very large log files according to certain specifications is a common task in debugging distributed embedded systems. E.g. in the automotive industry recording traces from communication buses and engine controlling processors during a test drive of some hours can easily produce log files in the size of petabytes. This thesis is about applying runtime verification techniques to such giant log files.*

## Problem Statement

We are working on runtime verification technologies with stream-based specification languages. With stream-based specification languages like LOLA [2] you can specify correctness properties, but also aggregate data and perform quantitative analysis on input streams. Log files can be used as input streams for a Runtime Verification [3] engine which then analyzes a log file with a given specification. The need to process giant log files with such approaches grows with the evolution of complex distributed systems. The goal of this thesis is to investigate in how parallel programming approaches like MapReduce [1] can be applied to enable the parallelization of such Runtime Verification engines.

## Goals

- Work your way into our specification language TeSSLa
- Investigate how MapReduce can be applied to checking large traces against TeSSLa specifications.
- Implement a prototype and evaluate the approach.

## Requirements

- You have some knowledge in declarative programming and formal languages and you either already know about runtime verification (e.g. CS4139) or you are at least interested in formal methods.
- You are either interested in or you already have some knowledge of parallel programming (e.g. CS3051).

## IMDEA Software Institute

This thesis will be done in close collaboration with the IMDEA Software Institute in Madrid<sup>1</sup>. The thesis can be combined with an ERASMUS stay in Spain.

## Literature

- [1] Jeffrey Dean, Sanjay Ghemawat: *MapReduce: Simplified Data Processing on Large Clusters*.
- [2] Ben D'Angelo, Sriram Sankaranarayanan, César Sánchez, Will Robinson, Bernd Finkbeiner, Henny B. Sipma, Sandeep Mehrotra, Zohar Manna: *LOLA: Runtime Monitoring of Synchronous Systems*. TIME 2005: 166–174
- [3] Martin Leucker: *Teaching Runtime Verification*. RV 2011: 34–48

---

<sup>1</sup> <https://software.imdea.org/>