

DADSTORM

A simple, fault-tolerant and real-time stream processing system

DAD 2016-2017
Group 12:
Daniel Fermoselle nº 78207
João Marçal nº 78471
Tiago Rodrigues nº 78692

Abstract

DADSTORM is a simple but reliable stream processing system. It's mainly used by Instituto Superior Tecnico Students.

The main features are: the 3 possible semantics of tuple processing it can have, fault-tolerance to $f+1$ faults per operator with a synchronous model of detection, 3 modes of tuple routing and last but the not the least 5 types of operators. This system is composed by a Puppet Master, Process Creation Service, Operators with their Replicas, Tuples and ThreadPools.

1 Introduction

Nowadays while streaming more and more information is added and we want to process it as fast as possible as well as to get the desired information even though existing the possibility of existing faults. In order to get that information in a reliable way we developed DADSTORM. This system process tuples based on

2 Programming Model

3 DADSTORM Abstractions

3.1 Tuple

3.2 Operators and Replicas

3.3 Puppet Master

3.4 Process Creation Service

4 Architecture and Implementation

4.1 Tuple

4.2 Operators and Replicas

4.3 Puppet Master

4.4 Process Creation Service

5 Discussion

6 Production Experiences

7 Evaluation

8 Conclusion