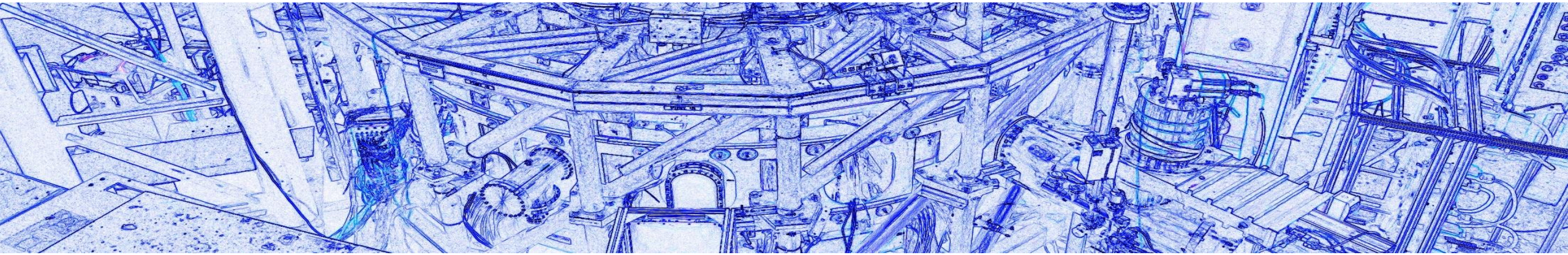
A Database Dedicated to the Development of Machine Learning Based Disruption Predictors

Wu Qiqi, Zheng Wei and J-TEXT team

IAEA-TM CODAC 2019

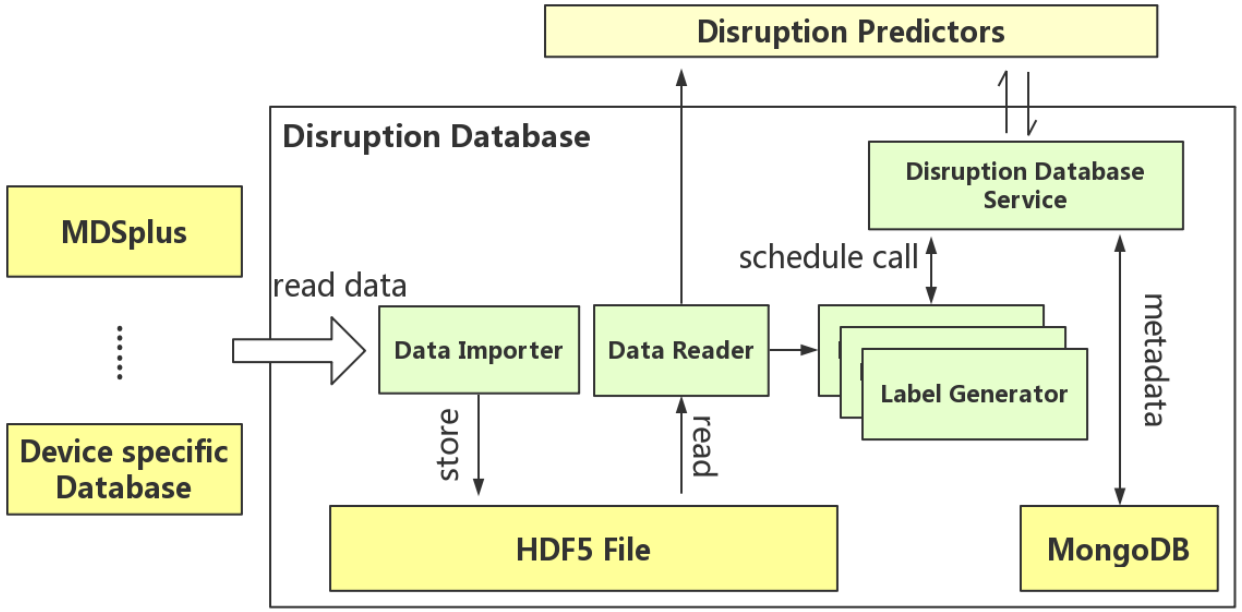


Introduction

- **In order to develop the disruption prediction algorithms conveniently, we created the disruption database(DDB).**
- **It allows developers to design disruption predictor without the consideration of complex processing of data**
- **It provides a data searching, data filtering and predictive performance evaluation function**

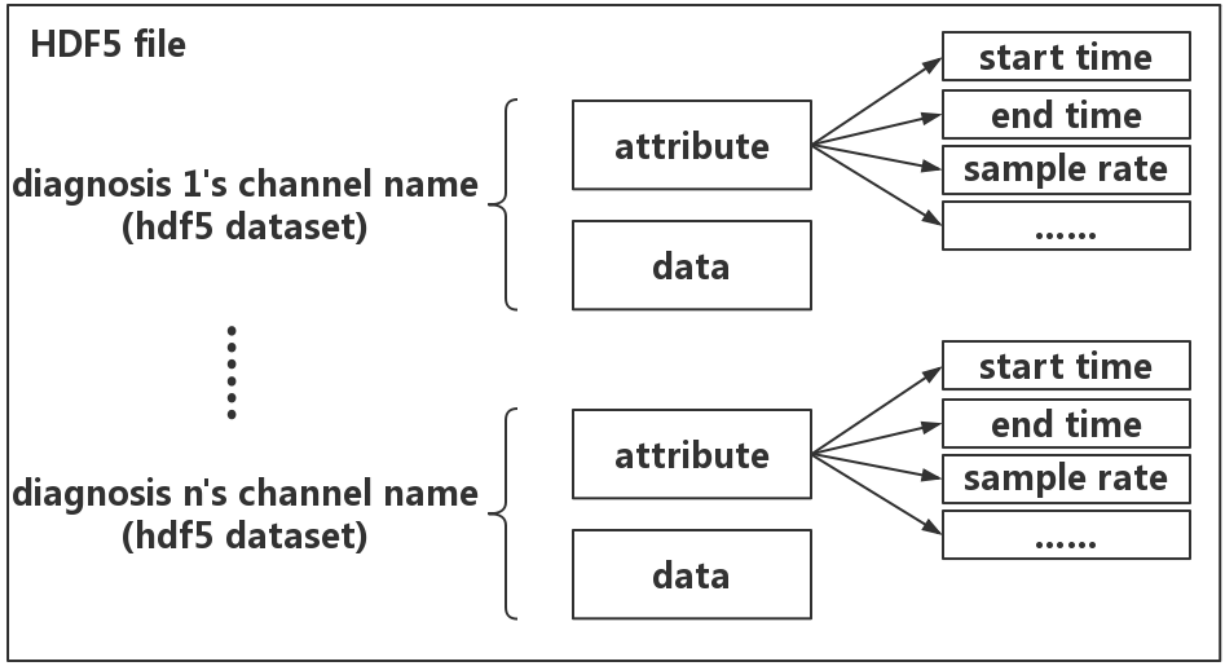
The structure of the disruption database

- **It works with a MongoDB and a file system**
- **It has 4 main components: Data Importer(DI), Label Generator(LG), Disruption Database Service(DDBS) and Data Reader(DR).**



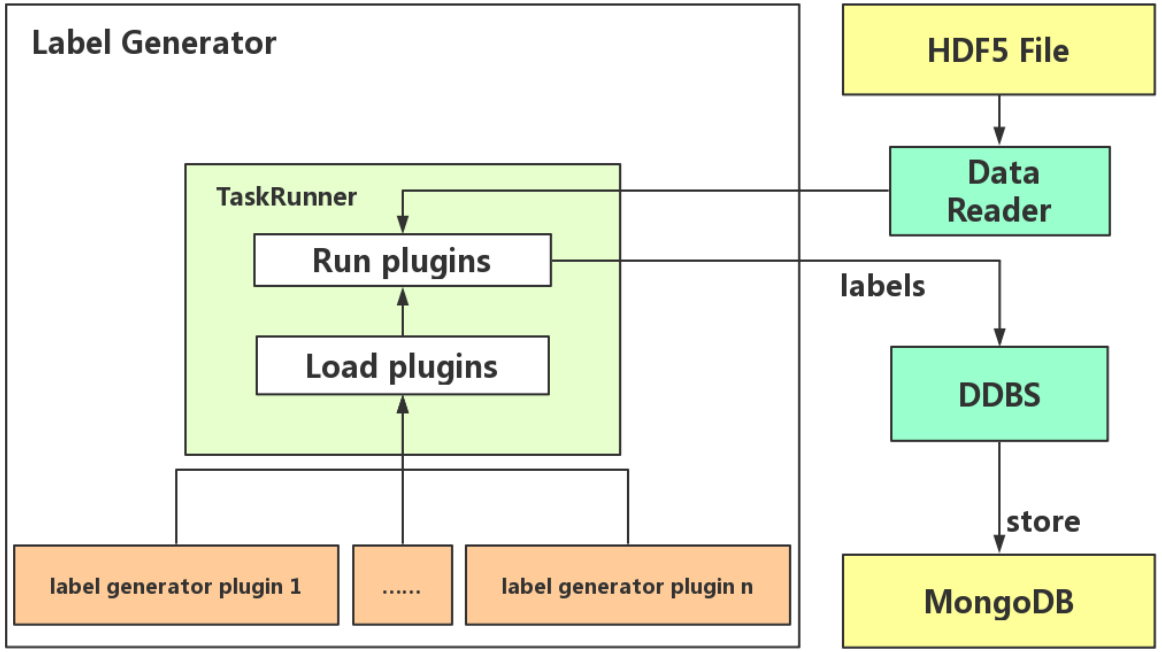
The main components of DDB

- **Data Importer**
  - DI provide API to write the data into unified format HDF5 file for storing data from different devices.



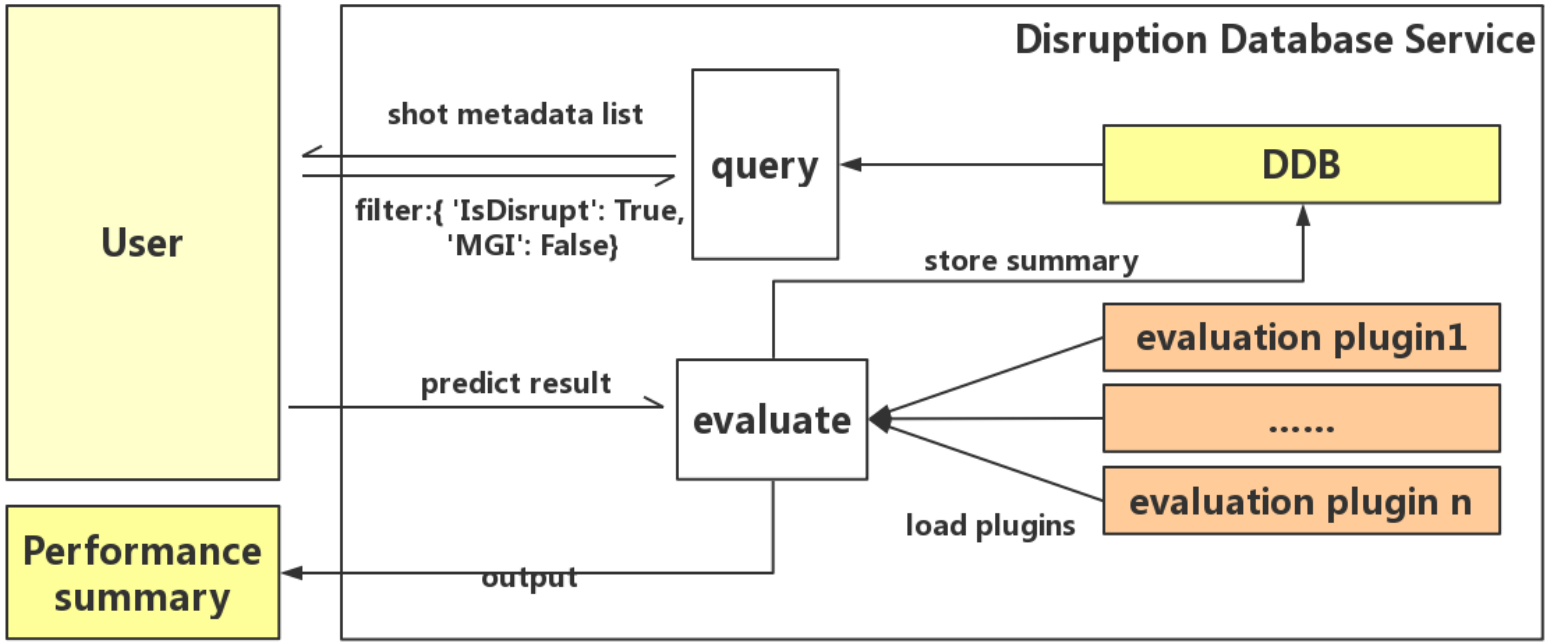
The main components of DDB

- **Label Generator**
  - Load different label generator plugins and generate a set of labels of the diagnostic data
  - We have already developed a few label generator plugins and worked out dozens of labels of the data from J-TEXT



The main components of DDB

- **Disruption Database Service**
  - ➢ It provides query and performance evaluation functions
  - ➢ You can get information of each shot and the summary of your predictor's performance by the API



Future work

- **A web UI for human analyzing and labeling the shots**
- **Open sources of the disruption database on github.**
- **Host a MongoDB server for various tokamak**