



Contribution ID: 599

Type: **Poster presentation**

Real time data access log analysis system of EAST tokamak based on spark

Thursday, 14 June 2018 15:50 (15 minutes)

The experiment data generated by the EAST device is getting larger and larger, and it is necessary to monitor the MDSplus data storage server on EAST. In order to facilitate the management of users on the MDSplus server, a real-time monitoring log analysis system is needed. The data processing framework adopted by this log analysis system is the Spark Streaming framework in Spark ecosphere, whose real-time streaming data is derived from MDSplus logs. The framework also makes use of key technologies such as log monitoring, aggregation and distribution with framework likes Flume and Kafka, which makes it possible for MDSplus mass log data processing power. The system can process tens of millions of unprocessed MDSplus log information at a second level, then model the log information and display it on the web. This report introduces the design and implementation of the overall architecture of real time data access log analysis system based on spark. Experimental results show that the system is proved to be with steady and reliable performance and has an important application value to the management of fusion experiment data.

Minioral

No

Description

Speaker

Feng Wang

Institute

IPP Hefei

Country

China

Primary authors: WANG, Feng; Mr ZHANG, Qihao (ASIPP); Ms WANG, Yueting (ASIPP); Ms CHEN, Ying (ASIPP); Dr YANG, Fei (Department of Computer Science, Anhui Medical University)

Presenter: WANG, Feng

Session Classification: Poster 2

Track Classification: Data Acquisition