

## INTISARI

### **IMPLEMENTASI *BIG DATA* PADA DATA TRANSAKSI TIKET ELEKTRONIK *BUS RAPID TRANSIT* (BRT)**

Salah satu bentuk angkutan massal yang ada di Indonesia adalah *Bus Rapid Transit* (BRT). Penerapan tiket elektronik pada BRT telah diterapkan sejak tahun 2013 hingga saat ini. Data yang dihasilkan dari transaksi tiket elektronik tersebut dapat dikelola untuk memperbaiki rute dan meningkatkan pelayanan sesuai dengan kebutuhan dari pengguna BRT. Namun, terdapat masalah yaitu data yang dihasilkan dari transaksi tiket elektronik tersebut sangat besar mencapai ratusan *gigabyte*. Oleh karena itu diperlukan teknologi yang mampu melakukan proses pengolahan, penyimpanan dan analisis data dalam beragam bentuk/format, berjumlah besar dan penambahan data yang sangat cepat yaitu *big data*. Dalam penelitian ini akan dibangun infrastruktur *big data* yang terdiri dari tiga teknologi utama yaitu Apache Kafka, Apache Spark dan Apache Cassandra. Integrasi dari ketiga teknologi tersebut akan membentuk *workflow* yang mampu mengelola data transaksi tiket elektronik BRT menjadi informasi mengenai asal-tujuan penumpang. API dari Apache Spark yaitu *Spark Streaming* sebagai data *streaming process* akan dikombinasikan dengan *java regular expression* untuk ekstraksi log data transaksi BRT. Data ekstraksi akan disimpan pada Apache Cassandra, diolah oleh Apache Spark menggunakan metode *trip chaining* untuk mendapatkan informasi mengenai asal-tujuan penumpang dalam bentuk matriks asal-tujuan (AT).

**Kata Kunci:** *Big data*, BRT, spark, matriks, asal-tujuan.

**ABSTRACT**

**IMPLEMENTASI *BIG DATA* PADA DATA TRANSAKSI TIKET ELEKTRONIK  
*BUS RAPID TRANSIT (BRT)***

*One of mass transit in Indonesia is Bus Rapid Transit (BRT). The application of electronic ticketing at BRT has been implemented since 2013 until now. Data generated from electronic ticket transactions can be managed to improve routes and increasing services in accordance with the needs of users BRT. However, there is a problem that the data generated from the electronic ticket transaction is very large reaches hundreds of gigabytes. Therefore it takes technology that can do the processing, storage and analysis of data in various forms / formats, large amounts and data increase is very fast i.e big data. In this research will be built big data infrastructure consisting of three main technologies namely Apache Kafka, Apache Spark and Apache Cassandra. The integration of these three technologies will create a workflow capable of managing BRT electronic ticket transaction data into information on passenger's origin-destination. API from Apache Spark that is Spark Streaming as data streaming process will be combined with java regular expression to log extraction of BRT data transaction in near-realtime. The extraction data will be stored in Apache Cassandra, processed by Apache Spark to get information passenger's origin-destination in the form of OD Matrix.*

**Kata Kunci:** *Big data, BRT, matrix, origin-destination*