

ABSTRAK

Universitas Pendidikan Ganesha saat ini menjadi anggota Konsorsium Automatic Identification System (AIS) ITS. Perangkat AIS yang dapat dipasang oleh anggota konsorsium mampu merekam transmisi dan menyimpan hasil transmisi perangkat AIS yang telah terpasang pada kapal laut. Saat ini data transmisi disimpan dalam satu perangkat server. Transmisi data yang dihasilkan dalam bentuk format data .csv (*comma separated values*). Infrastruktur penyimpanan data yang digunakan saat ini berorientasi pada penyimpanan menggunakan dedicated server. Ketika ukuran data tumbuh semakin besar maka diperlukan penambahan storage pada server tersebut. Data hasil rekaman AIS berpotensi pada bertumbuhnya data dalam jumlah besar (*Big Data*), jika dilihat dari infrastruktur penyimpanan data saat ini maka terdapat potensi semakin besar *workload* yang akan dialami pada *dedicated server* jika tidak disesuaikan dengan lingkungan kerja Big Data. Berdasarkan gambaran keadaan infrastruktur saat ini, maka tujuan umum pada penelitian ini melakukan eksplorasi rancangan pengembangan arsitektur lingkungan kerja Big Data menggunakan Hadoop Framework. Volume data yang mampu dikoleksi dari transmisi AIS berjumlah 9,613,762 *records*. Selain hasil data transmisi, diberikan volume data sintetik sebanyak 1.000.000 records untuk menambah volume data. Total volume data yang keseluruhan dalam penelitian ini adalah 10.613.762 records. Volume data yang relatif besar memberikan tantangan waktu akses yang relatif tinggi yang tersimpan pada suatu single server. Benchmark yang diperoleh dengan perancangan arsitektur *Big Data* menggunakan Hadoop memperoleh *improvement* kecepatan akses *query* sebesar **94,82%** dibandingkan dengan penggunaan dedicated server. Rata-rata performa *query* yang dicapai adalah **3,6392 detik**. Hal ini menunjukkan bahwa pengembangan arsitektur big data menggunakan Hadoop menjadi solusi yang menjanjikan untuk menyimpan dan memproses *big data*. Hasil benchmark dapat dijadikan referensi untuk memberikan gambaran mengenai kemampuan *Hadoop environment* untuk menangani studi kasus *Big Data* AIS.

Kata Kunci: AIS, Big Data, Hadoop, Distributed Computing, *Enterprise Architecture Planning*.

ABSTRACT

Ganesha University of Education is currently a member of the ITS Automatic Identification System (AIS) Consortium. AIS devices that can be installed by consortium members are capable of recording transmissions and storing the transmission results of AIS devices installed on ships. Currently, transmission data is stored on one server device in the form of data format .csv (comma-separated values). The current data storage infrastructure is oriented towards dedicated server storage. However, as the size of the data grows larger, additional storage is needed on the server. The data recorded by AIS has the potential to grow into large amounts of data (Big Data). When considering the current data storage infrastructure, there is potential for a greater workload to be experienced on a dedicated server if it is not adapted to the Big Data work environment.

Based on the description of the current state of the infrastructure, the general objective of this study is to explore the architectural development design of the Big Data work environment using the Hadoop Framework. The volume of data that can be collected from AIS transmissions amounts to 9,613,762 records. In addition to the data transmission results, a synthetic data volume of 1,000,000 records is given to increase the data volume. Thus, the total volume of data in this study is 10,613,762 records. The relatively large volume of data poses a significant access time challenge when stored on a single server.

Benchmarks obtained by designing the Big Data architecture using Hadoop show an improvement in query access speed of 94.82% compared to using a dedicated server. The average query performance achieved is 3.6392 seconds. These results demonstrate that the development of a Big Data architecture using Hadoop is a promising solution for storing and processing big data. The benchmark results can be used as a reference to provide an overview of the Hadoop environment's ability to handle AIS Big Data case studies.

Keywords: AIS, Big Data, Hadoop, Distributed Computing, *Enterprise Architecture Planning*.