Drive BELT: A Strategic Decision-Making Tool for Optimizing Naval and Rail Freight Logistics

Nicola Bassi[1] and Enea Vincenzo Napolitano[2]
in collaboration with Almaviva

[1] DigITAlog
[2] DIETI, University of Naples Federico II

Abstract. The Drive BELT project, developed by digITAlog, has the objective of developing an IT tool for the structured collection and advanced analysis of data, with the aim of optimising decision-making processes in the field of transport. This comprehensive solution includes a Monitoring System that provides regular, structured reports and a Strategic Dashboard for in-depth scenario analysis using Big Data. The Drive BELT system facilitates the analysis of vehicle and freight flows in ports, interports, and industrial districts by consolidating data from various public and private entities within the Digital Logistics Chain. This data is then analyzed to provide insights and recommendations that support strategic decision-making for institutional stakeholders. The project is divided into two phases. The first phase involves the creation of a data collection system from diverse sources. The second phase entails the development of monitoring and strategic dashboard tools to assist the Ministry of Infrastructure and Transport.

Keywords: Decision Support System · Logistics Optimization · Big Data Analytics.

1 Introduction

In the age of digitalization, advanced monitoring and analysis of key performance indicators (KPIs) has become critical to optimize decision-making in various sectors [1], including rail and shipping. Efficient management of transportation logistics is critical to the functioning of modern economies, especially in less developed regions where infrastructure and resources may be limited. In this context, an integrated IT system for structured data collection and advanced analytics can add significant value. Improved visibility of vehicle and freight flows in ports, interports and industrial areas can lead to greater operational efficiency, reduced costs and improved decision support for the institutional authorities involved.

The Drive BELT project, developed by digITAlog, was created to address these challenges. The project aims to provide a comprehensive monitoring system and strategic dashboard to support decision making in the rail and maritime sectors. Drive BELT aims to consolidate data from different public and private entities within the digital logistics chain, facilitating the analysis of vehicle and freight flows in ports, interports and industrial districts.

The project consists of two main phases: the first phase involves the implementation of a data collection system, while the second phase focuses on the development of monitoring tools and the strategic dashboard. These tools are designed to assist the Ministry of Infrastructure and Transport in decision making by providing a clear and detailed view of vehicle and freight flows.

This article provides an overview of the development and applications of the Drive BELT project. Section 2 presents related work, illustrating previous studies and projects related to ours. Section 3 describes the system architecture, focusing on the components and infrastructure used. In section 4, the flow of data is described in terms of both data sources and data preparation and processing. Additionally, methodologies for data visualization and utilization are presented. Finally, the Conclusion summarizes the results obtained and possible future developments of the project.

2 Related Work

The field of transportation data analytics has seen significant advances, with numerous platforms and systems developed to address various challenges in maritime and rail transportation. These studies provide a fundamental understanding and inspiration for the development of Drive BELT.

In the field of rail transportation, Gheorghe et al. [7] present an enhanced risk assessment solution for the transportation of dangerous goods by rail. Their software platform integrates the causes and consequences of containment loss with the system's infrastructure and environment, illustrating the critical role of integrated platforms in rail risk management. Similarly, Ma et al. [9] propose a region-wide web-based transportation decision system, DRIVE Net, that uses digital roadway maps as its foundation and integrates multiple data sources. DRIVE Net provides a practical method for data retrieval, integration, and visualization to improve transportation decision-making processes. This platform serves as an inspiration for BELT's web-based analysis and visualization capabilities, especially in integrating and analyzing diverse data sources.

Gu et al. [8] introduce a new energy-efficient train operation model based on real-time traffic information through a nonlinear programming method. Their work leads to an energy-efficient driving strategy with real-time interstation running time monitored by an automatic train control system, highlighting the importance of real-time data processing for energy efficiency in rail transportation.

In the maritime sector, Rodrigues et al. [13] present an optimization approach to the problem of routing and scheduling ships for oil cabotage. This work highlights the complexity of ship routing and the need for efficient scheduling solutions, emphasizing the importance of optimization techniques in maritime logistics. Xiao et al. [16] propose a knowledge-based methodology for maritime traffic forecasting based on vessel waterway patterns and motion behavior. They implement a web-based prototype platform for Singapore waters, demonstrating the feasibility of real-world maritime operations systems. This study demonstrates the potential of integrating knowledge-based methods and web technologies for maritime traffic management. Tang et al. [14] discuss the design and research of an integrated information platform for smart ships, emphasizing the importance of data collection and application. Their work highlights the need for advanced solutions to integrate data from different subsystems for monitoring and management, which is a core principle in the development of BELT. Chen et al. [2] present a simulation platform for multi-yard rail intermodal terminals, which is capable of providing quantitative assessments of rail intermodal terminals. The platform's flexibility in terminal layout design and its ability to accurately reproduce real system activities are particularly noteworthy for the development of simulation and analysis capabilities in BELT. Ding et al. [3] discuss top-level design considerations for multimodal transportation systems, focusing on the coordination of transportation organization under mixed time window constraints. Their study aims to improve transportation efficiency and establish an electronic platform for information interconnection between multimodal transportation stations.

In the context of urban rail systems, Mulerikkal et al. [10] describe an intelligent data analytics platform for metro rail transportation systems using Hadoop Distributed File System (HDFS) and Spark framework, along with tools such as Apache Hive and Spark ML-lib for real-time and legacy data processing. Their project for Kochi Metro Rail Limited demonstrates the effectiveness of combining big data technologies with real-time analytics for urban rail systems.

Eisenhauer et al. [5] analyzes the strategy for maintaining ship readiness through system maintenance activities. This study provides insight into the maintenance strategies essential to ensuring the operational efficiency of maritime systems, and highlights the need for comprehensive data analytics platforms to support such activities.

Collectively, these studies underscore the advances in maritime and rail transportation analytics, the integration of real-time data processing, and the use of web-based platforms for data visualization and decision support. Drive BELT aims to build on this foundation by providing a comprehensive solution tailored to the specific needs of maritime and rail data analytics.

3 Drive BELT Architecture

3.1 Components

The DriveBelt system has been designed with the intention of providing a robust and scalable platform for the management and analysis of data. The primary components of the system are delineated below, each with specific functionalities and configurations to ensure optimal performance.

- **WSO2**:

  - *Composition*: Includes API Manager, Identity Server, Enterprise Integrator, and RDBMS. These components work together to manage APIs, user authentication, and service integration.
  - *Deployment*: Deployment is possible on a single node with a minimum configuration of 4 cores, 8GB RAM, and 1TB HDD, allowing efficient resource management.
  - *Capacity*: Capable of handling up to 500 API calls per second and 500 authentication requests per day, ensuring a high level of performance and reliability.

- **Apache NiFi**:
  - *Architecture*: Scalable linearly with the addition of nodes, enabling easy expansion of processing capabilities.
  - *Minimum Configuration*: The minimum configuration includes 3 nodes, each with 8 cores, 8GB RAM, and 1TB HDD, supporting a throughput of 50MB/s or 1000 events per second.

- **Cloudera Enterprise Data Hub (EDH)**:
  - *Composition*: Comprises master nodes, worker nodes, and Edge/Utility nodes dedicated to tools like JupyterHub.
  - *Master Node Configuration*: Each master node is configured with 16 cores, 64GB RAM, and 7TB HDD to support high-intensity data operations.
  - *Worker Node Configuration*: Worker nodes are configured with 16 cores, 128GB RAM, and 12TB HDD to balance workloads between CPU, I/O, and memory.
  - *Storage*: Storage is distributed across worker nodes with a total capacity of at least 12TB, ensuring sufficient space for data storage.

- **Apache Kafka**:
  - *Minimum Configuration*: Includes 12 cores, 64GB RAM, and 1TB HDD with a throughput of 125MB/s, ideal for managing real-time data streams.

- **JupyterHub**:
  - *Configuration*: Configured with 8 cores, 64GB RAM, and 1TB HDD to ensure high performance during the execution of interactive notebooks.

- **PostgreSQL**:
  - *Usage*: Used for GEODB and WSO2 service schemas, providing a robust and scalable relational database.
  - *Configuration*: Configured with 4 cores, 16GB RAM, and 1TB HDD to support up to 100 concurrent connections, ensuring efficient transaction management.

- **Tableau**:

- *Configuration*: Each Tableau node is configured with 8 cores, 64GB RAM, and 3TB HDD to support intensive data visualization workloads.

Two main scenarios of computational capacity have been defined: the initial system launch and the projected growth within the first 5 years.

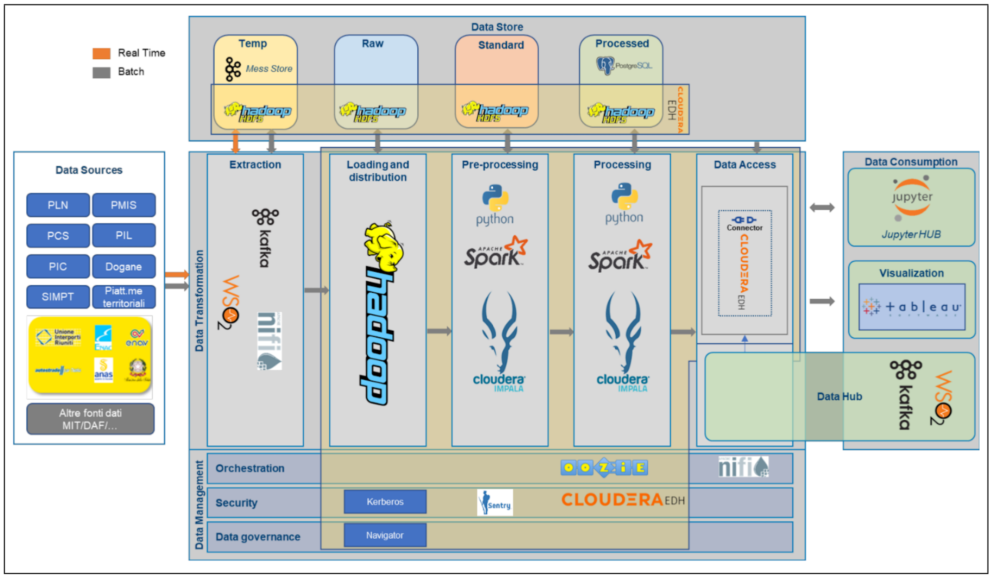

Fig. 1. Architecture of Drive BELT

3.2 Infrastructure

The technology and infrastructure underlying Drive BELT (3.1) have been chosen to ensure high performance, reliability, and scalability.

- **Virtual Machines (VMs)**:
  - *Usage*: All system components are implemented on virtual machines, offering flexibility and ease of management.
  - *Storage Features*: The VMs use spinning drives with a throughput of no less than 300 MB/s per node, ensuring fast data read and write operations.
- **Storage Components**:
  - *Configuration*: The storage is configured with spinning drives that ensure an overall throughput per node of no less than 300 MB/s.
  - *Distribution*: The storage is distributed across worker nodes to ensure efficient data management and optimal load balancing.

– **Network**:
  - *Connections*: The recommended network connections are at least 1 Gbps, ensuring fast and reliable communication between the various system components.

The components are configured to ensure maximum operational efficiency and flexibility, making Drive BELT an ideal solution for complex and dynamic data environments.

4 Dataflow

4.1 Data Sources

The Drive BELT platform integrates a wide range of data, primarily from maritime and rail transportation, structured to support comprehensive analytical scenarios. Data sources include detailed statistics, operational metrics and other relevant information from established databases and institutions.

In the maritime domain, data on cargo volumes and movements are sourced from national and international databases, including contributions from port authorities and shipping companies. This includes information on shipping routes and vessel movements, which is critical for analyzing trends in maritime logistics and assessing port efficiency.

For rail transport, the main data source is RFI (Rete Ferroviaria Italiana), which provides data on movements and distances traveled between Italian stations. The reference data come from ISTAT and Eurostat, which provide comprehensive import and export statistics. These datasets provide information on the volume of goods transported by rail, divided by type and origin/destination in Europe. The data could be further enriched by national transport data collected by the Connections Industrial Plan (CIP), which provides granular information on freight movement in Italy, focusing on specific corridors and regional breakdowns. Operational data, including metrics such as train-kilometers, travel times, and infrastructure utilization, are provided by the Ministry of Infrastructure and Transport (MIT) and the Rete Autostrade Mediterranee (RAM); these were not used in the implementation phase but could be added in the development phase.

4.2 Data Preparation and Processing

The collected data undergoes several stages of processing to ensure that it is suitable for in-depth analysis. The first step is data acquisition, where data is extracted from multiple sources using automated pipelines. This ensures that the data is timely and accurate. For example, ISTAT and Eurostat data could be retrieved on an annual basis, while CIP data could be updated monthly to reflect the latest trends. After data ingestion, the data is subject to a cleaning process. Records with significant missing information are excluded to maintain the overall quality and reliability of the dataset. After data cleansing, transformation processes are applied to create new metrics and dimensions that enhance analytical capabilities. Data aggregation is also performed to synthesize information at different levels of granularity, such as monthly summaries for rail freight volumes and annual statistics for import/export data.

The cleaned and transformed data sets are then integrated into a unified schema using tools such as Apache Spark. Finally, the processed data is stored in a structured format within the Drive BELT data warehouse [6].

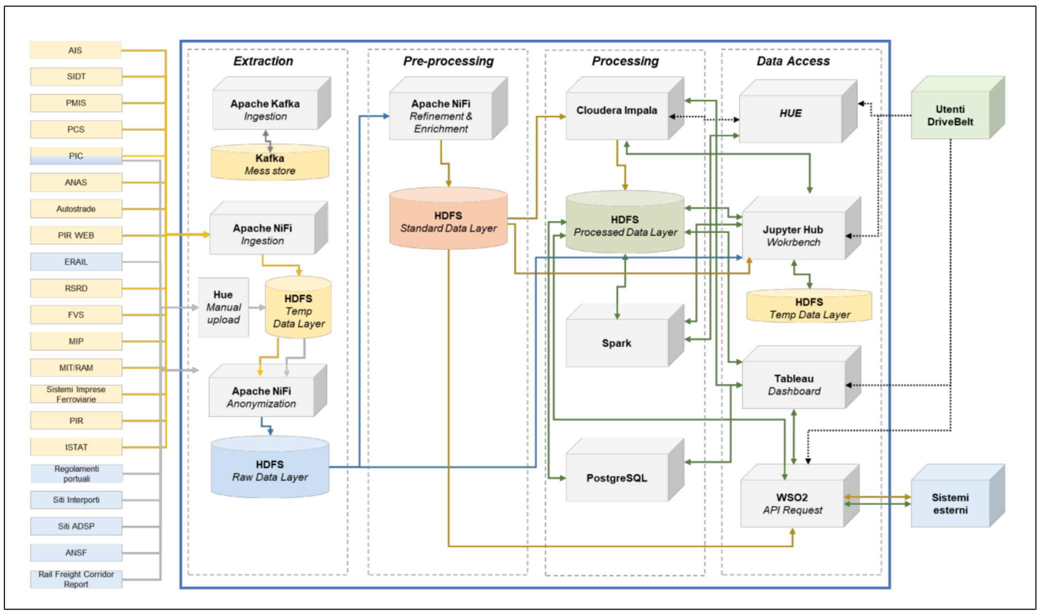

Fig. 2. Dataflow of Drive BELT

4.3 Methodologies

The analytical methods used in Drive BELT include a variety of techniques, including business intelligence (BI) [12], machine learning (ML) [15], and time series analysis [4]. Business intelligence techniques are used to create interactive dashboards and reports that provide insight into traffic trends, operational performance, and strategic decision-making. Tableau is integrated into the platform to facilitate data visualization, allowing stakeholders to easily interpret and act on findings.

Machine learning models are developed for predictive analytics, such as forecasting demand for rail and maritime transportation. These models use historical data and real-time inputs to make accurate predictions. Examples include regression models for predicting ship dwell time in port.

Time series analysis is used to monitor and forecast trends over time, particularly in the number of containers handled in a month. This approach can be used to anticipate trends and identify seasonal variations to support long-term strategic planning.

Scenarios The processed data and analysis methods are applied to different scenarios. These scenarios include:

1. **Rail Freight Demand Analysis:** This scenario involves the analysis of rail freight demand with a focus on volumes shipped, regional distribution, and trend analysis over time. The insights gained from this analysis support strategic planning and operational optimization.

2. **Freight Demand on Railways in Southern Italy:** This scenario examines freight demand on rail routes in Southern Italy. The analysis includes the identification of key trends and regional disparities, which are crucial for regional development and infrastructure planning.

3. **Rail Freight Supply Side:** This scenario analyzes the supply side of rail freight transportation, focusing on the availability of transportation services, infrastructure capacity, and service efficiency. The results help to understand the alignment between supply and demand in the rail freight sector.

4. **Rail Transportation of Hazardous Materials:** This scenario focuses on the transportation of dangerous goods by rail, integrating risk assessment models to ensure safety and compliance. The analysis provides a framework for identifying and mitigating risks associated with the transportation of dangerous goods.

5. **Maritime Freight Transportation:** This scenario examines maritime freight transportation by evaluating the performance of shipping routes, port operations, and the impact of external factors on maritime logistics. The analysis helps identify bottlenecks and optimize maritime operations.

By applying these methodologies and leveraging comprehensive data sets, Drive BELT provides a robust platform (Fig. 2) for transportation data analytics. The platform supports decision-making and strategic planning in maritime and rail transportation.

5 Conclusion

The Drive BELT system has proven to be a robust and scalable platform for managing and analyzing vehicle and freight flow data. With its modular architecture and high-performance components, the system is expected to help the Ministry of Infrastructure and Transport make informed decisions, improve operational efficiency, and manage resources. The implementation of advanced technologies such as WSO2 for API management and authentication, Apache NiFi for data processing, and Cloudera EDH for big data analytics enabled the creation of an integrated and flexible solution.

In the future, the Drive BELT project plans to expand its capabilities by integrating data from all transportation providers. This expansion will include not only commercial vehicles, but also all other modes of transportation, with an initial focus on trucks. The integration of provider data will provide an even more comprehensive and detailed view of traffic flows, improving the system's ability to support strategic and operational decisions.

In addition, the future of the project includes the implementation of new data analysis and visualization capabilities. These features will further enhance the system's ability to transform raw data into useful information that is easy for end users to interpret. The use of advanced tools such as Tableau for visualization and JupyterHub for interactive analysis will allow users to dynamically explore data and gain insights in real time.

Expected growth within the first five years of the project includes the expansion of compute capacity to handle an increasing volume of data and requests. This will be accompanied by continuous upgrades to the technology infrastructure to maintain high standards of performance and reliability.

In summary, Drive BELT is confirmed as an innovative and advanced solution for data management and analysis in the transportation context. The planned expansion and integration of new technologies and data will keep the system at the cutting edge, supporting the Ministry of Infrastructure and other agencies in the continuous improvement of operations and strategic planning. Drive BELT is thus a key element in the development of smart cities [11] and intelligent transportation systems.

Acknowledgement

This work has been supported by the project "REALIZZAZIONE, SPERIMENTAZIONE, FORMAZIONE, COMUNICAZIONE E DIFFUSIONE DEL SISTEMA DRIVE BELT A SUPPORTO DELLA CATENA LOGISTICA DIGITALE" funded by the Programma Operativo Nazionale (PON) Infrastrutture e Reti 2014-2020 (Asse prioritario II Linea d'Azione II.2.2) .

References

1. Anniciello, A., Fioretto, S., Masciari, E., Napolitano, E.V.: Digital twins for traffic congestion in smart cities: A novel solution using data mining techniques (2023)
2. Chen, X., He, S., Li, T., Li, Y.: A simulation platform for combined rail/road transport in multiyards intermodal terminals. Journal of Advanced Transportation **2018**, 1–19 (2018)
3. Ding, L.: Multimodal transport information sharing platform with mixed time window constraints based on big data. Journal of Cloud Computing **9**, 1–11 (2020)
4. Durango-Cohen, P.L.: A time series analysis framework for transportation infrastructure management. Transportation Research Part B: Methodological **41**(5), 493–505 (2007)
5. Eisenhauer, I.F., Walrath, B.D., Bebarta, V.S., Tadlock, M.D., Baker, J.B., Schauer, S.G.: Navy en-route care in future distributed maritime operations: a review of clinician capabilities and roles of care. Prehospital Emergency Care **27**(4), 465–472 (2023)
6. Fazzinga, B., Flesca, S., Masciari, E., Furfaro, F.: Efficient and effective rfid data warehousing. In: Proceedings of the 2009 International Database Engineering & Applications Symposium. pp. 251–258 (2009)
7. Gheorghe, A.V., Birchmeier, J., Vamanu, D., Papazoglou, I., Kröger, W.: Comprehensive risk assessment for rail transportation of dangerous goods: a validated platform for decision support. Reliability Engineering & System Safety **88**(3), 247–272 (2005)
8. Gu, Q., Tang, T., Cao, F., Song, Y.d.: Energy-efficient train operation in urban rail transit using real-time traffic information. IEEE Transactions on Intelligent Transportation Systems **15**(3), 1216–1233 (2014)
9. Ma, X., Wu, Y.J., Wang, Y.: Drive net: E-science transportation platform for data sharing, visualization, modeling, and analysis. Transportation research record **2215**(1), 37–49 (2011)
10. Mulerikkal, J., Thandassery, S., Rejathalal, V., Ayyappan, B., et al.: Jp-dap: an intelligent data analytics platform for metro rail transport systems. IEEE Transactions on Intelligent Transportation Systems **23**(7), 9146–9156 (2021)
11. Napolitano, E.V.: Intelligent technologies for urban progress: exploring the role of ai and advanced telecommunications in smart city evolution. In: European Conference on Advances in Databases and Information Systems. pp. 676–683. Springer (2023)
12. Reddy, C.S., Sangam, R.S., Srinivasa Rao, B.: A survey on business intelligence tools for marketing, financial, and transportation services. In: Smart Intelligent Computing and Applications: Proceedings of the Second International Conference on SCI 2018, Volume 2. pp. 495–504. Springer (2019)
13. Rodrigues, V.P., Morabito, R., Yamashita, D., Silva, B.J.d., Ribas, P.C.: Ship routing with pickup and delivery for a maritime oil transportation system: Mip model and heuristics. Systems **4**(3), 31 (2016)
14. Tang, Y.l., Shao, N.n.: Design and research of integrated information platform for smart ship. In: 2017 4th International Conference on Transportation Information and Safety (ICTIS). pp. 37–41. IEEE (2017)
15. Tizghadam, A., Khazaei, H., Moghaddam, M.H., Hassan, Y., et al.: Machine learning in transportation (2019)
16. Xiao, Z., Ponnambalam, L., Fu, X., Zhang, W.: Maritime traffic probabilistic forecasting based on vessels' waterway patterns and motion behaviors. IEEE Transactions on Intelligent Transportation Systems **18**(11), 3122–3134 (2017)