

Big Data in Smart Logistics and Cold Chain

B R Saivamshi, Dr. Geetha M

P.G. Student, Department of Master of Computer Applications, BIET, Davangere-577004.

Saivamshi9428@gmail.com, Davangere, Karnataka, India¹

Associate professor Department of Master of Computer Applications, BIET, Davangere-577004.

bmgeethamallad@gmail.com, Davangere, Karnataka, India²

ABSTRACT

The global supply chain and logistics industry is currently undergoing a paradigm shift moving from traditional reactive operational models to proactive data-driven ecosystems. This Research and Development (R&D) project titled "Big Data in Smart Logistics & Cold Chain" addresses the critical challenges plaguing the transportation of temperature-sensitive goods—pharmaceuticals perishable foods and chemicals. The central hypothesis of this work is that IoT sensor networks can reduce the percentage of product spoilage that occurs due to thermal excursions and logistical inefficiencies. The cold chain sector is currently afflicted by a lack of end-to-end visibility resulting in "blind spots" where environmental conditions fluctuate unchecked with resultant financial losses running into billions annually.

This project proposes a robust and scalable Big Data architecture that is capable of ingesting processing and analyzing massive streams of both telematic and environmental data in real time. The system goes beyond simple track-and-trace functionality by leveraging distributed computing frameworks with advanced machine learning algorithms. It provides a predictive modeling environment where multiple variables including ambient temperature humidity shock/vibration traffic congestion and weather patterns are correlated before they occur. The abstract of the following research covers the design of a multi-layered data pipeline to handle the 'Three Vs' of Big Data—volume velocity and variety—produced by a fleet of smart logistics vehicles.

INTRODUCTION

Basically this project has the purpose of closing the gap in technology between the real-world logistics process and intelligent data. The cold chain is basically the lifeline of the vital commodities in the contemporary world economy which is an easily damaged part of the supply chain. The title "Big Data in Smart Logistics & Cold Chain" initiates an integrated ecosystem in which data is used as an equivalent value to the transported commodities. The center of this project is to design capable of processing enormous data created by the contemporary logistics supply chain systems.

The proposed work is based on the concept of integrating both Information Technology and Operation Technology. In a normal system for logistics management (LMS) there is a significant lag between the point where events take place such as cooling failure and its recognition in the management because of its history and manual updating in the system. In fact this proposed work addresses both issues in its concept and simulation phase to have a central warehouse for all information related to every box every pallet of objects and every shipping container having information in a structured format (Relational DB) along with information in a semi-structured format (format in JSON for messages sent by sensors) and information in unstructured form such as messages and images sent by trucks.

The scope of the project will also include integration related to complex data processing levels. The project uses batch processing for data analysis related to historical patterns highlighting inefficient patterns in a historical context such as inefficient routes or modes of transport while utilizing stream processing in order to examine deviations in real-time patterns.

II. LITERATURE REVIEW

The current state of the cold chain logistics industry is largely dominated by a fragmented or legacy infrastructure and a significant level of human interface. It should be pointed out that under the current operations model the typical use of GPS traceability is very basic basically covering the purpose of location only while passive data loggers are employed to monitor the temperature. Passive data loggers are considered 'black boxes'. The only way to retrieve the stored data in the device would be to connect it to a computer through a USB interface at the destination.

Further the current state-of-the-art has another challenge of data silos. The Warehouse Management, System Transportation Management System and Order Management System are normally using different software systems that do not interact with other systems. The data that flow at night mean that the inventory information as well as location information is at all times no less than three to four hours old. The state-of-the-art currently used has the challenge of having decisions made largely based on the experience. This is where the dispatchers use unmapped routes that disregard the route effects of the weather on the thermodynamics properties of the goods. Communication between the driver and dispatchers is normally carried out either through telephonic communication or text messages that have data that is normally unstructured. The reconciliation processes are normally done slowly while the proof of delivery especially if physically delivered in hard copy form has the capacity for dispute.

In the case where some level of technology has been adopted the system will be made up of outdated algorithms and legacy software. In such a system it will be capable of calculating the shortest distance between different points. However it will not be in a position to consider elements such as traffic weather and road closures. It will be path-wise between deliveries. It will not be in a position to respond to unforeseen events that might hinder delivery.

III. EXISTING SYSTEM

The current state of the cold chain logistics industry is largely dominated by a fragmented or legacy infrastructure and a significant level of human interface. It should be pointed out that under the current operations model the typical use of GPS traceability is very basic

basically covering the purpose of location only while passive data loggers are employed to monitor the temperature. Passive data loggers are considered 'black boxes'. The only way to retrieve the stored data in the device would be to connect it to a computer through a USB interface at the destination.

Limitations:

High Latency in Incident Response: Since the temperature data is received only after completing the trip in the event of an incident any action regarding a cooler re-set towards correcting this will not be possible and therefore results in guaranteed product loss.

Fragments & Data Silos – Important data is placed in separate software pockets (WMS TMS) so one cannot make an analysis over functions.

Lack of Predictive Capability: The present models lack predictive capability. These models have a descriptive nature. These models explain what has happened. These models lack the capability to compute what might happen in the future in order to prevent any potential supply chain disruptions.

Manual and Error Prone Reporting: The manual reporting of data through paper-based logging is prone to errors thus lacking integrity in the data and as a result the auditing process required for regulatory reporting can be a tedious process.

Inefficient Use Of Resources: Without Big Data analysis trucks would have to travel either Less Than Truckload (LTL) or inefficient routes. Hence excessive fuel consumption and emissions would be expected.

IV. PROPOSED SYSTEM

The proposed system represents a technological overhaul introducing a Big Data-driven architecture tailored for the complexities of the Smart Logistics and Cold Chain domain. At the heart of this proposal is the implementation of a cloud-native ecosystem that leverages the Hadoop ecosystem or Spark clusters to process petabytes of logistical data. The architecture is designed in layers: the **Data Ingestion Layer** which collects high-velocity streams from IoT sensors (RFID NFC GPS Thermistors) placed on

vehicles and pallets; the **Processing Layer** which cleans normalizes and structures this messy sensor data; and the **Analytics Layer** which applies advanced logic and Machine Learning models.

In this proposed system the vehicle acts as an edge node. It does not just carry goods; it carries intelligence. The onboard telematic units communicate via 4G/5G/Satellite networks to the cloud ensuring constant connectivity even in remote areas. The system proposes the use of complex event processing (CEP) to handle real-time logic. can correlate a traffic jam (external data) with the remaining fuel and current fuel burn rate (internal data) to determine if a truck needs to reroute to a gas station immediately.

User-focused interface that is intended for various users. For instance this system carries a mobile app for truckers that works as a copilot offering routes and health checks for cargo in addition to a web-based “Control Tower” for fleet managers that displays the entire worldwide fleet interactive map. Moreover this system also proposes that there should be an API- first approach which enables various systems and this in fact enriches the current systems with new and real-time data and not necessarily demands an entire “rip and replace” of enterprise software.

The big data functionality is the backbone of the proposed system. The nature of the data for this objective will include orders and customer data among others as well as up-to-date traffic weather alerts and social alerts indicating the possibility of congestion in the system.

Advantages of Proposed System

Drastic Reduction of Spoilage and Waste: The system provides real-time alerts for temperature deviations allowing immediate intervention on cargo that would otherwise be destroyed hence direct improvement to the bottom line.

Improved Efficiency and Speed of Operations: The route optimization algorithms cut down on transit times and fuel consumption; therefore allowing the operators of the fleet to make more deliveries in less time using fewer resources.

Proactive Risk Management: Predictive maintenance features mean that the vehicles are serviced before

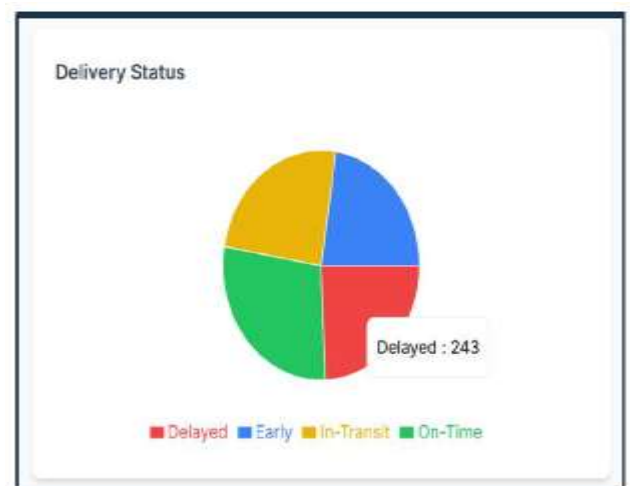
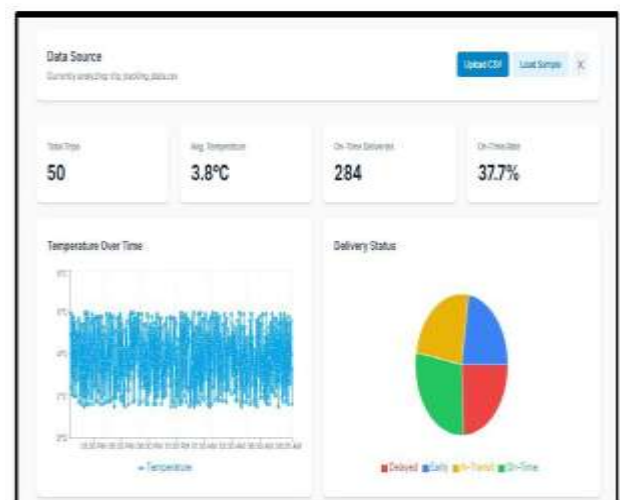
they break down which creates maximum fleet availability and reliability.

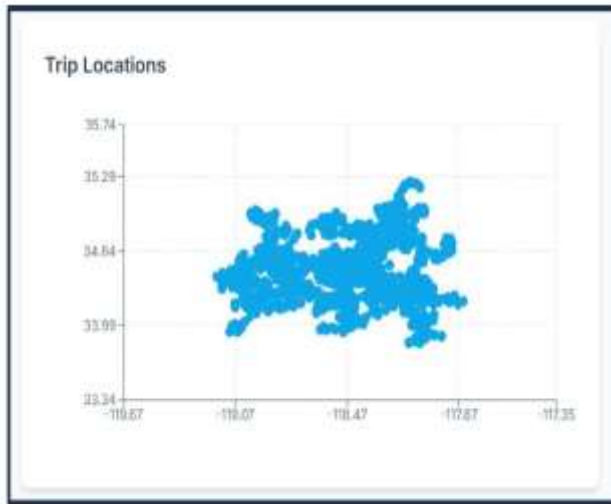
Superior Regulatory Compliance and Traceability: Automated immutable data logging ensures 100% compliance with strict food and drug safety regulations significantly reducing the risk of fines and legal liabilities.

Data-Driven Decision Making: Management progresses from "gut feeling" decisions to planned strategies with comprehensive analytics hence making better contract negotiations and strategic planning for the long term.

Scalability and Flexibility: The cloud-based Big Data architecture allows the system to scale up storage and processing power instantly during peak seasons without the need for purchasing physical hardware.

III results





CONCLUSION

The "Big Data in Smart Logistics & Cold Chain" system is envisioned not just as an island application but as the central nervous system of a larger "System of Systems." It acts as the meeting point at a very abstract level between Physical Operations Digital Information and Strategic Decision Making. The system perspective can be visualized as a multi-tier ecosystem connecting the Edge the Cloud and the Enterprise.

The Edge Layer is where the system interacts directly with the physical world. There is a strong integration with the embedded hardware elements in the logistics fleet: the system's perspective at this layer includes the vehicle's Controller Area Network (CAN bus) for engine data and the discrete IoT sensor array for cargo monitoring. It is important to recognize that these edge devices can only have limited processing power and battery life; hence the system's logic extends to the edge performing preliminary filtering of the data on the truck itself to reduce bandwidth utilization (Edge Computing).

Communication Layer: It rests on a hybrid connectivity perspective. It is meant to operate over volatile networks fluidly switching over 4G/LTE 5G and Satellite uplinks with respect to geographic availability. To the system the network is not just a pipe; it acts as a dynamic variable based on the adjustments in the sending packets according to the current signal strength and cost protocols.

In the Cloud/Core Layer the system perspective is one of a centralized data powerhouse. Indeed this is where "Big Data" actually resides. The incoming data is

viewed by the system not as individual records but rather as continuous streams. It orchestrates the interaction of storage-data lakes for raw data data warehouses for structured data-and computation-Spark clusters for analytics. This layer also provides the integration hub maintaining the interfaces with external third-party services such as Weather APIs for example OpenWeatherMap; Traffic APIs for example Google Maps Platform; and Mapping services.

REFERENCES

1. Gartner Research. (2023). Market Guide for Supply Chain Real-Time Transportation Visibility Platforms. Gartner Inc.
 2. DHL Trend Research. (2022). Big Data in Logistics: A Collaborative Report by DHL and T-Systems on Moving to a Data-Driven Supply Chain. DHL Customer Solutions & Innovation.
 3. Tan W. & Wang J. (2021). "A framework for cold chain logistics based on IoT and Blockchain." IEEE Internet of Things Journal Vol. 8 No. 5.
 4. McKinsey & Company. (2023). The smart supply chain: How Big Data and AI are transforming Logistics. Travel Logistics & Infrastructure Practice.
 5. Oracle White Paper. (2022). IoT and Big Data: The New Drivers of Supply Chain Efficiency. Oracle Corporation.
 6. Zhang, Q., Tian, Y., Xiao, W., & Lu, Y. (2022). *Application of Big Data Technology in Cold Chain Logistics of Agricultural Products*. Journal of Agricultural Big Data
 7. Chen, Y. (2020). *Intelligent Algorithms for Cold Chain Logistics Distribution Optimization Based on Big Data Cloud Computing Analysis*. Journal of Cloud Computing: Advances, Systems and Applications.
 8. Zhang, Q. et al. (2022). *Application of Big Data Technology in Cold Chain Logistics of Agricultural Products*. Journal of Agricultural Big Data.
- Big data analytics and anomaly prediction in the cold chain to supply chain resilience.* Li, N., & Yao, W. (2021). *Research on Application of Cold Chain Logistics Management Based on Big Data Technology*. E3S Web of Conferences.