THE OPEN GROUP®
INITIATE
Work Group

Irish Rail TRANSFORMATION

Team ArcVision

Rohan J S, Prajwal M, Kushaal S, Pranav D,
R.V College of Engineering

Mentors

Manuel Di Toma, Director, Bruhati
Hrishi Kuchibhotla, Director, Bruhati
John Kiernan, Chief Architect, Irish Rail
Daryl Sullivan, Irish Rail Architect
& Gavin Keating, Irish Rail Architect

THE OPEN GROUP®
INITIATE
Work Group

Table of Contents

1 Executive Summary

The following executive summary outlines the strategic vision, core challenges, proposed solutions, and anticipated business impact of the Irish Rail Transformation Project. It captures the essence of Team ArcVision's architectural approach to modernizing Iarnród Éireann's data ecosystem, aligning it with best practices in enterprise integration, real-time analytics, and regulatory compliance.

1.1 Overview of the Project

The Irish Rail Transformation Project, led by Team ArcVision, is a strategic initiative under the Enterprise Architecture competition, aimed at revolutionizing the data infrastructure of Iarnród Éireann (Irish Rail). With Irish Rail undergoing significant modernization across multiple verticals—ranging from passenger experience to operational efficiency—massive volumes of data are being generated from various sources including IoT sensors, legacy systems, mobile applications, control systems, and customer platforms.

Despite this growth, the current state of Irish Rail's enterprise architecture lacks a standardized, interoperable, and scalable framework to manage, analyse, and visualize data holistically. Fragmentation between systems leads to data silos, redundant data pipelines, inconsistent formats, and limited insights, ultimately constraining real-time responsiveness and strategic planning.

This project proposes a comprehensive Enterprise Reference Architecture (ERA) tailored to the unique operational, regulatory, and digital needs of Irish Rail. The architecture is modelled along the TOGAF framework, aligning the four key layers: Business, Application, Data, and Technology. The centrepiece of this architecture is the Unified Namespace (UNS)—a real-time messaging and integration backbone built using HiveMQ and MQTT protocols, facilitating seamless communication across distributed systems and devices.

Key objectives include:

- Establishing a single source of truth for all enterprise data.
- Designing a semantic hierarchy (based on ISA-95) for structured classification and discoverability.
- Design reference architecture for integrating real-time and historical analytics to support operational intelligence.
- Design reference architecture and timeline for implementing secure, scalable, and compliant systems for long-term transformation support.

The technologies considered are:

- Time-series databases (Influx DB) for sensor analytics.
- Structured databases (AWS RDS) for enterprise-grade applications.
- Data lakes (Amazon S3) for historical and unstructured data storage.
- A robust security architecture with compliance to GDPR, ISO 27001, NIS Directive, and EU Cybersecurity Act.
- The solution delivers an end-to-end digital transformation framework designed not only to support current needs but also to accommodate future scaling, integration with predictive systems, and adaptive control across Irish Rail's nationwide infrastructure.

1.2 Key Problem and Proposed Solution

Irish Rail is encountering significant challenges in effectively managing the growing volume and diversity of data generated across its modernization initiatives. The current landscape presents several opportunities for enhancement, including:

- Establishing a standardized enterprise data architecture to ensure consistency and coherence across initiatives.
- Enhancing integration capabilities to support the seamless connection of diverse data sources using various protocols and formats.
- Developing a unified semantic structure to enable efficient data classification and interpretation.
- Expanding real-time analytics capabilities to support more timely and informed operational decisions.
- Modernizing legacy systems to ensure scalability in line with growing data demands in terms of volume, variety, and velocity.

To address these challenges, the project introduces a comprehensive and scalable enterprise architecture framework designed to unify data sources, enhance interoperability. The proposed solution is as follows:

- Introduces a Unified Namespace (UNS) built on MQTT and HiveMQ to standardize communication and data flow.
- Implements semantic data classification using ISA-95 to organize all incoming data from sensors, applications, and control systems.
- Uses a multi-layered technology stack with Influx DB, AWS RDS, and S3 for optimized data handling.
- Enables real-time dashboards, predictive analytics, and historical trend visualization to support smarter, faster decision-making.
- Embeds a strong cybersecurity and compliance layer, ensuring secure data exchange and regulatory alignment.

1.3 Summary of Key Findings

- The current data architecture presents an opportunity to move towards a more connected and integrated framework, enabling efficient data management across various projects.
- A unified platform for managing data ingestion, classification, storage, and access would significantly enhance both operational efficiency and strategic decision-making.
- Enhancing real-time observability and control across systems can unlock the full potential of automation, predictive maintenance, and personalized customer services.
- Different stakeholder groups—such as traffic operators, analysts, customers, and regulators—have diverse data needs, highlighting the importance of building a flexible and scalable infrastructure to accommodate them.
- Implementing a standard semantic model can streamline the integration of data from numerous sensors and subsystems, improving accuracy and scalability.
- Introducing identity-based access controls and centralized monitoring mechanisms will strengthen the security and governance of sensitive data assets.

1.4 Business Impact

The proposed architecture delivers significant business value and transformation potential for Irish Rail:

- Operational Efficiency
  Real-time dashboards and anomaly detection systems improve train management, reduce delays, and allow better traffic flow and load balancing.
- Informed Decision-Making
  Centralized data visibility, historical insights, and predictive analytics enable smarter planning, budgeting, maintenance, and crisis response.
- Enhanced Customer Experience
  Personalized services, real-time journey updates, seamless ticketing, and better support systems improve the passenger journey across touchpoints.
- Cost Optimization
  Reduces manual interventions, avoids redundant system builds, and ensures efficient resource allocation through automation and system consolidation.
- Regulatory Compliance
  Ensures that Irish Rail meets evolving data privacy and safety requirements with full traceability, role-based access, and audit support.
- Future Scalability
  The modular architecture allows Irish Rail to integrate future services like AI-based forecasting, autonomous systems, or multimodal transport coordination without overhauling the core system.
- Cross-Stakeholder Collaboration
  Provides common access and shared views to customers, regulators, technical staff, and analytics teams, improving trust, transparency, and alignment.

2 Introduction

Iarnród Éireann Irish Rail is undergoing a major transformation to modernize its operations and infrastructure through multiple concurrent large-scale projects. These initiatives are generating vast and varied data, creating a strategic opportunity to enhance operational efficiency and enable informed, real-time decision-making. However, the lack of a standardized reference architecture hinders effective data integration, consistent collection, scalable storage, and meaningful analysis across the organization.

To address these challenges, this project proposes a comprehensive Enterprise Reference Architecture tailored for Irish Rail. The architecture includes a semantic hierarchy for organizing data and aligns with TOGAF's Business, Data, Application, and Technology layers. It serves as a blueprint for unified data integration, governance, and analytics—empowering Irish Rail with improved operational oversight and the agility to evolve into a data-driven enterprise.

2.1 Project Objectives and Goal

The primary objective of this project is to design a reference architecture that enables seamless integration, collection, storage, analysis, and visualization of data across the organization. This architecture will support the creation of a single, unified view of data across all systems, empowering Iarnród Éireann Irish Rail to make more coordinated and data-driven decisions. By achieving this, Irish Rail will be able to:

- Enhance operational efficiency through real-time data insights.
- Improve decision-making processes with accurate and comprehensive data analysis.
- Organize and categorize data in a meaningful way using a semantic hierarchy.
- Provide a historical view of the enterprise for trend analysis and reporting.

Projects goals include:

- Establish a unified data management platform that enables seamless integration of diverse data sources and provides a consistent enterprise-wide view of data.
- Ensure robust data collection mechanisms that preserve data accuracy while minimizing redundancy.
- Implement scalable and cost-effective storage solutions capable of handling high-volume, high-velocity data.
- Provide advanced analytics and visualization tools for actionable insights at all organizational levels.
- Deliver a target state architecture with clearly defined business, data, application, and technology components.
- Build internal capabilities for continuous monitoring and evolution of the architecture post-deployment.

2.2 Scope of the Project

The scope of this project encompasses the development of a comprehensive reference architecture for Iarnród Éireann Irish Rail. With multiple projects generating vast volumes of data, the goal is to build a standardized, future-ready architecture that enables efficient data management, meaningful insights, and improved enterprise performance.

The project will focus on the following key areas:

1. Enterprise Architecture Design
   Develop high-level EA models aligned with TOGAF, covering:
   - Business Architecture: Define core processes, organizational functions, and business capabilities aligned with strategic goals.
   - Data Architecture: Specify data sources, flows, storage, and implement a semantic hierarchy.
   - Application Architecture: Identify systems and tools for data processing, integration, and delivery, ensuring modularity and alignment with business needs.
   - Technology Architecture: Define infrastructure components including cloud, storage, and compute platforms to support scalable and secure data operations.
   - Security Architecture: Establish security principles covering identity management, data protection, and compliance with regulatory standards.
   - Source Data Policy: Ensures the collection, storage, and integration of accurate, secure, and standardized data from sensor networks and operational systems to support reliable decision-making and regulatory compliance.

2. Data Integration and Governance (Future Improvement Focus)
   - Design a framework for seamless integration of diverse data sources.
   - Ensure data consistency, security, and compliance with governance standards.
   - Define ownership, access controls, and quality metrics for data management.

3. Semantic Hierarchy and Data Organization
   - Develop a semantic model to organize, categorize, and interpret data uniformly.
   - Enable structured knowledge representation for improved interoperability and searchability.

4. Implementation Roadmap
   - Outline phases, milestones, and key activities required for architecture deployment.

This scope ensures that the resulting architecture framework is adaptable, scalable, and capable of guiding Irish Rail's transition into a data-centric organization, enabling both operational excellence and strategic foresight.

2.3 Research Methodology

The methodology for this project follows a structured, multi-phase approach, utilizing both qualitative and quantitative research methods to ensure a comprehensive and effective design of the reference architecture for Iarnród Éireann Irish Rail. The research methodology is outlined as follows:

1. Requirement Analysis

- Gather detailed requirements from key stakeholders (e.g., operational teams, IT, management). Document review of existing architecture and reports.

2. Literature Review

- Review best practices, industry standards, and frameworks related to Enterprise Architecture (EA), data management, and transformation projects.
- Review of academic papers, white papers, and industry reports on enterprise architecture (TOGAF), data integration, and semantic data models.
- Comparative analysis of similar case studies in the transport sector and large-scale data transformations.

3. Design and Framework Development

- Design the reference architecture based on the gathered requirements and existing frameworks (e.g., TOGAF).
- System modelling using ArchiMate for creating visual EA models.
- Development of a semantic hierarchy to organize and categorize data.
- Identification of key data sources and integration points.

4. Evaluation and Testing

- Evaluate the reference architecture against functional and performance benchmarks.
- Feedback collection from stakeholders to assess the effectiveness of the architecture and tools.
- Comparison with existing systems to ensure improvements in efficiency, scalability, and usability.

5. Documentation and Reporting

- Document the findings, design process, and outcomes of the project.
- Technical documentation of the reference architecture and its components.
- Final report presenting the implementation roadmap, architecture models, and governance policies.
- Presentation of results to stakeholders for feedback and final approval.

3 Problem Statement

Iarnród Éireann Irish Rail is currently undergoing a major transformation to modernize its operations and infrastructure through multiple concurrent, large-scale projects. These initiatives are expected to generate massive volumes of complex and diverse data originating from various sources—ranging from operational systems, IoT sensors, passenger services, and external providers such as weather data platforms. However, the organization lacks a standardized, scalable enterprise reference architecture to manage this growing data ecosystem effectively.

The absence of such a framework makes it extremely challenging to integrate data from different systems and protocols, collect it reliably, store it efficiently, and analyze it in a way that produces actionable insights. This fragmentation leads to inconsistent data handling, operational inefficiencies, and a lack of real-time and historical visibility into system performance. The existing reliance on legacy infrastructure further compounds these challenges, resulting in poor interoperability and no unified structure to classify and contextualize enterprise data.

3.1 Description of the Problem

The core of the problem lies in the inability to scale existing data systems and lack of uniformity in data management techniques. As Irish Rail's transformation progresses, the diversity of data formats and the volume of incoming information will continue to grow, demanding a solution that is robust, secure, and adaptable.

To address these challenges, the project necessitates the design of a comprehensive reference architecture—one that not only enables integration with a wide variety of data sources but also ensures compliance with regulatory and safety standards. This architecture must establish a “single source of truth”, leveraging semantic hierarchies to facilitate interoperability, real-time decision-making, and future scalability.

3.2 Importance of solving this problem

Solving this architectural gap is essential to Irish Rail’s long-term success. Without a cohesive and unified data framework, the organization will face significant limitations in executing its transformation strategy. The implementation of a robust reference architecture will:

- Enable real-time insights and monitoring, improving responsiveness across services.
- Support historical data analysis, trend detection, and performance benchmarking.
- Foster data governance and standardization, critical for collaboration and compliance.
- Unlock operational efficiency and agility, ensuring the system evolves with the organization’s needs.

- Lay the foundation for AI and predictive analytics, which require structured, high-quality data.

Ultimately, a well-implemented architecture will act as the backbone of a data-driven, future-ready railway enterprise.

3.3 Stakeholders Involved or Affected

- Irish Rail

1. Traffic management system operators: Operators in charge of signaling and switching process
2. Operations and maintenance team: Provide and access train schedules, train health status data, etc.
3. Customer experience and analytics team: Require data for growth rate projections, etc.
4. Other Irish Rail employees: Help desk authorities, ticketing officers, train staff.

- Customers of Irish rail

1. Passengers of Irish rail train services
2. Freight customers

- External entities

1. Regulatory and compliance authorities: Government, auditors, etc.
2. Weather data providers

3.4 Current Challenges with the existing enterprise architecture

The current architecture at Irish Rail is fragmented, outdated, and not scalable. It lacks:

- A standardized reference architecture for data management.
- A centralized semantic hierarchy for classifying and contextualizing data.
- An interoperable platform to connect diverse systems and protocols.
- Real-time data orchestration and historical data visibility.
- Resilience to handle growing demands and system complexity.

As a result, departments operate in silos, strategic decisions are delayed due to slow data availability, and integration with new systems becomes costly and time-consuming.

4 Current State Analysis

4.1 Description of the current EA in the Organization

The current Enterprise Architecture (EA) of Irish Rail is composed of a well-established set of systems and technologies that support the organization's daily operations. Key components such as sensors, control systems, PLCs, local databases, and actuators are actively utilized to manage infrastructure and ensure service continuity. While these systems function effectively within their respective domains, they primarily operate as standalone entities. The existing architecture reflects a distributed model, where data is stored and accessed locally across subsystems. The following diagram and table provide a detailed overview of these components, distinguishing between those already in place and new elements proposed in the enhanced architecture.

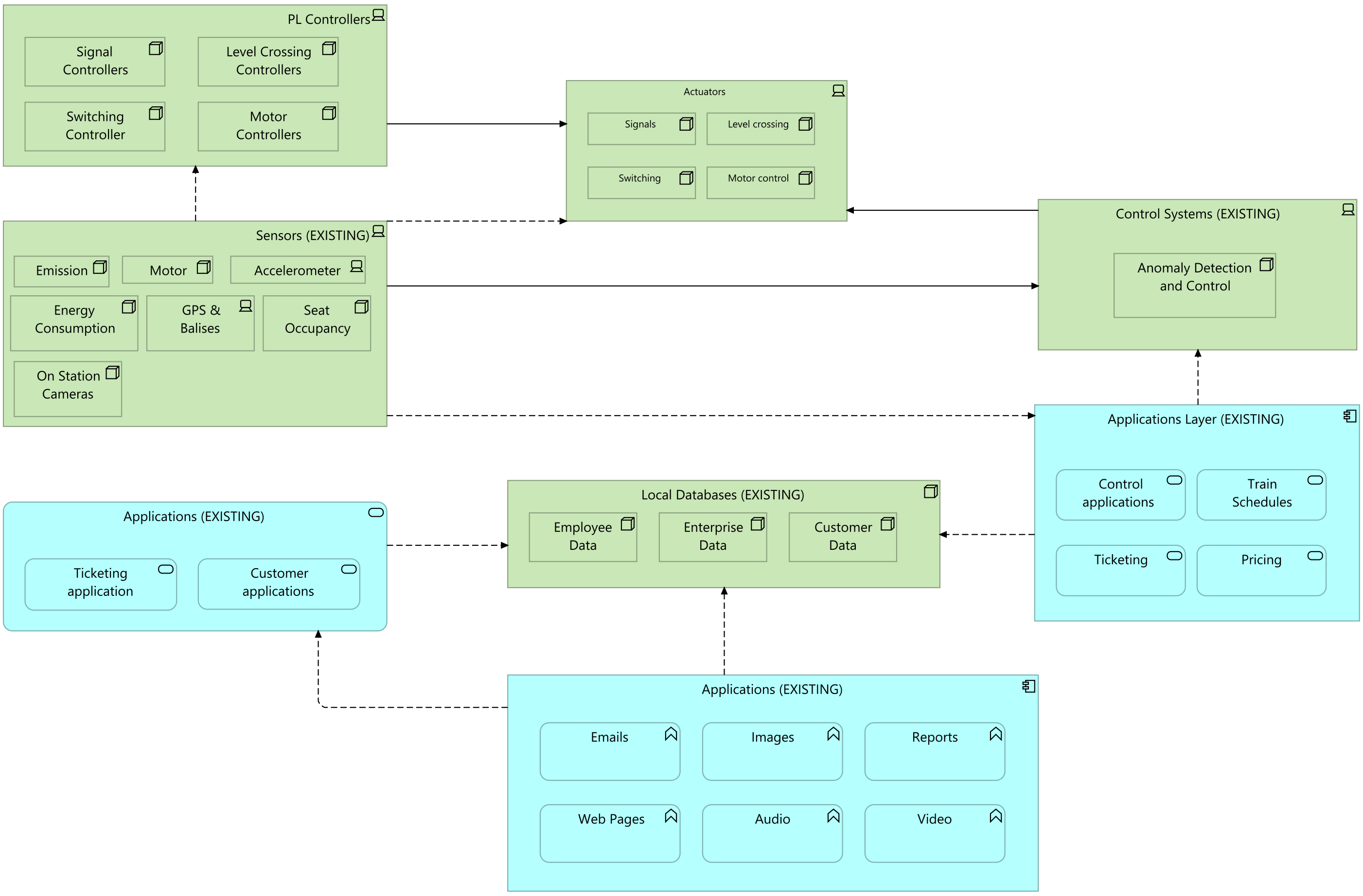

Figure 4.1 Existing Enterprise Architecture

To enable a more adaptive, data-driven environment, the enhanced architecture should incorporate a centralized data platform, real-time analytics capabilities, scalable storage solutions, standardized APIs for cross-system communication, and strong data governance practices. These improvements will support future scalability, operational agility, and improved customer experience across the Irish Rail network.

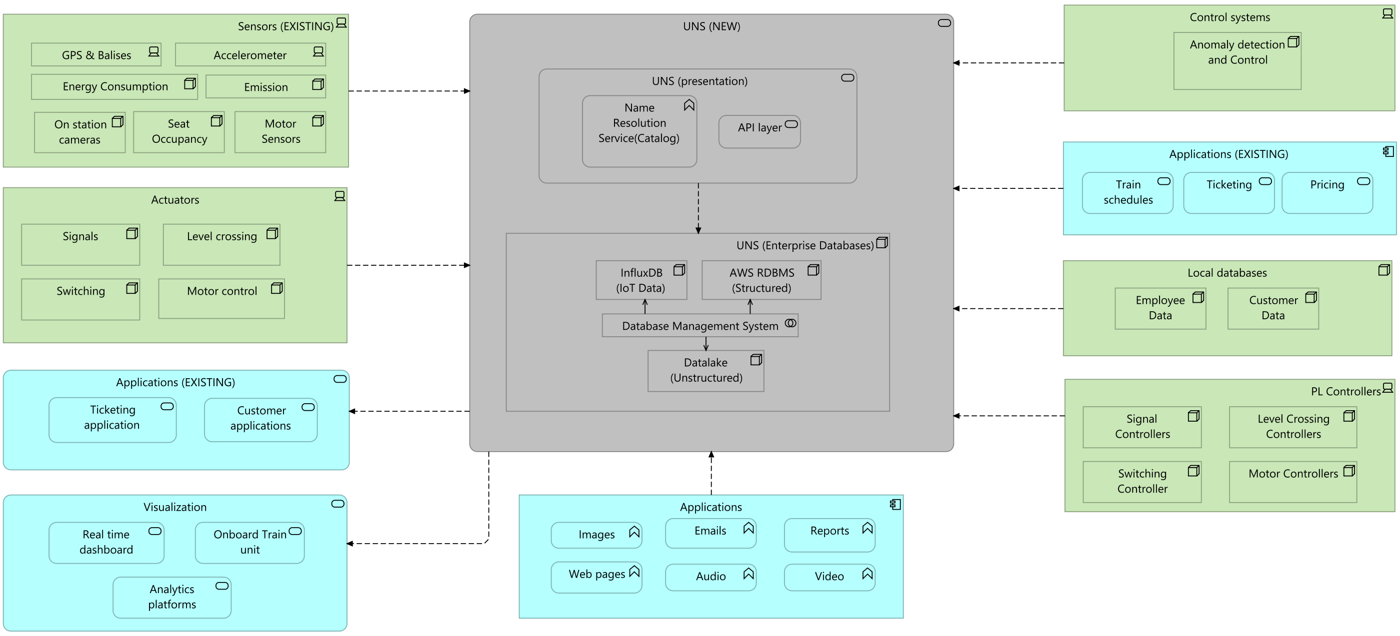

Figure 4.2 Proposed EA to bridge the gaps in the current EA (Condensed version)

Table 4.1 Component Status (Existing vs New)

Component	Existing / New	Description
Sensors	Existing	Sensors networks exist within the Irish rail enterprise but, they do not have a central data destination which they report readings to.
Control systems	Existing	Control systems exist in the Irish rail current architecture, but their state is only known at the control centres and isn't centrally available.
PLC controllers	Existing	Programmable logic controllers require data from sensors and need to trigger actuators but there lacks a central source of truth for this data.
Local databases	Existing	Existing local databases store local data and are not centrally available. To enable centralization of data they require to be connected to a central data store.
Actuators	Existing	Actuator systems exist within the current Irish rail network, but they require to be controlled through many sources. This process could be made more efficient with the use of the proposed UNS.
UNS (Enterprise databases)	New	The UNS enterprise database layer is one of the new additions of the proposed data architecture which improves the current Irish rail architecture by providing a single source of storage for historic data of the Irish rail enterprise.
UNS (Presentation)	New	The UNS presentation layer is absent in the current architecture and is part of the new proposed data architecture. It serves as a forwarding mechanism for all real time data within the Irish rail enterprise. It serves as a view of the current state of the entire enterprise at any point in time.

Applications	Existing/New	Applications exist within the current architecture of the Irish rail but there is no existing way to integrate them with other systems like sensor networks. New proposed architecture consists of visualization platforms to visualize data collected and stored.

4.2 Stakeholder Needs and Expectations

- Irish Rail
  1. Traffic management system operators: Require a central view of the status of all signals and locations of trains updated in real time
  2. Operations and maintenance team: Require access to the current status of all trains, crowd levels at stations, historic trends of users, train health data, etc. Customer experience and analytics team: Require access to customer footfall and growth data.
  3. Other Irish Rail employees (Help desk authorities, ticketing officers, train staff): Require access to live train and platform updates. Live train occupancy data and customer information to be available centrally.

- Customers of Irish rail
  1. Passengers of Irish rail train services: Require customer help applications that rely on new data systems, train occupancy levels pre boarding, platform crowd decongestion, etc.
  2. Freight customers: Require logs of payload data, data on emissions and energy consumed.

- External entities
  1. Regulatory and compliance authorities (Government, auditors, etc.): Require logs of data to make sure the organization meets all compliance norms.

5 Solution/Approach

This section presents the comprehensive solution designed to address the key challenges identified in Irish Rail's current data and systems architecture. It details the proposed Enterprise Architecture (EA) framework tailored to the organization's needs, outlines the envisioned future-state architecture, and highlights the tools, technologies, and methodologies selected to support its implementation. Additionally, it includes an analysis of cost-benefit considerations and evaluates the scalability and flexibility of the proposed solution to ensure long-term adaptability and value.

5.1 Proposed EA Architectures

The following section outlines the proposed enterprise architecture, structured in accordance with the TOGAF framework, and organized across the core layers—Business, Data, Application, Security & Source Data Policy and Technology. Each architecture is designed to address specific operational challenges while ensuring seamless integration, scalability.

5.1.1 Business Architecture

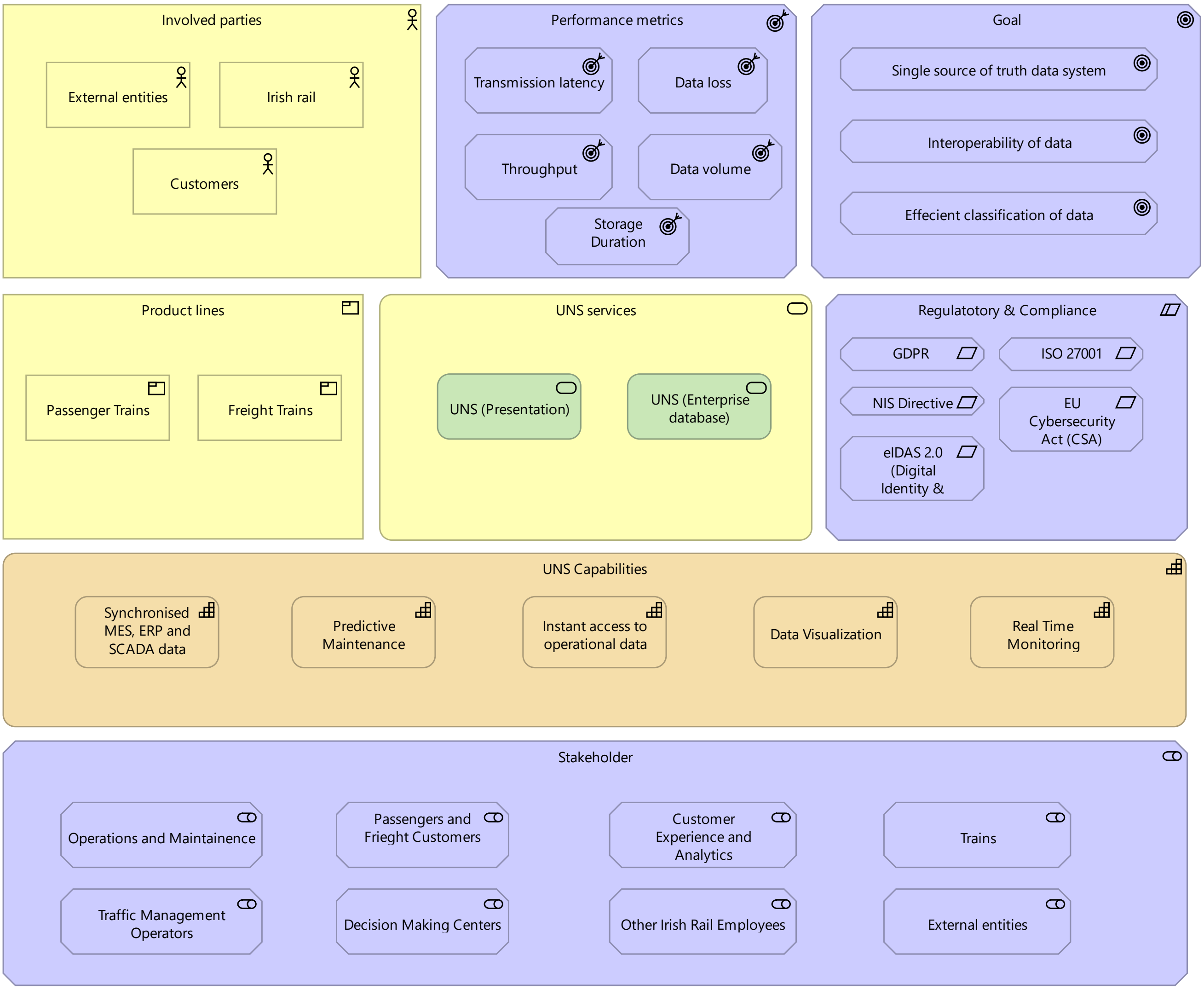

Figure 5.1 Business Architecture

The Irish Rail Business Architecture outlines the strategic, operational, and technological components that enable the delivery of efficient rail services. This architecture aligns stakeholders, capabilities, goals, and regulatory compliance under a unified operational framework, facilitating informed decision-making and real-time responsiveness. The Irish Rail Business Architecture facilitates a connected, data-driven operational environment, where passenger and freight services are monitored, analysed, and optimized in real time. The Involved Parties form the primary actors interacting with the system, Irish Rail (internal), customers (end users), and external entities (partners or authorities). The Product Lines represent the core service offerings.

The system's strategic Goals aim to establish a unified data ecosystem that provides consistency, enables seamless data exchange across systems, and supports automation via structured classification.

UNS Services act as the data orchestration layer, connecting disparate data sources into a unified presentation and enterprise repository. Through the UNS Capabilities, Irish Rail ensures intelligent, real-time insights through functionalities such as predictive maintenance, visualization, and synchronized operations with MES, ERP, and SCADA systems. Performance is governed by clearly defined Metrics, tracking aspects like latency and throughput to guarantee reliable digital operations. This is supported by robust Regulatory and Compliance mechanisms that ensure lawful, secure handling of sensitive data across national and EU standards.

Finally, the architecture is driven by a wide base of Stakeholders, including operations, customer analytics, and senior decision-makers, all of whom leverage real-time insights to drive value and enhance customer satisfaction.

The proposed business architecture introduces a more agile and integrated operational model that enhances value delivery across both passenger and freight services. It emphasizes stakeholder collaboration, improved customer experience, and streamlined operational workflows. By aligning business goals with unified data management through the UNS layer, Irish Rail is positioned to reduce operational silos, improve regulatory responsiveness, and foster data-driven culture across departments. This architecture enables Irish Rail to adapt swiftly to market changes, regulatory shifts, and technological advancements, forming the foundation for future digital transformation initiatives in the transport sector.

5.1.2 Business Capabilities

Business capabilities represent the core functions and competencies that an organization must possess to achieve its strategic objectives. In the context of Irish Rail's digital transformation, these capabilities define what the organization needs to do—independent of how it is done—enabling alignment between business goals and technology solutions. The proposed enterprise architecture enhances and enables these capabilities by providing the necessary data infrastructure, integration mechanisms, and analytics tools to support efficient, responsive, and customer-centric operations.

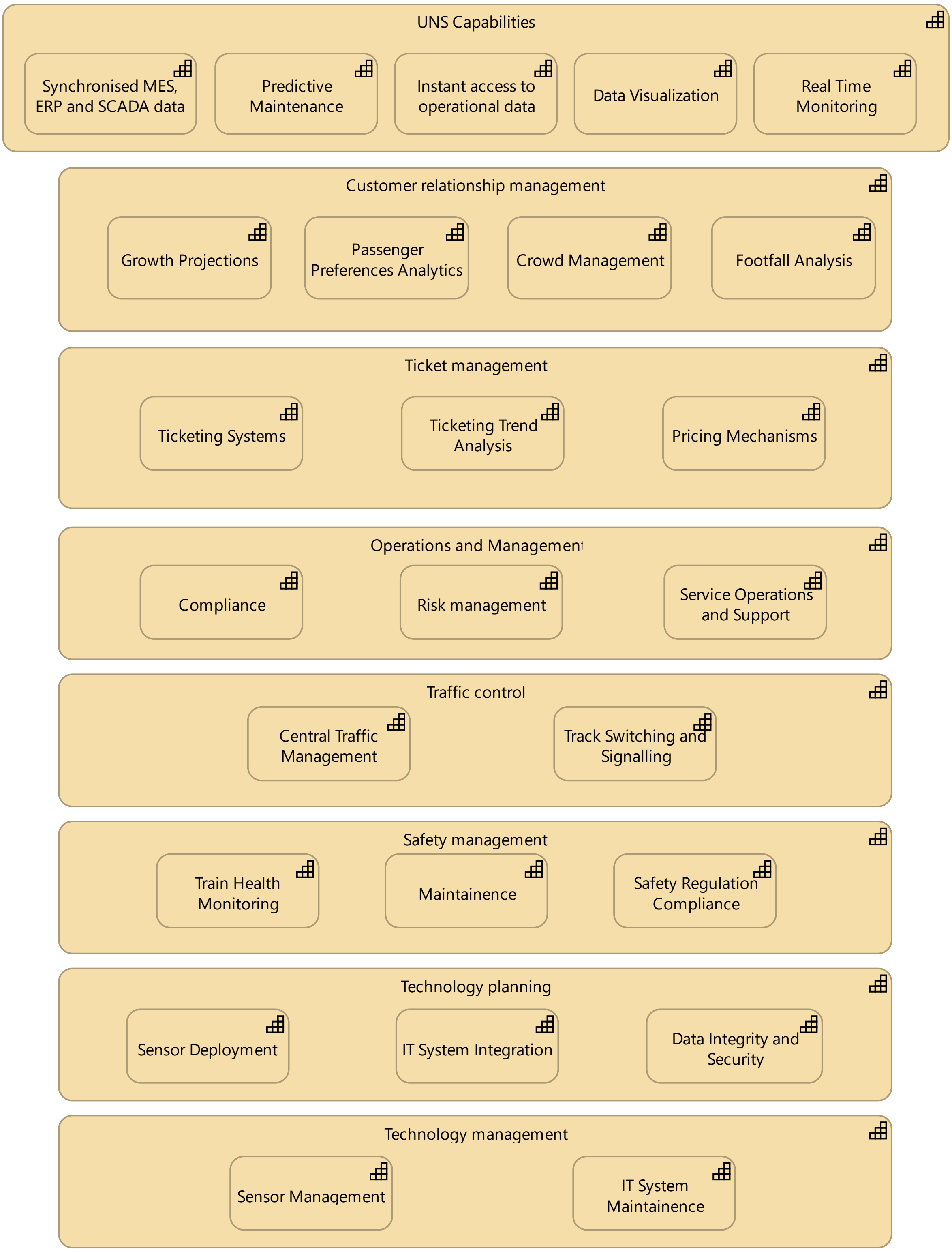

Figure 5.2 Business Capabilities

- Core Capability - Unified Namespace (UNS)

  The Unified Namespace (UNS) acts as the central nervous system of the Irish Rail data ecosystem, enabling synchronized, real-time access to all operational data across the enterprise. It facilitates seamless data collection, integration, processing, and visualization, empowering stakeholders at every level to make faster, more informed decisions.

  Key UNS Capabilities:

  1. Synchronized MES, ERP, and SCADA Data: Harmonizes industrial systems with business processes for end-to-end visibility.
  2. Predictive Maintenance: Enables early detection of equipment faults to prevent service disruptions.
  3. Instant Access to Operational Data: Provides real-time insights across control centres, customer services, and operations.
  4. Data Visualization: Supports decision-making through interactive dashboards and analytics.
  5. Real-Time Monitoring: Ensures proactive operational oversight, enhancing safety and efficiency.

- Supporting Capabilities

  These domains build upon the UNS foundation, ensuring that operational, customer-facing, and technical processes are optimized, secure, and data-driven:

  1. Ticket Management

  Manages core components of fare collection, pricing, and demand forecasting:

  - Ticketing Systems: Platforms for booking, validation, and issuance of tickets.
  - Ticketing Trend Analysis: Identifies purchase patterns and popular routes to refine services.
  - Pricing Mechanisms: Implements dynamic pricing based on demand and customer segmentation.

  2. Operations and Management

  Handles operational governance and daily service quality:

  - Compliance: Adheres to legal, safety, and environmental standards.
  - Risk Management: Proactively identifies and mitigates operational risks.
  - Service Operations and Support: Manages seamless station and train operations for improved customer experience.

  3. Traffic Control

  Ensures efficient and safe train movement through centralized control:

  - Central Traffic Management: Oversees and coordinates real-time train schedules.

- Track Switching and Signalling: Manages track paths and signal systems to prevent collisions and ensure on-time performance.

4. Safety Management

Focuses on continuous monitoring and compliance with safety standards:

- Train Health Monitoring: Tracks real-time train and equipment health.
- Maintenance: Supports predictive and scheduled maintenance activities.
- Safety Regulation Compliance: Ensures operational adherence to regulatory mandates.

5. Technology Planning

Lays the groundwork for scalable, future-ready systems:

- Sensor Deployment: Strategically places IoT sensors to gather environmental and operational data.
- IT System Integration: Seamlessly connects subsystems like ticketing, scheduling, and analytics.
- Data Integrity and Security: Implements encryption, access controls, and audit mechanisms for data reliability.

6. Technology Management

Manages the lifecycle and upkeep of deployed technology:

- Sensor Management: Oversees calibration, maintenance, and replacement of field devices.
- IT System Maintenance: Maintains operational uptime and reliability of critical applications and databases.

7. Customer Relationship Management (CRM)

This layer focuses on enhancing passenger satisfaction, engagement, and experience through data driven insights:

- Growth Projections: Forecasts passenger growth trends to assist in infrastructure and capacity
- planning.
- Passenger Preferences Analytics: Uses historical data to personalize services and understand
- customer behavior.
- Crowd Management: Monitors and controls crowd flow, especially during peak hours or
- emergencies.
- Footfall Analysis: Tracks station-level and train-level occupancy for better scheduling and resource allocation.

By positioning the Unified Namespace (UNS) as the primary enabler, Irish Rail's architecture ensures that all business, operational, and technical capabilities work in synchrony. The supporting capabilities strengthen the foundation laid by the UNS, collectively creating a modular, scalable, and data-driven smart railway system ready for future innovation.

5.1.3 Data Architecture & Permanent Data

The data architecture for Irish rail involves the design of a single source of truth (the UNS) for all data within the Irish rail enterprise. It involves defining how each data source publishes data to the UNS and how the data is organized. It also defines what data in the UNS which publishers will have access to. It also defines the semantic hierarchy based on which the data is classified using ISA95 framework.

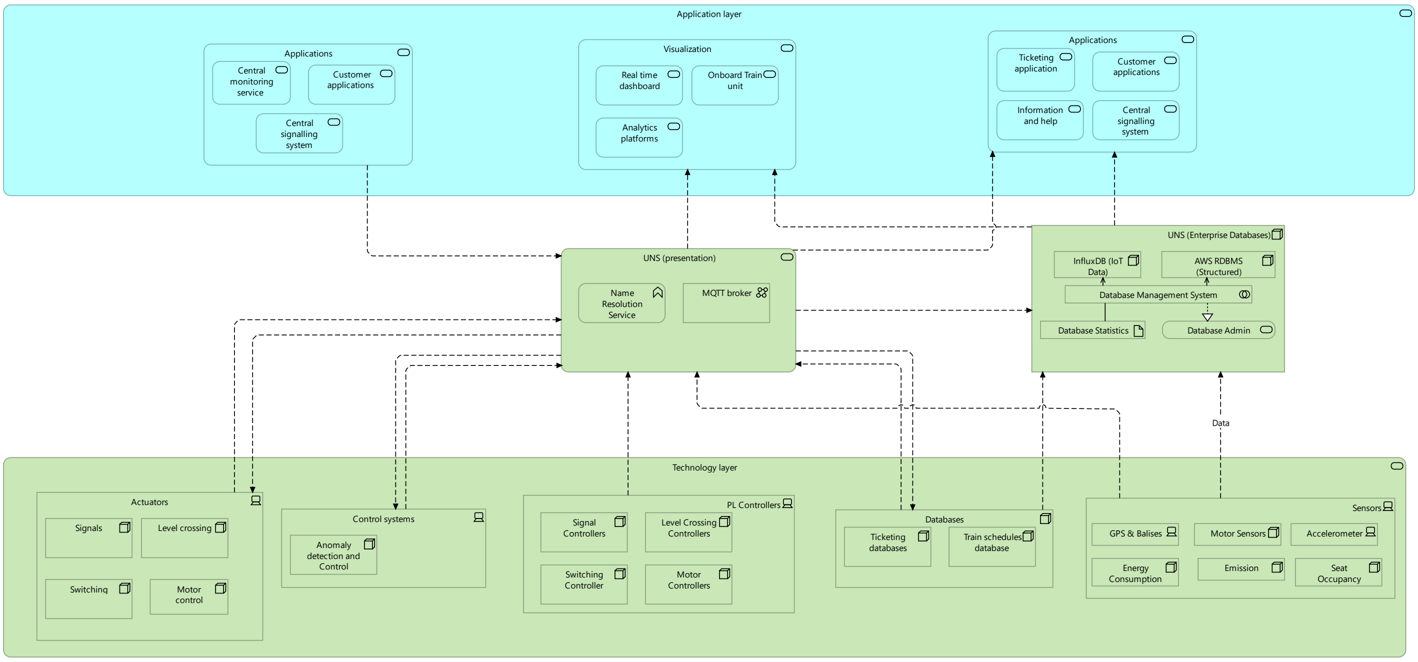

Figure 5.3 Data Architecture

While the presentation layer of the UNS acts as a routing mechanism for real time data without providing storage capability, the enterprise database component of the UNS provides a centralized storage mechanism for all data needed for the long term. The data flow diagram defines how data from each of the sources interacts with the data systems. It pictorially represents flow of data between application layer, technology layer, UNS presentation layer and UNS enterprise database layer. The below table gives all the elements and corresponding description to the components in the architecture.

Table 5.1 Components in Data Architecture

Component	Elements	Description
UNS (presentation)	Name resolution service	The function that resolves data and categorizes it into its designated position in the UNS.
	MQTT broker	The technology that facilitates the UNS service. Acts as a broker forwarding data between data publishers and data subscribers.

UNS (Enterprise database)	Time series database	Permanent data store for data with timestamps. Data on which time graphs need to be plotted.
	Structured data database	Database to store structured data whose format is predefined and known
	Database statistics	Element that logs database transaction related data for maintenance and management of the database systems.
	Database admin	The entity that manages and maintains the database systems and responds in case of failures.
Sensors	GPS and Balises	GPS and Balises are used to obtain real time train speeds and track speed limit data
	Motors sensors	Sensors that measure RPM voltage etc that motors operate at
	Energy consumption	Measures energy consumed by the train in standard units
	Accelerometer	Measures acceleration data of trains
	Emission sensors	Measures emission levels of different pollutants.
	Seat occupancy sensors	Weight sensors fitted on seats to determine if a particular seat is occupied or not.
PL controllers	Signal controllers	Controller input for on track signals for trains
	Level crossing controllers	Controller input for level crossings to regulate road-rail crossings
	Switching controllers	Controller inputs for track switches
	Motor controller	Controller input for train motors and switch motors
Actuators	Signals	The signal lights that are used to guide trains located on tracks
	Level crossings	Powered barriers that control level crossings
	Switches	Switches onboard tracks to switch trains to different tracks
	Motor control	Motor control electronics that control motors onboard trains and switches
Control systems	Anomaly detection and control	Monitoring and alarm systems that are used to flag down anomalies and abnormalities.
Local databases	Ticketing databases	Existing local databases that are being used in the current system for storing ticketing data.
	Train schedules databases	Existing local databases that are being used in the current system for storing train scheduling data.
Applications	Ticketing applications	Software applications that are employed to handle ticketing requirements of Irish rail
	Information and help applications	Applications designed to provide information to customers and help them when necessary
	Customer applications	Other applications offered by Irish rail used to help customers like Irish rail application, journey planning application, etc

	Central signalling system	Central application that is used to control all signals across the Irish rail
	Central monitoring system	Central application that is used to monitor all anomalies and abnormalities in sensors and actuators across the Irish rail
Visualization	Real time dashboard	Real time dashboard to visualize all the data from the UNS presentation layer to obtain a picture of the current state of the Irish rail at any point of time.
	Onboard Train unit	Visual display onboard trains that is used to display metrics like its speed, motor RPM, track speed limit, etc
	Analytics platforms	Platforms designed to perform analysis and extract meaningful insights from real time and stored data

While the UNS presentation layer deals with forwarding of real time data, it does not provide any data storage capabilities. To represent this, we define a permanent data architecture to represent flow of the data that requires to be stored for the long term.

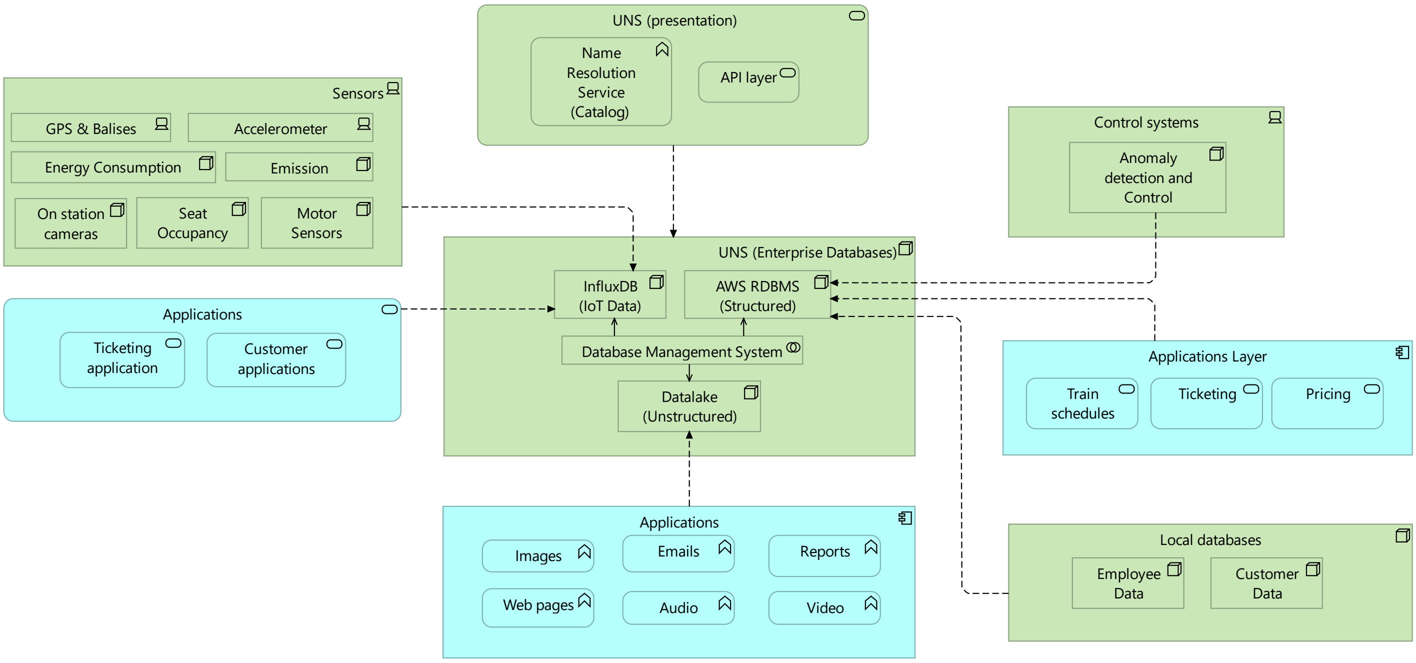

Figure 5.4 Permanent Data – Architecture

The below table gives all the elements and corresponding description to the components in the architecture.

Table 5.2 Components of Data Architecture

Component	Elements	Description
Sensors (Data to time series database)	GPS and Balises	Used to log train speed data, station arrival times and speed limit violations over time.
	Accelerometer	Used to store train acceleration data over time to aid in predictive maintenance
	Energy consumption	Energy consumption over time can be used to derive insights into efficiency
	Emission	Emission data over time is used for predictive maintenance and emission logging
	On station cameras	Data from on station cameras over time can be used to decongest congested platforms, derive station footfall statistics, etc
	Seat occupancy	Used to derive insights like coach occupancy, train occupancy levels.
	Motor sensors	Used to monitor train health
Applications (Data to time series database)	Ticketing applications	Ticketing trends over time used for growth projections and ticket price determination
	Customer applications	Historic data from other customer applications like usage data
Control systems (To structured data database)	Anomaly detection and control	Historic record of each anomaly, its level and its time of occurrence for safety analysis
Applications (Data to structured data database)	Train schedules	Train schedules data in a structured form for other applications to use
	Ticketing	Ticketing applications that store ticketing information in structured form
	Pricing	Historical pricing data for trend analysis
Local databases (Data to structured database)	Employee data	Employee data in a structured format stored centrally on enterprise databases
	Customer data	Customer data stored in structured format in enterprise databases.
Applications (Data to unstructured data database)	Images	Images from Irish rail websites/applications
	Emails	Important email data from Irish rail
	Reports	Growth reports, Footfall reports, etc
	Web pages	Web pages for Irish rail related websites
	Audio	Audio recordings of the Irish rail
	Video	Video recordings of the Irish rail
UNS(Presentation)	Name resolution service	The data from each level of the UNS presentation layer presented to the enterprise database
	API layer	Application programming interface component that facilitates communication between the presentation layer and the data layer of the UNS

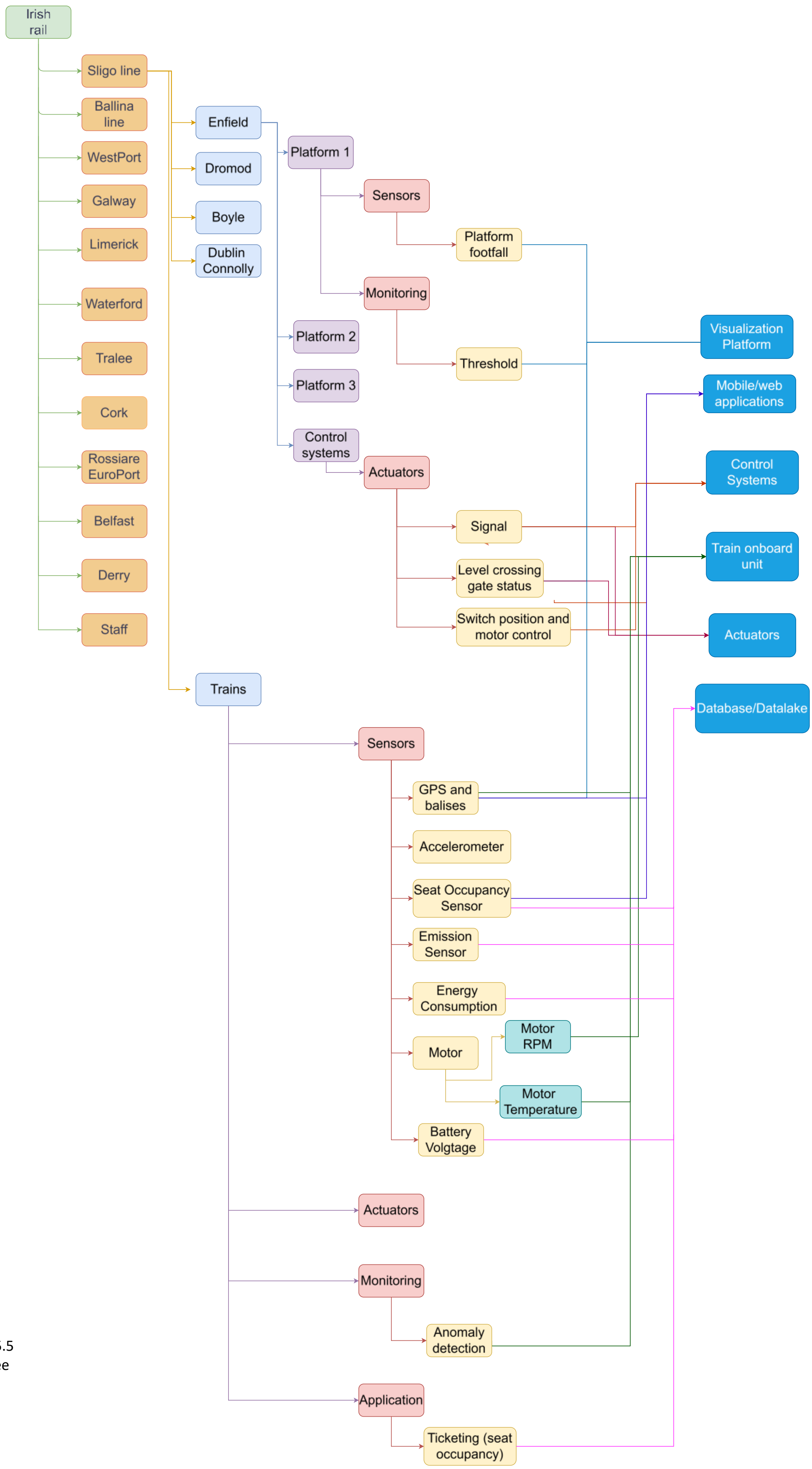

Figure 5.5
UNS Tree

The UNS topic tree is architected using the ISA95 standard as reference.

- **Level 1**: Irish Rail
  The first level consists of all the real time data under the Irish rail enterprise.

- **Level 2**: Lines
  The next level of classification is done based on the individual lines of the Irish rail network. The station data under each line is classified separately leading to the next level of the hierarchy

- **Level 3**: Stations and Trains
  Under each line data is further classified as belonging to each station or train.

- **Level 4**: Platforms and control systems:
  Under each station data is further classified into platforms on the station and the control and signalling systems under the control that station.

- **Level 5**: Classes of data
  Sensors: Data from devices that convert physical input into electrical signals
  Actuators: Data from devices that convert electrical input into physical output
  SCADA: Monitoring systems that trigger on threshold breach
  Applications: IT systems that provide as interface between users and data
  Databases: Stores of data stored in structured format

- **Level 6**: Individual data publishers

  Data subscribers
  Visualization platforms
  Mobile/web applications
  Control systems
  Onboard train unit
  Actuators
  Enterprise databases

Data published by each publisher is classified into this hierarchy and each subscriber subscribes to one or more of the topics in the UNS topic tree.

5.1.4 Security Architecture

The Irish Rail Enterprise Security Architecture provides a comprehensive, layered approach to ensure the integrity, confidentiality, and availability of operational systems and data across the railway network. The architecture integrates modern cybersecurity standards, operational control systems, secure communication protocols, and regulatory compliance mandates to protect critical rail infrastructure.

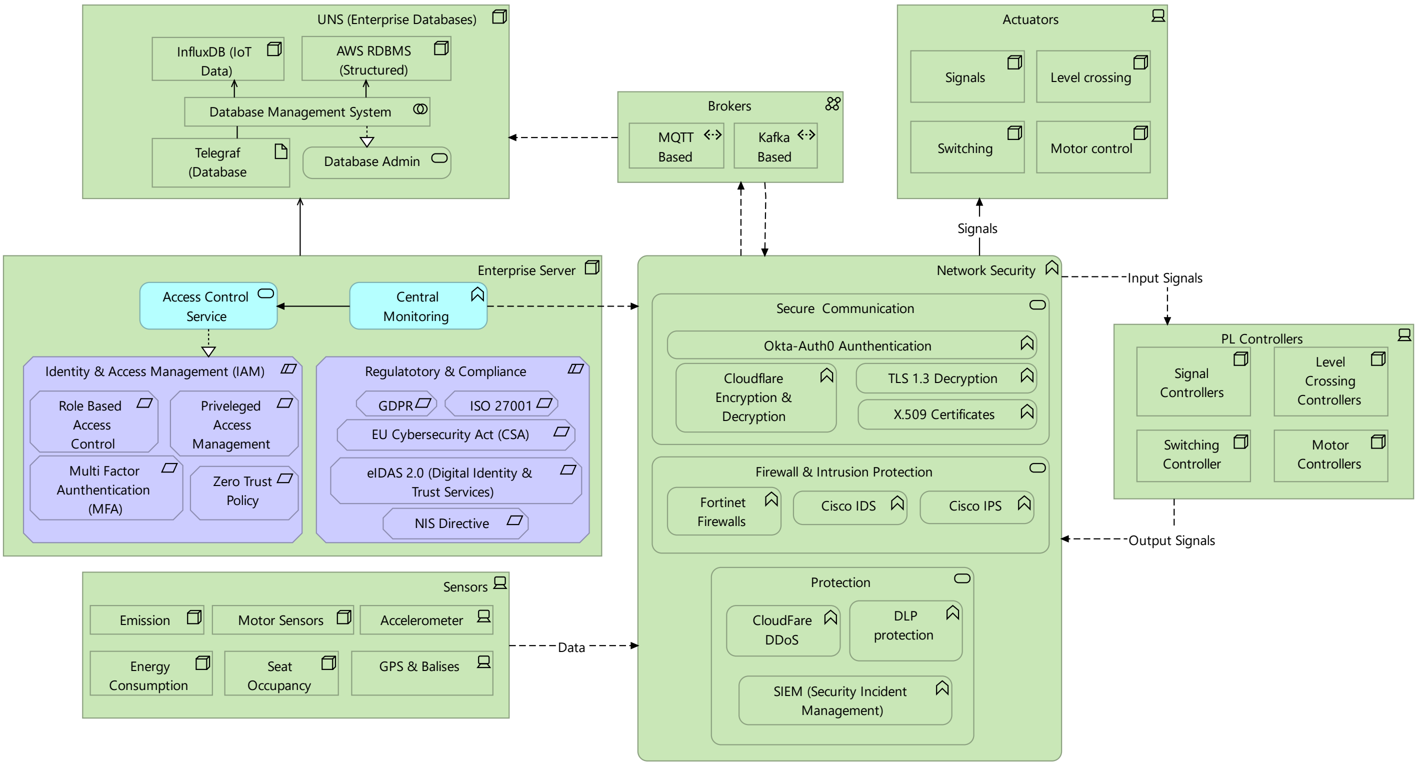

Figure 5.6 Security Architecture

At the heart of this architecture is the Enterprise Server, responsible for managing access control and central monitoring. It connects to various domains including sensors, databases, actuators, PL controllers, and brokers. These domains represent different functions in railway operations from data acquisition and control to real-time communication and automation.

Security is enforced through strong Identity & Access Management (IAM) mechanisms such as Role-Based Access Control (RBAC), Multi-Factor Authentication (MFA), Privileged Access Management, and Zero Trust policies. The architecture also supports compliance with global regulations such as GDPR, ISO 27001, and the EU Cybersecurity Act. Communication is safeguarded with TLS 1.3, X.509 certificates, and Okta-Auth0 based authentication. Network security components like Fortinet Firewalls, Cisco IDS/IPS, and Cloudflare DDoS Protection ensure intrusion detection and prevention. Additionally, SIEM tools help monitor and respond to security incidents in real time.
The table below explains all the elements with respect to the security architecture,

Table 5.3 Components in Security Architecture

Component	Elements	Description
Enterprise Server	Access Control Service	Controls which users/devices/services can access parts of the system based on their credentials and roles.
	Central Monitoring	Provides real-time visibility into system status, alerts, performance, and security events across the enterprise.

Identity & Access Management (IAM)	Role-Based Access Control (RBAC)	Users get permissions based on their job roles, limiting access to only what they need.
	Privileged Access Management	Special management for high-level users (like admins) who have deeper access, reducing insider threat risks.
	Multi-Factor Authentication (MFA)	Requires two or more forms of identity verification like password + mobile OTP for secure access.
	Zero Trust Policy	Assumes no user or system is trusted by default always verify before granting access, even inside the network.
Regulatory & Compliance	GDPR	General Data Protection Regulation – EU's law on data privacy and protection. Must be followed for any user data.
	ISO 27001	International standard for Information Security Management Systems (ISMS). Ensures security best practices.
	EU Cybersecurity Act (CSA)	Framework to enforce cybersecurity standards across digital services and devices in the EU.
	eIDAS 2.0	Digital Identity framework to securely manage digital trust services (e-signatures, identities, etc.).
	NIS Directive	Directive on security of network and information systems across critical sectors like transport.
Network Security	Okta-Auth0 Authentication	Cloud-based identity management platform — manages user login, authentication, and SSO (Single Sign-On).
	TLS 1.3 Decryption	Latest encryption protocol that protects data in transit (e.g., between sensors and servers) by encrypting the connection.
	Cloudflare Encryption & Decryption	Cloudflare ensures secure and encrypted traffic through its content delivery and security network.
	X.509 Certificates	Digital certificates used in public key infrastructure (PKI) for secure identity and communication verification.
	Fortinet Firewalls	Hardware or software appliances that control incoming/outgoing traffic based on security rules to prevent attacks.
	Cisco IDS (Intrusion Detection System)	Monitors network for suspicious activity and alerts admins of potential threats.
	Cisco IPS (Intrusion Prevention System)	Like IDS but actively blocks detected threats automatically.
	Cloudflare DDoS	Protects systems from Distributed Denial-of-Service attacks which try to overwhelm servers with fake traffic.
	DLP Protection	Data Loss Prevention – monitors and blocks sensitive data from leaving the network accidentally or maliciously.

	SIEM (Security Incident & Event Management)	Collects and analyzes security data from across systems to detect threats, create alerts, and support forensics.
Brokers	MQTT Based	Lightweight messaging protocol ideal for transmitting small packets (like sensor readings) over unreliable networks.
	Kafka Based	Distributed messaging system for handling real-time, large-scale data streams — ensures reliability and scalability.

5.1.5 Application Architecture

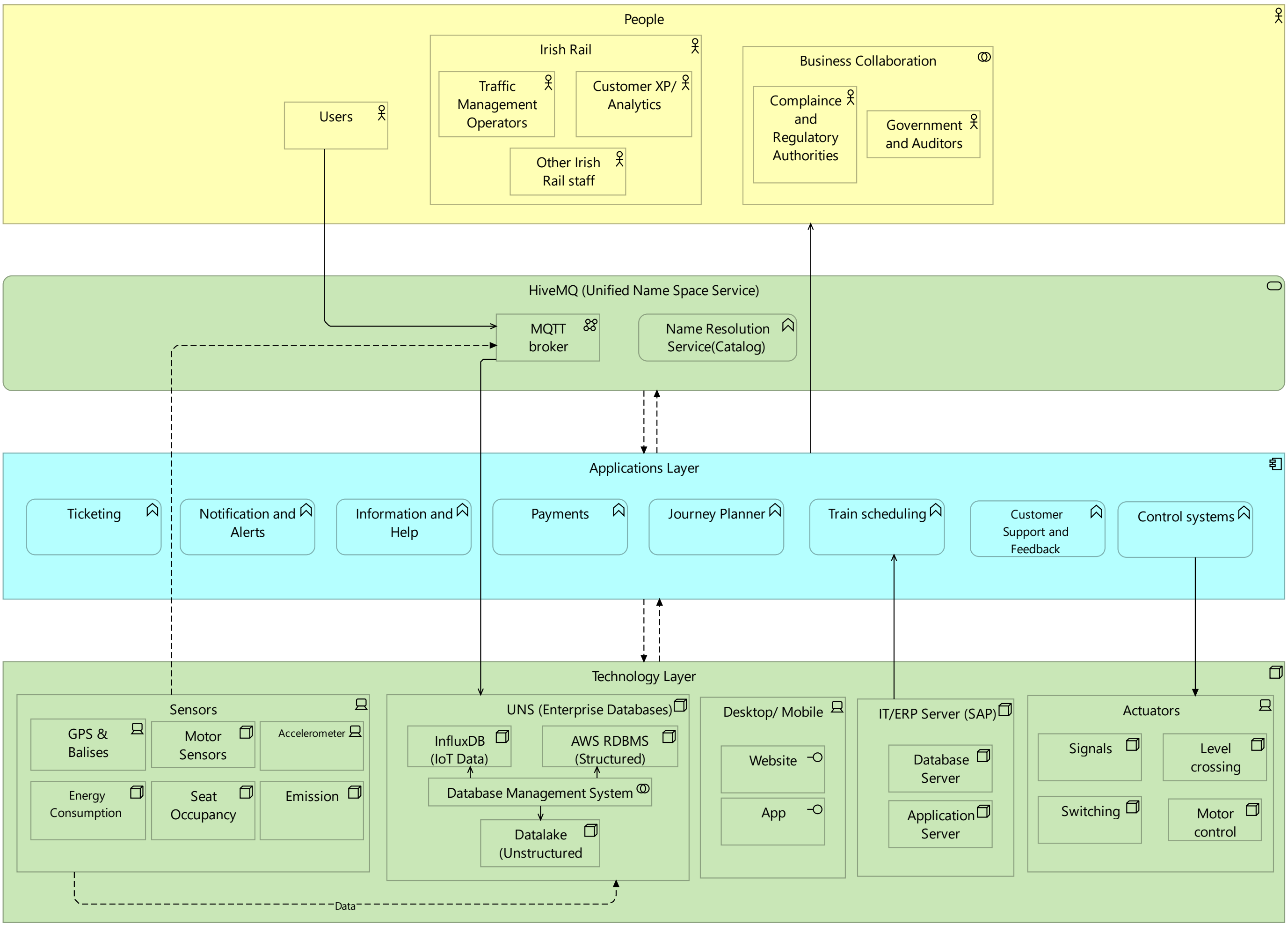

Figure 5.7 Application Architecture

1. People Layer

   Stakeholders Involved:

   • Users: General public or passengers accessing applications such as ticketing, journey planners, and support.

   • Irish Rail Staff:

- Traffic Management Operators: Monitor and control train movements.
- Customer XP/Analytics: Analyse user experience and feedback.
- Other Irish Rail Staff: Include administrative and support personnel.
- Business Collaboration:
- Compliance and Regulatory Authorities
- Government and Auditors

These stakeholders interact with the system through various interfaces or utilize insights generated by data analytics and system reports.

2. Integration Layer (HiveMQ - Unified Namespace Service)
Components:
• MQTT Broker: Enables lightweight, publish-subscribe messaging between sensors/devices and the cloud.
• Name Resolution Service (Catalog): Manages identification and service registration across the distributed system.
This layer acts as the communication backbone, facilitating secure and scalable messaging between edge devices, enterprise systems, and application services.

3. Applications Layer
This layer hosts all end-user and operational services, tightly integrated with backend systems and real-time data feeds:
• Ticketing: Booking, reservations, and travel authorization.
• Notification and Alerts: Real-time alerts on delays, maintenance, and emergencies.
• Information and Help: Support channels, FAQs, and system-related queries.
• Payments: Secure payment processing for tickets and services.
• Journey Planner: Route optimization, ETA predictions, and schedule lookups.
• Train Scheduling: Dynamic scheduling based on data analytics and operational demand.
• Customer Support and Feedback: Feedback collection and issue resolution.
• Control Systems: Direct interfaces to operational control and automation systems.

4. Technology Layer
Sensors collect raw operational data from physical train and track environments, including:
• GPS & Balises: Location and speed tracking.
• Motor Sensors: Mechanical performance.
• Accelerometer: Motion and vibration analysis.
• Energy Consumption: Power usage monitoring.
• Seat Occupancy: Passenger load.
• Emission Sensors: Environmental compliance.
These feed data into the system through MQTT-based communication.

Databases & Management Systems

• InfluxDB: Time-series database for IoT data like sensor streams.

• AWS RDBMS: Structured data handling for enterprise applications.

• Datalake: Storage for unstructured, historical, or batch analytics data.

• Database Management System (DBMS): Manages data integrity, access control, and schema evolution.

Enterprise Systems

• UNS (Unified Namespace): A logical layer that organizes and indexes data sources across the system for discoverability.

• IT/ERP Server (SAP):

  o Application Server: Runs backend logic.

  o Database Server: Handles structured business data (e.g., schedules, financials).

Interfaces (Desktop/Mobile)

• Website & App: Accessible UIs for passengers and staff.

• These applications retrieve real-time data from databases and cloud services to present updated information to users.

Actuators and Control

• Signals, Switching, Level Crossing: Receive automated or manual commands based on insights from control systems.

• Motor Control: Fine-tuned control for vehicle propulsion or braking systems.

5. Data Flow Summary

• Sensor data is collected and streamed via MQTT to the cloud.

• Data is parsed and stored in appropriate databases (time-series, structured, unstructured).

• Applications access and visualize this data for both end-users and staff.

• The system issues alerts or controls based on predefined rules or ML algorithms.

• Feedback loops enable users and authorities to refine service performance and reliability.

6. Benefits of Architecture

• Scalability: Modular components and MQTT-based messaging allow easy expansion.

• Real-Time Operations: Enables instant alerting and decision-making.

• Data-Driven Decisions: Centralized analytics with historical and live data sources.

• Security & Compliance: Ensures data integrity and accessibility for regulatory bodies.

5.1.6 Source Data Policy

The source data policy shown in the architecture prioritizes structured ingestion, secure processing, and compliance in all layers. The data comes from different sensors, PLC controllers, and enterprise systems such as UPS, HVAC, ticketing, energy, and security systems. They all flow into the system using data sources, both real-time sensor streams and batch data (such as ticketing or energy consumption). PLCs and control systems provide operational continuity and safety, whereas enterprise services add administrative information (e.g., access logs).

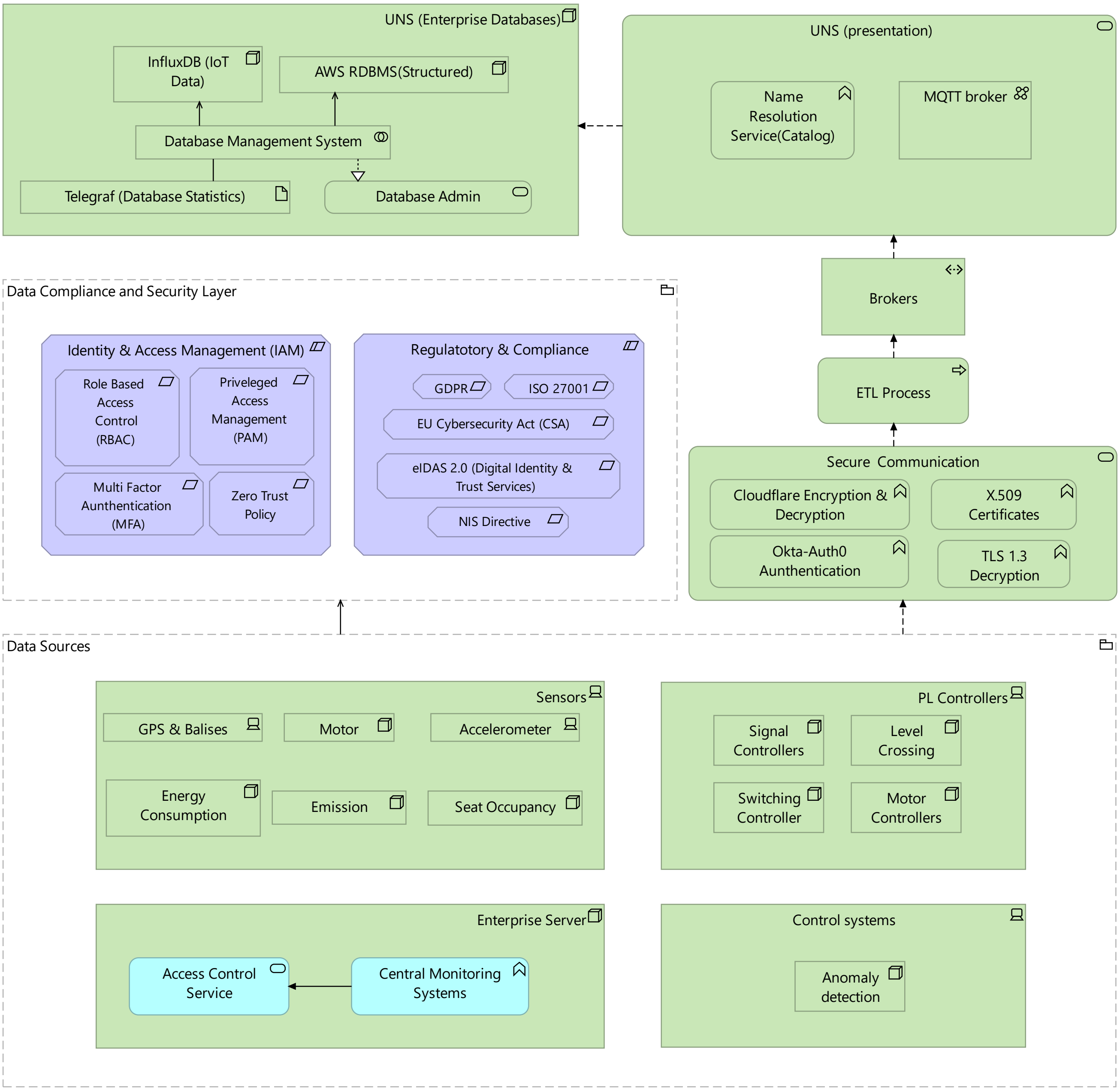

Figure 5.8 Source Data Policy

All incoming data enters a single namespace (UNS) through MQTT brokers and service catalog layers, guaranteeing consistency and discoverability. Based on their type, data is stored in either InfluxDB (time-series) or AWS RDS/Redshift (structured relational) for processing and analytics. ETL processes properly transform and route this data, overseen by specialized database admins and statisticians to ensure integrity and performance.

A robust security and compliance data layer envelops the pipeline to comply with regulation requirements such as GDPR and NIS2. Okta/Auth0 for identity, Cloudflare and TLS 1.3 for secure transport, and encryption protect data both in transit and at rest. SIEM and DLP systems keep an eye on activity and apply policy, while IAM platforms manage access control. This infrastructure makes Irish Rail's data capture secure, compliant, and dependable for real-time and historical insights.

5.1.7 Technology Architecture

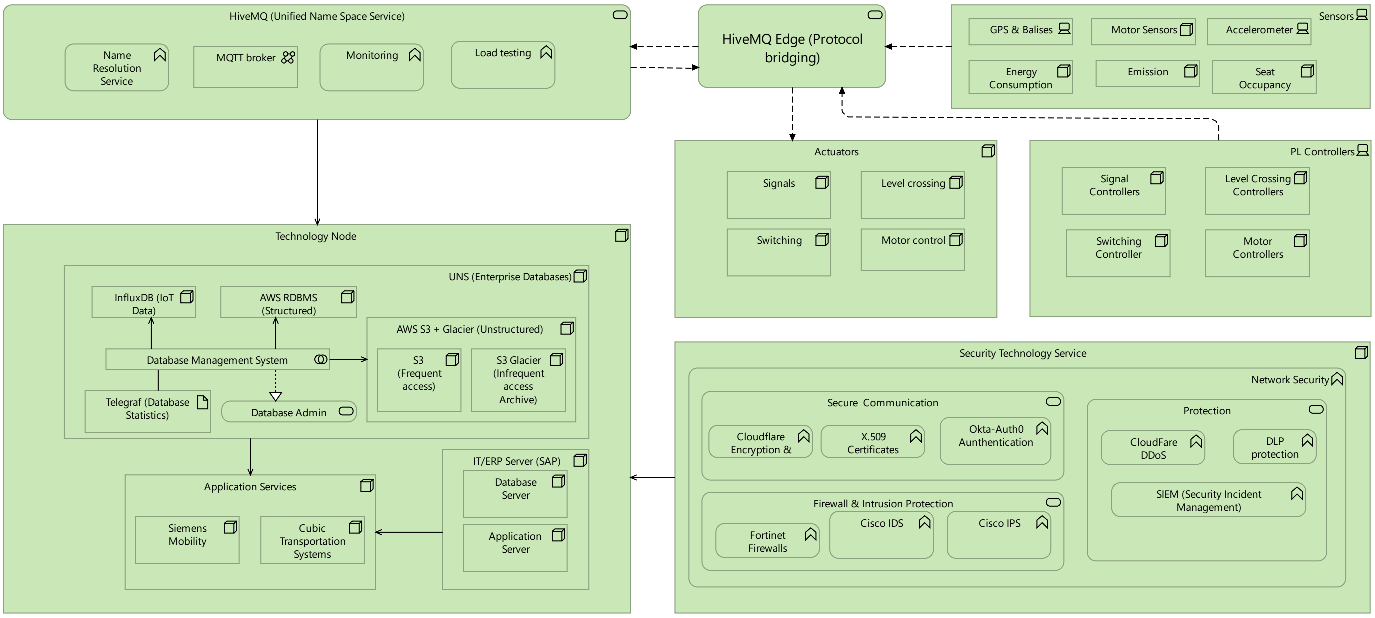

Figure 5.9 Technology Architecture

The above technology architecture is carefully designed and uses the cutting-edge technologies that can be easily integrated into the existing Irish Rail System. Each of the elements will be explained with their pros and cons and other alternatives below.

- Unified Name Space (UNS):
  The UNS makes use of HiveMQ, an industry-leading MQTT Broker, to enable reliable and secure communication across Irish Rail's infrastructure. It facilitates the seamless collection and distribution of sensor data from trains, IT systems and other data sources which publish data to the UNS.

Table 5.4 Brokers

Technology	Pros	Cons
Hive MQ	i. Flexible Deployment ii. High Throughput & Low Latency iii. Enterprise-grade Security with authorization, authentication and TLS	i. Enterprise features are licensed. ii. Requires Infrastructure Management when using without HiveMQ Cloud iii. Needs external tools (like Kepware) to connect legacy systems.
AWS Core IOT	i. No infrastructure to manage – auto-scales with demand.	i. Not ideal for systems needing local data access. ii. High Latency

	ii. AWS Ecosystem Integration. iii. Built-in Rules Engine. iv. Integration with AWS IAM for security	iii. Limited Edge Support iv. Cost may increase unpredictably with message volume.
EQMX	i. Open-Source Core ii. Supports multiple protocols and acts as a bridge iii. Lightweight, scalable MQTT for mixed edge/cloud systems	i. Requires more effort to configure and maintain ii. Not as mature as HiveMQ in terms of support iii. Some features are commercial

- HiveMQ Edge:
  HiveMQ Edge enables seamless integration between legacy rail systems and modern digital platforms by converting industrial protocols like Modbus and OPC-UA into MQTT. Deployed at the network edge such as stations, depots, or signal cabins it ensures low-latency, secure, and structured data flow into the Unified Namespace.

  This approach reduces complexity, enhances system resilience, and accelerates the adoption of real-time analytics. HiveMQ Edge plays a critical role in bridging operational technology with enterprise systems for data-driven transformation across Irish Rail.

- Time Series Database:
  A Time Series Database like InfluxDB is optimized for storing and querying large volumes of time-stamped sensor data. It enables fast analytics, trend detection, and real-time monitoring—critical for rail operations. InfluxDB supports data down sampling and retention, reducing storage needs over time. Its seamless integration with visualization tools makes it ideal for dashboards and alerting in transport systems like Irish Rail.

Table 5.5 Time-series Databases Providers

Technology	Pros	Cons
InfluxDB	i. Optimized for fast writes, handles millions of points/sec ii. InfluxQL and Flux support powerful queries and transformations iii. Ideal for managing high-volume time series iv. Flexibility to deploy self-hosted or use InfluxDB Cloud	i. InfluxQL is similar but not full SQL ii. True clustering is available only in the Enterprise/Cloud versions iii. Chronic Architectural Instability and Breaking Changes iv. Frequent query language changes

TimescaleDB	i. Full SQL support, joins, indexing, and PostgreSQL tooling ii. Good for complex queries iii. Supports horizontal scaling	i. Not as fast as InfluxDB for time-series ingestion ii. Higher storage footprint than columnar formats iii. Setup complexity
Amazon Timestream	i. Serverless ii. Native AWS integration iii. Secure and managed iv. SQL-like language	i. Vendor lock-in ii. Query limitations iii. Can become expensive for high-ingestion workloads

- Relational Database

  A relational database can manage train schedules, routes, and timetables with complex relationships between stations, trains, and services. It supports ticketing systems, handling user data, bookings, and payment transactions securely. It enables reporting and compliance tracking, with reliable SQL-based queries for audits and KPIs. It ensures data integrity and consistency across interconnected operational systems like rostering, maintenance, and crew management. Oracle offers rich features but comes with high licensing costs and complex management, which can slow down modernization efforts.

  AWS Aurora provides a fully managed, cost-effective, and cloud-native solution, ideal for scaling real-time rail data and passenger services. Aurora supports automated backups, replication, and high availability across multiple zones, reducing infrastructure overhead. For Irish Rail's digital transformation, Aurora enables faster innovation cycles and easier integration with other AWS services like analytics and IoT.

Table 5.6 RDBMS Providers

Technology	Pros	Cons
Amazon Aurora RDS	i. Fully managed by AWS ii. Auto-scaling read replicas iii. Highly available & fault-tolerant	i. Vendor lock-in ii. Limited extensions/support iii. Storage cost can grow quickly with large datasets
Oracle DB	i. Mature and highly reliable ii. Rich support for complex transactions iii. Fine-grained security & auditing	i. Expensive licensing ii. Complex management iii. Harder to scale horizontally

- Data Lakes:

  We chose Amazon S3 for Irish Rail because it offers granular storage tiers like S3 Standard, Deep Archive, and Glacier, enabling smart cost-optimization for both hot and cold rail data. Compared to Microsoft Azure and Google Cloud, S3 provides broader tooling, mature ecosystem, and tighter integration with AWS analytics and IoT services used across the rail infrastructure. S3's fine-grained access control, cross-region replication, and native support for Athena, Glue, and Redshift make it ideal for secure, scalable analytics.

Table 5.7 Data Lake Providers

Technology	Pros	Cons
Amazon S3	i. Pay-per-use pricing with Storage Classes depending upon the use case ii. Low latency and high throughput iii. Virtually unlimited storage iv. Robust security features	i. It can take time to process large datasets ii. Datasets grow, organizing and managing becomes more complex. iii. No built-in querying
Azure Data Lake Storage (ADLS) Gen2	i. Hierarchical namespace ii. Optimized for analytics iii. Multiple storage options (Hot, Cool, and Archive) to optimize costs	i. Azure ecosystem may require additional learning for teams ii. Limited compatibility with non-Azure tools
Google Cloud Storage (GCS)	i. GCS can scale to handle large volumes ii. Excellent integration with Google's BigQuery, Dataflow, and AI/ML tools	i. Limited ecosystem ii. Complex pricing iii. Less comprehensive native tooling for ETL

- Security Technology Service:
  The selected network security technologies were chosen to provide layered protection for Irish Rail's infrastructure. Cloudflare and Okta/Auth0 offer secure communication through encryption, authentication, and DDoS mitigation. Fortinet firewalls and Cisco IDS/IPS ensure deep packet inspection and intrusion detection/prevention. SIEM tools and DLP enhance real-time threat monitoring and safeguard sensitive passenger and operational data.

- Application Services:
  Siemens Mobility and Cubic Transportation Systems were selected for their proven expertise in intelligent transport and fare management systems. Siemens supports critical rail operations with signal control, scheduling, and mobility platforms. Cubic brings smart ticketing and passenger flow optimization, key to modernizing Irish Rail's customer experience. Both align with Irish Rail's goals for efficiency, scalability, and digital transformation.

- The SAP IT/ERP Server stack was chosen to centralize and streamline Irish Rail's financial, asset, and HR operations. The Application Server hosts critical business logic and workflows, ensuring process consistency. The Database Server provides secure, high-performance storage for enterprise data with strong SAP integration. This setup supports scalability, auditability, and integration with other rail systems.

5.2 Cost-Benefit Analysis

These assumptions form the basis for estimating infrastructure requirements, scalability needs, and associated operational costs under varying growth scenarios.

- **Timeframe**: Monthly cost (USD)
- **Users & Devices**:
  - Low Growth: ~5,000 connected devices/users
  - Central Growth: ~50,000 connected devices/users
  - High Growth: ~200,000+ connected devices/users
- **Data Rate**: 1KB/s per device average (adjusts cost on MQTT, DB, S3)

Table 5.8 Individual Cost Breakdown by Scenario

Service / Component	Low Growth (5K devices)	Central Growth (50K devices)	High Growth (200K+ devices)
HiveMQ UNS	$10,000	$20,000	$40,000+
Amazon RDS	$2,000	$5,000	$15,000
Amazon S3	$500	$2,000	$6,000
Amazon Glacier	$200	$700	$2,500
InfluxDB (Cloud)	$1,000	$3,500	$10,000
SAP ERP Licensing	$4,000	$12,000	$25,000+
SAP Training/Maintainence	$2,000	$6,000	$15,000
IoT Hardware (avg/month)	$5,000	$10,000	$25,000
Data Transfer (AWS)	$500	$2,500	$10,000

Growth Scenario	Estimated Monthly Total
Low Growth	**~$25,200**
Central Growth	**~$61,700**
High Growth	**~$148,500+**

Explanation of Costs

1. HiveMQ UNS

Enterprise pricing scales by:

- No. of messages per second
- No. of connected clients
- Throughput + Clustering

Central growth likely needs clustering, persistent sessions, and rule engine use.

2. Amazon RDS

- Low: Single instance (e.g., db.m6g.large, ~$0.10/hour)
- Central: Multi-AZ deployment (db.m6g.4xl+)
- High: Read replicas + failover setup, high IOPS SSD

3. Amazon S3 + Glacier

- Low: 1 TB (data logs, firmware)
- Central: 10–50 TB (historical + audit data)
- High: 100+ TB with lifecycle rules (S3 to Glacier)
- Glacier used for long-term archival (older sensor data, logs, etc.)

4. InfluxDB

- Ingestion = $0.0025/MB, Querying = $0.012/100 queries
- Low: Light analytics (500 MB/day)
- High: Full-scale dashboards, high ingest (10–20 GB/day)

5. SAP ERP

Costs vary by:

- User licenses (named vs concurrent)
- Modules in use (SCM, PM, MM, etc.)
- Integration effort with MQTT/Ingestion layers
- Maintenance & training grow significantly with staff + locations.

6. Industrial IoT Hardware

Monthly average includes:

- Spare parts, replacements and adding new PLCs, sensors
- Maintenance contracts
- High growth = scaled manufacturing plants, hence more PLCs/sensors

7. AWS Data Transfer

AWS offers first 100GB free each month. After that:

- $0.09/GB for first 10TB
- $0.085/GB up to 40TB, etc.
- High growth = large sensor fleets streaming real-time data

€688,572 yearly costs estimate for proposed architecture, calculated as the sum of cost of each individual technology.

The adoption of a modern digital architecture represents a strategic investment in the future. It enables enhanced operational efficiency, supports data-driven decision-making, and provides a scalable foundation for long-term innovation and growth. By improving real-time monitoring, predictive maintenance, and standardized data management, this initiative not only optimizes daily operations but also strengthens regulatory compliance and customer experience.

Below are the key benefits identified for the cost-benefit analysis.
1. Enhanced Operational Efficiency

- Enables real-time monitoring → fewer delays, optimized scheduling.
- Predictive maintenance → reduced downtime & emergency repairs.

2. Data-Driven Decision Making

- Accurate, centralized analytics → smarter strategic & operational decisions.
- Reduces resource waste and misallocation.

3. Semantic Data Organization

- Standardized classification → no data redundancy or misinterpretation.
- Supports future integration of AI, ML, automation tools.

4. Historical & Real-Time Data Views

- Full enterprise visibility → better trend analysis, forecasting.
- Essential for regulatory compliance & performance reporting.

5. Long-Term Strategic Benefits

- Future-proofing through scalable, modular architecture.
- Enables continuous innovation and digital transformation.
- Boosts stakeholder trust and improves customer experience.

The cost of implementing this architecture, with a recurring yearly estimate of €688,572 for technology operations and an additional one-time professional services cost of €1–2 million for infrastructure setup, is marginal when measured against the tangible operational savings, strategic agility, regulatory preparedness, and long-term scalability it enables. It's not merely a cost center—it's an enabler of enterprise-wide transformation and a cornerstone of Irish Rail's future-proof digital foundation.

5.3 Scalability and Flexibility

The primary purpose of the UNS is to enable easy scalability. In contrast with legacy data systems UNS enables interoperability across various data sources. Sources operating using different protocols ultimately use the same forwarding mechanism to forward data to the respective data consumers. The new proposed data systems also enable all data within the enterprise to be accessed from anywhere using a common interface. It also serves as a representation of the current state and the historic state of the entire enterprise. The protocol transformation protocol provides flexibility in enabling all data sources to interact with the data system. The proposed data system integrates IOT sensors, Control electronics and IT systems into a common central system providing flexibility in operating with various types of technology infrastructure.

6 Implementation Plan

6.1 Steps to Implement the Proposed Solution

The proposed solution is implemented in 4 steps. The first step is to set up the IT infrastructure and emulate the actual conditions of the train to refine the IT systems and load test them before investing a lot of money into upgrading the physical infrastructure. It also reduces downtime as the transformation can take place in a simulated environment without disrupting live operations. A pay as you go model offered by a lot of providers. The second step is to upgrade the physical sensor networks, which require a huge investment and a lot of planning. The third step involves the seamless integration of the physical sensor networks into the IT infrastructure, with a focus on maintaining uninterrupted operations and avoiding any system downtime.

Step 0: Finalise design and select product – Request for Proposal

Step 1: Implement the IT Infrastructure and Make Refinements:

- Deploy Core IT Infrastructure
- Integrate Simulated Sensor Inputs
- Perform System Integration Testing
- Conduct Load and Stress Testing
- Enable Monitoring and Logging
- Refine and Optimize
- Validate Against Operational Use Cases
- Prepare for Transition to Physical Deployment

Step 2: Deploy the Physical Sensor Networks:

- Finalize Sensor Specifications and Procurement
- Plan Deployment Strategy
- Install and Commission Sensors
- Validate Real-Time Data Integration
- Implement Ongoing Monitoring and Maintenance
- Scale Deployment Across the Network
- Leverage Sensor Data for Operational Insights

Step 3: Integration of the Sensor Networks

- Integration of Sensor Data Streams into Unified Namespace (UNS) and Enterprise Systems.
- Establishment of Low-Latency Data Ingestion Pipelines and Secure Data Handling.
- Validation of Integration Through Live Scenarios.

An alternative implementation approach could be to adopt an incremental, iterative strategy starting small and gradually expanding. This method allows for early demonstrations of value, helping to build momentum and stakeholder confidence. By focusing on manageable, low-risk components initially, the team can refine the solution based on real-world feedback before scaling up. This flexible approach complements the existing plan and can be considered as a parallel or phased pathway to full implementation.

6.2 Timeline and Key Milestones

This is structured over a 10-month timeline, with realistic buffers for procurement, testing, and integration. This consists of a phase-by-phase implementation map, but the work can also be implemented in a parallel with some milestones implemented in parallel with others to speed up the timeline.

Table 6.1 Timeline & Key Milestones

Month	Phase	Milestone
1	Step 1: IT Infrastructure Setup	- Provision cloud services (Aurora, S3, InfluxDB, HiveMQ) - Deploy virtual environments and simulate sensors
2		- Complete system integration of services (Cloudflare, Okta, Fortinet)
3-4		- Conduct load and stress testing; optimize based on feedback
4		- Validate simulated environment against real use cases
4	Step 2: Physical Sensor Deployment	- Finalize procurement of IoT sensor hardware and edge devices (HiveMQ Edge)
5		- Begin sensor installation on selected pilot routes
5-6		- Complete commissioning and test real-time data streams
6-7	Step 3: Integration and Scaling	- Integrate live sensor feeds into cloud IT stack
7-8		- Monitor performance and uptime across end-to-end system
9-10		- Conduct full system go-live across all routes and issue operational report

6.3 Resources Required

1. People

- Project Manager: To oversee planning, coordination, execution, and delivery of the overall architecture implementation.
- Solution Architects (2): To design and validate the technical architecture, ensuring scalability, security, and alignment with enterprise standards.
- Developers (4): To build, integrate, and customize software components across systems and interfaces. Data Engineers & Integration Specialists, Database Administrators, Security Engineers and DevOps Engineers are preferred.
- Testers (2): To design and execute comprehensive test plans, ensuring system reliability, performance, and compliance with requirements.

2. Tools

- HiveMQ & HiveMQ Edge: For MQTT-based unified namespace and edge protocol bridging.
- InfluxDB: High-performance time-series storage for telemetry and sensor data.
- Amazon Aurora RDS: For structured operational and transactional data handling.
- Amazon S3: Cost-effective storage for raw, transformed, and archived datasets within a data lake architecture.
- Fortinet & Cisco Security Appliances: For perimeter defense and real-time intrusion detection.
- Cloudflare, Okta/Auth0: For encrypted communication, access control, and DDoS protection.
- SIEM, IAM, and Compliance Tools: For maintaining regulatory standards (GDPR, NIS2, ISO 27001).

6.4 Potential Risks and Mitigation Strategies

Table 6.2 Risk and mitigation stratergies

Risk Area	Specific Risks	Mitigation Plan
Resource Requirement	High resource demands for enabling technologies and data architecture setup	Staggered implementation, budget and procurement planning
Time Requirement	Long implementation timelines requiring consistent effort	Use of detailed timeline for phased technology planning
Legacy System Integration	Difficulty in integrating legacy systems into modern architecture	Use of protocol bridges; phased migration to new technologies
Stakeholder Resistance	Reluctance or lack of buy-in from stakeholders	Effective communication of benefits, structured knowledge transfer
Technical Failures	Reliability issues and possible technology failures	Use of redundant systems; procurement from reliable, proven vendors

7 Outcome

7.1 Anticipated outcomes of the proposed architecture-based solutions

- Single source of truth for all data produced by the Irish rail. This data system should be accessible across the Irish rail enterprise.
- Interoperability of different types of data. Existing systems using different communication protocols will be able to communicate with each other using the central interface.
- Efficient classification of data within the data systems to ensure accurate and seamless data publishing and retrieval with minimal overhead.

7.2 Hypothetical Results if implemented

- Improved Resource Utilization: Effective implementation could lead to more efficient allocation and use of resources due to better data visibility and streamlined operations.
- Reduced Time to Insight: A unified data framework can significantly decrease the time required to access, process, and analyse data for decision-making.
- Enhanced System Compatibility: The architecture aims to improve the integration of various systems, including legacy technologies, leading to a more cohesive and functional IT landscape.
- Increased Stakeholder Confidence: Successful implementation and demonstrated benefits can lead to greater buy-in and support from stakeholders across the organization.
- Greater System Stability and Reliability: By addressing existing architectural fragmentation, the proposed solution can result in a more robust and dependable overall system.

7.3 Challenges (hypothetical) that might arise during the implementation of the solution

- Resource requirement: The transformation plan and establishment of all the enabling technologies of the data architecture are resource intensive. Time requirement: The transformation plan requires a good amount of time to be implemented and requires continual work over this period.
- Integration of old technologies: Integration of old technologies and legacy systems into the modern data architecture could be a challenge.
- Resistance to change from stakeholders: All stakeholders must be convinced of the benefits of the transformation plans.
- Technical failures: Reliability issues and failures of the technologies and systems involved could be disruptive.

7.4 Post-Implementation Monitoring and Evaluation

- Transmission latency: Time taken for data to travel from data publishers to the data system and finally the data subscribers
- Data loss: Bytes of data lost as a percentage of data sent
  Throughput: Rate of flow of data supported by the system.
- Data volume: Maximum data capacity of the data system
  Storage duration and IO: Maximum storage duration and frequency of access of stored data.

8 Conclusion

8.1 Summary of the Problem and Proposed Solution

Iarnród Éireann Irish Rail's transformation initiatives are generating increasingly vast and varied data streams, yet the existing enterprise architecture lacks the scalability, interoperability, and standardization needed to manage them effectively. Legacy systems, fragmented data silos, and the absence of a unified data framework have hindered integration, governance, and strategic use of information.

To address these challenges, we proposed a comprehensive Enterprise Reference Architecture grounded in TOGAF principles. The core of the solution is the Unified Namespace (UNS), built on HiveMQ and MQTT, enabling seamless communication between diverse systems and protocols. By incorporating semantic hierarchies, secure data pipelines, and modular scalability, the architecture establishes a centralized data infrastructure—providing real-time visibility, operational intelligence, and long-term adaptability.

8.2 Key Findings and Insights

- Irish Rail's existing architecture does not yet incorporate a 'single source of truth,' which limits the ability to monitor, analyze, and optimize operations effectively.
- Stakeholders across departments have diverse data needs, but current systems do not support interoperability or consistent access control.
- The proposed solution enables real-time analytics, predictive maintenance, and historical trend evaluation—key to modernizing railway operations.
- Our cost-benefit analysis shows that the architecture is financially sustainable across growth scenarios and can significantly enhance long-term efficiency and service delivery.
- By aligning business, data, application, and technology layers, the architecture provides a strong foundation for AI-driven insights, automation, and future innovations.

8.3 Recommendations for Future EA Improvements

- Establish a permanent EA governance body to oversee implementation and ensure architectural alignment across departments and vendors.
- Gradually transition legacy systems to modern, MQTT-compatible devices using protocol bridges like HiveMQ Edge to reduce risk during migration.
- Implement continuous feedback loops from operational stakeholders to iteratively refine semantic hierarchies and access structures.
- Leverage the UNS and real-time data streams to develop AI-based use cases such as passenger flow optimization, energy efficiency tracking, and predictive maintenance scheduling.
- Expand integration with external systems (e.g., weather, GIS, multimodal transit APIs) to enrich decision-making and cross-organization collaboration.

9 References

1. K. Manditereza, "What is Unified Namespace (UNS) and Why Does it Matter?," HiveMQ, Dec. 7, 2022. [Online]. Available: https://www.hivemq.com/blog/what-is-unified-namespace-uns-iiot-industry-40/
2. Iarnród Éireann, *Annual Report 2023*, Irish Rail, 2023. [Online]. Available: https://www.irishrail.ie/Admin/IrishRail/media/Content/About%20Us/Irish-Rail-Annual-Report-2023.pdf
3. The Open Group, "TOGAF® Standard," [Online]. Available: https://www.opengroup.org/togaf
4. K. Manditereza, "Designing Your UNS Semantic Information Hierarchy," HiveMQ, Mar. 19, 2025. [Online]. Available: https://www.hivemq.com/blog/designing-your-uns-semantic-information-hierarchy/
5. Archi, *Archi - Archimate 5 Minute Guides* [YouTube Playlist]. [Online]. Available: https://www.youtube.com/playlist?list=PLUxDCM4ujDqXDI2P2Vm2Ruoj6VfQVJRvv
6. Iarnród Éireann, "Irish Rail – Official Website," [Online]. Available: https://www.irishrail.ie/en-ie
7. Amazon Web Services, "AWS GDPR Center," [Online]. Available: https://aws.amazon.com/compliance/gdpr-center/.
8. Amazon Web Services, "Amazon Aurora," [Online]. Available: https://aws.amazon.com/rds/aurora/.
9. InfluxData, "InfluxData," [Online]. Available: https://www.influxdata.com/.
10. Timescale, "What InfluxDB Got Wrong," [Online]. Available: https://www.timescale.com/blog/what-influxdb-got-wrong.
11. InfluxData, "InfluxDB Pricing," [Online]. Available: https://www.influxdata.com/influxdb-pricing.
12. InfluxData, "Telegraf," [Online]. Available: https://www.influxdata.com/time-series-platform/telegraf/.
13. Amazon Web Services, "AWS Pricing Calculator," [Online]. Available: https://calculator.aws/#/addService.
14. Amazon Web Services, "Amazon S3 Storage Classes," [Online]. Available: https://aws.amazon.com/s3/storage-classes/.
15. HiveMQ, "HiveMQ," [Online]. Available: https://www.hivemq.com/.
16. HiveMQ, "Designing Your UNS Semantic Information Hierarchy," [Online]. Available: https://www.hivemq.com/blog/designing-your-uns-semantic-information-hierarchy/.
17. Siemens Mobility, "Rail Services – Up and Running Guaranteed," [Online]. Available: https://www.mobility.siemens.com/global/en/portfolio/rail-services.html.
18. Cubic Corporation, "Cubic Corporation," [Online]. Available: https://www.cubic.com/.