WeAre

CASE STUDY

WeAre Solutions Oy www.weare.fi/en/contact-us sales@weare.fi

In this project, we supported a financial sector client to **improve their log management systems.** Our team implemented a modern **Splunk and Azure framework** to ensure compliance and enable future data analytics – a particularly important consideration in such a sensitive industry.

Log and data management with Splunk - Case Study

The challenge - log management and compliance

The collaboration started with a clear goal: **to implement a unified solution that would consolidate audit data and guarantee full regulatory compliance.**

At the time, audit logs were being generated across multiple systems and were primarily used for operational monitoring. While some data was already being processed in **Splunk**, there was no dedicated, central solution specifically designed for long-term, immutable retention.

In light of the fact that **regulatory requirements** stipulate that audit log events must be stored in a tamper-proof manner for over a decade (depending on the source system), the organisation decided to establish a structured, future-proof archival approach.

The customer **had been using Splunk for several years** for operational log management and analytics. We designed the solution so that Splunk focuses on what it does best: operational analytics and actionable insight, **with a typical retention of 6-12 months.**

For long-term archival, we needed to store log events in their raw format and enforce WORM (Write Once, Read Many) immutability to guarantee tamper-proof retention. **Azure Storage Account**, which the customer already had, provides these capabilities as a cost-effective storage layer, ensuring the overall architecture remains scalable and efficient.

Depending on legal and operational requirements, data may be stored for varying periods of time. One way to categorize it is using the four-tier hot – warm – cold – archive data model.

> *The **hot-warm-cold-archive** model is a data tiering strategy that optimizes storage costs and performance by categorizing data based on how frequently it is accessed. In this framework, "hot" data remains on high-performance hardware for immediate retrieval, while "warm","cold" and "archive" data are progressively moved to more economical, slower storage as they age and their relevance decreases. Read more about it on our dedicated solution,* ***Observability Hub Log Management & Archive***

The project goal

The primary driver was achieving regulatory compliance by establishing **immutable, long-term storage of audit events**. Beyond that, the project delivered significant additional value: **more visibility**. Splunk enables developers to map all sources and retrieve data, even from the least convenient ones. Previously invisible log sources were integrated into a centralized pipeline, improving security visibility and incident response capabilities.

The architecture also delivered significant cost savings. **Storing up to 1TB per day** with a retention period of over 10 years allows for significant savings with the proposed setup.

Our solution

The technical framework of the solution is centred on a dual-component Kafka architecture designed for reliable data flow. It also uses Splunk components for data retrieval and an Azure Storage Account for archiving.

At the core, **Managed Kafka Brokers** serve as the primary event streaming layer, managing all message routing and buffering. Complementing this is a **self-managed Kafka Connect** cluster, which acts as the integration framework running source and sink connectors for data ingestion and distribution.

Source systems are integrated using a variety of predefined connector and integration types, including:

- **Splunk components** – HEC, S2S (Edge Processor, Heavy/Universal Forwarders), and DB Connect for database-level log extraction
- **Native Azure services** – Event Hubs, direct Storage Account integrations, and other Azure-native ingestion methods

New integration types are added as needed when onboarding new sources, and the Kafka connectors ecosystem provides a wide range of ready-made connectors to support virtually any integration requirement.

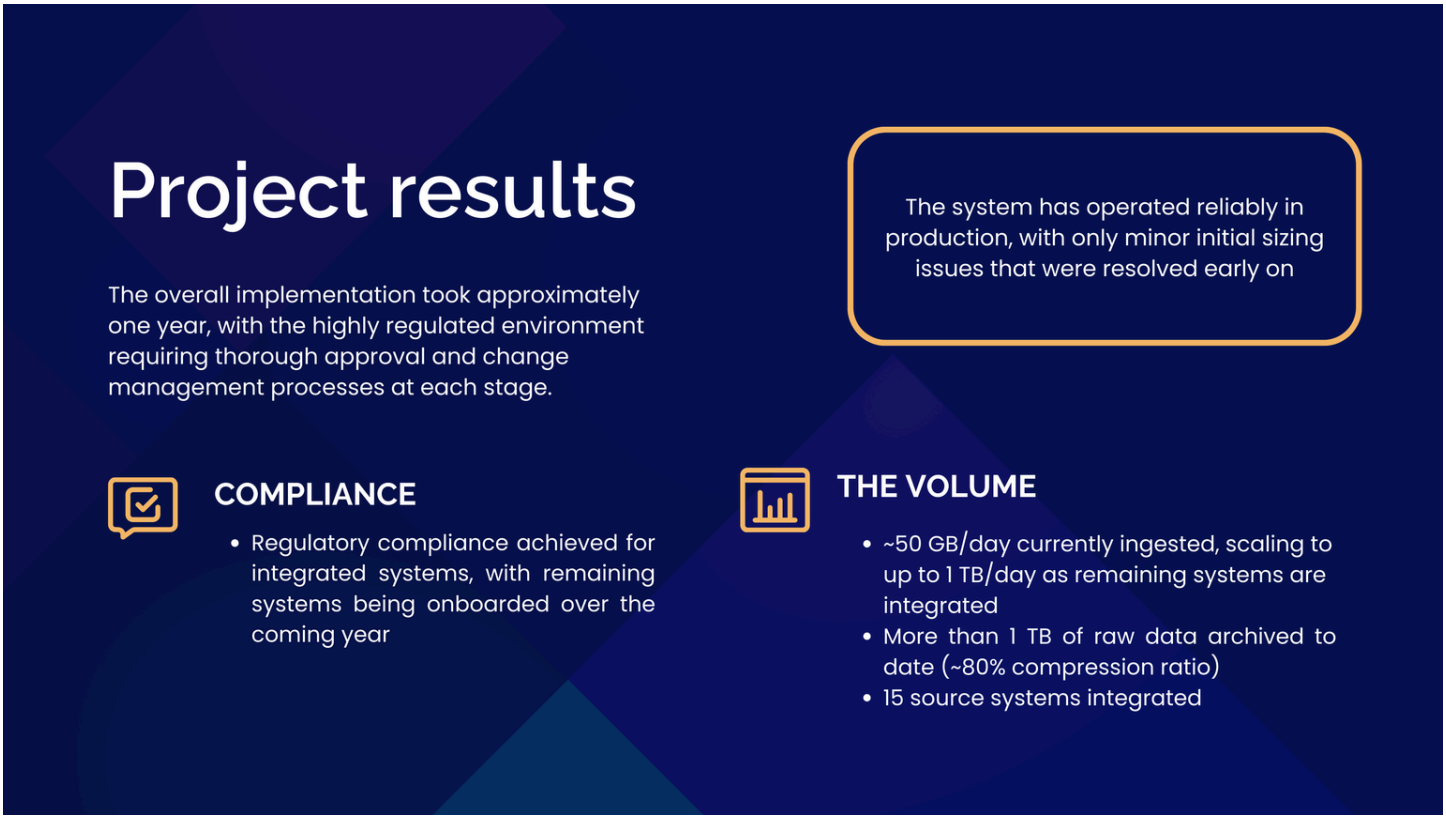
Once data enters the Kafka pipeline, sink connectors distribute it to multiple outputs per topic:

- **Azure Storage Account** with WORM immutability policies – this is always an output for every integration, ensuring compliant long-term archival
- **Splunk or other destinations** – optionally added when a party needs to perform operational analytics or further processing on that data

Delivery approach and team

The team consisted of **5 – 6 people in total**, with 2 developers doing the hands-on implementation. Both developers operated in combined architect/developer/operations roles, covering everything from solution design to infrastructure deployment. The project manager and product owner were internal customer employees.

The work was organized using **the SAFe (Scaled Agile Framework) methodology** with Agile Release Train cadence. Task management and tracking was handled through **Jira**, and day-to-day communication took place via Microsoft Teams.



The project results

This project significantly **improved log management and log archival** for our client and ensured regulatory compliance achieved for integrated systems. With **~50 GB/day currently ingested**, the data volume will be scaling to up to **1 TB/day** as remaining systems are integrated. To this date, we have archived more than **1 TB of raw data** (with an impressive **~80% compression ratio**).

We've introduced the immutable retention of **10+1 years** guaranteed via Azure WORM policies, which creates significant cost savings at this volume and retention period.

The system has **operated reliably in production**, with only minor initial sizing issues that were resolved early on.

Currently **15 different source systems** are integrated (with remaining systems being onboarded over the coming year), each potentially having multiple distinct log sources. These cover areas such as core business applications, transaction processing, risk management, CRM, and platform audit logs.

The project was delivered iteratively, with **new integration types and source systems added in phases.** The overall implementation took approximately one year, with the highly regulated environment requiring thorough approval and change management processes at each stage.

Why log management and observability?

The number of systems that produce data is not decreasing. In fact, there are more and more of them, and it is thanks to these sources that companies can monitor their systems and make informed business decisions. Not all companies have their own team of IT experts and systems for storing and analyzing data.

This is why we have created our **WeAre Observability Hub** service, which serves companies at various stages of their observability journey.

Conclusion: Log Management and Archiving as a Strategic Foundation

Modern **log management** is no longer only about collecting system data. For regulated industries, it is a critical part of compliance, security, and long-term operational resilience. As this case shows, achieving regulatory requirements for log archiving requires more than storing data. It demands a structured architecture that balances cost efficiency, scalability, immutable retention, and operational visibility.

By combining **Splunk log management for real-time analytics with Azure-based long-term archival storage**, we helped our client build a compliant and future-ready logging platform. This approach ensures that operational teams maintain visibility into live systems, while regulatory audit logs are securely stored in a tamper-proof environment for more than a decade.

The result is not only regulatory compliance. Our client gets improved security posture, centralized log governance, and a scalable log management architecture that supports hybrid and cloud environments.

This is the core principle behind our **Observability Hub: Log Management & Archive solution** package, where we help organizations design and implement:

- Centralized log management across hybrid infrastructure
- Cost-efficient long-term log archiving with immutable storage
- Structured log pipelines for compliance and audit requirements
- Scalable architectures that prevent excessive Splunk storage costs
- Visibility that supports both operations and regulatory reporting

If your organization is facing increasing data volumes, long retention requirements, or regulatory deadlines, now is the time to evaluate your log management strategy.

About WeAre Solutions Oy

WeAre Solutions Oy is a Finnish **observability-focused consultancy and reseller**. WeAre helps organizations monitor, understand, and optimize their digital systems through observability consulting and implementation. With **60+ professionals** and **€7.9 million in revenue in 2024**, WeAre's mission is to make complex systems **transparent and performant**, ensuring reliability, scalability, and operational insight across infrastructure and applications.

Kimmo Kärkkäinen

Business Development Manager

kimmo.karkkainen@weare.fi

Rami Rantala

CEO

+358 40 027 0781

rami.rantala@weare.fi

Juha Ahlgren

Business owner, Splunk services

+358 44 504 4828

juha.ahlgren@weare.fi