influxdata®

AN INFLUXDATA CASE STUDY

How MuleSoft Uses InfluxData to Achieve Better Visibility and Business Impact

Gagan Kanwar

Director of Partnerships,
MuleSoft

MuleSoft®

NOVEMBER 2018

Company in brief

MuleSoft, a Salesforce company, provides a market-leading integration platform and API management solution that works with over 1,500 large enterprises globally.

MuleSoft's Anypoint Platform™ is the world's leading integration platform for SOA, SaaS, and APIs. MuleSoft provides exceptional business agility to companies by connecting applications, data, and devices, both on-premises and in the cloud with an API-led approach. By leveraging Anypoint Platform™, companies can re-architect their SOA infrastructure from legacy systems, proprietary platforms, and custom integration code to create business agility. They can migrate technology infrastructure to the public or private cloud and prioritize adoption of SaaS applications and other cloud technologies.

Mule ESB, CloudHub iPaaS, API Manager, and hundreds of SaaS and on-premise connectors and templates are the building blocks of the only unified connectivity platform for integration that allows companies to deploy their integrations in a hybrid environment, connecting to both SaaS applications and on-premise systems seamlessly.

Case overview

MuleSoft wanted to build a platform that would provide its customers — large enterprises and organizations often with stringent security requirements — with deep visibility into APIs and integrations. On average, a business supporting digital transactions now crosses 35 back-end systems, and legacy tools haven't been able to keep up. The growing complexity of enterprise systems poses observability challenges as well as business system performance risks.

MuleSoft uses InfluxDB Cloud to power its monitoring and diagnostic solutions as well as provide end-to-end actionable visibility to APIs and integrations to help customers identify and resolve issues quickly.

MuleSoft built Anypoint Platform™ using InfluxData as its core data infrastructure. The business planning between InfluxData and MuleSoft teams involved not only the technology and visibility that InfluxDB can provide but also the use of InfluxData technology to create business impact for MuleSoft.

MuleSoft's Anypoint Platform™ powered by InfluxData

"

We didn't have a great way to figure out and quickly diagnose what endpoint was driving that problem... That's how we started thinking about monitoring APIs. And that's what led us down this journey with InfluxData.”

Gagan Kanwar, Director Of Partnerships

The business problem

In increasingly complex and dynamic enterprise environments, integration is a challenge both to achieve and to monitor. MuleSoft was successfully delivering connectivity to its customers but needed to provide visibility, alerting and efficient resolution of issues detected. The monitoring platform MuleSoft sought to develop was based on the nature of its solutions and the size of its customers.

MuleSoft's solutions span three areas:

- **API's** – Provides a full-scale API management system
- **Orchestration** – Enables customers to connect systems and have different processes, do automated things
- **Connectivity** – Delivers out-of-the-box connectivity to systems ranging from SaaS systems to mainframes to modern systems within AWS

APIs
Proxy and Manage APIs
Design & Build APIs
Engage app developers
Enable secure access

Orchestration
Map and transform data
Route messages & events
Real-time & batch
Infuse business logic

Connectivity
Connect to legacy
Connect to SaaS
Connect to SOAP & REST
Build your own w/ SDKs

MuleSoft's customers comprise very large organizations such as banks, hospitals, government departments, universities, and manufacturing companies. That scale of customers using MuleSoft to create application networks resulted in a very high volume of API calls between systems.

Trusted by 1,500+ Customers Across Industries

Using APIs to drive simplicity & consistency

The typical enterprise has numerous systems, and connecting them involves considerable complexity. As large companies think about how to compete in the new digital landscape, they need connectivity to remain competitive and deliver that user experience. MuleSoft strives to achieve that for them. Below is a simple example of a flights API. MuleSoft has its own API modeling language: RESTFul API modeling language (RAML).

4

How One Service Makes a Request to Get Flights Using RESTful APIs
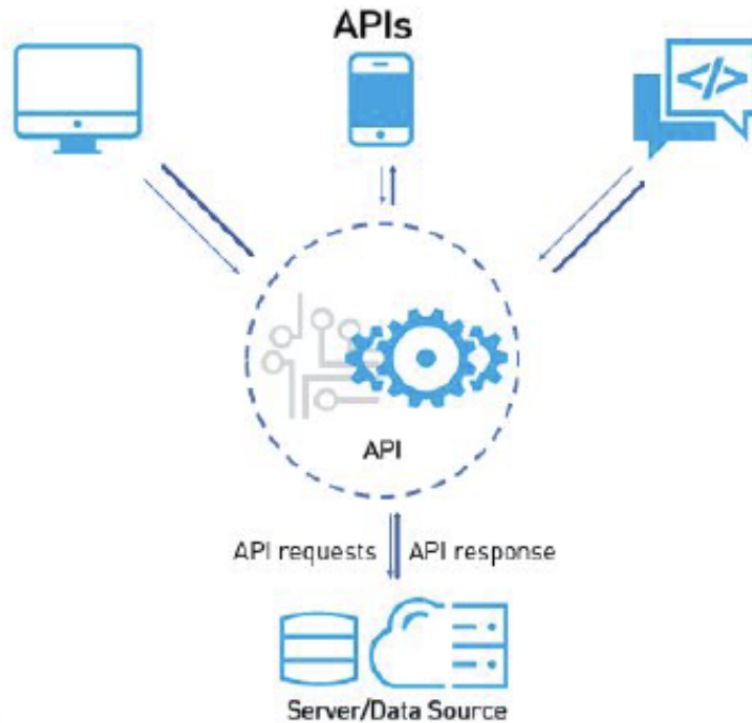
The API response depends on how the API spec is defined, which has many implications in terms of data storage and cardinality.

The API Response

MuleSoft's approach to the connectivity problem

MuleSoft realized that customers were building point-to-point connectors between their systems such their website and SAP or NetSuite (to connect their backend inventory management system to their website, for example, in order to ensure that an item being sold is actually in stock). Over time, point-to-point connectors would make the organization and code base very brittle because they were hard to maintain. That slowed the organization down when they tried to move forward and deliver it at a faster pace. This led MuleSoft to develop its three-tiered "API-led connectivity approach," which is critical to what it does and how API management and API monitoring play into that. It comprises:

1. **System APIs** - These APIs model the data in core systems of record (such as CRM system, order management system, supply chain management system). System APIs sit at the layer (shown below) that provides a RESTful interface to access the data from the core system of record that the enterprise is using for that specific function. Making it a RESTful interface enables any other developer in the organization to quickly access that data without having to worry about learning what SAP looks like or learning what Salesforce looks like.

2. **Process APIs** - Process APIs combine things together (for example, customer data sitting in Salesforce but also in SAP). So for the developer who's building that mobile app, who needs information about both what the customer has bought and maybe where they live, that customer API is a Process API that provides access to that customer object in a more logical fashion, so you don't have to worry about the underlying implementation of how that data is stored in those systems. You can also have an order status API and inventory management API.

3. **Experience APIs** - These APIs enable a developer who is building new experiences (mobile app, website, some IoT type of interface or vending machine) to quickly plug into those already developed APIs. The value created is decoupling these systems in a way that drives up resilience for the organization and developer productivity, because developers don't need to worry about learning SAP or getting an SAP expert. They can use that RESTful API interface to build whatever they need.

The API-Led Connectivity Approach
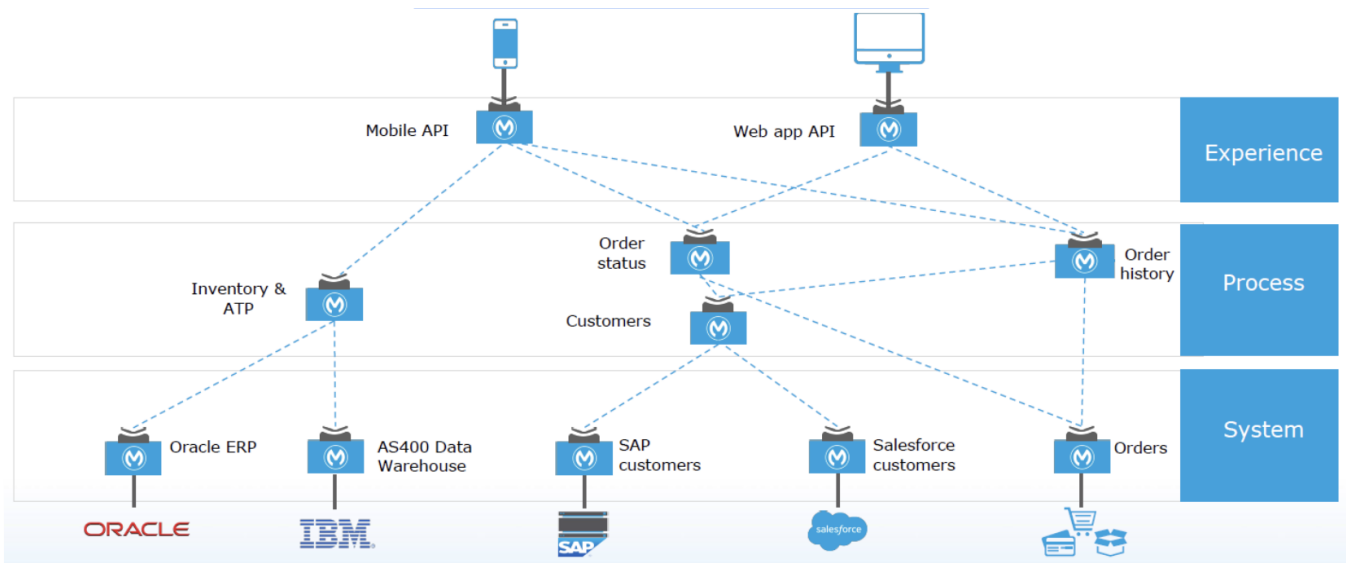
MuleSoft also has an exchange: a central hub for saving, discovering, and reusing APIs and integration assets that makes collaboration easy, boosts productivity, and make it easier for companies to build faster.

Though the large number of APIs and services talking to each other achieves decoupling and developer efficiency, one of the unforeseen consequences is loss of observability. Knowing where an actual issue (on what layer) is happening became difficult to diagnose for MuleSoft and its customers, which led to unforeseen downtime. Customers would often tell MuleSoft: "Your APIs aren't working."

The enormous complexity of enterprise systems

A single web or mobile transaction crosses an average of 35 different systems or components. Complexity introduces:

- Gaps in observability
- Increased time to identification
- Increased time to resolution
- Poor customer experience
- Lost revenue, etc.

The solution driver for MuleSoft was their need to monitor the application network, given the size and requirements of their customers' data systems:

- **Very large customers** with very stringent SLAs (99.999% uptime)
- **High product usage & data volumes** requiring a scalable adaptive system
- **High & unpredictable cardinality** resulting from various APIs with values that greatly varied
- **Complex heterogeneous environments** that MuleSoft needed to support (public cloud, private cloud & GovCloud; hybrid; fully on-prem for customers with very high regulatory requirements)

The below typical data volumes per single "largi-sh" MuleSoft customer demonstrate why scalability of the needed solution was extremely important:

- 2017 - Total # of API calls: 9 billion
- Average # transactions/second: 275
- Year-over-year transactions growth rate: ~ 30%
- Tracking to 12 billion transactions (2018)

8

The technical problem

To build their product, MuleSoft sought a data infrastructure product/partner that could meet the following technical requirements:

1. **Natively manage time series data** – With enterprise monitoring as their core (and highly temporal) use case, MuleSoft needed to know what happened at a given point in time to a given machine or API, to learn the relationship between those components in terms of time, and establish causality.
2. **Deal with very high cardinality data sets** – API calls have well-defined formats, but some fields (e.g. IP) can very quickly drive up cardinality. If your API call has a field for IP address, and that IP address is dynamically assigned because it's sitting in a virtual machine or container, that address will constantly change, resulting in high cardinality for your data set.
3. **Support cloud, hybrid and fully on-prem environments** – MuleSoft's partner would have to support their core use case across all environments, and especially on-prem, since their biggest customers are on-prem.
4. **Deploy in both multi-tenant & single tenant** – Business model optionality was also a requirement for MuleSoft, providing customers the option of multi-tenant or single-tenant scenario.
5. **Very specific per-metric price range** – As a public company at the time (prior to being acquired by Salesforce), MuleSoft needed to tightly manage cost of goods sold (COGS). When you consider alternatives, a very specific price range emerges. But when you want to capture 40 metrics instead of two, those solutions can become prohibitively expensive. MuleSoft had to consider the cost basis, because the cost of goods sold for a software company has to be managed very carefully.
6. **Host the solution in MuleSoft's own AWS instance** – Many MuleSoft customers have very stringent requirements around data residency and access controls, so MuleSoft had no choice but to host the solution in their own AWS instance.

The solution

"

Our customers didn't want us to put agents within our solution, and so we didn't have a choice there. We had to build our own."

Why InfluxDB Cloud?

Below were MuleSoft's decision parameters and process for leveraging InfluxDB Cloud to build their integration platform:

1. **Business case** (defining business objectives, whether embedding InfluxData technology will save or make them money, and considering the target market and customer demand)
2. **Product** (determining whether the product does what MuleSoft needs and whether it will fit in with their broader infrastructure)
3. **Roadmap** (evaluating the product in terms of thought leadership and whether InfluxData is heading in the same direction; for example, able to support large government agencies who have very stringent requirements such as hosting their applications and databases only in a GovCloud certified data center)
4. **Proof of concept to validate scalability** (addressing scale & latency needs and building internal stakeholder alignment)
5. **Competitive evaluation** (considering competitors such as SignalFX, DataDog, CloudWatch)
6. **Unit economics & deal structure** (meeting COGS requirement, economic alignment between MuleSoft and InfluxData, creative thinking to embed InfluxData product into MuleSoft platform)
7. **The team** (examining the InfluxData team's technical chops, ongoing product support, and executive alignment between leadership of both companies)
8. **Customer references** (finding out, based on customer references, whether InfluxData can deliver, whether its service is getting better, and what the product downtime is)
9. **Legal (planning to OEM**, not just license, InfluxData's technology involved extra legal & commercial complexity)
10. **Confidence** (making the decision to strike a strategic deal that seemed to be high-risk but high- reward)

Proceeding from the above, MuleSoft picked InfluxData, in a partnership they considered strategic, high-risk but high-reward.

MuleSoft's goal was to give customers deep visibility into the data that's relevant to them via APIs, compared to what a traditional monitoring system would be able to do just out of the box, especially since managing and monitoring APIs has different requirements than managing a typical application. Yet MuleSoft didn't want to put a third-party agent into their systems because that creates security issues and adds risk in terms of their infrastructure. So they had a business and technical objective to build their own product: Anypoint Platform™.

Anypoint Screenshot
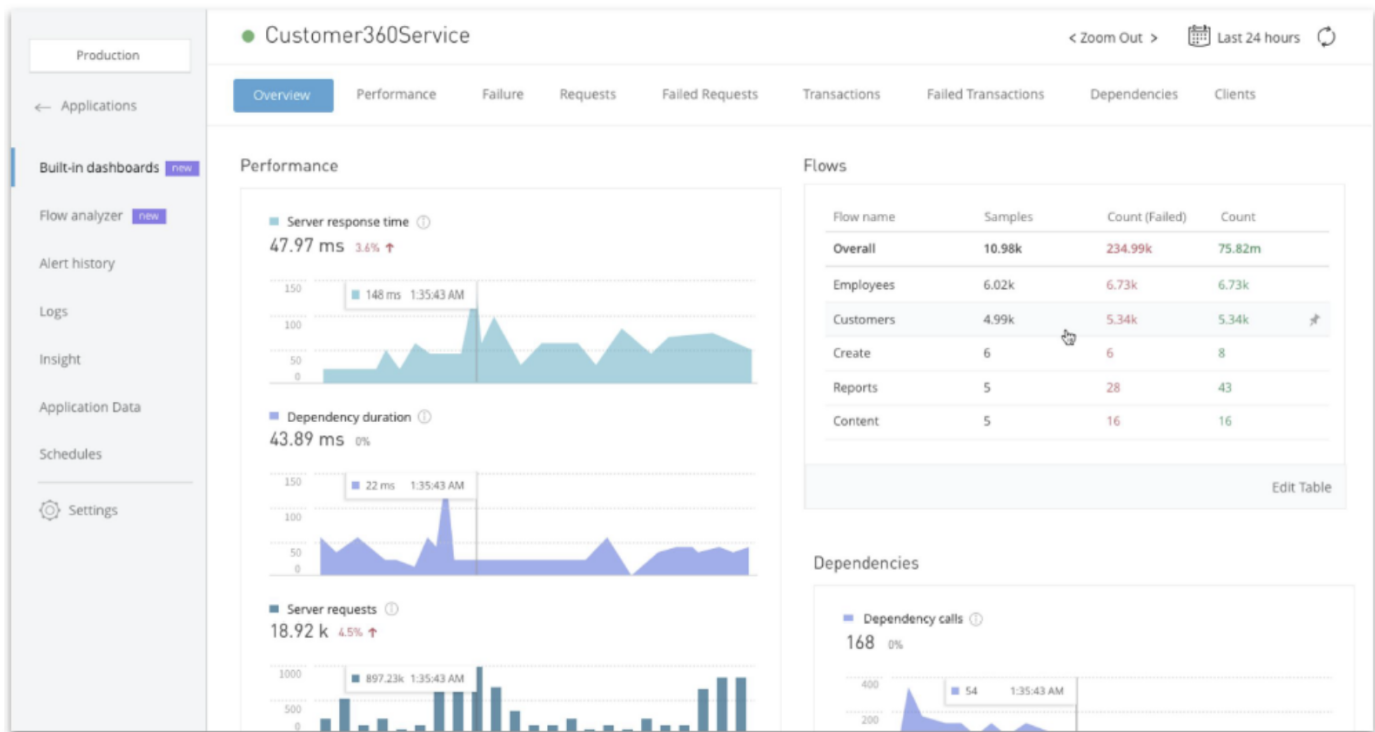
In corporate environments where an application network spans hundreds if not thousands of applications and services talking to each other, Anypoint Platform™ provides end-to-end actionable visibility for APIs and for any integrations that MuleSoft provides as part of that application network:

- 360-degree functional APM (Application performance monitoring)
- Two product flavors — Included (multi-tenant) and premium (single-tenant that offers features around longer data retention and no limits on data)
- Log management
- Integrated visibility — Ability to view logs and metrics together, and switch views between them
- Advanced alerting — Ability to know when certain thresholds have been exceeded or when certain components aren't responding

MuleSoft built Anypoint Platform™ on a dedicated instance of InfluxDB Cloud, which sits in their AWS instance. They use a multi-tenant version of InfluxDB Cloud for the multi-tenant free version of Anypoint Platform™ and a single-tenant version of InfluxDB Cloud for the single-tenant version.

Anypoint Platform™ helps companies connect applications, data, and devices to make them pluggable and reusable. It's no longer about creating rigid point-to-point connections, but about building an enduring application network that allows reuse and self-service — a network that bends, not breaks, as your business grows. Anypoint is the only platform that combines full API lifecycle development and management with enterprise-grade connectivity. It allows you to design APIs for your services and validate them before you even build them:

- Build, edit, test, and debug integrations graphically in an intuitive drag-and-drop environment.
- Easily connect to any data source from mainframes and databases to ERPs and SAS apps, mobile apps, and IoT devices.
- Head to the exchange and build your services with a pre-built connector, or compose an integration flow with a pre-built template.
- Visually map your data with real-time data previews.
- Securely manage APIs and operate services in private or public clouds; the runtime engine combines real-time application integration and orchestration.

Integrations can be deployed on-premise, in the cloud, or both — in real-time or in bulk on schedule — and managed from a single web interface.
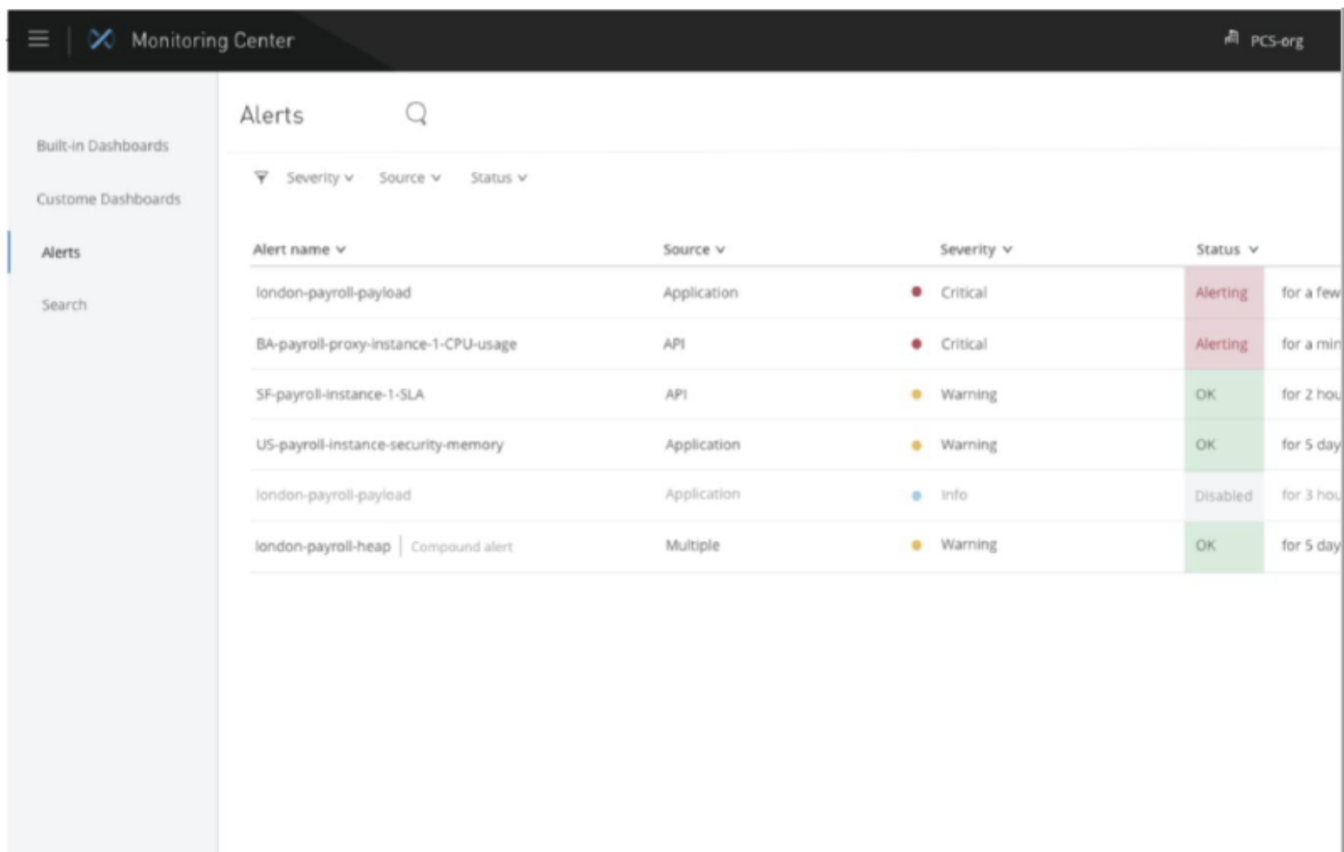
Results

"

You can track that transaction through all the different APIs and systems that it touches, and then be able to see exactly where the slowdowns or latency might be occurring."

Anypoint Platform™ is used to build a structured application network that connects applications, data, and devices with reusable APIs. Anypoint monitoring is being used by MuleSoft customers in various scenarios to achieve visibility, analysis and security. Three of these scenarios are outlined below.

1 - Reducing MTTR

Enterprise alerting – track trends and create alerts to identify that abnormal behavior and problems in your application network. For a sysadmin or someone who's monitoring this application network, it now makes it a lot easier for them to do useful things with that data and quickly diagnose what's going on. Customers want to reduce MMTR (meantime to resolution) and solve problems more quickly.

Sample Alerts Set Up Showing Source, Severity and Status

Advanced data analysis – Leverage templated dashboards and templates or customize your own to provide instant visibility into your applications' performance and detailed analysis of your network. You can drag-and-drop into screens such as the below and also monitor very specific scenarios.
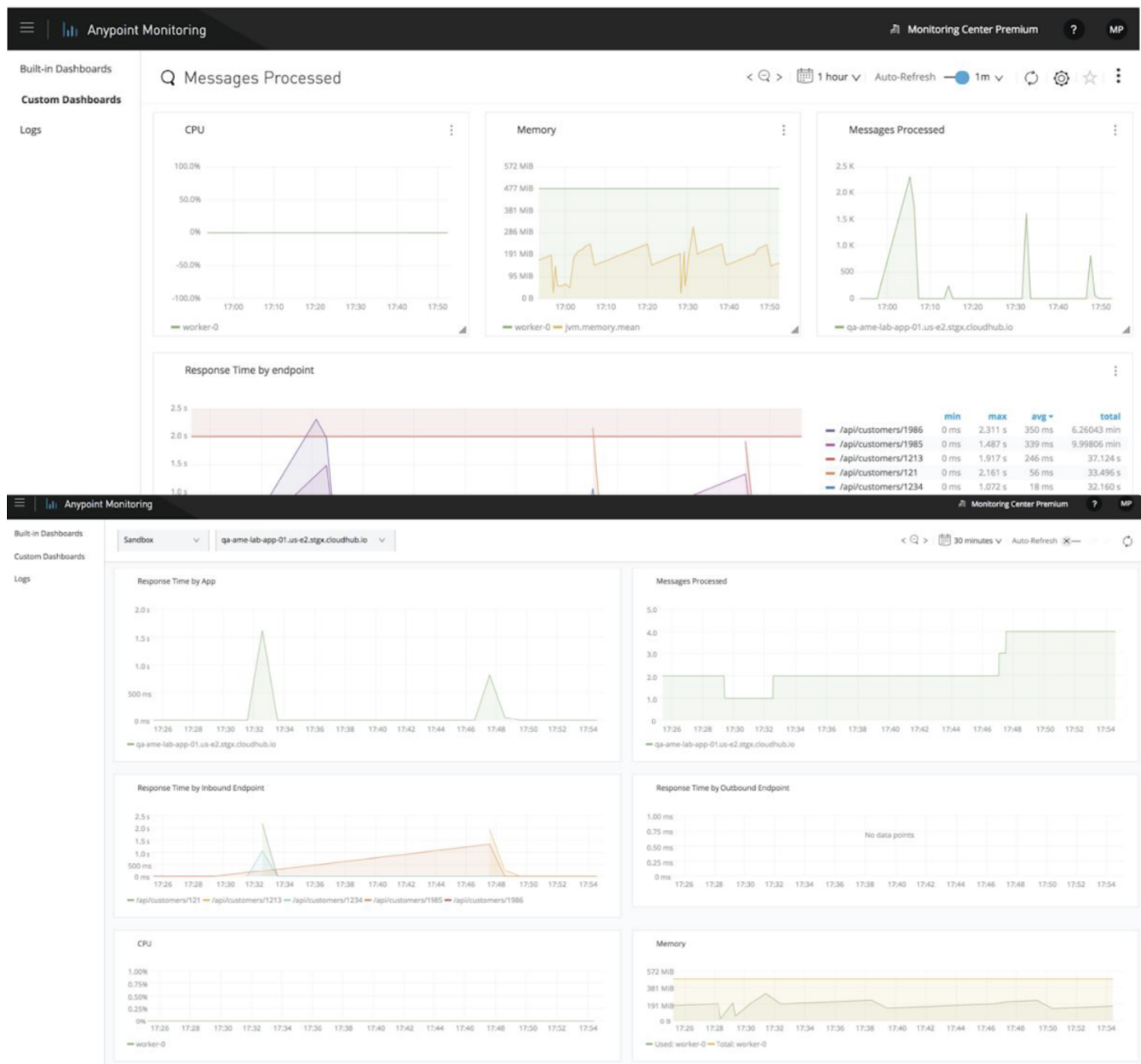
End-to-end transaction tracing — Identify and document the health of every consecutive component within an application in your network; diagnose and prescribe solutions to broken components with ease.
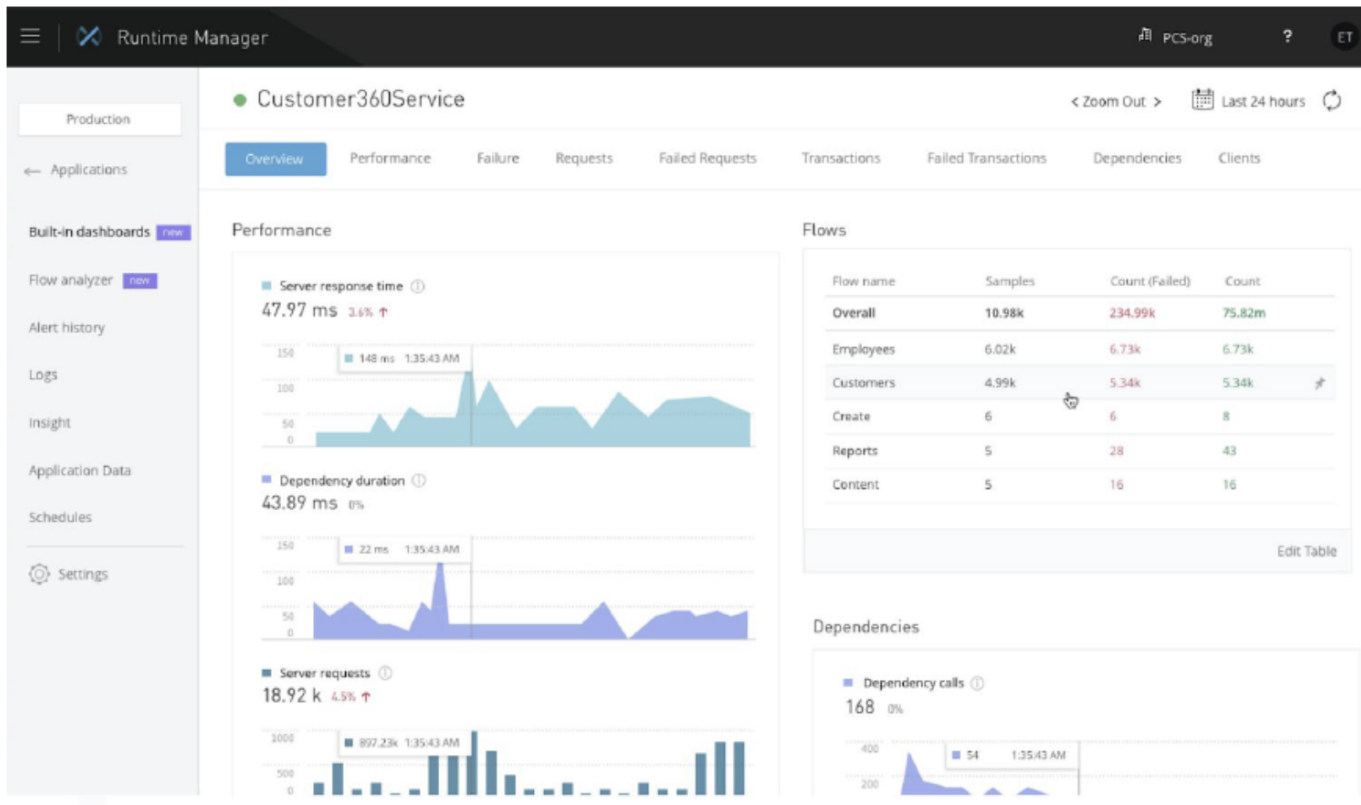
Distributed log management and search – Search raw log and event data from across the application network to pinpoint the root cause of a problem. This was key to MuleSoft customers, since they often use their own tools for monitoring (such as New Relic) and logging (such as Splunk). Yet those tools didn't necessarily work within MuleSoft's environment, and customers didn't want MuleSoft to put agents within their solution. MuleSoft had to build their own and realized that they can have one product that performs both application performance monitoring and log management.
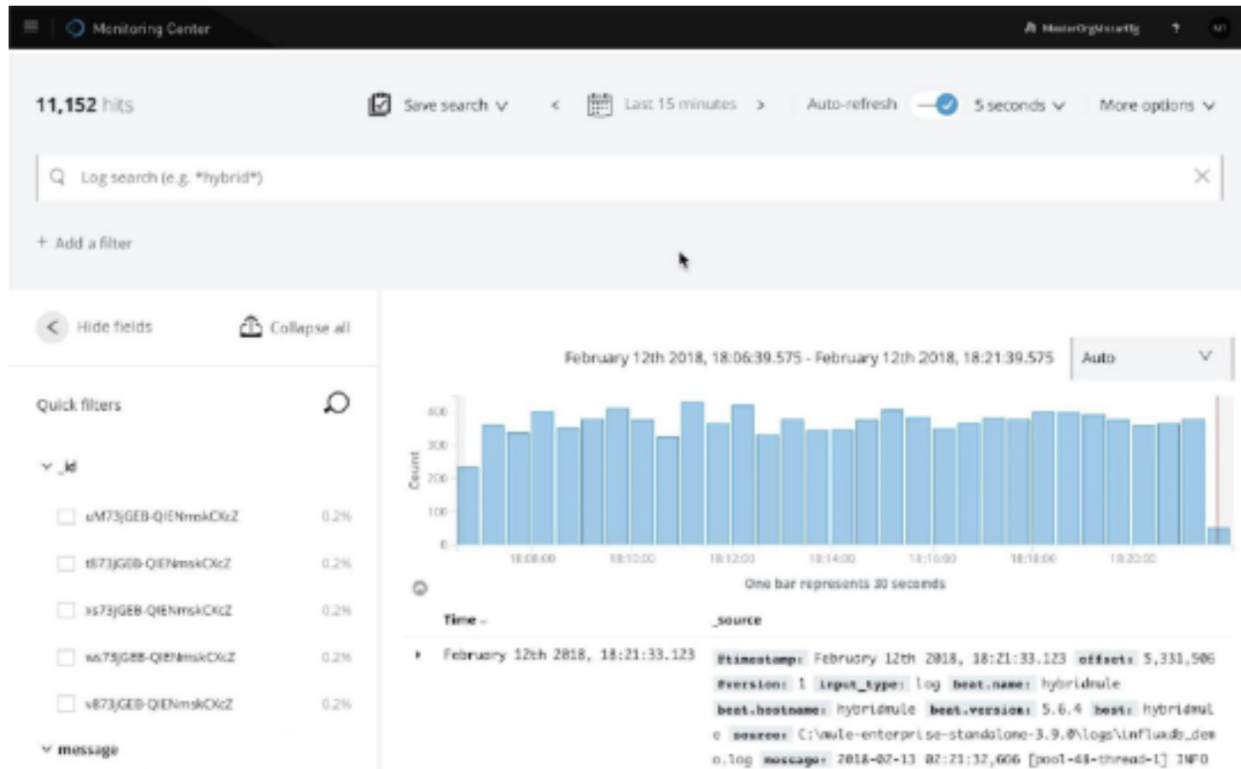
2 - Protecting PII and sensitive log data

MuleSoft has customers that have personally identifiable information (PII) — private, confidential data. Mulesoft needed to be able to tokenize credit card numbers before they go into log files, so that the log files wouldn't be subject to compliance and related issues. Log tokenization:

- Brings a bank-grade level of security to logs
- Supplements encryption in transit and encryption at rest
- Removes sensitive log data from the scope of compliance
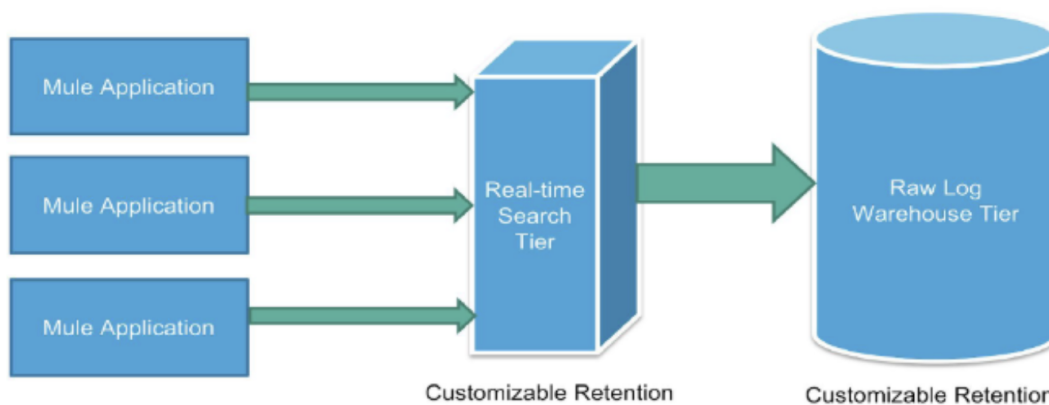
Format Preserving Tokenization

Application data validation logic work "as is"
No downstream application changes needed
Compliance scope reduction (PCI, HIPAA, GDPR)

Log Tokenization

Credit Card #
4111-1111-1111-1111

Token Value
3594-6249-5432-1111

Preserve – Last 4 of CC #

3 - Warehousing logs for auditing, compliance and security

Log data is valuable yet prohibitively costly to retain in traditional systems. MuleSoft's two-tier storage data architecture enables unique flexibility since it allows setting retention for real-time search and raw data storage tiers independently. This architecture helps address both use cases where customers want to keep the data for a long time but also want to ensure that they can access what they need — such as data from a recent timeframe — much faster.

Anypoint Platform's Two-Tier Storage Data Architecture

Mule Application

Mule Application

Mule Application

Real-time Search Tier

Raw Log Warehouse Tier

Customizable Retention

Customizable Retention

Anypoint Platform™ today is one of the world's leading platforms for building application networks that connect enterprise apps, data and devices, across any cloud and on-premise. It enables designing, building, and managing APIs and integrations in a single product. Anypoint Platform™ provides a flexible, unified software platform that enables organizations to easily build application networks using APIs — the digital glue that allows applications to talk to each other and exchange data. Powered by InfluxDB Cloud, Anypoint Platform™ is fulfilling MuleSoft's promise of "Connect anything. Change everything."

About InfluxData

InfluxData is the creator of InfluxDB, the leading time series platform. We empower developers and organizations, such as Cisco, IBM, Lego, Siemens, and Tesla, to build transformative IoT, analytics and monitoring applications. Our technology is purpose-built to handle the massive volumes of time-stamped data produced by sensors, applications and computer infrastructure. Easy to start and scale, InfluxDB gives developers time to focus on the features and functionalities that give their apps a competitive edge. InfluxData is headquartered in San Francisco, with a workforce distributed throughout the U.S. and across Europe. For more information, visit influxdata.com and follow us @InfluxDB.

![influxdata logo]

Try InfluxDB

Get InfluxDB

Contact us for a personalized demo influxdata.com/get-influxdb/