influxdata®

AN INFLUXDATA CASE STUDY

How a German Print Manufacturer Works with Consultants to Use Untapped Data to Improve Operations

Bastian Mauser

IT Consultant, NETZConsuit

SEPTEMBER 2019 (REVISION 1)

Company in brief

NETZConsult, a company based in Germany, helps firms improve processes and aid in digital transformations. As external consultants, they promote the right tool, regardless of vendor. Their consultants work with clients to enhance innovative processes while accommodating customers' specific internal policies. They embrace clients' employees input as it helps with adoption of changes. NETZConsult's goals include saving their customers money, enabling them to escape vendor lock-ins, and improving operations.

One of NETZConsult's customers is the German manufacturing company Fr. Ant. Niedermayr. Based in Regensburg, Bavaria, Germany, it was founded in 1801 and has been family-owned for six generations. It started out as a graphic arts institute and has evolved into "a media services company focusing on industrial web offset printing". To survive since the 19th century, they've had to embrace new technologies. These technological advances led to Fr. Ant. Niedermayr having "the world's first 120-page web offset print".

Case overview

Fr. Ant. Niedermayr needed to gain better insights into their operations. They desired to learn more about individual machines in their factory and to understand the correlation between different equipment and materials. The German company was faced with equipment that had valuable untapped data; they knew Excel spreadsheets and a Postgres database weren't sufficient to answer their questions about their own business. They turned to NETZConsult to address their requirements. As external consultants, Fr. Ant. Niedermayr knew NETZConsult would implement the correct tools given their challenges, regardless of vendor.

The business challenge

The challenges Fr. Ant. Niedermayr needed to address aren't entirely specific to the printing industry. Manufacturing organizations are going through a digital revolution that is known as Industry 4.0. Companies who are keen on Industry 4.0 have to understand that it isn't an overnight process and can be quite costly. In addition to every IoT sensor on a manufacturing floor constantly generating data, organizations are adding even more sensors and external data sources to the mix. Therefore, their infrastructure has to be able to handle the increase in data.

As Bastian Mäuser from NETZConsult points out, businesses often rely on assumptions, rather than actual figures. As he further explains, Fr. Ant. Niedermayr had a "black box" situation where machines are producing and working together, but there's no data on their efficiency. Their customers had

questions about their operations that Fr. Ant. Niedermayr couldn't answer, and they turned to NETZConsult to help address their concerns.

Fr. Ant. Niedermayr knew they used approximately 100,000 metric tons of paper per year. As the offset printing process includes waste, it was in the company's own interest to better analyze the waste.

- Location of waste
- Correlation between the type of paper used and type of ink used
- Causes of waste
- Other environmental considerations
- Impact on business

What is offset printing?

Fr. Ant. Niedermayr has "the largest, highest, and most variable single-web offset presses". Their machines can print over six million pages per hour. In print manufacturing, there are two different styles: offset printing and digital printing. Digital printing is considered better for short run print where a single copy might be needed. Offset printing is ideal for large-scale print jobs — for example, when 2,000 identical copies are needed.

Offset printing can handle larger paper sizes, and color interpretation can differ from digital printing. As offset printers use Pantone colors, it is more precise and useful for imagery. Color consistency can be a key decision factor for customers. Offset printing has higher setup costs and is more time-consuming. Web offset printing is when the paper is fed into the machine as a continuous roll of paper. The roll is referred to as a web and it is cut into sheets and then cut to size.

The technical problem

> *"They have to order stuff from different suppliers...They all have their own subsystems and their own development philosophies."*
>
> ***Bastian Mäuser**, IT Consultant*

Fr. Ant. Niedermayr equipment includes one Manroland Lithoman IV, two Lithoman S and a few other smaller printing machines. The Lithoman IV is an 80-page machine and the Lithoman S machines are 96-page offset printing machines. The Lithoman S can output about 4.8 million A4 (8.5"x11") pages per hour per unit. Their entire setup often runs 24/7; however, there are maintenance windows on the weekends.

The printing industry is notorious for disparate data sources and excessive protocols. To ensure the highest level of efficiency, with the least amount of downtime, organizations often choose specific tools and machines regardless of suppliers. As they have to work with multiple vendors, there's an increase in complexity. Fr. Ant. Niedermayr has to work with approximately 15 different suppliers who all provide different components for their operations. Some of their machines include:

- Splicer
- 4x Inking
- Dryer
- Remoistener
- Web Cutter
- Folder
- Stitcher
- Conveyor
- Trimmer
- Stacker
- Strapper
- Palettizer
- Foliator

Prior to implementing InfluxData's TICK Stack, their technical director had to compile data from numerous reporting systems that came with the units. All the reports came in spreadsheet format, which Mäuser described as "Excel spreadsheet of hell". They had to aggregate the data into one report, which was often full of errors and would take weeks to collate.

As it took time to compile the data, it was often outdated. A report that states how the manufacturing plant was doing six months ago isn't valuable. Reporting was only available after the job was done. Jobs run between one to three days, and it became clear that they needed to detect errors when they happened, not after the job was completed. A plethora of paper and ink were wasted as a result of errors.

Fr. Ant. Niedermayr knew Excel spreadsheets weren't going to work moving forward. They were using RRD Collectors to consolidate statistical data, but it wasn't scalable because it isn't flexible given the

types and amounts of data. Their primary plant control system is called Pecom PMI. It was a Postgres relational database.

They decided against the ELK Stack as it's better suited for alphanumeric information. Fr. Ant. Niedermayr's data set was predominantly numbers, and therefore, Elasticsearch wasn't a good fit. They knew they wanted to use Grafana for visualization. It came down to whether to use Grafana with Graphite, or Grafana with InfluxDB.

During the selection process, they confirmed InfluxDB could handle any amount of data their use case could throw at it. They determined InfluxDB's TICK Stack could scale well with high ingest rates. It could handle more than 500,000 data points per second on a single instance. Currently, Fr. Ant. Niedermayr collects 800 data points per machine per second. The TICK Stack's storage engine was appealing as its speed, ingest rates, compression rates and reclaim and retention policies met their current needs and future requirements.

They appreciated that if the data is deleted because of the set retention policy, it will delete the file from the hard disk and the space will be available again. In contrast, in a SQL database when records are deleted, the space isn't reclaimed. The space can be used for new data; however, reclaiming the disk space won't necessarily work in a typical SQL database.

Over the years, Fr. Ant. Niedermayr has created a blend of machines that works best for them, regardless of vendors. They would rather have the right machine for the job than be vendor-locked with a specific company. Mäuser points out that this is challenging as vendors expect clients to use everything in their ecosystem down to the sub-instrumentation. Vendors don't want to have to work with other companies' devices.

MQTT has become a standardized protocol used by machines to communicate with each other; however, not all organizations have embraced it. On top of that, some components are still running Windows 98 or Windows XP. They were aware the software likely wasn't going to be updated, and determined the best workaround was to see if the vendors could provide an interface to enable them to extract the data. While the companies are using outdated software and hardware, it was still possible to extract untapped data from the machines.

They had to reverse engineer some machines to understand how they worked and determine how to incorporate data from them. There aren't a lot of software developers for the machines; the software can be up to 20 years old. Not much has changed since, and they aren't open to adopting new technologies or approaches.

The InfluxDB platform is built for developers, and they valued the extensive ecosystem of Telegraf plugins. Telegraf is like "a swiss army knife for connecting various MQTT sources and OPC UA sources".

As InfluxDB is purpose-built for metrics and events and was proven to be production-ready, InfluxData was chosen as the time series database provider.

Validating the authenticity of the data collected presented Fr. Ant. Niedermayr with another challenge. There were incidents where the same value was coming from three units and all three values were different. With the help of NETZConsult, they determined the logic behind the discrepancies and which value is appropriate to use. The same data can be collected multiple times with slightly different timestamps. Sometimes the raw data is collected, and sometimes it has been extrapolated and used in another step.

Furthermore, data was coming from multiple sources using a multitude of collection mechanisms making the entire system complicated and error-prone. Data might be pulled from their ERP, the plant control system or a machine or sensor directly. And they could all be using a different set of protocols — from a REST API, SQL connectors or even ODBC — which further contributed to variances in data that could misrepresent what is actually happening in the plant.

Businesses are realizing that rather than collecting data about every aspect, they should only collect useful data. Otherwise, as Mäuser points out, an organization will be "drowning in a sea of data which you may not use or you may not need...you don't need to monitor everything". Identifying data sources that matter is key; otherwise, a company can drown in data. Often the data can be hard to work with.

The solution

Why InfluxDB?

NETZConsult quickly provided Fr. Ant. Niedermayr with a proof of concept, and within a matter of days, with the TICK Stack. Most of this time was spent interpreting and validating the data and looking into the processes used to collect metrics and events. They use two different versions of Chronograf. They use it for data exploration and for creating machine learning models with LoudML. They use Kapacitor for alerting; emails are sent out when key events occur.

From the start, they were smart about selecting what data to collect. Fr. Ant. Niedermayr determined some data wasn't useful as it was hard to work with and decided to avoid wasting time collecting that unnecessary data. Additionally, they pinpointed where they needed more visibility. As an example, Fr. Ant. Niedermayr chose to add IoT sensors to monitor power usage and pressurized air.

The plant uses their data to help benchmark the business, and they were able to create key performance indicators (KPI's). NETZConsult worked with management and employees to determine the data that should be used to create KPIs. They needed to work with everyone to determine the metrics that mattered and the data needed in real time.

As with any organization, having proper security procedures in place is important for the longevity of the business. NETZConsult recommended they implement security measures right away rather than as an afterthought. They used MQTT to connect the myriad of data sources; it's easier to implement authentication requirements, like TLS client certificates, at the beginning of the project. They used separate VLANs to keep proprietary data away from the internet and to provide an additional level of security.

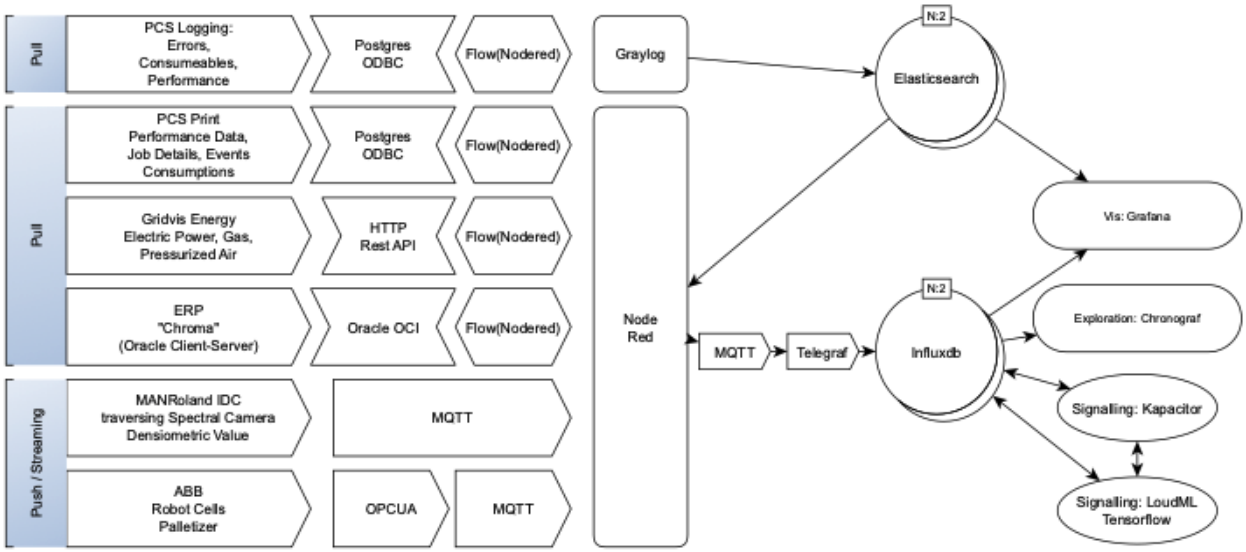
Providing the organization with an interpretation document is important for everyone involved to understand the workflows and the data shown in dashboards. Making sure the team understands where measurements are being calculated is important. The guide should also include the locations of the IoT sensors. If the raw data has been extrapolated for another use, it's important to include information on where the raw data is gathered and the formula used to alter it. NETZConsult helped them create a wiki with all pertinent information. The resource guide provides Fr. Ant. Niedermayr with the necessary structure and standards for collecting data. Rather than having custom Python scripts for every data source, they have been developed in Node-RED, as it's scalable.

Using telemetry data to improve efficiencies

Offset printing machines need to be washed after certain intervals of time and use. Fr. Ant. Niedermayr knew they could save money by only washing when required, based on telemetry data from the printers. Management wanted to take the guesswork out of the washing cycles and reduce the chance of unplanned downtime. Washing cycles are required only when specific parameters are met, like dot gain rate changes. Prior to using telemetry data, washing would only occur every 1-2 reels. However, by better understanding different paper and ink combinations, they realized that for some it's enough to wash every fourth reel. It's important to know that washing involves waste of paper and energy, so it's good for savings and for the environment to stretch those cycles.

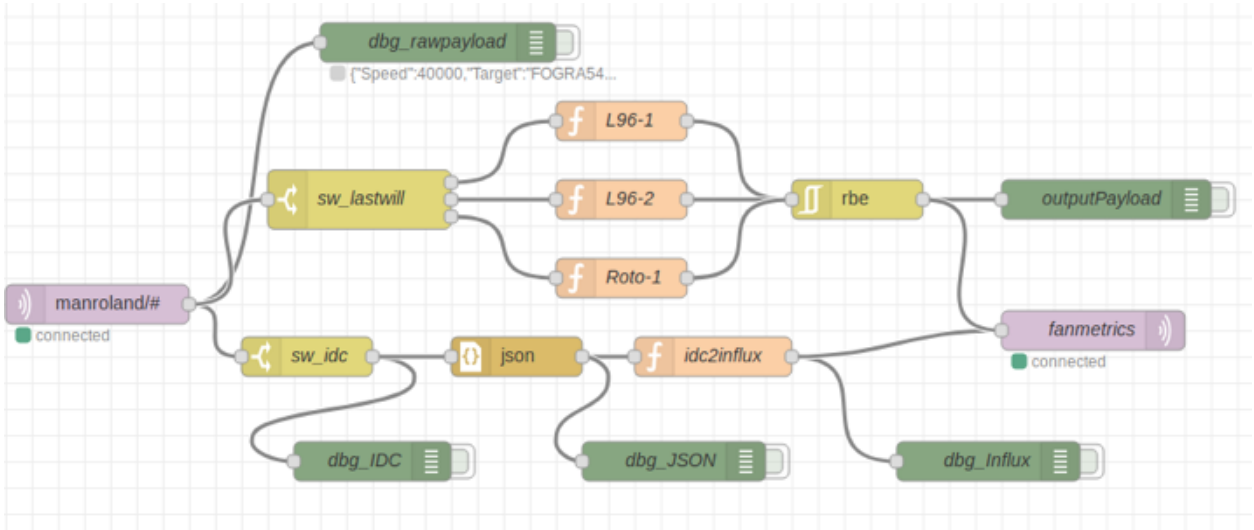
Technical architecture

InfluxDB with Node-RED at Fr. Ant. Niedermayr



- Data is pulled into their systems via a REST API, SQL connectors, or ODBC.
- Their ERP system's data is pulled in through an Oracle connector. Through their MQTT broker, they get data from densitometric cameras every second. They use two ABB robots for their palletizing process which has an OPC UA server.
- The team found a useful tool on GitHub which they modified to suit their needs. They receive data from OPC UA brokers, push it to MQTT, transform it in NodeRED (to match their own internal standard), push it back to MQTT, and use Telegraf to push it to InfluxDB to be stored.

Example of a Node-RED Flow for IDC



On the left side of the chart as shown above, the MQTT comes from the IDC unit from the machine. Next, they get a JSON dataset — all the green nodes are for debugging and turned off. By using roughly 10 lines of JavaScript, they are able to transform the data or extract the data needed from the JSON object. These objects are converted to InfluxDB line protocol, and through a MQTT broker, the data is put back into Telegraf.

Results

> *"Telegraf is like a swiss army knife for connecting various MQTT sources and OPC UA sources."*

After implementing InfluxDB and Grafana into their operations, Fr. Ant. Niedermayr has a better tactical overview of four of their biggest printers. They've improved transparency, as now the data is available to everyone across the organization. Their team now better understands their machines, operations and outcomes.

The customer determined one machine, the dryer, used €300 of gas during a 36-hour period. However, if the dryer is running, the cooling tower is idling. During the same time period, the cooling tower used 300-350 kilowatts (close to €10,000) of energy. If the dryer is shut down for maintenance, it makes fiscal sense to turn off the cooling tower as well. If a machine is shut down, the related machine should be turned off too. They realized that not everything can be predicted and sometimes a machine has to run separately due to maintenance requirements. Previously, Fr. Ant. Niedermayr didn't have the data available to discover these kinds of cost savings.

Idling machines over weekends use more energy than realized. There were machines using 300-400 kilowatts of energy, even when not in use. Over weekends, the dryers were constantly using around 350 kilowatts of energy. As the machines have timers, they can be turned off on Friday, and automatically on early Monday, to allow enough time to heat up prior to being needed. By turning off machines over the weekend, they saved a high six-figure sum of Euros per year.

They have created a dashboard with dot gain data, collected from their densitometric web surveillance System (IDC), which provides the company with information on print quality. This includes data on expected results and actual results. This helps management create accurate quality targets and set realistic goals. The green bars below represent the range they want to be in. The green bar isn't fixed —

it will move based on their tolerance levels, which are based on the standards specific for a job and on the paper and ink used. The data is extracted through MQTT from the plant control system. The data is visible to everyone, so they can react as soon as the values go beyond the non-static green bars.

Print Quality Dot Gain

- **Blue lines:** It is important to understand the rhythm of washing cycles. They can change the frequency depending on the specifics of the job running to improve efficiencies and to reduce waste.
- **Red lines:** They need to understand when there are wet cuts occuring. Understanding how their machines are working with certain papers enables them to work with paper suppliers to find the best mix.
- **Green bars:** These bars display their tolerance level for dot gain and signify when a job is running. Dot gain is a variable in printing that must be controlled as printed materials can come out darker than intended. The green bars are not static; they move according to job specifications. There are density cameras that move during the print job to monitor.

Fr. Ant. Niedermayr now has the ability to drill down into their data and look at a job that ran six months ago. They are able to compare efficiencies based on factors like paper type and ink type. Different combinations of paper and ink can result in different qualities of the product. They correlated the frequencies of washes to the amount of waste. It was determined that they used 179 tons of paper, but only 174 tons were usable.

In the past, they would cope and work past the issues. This would lead to more cuts, more waste and more quality issues. Moving forward, they are able to analyze a paper and ink combination and decide if the extra time and money is worth it. There are aspects of production that can be adjusted to reduce

the likelihood of problems. Fr. Ant. Niedermayr wants to reduce the number of cuts as these costs are directly absorbed by them.

Some of their improvement projects are ongoing as they discover more. They have automated specific areas of their operations. All efforts have resulted in cost savings, reductions in waste, reductions in carbon dioxide, and electricity savings. Fr. Ant. Niedermayr wanted to improve their production, rather than replacing it with different solutions.

The German print manufacturer has escaped their vendor-lock. The vendor-lock made them inflexible. The TICK Stack and Grafana have provided them with visibility into their untapped data. They've improved transparency across the organization and during every stage of the manufacturing process. This has resulted in improved performance and team cohesion.

Where did the ink go?

In 2018, they used about 2,700 metric tons of ink and Fr. Ant. Niedermayr wanted to validate the usage. Since implementing InfluxDB as their time series database, they have been able to better understand when they need to order ink. With a simple algorithm, they are able to determine when the ink tanks will be empty, based on historical data. As customers can request specific Pantone colors, having insight into ink used and needed is vital to the business. They discovered an 800 ton discrepancy in their calculations. Based on data in their ERP system, the printers used 800 tons more ink in 2018.

It turned out that the volumetric counters were incorrectly measuring how much ink was being used. When new, these counters produce 250 ticks per one metric liter of ink, but because these counters are mechanical, they degrade over time. In addition, ink color can vary in viscosity which can lead to miscounting the amount of ink used. Understanding this helped them to see that 800 tonnes of ink weren't really missing and they were able to build a correction key that is used to get the real ink values. They are now able to visualize ink usage through Grafana dashboards. They purposefully use CMYK colors to represent the ink colors. This way, it's intuitive and plant operators can see if everything is running smoothly.

CMYK Ink Usage

What's next?

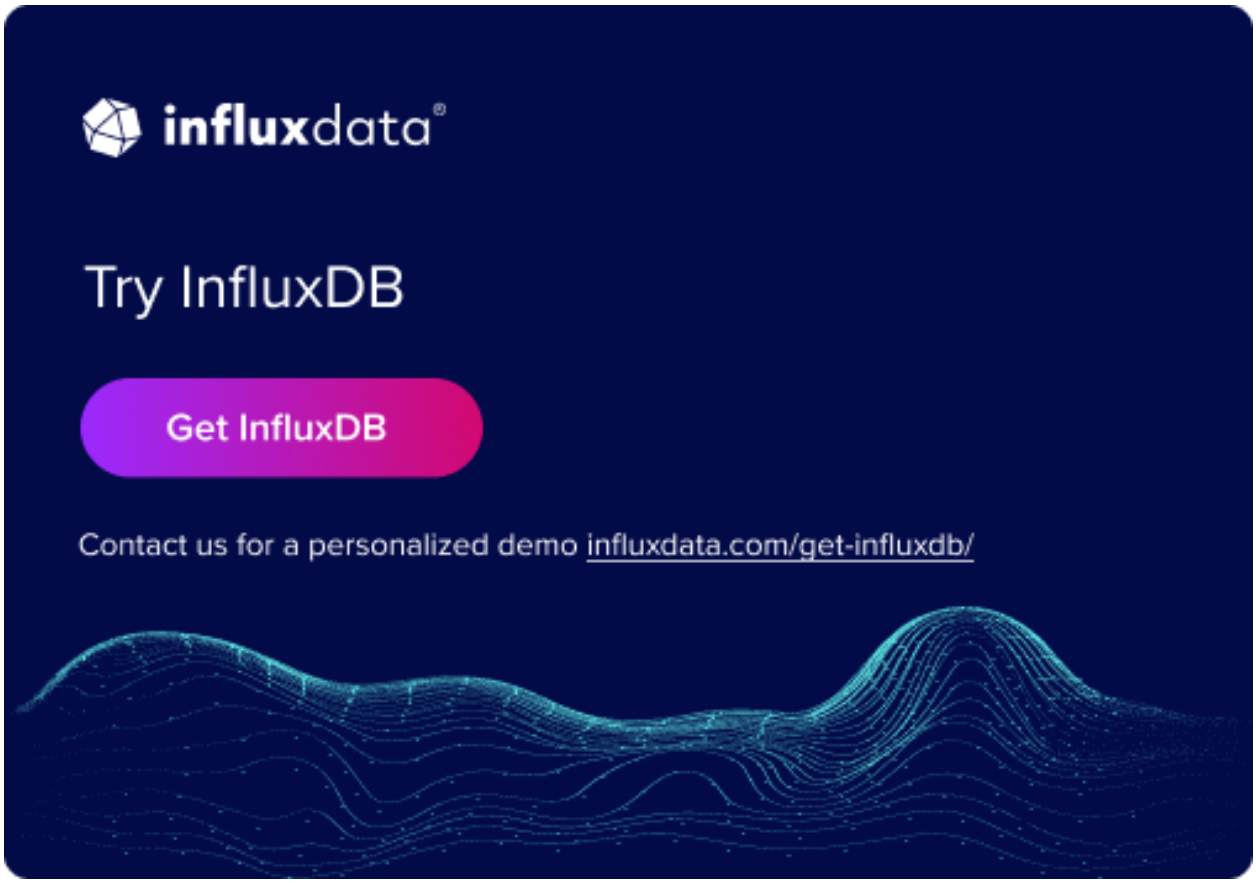
Fr. Ant. Niedermayr. is hoping to create a fluid management system. In addition to ink, there are fluids for remoistening the web when the product is removed from the dryer. Fr. Ant. Niedermayr has an ink supply system in addition to central ink supply systems. They use a lot of ink and they want to know where the ink is going and where it's being used. They want better visibility into miscellaneous fluids and materials needed during the offset print manufacturing process.

At the point of post-processing, the print products are picked from a set of conveyor belts. They gathered a set of metrics from these conveyor belts (torque, revolutions per hour, power amperage or the voltage amperage, temperature, fan speed) and noticed a slight deviation in the amperage from the motor of the conveyor belt. Not convinced that this is a fluke, they have started to put this data in a ML framework and hope to reduce machine downtime by being able to predict when a motor will fail.

InfluxDB and Grafana have created endless opportunities for Fr. Ant. Niedermayr. NETZConsult hopes to improve predictive maintenance for their machines. They are realizing that to achieve their desired outcomes, they will need to purchase specialized hardware. The team is continuously testing ways to improve their operations.

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.



548 Market St. PMB 77953, San Francisco, CA 94104