



AN INFLUXDATA CASE STUDY

How Rise Uses Time Series Data and Synthetic User Monitoring to Improve Web Performance

Matthew Van Tassel

Chief Technology Officer,
Rise (Build with Rise)



DECEMBER 2019 (REVISION 1)

Company in brief

Rise ([Build with Rise](#)) is based in Fredericton, New Brunswick, Canada and helps homeowners build and renovate more sustainably. Rise was founded in 2015 after its CEO renovated a home and realized that despite investing in his home, it wasn't running any more efficiently. Rise's platform connects homeowners with professionals (i.e. architects, contractors, plumbers and electricians) in their area who also have sustainability in mind. Rise's platform also provides customers with rebates, incentives and discounts to help encourage sustainable home construction and renovation.

In 2019, Rise became a certified [B Corp](#). Organizations who are B Corp Certified are legally required to consider the impact they are having on their employees, customers, suppliers, community and the environment. B Corps are striving to globally shift companies' definition of success by solving social and environmental problems. They use their profits to positively impact their employees, communities and the environment. From Rise's inception, CEO Matt Daigle knew becoming a Certified B Corp was going to be [important to Rise's success](#).

Rise became a member of [1% for the Planet](#) in 2019. Members are committed to giving 1% of their gross sales annually to approved nonprofit organizations. A quarter of a company's annual donation can be in the form of employee volunteering or product and service donations. Rise donates 1% of its annual sales to environmental nonprofits and hopes to create positive change through its actions and investments.

Case overview

Organizations are always looking to optimize their online presence with their website. Rise is no different — its team works hard to ensure the company website is working as efficiently as possible for their customers and professionals who use the platform to connect with clients. By improving their development process, they have better insight into what changes caused any website performance degradation. Rise CTO Matthew VanTassel is always encouraging his team to improve their DevOps practices — there is always room for improvement. By reducing website load times and improving development cycles, Rise has significantly improved the customer experience across its entire platform.

The business challenge

Rise's team understands that it can be overwhelming for the average homeowner who wants to build or renovate their home sustainably. Rise is a "web platform that educates and inspires everyone to make better, more sustainable home choices". This includes purchasing better building materials and

choosing better fixtures, like low-flow shower heads. Rise hopes to help homeowners from the raw construction of a home to picking the finishing touches.

Like most organizations, Rise relies on its website to attract customers. Rise's professional clients rely on its platform to connect with homeowners. User analytics are key for Rise to understand how customers are interacting with its website. The rise of data privacy laws influence how organizations can collect user data. As a Canadian company, Rise has to contend with Canada's Anti-Spam Laws (CASL). CASL's laws impact all companies who send electronic messages (email, text) to customers within Canada. Rise's current clientbase is primarily in North America; however, as Rise expands into other markets, other data privacy laws like GDPR will become increasingly important.

The technical challenge

Rise needed to understand its website's performance better. In addition to being a marketing tool, the Rise website is a platform that enables homeowners to learn how to best consider the environment while building or renovating. The website hosts content, photos and articles which all help customers get inspired, learn about the best products, discover ways to save money and find pros that can help. All this content is updated constantly to keep up with the trends, new materials, and new builders which adds to the challenge of keeping the site performant.

Organic website traffic

Rise needs to make sure that their site, BuildWithRise.com, is consistently ranking a high domain authority. Domain Authority (DA) is a score which helps determine how well a website will rank within search engine results. VanTassel's team needs to continuously confirm that their webpages have lots of backlinks, that the backlinks are functioning, as well as maintain a large set of reputable backlinks. The more reputable backlinks that direct to Rise's website will result in a higher search engine rank; they help make websites appear authoritative to the search engines. In addition, since the number of pages and content are always increasing, the task of managing backlinks is never done.

Content

Rise's team needs to ensure they have relevant content that people are looking for. The content is enriched with structured data so that the customer experience is rich and beautiful. Rise's [Lookbook](#) is a content-rich area of their website which provides customers with photos and articles for inspiration. The [Stories](#) section of their website is abundant with articles filled with tips, advice, product guides, lifestyle information, etc. Through Rise's [Rebates](#) page, consumers can find rebates, incentives and

discounts specific to their geographical area. There are incentives for commercial and industrial companies who want to operate with the environment as top of mind.

Website performance

Common metrics used to monitor website performance are First Contentful Paint and First Meaningful Paint. First Meaningful Paint (FMP) is the time it takes for a webpage's primary content to load. It helps determine the user experience (UX) of a webpage. For example, if the image is a webpage's primary purpose, First Meaningful Paint is the time it takes for an image to load — headers, navigation bars, etc. do not count when determining FMP. First Contentful Paint (FCP) is the time it takes for content to load; content can include images, text etc. First Contentful Paint is meant to be more impactful than FMP; however, they might be very similar, especially for simpler websites. As websites become more complex, the two scores will differ more. With more complex websites, First Contentful Paint will become the more useful metric. Similar to the DA issue, keeping the website performant is a never ending task.

Developer workflows

VanTassel needed to determine the current status of his team and document their goals. As a startup, Rise was often more concerned about getting new code released. Previously, Rise's engineers would get work assigned and development would immediately begin. The engineers may or may not create a PRs. Even if they created a PR, it wouldn't necessarily get reviewed. Code changes were deployed to production at the engineer's discretion.

As Rise's CTO Matt VanTassel said, "this was the modus operandi — we just wanted to get things out as fast as possible". Their small team would push out changes as they went. This approach resulted in large performance regressions. Dependencies were added which reduced their website's speed tremendously. There were dependencies that were increasing their bundle size, increased the time to interactive and increased the speed index. VanTassel knew they needed to improve their monitoring of these key metrics to ultimately improve their product offerings and customer experience.

Their developers were able to run audits on specific webpages to understand the impact of changes made. VanTassel's team can understand the impact they are having as individuals. However, VanTassel is also concerned about testing their website as a whole. They need to account for variability across all of their pages. They needed to be able to review all changes in a holistic way that took into account all historical changes. VanTassel realized that they needed to improve their performance monitoring.

Security vulnerabilities

When VanTassel's team did a preliminary scan, they had around 480 security vulnerabilities just within their UI application. They knew the list of security issues would just grow — and were aware that they lacked functional tests. As the team planned on adding new features, they were also aware that they'd introduce new bugs as part of new features and fixes. They were lacking consistent and immediate feedback that would alert them of regressions. To decrease security threats, Rise needed to reduce the number of vulnerabilities and decided to implement continuous integration procedures into its development workflows to help with this.

The solution

Rise's team found various tools to address their needs. They gravitate towards open source solutions. They needed to understand their development processes better to improve them. As their website is the core way that people interact with their platform, they needed to improve their website's performance as much as possible.

Continuous integration

By using npm audit and [Snyk](#) CI checks, they have immediate feedback for all PRs. Npm audits are short for node package manager and are commands that enable moment-in-time security reviews of a project's dependency tree. Npm audits are package managers for JavaScript and Npm audits are the primary package manager for Node.js environments. Audit reports provide Rise's team with information about security vulnerabilities within their dependencies. They also help fix security bugs with easy npm commands and suggestions for further troubleshooting.

Rise uses Snyk to automatically detect potential bugs and vulnerabilities and also to help them expedite fixes through their development process. They have written functional tests and end-to-end tests in Cypress, and the list of tests grows every week. They have Cypress tests running in [CircleCI](#) which have been instrumental for Rise's team. As their CTO says: "We have feedback in near real time for anything we add to our repository". While their new CI methodology doesn't directly improve performance, it has helped by addressing their bundle size with a better CI process.

Performance testing

Rise has automated their performance testing. All PRs created are reviewed, and at least one reviewer has to approve the PR before it can be added into the master. The master is deployed to their staging

environment, which has end-to-end tests running. They receive nightly feedback, and by so doing, are able to determine if their deployments are better than previous ones.

Rise uses Google's open source Lighthouse tool to measure and improve website performance. [Google Lighthouse](#) focuses on five key areas: performance, progressive web apps, accessibility, best practices and search engine optimization (SEO). The solution provides Rise with a score for every section using documented weighted scoring that's based on different metrics. The metrics for performance include first contentful paint, speed index, time to interactive, CPU idle, input delay and First Meaningful Paint.

The team needed to automate their Lighthouse testing, rather than rely on manual processes. For debugging purposes, they needed to maintain the HTML of the Lighthouse reports. If they were to just log the scores into InfluxDB without any contextual information, the score isn't very useful. For any metric pushed to InfluxDB, there is a link to a report which enables the team to dive into the results.

They are able to answer questions including:

- Why is this value really high?
- Why is this value really low?
- What happened on a particular day?

Measuring performance

To measure performance, the Rise team uses Garie web performance software, which provides insights into website page speeds. They love that Garie is built on open source tools. VanTassel explains that it's easy to pull Garie's components apart and update specific pieces as needed. They got Garie running within 15 minutes and have been tweaking it ever since. It's on Docker so it's easy to deploy anywhere; all they had to do was configure and run the container.

A key criteria for Rise's team was to have access to all of the logs for the disparate servers. Rise relies on their logs for deeper analysis, and they need tools that enable engineers to manage and maintain their systems efficiently and quickly. They use Docker Compose to define restart policies. When they face issues with a particular container, it's helpful being able to reset just the specific container.

Visualization

Once their website metrics are stored in InfluxDB, Rise relies on Grafana for visualizations. They are able to quickly glance at graphs and scores to determine the current state of important KPIs. Rise's team is interested in understanding how their website is performing in the last week. This dashboard uses Google Lighthouse's standard colored weighted system to show the median value for the last 7 days. By having the scores at the top of every dashboard, they are able to toggle between different environments to see the scores for that area of their website and the specific environment.

Important metrics:

1. Performance
2. SEO
3. Accessibility
4. Best Practices

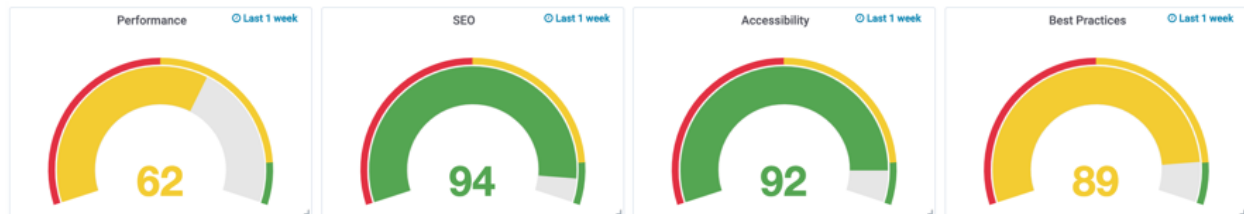


Figure 1: Google Lighthouse dashboard

Query: `SELECT median("value") FROM "performance-score" WHERE ("url" =~ /$env/) AND $timeFilter GROUP BY time (1w)`

Version 1 - Visualizing all Lighthouse data

They started with the default graph found in Grafana. It's very hard to understand this amount of time series data, so they needed to be able to drill down into specific time periods for root cause analysis. The graph below shows that at the beginning of this project, their scores were around 25% (far left) and now they are close to 75% (far right). It's great seeing that there's been overall improvements on performance, but it isn't helpful in determining root cause analysis and ways to fix performance issues.

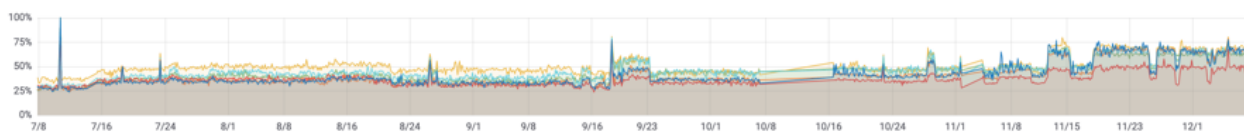


Figure 2: Original Grafana graph

Version 2 - Day over day

By segmenting their data by day, data collection issues were more pronounced. Rise's leadership team can also tell that they're doing much better in December than at the beginning of the year. They were able to tie the October and November issues back to a crash and a power outage in the office. The crash led to their systems not running for a week. November's power outage was a weekend-long outage. VanTassel learned the valuable lesson that in addition to monitoring key systems, it's also important to monitor the monitoring systems.

"We need to set up alerts for when things go wrong... It's one of those things you don't think of until it happens."

Matthew VanTassel, CTO, Rise (Build with Rise)



Figure 3: Day-over-day data buckets

Version 3 - Week over week

As Rise is on a weekly deployment schedule, VanTassel decided to change the daily buckets to weekly buckets. By finding the median score of each week, Rise was able to see that over a six-month period, they got close to doubling their scores in some areas. VanTassel likes to point out all of the hard work they've done and how far they've come so far. Being able to show his team that all of their hard work has paid off has helped encourage his team to continue working hard. As he said "it's really good to have the [historical] context, especially when you're trying to push to the next level".

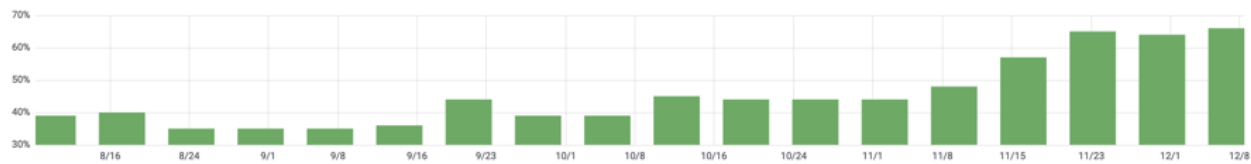


Figure 4: Week-over-week data buckets

Deployment

These changes have culminated in Rise creating a deployment schedule for the entire organization to follow. Thanks to having a schedule, they have a process which includes testing and monitoring. Every week, Rise's team knows there will be an increase in fixes and features that will benefit their customers.

Why InfluxDB?

There are a myriad of reasons why Rise chose InfluxDB as their time series database of choice. InfluxDB's high ranking on DB-Engines.com was a contributing factor. The Rise team values that there is a thriving online community of InfluxDB users and that there were great recommendations online.

VanTassel was very familiar with InfluxDB, [having used it at home](#). InfluxDB's ease-of-use and quick implementation process factored into Rise's decision.

Rise values that InfluxDB's query language is similar to SQL as there are people on Rise's team with PostgreSQL and MySQL experience. They were able to very quickly understand InfluxDB and InfluxQL. Adoption of new technology is often one of the biggest challenges of any tool, and VanTassel's team was able to quickly adapt and start using InfluxDB. By using it habitually, they have been able to maintain and improve it easily and on a regular basis.

"InfluxDB's documentation for all of the client libraries is stellar. This makes a huge difference when you're getting into a new technology."

Matthew VanTassel, CTO, Rise (Build with Rise)

Technical architecture

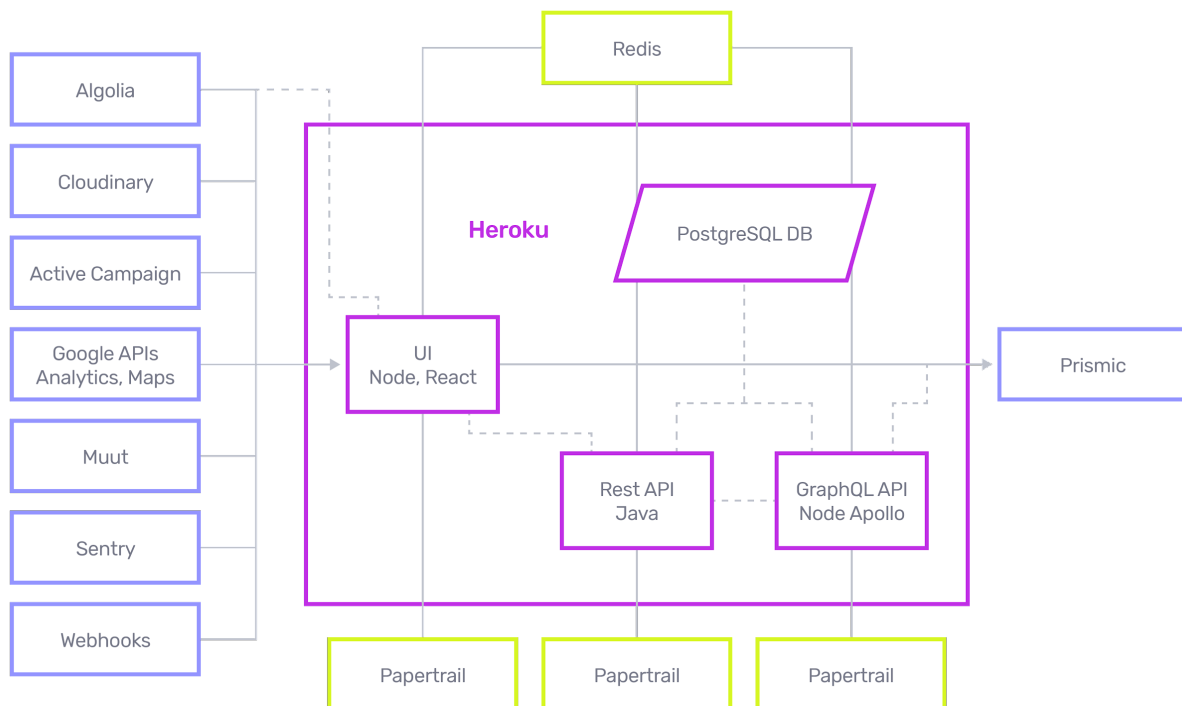


Figure 5: Rise's platform architecture

- Node is used for their front end UI and is built on React.
- They are migrating from their legacy Java REST API to GraphQL API (Node Apollo).
- Rise uses PostgreSQL as their relational database and also uses Redis.
- Papertrail is used for log management.
- Algolia has been chosen for search indexing.
- Cloudinary is used for image hosting.
- ActiveCampaign is used for email campaign lists.
- Rise uses various Google products including Google Analytics, Google Maps and Google Places.
- Sentry is used for error tracking within their UI and API.
- Webhook automations have been set up to enable notifications, background processes and alerting in Slack.
- Prismic is their content management system for articles and videos available on their website.

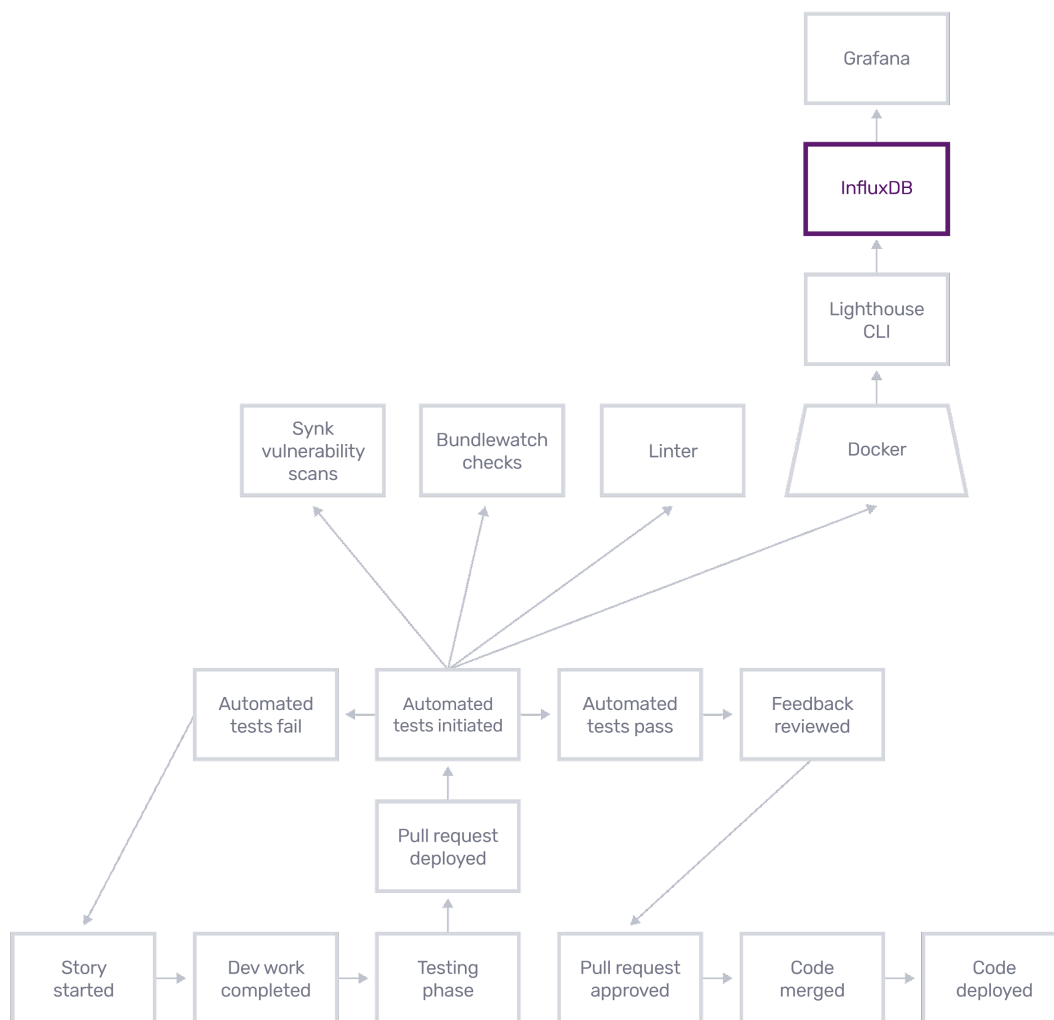


Figure 6: Rise's development architecture

InfluxDB is integral to Rise's internal development process. All metrics collected through the entire process are stored in InfluxDB and then visualized in Grafana. InfluxDB's integration with Grafana was very important in their decision-making process. Rise's team was already quite comfortable using Grafana as a visualization tool. Rise realized that if they couldn't push everything to InfluxDB, they could combine data sources, create one dashboard in Grafana, and visualize different data sets in one place. The combination of InfluxDB and Grafana is pain-free for Rise; they are able to run it in Docker and can manage it using Portainer. It also reduces the work required to manage everything.

Results

InfluxDB has enabled Rise to understand their backend systems better, which has ultimately improved website performance. The team recognizes they aren't where they want to be just yet and they have aggressive goals in mind and are always working to improve their website's performance.



Figure 7: Dashboards using Lighthouse metrics

- JavaScript bundle is 165% smaller.
- CSS bundle is 185% smaller.
- First contentful paint and meaningful paint times on average are down 2 seconds, and some are down 5 seconds.
- The speed index across the entire site is down by 8 seconds.
- Performance score went up 170%.
- Time to interactive went down 200%.

If VanTassel were to only look at the previous Lighthouse dashboards, he would see that First Meaningful Paint started at 3.5 seconds, went up to 4.5 seconds and back down to 1.5-2 seconds. Using the dashboard below, Rise can drill into the difference between multiple pages, perform root cause analysis, and determine which specific web page components or containers of components are negatively impacting First Meaningful Paint. It is VanTassel's recommendation to have different visualizations of the same data so that you can quickly dig into your data, triage and ultimately understand your data better.

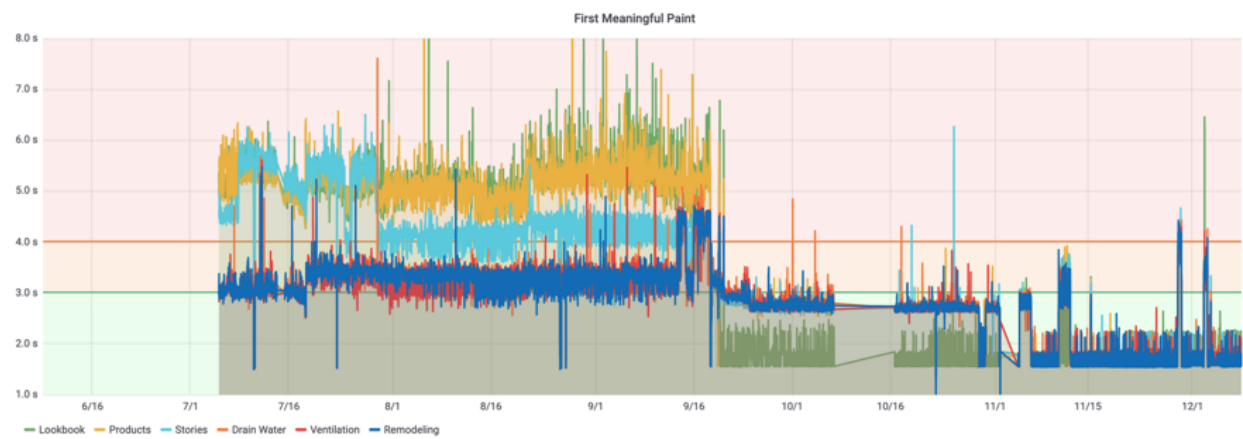


Figure 8: First Meaningful Paint dashboard

VanTassel developed a dashboard to understand how performance has shifted in a specified time frame. As seen below, the graph on the left shows the current First Meaningful Paint times, and the graph on the right demonstrates where they were 5 months ago. The thresholds are set within Lighthouse, which is denoted by the green areas in the graphs. VanTassel is proud to point out that all of their web pages are now performing well within the green threshold. The upper limit of the green threshold is currently 3 seconds. VanTassel hopes to reduce it to 2 seconds.



Figure 9: First Meaningful Paint comparison

Lessons learned

VanTassel likes to reinforce that organizations should be continuously tweaking and optimizing their visualizations. Parameters that were set 6 months ago might not make sense today. Sometimes they need to be able to dive deeper into their data. However, if the boundaries are set too wide, anomalies might not be apparent. Dashboards might not be configured correctly to see all variances in datasets. He recommends making sure that data collection allows you to zoom into the data as needed.

Having alerts for everything, including alerts for no data, is an important lesson that Rise had to learn the hard way. It's better to properly set up alerting at the beginning. Rather than focusing on specific tests, it's better to include everything at the beginning. Only time will provide organizations with insight into important metrics.

It's better to perform the same test multiple times and use the average of all of the tests. When relying on just one test, it is possible there could be a slow network connection at that moment, or maybe there is a high CPU and memory load on the machine. Rise's rule-of-thumb is to measure performance based off of the median scores collected from multiple tests – sometimes even up to 10. It is foolish to run one test and assume that performance has not been impacted.

VanTassel likes to remind himself to experiment with his data. His recommendations include cloning existing dashboards that are used frequently, and then experimenting with different ways to visualize the same data and to play with different data segments.

It's also important to get really familiar with your data – to look at it everyday – so that it's easy to understand shifts in your data. The best way to continuously gain value from the data and dashboards is to experiment and fine-tune existing dashboards.

What's next

Rise's future plans include pushing all metrics into InfluxDB that make sense for monitoring. They'd like to ingest their dyno metrics into InfluxDB. Their dyno metrics include LOAD, MEM, RPS, errors, etc. In the future, they would like to add their continuous integration job statuses into InfluxDB including bundlesize, Snyk, Cypress and Garie data.

By starting with page view data, Rise hopes to begin adding Google Analytics metrics into their time series database. They want to be able to correlate performance and other disparate metrics against actual website traffic data. They hope to annotate all dashboards with corresponding deployments so they know who the author was, the release and the build ID. By doing this, they hope to be able to quickly resolve issues related to particular builds.

They use Sentry as their error logging tool. By pushing all error logs into InfluxDB, they will begin to visualize errors against potentially high CPU loads — potentially high loads on their UI dyno or their REST API dyno. As CTO, VanTassel wants to be able to see, at-a-glance, if there is a systemic reason for performance issues.

VanTassel wants to ensure dashboards are dynamic so that team members can drill into the logs or the Sentry errors for debugging purposes. Dynamic tools for his developers are key, as they will be beneficial when the team is deep into triaging issues.

They want to create disposable environments, using Heroku Review Apps, to replace what they're currently doing in their staging environment. They want to empower their developers to create a pull request, have it spin up a testing environment, run end-to-end tests, run performance tests, and provide a functional environment for manual functional testing. The reviewer should be able to immediately deploy code changes to the staging environment or run it locally through functional tests.

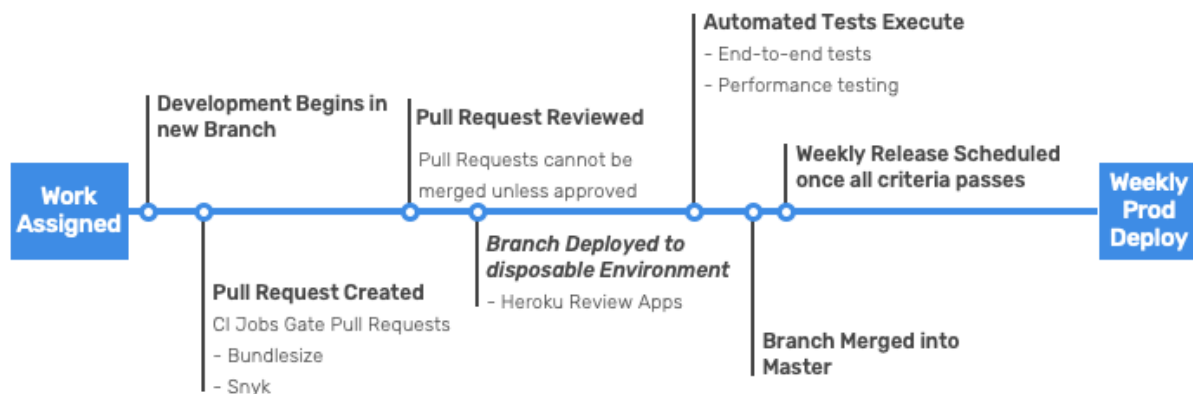
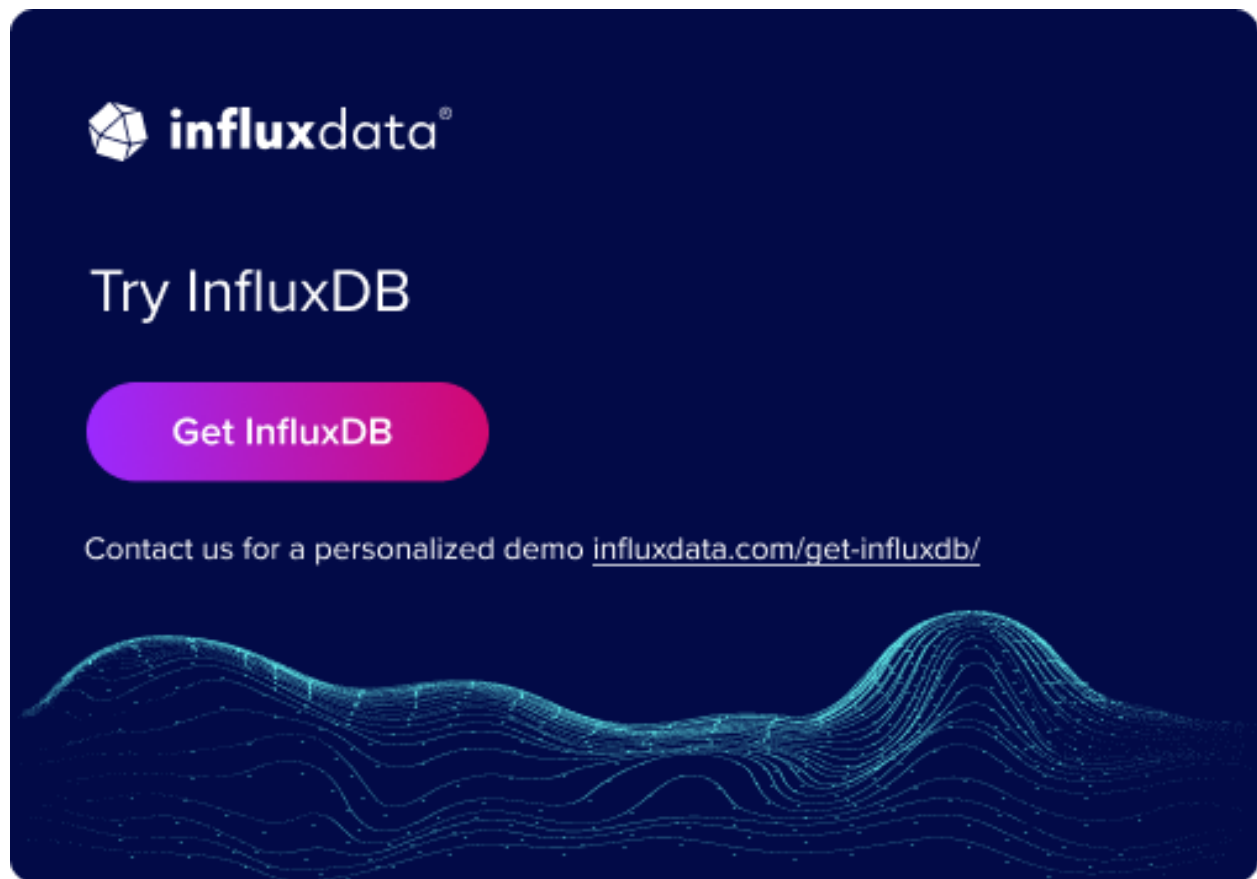


Figure 10: New development and deployment process

Rise provides professional and manufacturer clients with dashboards that are built with PostgreSQL. They are facing row count limit issues. They have aggregated their data, which enabled them to create 30 days' worth of high-precision data; all of the other data will go into total summaries. They would like to be able to answer questions like "How did I perform last Christmas?" and "How am I doing week over week?". Rise knows InfluxDB will be able to address some of these issues. They will have to start with mirroring this data into InfluxDB for at least 30 days and beyond. They will also have to determine how to provide loyal customers with long-term data. They need to make sure they are able to provide a solution that is scalable, maintainable and flexible. Before offering them an improved platform, Rise wants to make sure they've thought of all of their needs, including custom data ranges so they can look back over months or even years.

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](https://twitter.com/InfluxDB).



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