

Wolfram Wingerath & Sebastian Grebasch

Let's Agree to Disagree: Why Google's CrUX Results Are Not Reproducible With Your Real-User Monitoring

Design & Experience



code.talks

Measure user experience on the web.



Why should you care?





50 ms

Site speed improvement of just
0.1s
leads to an increase in conversion
rates & average order value

Google/55/Deloitte, [Speed Impact study](#), EMEA and US, Oct-Nov 2019, n=37 brand sites analysed hourly over a 30 day period totalling 30.5m sessions.

Update in Google Search - May 2021

Home > Google Search Central > What's new > Google Search Central Blog

Rate and review  

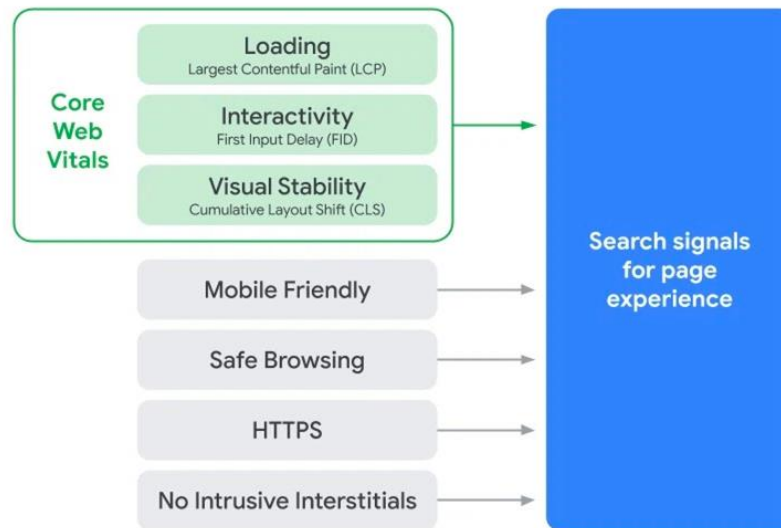
Timing for bringing page experience to Google Search

[Send feedback](#)

Tuesday, November 10, 2020

This past May, we announced that [page experience](#) signals would be included in Google Search ranking. These signals measure how users perceive the experience of interacting with a web page and contribute to our ongoing work to ensure people get the most helpful and enjoyable experiences from the web. In the past several months, we've seen a median 70% increase in the number of users engaging with Lighthouse and PageSpeed Insights, and many site owners using Search Console's Core Web Vitals report to identify opportunities for improvement.

Today we're announcing that the page experience signals in ranking will roll out in May 2021. The new page experience signals combine Core Web Vitals with our existing search signals including [mobile-friendliness](#), [safe-browsing](#), [HTTPS-security](#), and [intrusive interstitial guidelines](#).

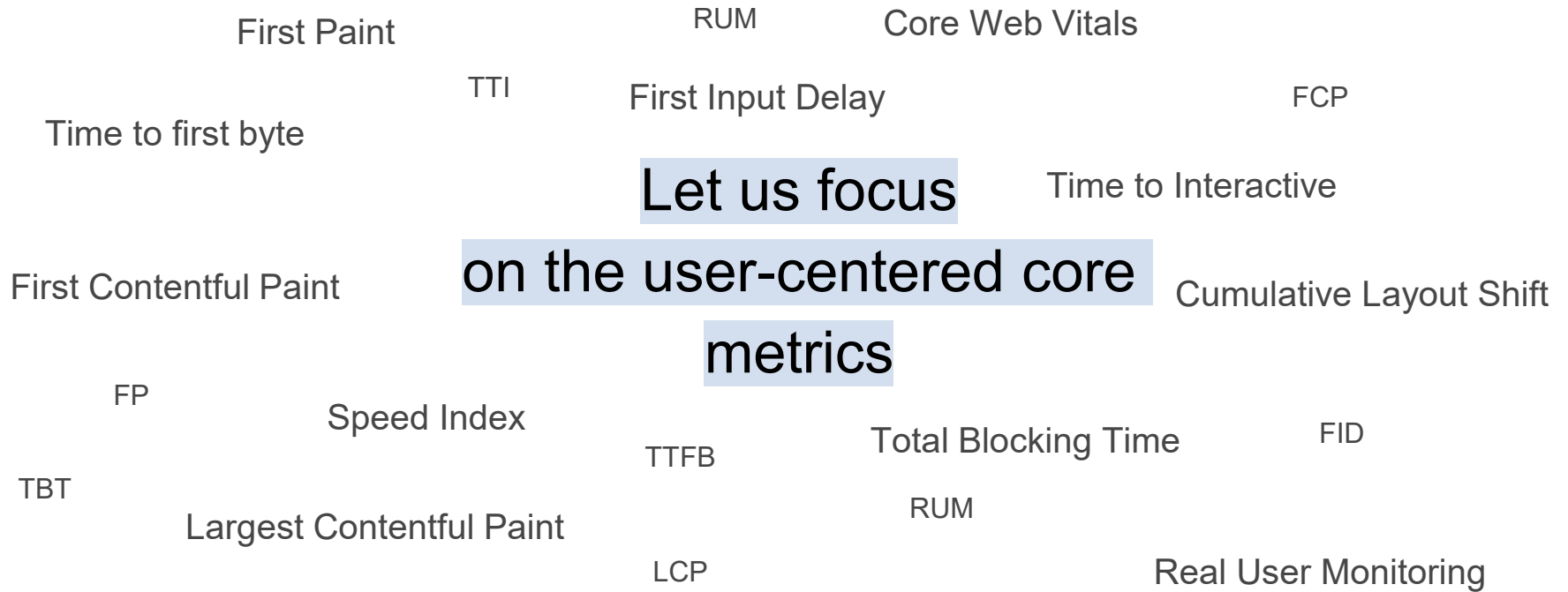


<http://goo.gle/page-experience-timing>

Core Web Vitals

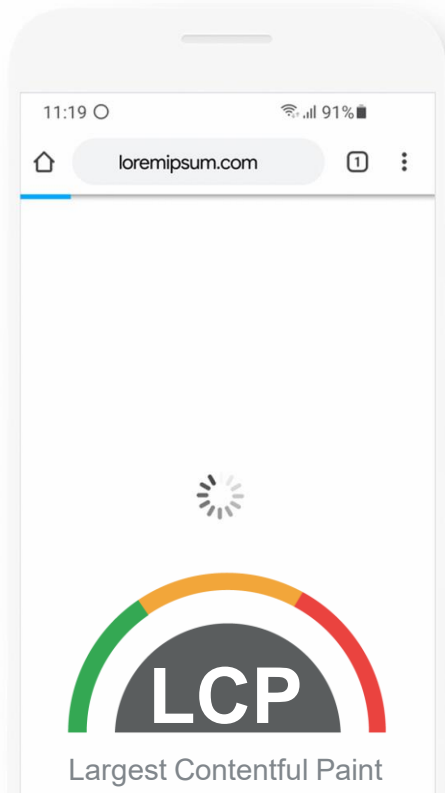
Providing unified guidance for quality signals that are essential to delivering a great user experience on the web.

Performance Metrics

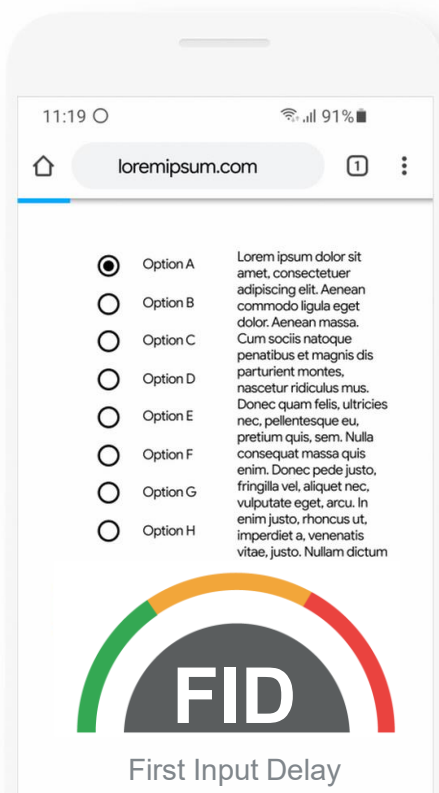


Prioritising the User Experience

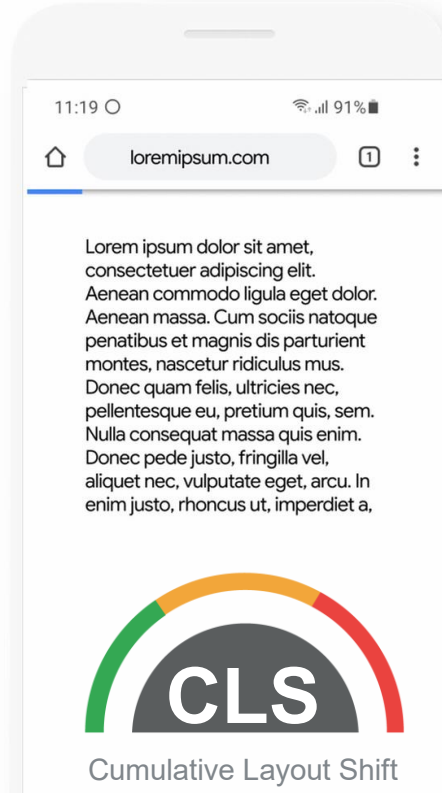
Is it **happening**?



Is it **responsive**?



Is it **stable**?



Tool Taxonomy and CRUX

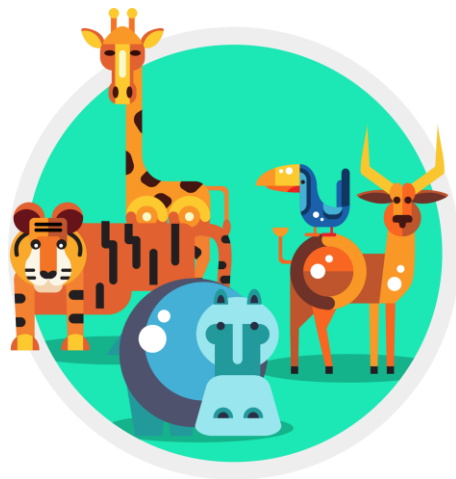
How to think about Lab vs. Field speed tools...

Taxonomy of Speed Tooling @ Google



Lab data

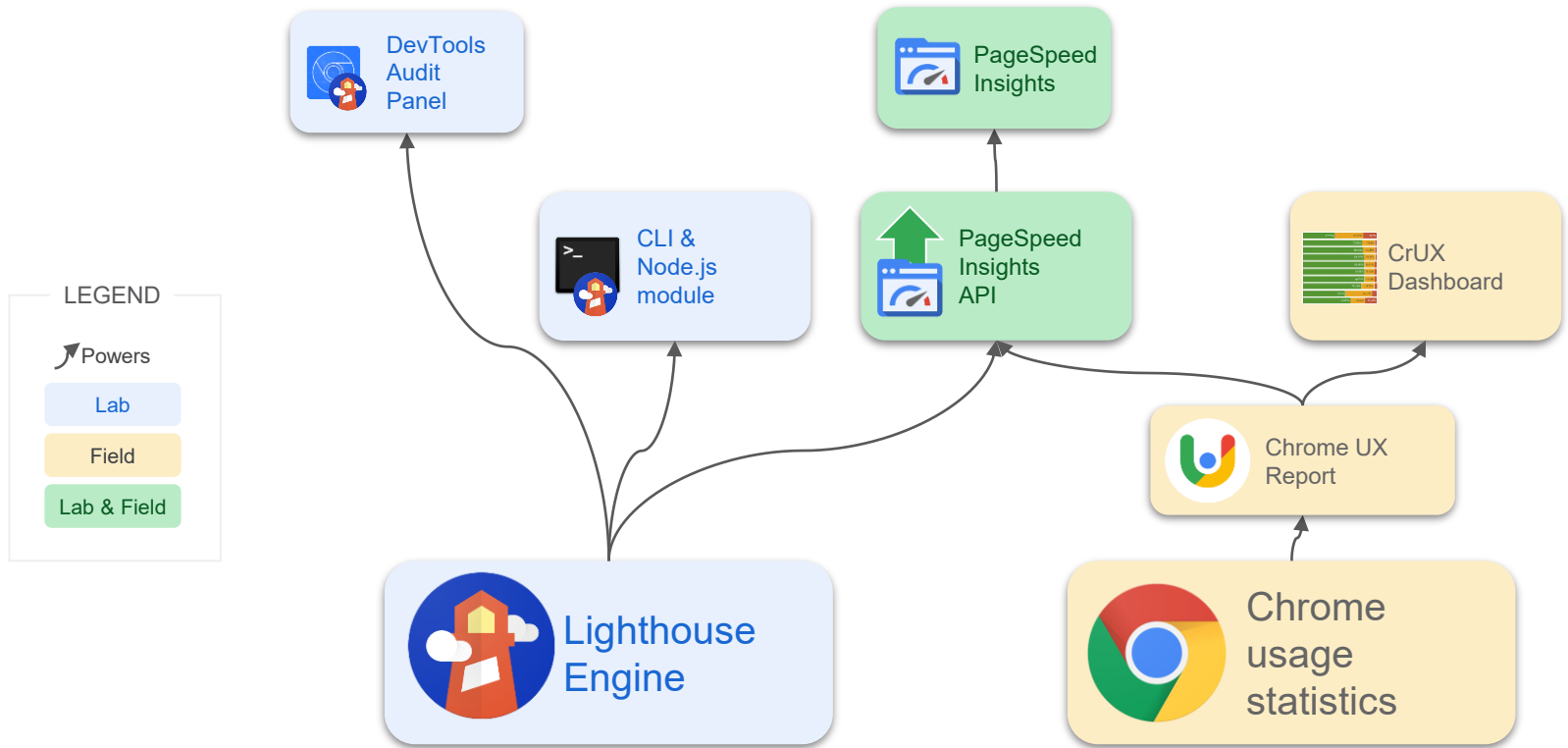
(aka metrics from the lab)



Field data

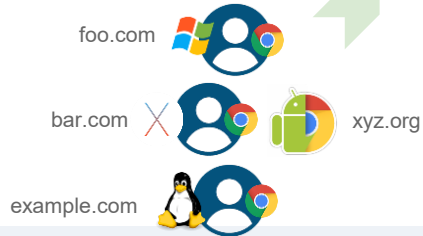
(aka metrics from the wild)

Taxonomy of Speed Tooling @ Google



How does CrUX work?

CrUX provides “field data” for top sites on the web. Field data is performance data collected from real page loads users are experiencing in the wild.



1

Real-world user experience data (real devices, network conditions, etc) is collected from opted-in Chrome users as they browse the web, and uploaded to Google’s URL Keyed Metrics service.

2

UKM data is anonymized, non-public URLs are filtered out and URLs and origins with insufficient number of samples are removed. The end result is consumable, regularly updated data on how real-world users experience sites on the web.

3

CrUX data is made public via a variety of channels: BigQuery, PageSpeed Insights, Search Console, and other 1P & 3P products.

Case Study

Vodafone A/B Testing AD Landing Pages

Optimizations for 50% of traffic via different traffic sources

- Image Optimization
- Server-side Rendering of critical HTML & Widget Optimization



Learn more @ <https://web.dev/vodafone/> Screenshot Source: <https://www.web.dev/vodafone>



Speaker Shuffle



Sebastian

Speaker Shuffle



Wolle

Real-User Monitoring



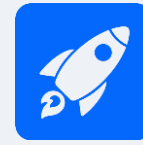
We bring performance research to practice



30+ man-years of **web performance research** at University of Hamburg



Novel technology for **caching dynamic data** went into Baqend in 2014



Baqend **launched Speed Kit** as the all-in-one page speed platform in 2018

7,000

customer websites are already using Speed Kit

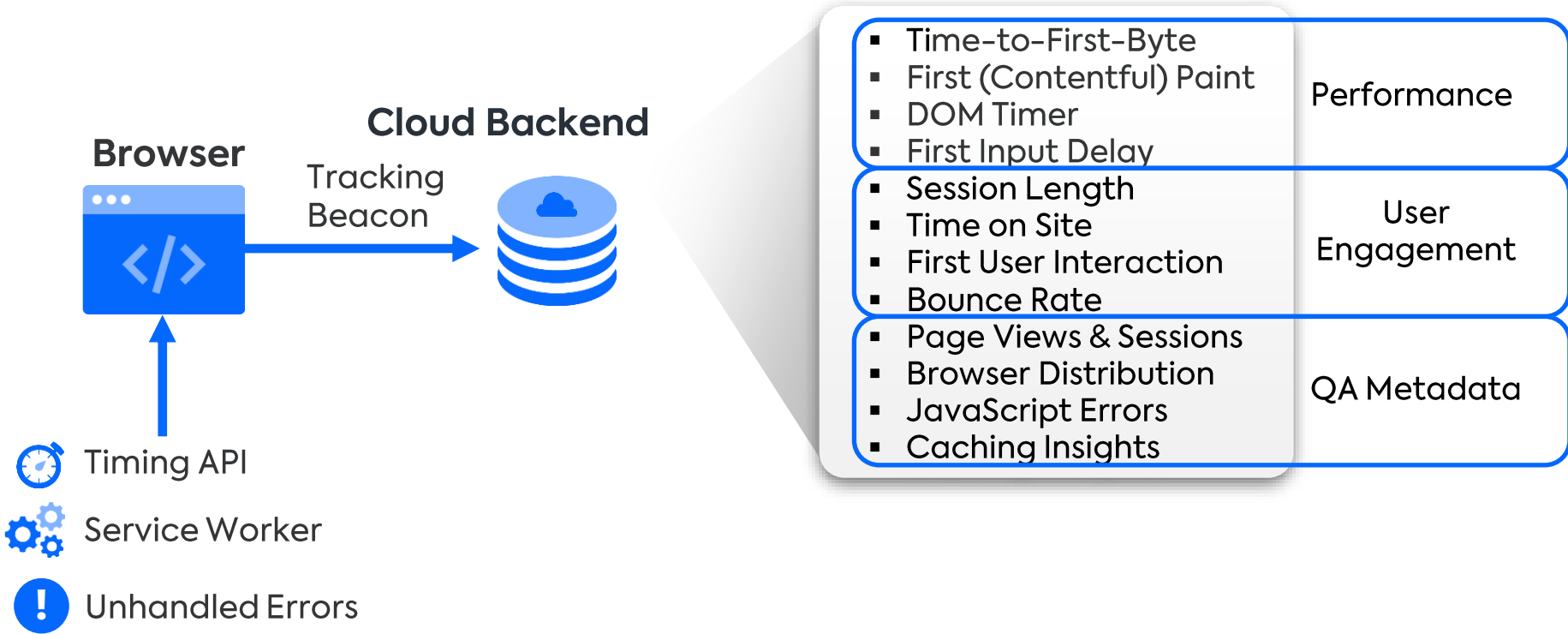
\$2.6 billion

in annual revenue runs with Speed Kit

160 million

users per month benefit from Speed Kit

The Basic Idea



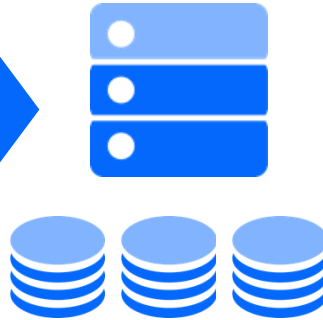
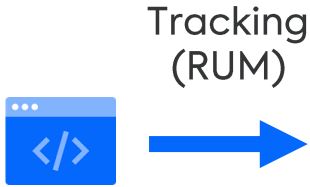
Industry Example

Collection

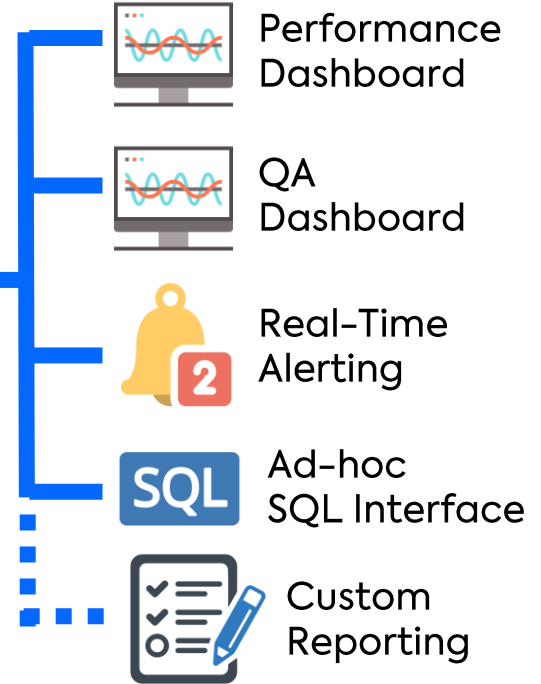
Ingestion

Analytics

Reporting



SQL Interface



- Raw PI tracking & meta data
- Custom tracking

- Materialized views & aggregations
- Historical data

Industry Example

Collection



Tracking (RUM)



Ingestion



- Raw PI tracking & meta data
- Custom tracking

Analytics



- Materialized views & aggregations
- Historical data

Reporting

SQL Interface



Performance Dashboard



QA Dashboard



Real-Time Alerting

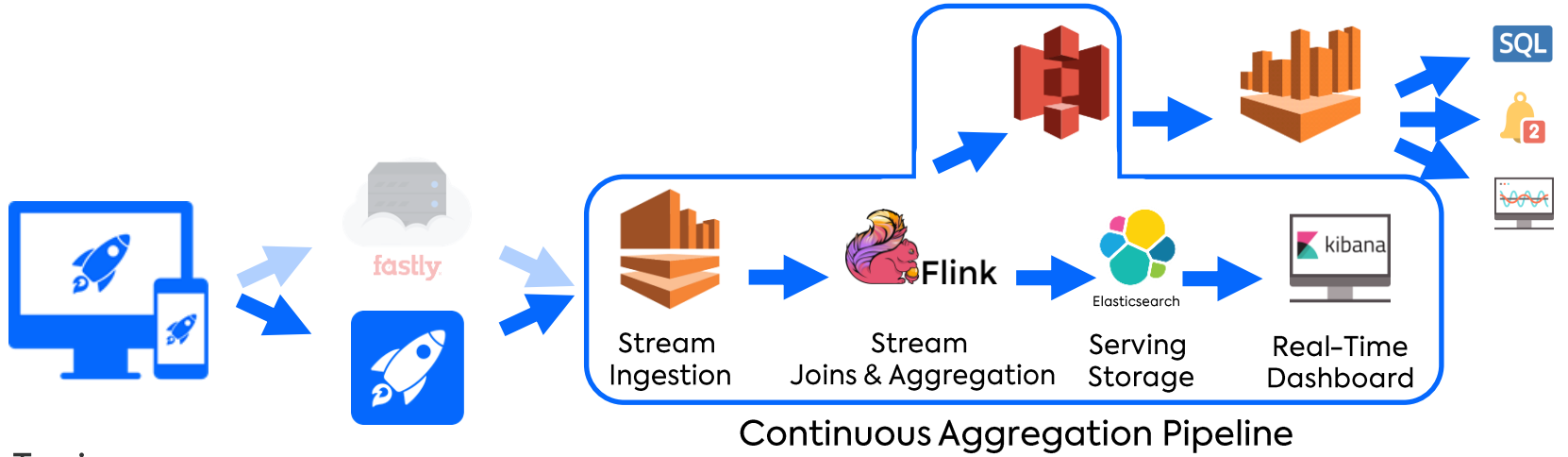


Ad-hoc SQL Interface



Custom Reporting

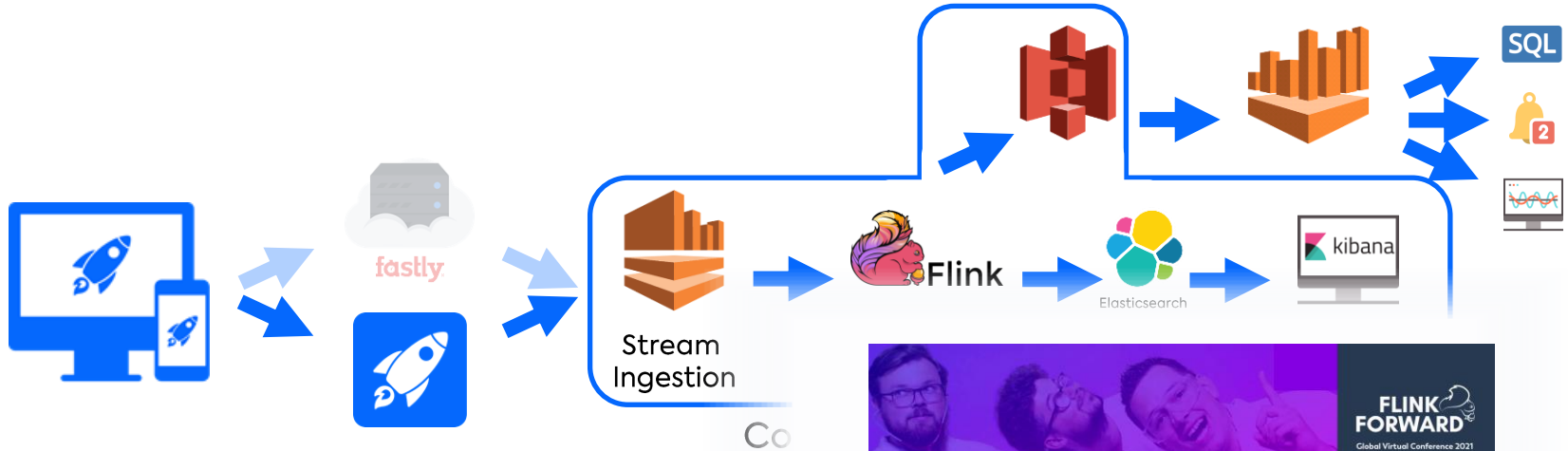
Beaconnect: Continuous Processing



Key Topics:

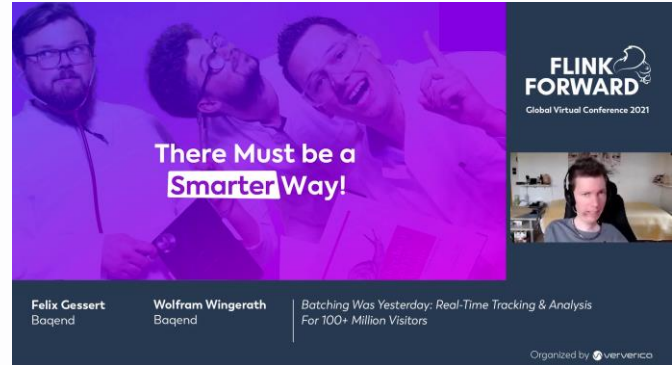
- ✓ Continuous Aggregation
- ✓ Real-Time Reporting & Analytics
- ✓ Extreme Scalability

Beaconnect: Continuous Processing



Key Topics:

- ✓ Continuous Aggregation
- ✓ Real-Time Reporting & Analytics
- ✓ Extreme Scalability



F. Gessert, W. Wingerath, *Batching Was Yesterday: Real-Time Tracking & Analysis For 100+ Million Visitors*, Flink Forward (2021)

Analyzing Performance Data



Split Testing for Web Performance

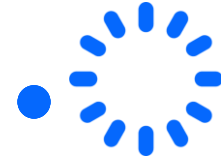
Speed Kit Users



vs.



Normal Users



- Speed Kit enabled

- **Measurable uplift:**
 - + Performance
 - + User engagement
 - + ...

- Speed Kit disabled
(no acceleration)



W. Wingerath, B. Wollmer, M. Bestehorn, S. Succo, F. Bücklers, J. Domnik, F. Panse, E. Witt, A. Sener, F. Gessert, N. Ritter. *Beaconnect: Continuous Web Performance A/B-Testing at Scale*. VLDB 2022

Split Testing for Web Performance

Speed Kit Users



Tracking
→

vs.



←
Tracking

Normal Users



- Speed Kit enabled

- **Measurable uplift:**
 - + Performance
 - + User engagement
 - + ...

- Speed Kit disabled
(no acceleration)



W. Wingerath, B. Wollmer, M. Bestehorn, S. Succo, F. Bücklers, J. Domnik, F. Panse, E. Witt, A. Sener, F. Gessert, N. Ritter. *Beaconnect: Continuous Web Performance A/B-Testing at Scale*. VLDB 2022

3 Levels of Aggregation

Partial Page Impressions (PPIs)
Enhanced Data Beacons

Time	Browser	Device	Test Group	First Contentful Paint (FCP)
11:05:04.578	Firefox	Mobile	Speed Kit	127ms
11:06:48.139	Chrome	Mobile	Original	958ms

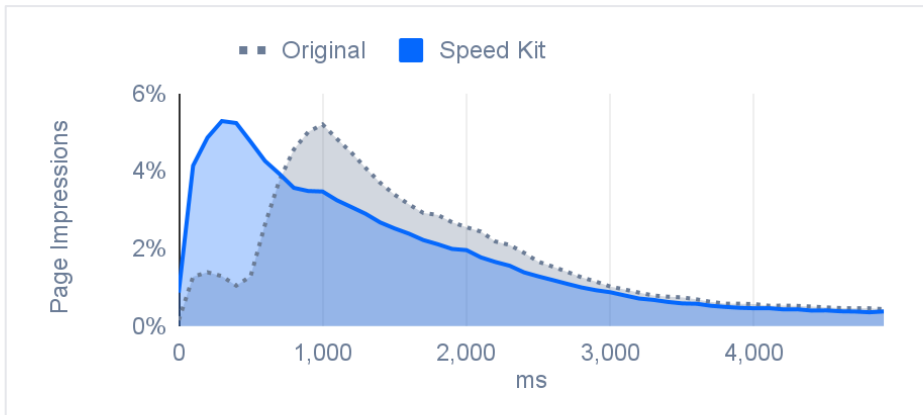
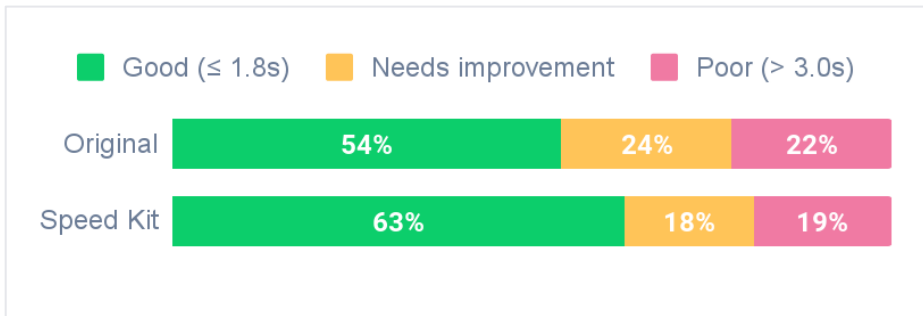
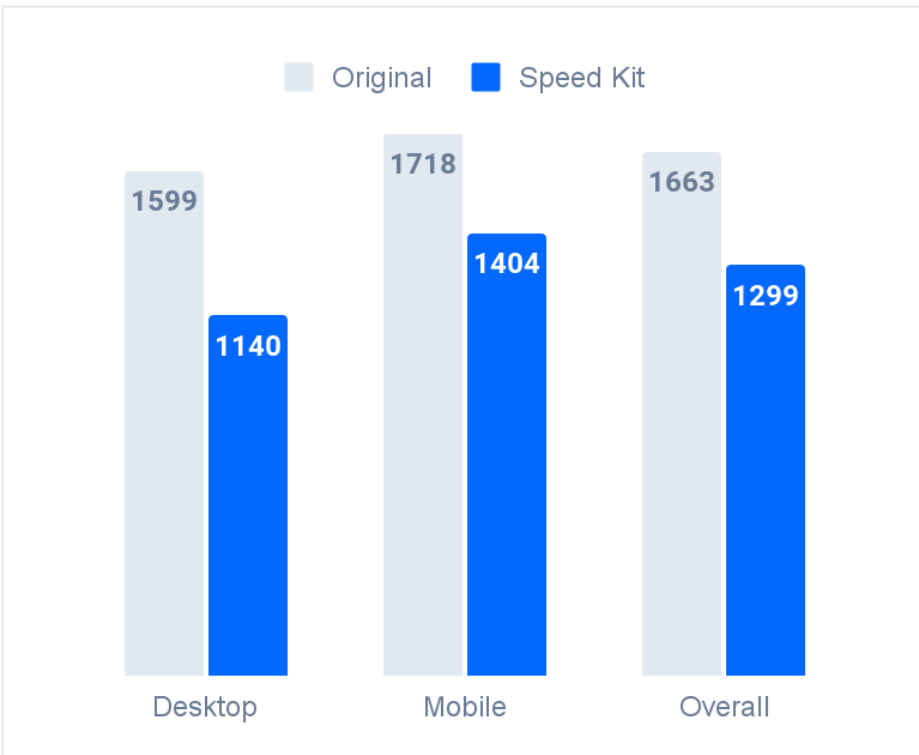
1-Min. Time Windows
Immediate Aggregates (Storage)

	Browser	Device	Test Group	First Contentful Paint (FCP)
11:05	Firefox	Mobile	Speed Kit	{200ms: 1, 500ms: 2}
	Firefox	Mobile	Original	{600ms: 2, 800ms: 5}
	Safari	Desktop	Original	{1100ms: 1}
11:06	Firefox	Mobile	Speed Kit	{200ms: 3}
	Chrome	Mobile	Speed Kit	{400ms: 2}
	Opera	Tablet	Original	{700ms: 1, 1300ms: 2}
	Safari	Desktop	Original	{600ms: 4, 900ms}

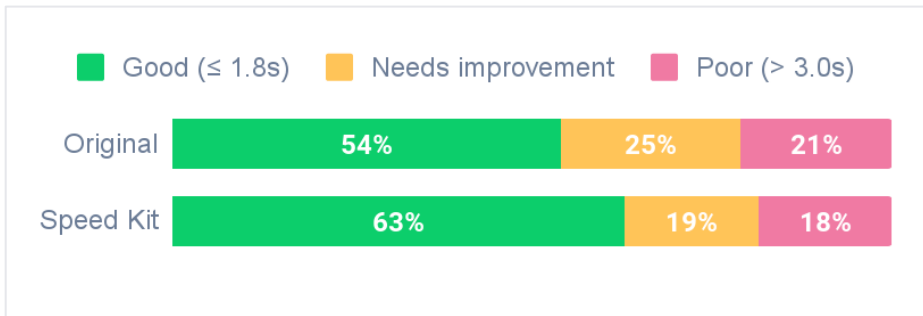
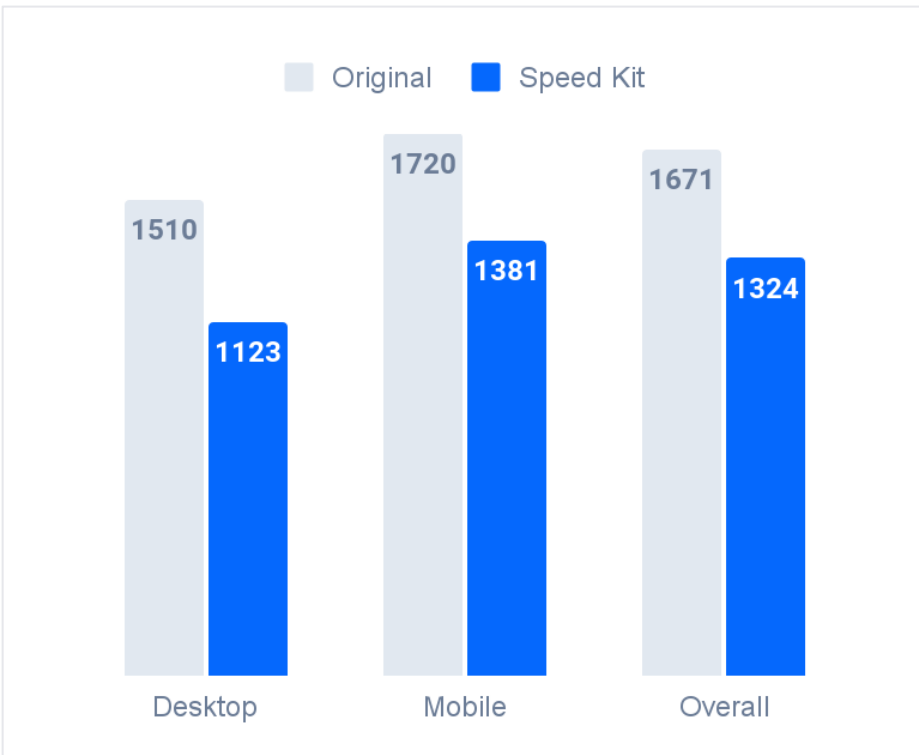
Arbitrary Time Windows
Real-Time Reporting (Dashboard Queries)

	Browser	Device	Test Group	First Contentful Paint (FCP)
11:05				
-				
11:06	Firefox	Mobile	Speed Kit	{200ms: 4, 500ms: 2}

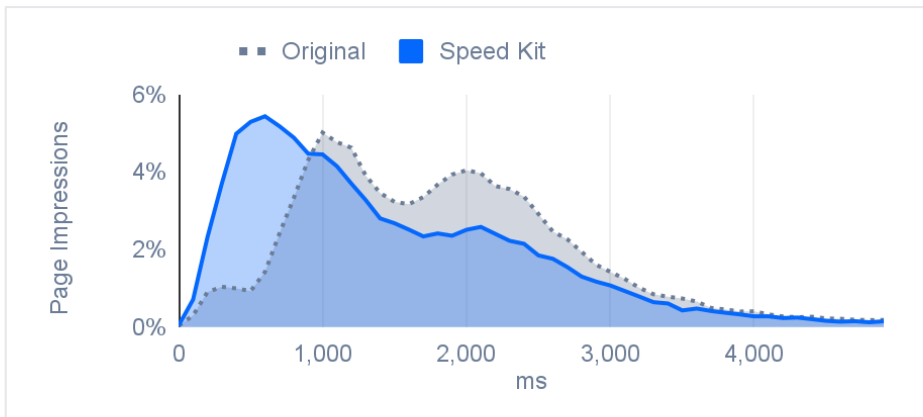
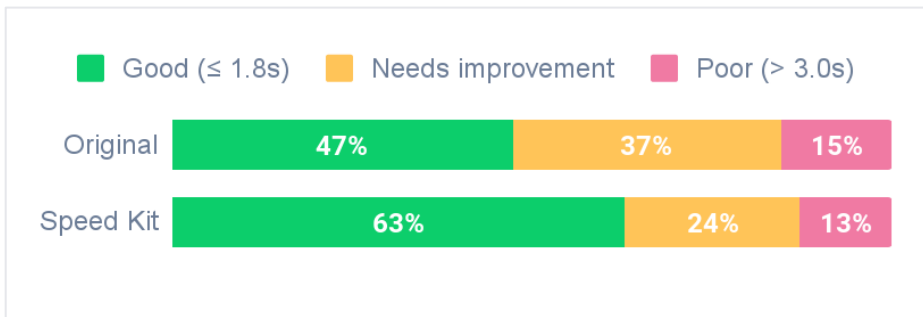
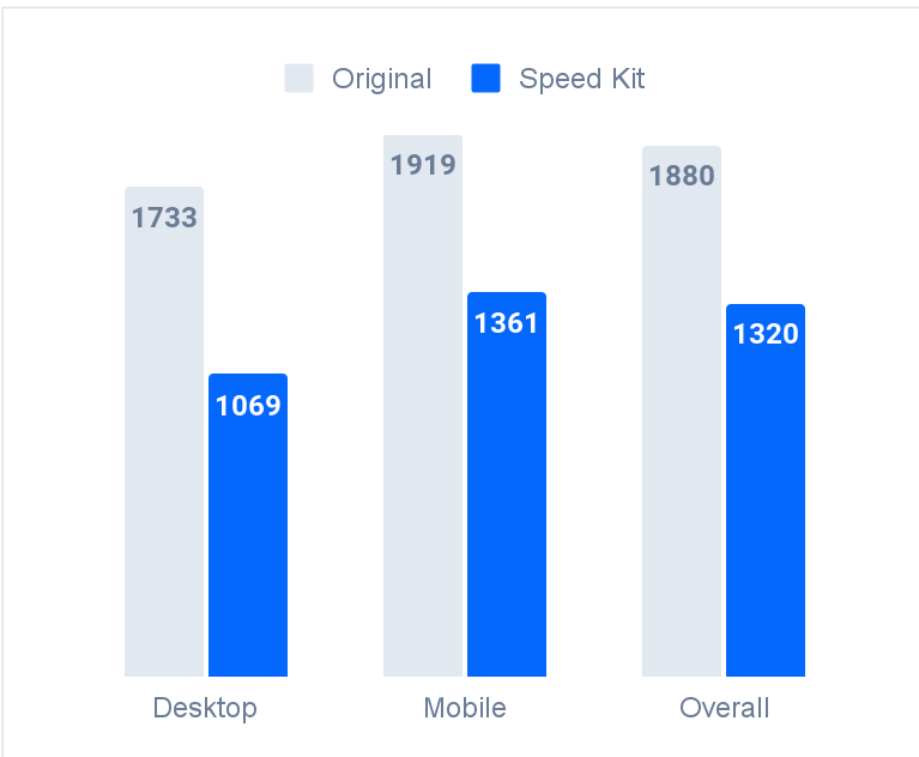
Applying Dimension Filters: All Users



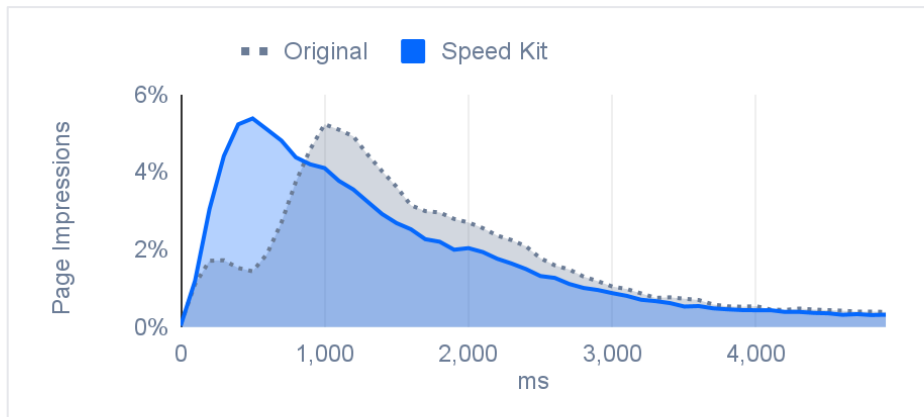
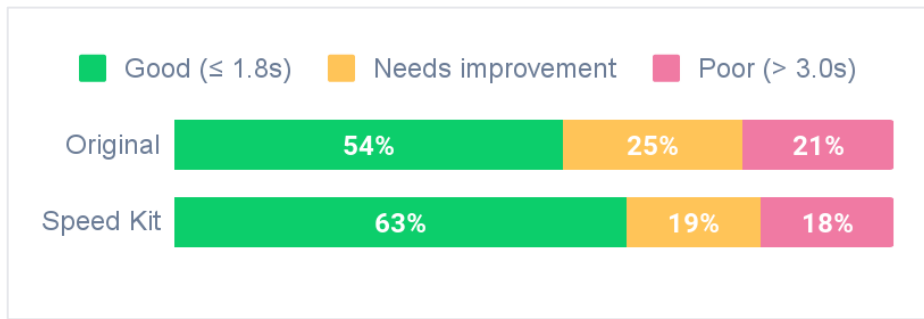
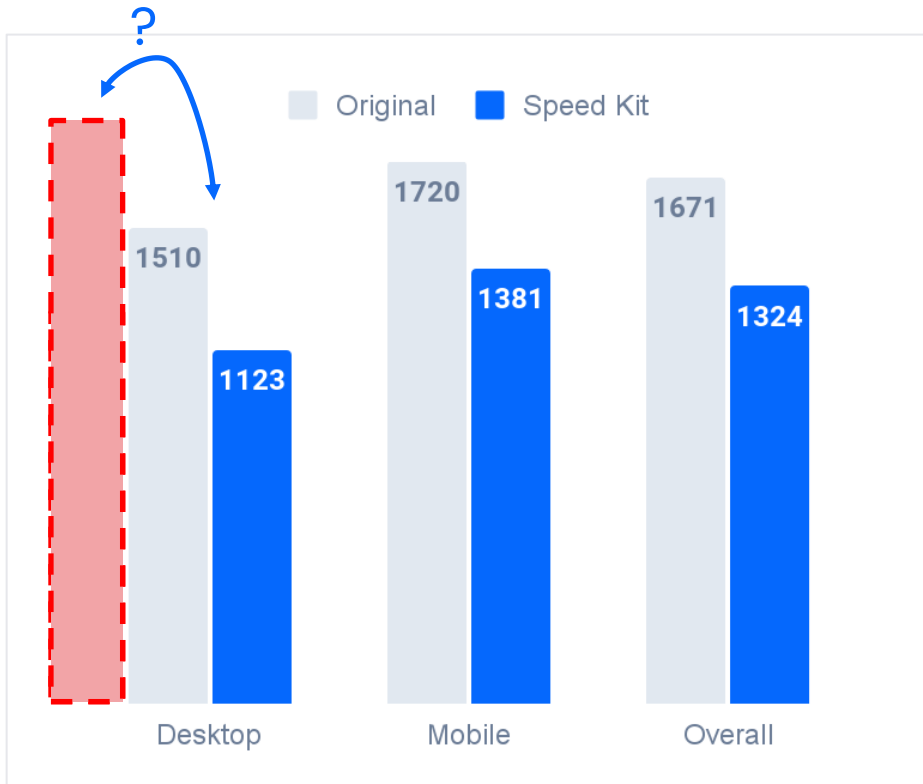
Applying Dimension Filters: Chrome



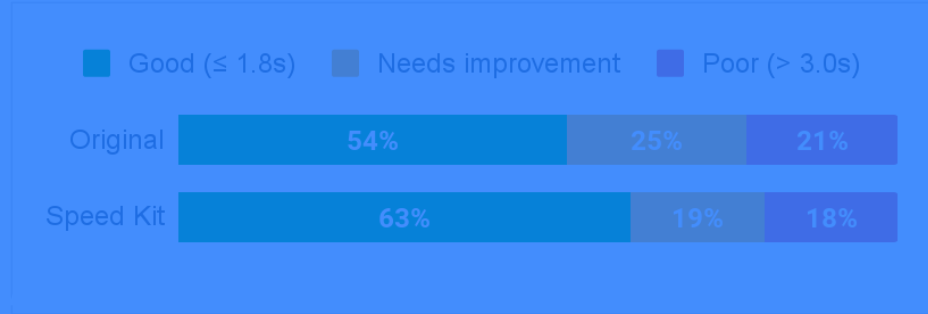
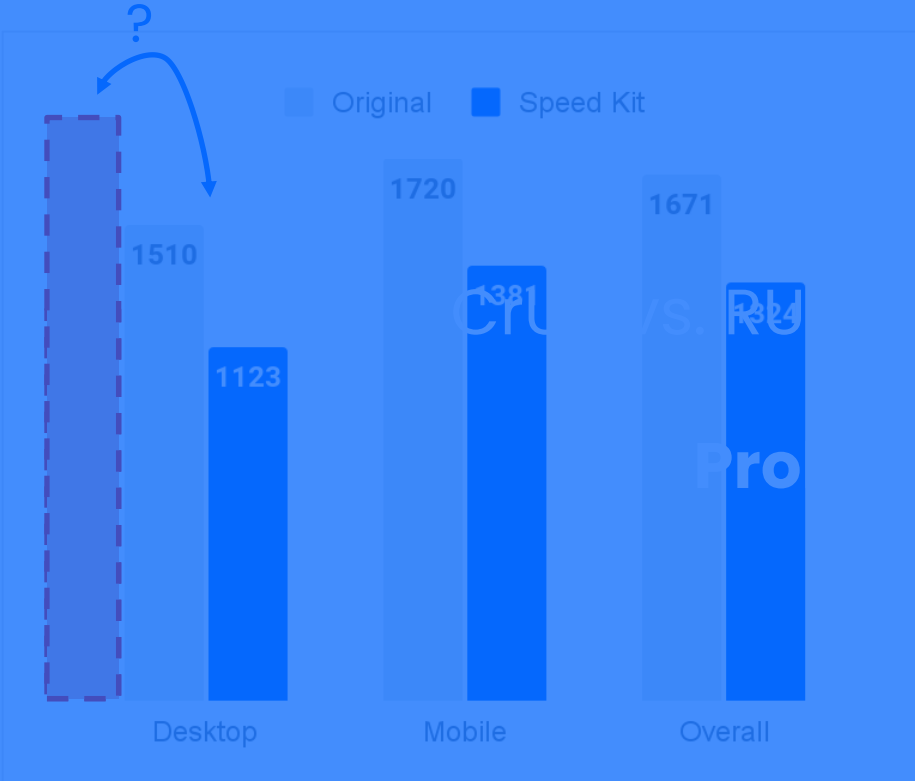
Applying Dimension Filters: Chrome, Product Pages



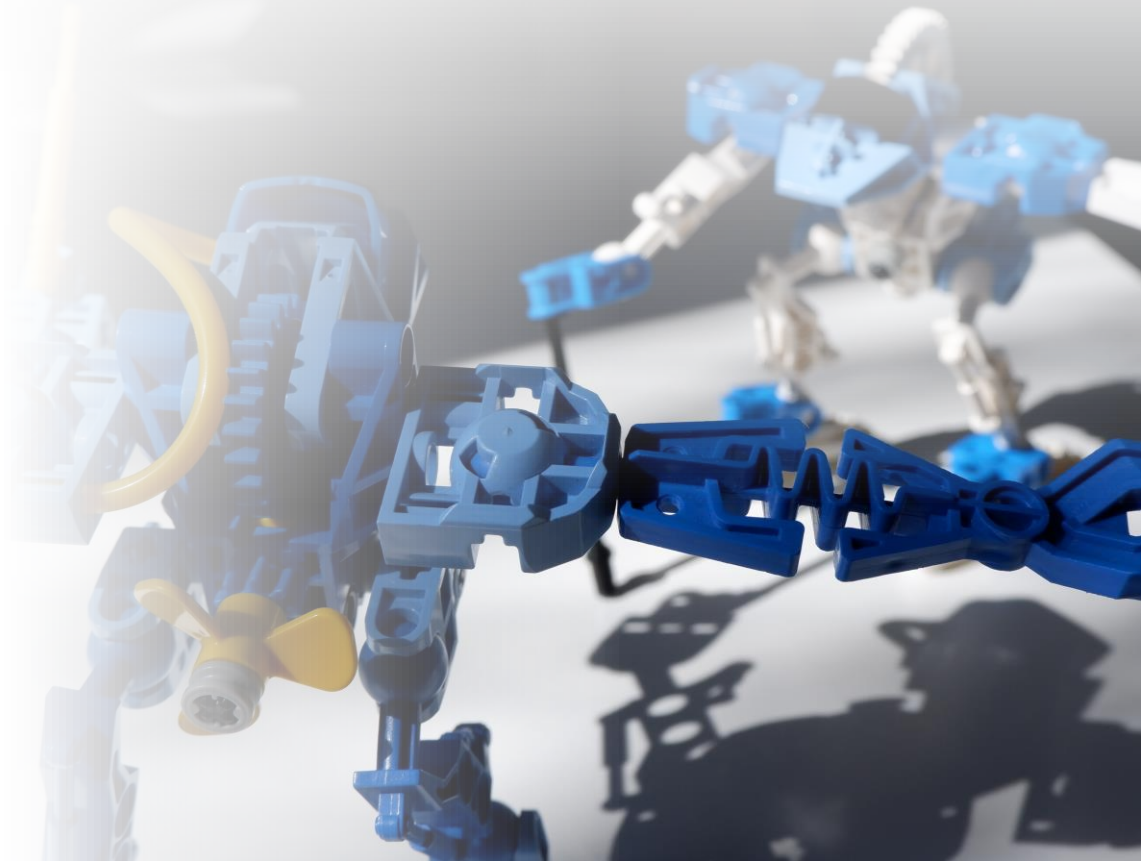
Applying Dimension Filters: Chrome



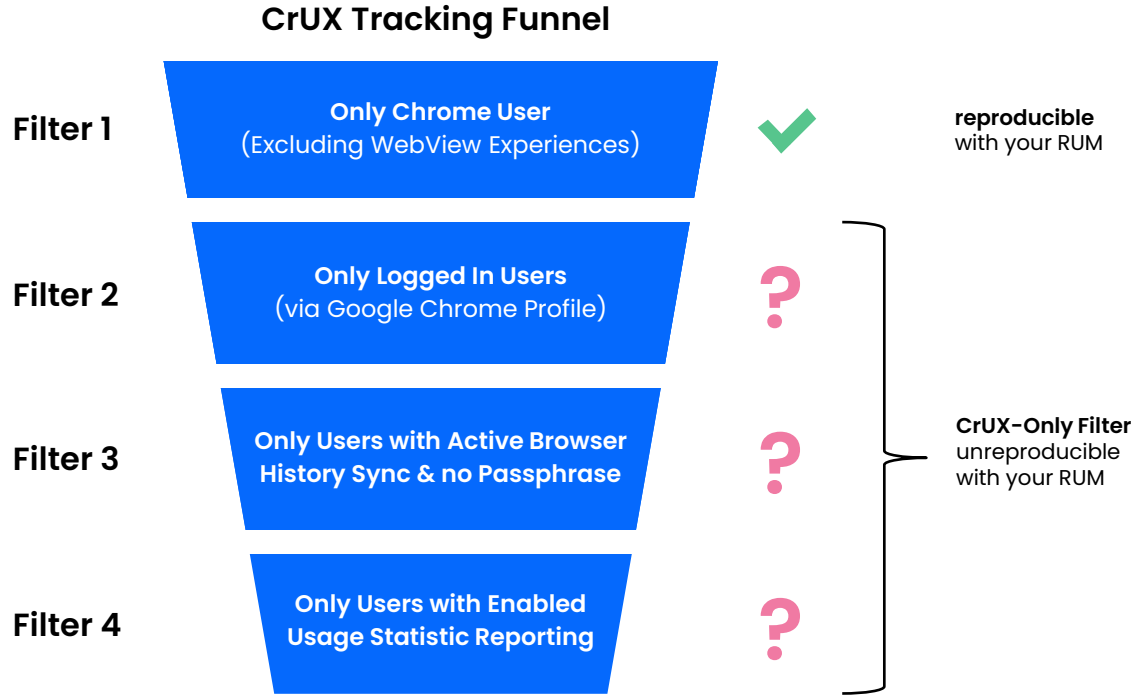
Applying Dimension Filters: Chrome



CrUX vs. RUM



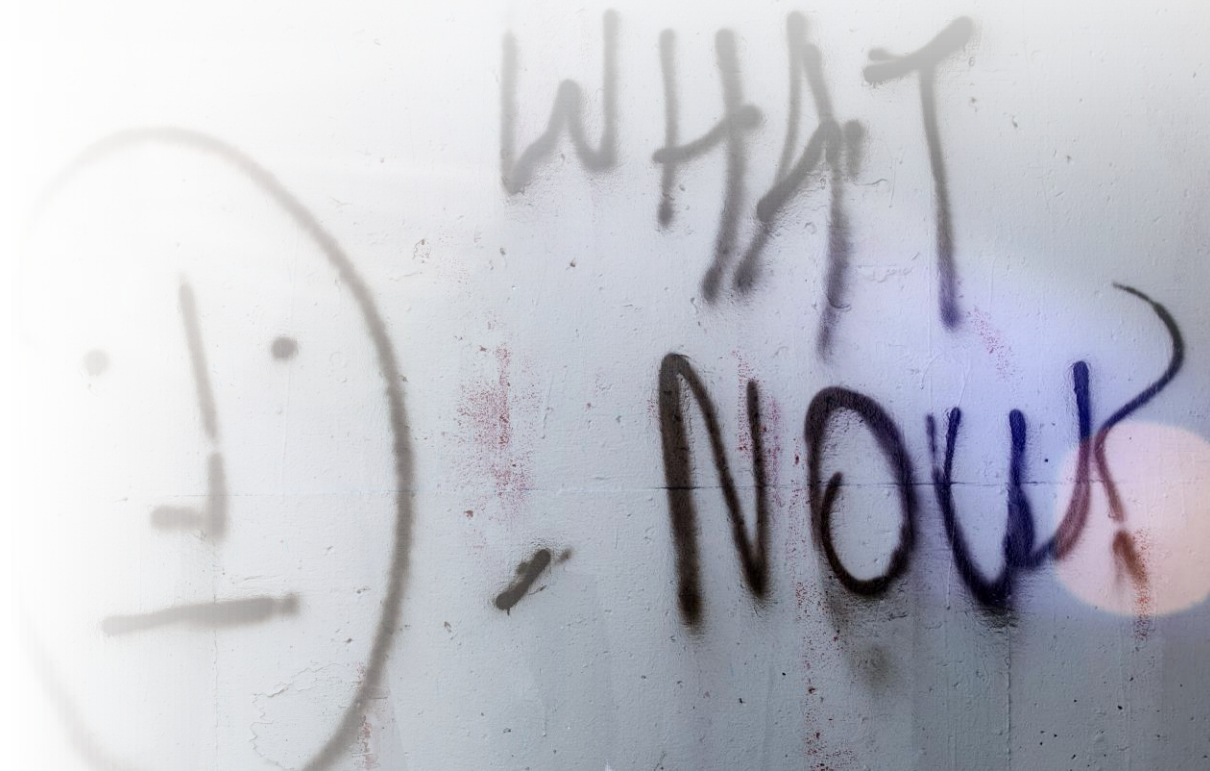
Why Google's CrUX Data is a Black Box



Sources: <https://developers.google.com/web/tools/chrome-user-experience-report>,
<https://groups.google.com/a/chromium.org/g/chrome-ux-report/c/i3ERRf7Mqio> (Access May 11, 2022)

 S. Ferrlein, W. Wingerath, B. Wollmer. Why Google's CrUX Results Are Not Reproducible With Your Real-User Monitoring. Baqend Tech Blog, 2022

Wrapup



Wrapup

CrUX

vs.

RUM

Based on field data

Public data

=> zero-effort

=> competitor data available

Fixed granularity

=> only by month / last 28 days

=> only fixed dimensions

=> only (part of) Chrome users

Based on field data

Custom deployment

=> complex to operate

=> Just your own website

Complete freedom

=> real-time / full detail

=> custom dimensions

=> all browsers

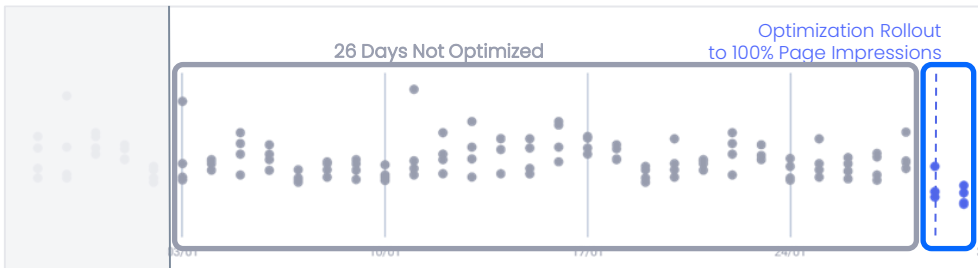


Neither gives you the full picture!

CrUX + RUM!



CrUX 28 Days Rolling Window: After **2 Days (7%)**



Still looks bad!

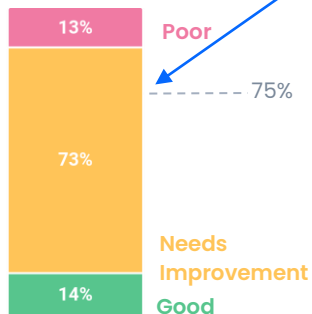
Looks good already!

After 2 Days of Optimization

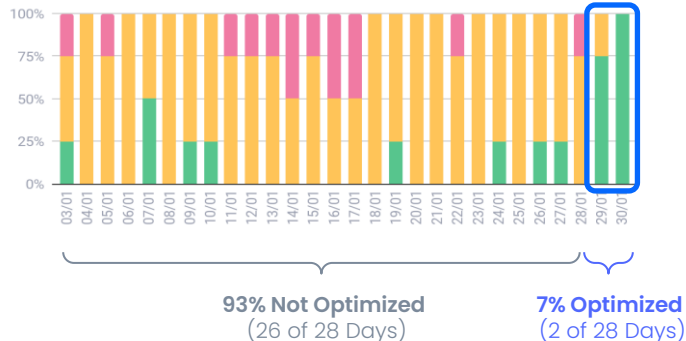
Two days after the 100% Rollout of an optimization, the Google CrUX 28 Days Report still includes **26 days (93%)** of not optimized performance.

The **2 Days (7%)** of optimized performance can only impact the overall result slightly.

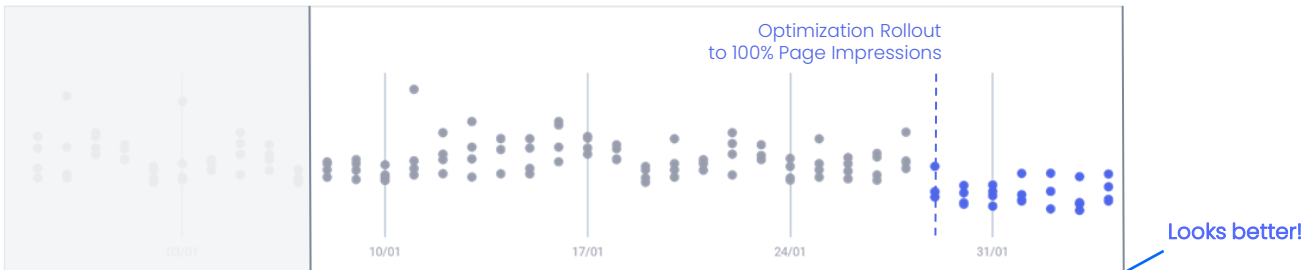
What Google Reports That Day



What Google Aggregates In Its Report That Day (But doesn't report in that detail)



CrUX 28 Days Rolling Window: After 7 Days (25%)

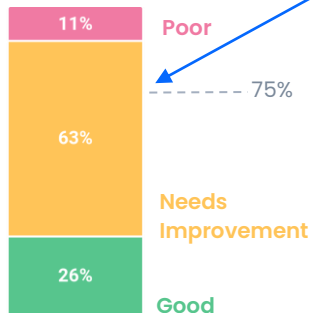


After 7 Days of Optimization

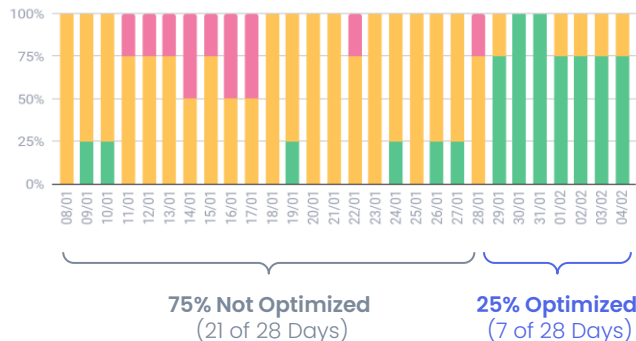
One week after the 100% Rollout of an optimization, the Google CrUX 28 Days Report still includes **21 days (75%)** of not optimized performance.

The **7 Days (25%)** of optimized performance are still not dominant in the overall result.

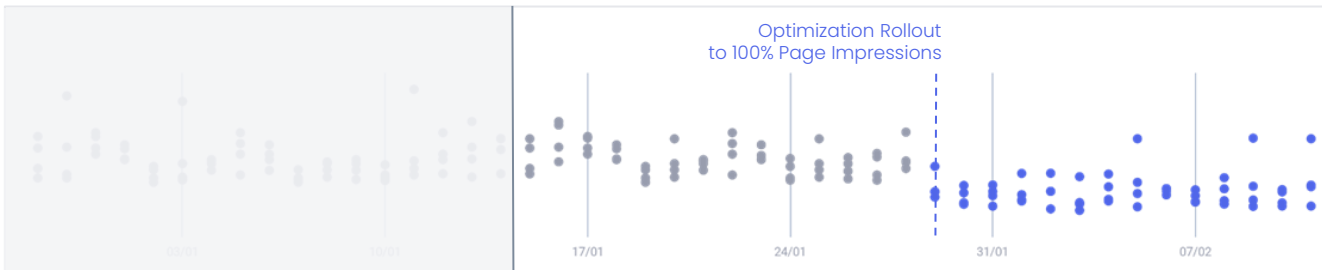
What Google Reports That Day



What Google Aggregates In Its Report That Day (But doesn't report in that detail)



CrUX 28 Days Rolling Window: After 14 Days (50%)

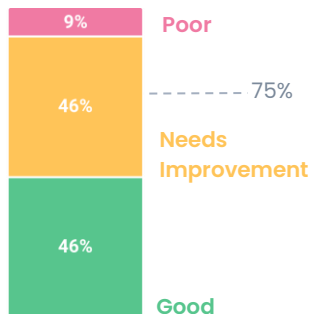


After 14 Days of Optimization

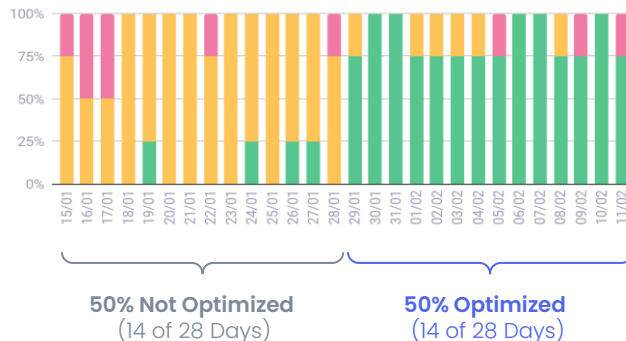
Two weeks after the 100% Rollout of an optimization, the Google CrUX 28 Days Report contains **14 days (50%)** of not optimized and **14 Days (50%)** of optimized performance.

From now on the impact of the optimization will be at least dominant in the overall result.

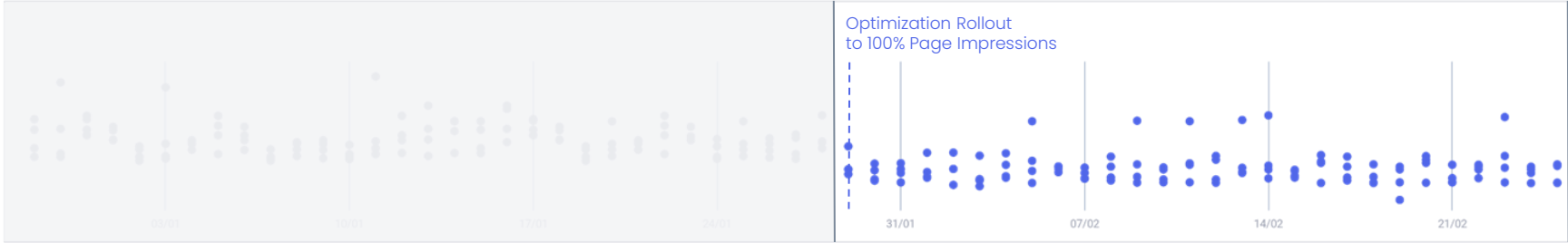
What Google Reports That Day



What Google Aggregates In Its Report That Day (But doesn't report in that detail)



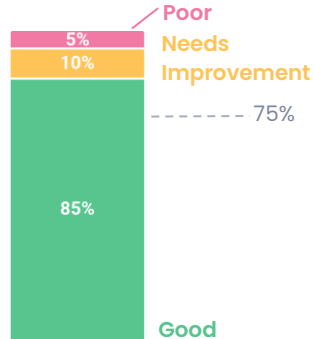
CrUX 28 Days Rolling Window: After 28 Days (100%)



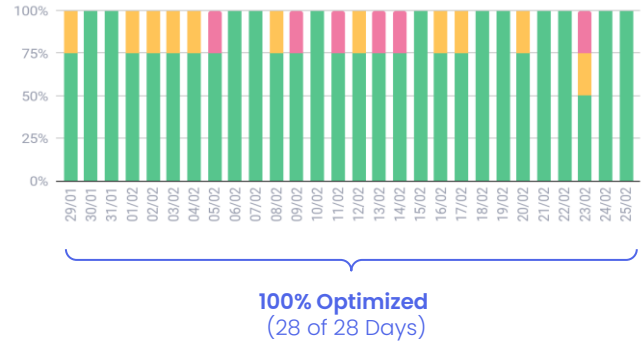
After 28 Days of Optimization

Only when the page optimization is live for **28 days (100%)** the effect will be fully reflected in the Google CrUX 28 Days Report.

What Google Reports That Day



What Google Aggregates In Its Report That Day (But doesn't report in that detail)



CrUX 28 Days Rolling Window: FCP CrUX Uplift Example

Before Optimization

Before the rollout of the optimization, the First Contentful Paint (FCP) of the last 28 Days is obtained from the CrUX API.

Sep 1 2022
CrUX API



CrUX FCP
2,301 ms

28 Days
Not Optimized

7 Days after Optimization

After 7 Days of optimizations, the 28 Days FCP reported by the CrUX API is already **202 ms faster**.

Sep 7 2022
CrUX API



CrUX FCP
2,100 ms

21 Days
Not Optimized

7 Days
Optimized

-202 ms

28 Days after Optimization

Extrapolated to 28 Days of Optimization (7 Days Uplift * 4) the FCP is expected to become **806 ms faster** compared to the FCP before the optimization.

Sep 28 2022
FORECAST

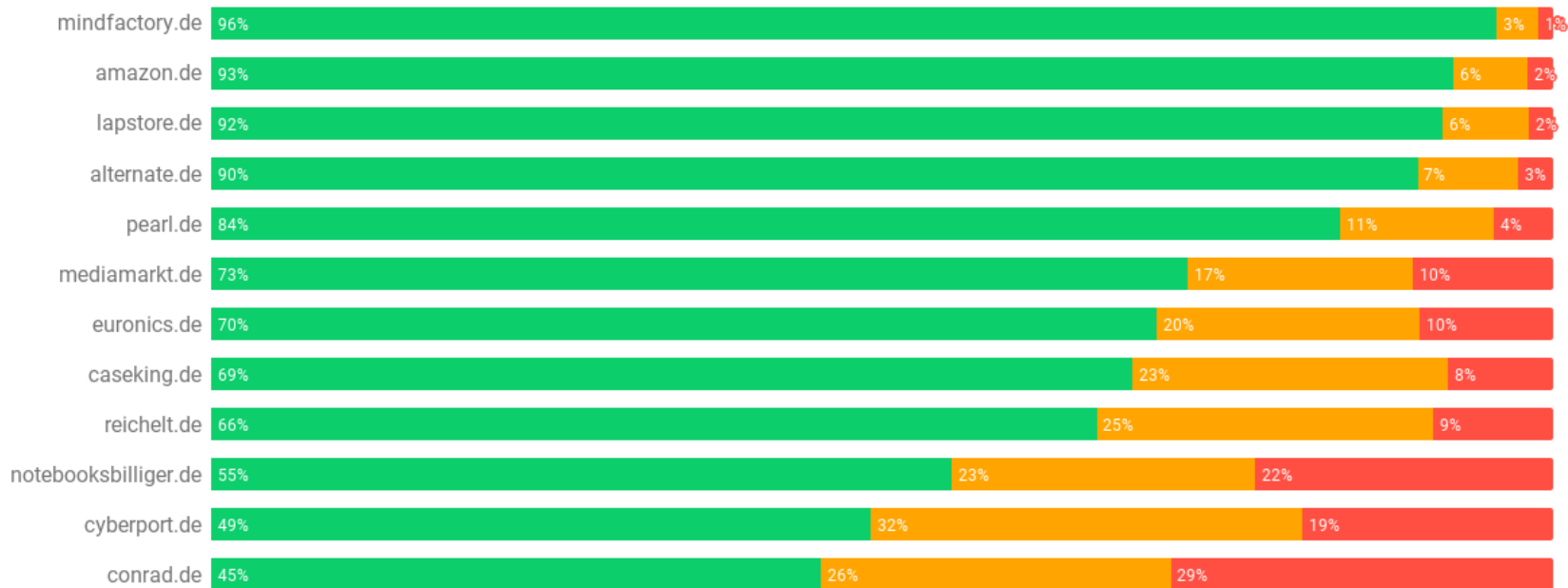


CrUX FCP
1,495 ms

28 Days
Optimized

-806 ms

Rank Your Score & Forecast With Competitors!



Roadmap Features



Dashboard KPI Overview

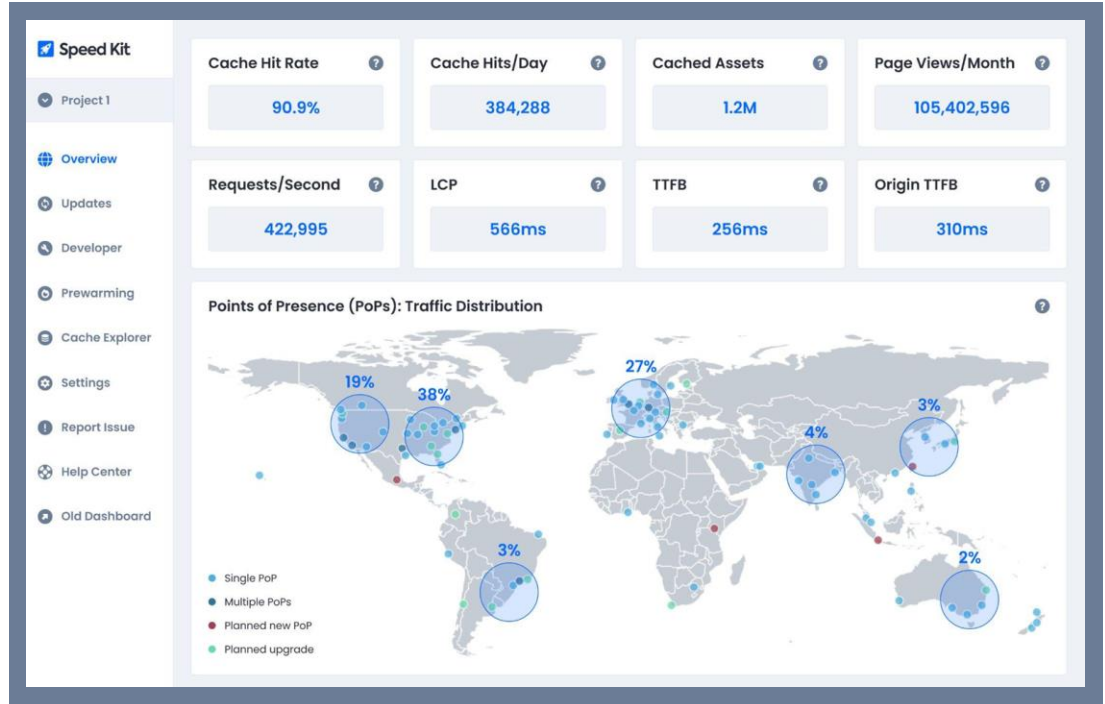
Purpose

Simple overview of the technical Speed Kit KPIs

Approach

The dashboard will provide a view to display the technical KPIs:

- See your current traffic usage and how many cache hits you have
- See how the performance of your site behaves at a single glance



Speed Analytics Dashboard

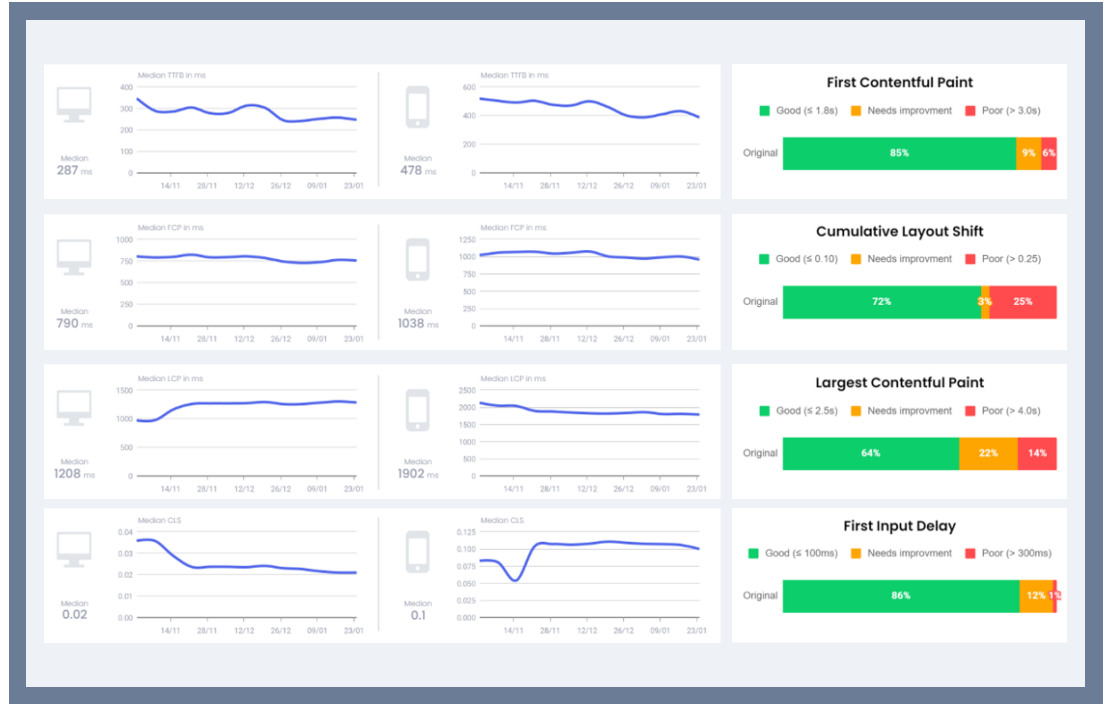
Purpose

A simple overview of the main performance metrics

Approach

An easy-to-use performance dashboard based on monthly RUM data:

- Status quo and over time development of core web vitals
- Drill-down by origin, device and page types
- Can be extended by SQL workbench for drill-downs



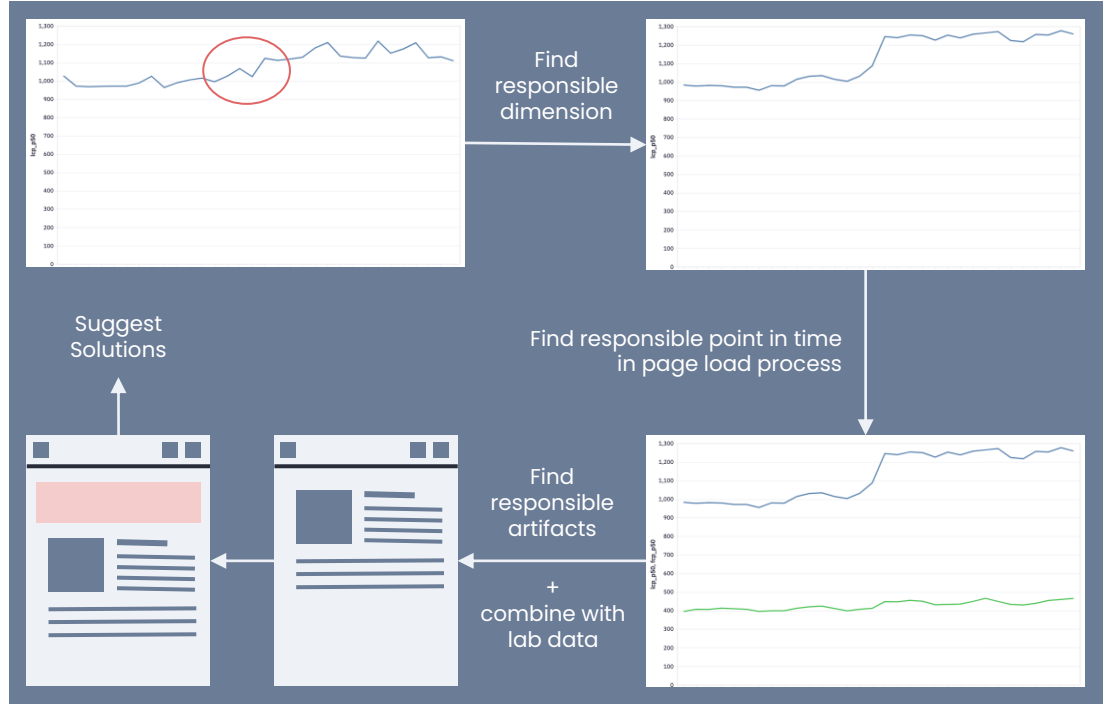
Speed Guard

Purpose

Detect performance issues with exact cause and solution

Approach

1. Monitor performance and detect anomalies
2. Find cause by automated dicing and combining RUM and lab data
3. Suggest solution based on caused and web performance best practices



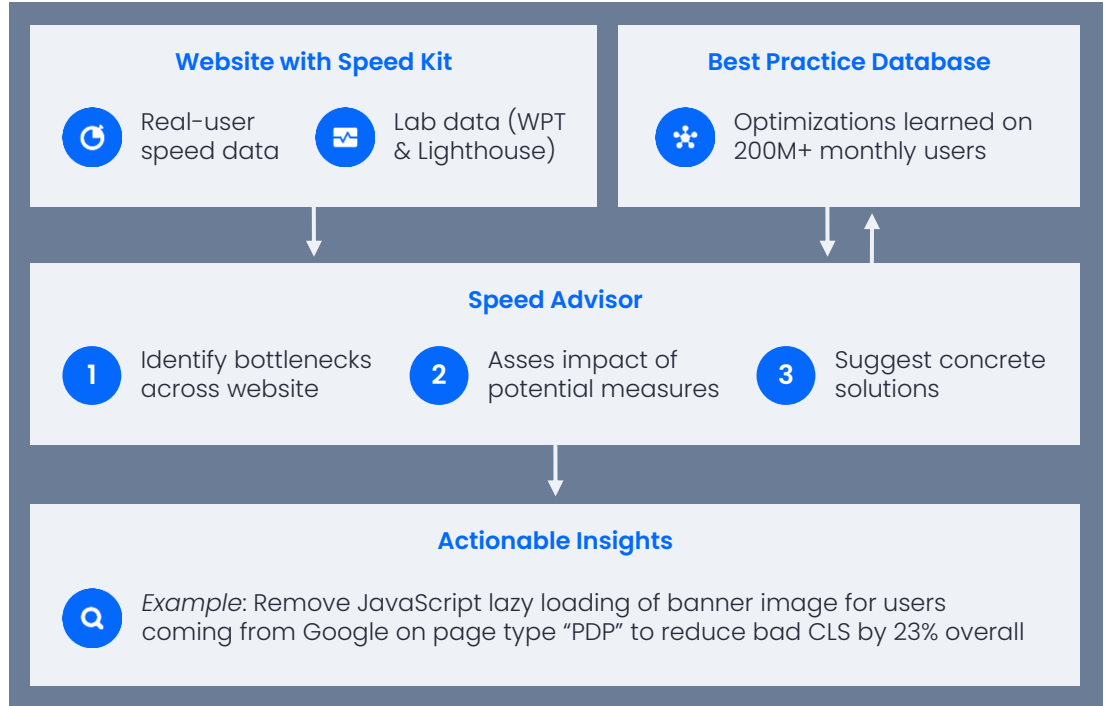
Speed Advisor


Purpose

Suggested speed improvements beyond Speed Kit

Approach

Speed Kit's real-user monitoring and synthetic testing collect detailed speed data. The advisor analyzes various dimensions (e.g. browsers) and metrics (e.g. CLS) and suggest applicable optimizations.



 Wolfram Wingerath, Felix Gessert, et al.: Speed Kit: A Polyglot & GDPR-Compliant Approach For Caching Personalized Content, ICDE (2020).

Join Baqend's Mission of Building a Faster Web



Product Integration

Team Lead, Seniors & Juniors



Performance Engineering

Seniors & Juniors



Fullstack Development

Seniors & Juniors



Data Engineering

Seniors & Juniors



www.speedkit.com/careers

Thanks!

Questions?



research@baqend.com

baqend.com/publications