Not Your Father's Data Stack

Batch-Free Tracking & Analytics For 100+ Million Users

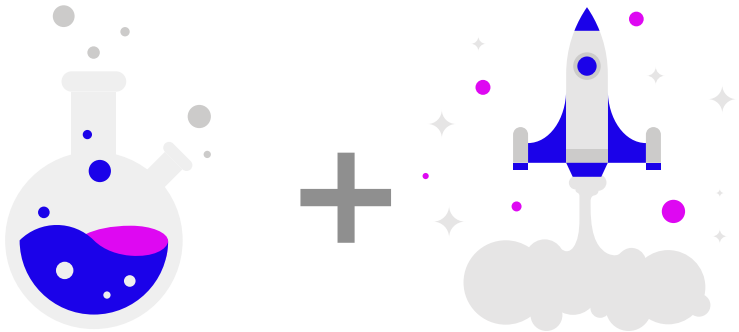
DBDC / Munich, Germany
Wolfram Wingerath
Dec. 8, 2020

Baqend

Who is This Guy !?

Wolle
Data Engineering

Research:
- Web Caching
- Real-Time Databases
- NoSQL & Cloud Systems
- ...



Practice:
- Website Acceleration
- Real-User Monitoring
- ...

Baqend

Table of Contents


Why Do Businesses
Care About Web Performance?

You Heard The Stories

amazon — **100 ms** slower → **-1%** Conversion Rate

zalando — **100 ms** faster → **+0.7%** Revenue Per Session

Walmart — **100 ms** faster → **+1%** Revenue

Greg Linden. Make Data Useful. Stanford Data Mining Class CS345A, 2006

Shuhei Kagawa, Jeff Cybulski, David Martin Jones, et. al.. Loading Time Matters. Zalando Tech Blog, 2018

C. Crocker, A. Kulick, B. Ram. Real-User Monitoring at Walmart, SF & SV Web Performance Group, 2012.

You Heard The Stories

amazon

100 ms slower → 1% Conversion Rate

PageSpeed

=

zalando

100 ms faster → +0.7% Revenue Per Session

Money

Walmart

100 ms faster → +1% Revenue

Greg Linden. Make Data Useful. Stanford Data Mining Class CS345A, 2006

Shuhei Kagawa, Jeff Cybulski, David Martin Jones, et. al.. Loading Time Matters. Zalando Tech Blog, 2018

C. Crocker, A. Kulick, B. Ram. Real-User Monitoring at Walmart, SF & SV Web Performance Group, 2012.

Delay Psychology : Rules of Thumb

Delay	User Perception
0 – 100 ms	Instant
100 – 300 ms	Small perceptible delay
300 – 1000 ms	Machine is working
1+ s	Mental context switch
10+ s	Task is abandoned

Stay under 1000 ms to keep users' attention

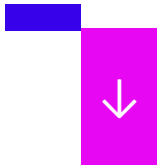
I. Grigorik, High performance browser networking, O'Reilly Media, 2013

Jakob Nielsen, Usability Engineering, Morgan Kaufmann, 1994

Speed in Corona Crisis: Why Now

Congested Networks

Traffic spiked by 32% to 109%, download speeds dropped **up to 35%.** [1]



E-Commerce Performance Drop

85% of shops have seen a steep **performance decline** since January. [2]

Anxious User Behavior

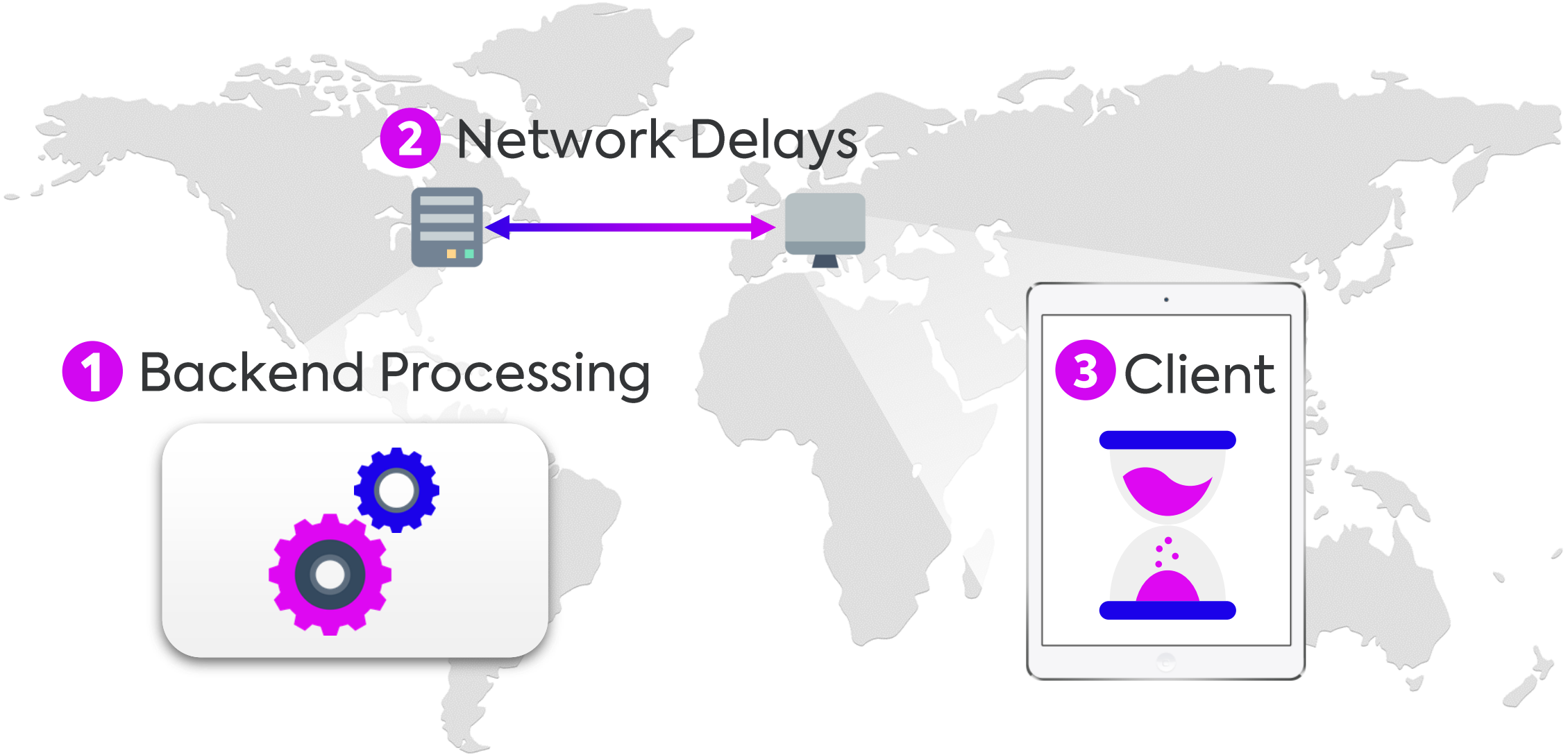
When under pressure, **44% of users** perceive websites to be **slower.** [3]

Artur Bergman & Jana Iyengar. How COVID-19 is affecting internet performance. Fastly Blog, 2020

Felix Gessert. E-Commerce Loading Times in the COVID-19 Crisis. Baqend Blog, 2020

Brain Food! Vol. 3. Speed Matters: Designing For Mobile Performancence. Awwwards.org, 2019

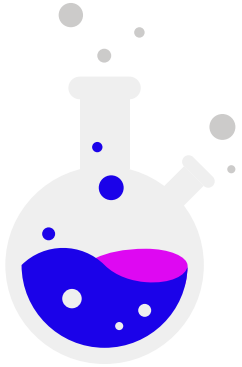
3 Things Make Your Website Slow



We Are Baqend

We bring performance research to practice.

40+ man-years of **web performance research**



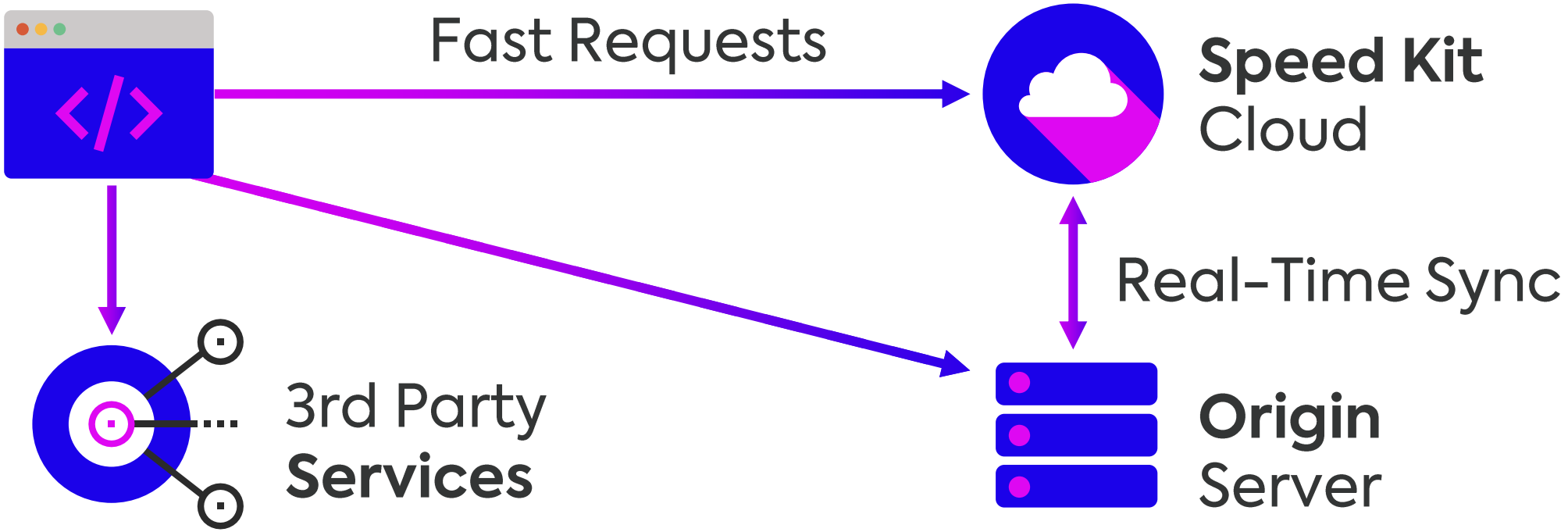
Novel technology for caching **dynamic data**

Speed Kit – SaaS for e-commerce speed


We Kill
Load Times

Speed Kit Makes Websites Fast

Website

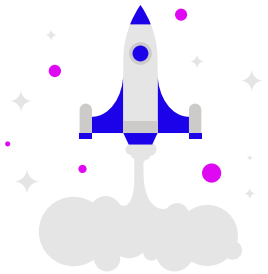


Measuring the Uplift – With SCIENCE

CDNs, Manual Optimizations

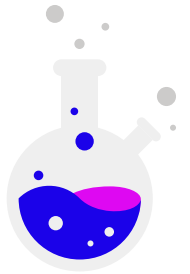
- Only before-after comparison

Speed Kit



- Statistically sound **split testing**
- Clean measurement of performance & business uplifts

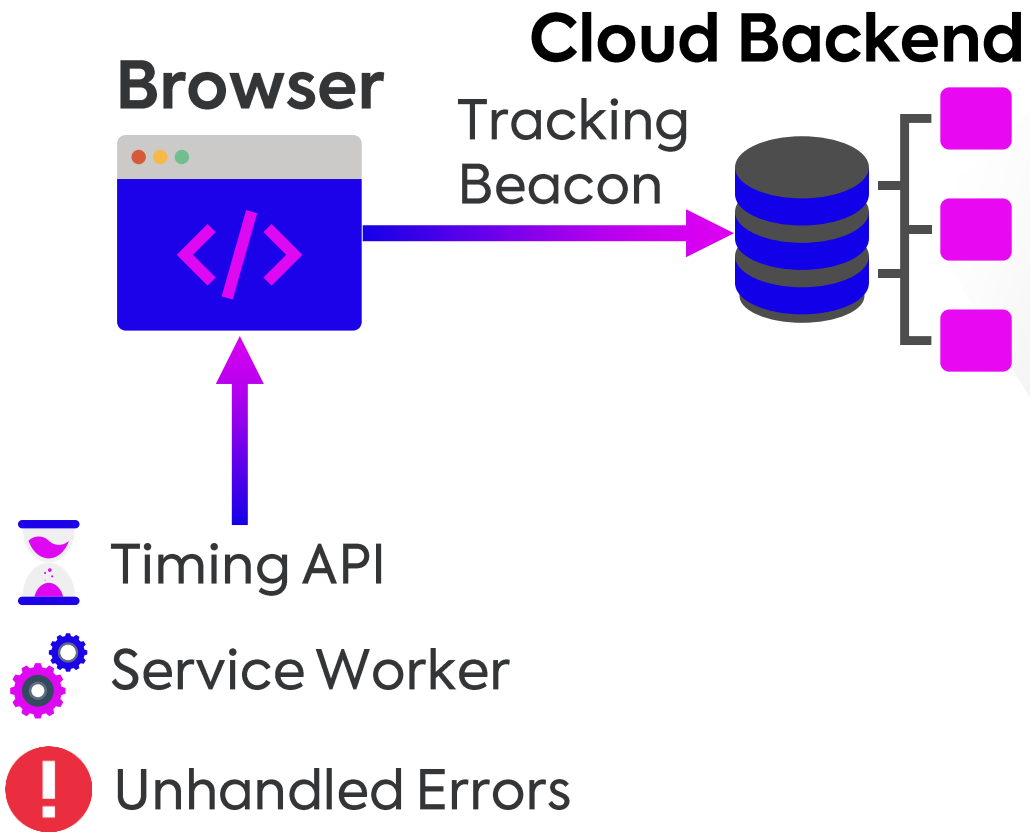
Application Features



- Measurable business impact through A/B tests


How Do We Measure
Web Performance?

Goal: Performance & Business Insights



▪ Time-to-First-Byte ▪ First (Contentful) Paint ▪ DOM Timer ▪ First Input Delay	Performance
▪ Session Length ▪ Time on Site ▪ First User Interaction ▪ Bounce Rate	User Engagement
▪ Cart Size ▪ Transactions ▪ Conversion Rate ▪ Revenue	Business KPIs
▪ Page Views & Sessions ▪ Browser Distribution ▪ JavaScript Errors ▪ Caching Insights	QA Metadata

Real-User Monitoring (RUM)

Collection

Tracking (RUM)

Ingestion

S3

MongoDB

Kinesis

- Raw PI tracking & meta data
- Custom tracking

Analytics

Athena

Flink

Kinesis Data Analytics

- Materialized views & stream aggregations
- Historical data & real-time updates

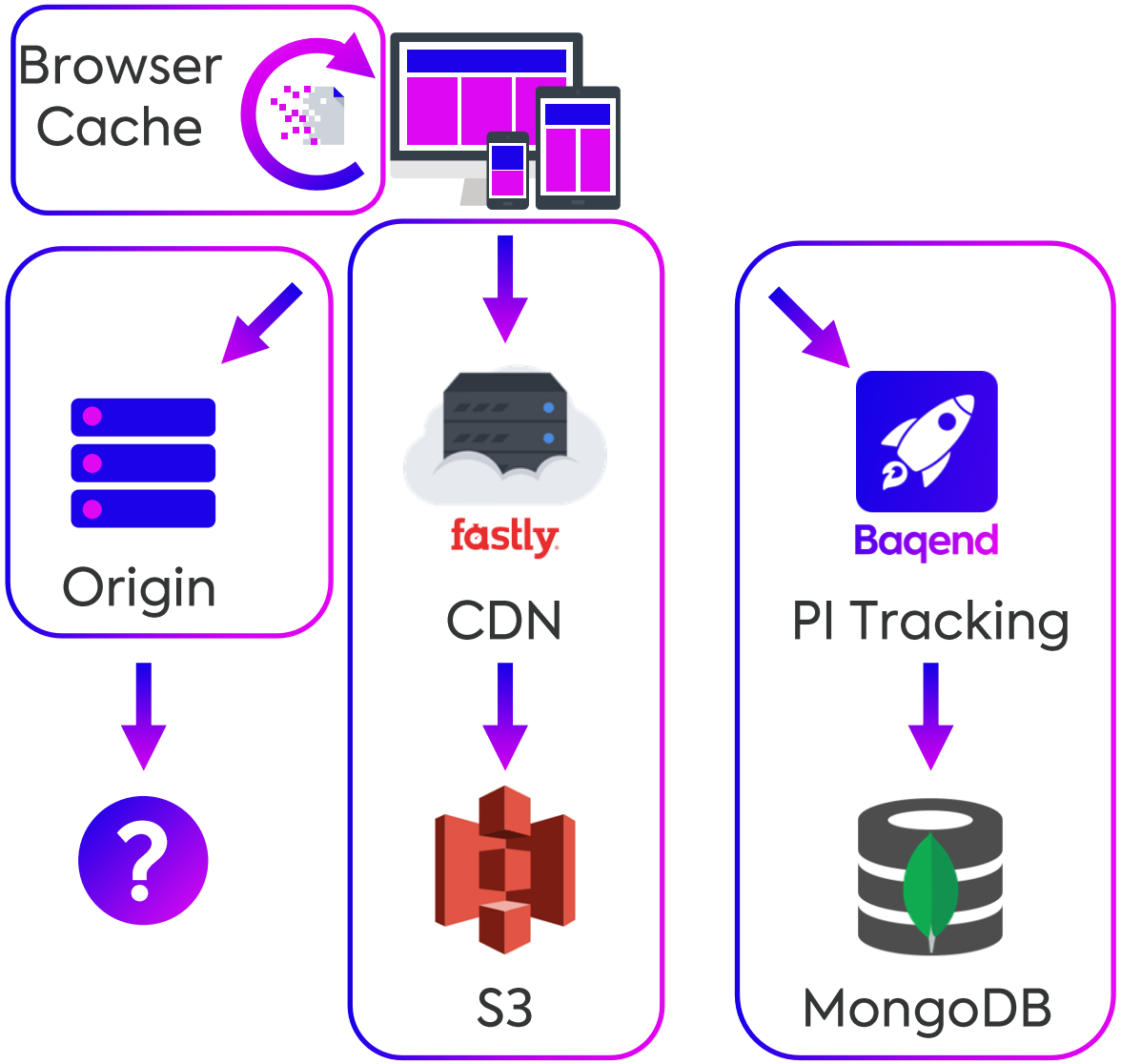
Reporting

Performance Dashboard

Real-Time Alerting

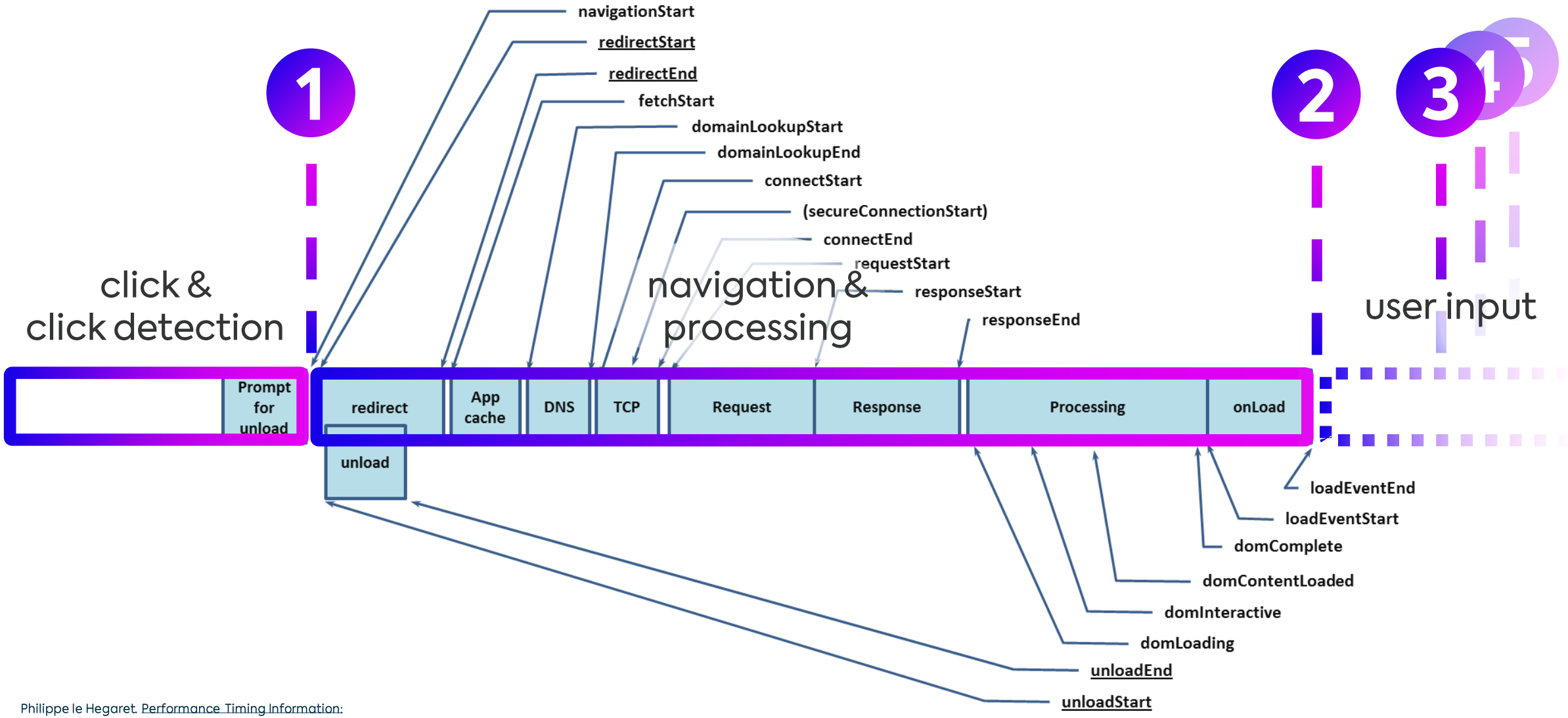
Custom Reporting

How to Collect the Performance Data?



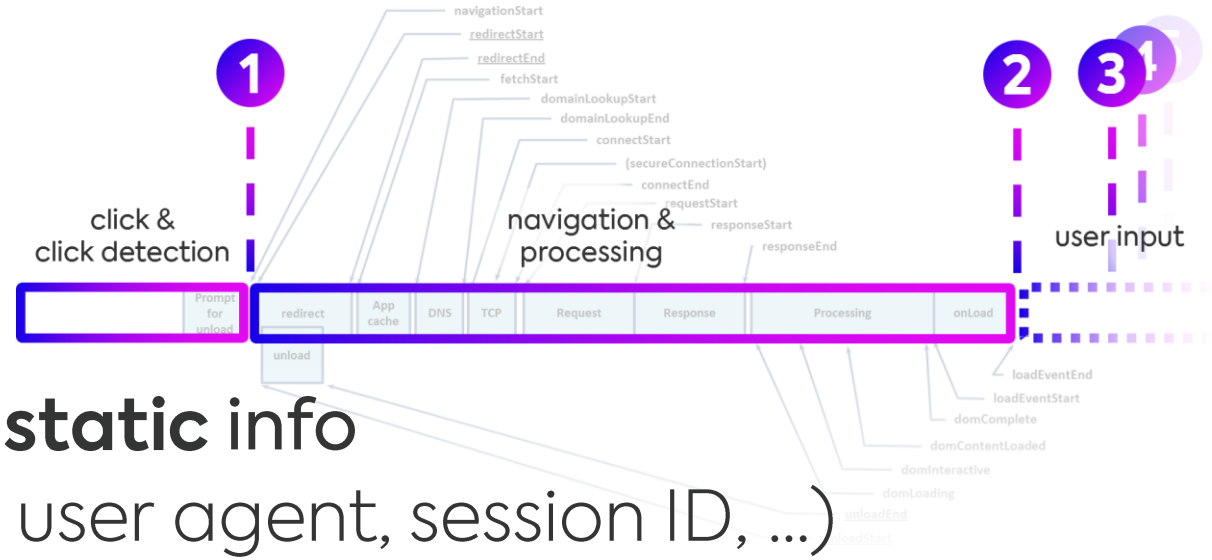
- Logging requests is not enough:
  - ✗ User? Rendering? …
  - ✗ Browser cache (invisible)
  - ✗ Origin requests (no logs)
  - ✓ CDN requests
- Solution: **Tracking every PI** (page impression)

When to Send Data Beacons?



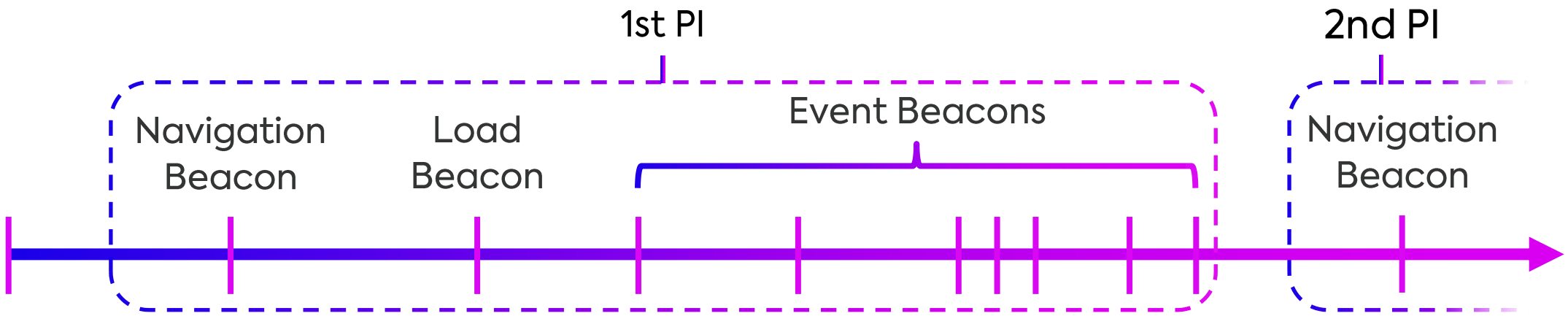
Philippe le Hegaret. Performance Timing Information: Part 1 – Navigation Timing, W3C Blog, 2012

Types of Data Beacons



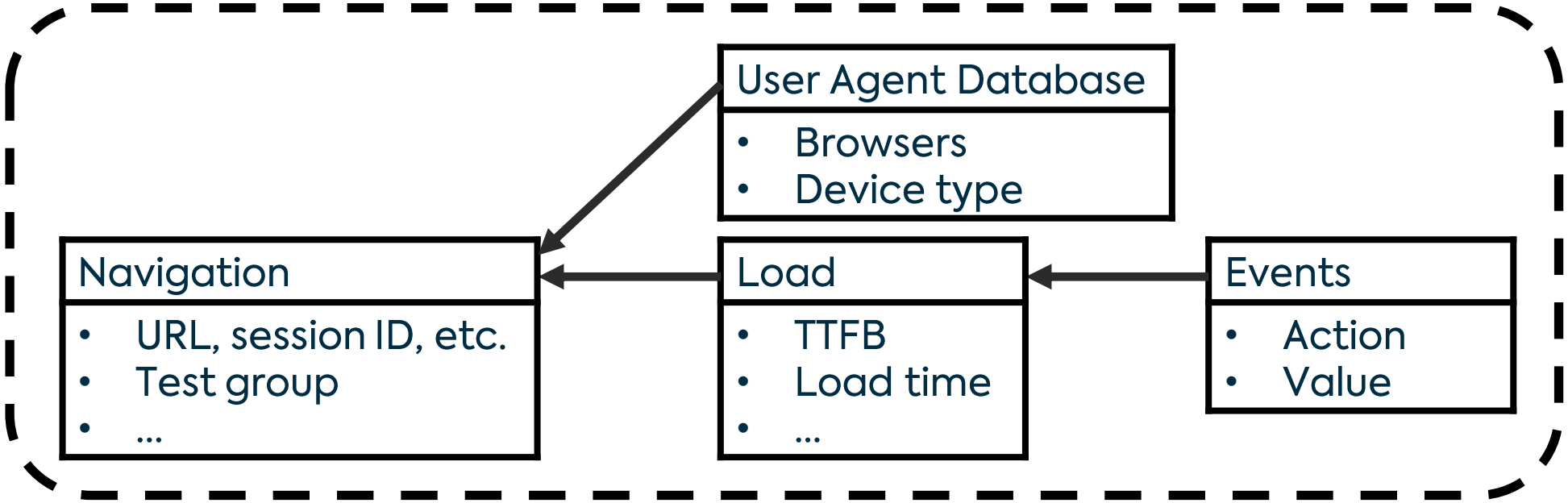
1. 1 for **static** info
   (URL, user agent, session ID, …)
2. 1 for **timings**
   (TTFB, load time, FCP, …)
3. 0–n for **events**
   (first input, add-to-cart, …)

Schema:PI



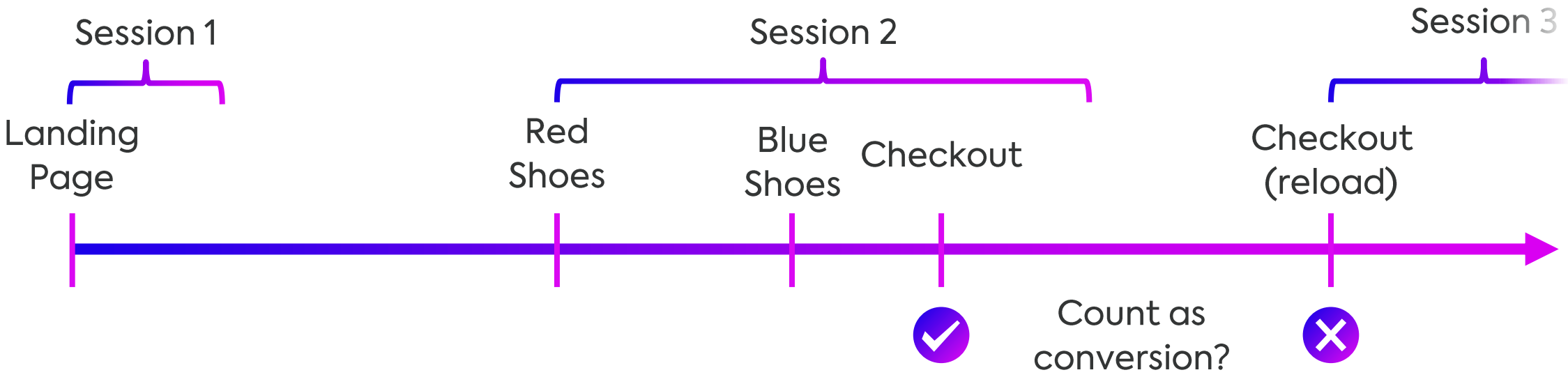
- **Beacon Join → PI**: How do we handle events that come late?
  - Simply wait 5 minutes?
  - Wait for next PI or session timeout?
  - ...?
- How to resolve **user agents**?

Schema:PI



- **Aggregate events**: collect all events per PI
- **Join 3 Collections**: put together PI from navigation/load/event beacons
- **Resolve User Agents**: derive browser, device, etc. from UA string

Schema: Sessions



- **Unique conversions**: remove phantoms
- **Session timeout** after 30 minutes of *inactivity*

MongoDB Aggregation Pipeline: Problems

Indexing

Queries over non-indexed attributes were infeasible

Runtime

Even with indexes in place, queries could take 30+ min.

Scalability

Queries got slower with increasing amounts of data

Complexity

MongoDB aggregation pipelines become sophisticated quickly


Fixing My Life With
~~Flex Tape~~ Athena

The „A“ Stands for „AWSome“

Desperate **attempt**:

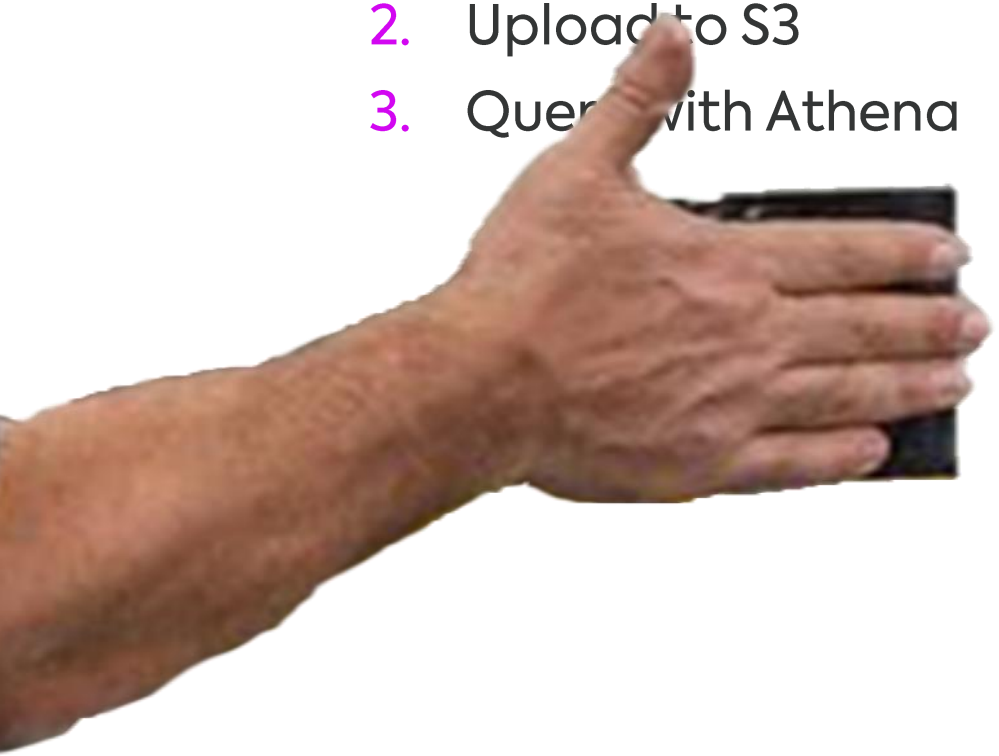
1. Dump MongoDB collection
2. Upload to S3
3. Query with Athena

- Typical analysis:
  - 1 equi-join
  - 3 mio. PIs
  - ~10 **seconds**

The „A“ Stands for „AWSome“

Desperate **attempt**:

1. Dump MongoDB collection
2. Upload to S3
3. Query with Athena

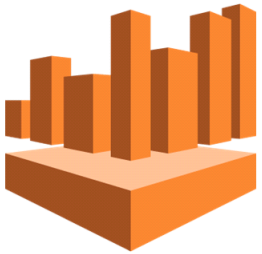


- Typical analysis:
  - 1 equi-join
  - 3 mio. PIs
  - ~15+ min.

The „A“ Stands for „AWSome“

✓ ~~Desperate **attempt**~~:New best practice:

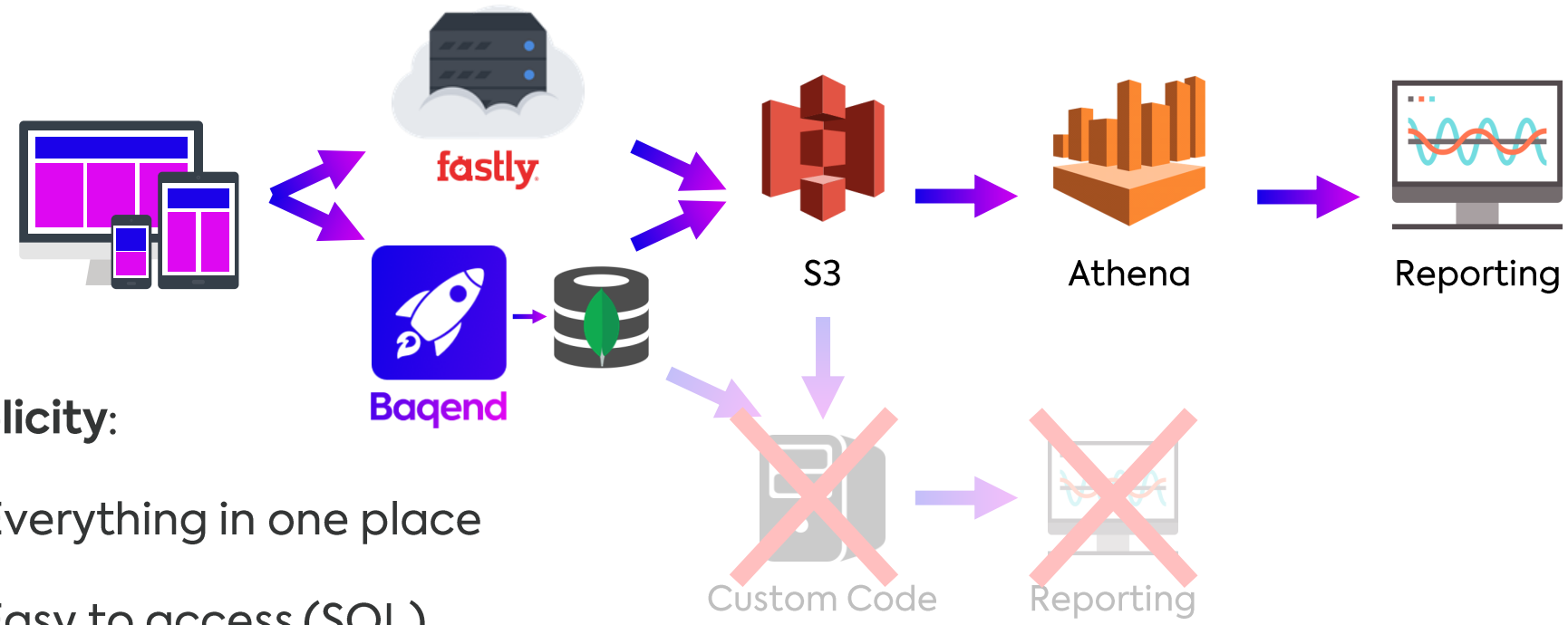
1. Dump MongoDB collection
2. Upload to S3
3. Query with Athena



AWS Athena

- Typical analysis:
  - 1 equi-join
  - 3 mio. PIs
  - ~10 **seconds**

Upgrading Our ETL Pipeline



- **Simplicity**:
  - Everything in one place
  - Easy to access (SQL)
- **Scalability & efficiency**:
  - Hundreds of gigabytes scanned in a query
  - Response time on the order of seconds

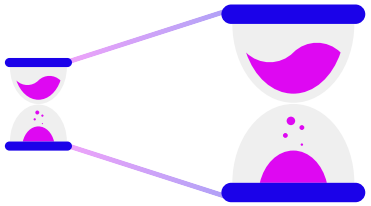

So Where is the
Problem?

Processing Stages & Latency

Alerting

- Simple metrics / little context
  - Counters
  - Extreme values
  - Specific errors

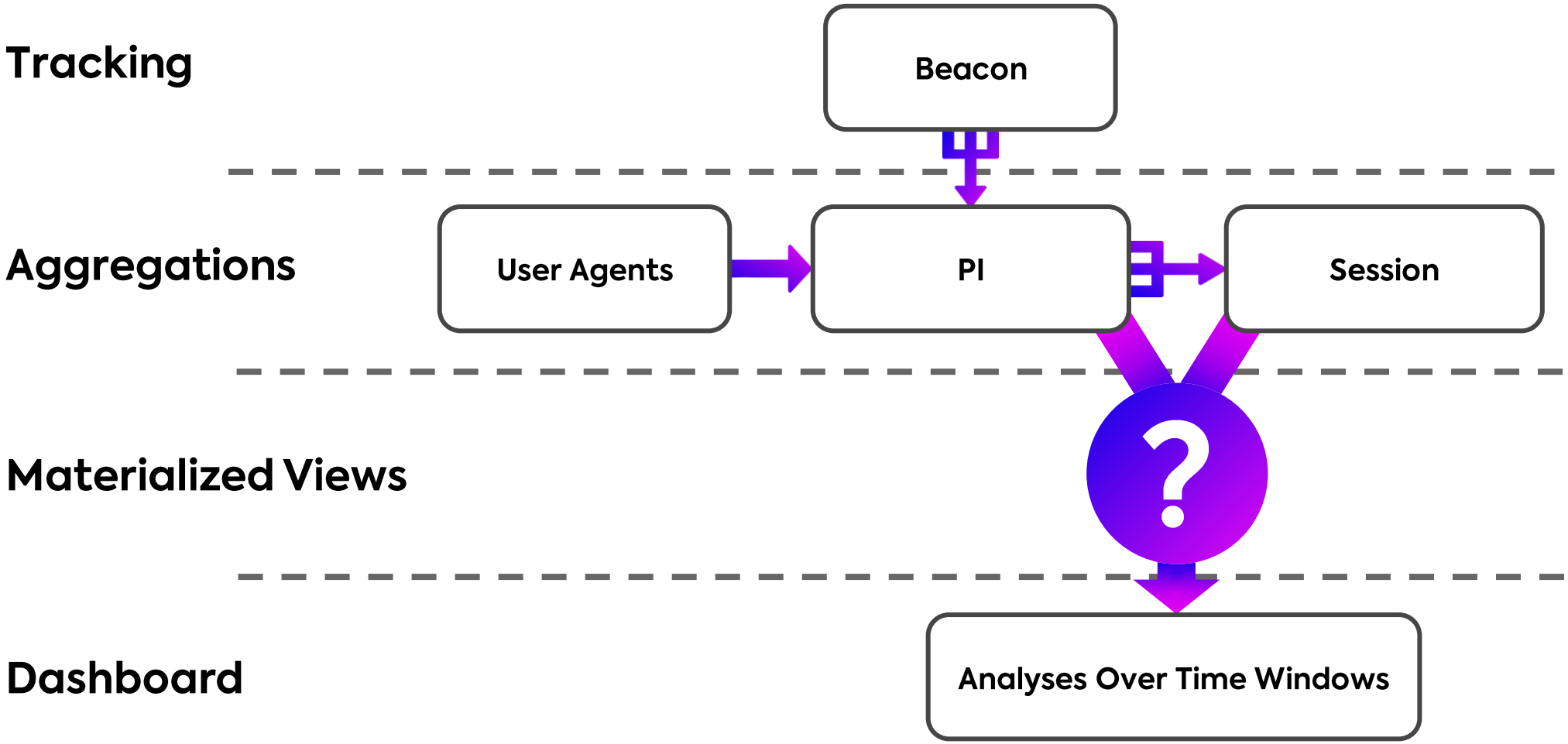
Processing Time



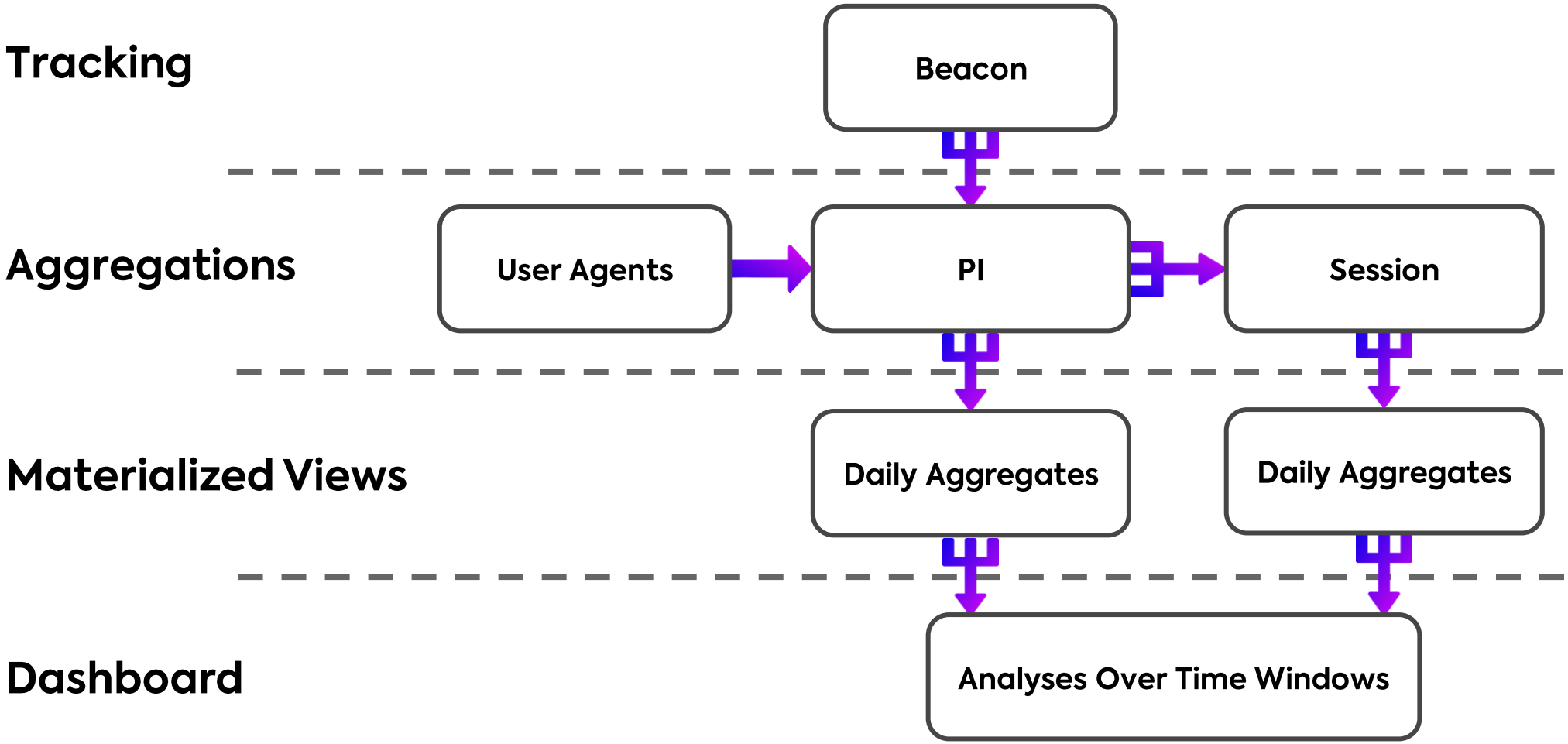
Trend Analysis

- Complex aggregations / huge time windows
  - Conversion rate
  - Performance by month
  - Seasonal effects

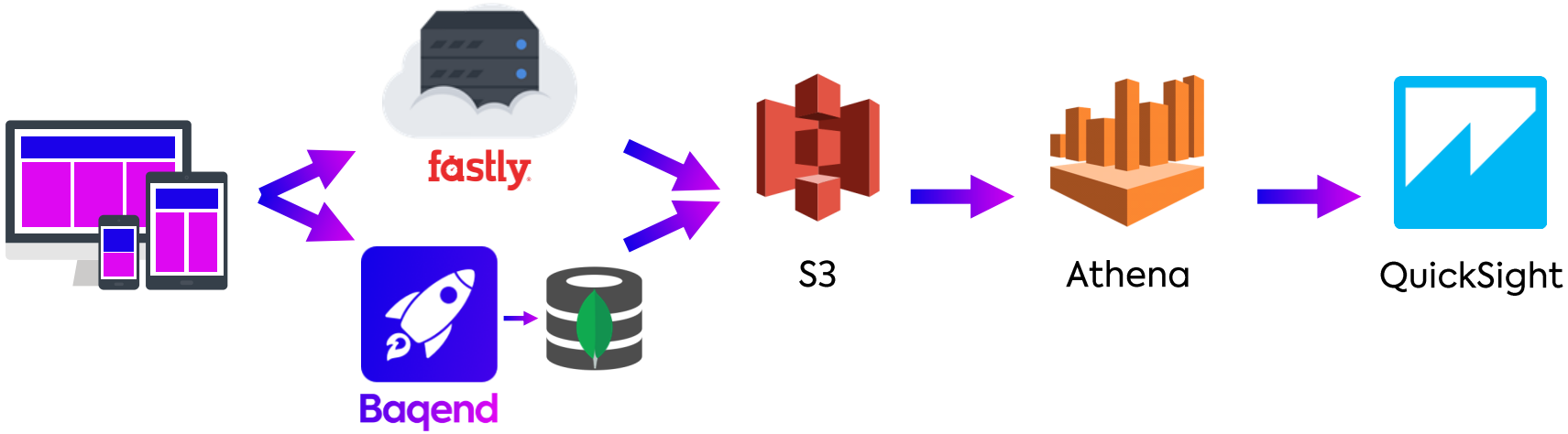
Schema Overview



Schema Overview

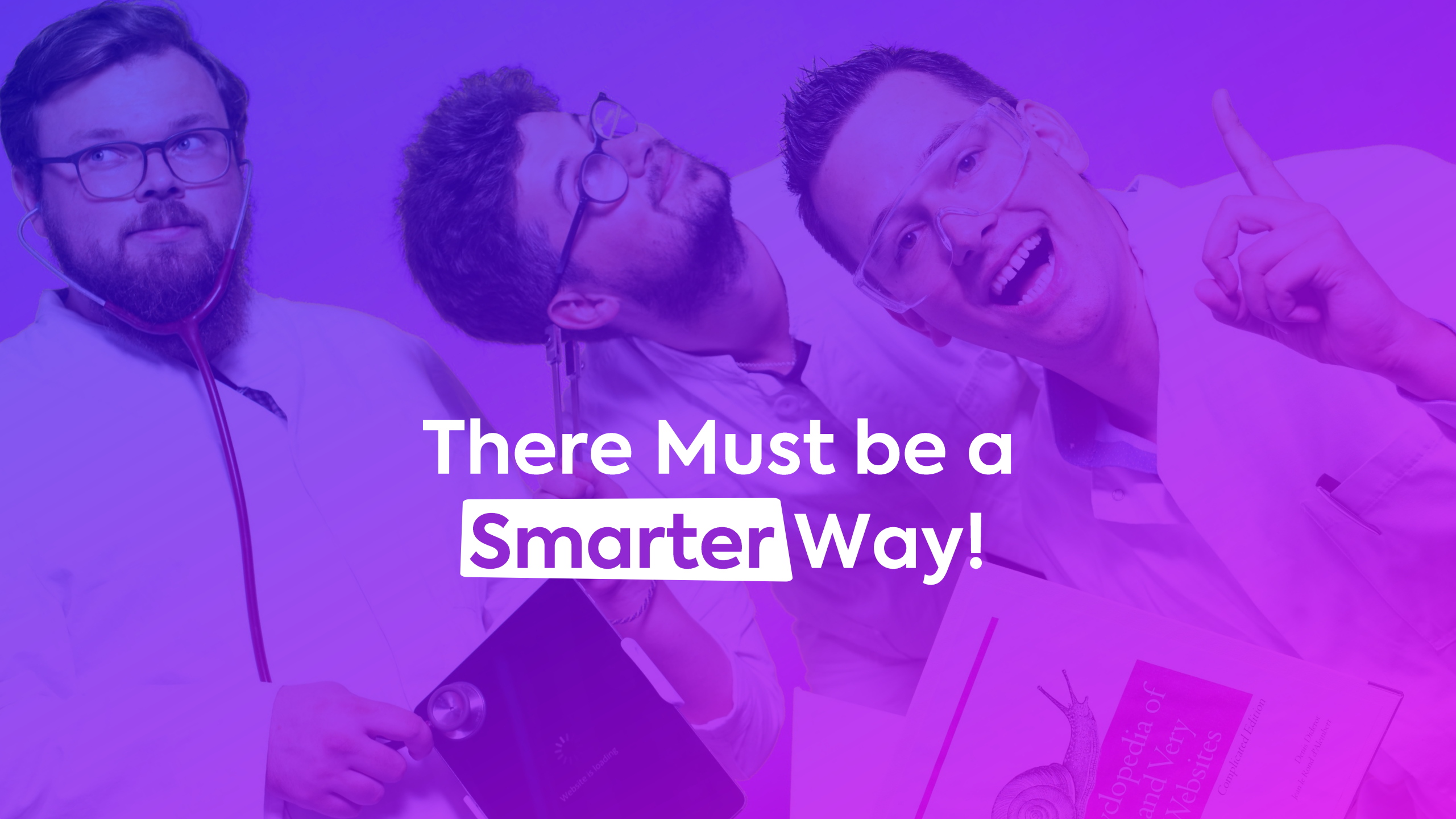


Our Batch Analytics Tech Stack

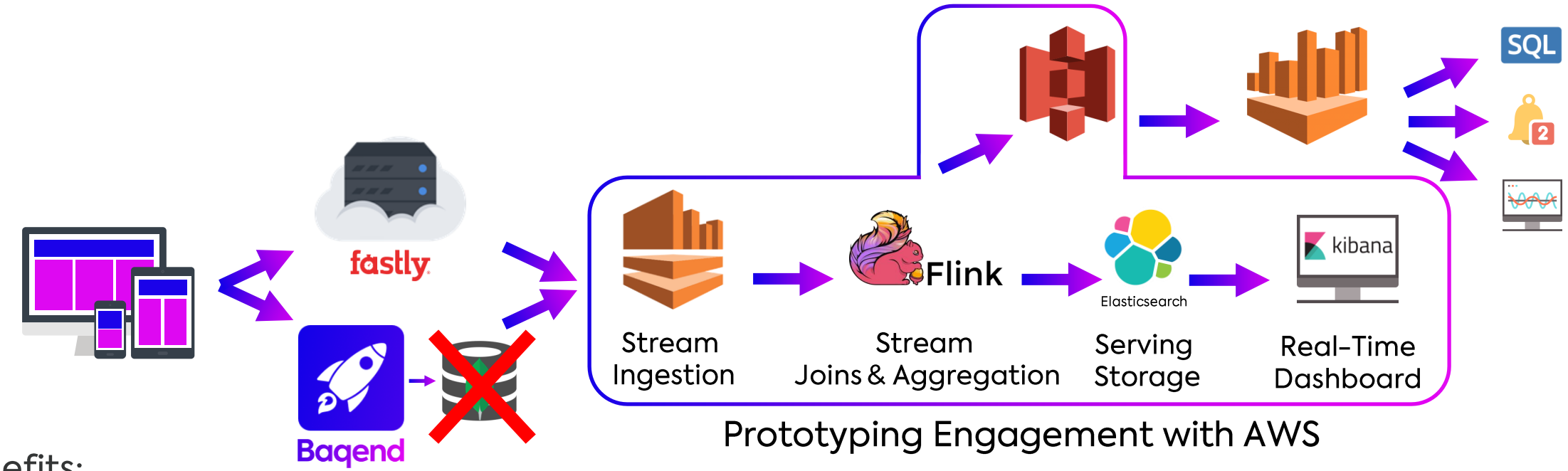


Issues:

- ✗ Many joins → slow queries
- ✗ 90 minutes discovery time
- ✗ No continuous dashboard (daily materialization)


There Must be a
Smarter Way!

Early 2020: AWS Prototyping

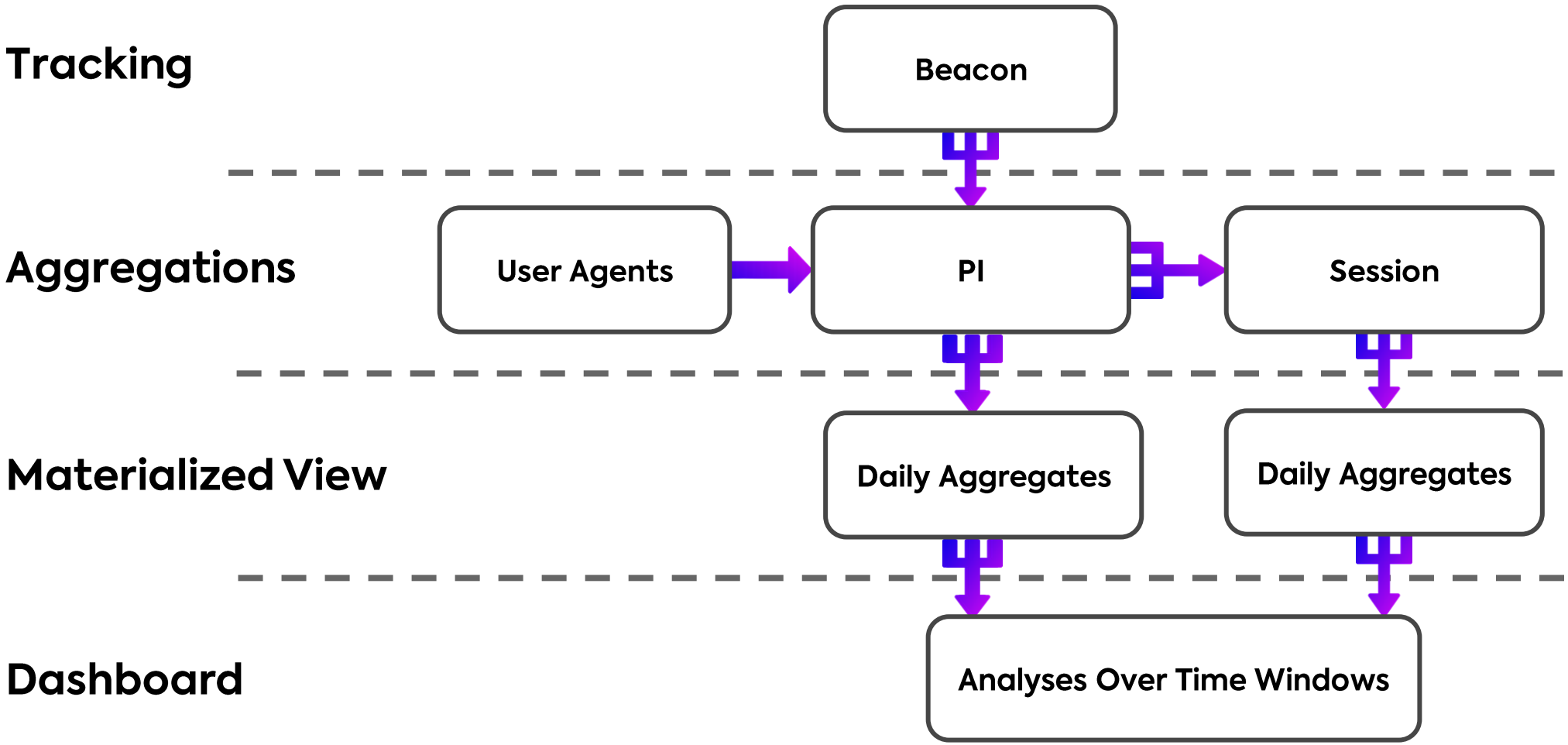


Prototyping Engagement with AWS

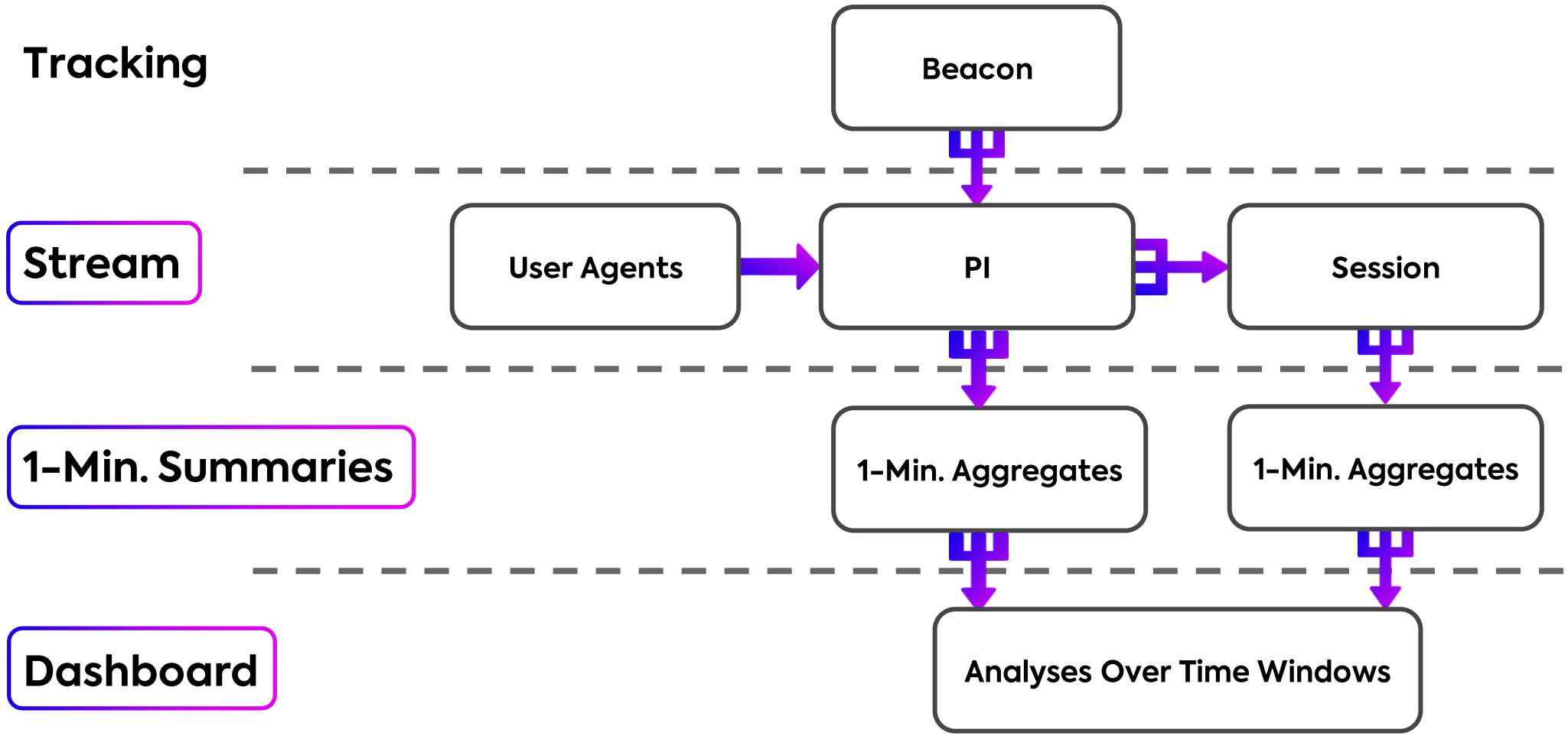
Benefits:

- ✓ No legacy tech → stability & efficiency
- ✓ Faster ingestion → Live performance charts
- ✓ Fewer joins → faster analytics

Old & Lame Schema



Shiny & New ~~Old & Lame~~ Schema



3 Levels of Aggregation

Page Impressions
Stream

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

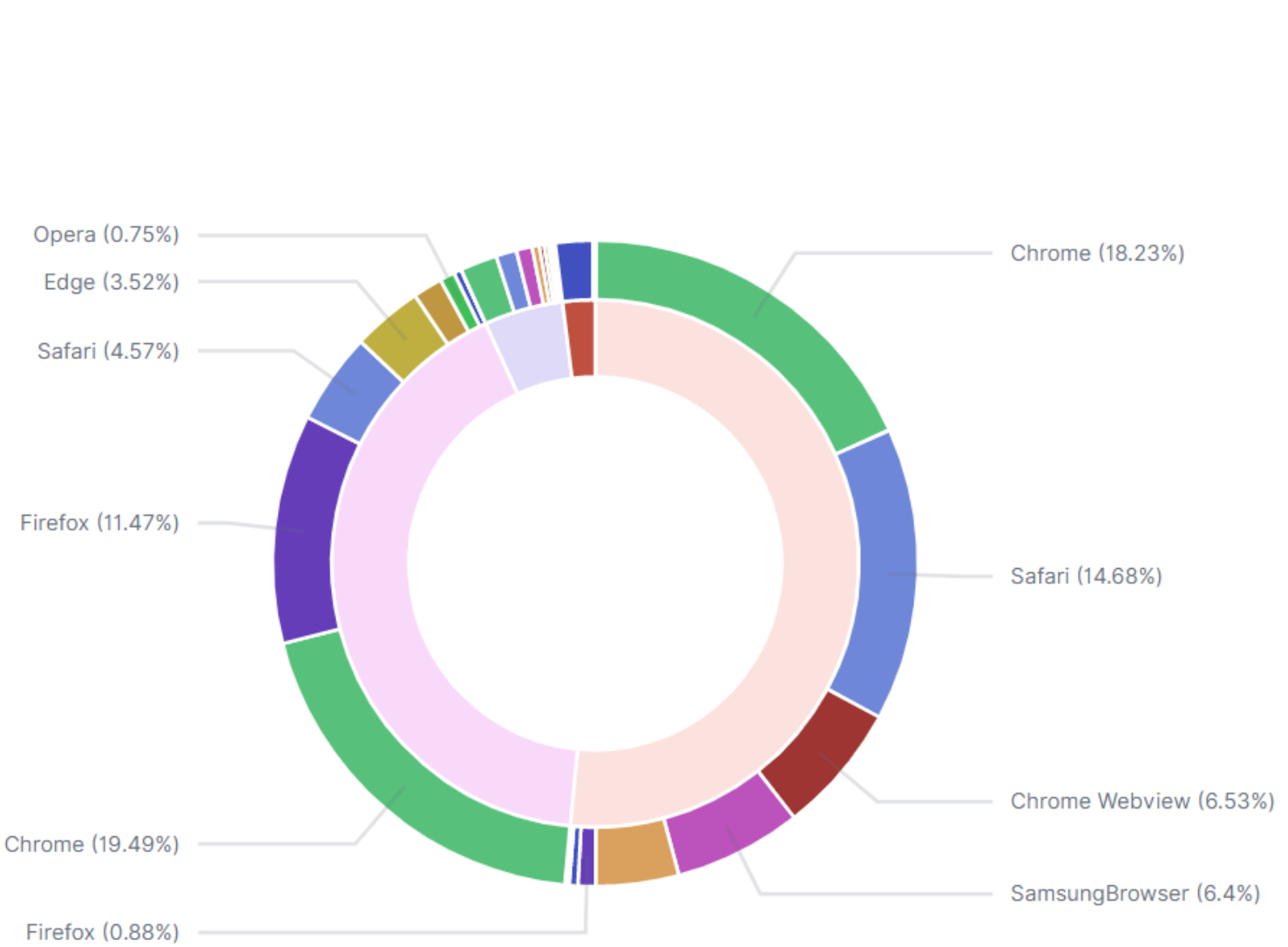
1-Min. Summaries
Elasticsearch

	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}

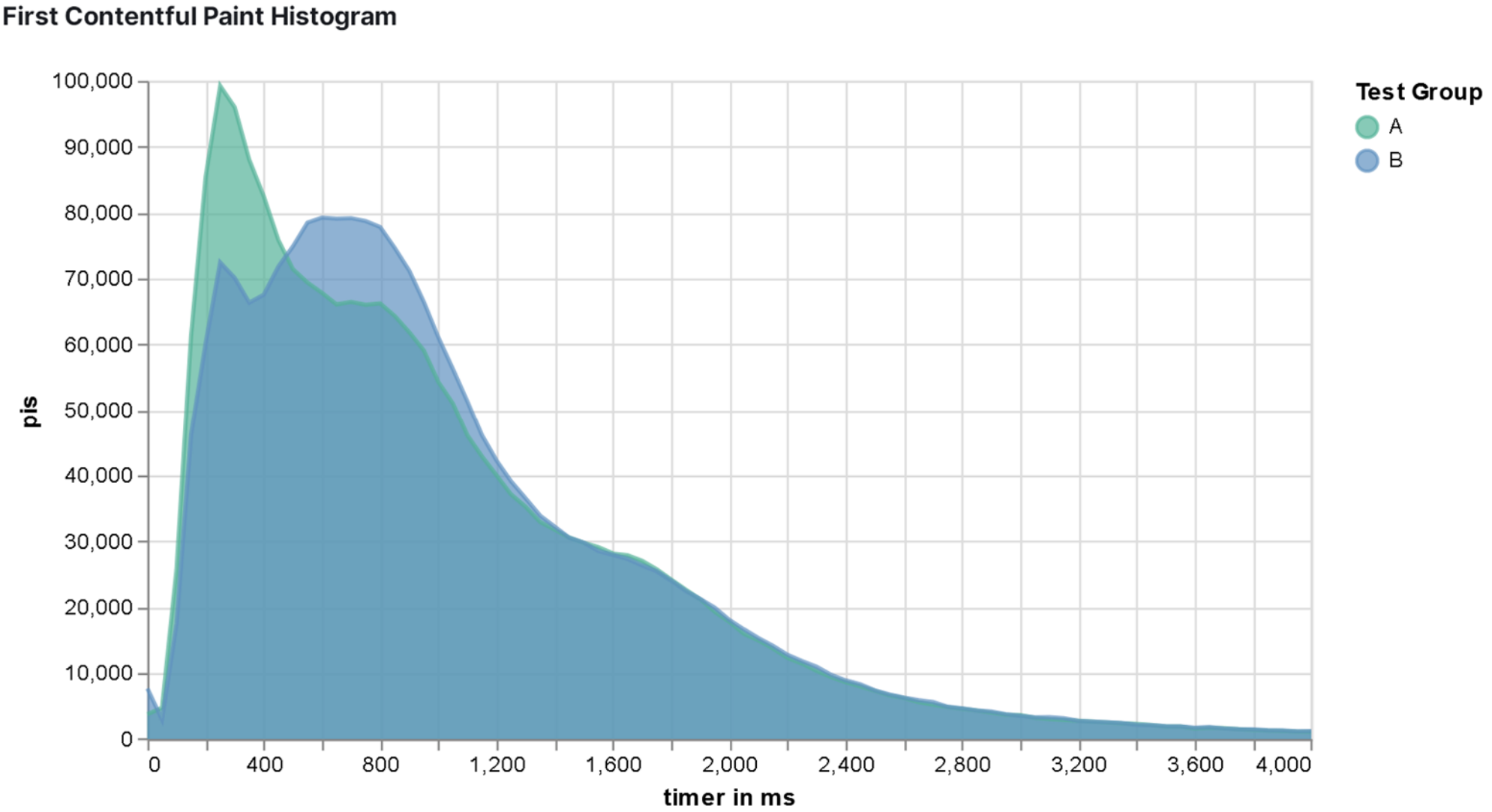
Arbitrary Time Windows
Dashboard

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

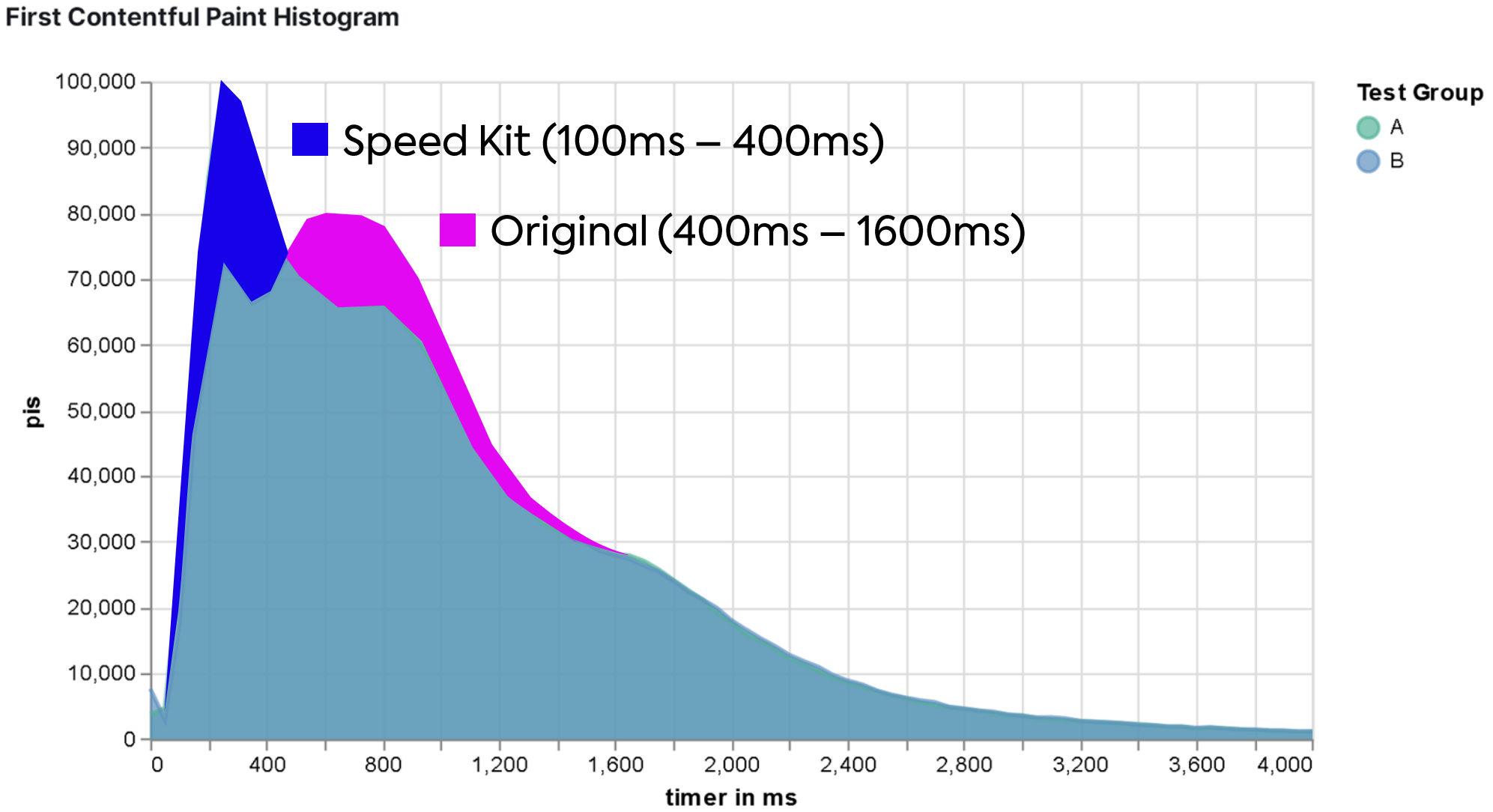
Kibana Dashboard: User Group Distribution



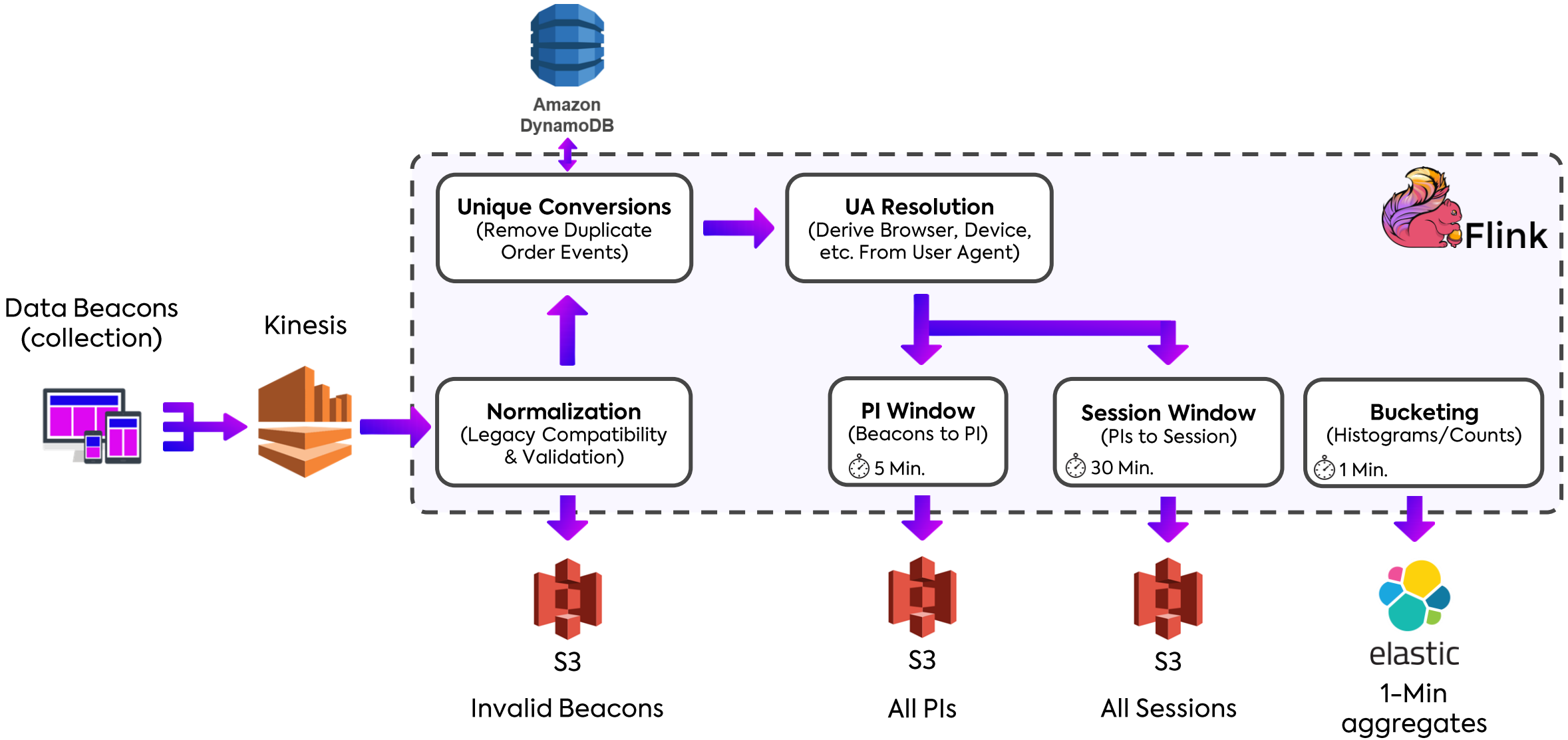
Kibana Dashboard: Performance Uplift



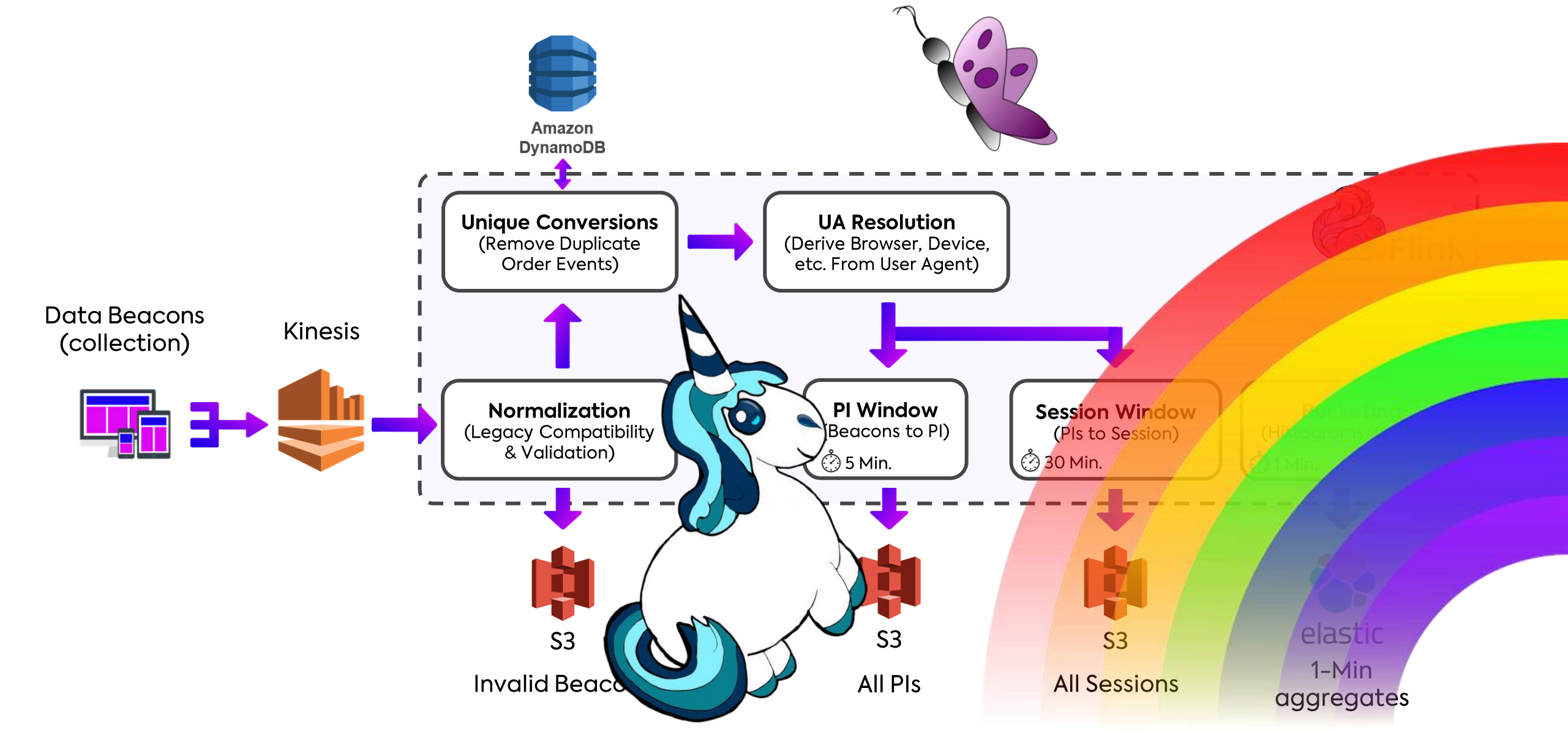
Kibana Dashboard: Performance Uplift



Zero-Latency Analytics



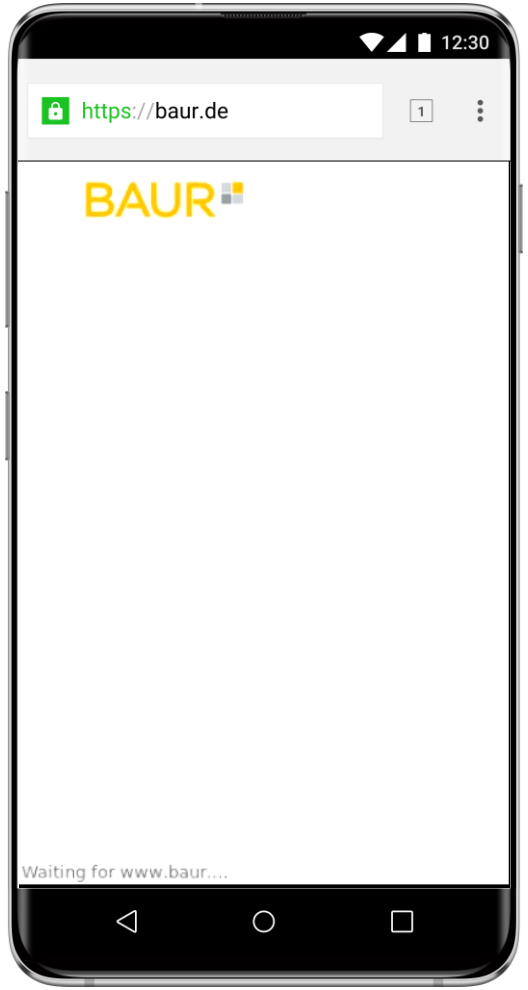
Zero-Latency Analytics



Summary & Outlook

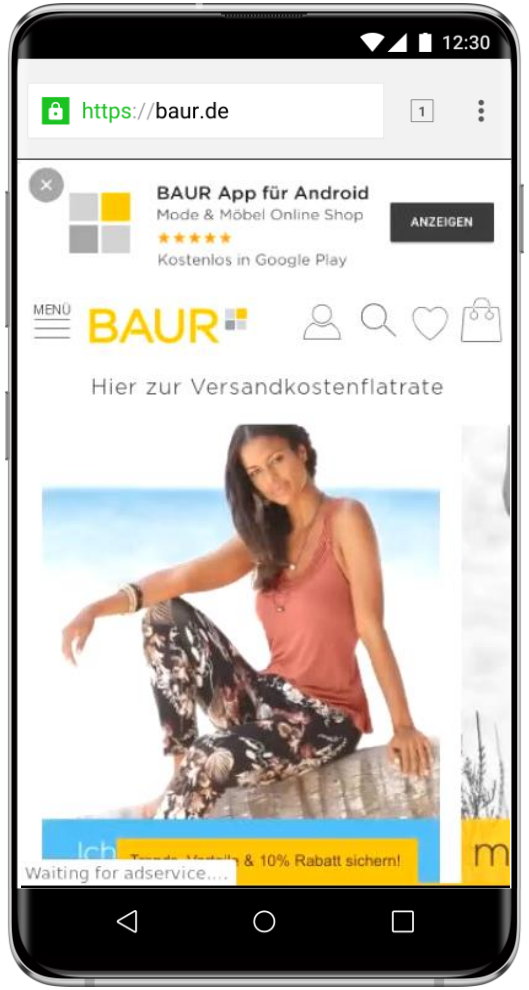
Baur.de: Speed Kit Acceleration

Before
Speed Kit

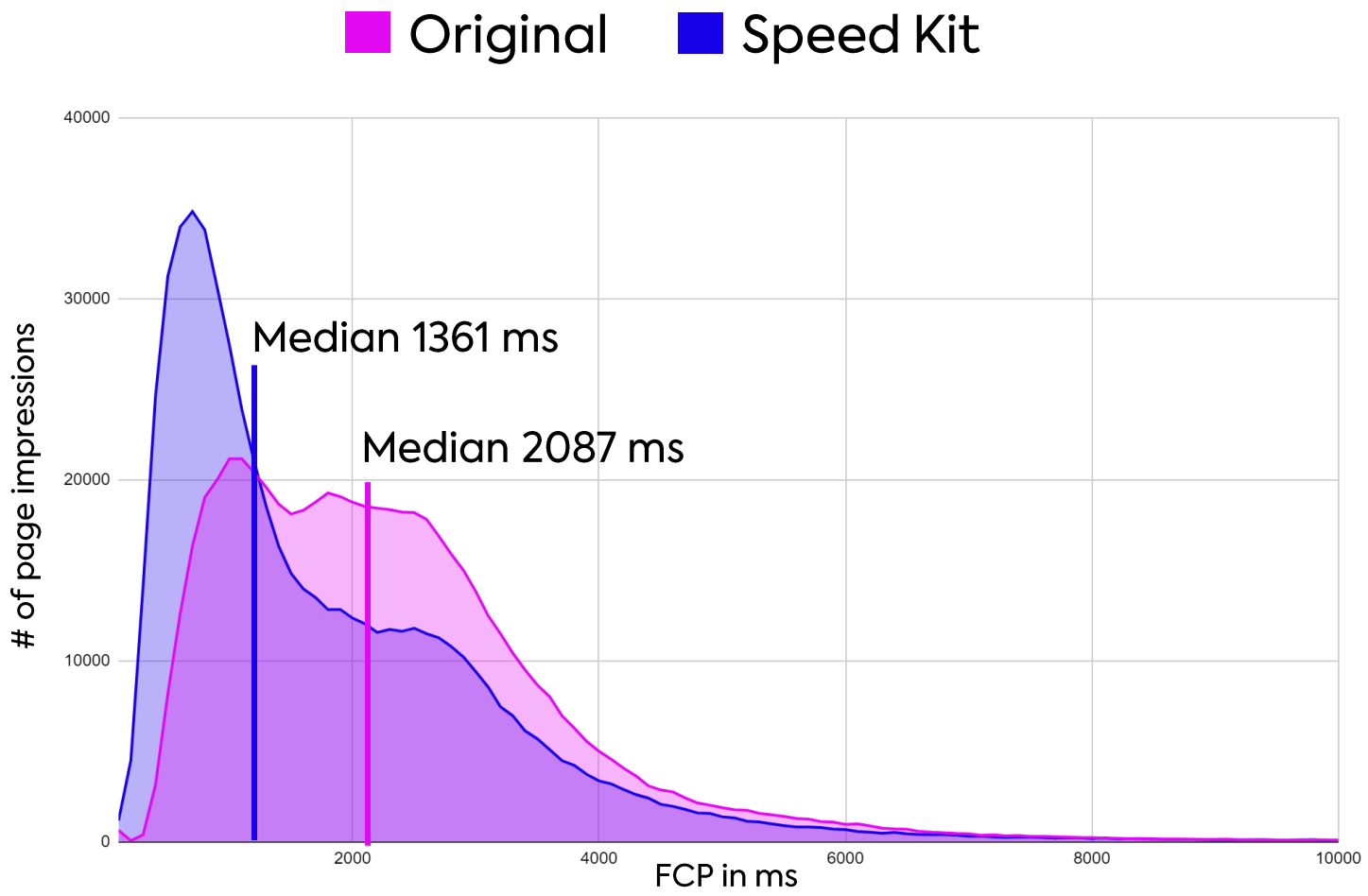


1.5x faster

After
Speed Kit



Baur.de: Overall Performance Uplift

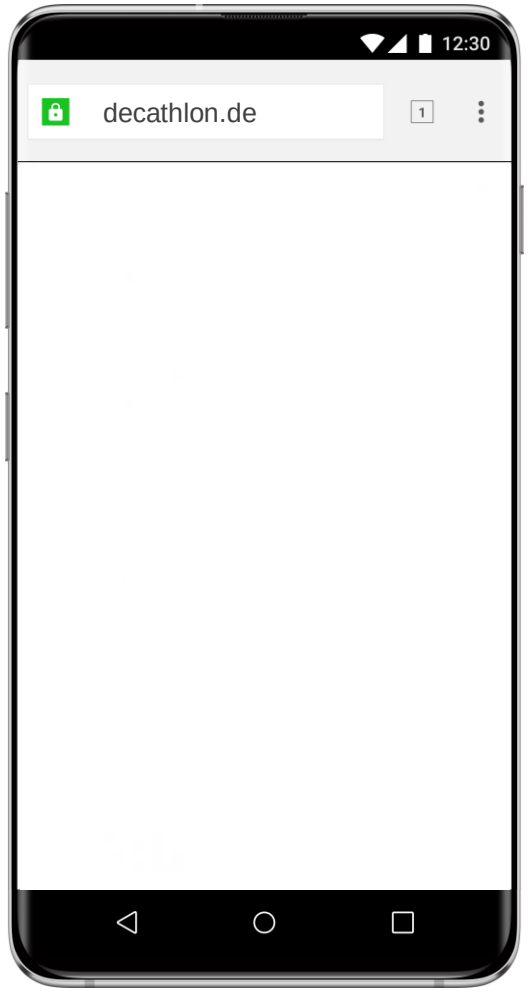


First Contentful Paint Histogram

*Histogram of first contentful paint on PDV pages compared between the two A/B test split groups

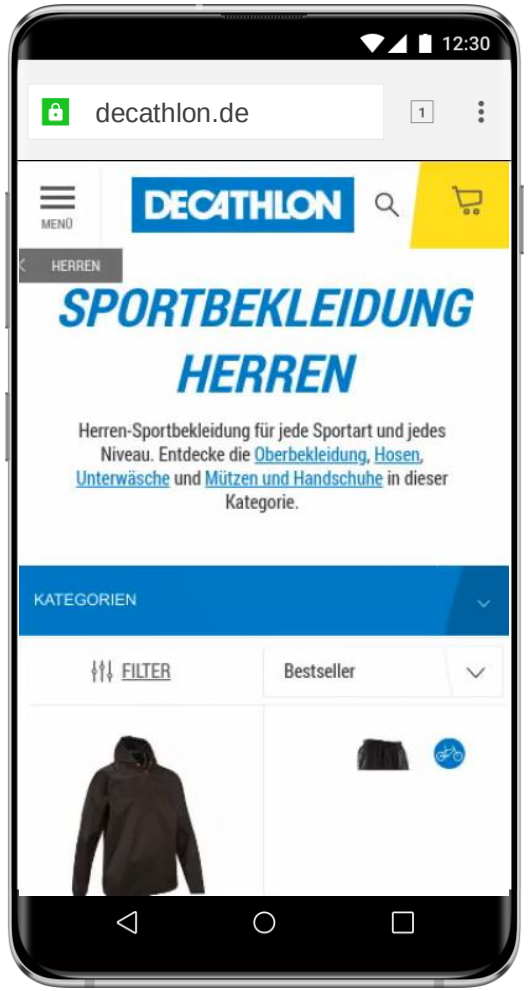
Decathlon.de: Speed Kit Acceleration

Before
Speed Kit

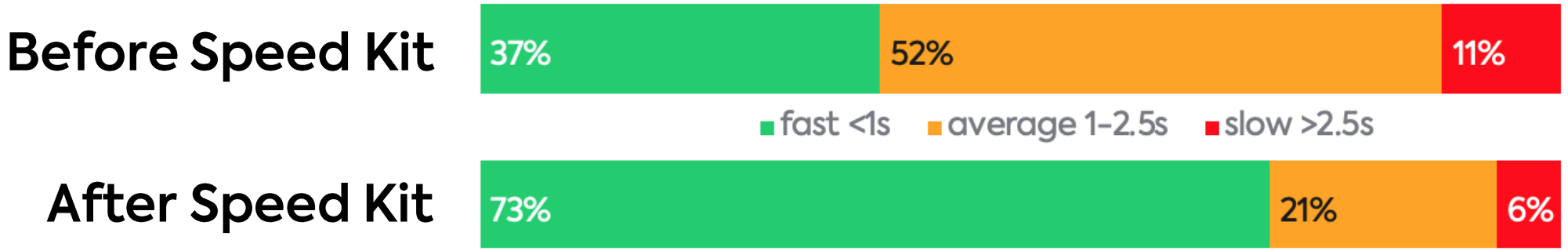


2.5x faster

After
Speed Kit

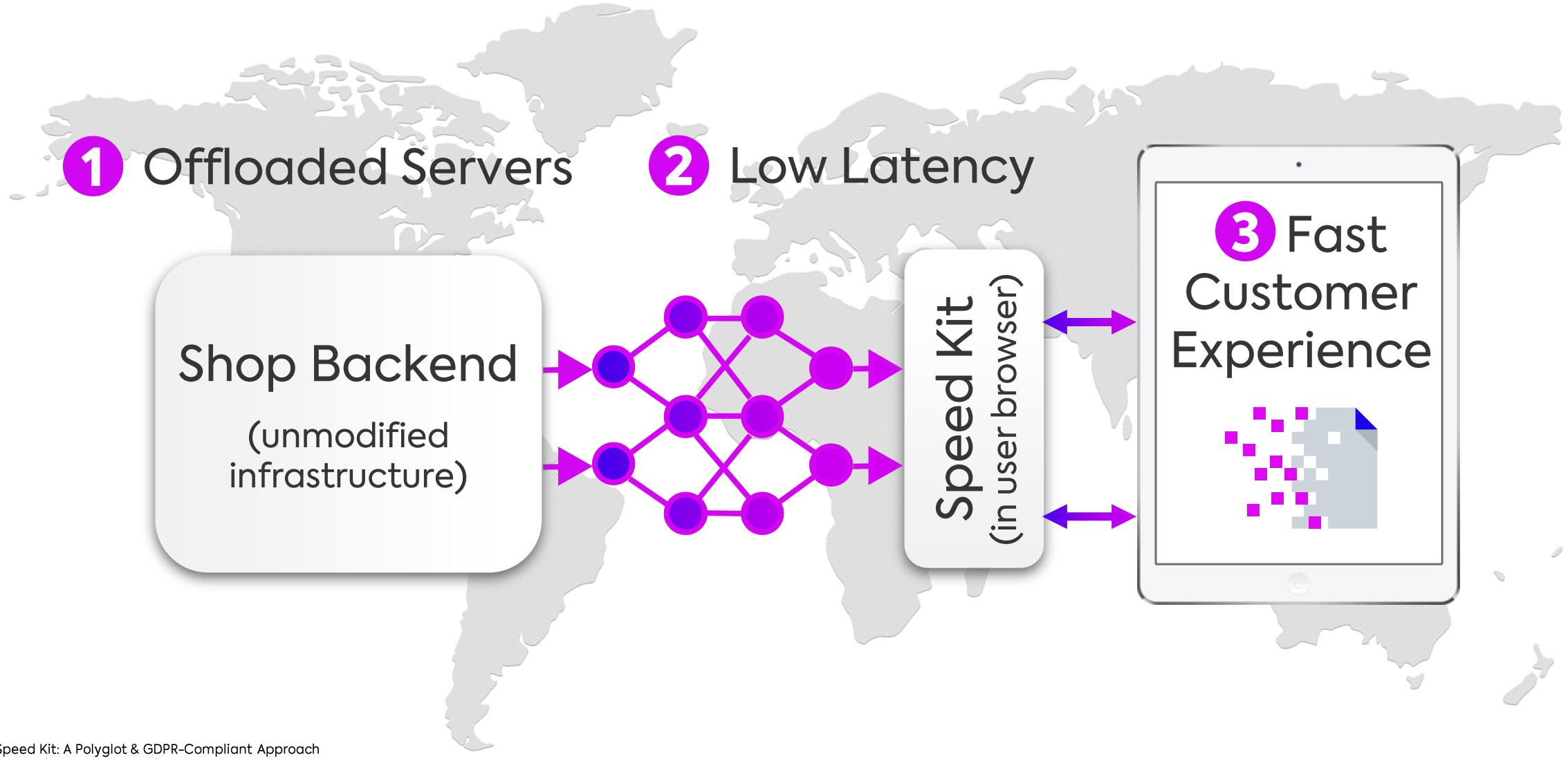


Decathlon.de: Uplift According to Google



*time until First Contentful Paint according to Google's Chrome User Experience Report (CrUX)

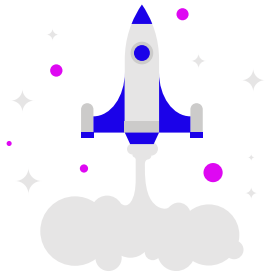
Speed Kit Optimizes End-To-End



Speed Kit: A Polyglot & GDPR-Compliant Approach For Caching Personalized Content, ICDE 2020.

Split Testing for Web Performance

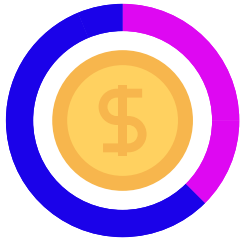
Speed Kit Users



- Speed Kit enabled

Tracking →

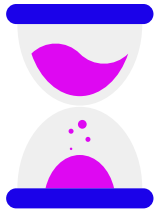
vs.



- **Measurable uplift**:
  - \+ Performance
  - \+ User engagement
  - \+ Business success

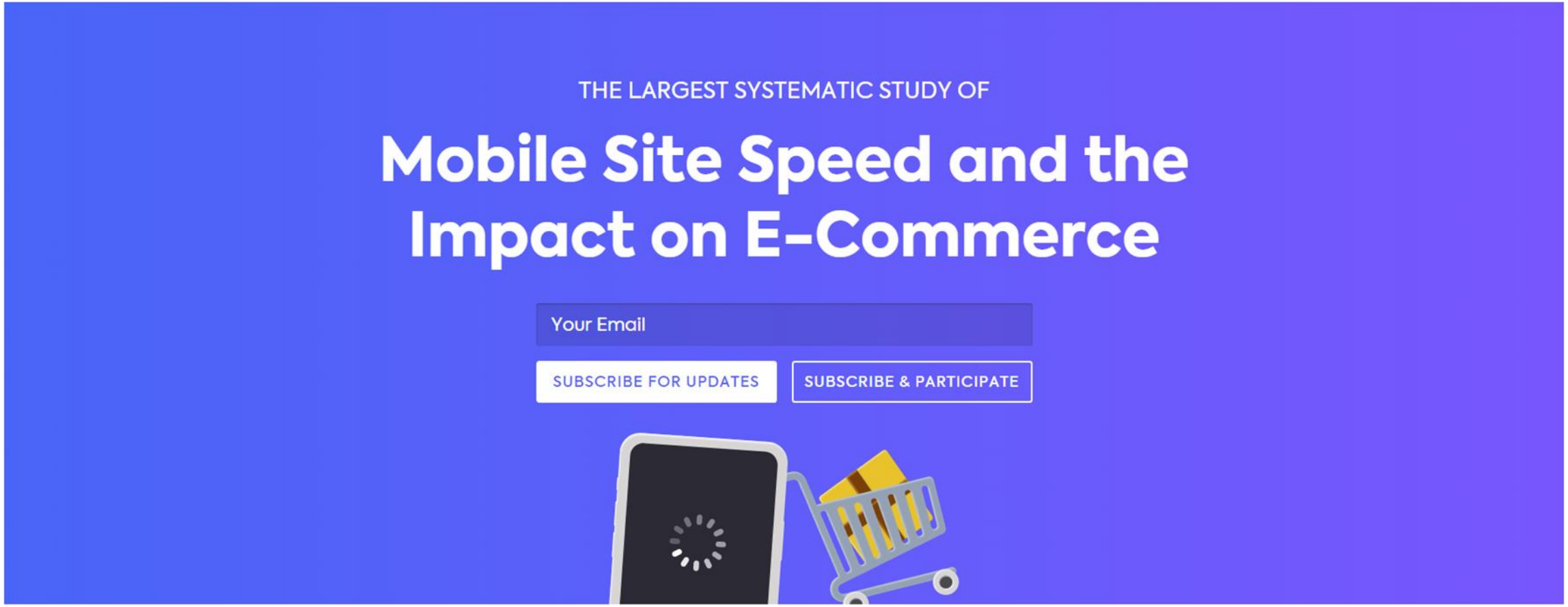
← Tracking

Normal Users



- Speed Kit disabled (no acceleration)

Join & Learn More: speedstudy.info



Google

Baqend

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Ready to Load Instantly?

Join the study!

Details & newsletter on

speedstudy.info

Wolfram Wingerath wolle@baqend.com