Third-party Content *a deep dive*

Simon Hearne

Web Performance Solutions Engineer @ Akamai

Slide notes

Welcome to this talk on third-party content.
If you are viewing this live, use keys 'n' and 'p' to navigate the slides. Use key 'o' to view an overview.
Click here to hide these notes.

@SimonHearne

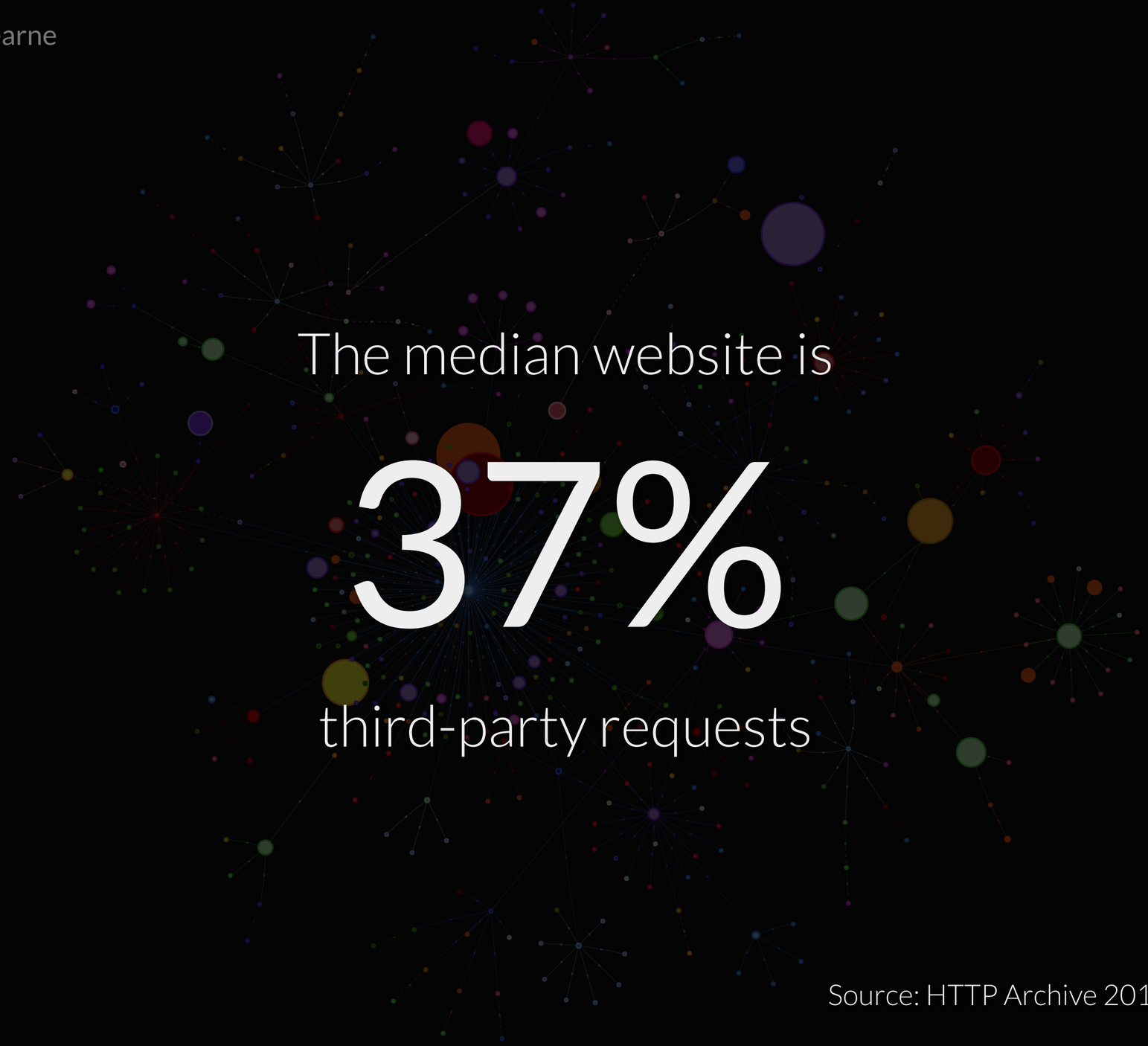
Third-parties?

Content served from a domain outside of your control

- Analytics
- Advertising
- Optimisation
- JS CDNs
- Tag Management
- …

Slide notes

Third-party content is stuff that's on your apps but outside of your control.
They're offten added after release, by teams not in the development side of the house.



The median website is

37%

third-party requests

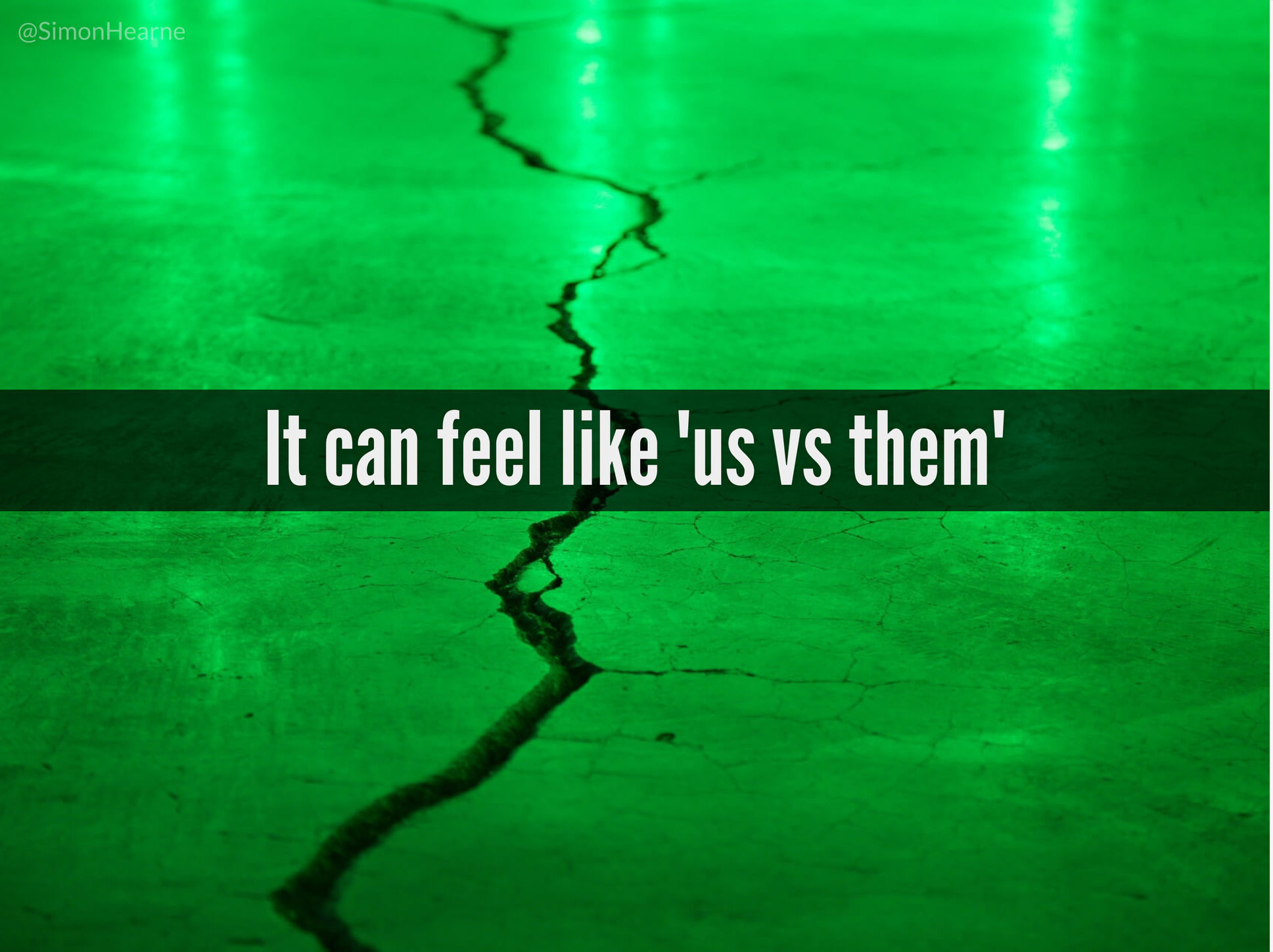
Source: HTTP Archive 2019-10-01

Slide notes

Third-parties add value to websites, improve customer insight and allow monetisation.
So it's no surprise that there are a lot of them!

Slide notes

We've been talking about third-party performance for years! In my case it's been five years, to the day!
I've given a few talks about third-parties, I've given workshops, audits and reports.
Unfortunately not a huge amount has changed, we're using more third-parties than ever.



Slide notes

It feels like there's a disconnect between the many folks who have ownership of a web app

everyone who works on a web product

shares ownership of performance and security

(whether they know it or not)

Slide notes

No notes on this slide.

Slide notes

As you are viewing this slide deck, I'm going to make an assumption that you are hands-on.
How many in the audience actively work on a web product (developer / product / ops / marketing)?
Keep your hand up if you touch code.

So the majority of us here actively work on code. There will be very few representatives from User Experience, Marketing, Product, BI & Analytics in this room.

Slide notes

That means that for most of you, third-party content is more of an irritation and frustration rather than a source of revenue, insight or joy.

Regulations are increasing visibility

Slide notes

While regulations such as GDPR and cookie consent laws have caused a lot of frustration, they have the positive effect of bringing third-party accountability to the foreground.
Plus the unfortunate acronym for China's Personal Information Security Specification.

@SimonHearne

Tag managers are enablers

Slide notes

We have developed tools to move ownership of third-parties away from technical teams.
While this is a good thing for velocity, it has diluted ownership of the front-end.

@SimonHearne

Fancy a super power?

Slide notes

It's unlikely that anyone will become a full-time third-party regulator. What you can become is the subject matter expert.
I've known a number of engineers who have become the third-party SME, accidentally or purposefully, and it has been a net positive on their careers and the products they work on.
In the next half hour I will try to arm you with the stories, tools and techniques to manage third-party content. I'll do this through post-mortems of some recent incidents.

@SimonHearne

Incident 1: the 30s wait

Slide notes

This is the incident that got me interested in third-parties in the first place.
A client had just launched a new version of their site across asia
All the pre-production tests worked fine, they tested from servers in the target geos and performance was good

Launched in China 🇨🇳

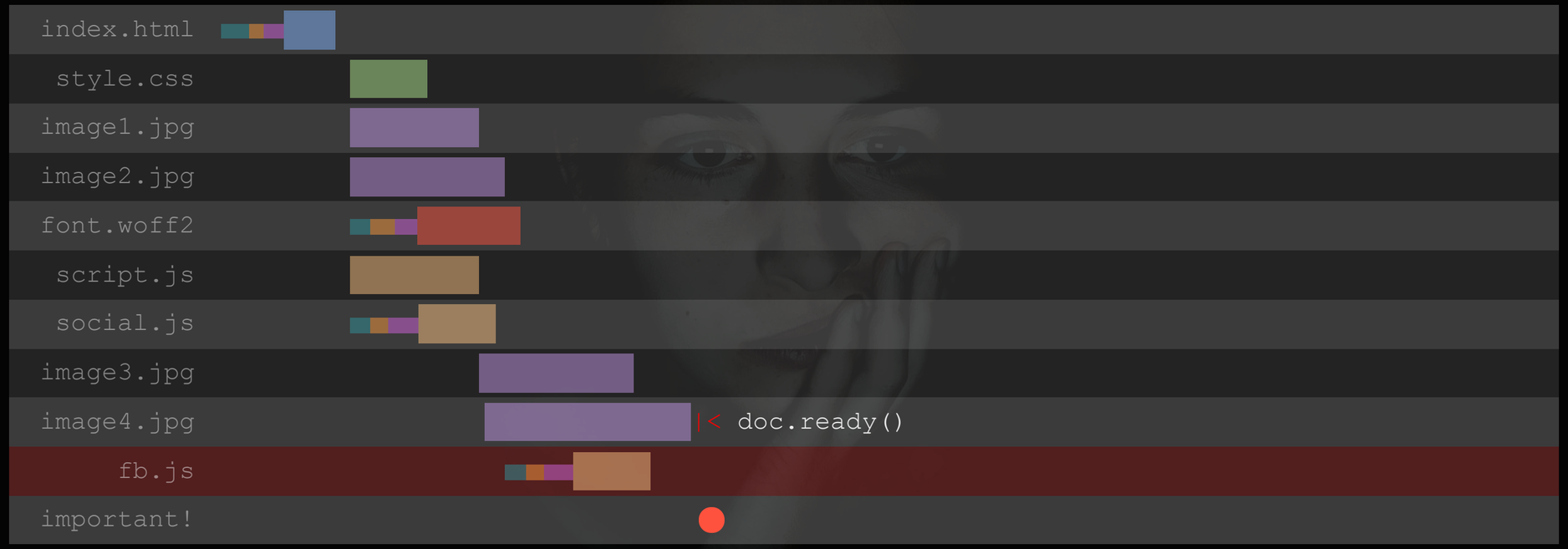
👍 localisation
👍 dev testing
👍 QA in China

Slide notes

The launch was continued to China.
Feedback from local users was extremely poor and analytics showed page load times of over 30s.

@SimonHearne

The problem



Slide notes

The problem was an assumption that the document.ready event would not be impacted by the script.
The application design had critical event management logic on document.ready so the application was unavailable until it fired!

@SimonHearne

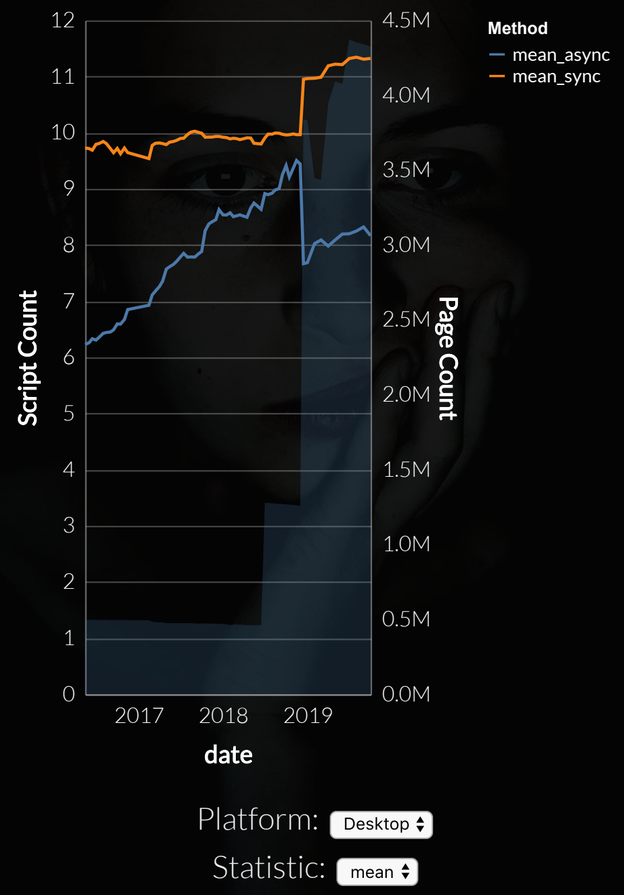
Don't worry, it's async!

Slide notes

No notes on this slide.

@SimonHearne

sync is still king(!)



Platform: Desktop

Statistic: mean

Source: HTTP Archive. 2016-05-15 to 2019-10-01

Slide notes

No notes on this slide.

The defense 👮

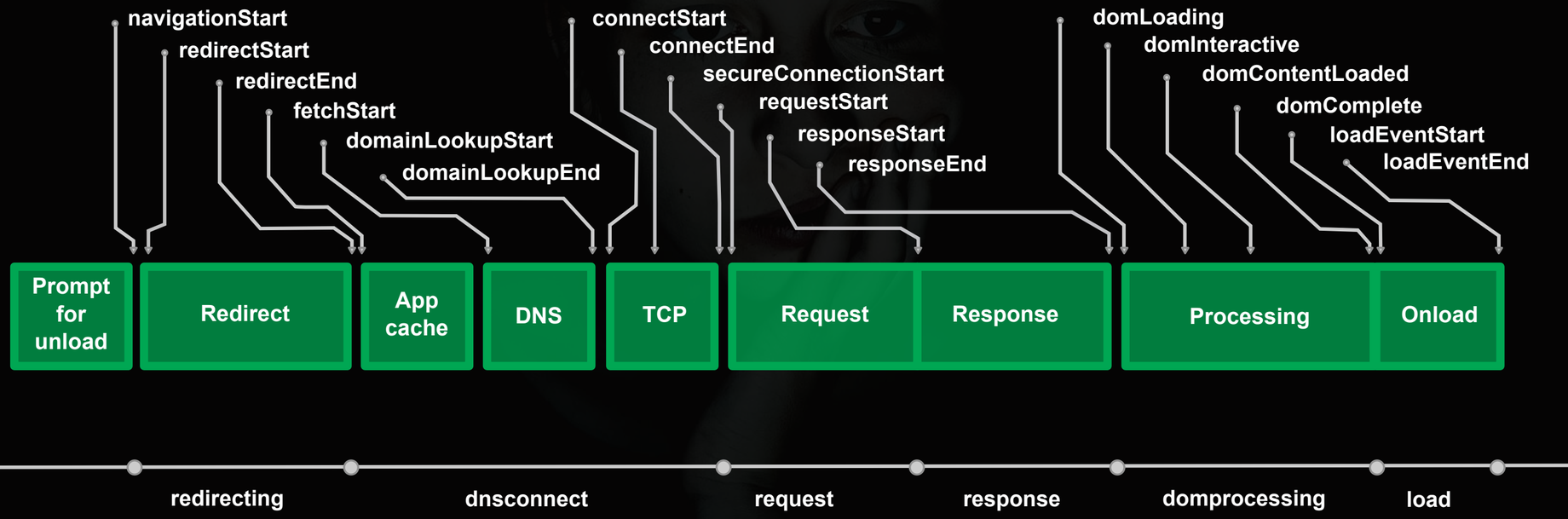
- 🇨🇳 feature flag
- don't block on onload
- `<script defer >`

Slide notes

No notes on this slide.

@SimonHearne

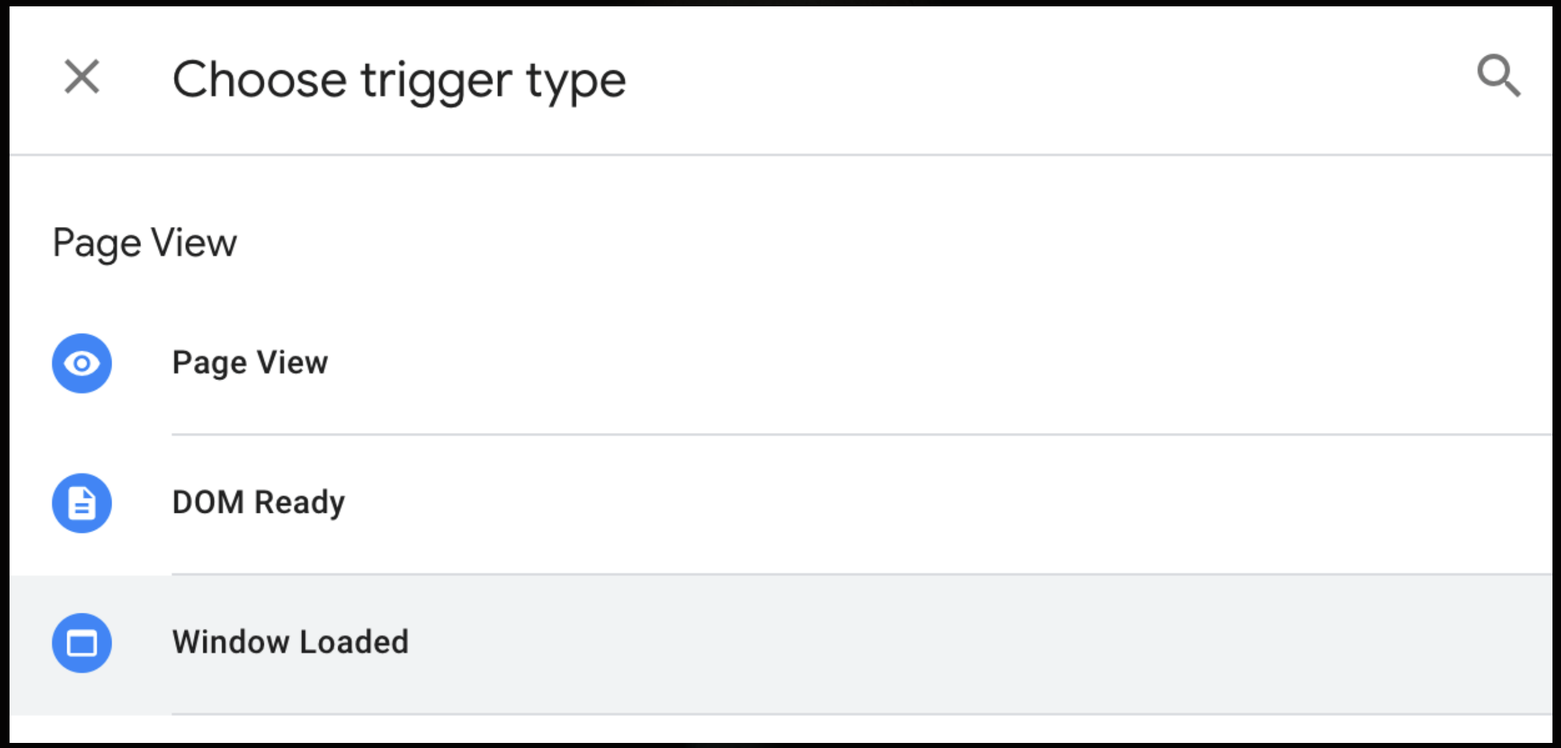
Beware your script loading strategy



Slide notes

Deferred scripts run immediately after dominteractive: "when the parser finished its work on the main document"

Tag managers allow for this

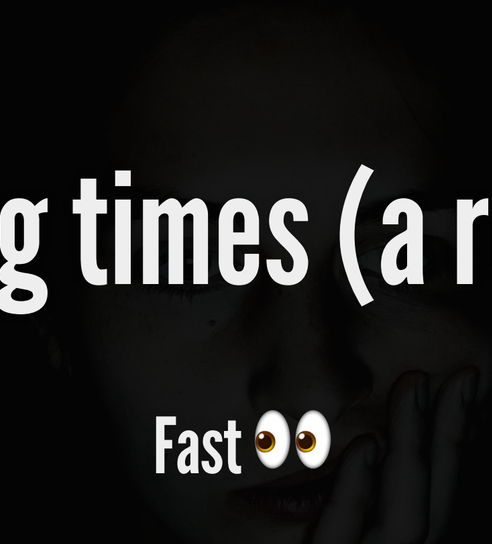


Load scripts as late as possible

Slide notes

Tag managers offer flexibility to define when scripts are loaded.
By default, you'll probably find them loading too early (marketers always want their scripts to execute as soon as possible).
It is simple to evaluate when a script needs to evaluate, so make it a point to load scripts only when necessary.

Tag loading times (a rough guide)



Immediate ⚡

Experimentation
Tag Manager

Fast 👀

Session Tracking
Performance Monitoring

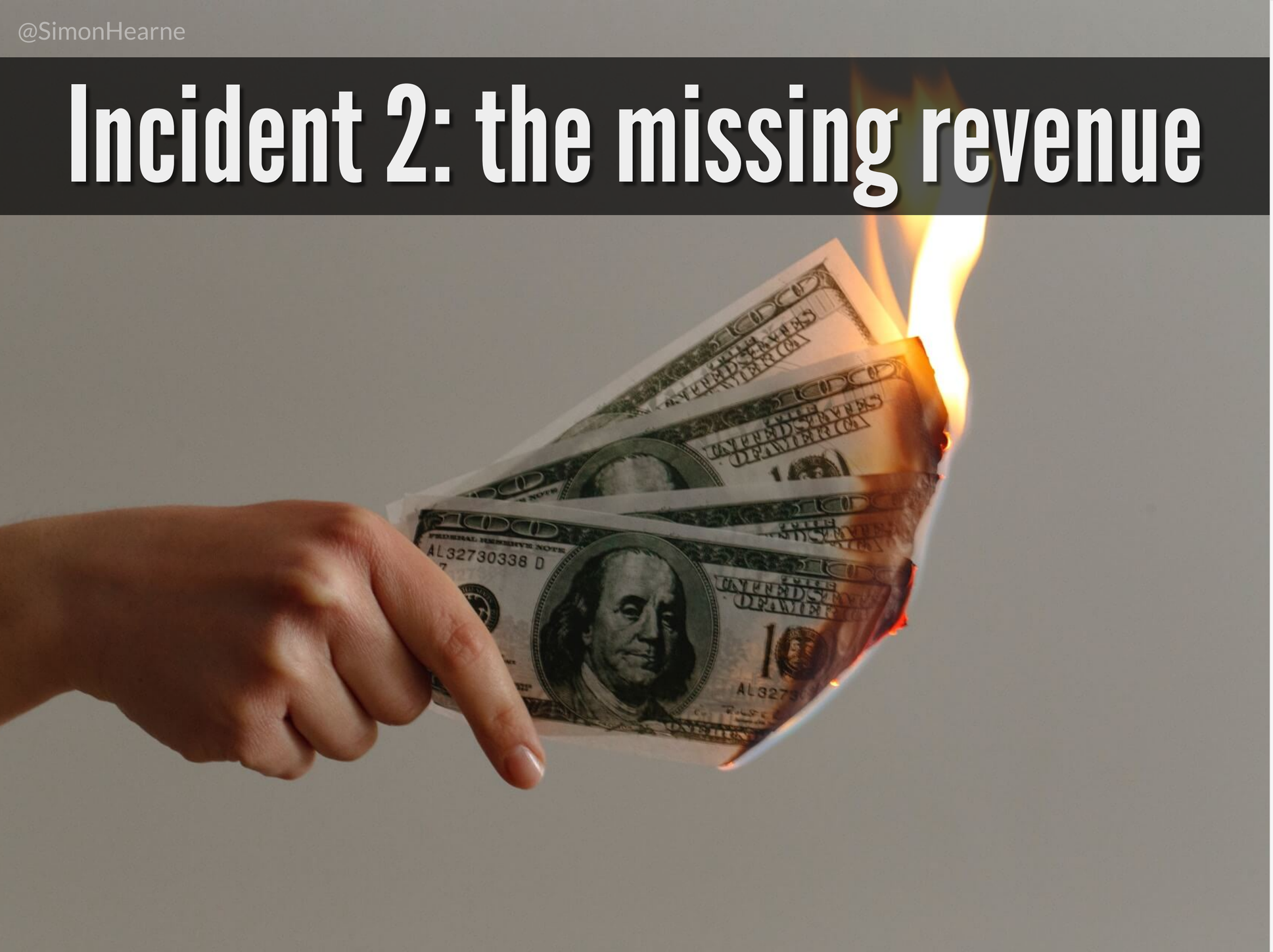
Late 🐌

Ads(!)
Reviews & Ratings
Live Chat & Support

Slide notes

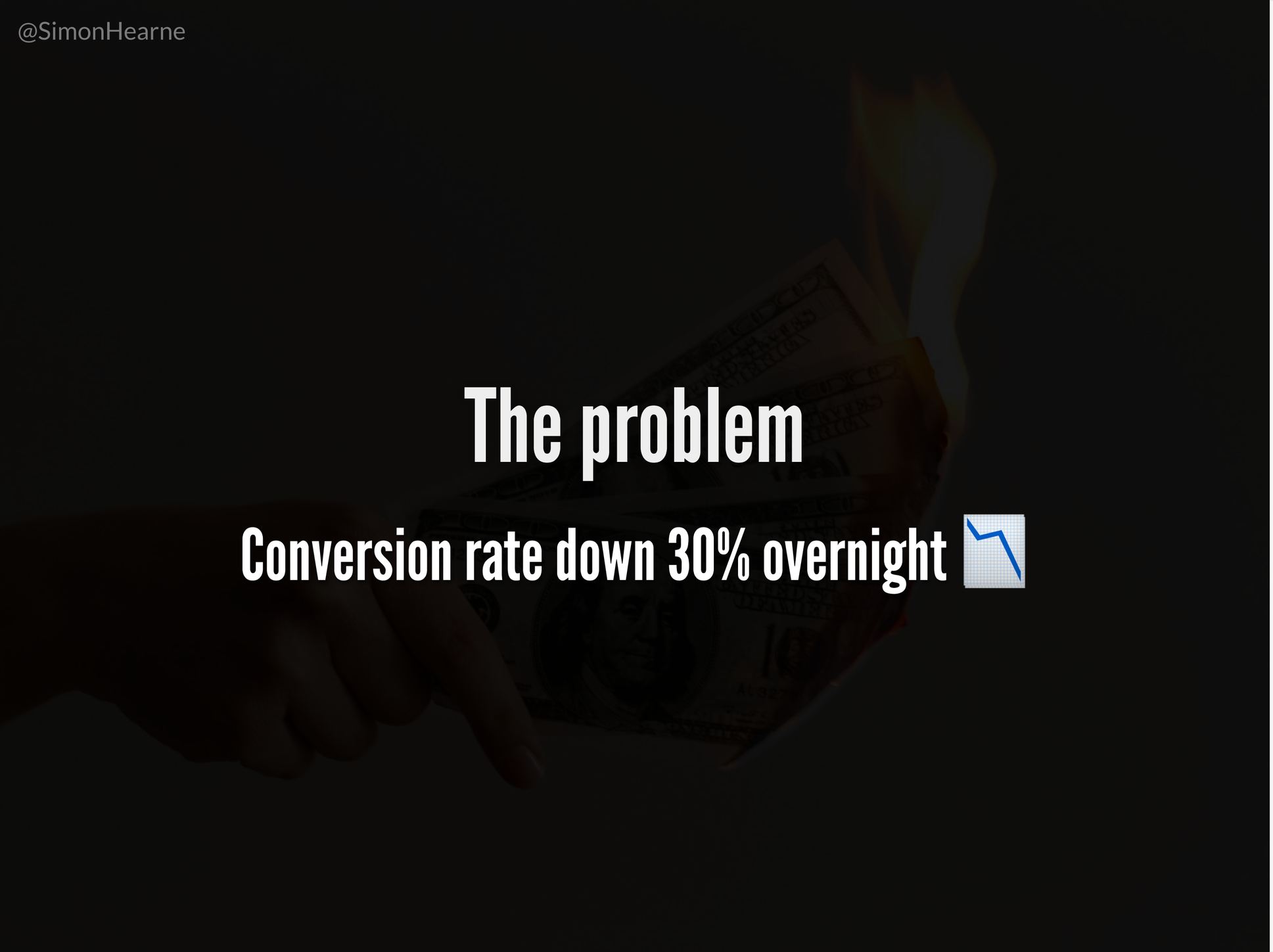
Don't treat your tags the same, some have to load early but most can be loaded late.

Incident 2: the missing revenue



Slide notes

I was working with a UK retailer who had initiated a war room because their conversion rate had dropped by 30%.
Their automated tests were showing 100% success, they had not released any code and there were no stock changes.



The problem

Conversion rate down 30% overnight 📉

Slide notes

No notes on this slide.

The discovery

```
/**
* This file uses jQuery materials and hashchange plugin
* Portions, Copyright (c) 2010 "Cowboy" Ben Alman
* http://benalman.com/projects/jquery-hashchange-plugin/
* Portions, Copyright 2012 jQueryFoundation and other contributors
* http://jquery.com/
```

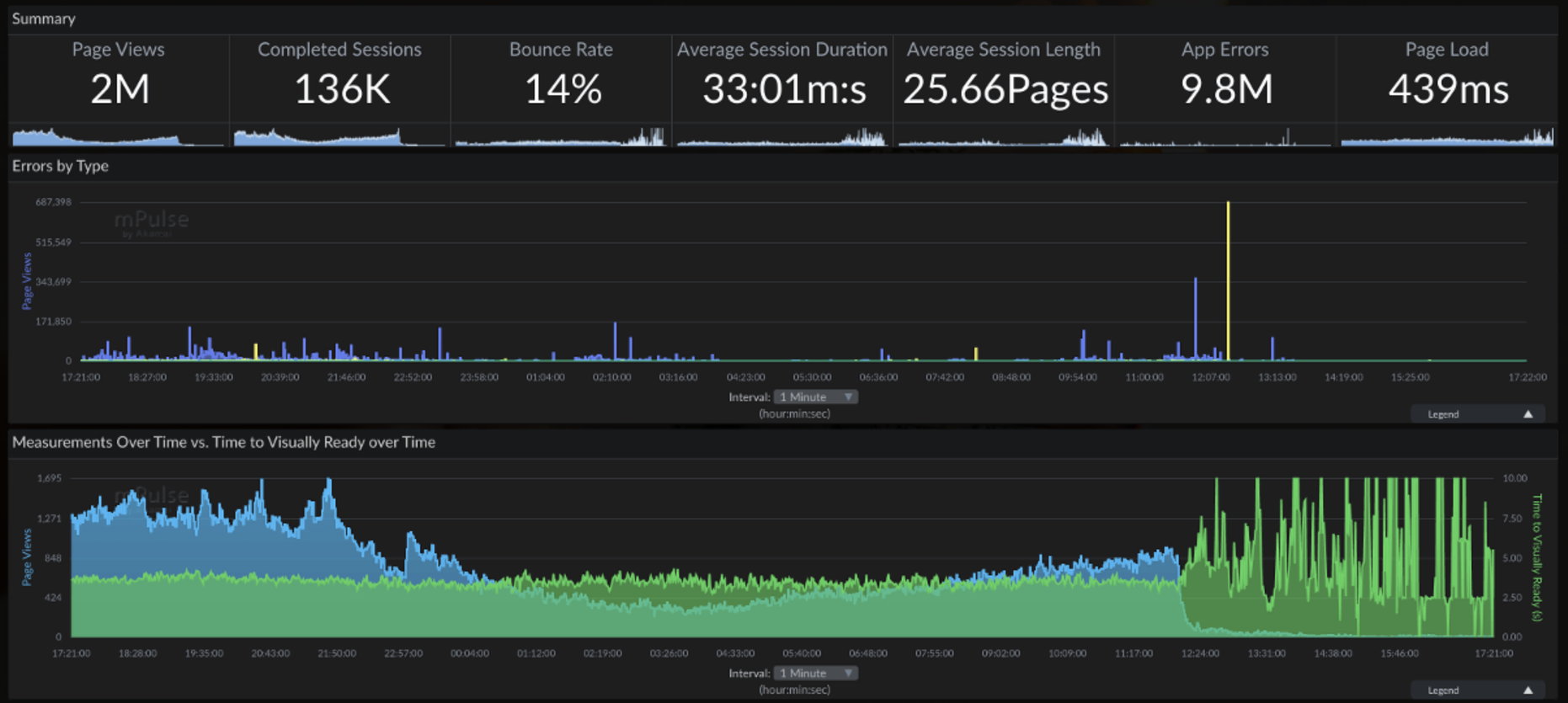
```
<script async src="//thirdparty.com/embed.js">
```

Slide notes

The issue was a change in a third-party script. This was not announced publicly, and as the URL was absolute the change was pushed immediately.
The update conflicted with an event listener (poorly implemented by the client) which broke the quick checkout button.

@SimonHearne

The discovery



Slide notes

The issue went unnoticed by operations until questions were raised by ecommerce.
All testing was reporting 100% success, but there were some errors being logged to the console.
Due to alert fatigue, no-one was informed of the increase in errors. In this example we can see a massive spike before a huge drop in traffic, caused by a misconfigured third-party script.

Subresource Integrity

Only execute code if it matches the hash

```
<script
src="//thirdparty.com/embed.js"
integrity="sha256-ivk71nXhz9nsyFDoYoGf2...="
crossorigin="anonymous">
```

Currently used on *5.15%* of pages

(HTTP Archive - 2019/10/01)

SRI on MDN

Slide notes

SRI enforces a version of a third-party asset. If the asset is changed without your knowledge it will not execute in the browser.
This is best used for non-critical assets, such as analytics. The failures will be silent, but can be reported with Content-Security-Policy.

Back to <img> pixels?

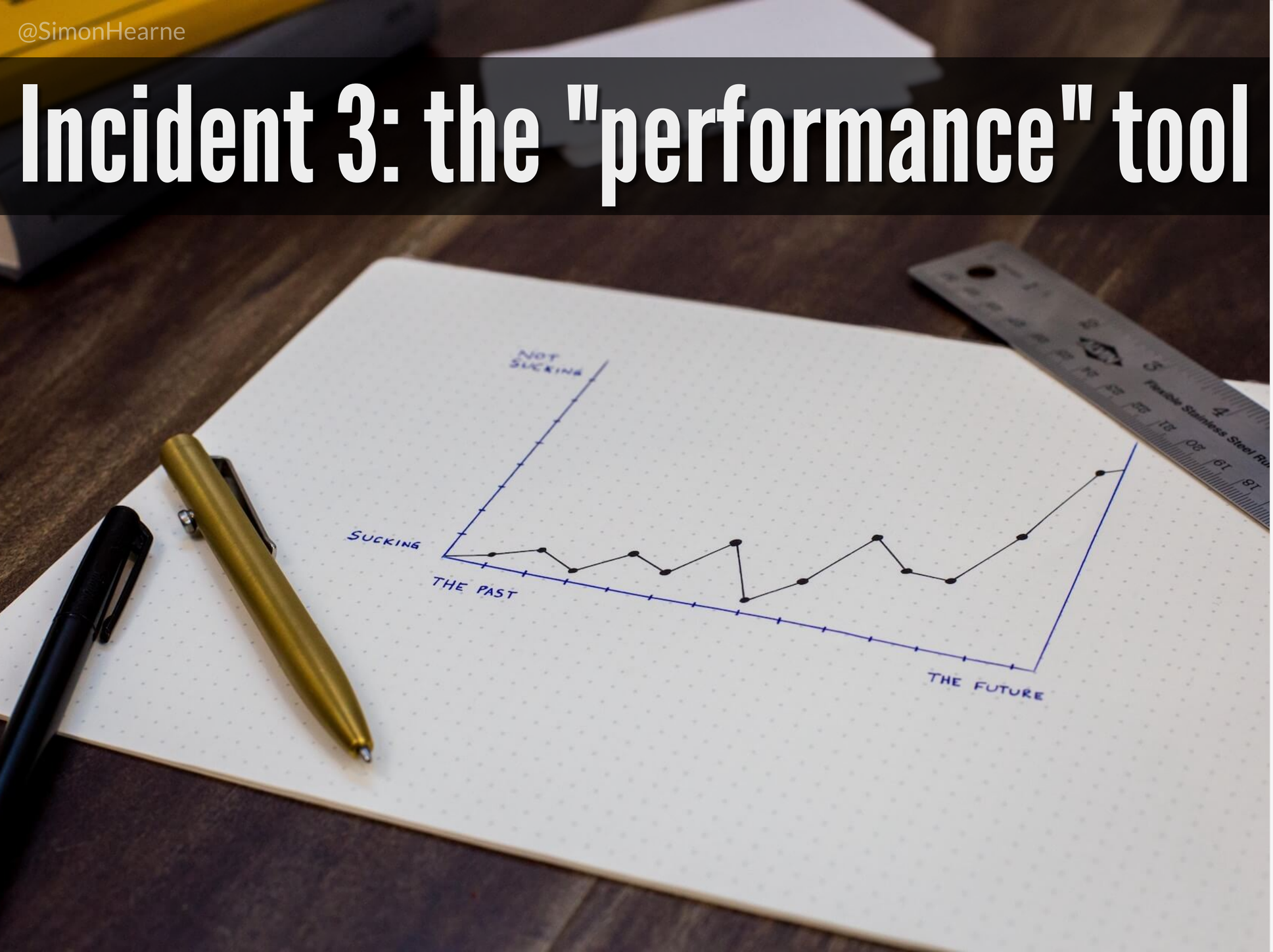
```
<!-- Facebook Pixel Code -->
<script>!function(f,b,e,v,n,t,s){...}();</script>
<noscript>
<img height="1" width="1" style="display:none"
src="https://www.facebook.com/tr?id=...&noscript=1"
/>
</noscript>
```


Slide notes

Reduce the impact of third-party scripts, by swapping them for the fallback image pixel!
This will reduce the amount of data that can be collected, but that's not necessarily a bad thing.

@SimonHearne

Incident 3: the "performance" tool



Slide notes

This was a performance monitoring tool that runs in the client, like mPulse RUM!
After implementatin there were a number of complaints from customers...

The problem

> *The site feels "laggy"*

🧐

Slide notes

The feedback was very subjective, and nothing in the data showed a performance problem.

@SimonHearne

The problem

`clicked!`

instant

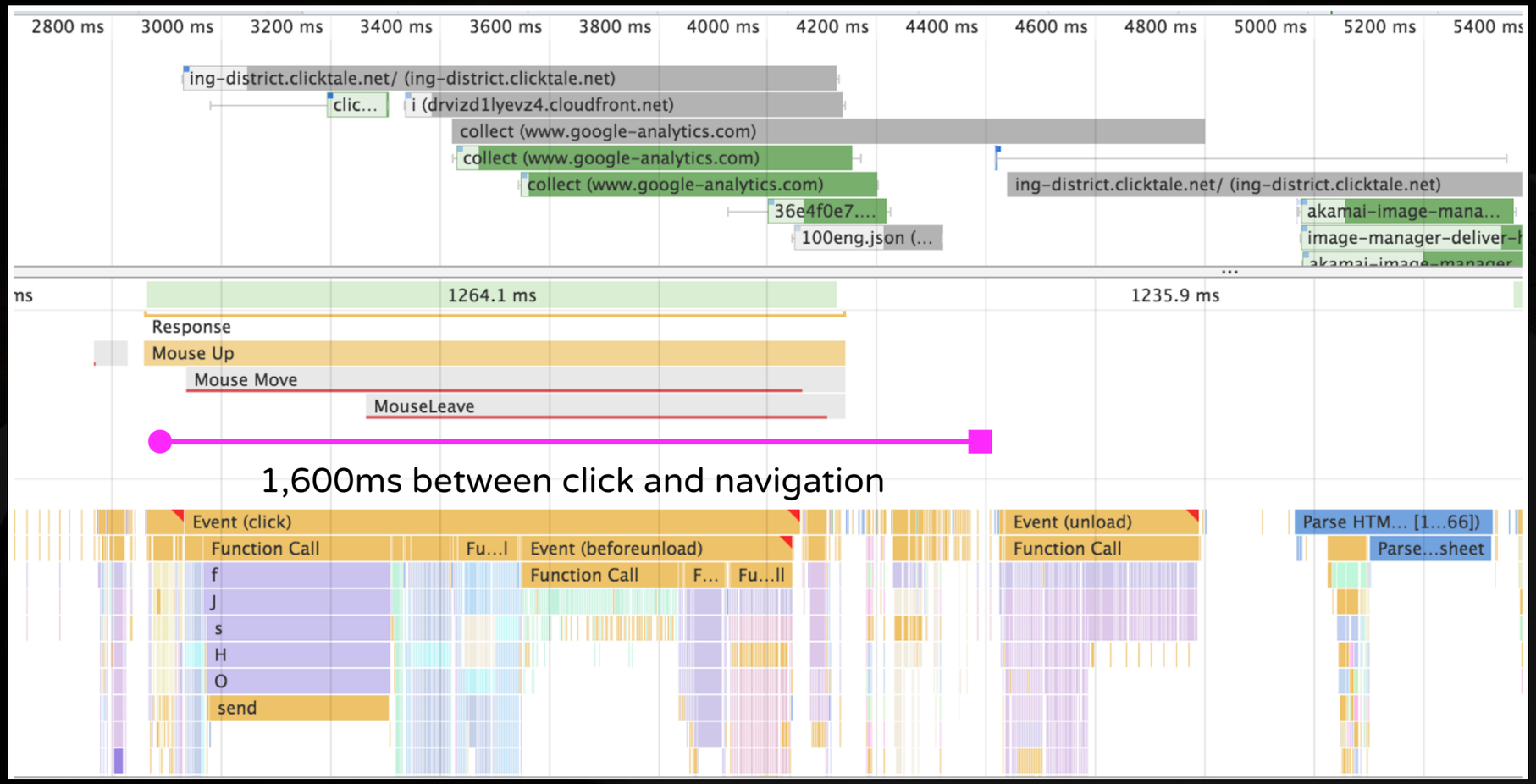
500ms delay

Slide notes

No notes on this slide.

@SimonHearne

The discovery



Slide notes

The issue was caused by a combination of third-party event listeners.
In this example we can see ClickTale adding around a second to the unload event - this blocks the next navigation and is perceived as latency.

@SimonHearne

The discovery

95th %ile Unload Time by OS - Top 5

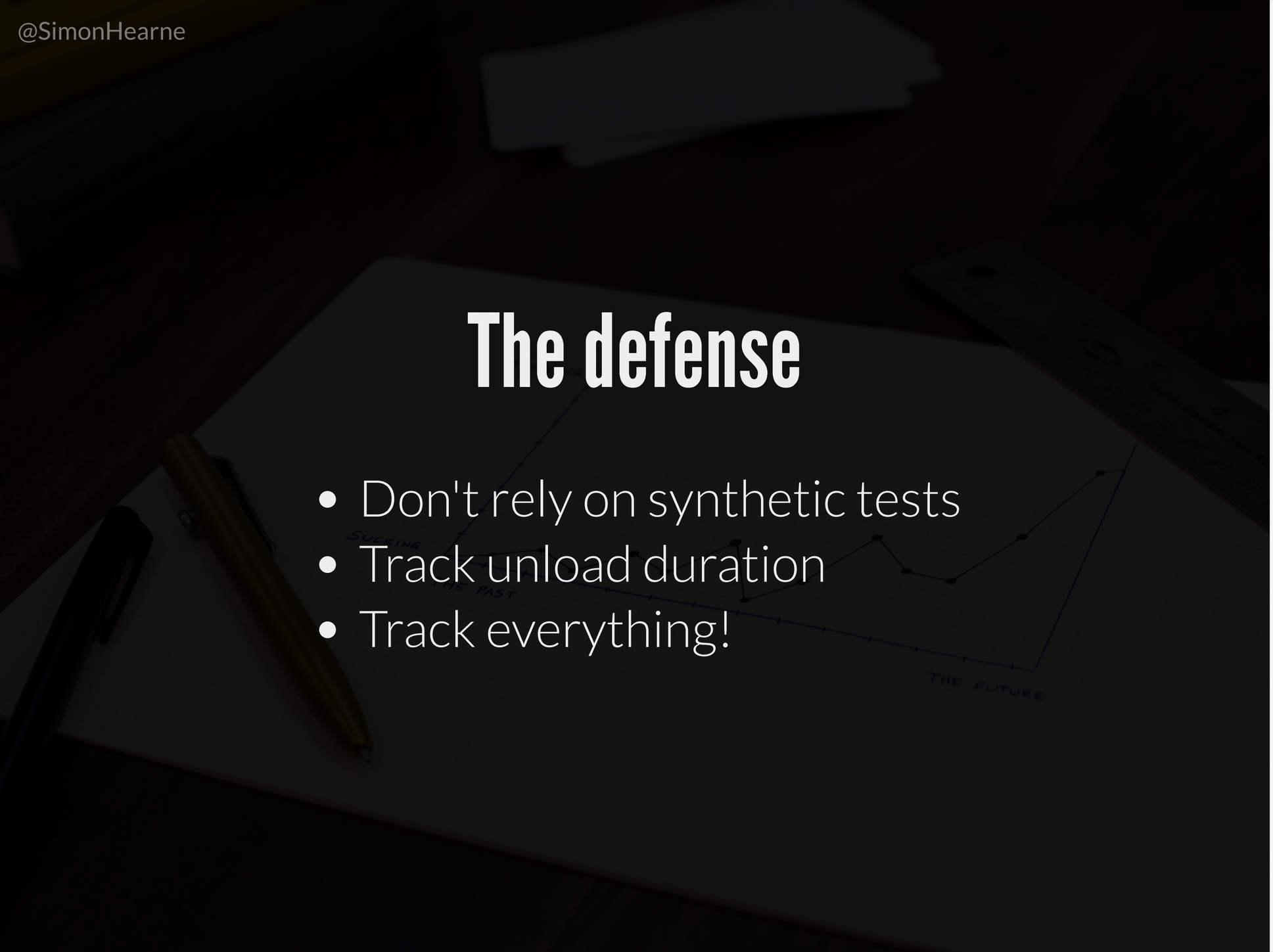
Row	Operating System	Unload	Percent ▼	Session Length	JavaScript Decoded	Time to Interactive
1	▶ Windows	77 ms	54.26%	4.15	2.43 MB	16.75 s
2	▶ Mac OS X	511 ms	19.88%	3.19	2.65 MB	13.13 s
3	▶ iOS	792 ms	15.41%	1.80	2.50 MB	19.03 s
4	▶ Android OS	142 ms	8.57%	1.75	3.09 MB	31.67 s
5	▶ Windows Phone	278 ms	0.01%	1.82		1:13 m:s
Othe		98 ms	1.87%	2.51	3.22 MB	35.02 s
Total		374 ms	100.00%	2.78	2.70 MB	20.85 s

Slide notes

This can be tracked in tools like mPulse, and correlated with things like bounce and conversion rate.

The defense

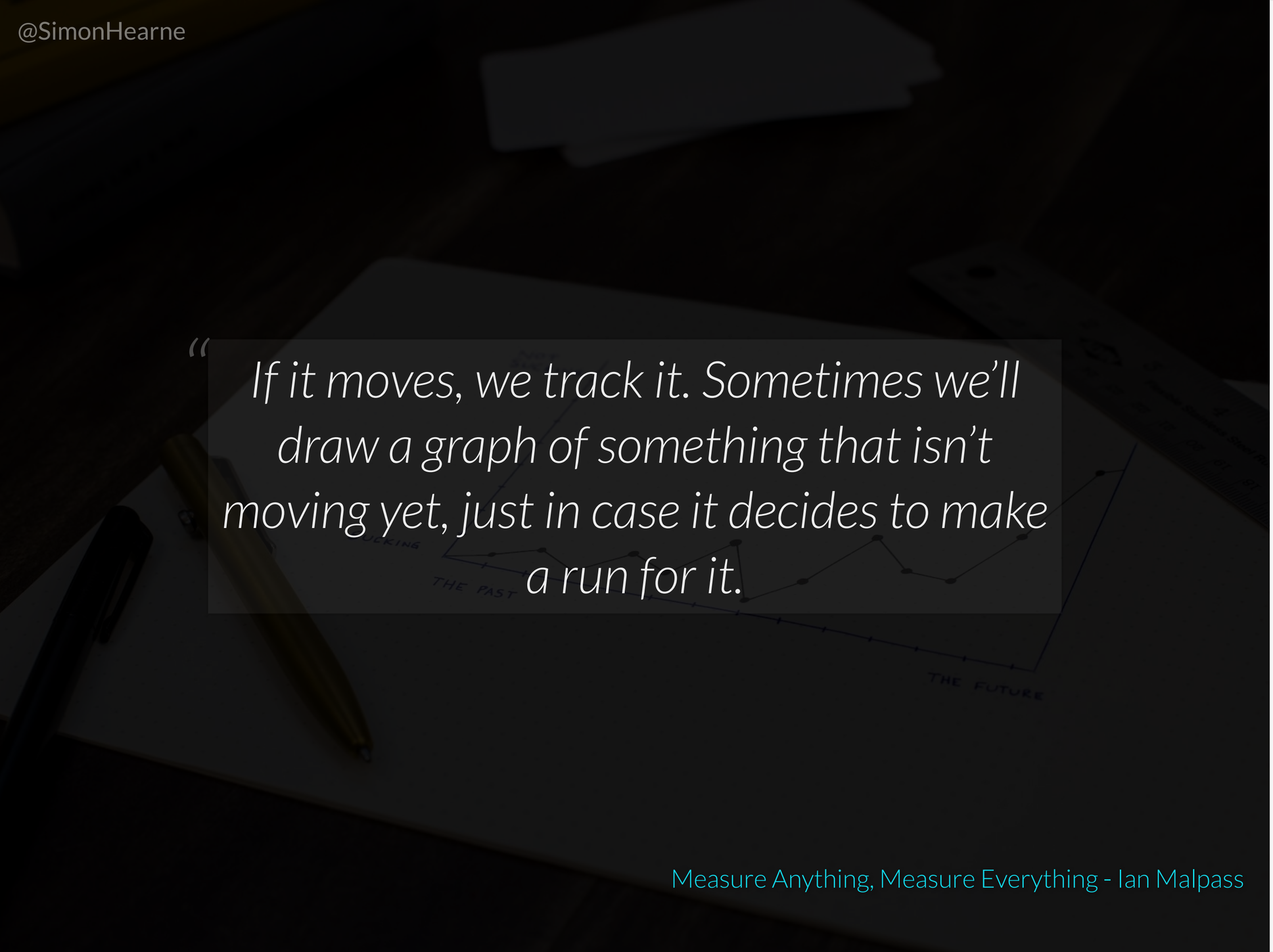
- Don't rely on synthetic tests
- Track unload duration
- Track everything!



Slide notes

No notes on this slide.

@SimonHearne



> *If it moves, we track it. Sometimes we'll draw a graph of something that isn't moving yet, just in case it decides to make a run for it.*

Measure Anything, Measure Everything - Ian Malpass

Slide notes

No notes on this slide.

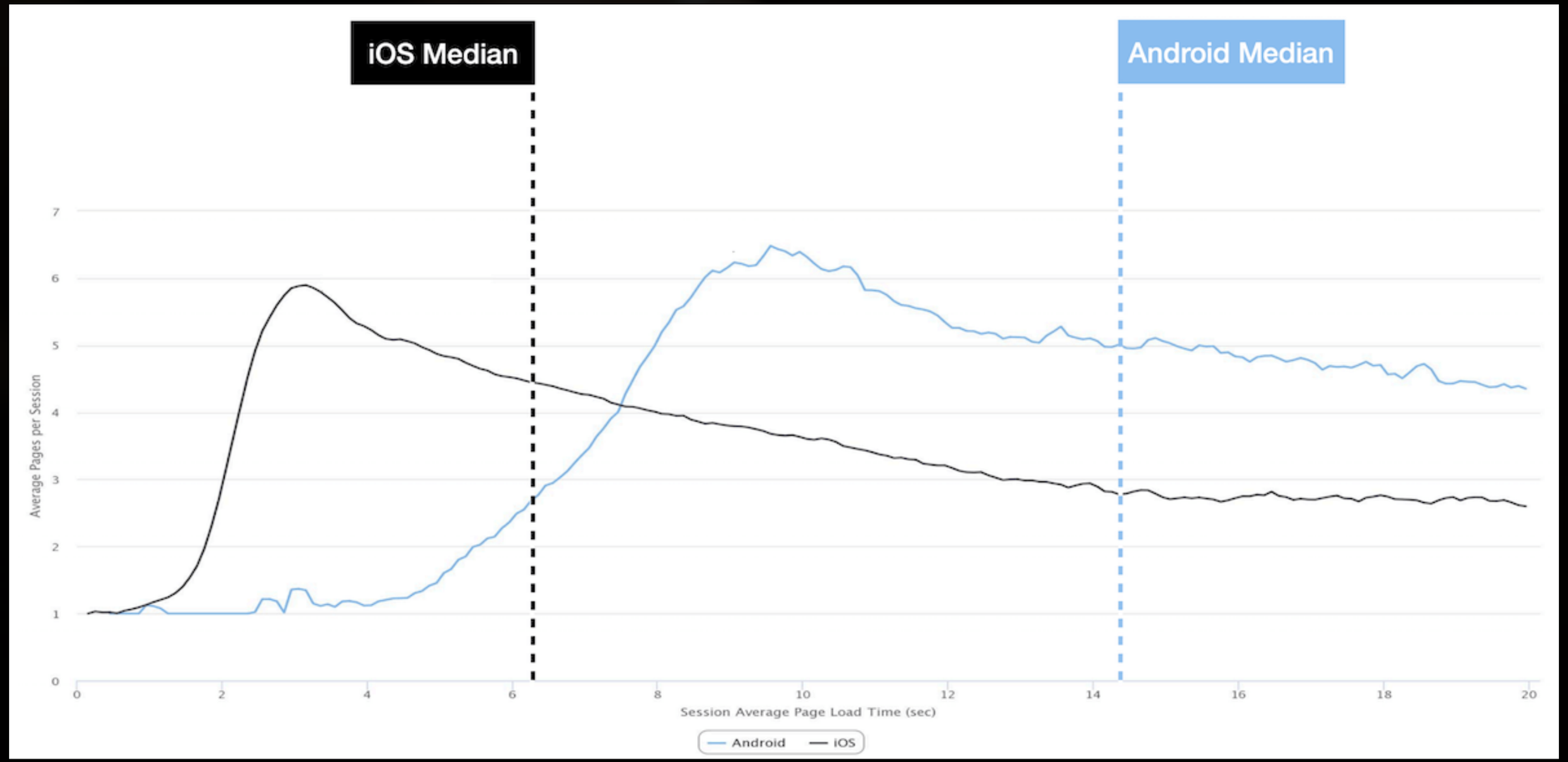
@SimonHearne

Incident 4: the review provider

Slide notes

Customers love ratings and reviews, and prefer independent sources.

The problem



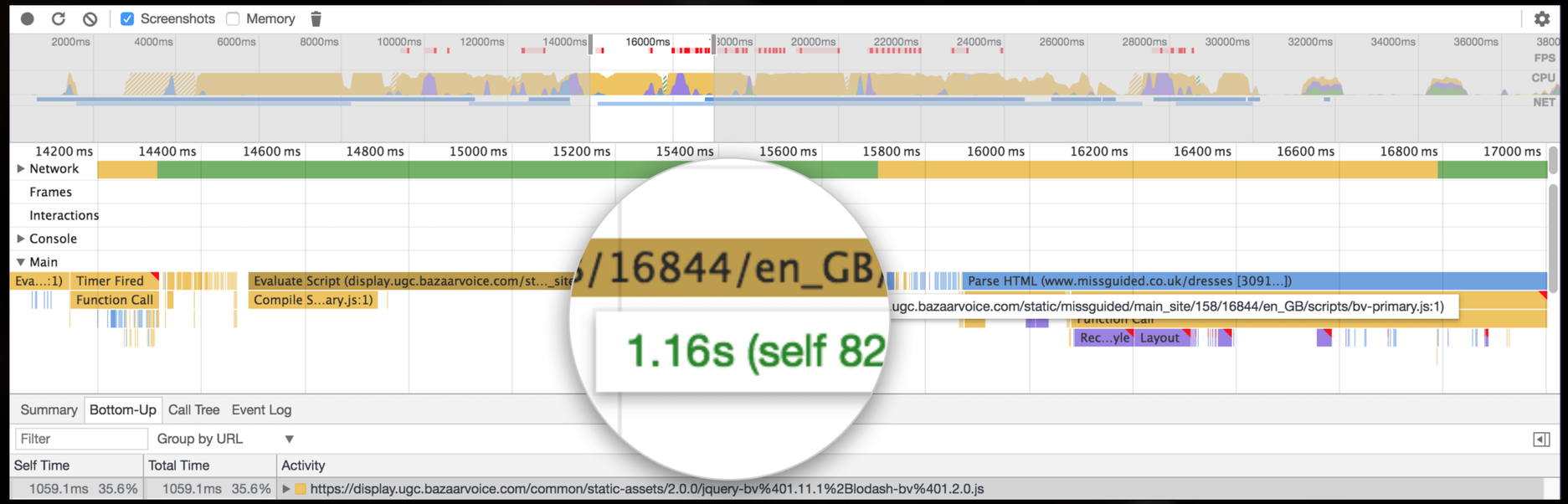
How Missguided revolutionised thier approach to site speed

Slide notes

The problem in this case was that Android users had a significantly poorer experience.
As a result, the conversion rate for Android users was much lower than iPhones.

@SimonHearne

The discovery

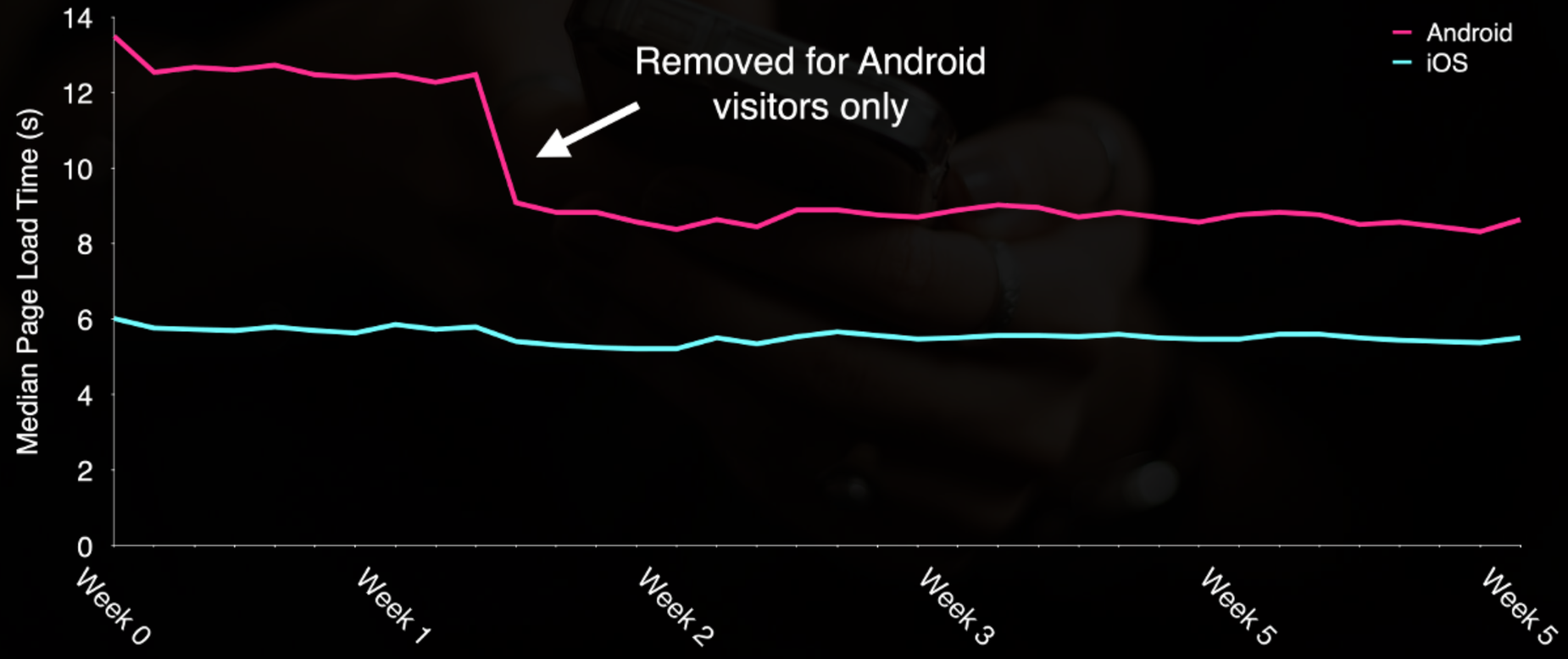


Slide notes

A little bit of digging showed that the review provider was adding significant time to JS execution.

@SimonHearne

The defense



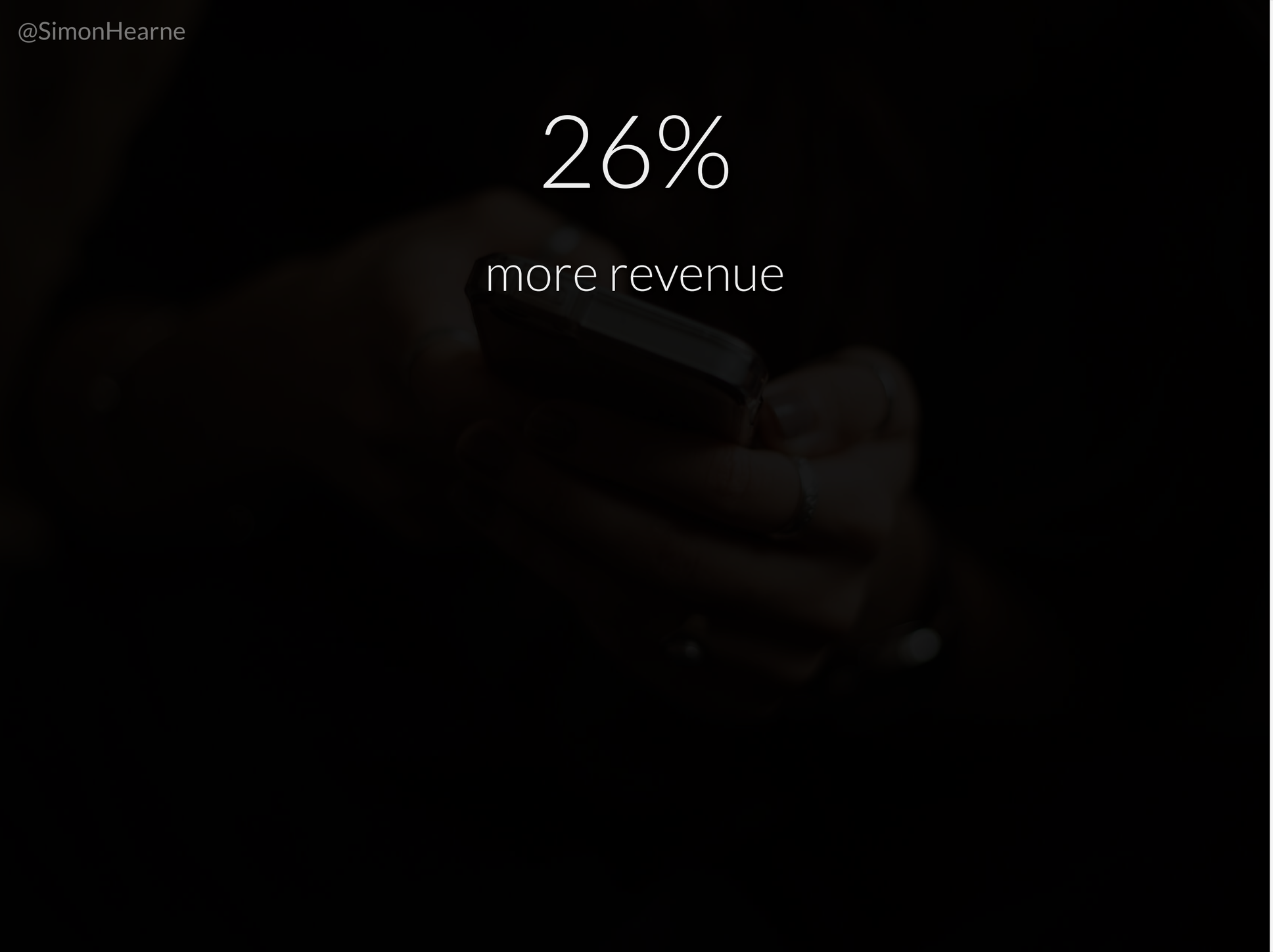
Slide notes

To evaluate the impact further, the reviews were removed from the website for Android visitors.

@SimonHearne

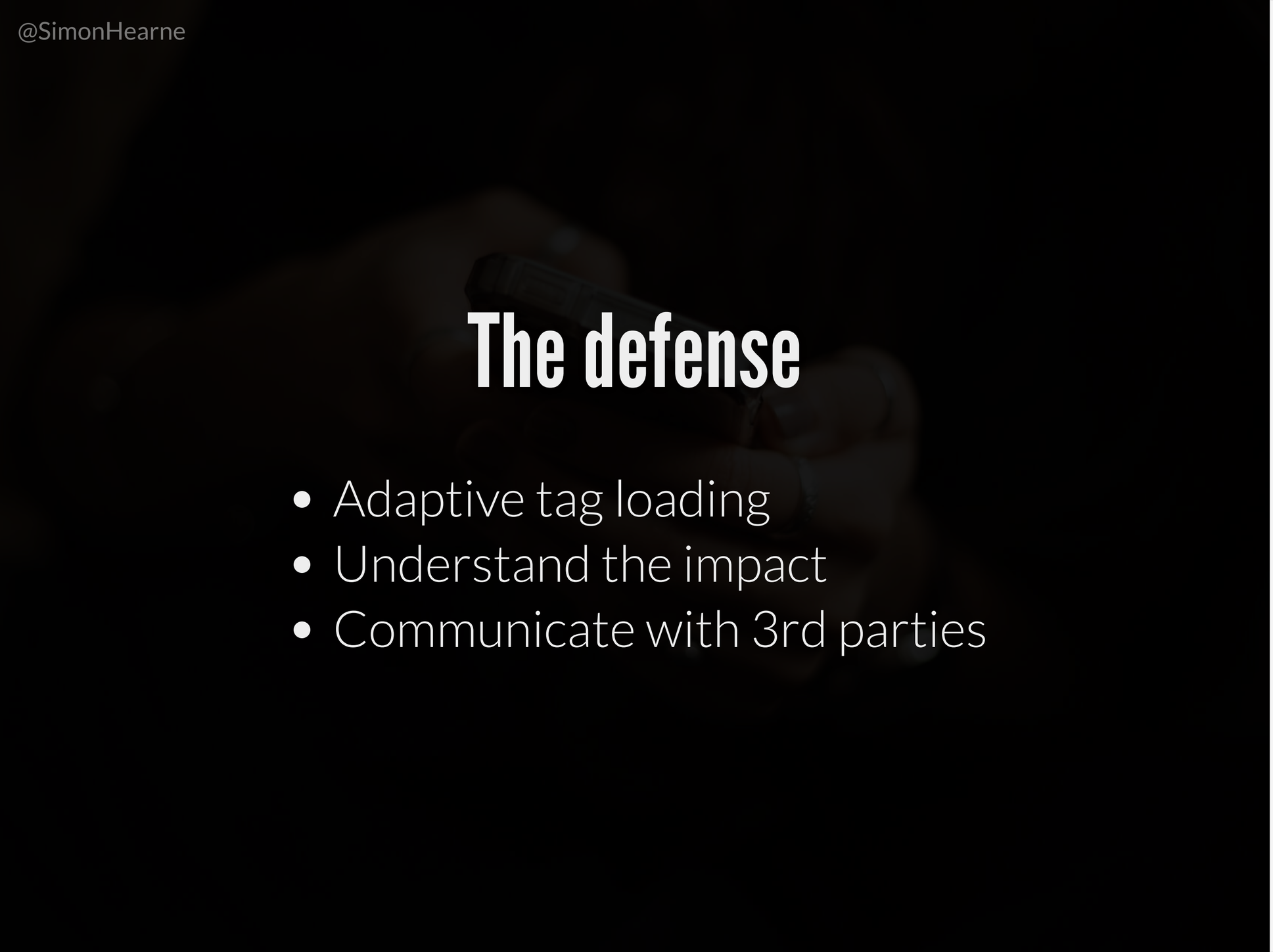
26%

more revenue



Slide notes

The result was an increase in revenue of 26% from Android visitors.



The defense

- Adaptive tag loading
- Understand the impact
- Communicate with 3rd parties

Slide notes

This bit is really tricky - attribution is not something that is made available through any APIs so it can take manual measurements and intervention.

@SimonHearne

Incident 5: the malicious ads

Slide notes

Malvertising is a surprisingly common problem. While not necessarily a performance challenge, it's still a major concern.

@SimonHearne

The problem

Scan now

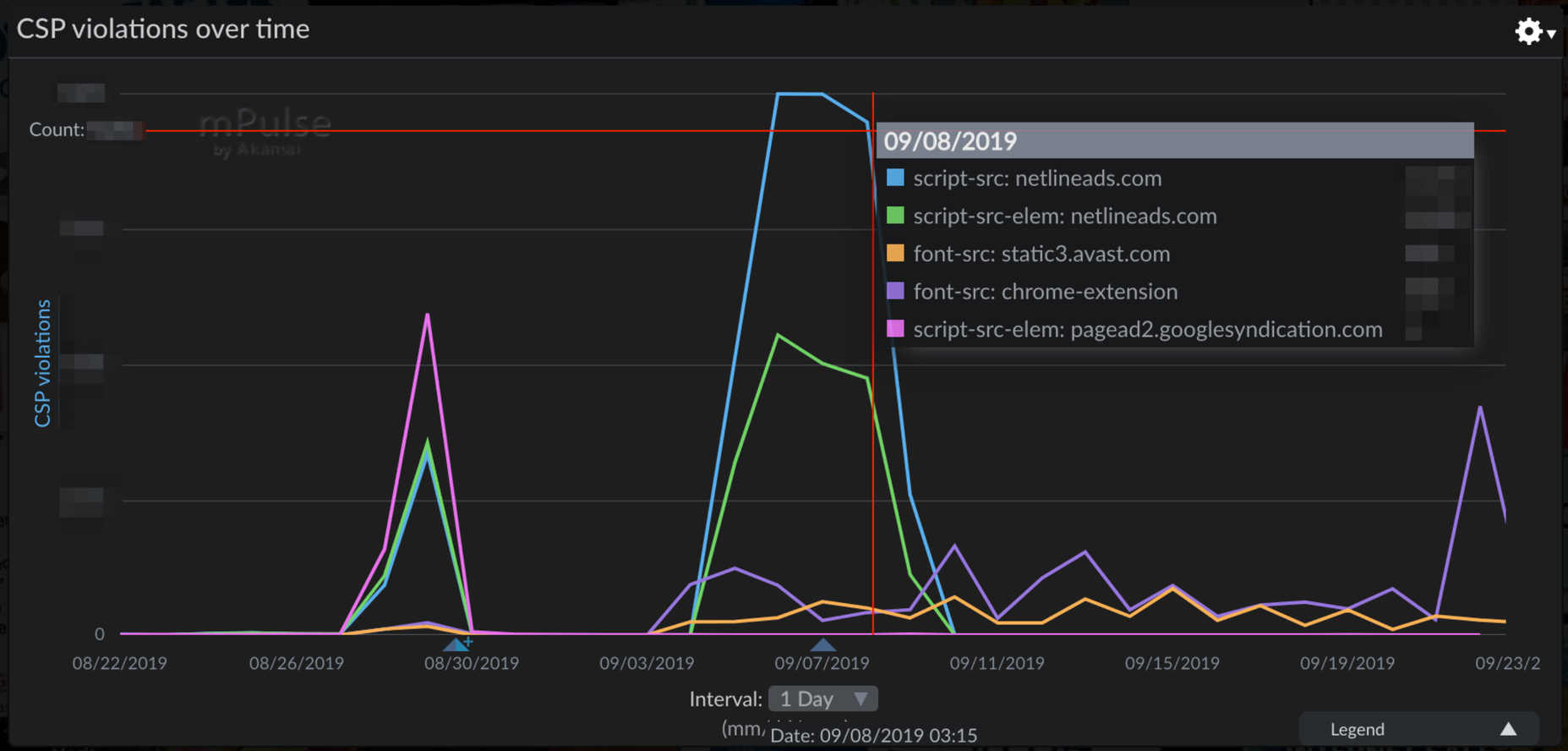
Update your antivirus (free)

Slide notes

No notes on this slide.

@SimonHearne

The discovery



Slide notes

This issue was captured in mPulse.

@SimonHearne

Content Security Policy

```
content-security-policy: default-src *;
script-src 'self' 'unsafe-inline' 'unsafe-eval' *.go-mpulse.net
maps.googleapis.com;
object-src *;
style-src 'self' 'unsafe-inline' fonts.googleapis.com;
img-src 'self' data: img.youtube.com *.akstat.io *.gstatic.com
*.googleapis.com *.google-analytics.com *.ytimg.com;
media-src 'self';
frame-src 'self' *.youtube-nocookie.com;
font-src 'self' *.gstatic.com data:;
connect-src 'self' *.akstat.io *.go-mpulse.net *.google-analytics.com;
report-uri https://akstat.io/report/<api-key>
```

Currently used on *6.11%* of pages

(HTTP Archive - 2019/10/01)

Slide notes

No notes on this slide.

@SimonHearne

Content Security Policy

CSP Directive	HTML / JS Features
default-src	*
connect-src	XMLHttpRequest(), WebSocket(), EventSource(), sendBeacon(), fetch()
style-src	<link rel="stylesheet">
script-src	<script>
form-action	<form>
font-src	@font-face
child-src	<iframe>, Worker()
object-src	<object>, <embed>
media-src	<video>, <audio>
img-src	<img>
manifest-src	<link rel="manifest">

Slide notes

No notes on this slide.

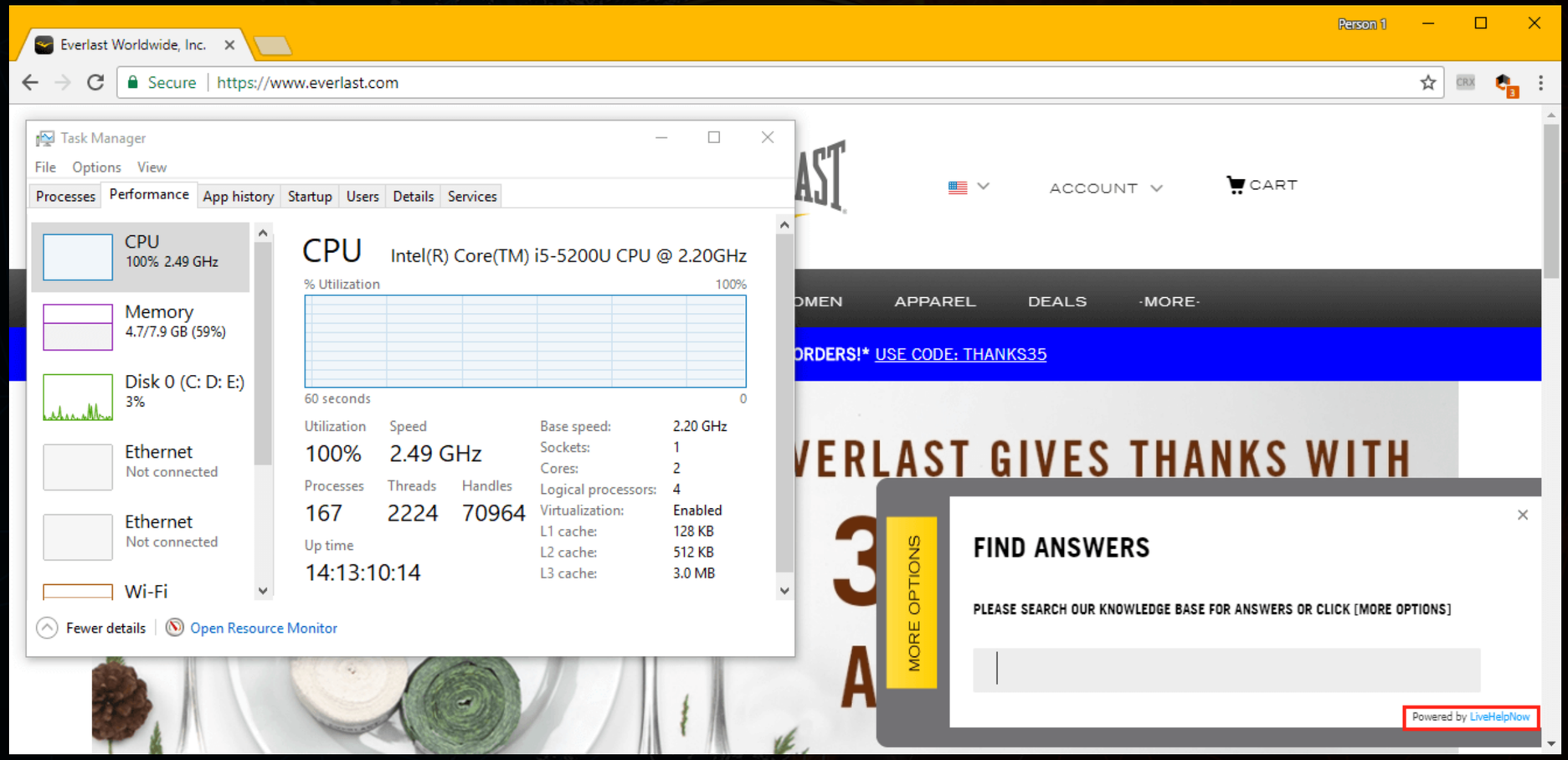
@SimonHearne

Incident 6: the crypto miner

Slide notes

This was reasonably well covered in the press.

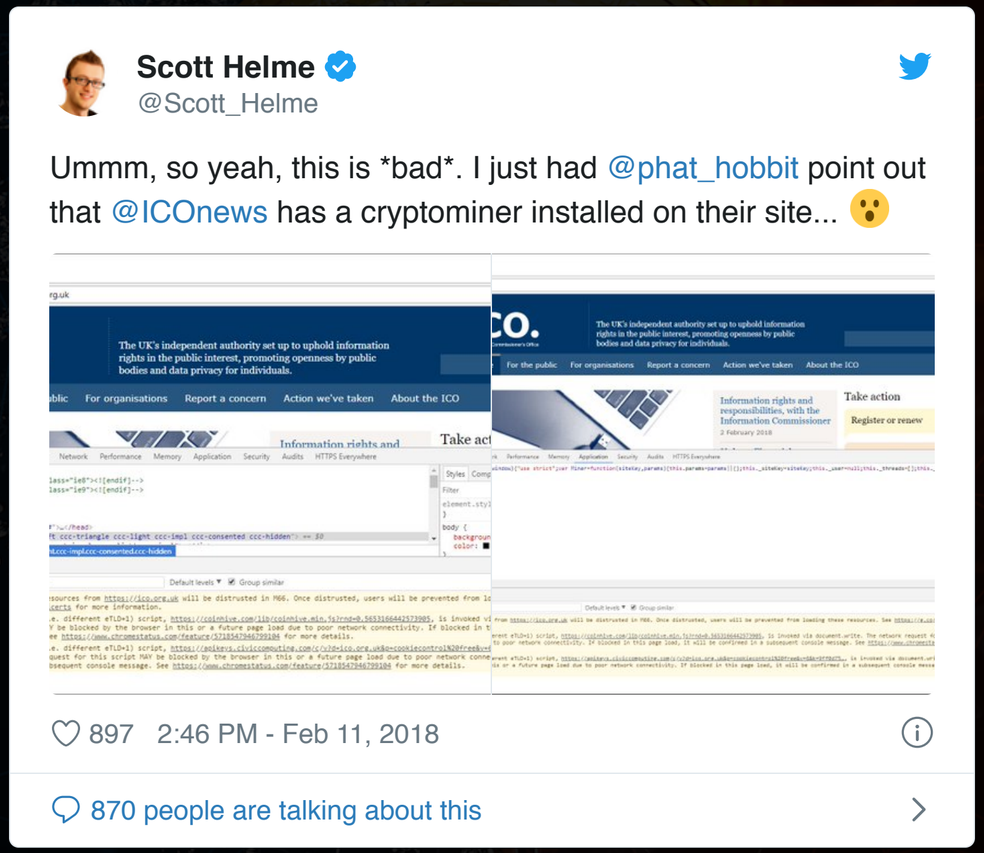
The problem



Slide notes

The issue was caused by crypto miners, written in JS, snuck into third-pary libaries.
The miners pegged the CPU and made the user experience awful.

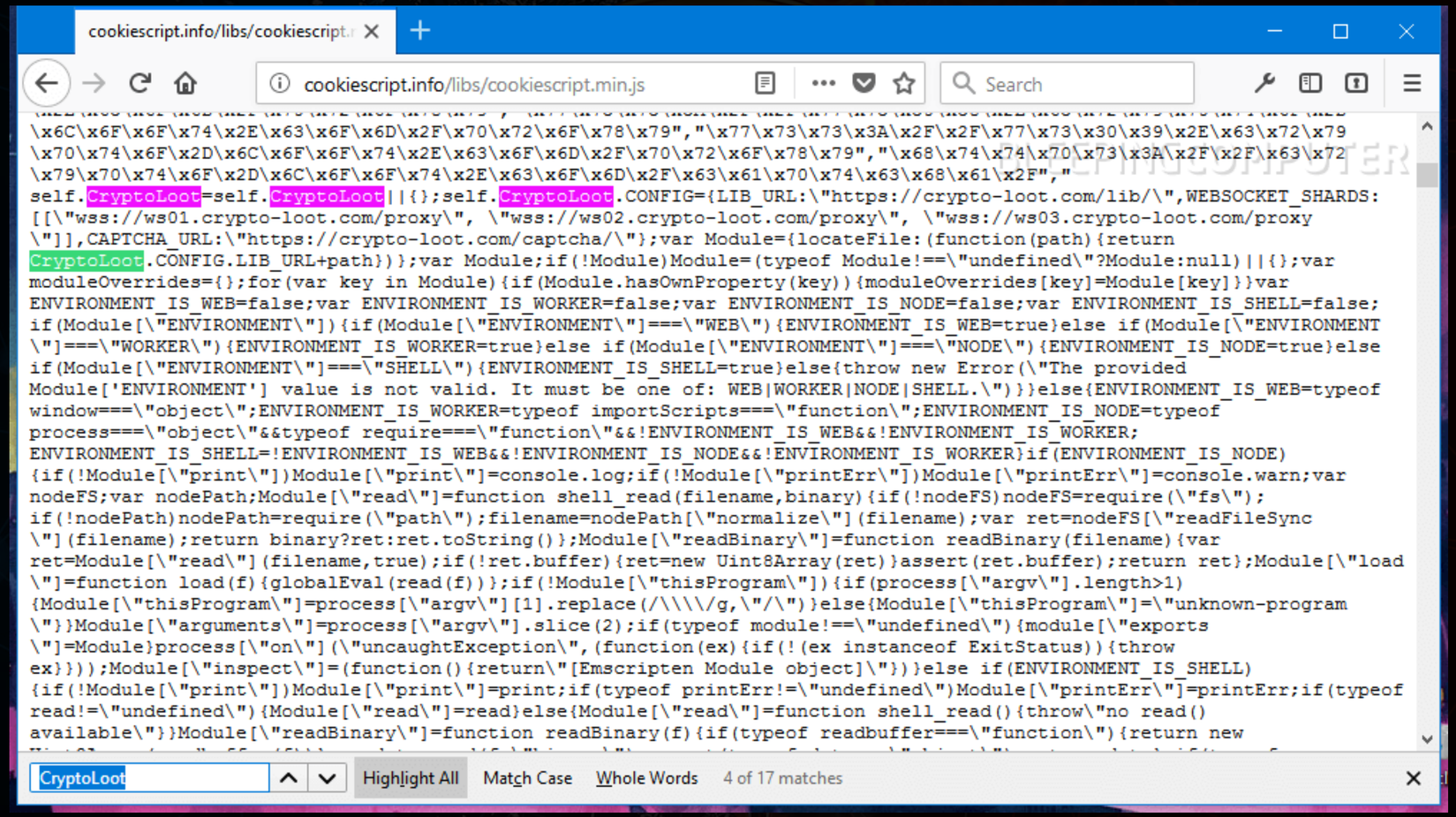
The discovery



Slide notes

Unfortunately for most site, the discovery was through twitter.

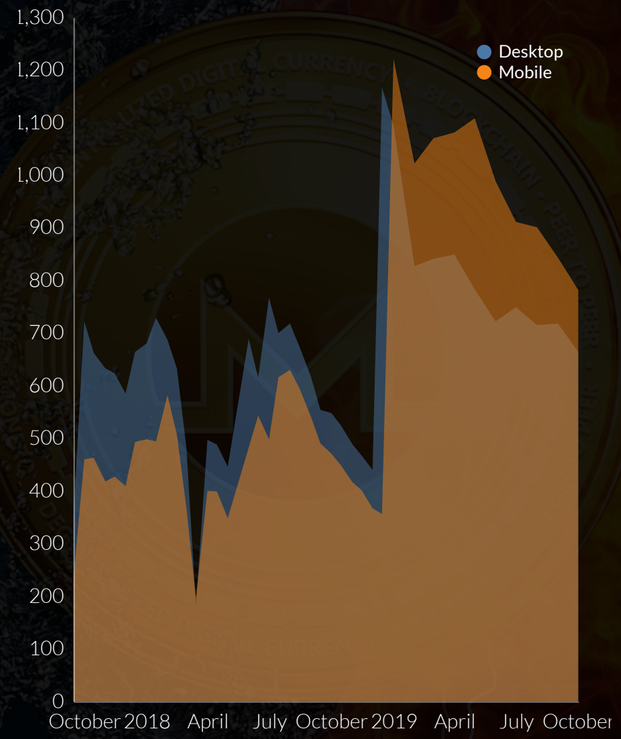
Many compromised services...



Slide notes

@SimonHearne

CryptoMiner Use

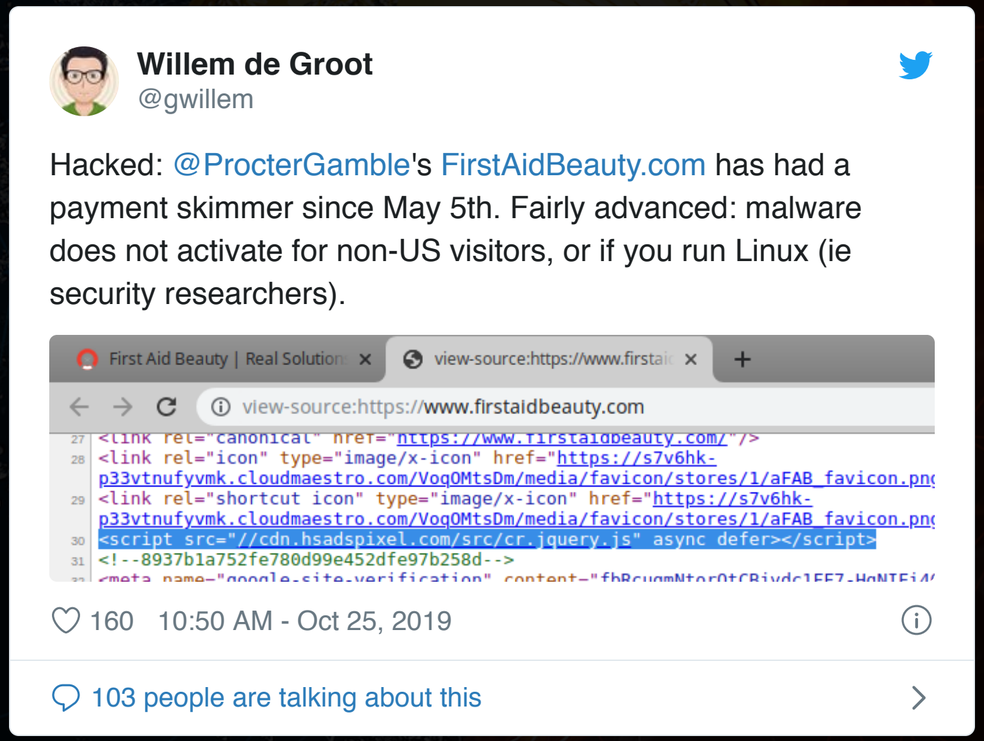


Source: HTTP Archive. 2017-09-15 to 2019-10-01

Slide notes

While the presence of crypto miners is apparently reducing...

Static analysis is not enough



Slide notes

They could just be getting better at hiding.

The defense

- Subresource Integrity
- Content-Security-Policy (connect-src)
- Track CPU metrics (TTI / CPU Idle)

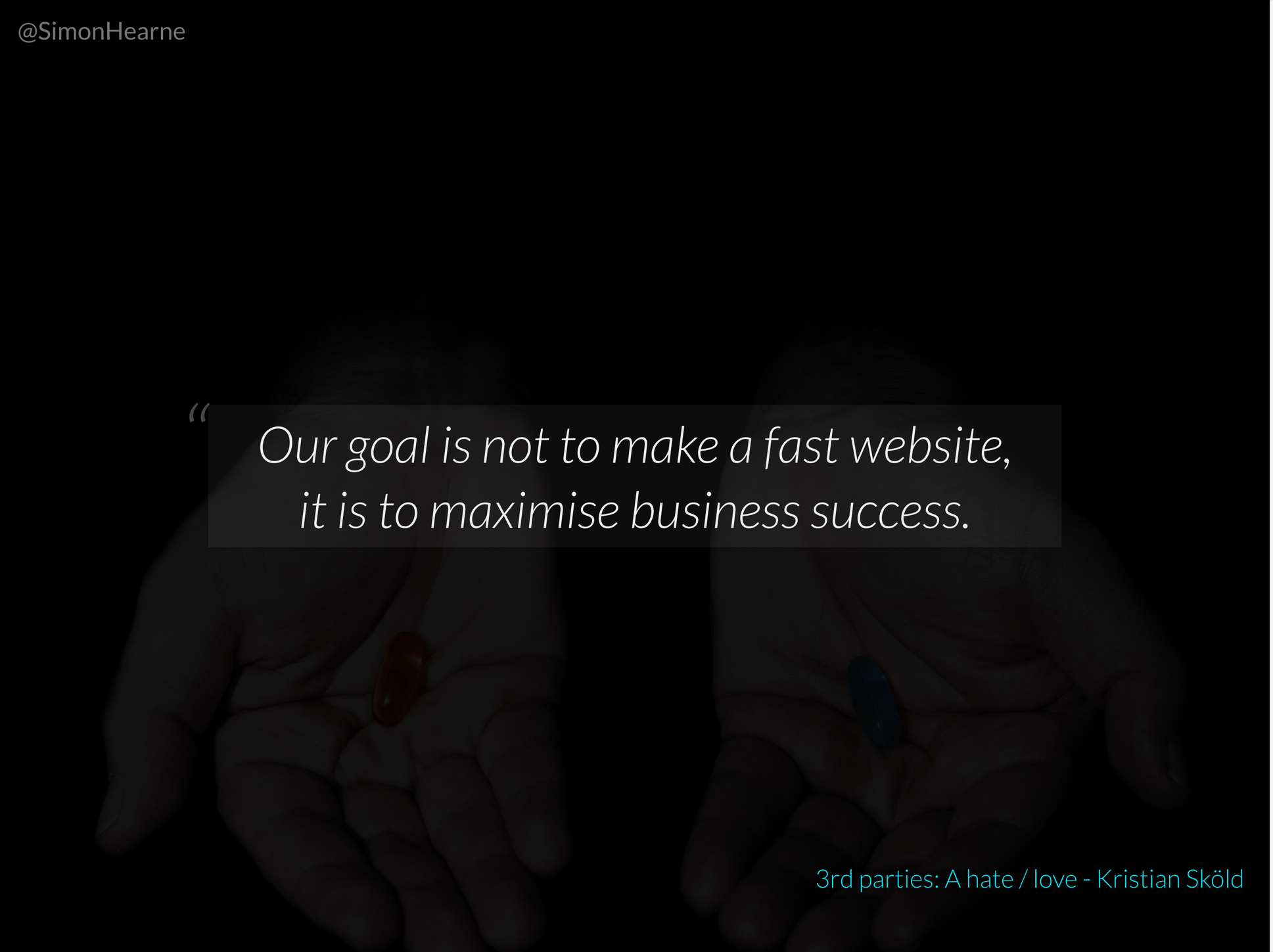
Slide notes

No notes on this slide.

@SimonHearne

Incident 7: the "optimisation"

Slide notes



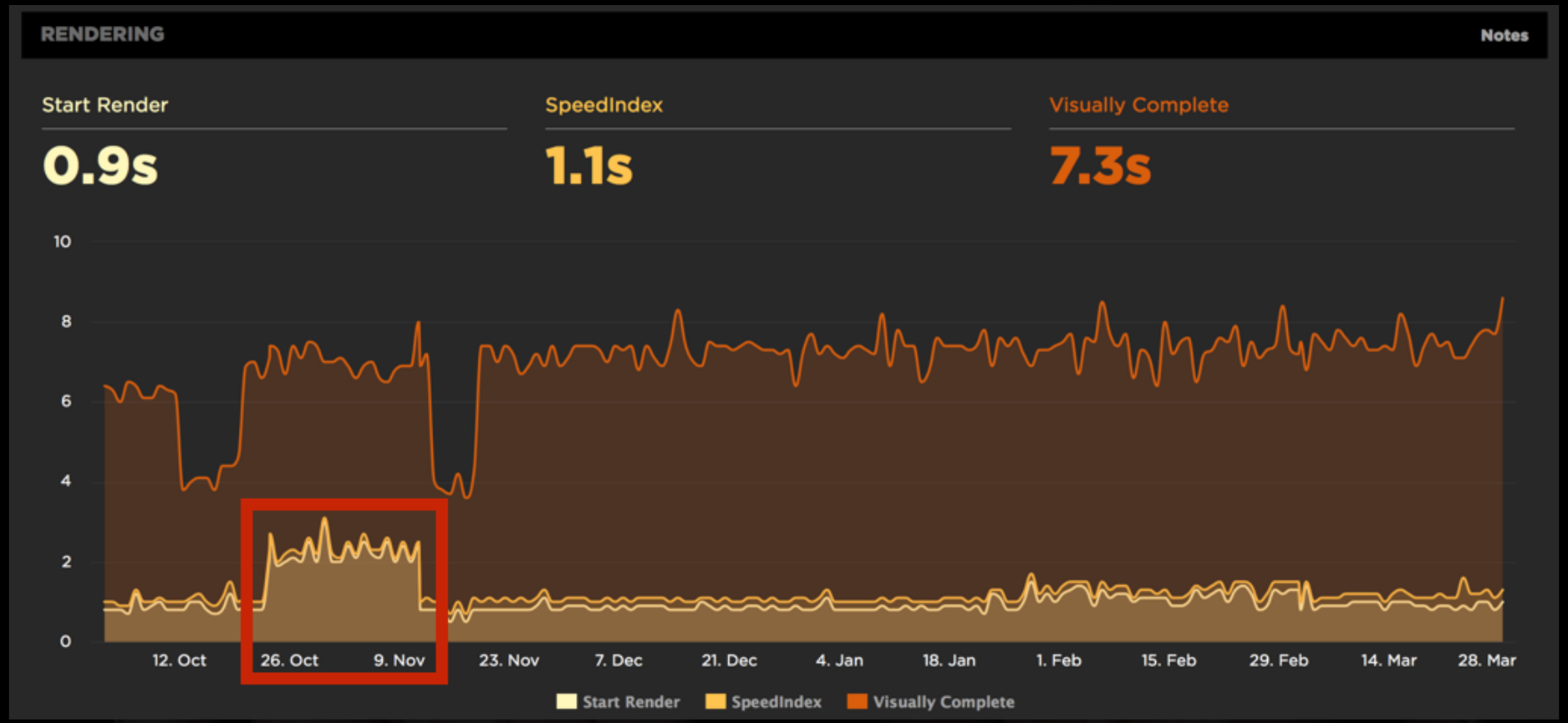
> *Our goal is not to make a fast website, it is to maximise business success.*

3rd parties: A hate / love - Kristian Sköld

Slide notes

No notes on this slide.

@SimonHearne



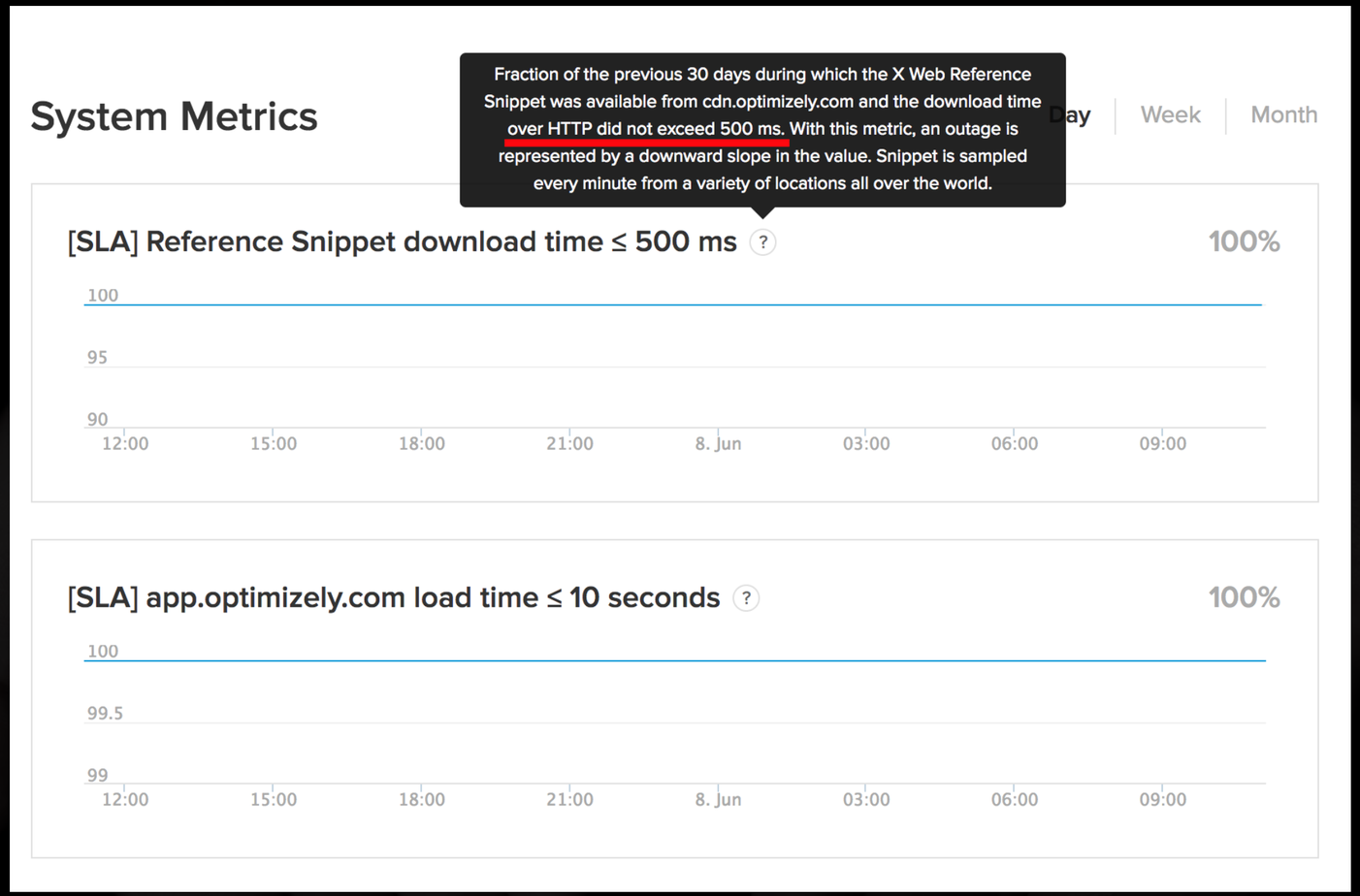
How A/B testing tools worsen your site speed - orangevalley

Slide notes

The customer implemented the performance optimisation tool and performance metrics tanked.

@SimonHearne

What's 500ms worth?



status.optimizely.com

Slide notes

Optimizely's default implementation is a render-blocking script. Their SLA for performance is 500ms. Half a second is a long time to block the user experience.

Proxy through your domain

Optimizely self-hosting for Akamai users

Last updated: Mar 27, 2019

THIS ARTICLE WILL HELP YOU:

- **Set up CDN self-hosting** using Akamai

Self-hosting is ideal for customers who are using both HTTP/2 to serve their website and a CDN. By self-hosting, you can eliminate a SSL connection to Optimizely while using multiplexing to request the snippet faster.

Optimizely self-hosting for Akamai users

Slide notes

The biggest win here is to serve the snippet over your connection. This removes at least three round-trips from the critical path.

Evaluate performance

Top 9 Fastest AB testing tools

Ranked by time to complete test variation

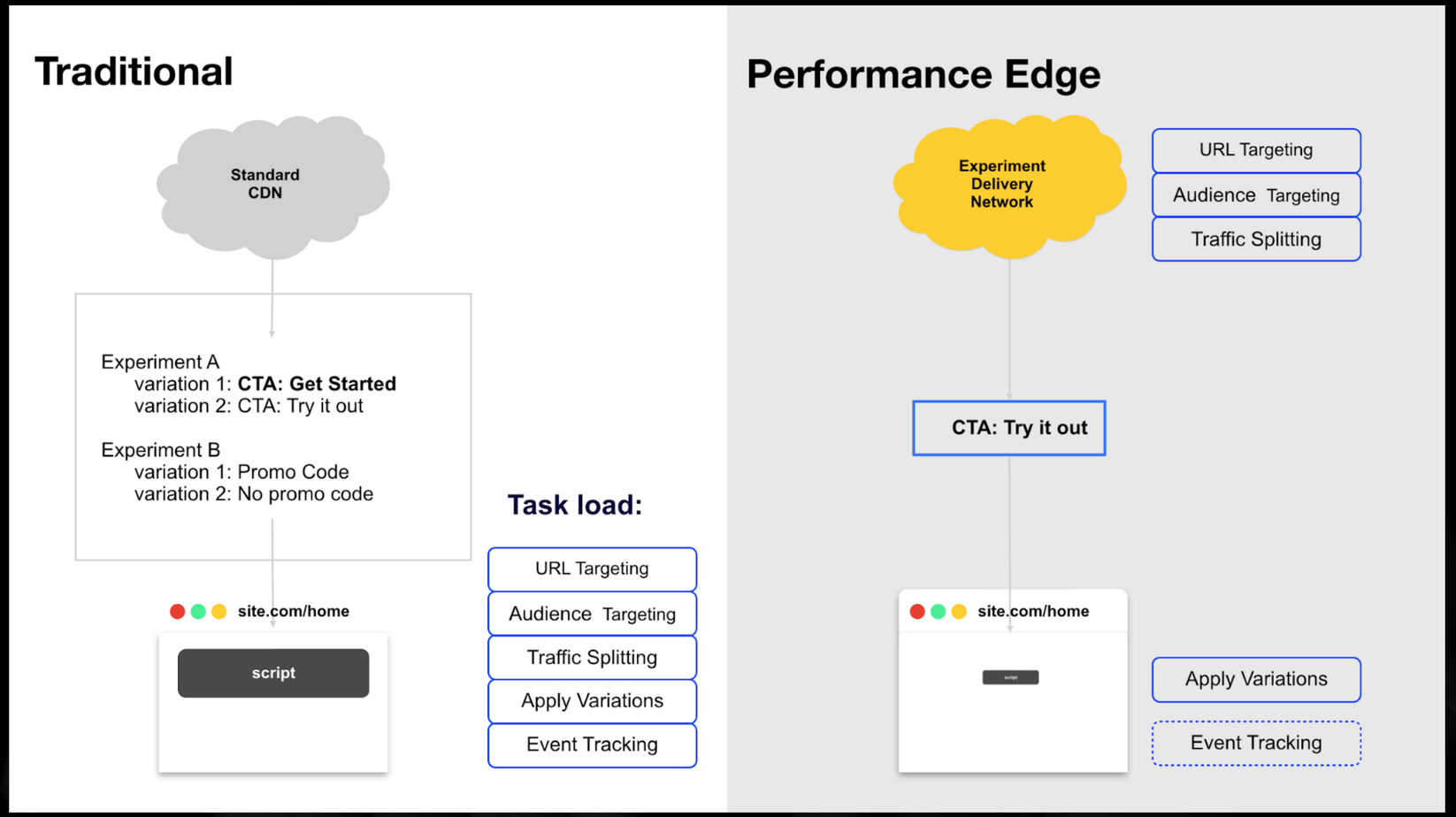
Rank	Tool	Technology	CDN	Snippet type	Snippet placement	Variation complete (s)
1	SiteSpect	Server-side	n/a	n/a	n/a	1.33
2	Maxymiser	Client-side	Akamai	Synchronous	<head>	1.70
3	Convert	Client-side	Akamai	Synchronous	<head>	1.81
4	Optimizely	Client-side	Akamai/Edgecast	Synchronous	<head>	1.86
5	Optimost	Client-side	Akamai	Synchronous	<head>	1.93
6	Marketizator	Client-side	Cloudfront	Synchronous	<head>	2.13
7	Qubit	Client-side	Cloudfront	Synchronous	<head>	2.73
8	AB Tasty	Client-side	Amazon/MaxCDN /Cloudfront	Synchronous	<head>	2.88
9	VWO	Client-side	Private CDN	Asynchronous	<head>	4.29

How A/B testing tools worsen your site speed - orangevalley

Slide notes

Not all tools are created equally, so make sure you evaluate based on performance.

Use the platform



Slide notes

Optimizely knows the issue, and have released a product which dynamically generates the script at the edge to reduce bundle size.

What can we learn?

Slide notes

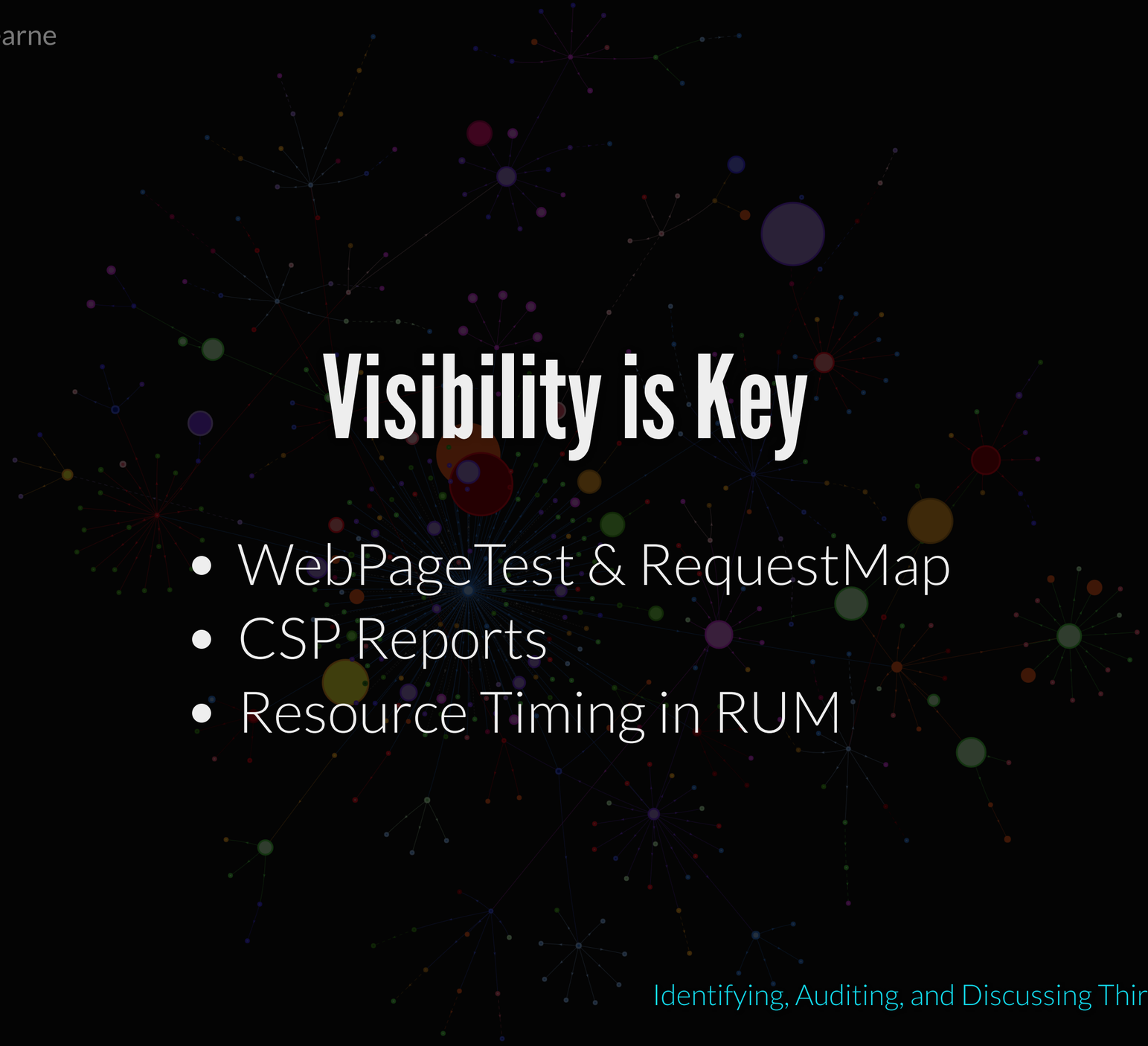
No notes on this slide.

> *Everything should have a value*
> *because everything has a cost*

Tim Kadlec

Slide notes

No notes on this slide.



Visibility is Key

- WebPageTest & RequestMap
- CSP Reports
- Resource Timing in RUM

Slide notes

No notes on this slide.

Determine the risk & value

- WebPageTest Block & SPoF
- RUM Correlations
- Webbkoll by Dataskydd

Slide notes

No notes on this slide.

Remove the unnecessary

- Ownership & ROI
- Adaptive Loading
- Server-side equivalents

Slide notes

No notes on this slide.

Immunise against the necessary

- Proxy through your CDN
- Content-Security-Policy
- Subresource Integrity
- Minimise Customisation

Slide notes

No notes on this slide.

Lock it down!

- Lock down tag manager
- Business process for tags

Adding Controls to Google Tag Manager - Barry Pollard

Slide notes

No notes on this slide.

Thank you

@SimonHearne

simon@hearne.me

simonhearne.com/talks/

noti.st/simonhearne

Slide notes

No notes on this slide.