Reliably Measuring Responsiveness in the Wild



Shubhie Panicker

@shubhie



Nic Jansma

@nicj

When is **load**?



Interimeter and

T theorem

wanted.



media (pro-

pera la comercia.

water and

second data and second

Index Identificant

ALL PARTY



(min) a second second



46.10



13. java - 28. j	1.66	1.1184
ingen (1991)	in	1.294
16. jún († 69)	44	1.00
nijes (. n. j	16	1.009.5
11.1000 1.100	1977	1000
ii.jpe 14	14	2009
14. Java 1. 29	14	1084
11.34mm > 24 - 3	10	1 (144
tobal (1960)	144	1500031

ITT: - Applying Coll. Nan-Thatagling-reactor (spins)
 Theys - galge(TTT limeting impression). Length
 Theys 9 - Theys Spins / ITT Spins

I need with an exampling the interface or constrained and playing the result gost depicts both the first full uner Tricing Constrained in the types. Here, in general, gost is 10% larger than User Tricing Constrained. There uses a first public defect, including is well alone with a result including to the play

Addressing, spatients processing and an include a positivity of a statistic statistic and a statistic statistic and a statistic dependence of the statistic dependence of

Finally, Pyeces using gap compression, you care you call the gap date roots Query limits, an old, secondly 40 breasts to the revealently. Old load metrics don't capture user experience.

We need to rethink our metrics and focus on what matters.



Performance only matters at Load time

My app loads in X.X seconds



Load metrics are NOT a single number





Performance in the **Lab**

Performance in the Real World

Key questions

- What metrics accurately capture responsiveness?
- How to measure these on real users?
- How to analyze data and find areas to improve?



Responsiveness Metrics



Queueing Time



Millions of Long Tasks

Long Tasks on 3 customer sites (daily average)

- Site 1 (Travel): 276,000
- Site 2 (Gaming): 200,000
- Site 3 (Retail): 593,000



What Are LongTasks?





60 fps: An Elusive Dream



Real User Measurement (RUM)

Real world measurement with Web Performance APIs



New Performance APIs and Metrics

- Performance Observer
- LongTasks
- Time to Interactive
- Input Latency



PerformanceTimeline vs PerformanceObserver

// PerformanceTimeline

var entries = performance.getEntriesByType("resource");

```
// PerformanceObserver
var entries = [];
const observer = new PerformanceObserver((list) => {
  for (const entry of list.getEntries()) {
    entries.push(entry);
  }
});
```

observer.observe({entryTypes: ['resource']});



https://github.com/w3c/longtasks

Bad Workarounds

- Timeout polling
- rAF loop

Issues

- Performance overhead
- Battery drain
- Precludes rIC
- No attribution



LongTasks via PerformanceObserver

```
const observer = new PerformanceObserver((list) => {
  for (const entry of list.getEntries()) {
    sendDataToAnalytics('Long Task', {
      time: entry.startTime + entry.duration,
      attribution: JSON.stringify(entry.attribution),
    });
}
```

observer.observe({entryTypes: ['longtask']});



Render & Jank

1000 1000	34'0 0100	
 M. M. M. 1998, J. M. 1998, M. M. 1998, M. 1998,	as task over Taken descents. When more all	
MM Model Sector	A OW, MAD . STATUS OF ANY	
 Janel, Janel, J., Walk, J.,	analysis fills waar Starvast	
Mark 1992 A Mark 200	9,818 (Har, 1759), Aware, 7270, 1896; 831	
Mail Action Mail <	g tokl met 2000, Autom, 13500, tarm str	
 Lind J. Mark. 2000. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and T. Lind J. Mark. 2007. Antonia of Them. Sectors and Basel and Them. 2007. Antonia of Them. Sectors and Basel and Ba	g took! start 20020, Augustur, Ulway, name out	
1014 1014 2014 <td< td=""><td>a bisk l share NYMA down ine 17 has name and</td><td></td></td<>	a bisk l share NYMA down ine 17 has name and	
Mar. 100, 100, 100, 100, 100, 100, 100, 100	A DE LA COMPANY OF A DESCRIPTION OF A DE	
Mar. Mar. Mar. Series (1999). 107 (1999).	KIND DR. CONTRACTOR STR. 1949-194	
Name Name <th< td=""><td>Conversion Conversion Conversion and</td><td></td></th<>	Conversion Conversion Conversion and	
Mar. And Target, Mar. Annual, Mar. Mar. Mar. Mar. Mar. Mar. Mar. Mar.	MAALINE PETTANANAN ATANA MENTAN	
 Half, Tardi, T., W.M., Alarakan, Norsa, and J., and J. S. Alaraka, and J. S. Alaraka, J. S. S. Alaraka, J. S. S.	a Mek Lylam, 17797, Aurilian, 100799, Openic, and	
Mark 1985 - 424 Antones Antones and Antones and Antones and Antones An	tisk rist, 19639, Januar, 2008, anni: art-	
Nel and All All All All All All All All All Al	a taski men 1998, drawner 264mi, meter 448	
Mark More, 1970 - Alexander Company, Mark More, 201 Mark More,	a trak! stare 1.4817, donation. 18 year, manue will	
Mar 1992. "Control of the March 1995. Mar 1995. Mar 1992. "Control of the March 1995. Mar 1993. "Control of the March 1995. Mar 1993. "Control of the March 1995. Mar 1995. "Control of the March 1995." Mar 1995." Mar 1995." Mar 1995." Mar 1995." Mar	a high falses. [17] of Assessing form, since will	
Med. Jane 1997. Januara Januara Jana Jana Jana Jana Jana Jana Jana J	PRETRICT CONSTRUCT OF BRITE REPORT	
Mark and The Constraints of Constraints and Mark and The Constraints and Mark and The Constraints of Constraint	THE CALL PROPERTY OF AND AND ALL	
Mark and The Contract Systems and Mark and Mark and Mark and Systems and Sy	A REAL PROPERTY AND A REAL	
Med Ant: 1017, Junites 299, Antone and Med Ant: 1017, Junites 2018, Antone and Med Ant: 1017, Junites 2018, Antone and Med Ant: 1017, Junites 2018, Junites 2018, Med Antone and Med Ant: 1017, Junites 2018, Junites 2018, Med Antone and Med Ant: 1017, Junites 2018, Antone and Med Ant: 1017, Junites 2018, Antone and Med Ant: 1017, Junites 2018, Junites 2018, Med Antone and Med Ant: 1017, Junites 2018, Junites 2018, Med Antone and Med Antone 2017, Junites 2018, Junites 2018, Med Antone and Med Antone 2017, Junites 2018, Junites 2018, Med Antone and Med Antone 2017, Junites 2018, Junites 2018, Med Antone and Med Antone 2017, Junites 2018, Junites 2018, Junites 2018, Med Antone and Med Antone 2017, Junites 2018, Junites 2018, Med Antone and Med Antone 2018, Junites 2018, Junites 2018, Med Antone and Med Antone 2018, Junites 2018, Junites 2018, Junites 2018, Med Antone 2018, Junites 2018, Junit	a total start, 12235, duration, 178mm, name, sold	
Next NEXT With Johnson Proc. Letter, NY, Next NEXT WITH JOHNSON PROC. NEXT Next NEXT WITH AND PROC. NEXT NEXT NEXT Next Next WITH AND PROC. NEXT NEXT NEXT NEXT NEXT NEXT NEXT NEXT	stablater 11045, developer 200mm energy and	
NAX 2012 IN A DECEMPTING OF A	this statt will cannot blow turns will	
MAR AND TAX DEVELOPMENT FOR ANY	task over Dist. Provide: Miles water all	
Mark 1987, Tel Jarone, Tel Anne, 197 Mark 1987, Tel Jarone, Mark, 1997 Mark 1987, Tel Jarone, Mark, 1997	A STATE OF THE PARTY OF THE PAR	
Mark 100, 100, Markes, 2010, Mark 201 Mark 100, 201, Annual 2017, Mark 201 Mark 100, 201, Annual 2017, Mark 201 Mark 100, 201, 2010, 101, Mark 201	A MAY THE STATE OF A DESIGN AND A STATE OF A	
Mark Ture, 2125, Decemen, Cons., and Const., and Const	ANALASS, AND AND AND A COMPANY AND	
MAX AND STATUS AND	g hist rise; 5178, datasion; 1878a, name; 167	
Mark June, 2002, January, Chron, Anne, and Mark 1992, 2005, Allena, Anne, and Mark 1992, 2007, Dannan, Marka, Anne, and Mark 1992, 2007, Dannan, Marka, Anne, and Marka, 1992, 2007, Dannan, Marka, Anne, and	g task? more 5942, months: 212.00, motor: with	
Incl. Law, 1995, Lawler, S. Clam, Marr, and Ball, and D. C.V. Borner, 1975, Marr, 197 Wall and T-Hill, Improve Silver, Awar, eff	a book? starr 4771, formour \$ lact, torne will	
Bak and July doctor. In the same at We and the second line, and at We are the second line, and at	a tick/ slaw: 1890, departure: 247mm, marrier and	
	a tage start. 17 M. datastor 14 Terr same and	
	VARATION SPACE ACCURATE ACCURA	
	THAT HAS THE DESIDER THAT WERE NOT	
	AVAI 1002 - HUL, DEDEN SLAN, AND WE	

https://w3c.github.io/longtasks/render-jank-demo.html



Multiple sub-tasks (scripts) within a long task

Attribution: Who?

"Minimal Frame Attribution" with name

 self, same-origin-ancestor, same-origin-descendant, cross-origin-ancestor, cross-origin-descendant, multiple-contexts, unknown etc.

Attribution: Who And Why?

Detailed attribution with TaskAttributionTiming

- attribution[]
 - containerType: iframe, embed, object
 - o containerSrc: <iframe src="http://..." />
 - o containerId: <iframe id="ad" />
 - o containerName: <iframe name="ad-unit-1" />

More Attribution: Coming soon!

Detailed attribution with TaskAttributionTiming

- attribution[]
 - containerType: iframe, embed, object
 - containerSrc: <iframe src="http://______Long Tasks V2!
 - o containerId: <iframe id="ad" /-</p>
 - containerName: <iframe_name="ad-unit-"
 - o scriptUrl:

https://connect.facebook.net/en_US/fbevents.js:97

LongTasks: Usage Tips

- Measuring during page load: Turn it on as early as possible (e.g. <head>)
- Measuring during interactions with a circular buffer
- First-party ("my frame") LongTasks give only duration
- Third-party other-frames provide attribution if the IFRAME itself is annotated via id, name or src.

Time to Interactive

Is it Usable?

Time to Interactive

User Sees Content

Not Interactive Until Here

🦁 Developer Tools - http://www.cnn.com/					- 🗆 🗙
🕞 🖬 🛛 Elements Console Sources	Network Performance Me	emory Application Se	curity Audits EditThis	Cookie »	8 14 ▲ 1 ↓ 1
● C 🛇 🗹 Screenshots 🗎 Memory	1				\$
1000 ms 2000 ms	3000 ms 4000 ms	5000 ms	6000 ms 7000 ms	8000 ms	9000 ms
and the second	A IA	h maleda	Lilla all		СРО
				an Brit In Brit In B	NET
1000 ms 2000 ms 3	3000 ms 4000 ms	5000 ms 60	00 ms 7000 ms	8000 ms	9000 ms
Frames 633.4 ms 116 <mark>2</mark> .1 m	ıs 457.1 ms 437.9 ms	439.6 ms			^
► User Timing					
▼ Main Pars1]) Ev1) Ev1 Ev1) Ev1 Ev1) Ev1 Ev.		Ti) Rs (a) (a) i R			

Time to Interactive: Lower Bound

When does the page appear to the visitor to be interactable?

Start from the latest *Visually Ready* timestamp:

- DOMContentLoaded (document loaded + parsed, without CSS, IMG, IFRAME)
- First Paint, First Contentful Paint
- Hero Images (if defined by the site, important images)
- Framework Ready (if defined by the site, when core frameworks have all loaded)



Time to Interactive: Measuring

What's the first time a user could interact with the page and have a good experience?

Starting from the lower bound (Visually Ready) measure to *Ready for Interaction* where none of the following occur for your defined period (e.g. 500ms):

- No Long Tasks
- No long frames (FPS >= 20)
- Page Busy is less than 10% (setTimeout polling)
- Low network activity (<= 2 outstanding)

github.com/GoogleChrome/tti-polyfill

github.com/SOASTA/boomerang/tree/continuity

Input Latency

Measuring bad user experiences

- Interactions (scrolls, clicks, keys) may be delayed by script, layout and other browser work
- Latency can be measured (performance.now() event.timeStamp)
- Latency can be attributed via LongTasks

Measure input latency: event.timeStamp and performance.now()

const subscribeBtn = document.querySelector('#subscribe');

```
subscribeBtn.addEventListener('click', (event) => {
    // Event listener logic goes here...
```

```
const lag = performance.now() - event.timeStamp;
if (lag > 100) {
   sendDataToAnalytics('Input latency', lag);
}
```

});

Input Latency

Determining the cause via LongTasks:

- 1. Turn on PerformanceObserver
- 2. Watch for input delays
- 3. Find LongTasks that ended between event.timeStamp and performance.now()

Sample code:

github.com/nicjansma/reliably-measuring-responsiveness-in-the-wild/

14141WNEARD

Real World Data

Case Studies

- 3 sites over 1 month
 - Site 1: Travel (ads, social)
 - Site 2: Gaming (ads, social)
 - Site 3: Retail (social, 3p, spa)

18+ million LongTasks

LongTask Duration

1000000

750000

500000

250000

Count



- 50th: 106 ms
- 75th: 208 ms
 - 90th: 427 ms
- 95th: 666 ms
 - **99th: 1,712 ms**
 - Range: 50 to 10+ seconds



Bucket (ms)

LongTasks as % of Front End Load Time



LongTasks directly delay Time to Interactive.

First Page TTI vs Conversion Rate



First Page TTI (s)

Time to Interactive has high correlation with overall conversion rate.



First Page LongTask Time vs. Conversion Rate

First Page LongTask Time vs. Session Length

First impressions matter: as first-page LongTask time increased, overall Conversion Rate decreased

Desktop vs. Tablet vs. Mobile



Site 1 Site 2 Site 3

Mobile devices could see 12x LongTask time as Desktop.



LongTask • Other

Older devices could be spending half of their load time on LongTasks.







Optimizing Performance

Every site is different.

Identify your core metrics.

Minimize the time to TTI

Consider Mobile traffic Ship less JS Break up existing JS with Code Splitting

Reduce Long Tasks

Mobile is especially hurting Break up JS

Move intensive work off the main thread to workers

Hold Third Parties Accountable

Identify the worst offenders Evaluate impact on TTI & business metrics

Looking Ahead

- Long Tasks V2
- Input Latency + Slow Frames
- Long Tasks is not Panacea

Thank You



Shubhie Panicker @shubhie



Nic Jansma