“...we don’t plan on stopping until the web is instant...”

/15

URS HOELZLE
Pick a query, any query. ‘Weather, New York City.’ ‘Nineteenth-century Russian literature.’ ‘When is the 2012 Super Bowl?’ Now type it into a Google search box. As you type, we predict the rest of your query, comb through billions of web pages, rank the sites, images, videos, and products we find, and present you with the very best results. The entire process takes, in many cases, less than a tenth of a second – it’s practically instant.

If it isn’t, we’ll suffer. Our research shows that if search results are slowed by even a fraction of a second, people search less (seriously: A 400ms delay leads to a 0.44 percent drop in search volume, data fans). And this impatience isn’t just limited to search: four out of five internet users won’t bother waiting if a video stalls while loading – they’ll simply click away. But even though the human attention span has become remarkably fickle, much of the web remains slow. The average web page takes 4.9 seconds to load – in a world where fractions of a second count, that’s an eternity.

The web has become a critical hub for politics, schools, and entertainment. Every business is a digital business; large or small, local or multinational. So why is it okay for a web page to take five seconds to load? There are 245 million internet users in the US, and if every one of them has to wait five seconds, we just wasted nearly 39 years of their time.

‘Fast is better than slow’ has been a Google mantra since our early days, and it’s more important now than ever. The internet is the engine of growth and innovation, so we’re doing everything we can to make sure that it’s more

‘Fast is better than slow’ is a cornerstone of Google’s philosophy. Here, search guru and SVP of Infrastructure Urs Hoelzle, explains why.

WORDS BY Urs Hoelzle
ILLUSTRATIONS BY Adam Simpson
Formula 1 than Soap Box Derby. Speed isn’t just a feature, it’s *the* feature.

We have one simple rule to support this Gospel of Speed: Don’t launch features that slow us down. You might invent a great new feature, but if it slows down search, you have to either forget it, fix it, or come up with another change that more than offsets the slowdown. We have what we call a ‘fixed latency budget,’ which is sort of like a family budget. If you want to go on a nicer vacation but your budget doesn’t stretch, you need to cut back somewhere else.

This simple concept drives legions of Google engineers and product managers to do some pretty amazing things. It’s why, when you do a Google search from some remote corner of the world, your results are most likely served to you from nearby computers. We work to cache data in local facilities, with the objective of making Google nearly as fast in San José, Argentina, or San José, Costa Rica, as it is in San José, California.

It’s why we have live performance dashboards on big screens in many of our engineering offices, so that teams can see latency levels across our services. It’s why, a few years ago, when we failed to live up to our principles and things started to slow down, we called ‘Code Yellow!’ and directed engineers and product managers on major product teams to drop what they were doing and work on making stuff faster. Speed is simply part of our engineering culture.

Of course, it doesn’t really matter how fast search is if, when you click on a result, you immediately move back into the slow lane. That’s why we invest so much in helping the rest of the web speed up, too. Google Analytics measures a site’s speed and how it impacts engagement. We’re spearheading Page Speed, an open-source project that helps webmasters speed up their sites—it can even re-write pages to boost performance. We’re also experimenting with a Page Speed Service that automatically accelerates page loads without any code changes required. Just route your page through the service and it gets faster.

There’s lots of other things we’re doing to make the web faster, including working with the web community to update standards like HTML and TCP/IP alongside core network protocols such as DNS, TCP, SSL, and HTTP, or improving the speed of JavaScript.

Our open-source Chrome browser is now six times faster than when it first shipped three years ago, and pre-fetches certain web pages during searches so that when users click on those links they load instantly. Other popular
browsers such as Firefox, Explorer, and Safari have all upped their speed since Chrome entered the race.

And earlier this year we announced plans to build an ultra-high bandwidth fiber network in Kansas City (both Missouri and Kansas), giving its citizens internet access that’s more than 100 times faster than most Americans enjoy today. Our hope is that, like Chrome, this project will motivate other internet providers to crank it up, too.

All this speed makes a difference. When Edmunds, a leading car review destination, re-engineered its insideline.com site to reduce load times from nine seconds to 1.4 seconds, ad revenue increased three percent, and page views-per-session went up 17 percent.

When Shopzilla dropped latency from seven seconds to two, revenue went up seven-12 percent and page views jumped 25 percent. (By the way, they reduced their hardware costs by 50 percent.) When you speed up service, people become more engaged – and when people become more engaged, they click and buy more.

Even so, we need to set our expectations higher. At Google, we don’t plan on stopping until the web is instant, so that when you click on a link the site loads immediately, or a video starts without delay. What amazing things could happen then? What else could be invented?

I don’t know. All I can say for sure is that we want to get to that future, fast.