

What Does A Google Search Really Cost?

Mary V. Connolly
Mathematics and Computer Science
Saint Mary's College
Notre Dame, IN 46556
connolly@saintmarys.edu

Abstract

An article posted in TimesOnLine early in January claimed that two Google searches generated about the same amount of carbon dioxide as boiling water for a cup of tea. Google was quick to respond and to disagree. What are the real costs of such a search? Should the IT department at a college campus even worry about this? This paper will explore some of the issues which should be considered.

The Debate

On January 11, 2009 the Sunday Times of London published an article claiming that two Google searches performed on a desktop computer generate as much carbon dioxide as boiling water for a cup of tea. [7] The article cited the research of Alex Wissner-Gross, who is an Environmental Fellow at Harvard with special interests in green computing, energy and computation. Wissner-Gross, together with Tim Sullivan, developed a Web site, CO2stats.com, designed to educate people about energy efficiencies. They hope to leverage the growing energy consumption of the internet to the increasing demand for renewable energy. Although the Times article linked its claim about the environmental cost of a Google search to Wissner-Gross's research, in fact his research never mentions Google and focuses instead on the Web overall. He found that it takes on average about 20 milligrams of CO₂ per second to visit a Web site. [9] Google was quick to dispute the Times' claim as well. In fact the founders of Google, Larry Page and Sergey Brin, are really dedicated to green initiatives. Google claims that an average query uses .0003kwh. In terms of greenhouse gases, this is equivalent to about 20 milligrams of CO₂ (in contrast to the 7grams that would be equivalent to the cup of tea example).The Times issued a clarification on January 16 and accepted the 20 milligram figure for a one-hit search taking less than a second.

The Real Issues

The important question is not really the energy used for a single "average" search, but rather the environmental consequences of the internet as a whole. A recent report by the American research firm Gartner suggests that the global IT industry causes 2% of global emissions, exceeding that produced by the world's airlines. Estimates of the number of internet searches done daily across the globe range from 200 million to a billion. The one billion estimate comes from Google, which estimates the environmental cost as equivalent to driving a car 1,000,000 km. What is needed for a single search? If you wish to perform a search, you first need to type in the search, assuming that your computer is already turned on. Hence there is an energy cost for the personal computer being used. There is also the energy needed for the network. The search request then probably goes to multiple servers in multiple data centers. Each server must be prepared to receive your request, so there is work to be performed before the search

2009 ASCUE Proceedings

starts, work such as building a search index. In addition to the power used for each server, there is the overhead power needed to run the data centers. All of this has to be factored into any calculation on the total amount of power used for a search. Of course, searches vary widely in their complexity, with some needed far more time than others. Google's estimate of 20 milligrams of CO₂ for an average search only takes its part of the equation into account. The company in fact works hard to guarantee that its searches are efficient.

Among the company goals are goals to minimize the electricity used by its servers and reduce the energy used by the data center facilities themselves [3]. Of course, as Nicholas Carr (author of the Atlantic Monthly article 'Is Google Making Us Stupid') pointed out, Google has a bit of a problem. The company is definitely dedicated to energy efficiency but also to getting people to spend as much time on the internet and their computers as possible. After all, Google works on an ad-based business model. [2]

The Nature of a Search

Consider what happens when you need a particular piece of information. Before the age of computers this would have involved a phone call or a visit to the reference desk of a library. Now that computers can store vast amounts of information we expect to get online and find what we need. Google's mission is to organize the world's information and make it universally accessible and useful. A typical search request does not go just to one data center, but to several. Google does not make public the number and locations of its data centers, considering this to be proprietary information. The company does claim that the first hit is generally found quickly, but note that speed is accomplished by using multiple servers rather than just one for a single search. However some searches are inherently more complicated than others. Also, the Web was not set up with efficiently searching as the goal. It is almost fair to say that we set up the haystack first and now we want to search it. Many publishers have multiple URLs that all point to the same page, causing search engines to index the page multiple times. A good search engine will have some way to detect duplicates. In fact, Google and its rivals Yahoo and Microsoft have agreed to support a new standard (the Canonical Link Tag) that would allow publishers to get rid of the duplicate pages. Adherence to such a standard would allow both more efficient and more comprehensive searches. [5]

We, of course, don't always go to Google. There are plenty of sites dedicated to particular needs that also run search engines, searching databases that store focused information such as flight schedules, store catalogs or financial information. A company like Google does collect an enormous number of addresses of Web pages (adding its one trillionth address a year ago), but the search strategies used work best for the surface Web (gathering information by following the trails of hyperlinks) as opposed to searching dedicated databases (with something called a Deep Web search strategy). Google is indeed exploring Deep Web search strategies, using a strategy that involves sending out a program to analyze the contents of any database it encounters. Others are also involved in developing these newer search strategies, which pose thorny computational challenges. As search engines begin to incorporate Deep Web content into the search results, there is another problem. How does one present different kinds of data without making the search pages impossibly complicated? [10]

2009 ASCUE Proceedings **Thinking Green**

There are multiple ways in which to reduce the environmental impact of a Google (or Yahoo or Microsoft) search. Although clearly efficient search algorithms are important, one also has to consider the resources used to run the data centers. The EPA, in an August, 2007 report, estimated that energy consumption at servers and data centers has doubled in the past five years and will almost double in the next five years. It has been developing Energy Star standards for servers, focusing first on the efficiency of the servers' power supply and its energy consumption while idle. [6] A growing trend is to use virtualization to improve data center efficiency. Data centers often have underutilized servers. Consolidation using virtualization increases server efficiency and lowers energy consumption. [4] Hewlett-Packard has been able to consolidate 86 data centers into three, with three backups by using virtualization. [1] Google claims that the data centers it has designed use considerably less energy than a typical data center. Given that the company runs multiple data centers, this clearly is in their best interest. Google uses customized evaporative cooling in its data centers. Two of their facilities currently run on 100% recycled water, and by 2010 the company expects recycled water to provide 80% of the total water consumption at their data centers. The company claims that among its goals are goals to minimize the electricity used by its servers and to reduce the energy used by the data center facilities themselves. [3] Microsoft has built data centers in central Washington powered by hydroelectricity, power produced by the two dams in the region. Another Microsoft data center in Dublin, Ireland (expected to be operational this year) will be air cooled, due to the moderate climate in Ireland. Yahoo's data centers are carbon neutral due in part to carbon offsets. [6] Yahoo also took advantage of the Pacific northwest climate to build a new data center in Quincy, Washington knowing that it would use less air conditioning.

We also must consider what happens at the beginning end of a search. Clearly someone has to be using a computer, which may or may not be turned on when the information for which we are searching is needed. Once the computer is powered, we need to start with some kind of interface. Google's screen typically has a white background. A number of groups have worked to make the interface more environmentally friendly. For example, Greenlinking.com offers a wrapper that let's you use Google alone or Yahoo, ninemsn and Google combined. The site purchases carbon credits via Carbon Planet Offset of 20kg. of greenhouse gases per user per month. Ecocho offers a wrapper over a Google search which gives the users the opportunity to purchase carbon offsets by performing searches. Their estimate is that they grow 2 trees for every 1000 searches. Another approach is to make the wrapper itself more environmentally friendly. Both Blackle and Eco-Find offer black backgrounds for Google searches. The savings are achieved if the user has a CRT monitor, but there is probably not a significant savings for an LCD screen. [8]

Conclusions

If Google, Yahoo and Microsoft all wanted to make reducing energy consumption their number one goal, they could encourage us to reduce our time on line. Perhaps they could even give us a tool on the toolbar that would keep track of the grams of CO₂ we emit as we are on line, as Nicolas Carr suggested (somewhat facetiously). Of course, this isn't likely to happen. These companies are in business to make money, and the business requires that we be on line. It is a good thing that all three companies are working on controlling the environmental impact of the business at their end. That leaves the consumers to consider their end of the equation. Just maybe not all those hours of recreational searching are that

2009 ASCUE Proceedings

critical. Certainly we can be sure that our computers are appropriately controlled when not in use, and turned off at the end of the day. The idea of that tool on the toolbar is a bit frightening for what it might reveal. Perhaps we are not that far away from such a tool. Wissner-Gross's company, CO2stats, has developed a software suite that monitors a web site's energy usage, supplies hints on how to make the site more energy efficient and purchases renewable energy from wind and solar farms. A student, when told about the topic of this paper, asked "Just how guilty do I have to be?" The correct answer lies not with guilt but with awareness. If all of us paid more attention to the environmental consequences of our searches, not just leaving the concerns to the Googles of the world, we could make a significant difference.

References

- [1] Biello, David. (2008) "Digital Diet." Scientific American. April, 2008: p. 26.
- [2] Carr, Nicholas. (2009) "Strip mine media." Accessed 11 January 2009.
www.rougtype.com/archives/2009/01/the_dead_coal_n.php
- [3] Google. (2009) <http://www.google.com/corporate/green/datacenters>
- [4] The Green Grid. www.thegreengrid.org
- [5] Helft, Miguel. (2009) "Google, Yahoo and Microsoft Collaborate to Clean Up Web." New York Times. 12 February 2009. Accessed 12 February 2009
<http://bits.blogs.nytimes.com/2009/02/12/google-yahoo-and-microsoft-collaborate-to-clean-up-web/>
- [6] Kurp, Patrick. (2008) "Green Computing." Communications of the ACM. October, 2008: p.12.
- [7] Leake, Jonathan, and Woods, Richard. (2009). "Revealed: the environmental impact of Google searches." The Sunday Times. 11 January 2009. Accessed 12 January 2009
<http://technology.timesonline.co.uk/tol/news/tech_and_web/article54>
- [8] Perez, Sarah. (2008) "15 of the Best Green Search Engines." ReadWriteWeb. 22 April 2008. Accessed 9 April 2009
www.readwriteweb.com/archives/15_of_the_best_green_serach_engines.php
- [9] San Miguel, Renay. (2009) "Harvard Physicist Sets Record Straight on Internet Carbon Study." TechNewsWorld. Accessed 12 January 2009
<www.technewsworld.com/story/Harvard_Prof_Sets_Record_Straight_on_Internet_Carbon_Study_65794.html#>
- [10] Wright, Alex. (2009) "Exploring a 'Deep Web' That Google Can't Grasp." New York Times. 23 February 2009: pp B1, B4.