

Cookies

The cookies technology was developed at Netscape. We begin by discussing how and why Netscape developed cookies. Netscape's cookies technology led the Internet Engineering Task Force to develop a precise technical standard for cookies. This effort by a consortium is examined in the second section.

Netscape's Cookies

In December 1993, a bitter Andreessen graduated from the University of Illinois. By March, he was talking to Jim Clark about a potential new Internet company (Clark, 1999). Andreessen next persuaded almost all the core developers of NCSA Mosaic to leave NCSA and to join him at Mosaic Communications Corp., which eventually become Netscape Communications.

The new company would make money by selling web servers. According to Jim Clark, the profit margin on web browsers was slim, but significant on 50,000 secure server applications. These secure web servers would be in demand by corporations seeking to make money over the Internet. This business decision led to an emphasis on security, commerce, and performance in both web servers and browsers. This led Netscape to develop new technologies such as cookies, continuous document streaming, and Secure Sockets Layer (Reid, 1997). These new technologies would be incorporated in the new Netscape Enterprise Server as well as in the new browser code name Mozilla. This new "killer" browser was named Mozilla, using a mixture of Mosaic and Godzilla.

The cookies technology was the most innovative feature and one that would forever alter the web. According to Lessig, "before cookies, the Web was essentially private. After cookies, the Web becomes a space capable of extraordinary monitoring"

(Schwartz, 2001). In early web browsers, the Internet was a stateless place as intended by Berners-Lee (1991). A stateless web is analogous to a vending machine. It has little regard for who you were, what product you are asking for, or how many purchases you have made. It has no memory. Statelessness on the web made commerce difficult. Without a state mechanism, buying goods is analogous to using a vending machine. You could not buy more than one product at a time and there would be no one-click automated shopping feature that remembers your personal information.

The statelessness problem concerned the Netscape Enterprise Server Division, which was working on a contract for a new shopping cart application for online stores. A shopping cart would allow a web site to keep track of multiple items that a user requested. The current methods for shopping carts involved storing state information in the web address or Uniform Resource Locator (URL). However, these methods did not work very well, so the server division was open to ideas. This led to the idea state that the state data needs to be stored somewhere else other than the URL (Montulli, 1999). Over the next few weeks, Lou Montulli and John Giannandrea refined their ideas into a solid working concept termed, Persistent Client State HTTP Cookies (Netscape Communications Inc., 1994; St. Laurent, 1998). Over the next year, they would quietly prepare and submit this idea for a patent, which they received in 1998 (Netscape Communications Inc., 1998).

Programmers used the term cookies to refer to a small data object passed between cooperating programs. Similarly, Netscape would use cookies to pass information between a user's computer and the web site they were visiting. Nowadays, Lou Montulli is known as the "Father of the Web Cookie" (Bayers, 1998). Cookies were first used by

Netscape to determine if visitors to Netscape's web site were repeat visitors or first time users (Vieux, 1995).

The privacy risks of cookies are considerably heightened when combined with referrer information. The referer [sic] field is part of the HTTP protocol as advocated by Berners-Lee in 1992 (Berners-Lee). It provides a website with the previous URL visited by the person. Its intended purpose was to allow web sites to detect web sites that had linked to them with the hope that they would then add links back to the referring sites (Hallam-Baker, 2003). This would strengthen connections across the web. However, the combination of cookies and referrer information allows web sites to easily track a person's movement through their web sites. Web sites can then acquire considerable information about the long-term habits of their visitors. This ability to monitor and remember a users' movement is a central concern of privacy advocates.

The development process at Netscape was focused heavily on speed. According to Andreessen, the Netscape team:

cranked out the first clients and servers in the first two months or so. We tried to just blow this out the door. . . If you took two years to get it out, the product would be far more technically advanced. But it's more important to get it out there fast so people begin using it and begin to integrate the technology as rapidly as possible. (Allison, 1995)

This pace left cookies as a technological kludge. According to cookies expert Simon St. Laurent, the cookies technology "kind of works, but it's definitely concocted overnight" (Schwartz, 2001). This kludge was a natural consequence of the relentless pace that Netscape was undergoing.

The rapid development cycle and the emphasis on commerce led to cookies being stealthily introduced in Netscape's first web browser in four ways. Netscape turned the feature on by default without notifying or asking the consent of users (Millett, Friedman, & Felten, 2001). Secondly, there was no notification mechanism to alert people when cookies were being placed on their computer. Users did not know that information about them was being saved. Third, the cookies technology was not transparent. Examining a cookies file provides no information about what is stored in the cookie file. Fourth, there was no documentation available that explained what cookies were and their privacy implications (Hedlund, 1995).

Netscape incorporated cookies into its web browsers released in 1994. However, it was not until early 1996 that the public became aware of cookies. The *Financial Times* broke the story on February 12, 1996 with an article on cookies and privacy (Jackson, 1996). This was followed by a story the next day in the United States, "Web 'Cookies' May Be Spying on You" (Gomes, 1996). The article immediately drew attention to cookies and resulted in a great uproar over the use of cookies. Over the next few years, cookies became one of the top Internet privacy issues.

The IETF's Standard for Cookies

The development of cookies by Netscape led the Internet Engineering Task Force (IETF) to develop a standard for state management on the Internet. The IETF, as the de facto Internet standards body, sought to ensure there was a complete technical specification on state management. When the IETF began its work in mid 1995, it was not clear that Netscape's cookies specification would become the basis for the IETF's standard (Kristol, 2001).

The first proposed standard was based on a different technology from cookies that was more sensitive to privacy. The original basis of the IETF's effort was Kristol's State-Info proposal (Kristol, 1995). Kristol's proposal limited the state information to a browser session. In contrast, for Netscape's cookies there is no requirement that cookies be destroyed at the end of the browser session. Netscape's cookies can persist for many years. For example, the google search engine (www.google.com) routinely sets an expiration date in the year 2038 for cookies it creates, thus effectively making the cookies last forever.

However, the IETF eventually switched to the Netscape cookies model. This was largely because the Netscape version was a ubiquitous working model that had become a de facto standard. The IETF's goal shifted to developing a more precise standard for cookies than Netscape's one page draft standard. However, the standards process soon ran into problems. The IETF found that Netscape's implementation of cookies was fraught with privacy and security problems.

The most serious problem was third party cookies. The intent of Netscape's cookies specification was to only allow cookies to be written and read by the web site a person was visiting. For example, if the New York Times placed a cookie on a computer, Amazon.com could not read or modify the New York Times cookie. This provided security and privacy by only allowing sites access to information they authored. However, Netscape's cookies specification allowed third party components of a web page to place their own cookies. This created a loophole by which third parties could read and write cookies. This security and privacy defect was the outgrowth of the rapid

development and incorporation of the cookies technology. This loophole has led to a new breed of businesses, the online advertising management companies (Schwartz, 2001).

Third party cookies can be used by online advertising companies to create detailed records on a person's web browsing habits. Many sites contract out their banner advertising to advertising management companies. These companies find advertisers for web sites and ensure that their banners appear on the web site. For example, DoubleClick sells advertising space on sites such as ESPN and the New York Times. DoubleClick is also responsible for placing the banner advertising on their client's web site. Through the loophole of third party cookies, DoubleClick uses its advertising banners on an ESPN web page to place a cookie when a person visits ESPN. DoubleClick can then read and write to that same cookie when the same person visits the New York Times web site. This allows DoubleClick to aggregate the information about a person's web surfing from its client web sites. They can then create a detailed profile of a person's surfing habits. This has obvious and serious privacy implications.

The IETF's cookies standard is critical of third party cookies and states that third party cookies must not be allowed (Kristol & Montulli, 2000). It does allow an exception if the program wants to give the user different options. However, the baseline default must be set to off. It also requires the user be able to disable cookies, determine when a stateful session is in progress, and control the saving of cookies by web site. This last one is especially significant, because it allows users to manage what sites can and can't place cookies.

The first IETF specification for state management was published in February 1997 (Kristol, 1995). The work had taken almost two years largely because of privacy

problems with third party cookies. Members of the Internet Engineering Steering Group (IESG), which monitors and administers the IETF's activities, felt that third party cookies were a security and privacy issue (Kristol, 2001). They insisted the standard address these issues. However, this standard quickly became inadequate because of compatibility problems in its implementation. This meant a revised standard was necessary.

It took the IETF three years to develop the revised standard. This was again largely because of issues with third party cookies. After the initial standard, RFC 2109, the IETF found a new opposition force. The web advertising networks sought to ensure that consumers could receive third party cookies. However, members of the IETF maintained their support for disabling third party cookies by default. These privacy issues slowed the development of the standard. The IESG insisted on developing strong guidelines for the use of cookies before a new cookies specification would be approved (Moore, 2000). The final standard for cookies was eventually published as RFC 2965 in October 2000 (Kristol & Montulli, 2000).

The long delay in the IETF standard has marginalized the use of this standard, but not its importance. It is unlikely that web sites and web browsers will adopt the standard specified by the IETF anytime soon. Nevertheless, the standard does meet the IETF's goal of developing the best technical standard, even if it will not be adopted in the near term. Moreover, the process of developing the standard heightened public discussion on cookies.

The public discussion on cookies was manifested in the media uproar over the privacy problems, which led to hearings by the Federal Trade Commission (FTC) in June 1996. The hearings only skimmed the surface of the privacy issues and related technical

considerations. In fact, the lack of technical sophistication by the FTC allowed Netscape to pull the wool over their eyes. For example, Netscape stated they would follow the IETF's cookies standard and they would not allow third party cookies. At the FTC's Workshop on Consumer Information Privacy on the Global Information Infrastructure, Peter Harter the Global Policy Counsel for Netscape stated the following regarding third party cookies:

“Mr. Harter: Our position is we are not in favor of allowing third-party domains to pass through. Basically the user couldn't tell if I go to CNN or Outbounders and a cookie is being passed through from the promoter of the ad banner, advertising firms that handle putting up ad banners in multiple sites also want to collect data about who passes over their banners and aggregate that data and report it to advertising for Chrysler or whatever company sees the ad, it is their advertising agency or aggregator. And certainly if they can have a cookie that follows you around and enables you to see a cookie from “cnnnews.com” and a variety of other news sites and sees that you have seen all the different Chrysler ads at different sites during that period of time, they can create some user demographics and surfing behavior data about that particular user. And that's the concern. And that was probably the most controversial issue asked about cookies and this RFC at the Austin meeting.

Mr. Medine: To clarify, Netscape's position is those third parties should not be able to place a cookie?

Mr. Harter: Right. (Federal Trade Commission, 1996)

Netscape never fully followed the IETF standard for cookies and instead their browser by default allowed third party cookies. Kristol's explanation was that the customers of the browser makers were not consumers using free web browsers, but web sites paying for the web server software. These customers wanted to use third party advertising, and the browser makers did not want to alienate their customers. The latest version of web browsers in 2004 by Netscape and Microsoft still accept third party cookies by default to satisfy the advertising management companies.

Nevertheless, the government investigations pushed the browser makers to provide cookie management tools and improved documentation on cookies. In the first versions of Netscape the user could not set cookies preferences. At the FTC hearings in June of 1996 on consumer privacy, Netscape announced that the next version of Netscape (version 3.0) would allow users an option to be alerted whenever a cookie is placed on their computer (Federal Trade Commission, 1996).. At the 1997 FTC Workshop, Netscape announced that its latest browser (version 4.0) would provide the user with the following cookie choices: Accept all cookies; Accept only cookies that get sent back to the originating server; Disable all cookies; Warn me before accepting a cookie (Federal Trade Commission, 1997). It would not be until Netscape 6.0 released in November 2000, that users would be able to fully manage cookies by web site.

References:

- Allison, D. K. (1995). *Interview with Marc Andreessen*, from <http://americanhistory.si.edu/csr/comphist/ma1.html>
- Bayers, C. (1998, May). The Promise of One to One (A Love Story). *Wired*, 6.05.
- Berners-Lee, T. (1991). *HyperText Transfer Protocol Design Issues*. Retrieved Sep. 21, 2001, from <http://www.w3.org/Protocols/DesignIssues.html>

- Berners-Lee, T. (1992). *Basic HTTP as Defined in 1992*. Retrieved Jan. 20, 2003, from <http://www.w3.org/Protocols/HTTP/HTTP2.html>
- Clark, J. (1999). *Netscape Time: The Making of a Billion-Dollar Start-Up That Took On Microsoft*. New York: St. Martin's Press.
- Federal Trade Commission. (1996). *Workshop on Consumer Privacy on the Global Information Infrastructure*.
- Federal Trade Commission. (1997). *Consumer Information Privacy Workshop*.
- Gomes, L. (1996, February 13). Web 'Cookies' May be Spying on You. *San Jose Mercury News*.
- Hallam-Baker, P. (2003). History of Referer.
- Hedlund, M. (1995, 1 Nov). *State Wars, part XI*. Retrieved 1999, 3 Dec, from <http://www.ics.uci.edu/pub/ietf/http/hypermil/1995q4/0161.html>
- Jackson, T. (1996, Feb. 12). This Bug in Your PC is a Smart Cookie. *Financial Times*.
- Kristol, D. M. (1995, August 25). *Proposed HTTP State-Info Mechanism*, from <http://www.kristol.org/cookie/draft-kristol-http-state-info-00.txt>
- Kristol, D. M. (2001). HTTP Cookies: Standards, Privacy, and Politics. *ACM Transactions on Internet Technology*, 1(2), 151-198.
- Kristol, D. M., & Montulli, L. (2000, Oct.). *RFC 2965: HTTP State Management Mechanism*, from <ftp://ftp.isi.edu/in-notes/rfc2965.txt>.
- Millett, L., Friedman, B., & Felten, E. (2001). *Cookies and Web Browser Design: Toward Realizing Informed Consent Online*. Paper presented at the Conference on Human Factors in Computing Systems.
- Montulli, L. (1999). Personal Interview.
- Moore, K. (2000, November 18). *RFC 2964: Use of HTTP State Management*. Retrieved March 15, 2004, from <http://rfc.sunsite.dk/rfc/rfc2964.html>
- Netscape Communications Inc. (1994). *Persistent Client State HTTP Cookies*. Retrieved Sep. 19, 2001, from http://home.netscape.com/newsref/std/cookie_spec.html
- Netscape Communications Inc. (1998). *Persistent Client State in a Hypertext Protocol Based Client-Server System*. USA.
- Reid, R. H. (1997). *Architects of the Web: 1,000 Days that Built the Future of Business*. New York, NY: John Wiley & Sons, Inc.
- Schwartz, J. (2001, Sep. 04). Giving the Web a Memory Costs Its Users Privacy. *New York Times*, p. A1.
- St. Laurent, S. (1998). *Cookies*. New York: McGraw-Hill.
- Vieux, A. S. (1995, Nov. 1). The Once and Future Kings. *Red Herrinig*.