The WPAD name collision vulnerability in new gTLD era: a threat crying for urgent solution

A WhoisXML API Customer Success Story

October 2017.

Sometimes certain comfortable and seemingly innocent protocols can introduce significant security risks, especially when the systems' environment changes. The present document describes such a case.

The WPAD (Web Proxy Autodiscovery) protocol is prevalently used to configure the web proxy settings of end systems such as desktops and other devices belonging to an administrative domain, e.g. a corporate's network. The benefit of this solution is that system administrators can deploy local web proxy settings essentially without any user interaction. Due to an otherwise very progressive change in the domain registration policies, the otherwise very useful WPAD protocol has introduced the possibility of new and a very dangerous man-in-the middle attack. *In 2016 the researchers estimated that at least 6.6 million end-users are at a serious risk [1].*

There is no perfect protection against this threat yet. However, there do exist some efficient remediation strategies, and those which are the easiest to deploy rely highly on domain registration data.

The new gTLD program

In order to facilitate the enhancement of innovation, competition and consumer choice, the Internet Corporation for Assigned Names and Numbers (ICANN) has announced the new gTLD program [2] in 2012, enabling the largest expansion of the domain name system. Beside the few legacy generic TLDs such as .com, more than 1200 new gTLDs have been delegated since.

WPAD-based MitM attack based on name collision

The idea of the WPAD protocol is to obtain a proxy configuration file for the client's browser via a HTTP request. The proxy file location is inferred from this name and fetched using HTTP request to an URL deduced from the client's internal domain. This latter is deduced from a DHCP or DNS query. E.g. if the internal domain is `company.ntld`, the file will be searched for at the URL `http://wpad.company.ntld/wpad.dat`, involving a DNS query for `wpad.company.ntld`. And here the problem comes in: since the advent of the new GTLD era, the company's ntld can be also registered as a new gTLD, thus a *name collision* occurs. Assuming that the adversary can register hosts under the given gTLD (either by delegating the domain for the purpose of the attack or realizing this opportunity later on), the following man-in-the-middle attack becomes feasible:

- From the leaked DNS query of the WPAD protocol the adversary deduces the URL of the proxy configuration file.
- Based on this information, he sets up the appropriate resource at his server.
- The client will download the proxy file prepared by the adversary.
- All the web traffic of the client will now be sent through the adversary's proxy.

A recent detailed study in Ref. [1] is pioneering in the analysis of this threat and quantifying the attack surface. The analysis is partly based on WhoisXML API data. As already mentioned, a dramatically large number of users are at risk because of this vulnerability. And although in 2016 the researchers "did not find strong evidence of adversaries actively registering attack surface domains, but do observe potential blind attack registrations". This lucky situation may, however, change anytime.

Remediation strategies

According to Ref. [1] there is no perfect remediation strategy at the moment. An obvious solution is to treat the problem at the client's side by disabling the WPAD service, upgrading operating systems and revise their settings, or filtering device-level leaks. Owing to the huge number of potentially affected clients, however,

this approach has serious deployment issues. The other extreme level would be at the new gTLD registry, to scrutinize the registration of the union set of highly-vulnerable domains, which would be an efficient approach. It requires, however, a collective effort.

The maybe most viable strategy is thus at the enterprise's level. The filtering of highly vulnerable domains can lead to an efficiency of 97.4%. This is feasible by purchasing accurate WHOIS data, and using [1] as a guideline.

References

[1] Qi Alfred Chen, Eric Osterweil, Matthew Thomas, and Z. Morley Mao. Mitm attack by name collision: Cause analysis and vulnerability assessment in the new gtld era. *2016 IEEE Symposium on Security and Privacy (SP)*, May 2016. http://dx.doi.org/10.1109/SP.2016.46

[2] Internet Corporation for Assigned Names and Numbers (ICANN). The new generic top level domain program. https://newgtlds.icann.org/en, visited on 2017.10.06., 2012.