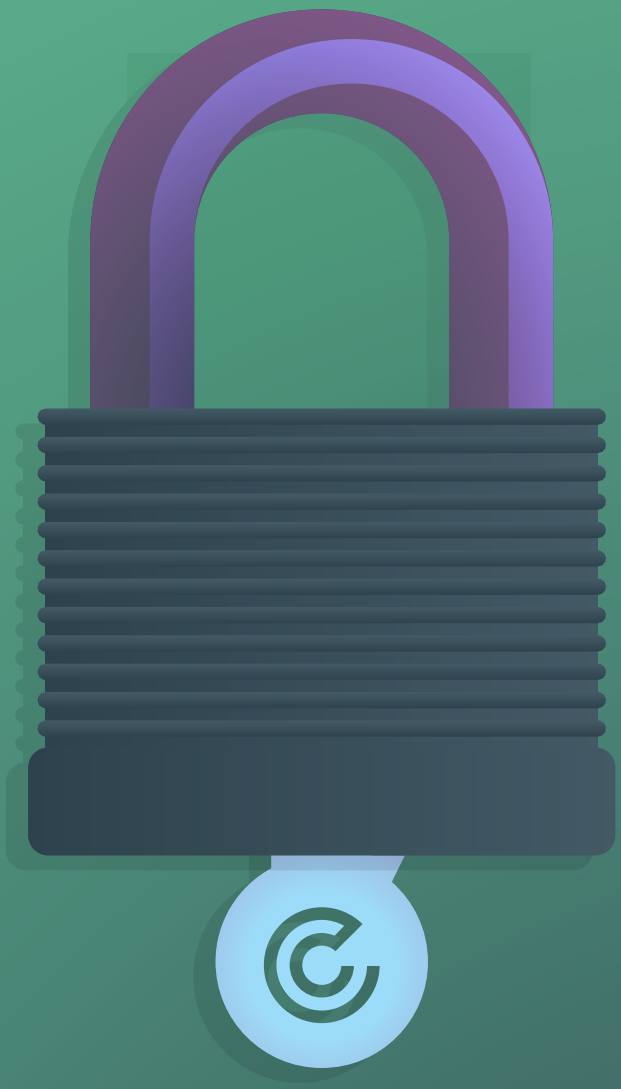
CLICK.CO.UK©

HTTPS WEBSITE MIGRATION



INTRODUCTION

One of the most common questions that we get here at Click Consult surrounds security and the migration of websites from HTTP to HTTPS. With that in mind we thought it best to look at how you can make the switch, the benefits of doing so and the pitfalls of neglecting this approach.

There is a glossary at the end of this ebook so that you can refer to it as you're reading.

WHAT IS THE DIFFERENCE BETWEEN HTTP AND HTTPS?

HTTP (Hypertext Transfer Protocol) is the set of rules for transferring files (text, graphic images, sound, video, and other multimedia files) on the world wide web. As soon as a web user opens their web browser, the user is indirectly making use of HTTP.

The 'S' at the end of HTTPS simply stands for 'Secure' and means that all communications between your browser and the website are encrypted. Using HTTPS, the computers agree on a 'code' between them, and then they scramble the messages using that "code" so that no one in between can read them. The code is used on a Secure Sockets Layer (SSL), sometimes called Transport Layer Security (TLS) to send the information back and forth. This keeps a business' information safe from potential hackers meaning a website that runs on HTTPS can benefit in many ways from both a user and SEO perspective.

Once everyone knew how to exchange information, intercepting on the internet was not difficult. So knowledgeable administrators agreed upon a procedure to protect the information they exchanged. The protection relies on SSL Certificate to encrypt the online data. Encryption means that the sender and recipient agree upon a "code" and translate their documents into random-looking character strings.

The procedure for encrypting information and then exchanging it is called HyperText Transfer Protocol Secure (HTTPS).

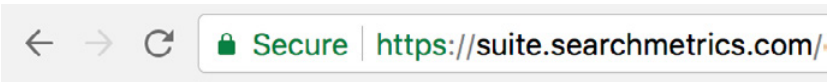
With HTTPS if anyone in between the sender and the recipient could open the message, they still could not understand it. Only the sender and the recipient, who know the "code," can decipher the message.

Humans could encode their own documents, but computers do it faster and more efficiently. To do this, the computer at each end uses a document called an SSL Certificate containing character strings that are the keys to their secret "codes."

SSL Certificates contain the computer owner's 'public key.' The owner then shares the public key with anyone who needs it. Other users need the public key to encrypt messages to the owner. The owner sends those users the SSL Certificate, which contains the public key. The owner does not share the private key with anyone.

It is especially recommended that sites migrate to HTTPS where they have any type of form or checkout process that transmits sensitive information.

You can easily check if a site is HTTPS secure thanks to the search bar and the top of the screen and the secure 'padlock' icon in search, as seen below:



WHY SWITCH TO HTTPS?

There are a lot of important reasons to consider making the switch to HTTPS from a user, security, data and SEO perspective. If you weren't sold by the importance of HTTPS at the start of this eBook then read on… we switch because it's SAFER!

From a user and security perspective:

- HTTPS secures any data transmitted from your browser to the website's servers. This includes your name, address, contact details, passwords and any credit card details that you enter on the website.
- HTTPS shows the padlock icon indicating that the site is secured and this gives users' peace of mind that their data is safe from third party viewing.
- Google has started showing HTTP websites and pages that collect any passwords or credit card details as non-secure. This is part of their plan to eventually show all non HTTPS websites as not secure.
- Data integrity over HTTPS is maintained so data cannot be modified or corrupted without being detected.
- Authentication prevents man-in-the-middle attacks so only you can see the data.

From an SEO perspective:

- Google began using HTTPS as a ranking signal in their search algorithms.
- On 16th December 2015, Google said they were going to start crawling HTTPS equivalents of HTTP pages, even when the former are not linked to from any page. They also began to prioritise HTTPS pages in search results.
- There is also a prerequisite for push notifications in the run up to switching to HTTP/2 which relates to site/page speed.

From a data perspective:

Making the switch to HTTPS helps with the loss of referral data that happens when switching from a secure website to an unsecured website. This happens during the switch as the referral header is dropped. This causes analytics programs to attribute traffic without the referral value as direct instead, which accounts for a large portion of what is called dark (or non-traceable) traffic.

SETUP

There are a significant number of checks and changes that need to be made in order to switch to HTTPS before businesses put it visible on the web.

Before you begin, it is essential that businesses have a copy of their website available for editing on a test server. Because the changes being made will be relatively extensive, and not all can be done immediately, you may end up with a partially broken website if you attempt to do this on a live website. Another important consideration is having your site crawled. This is important as a starting point because it allows businesses to gather vital information including their current URL's and internal links.

If businesses have a CDN (content delivery network) they will need to understand how that system works when it comes to using HTTPS and be able to make changes where necessary. All future work will require planning in terms of HTTPS and need specialists to implement the changes.

A lot of websites use absolute URL's in the content (direct reference to the full URL, not just the folder and file). This means there will be a lot of hard coded HTTP links and images within your content that will need changing to HTTPS or a relative URL.

As well as the content it is vital to think about the templates too. Businesses will need to be able to update all of their website templates that use absolute URL's to either HTTPS or to relative URL's and this also includes scripts and images.

Some external scripts that are included may not support HTTPS. If they don't then the owner will get an insecure content warning when the page is loaded and it will be classified as not secure.

Check that all external scripts are running in HTTPS. If they are not, change them.

Purchase and install a security certificate on the server

The next thing to do is for businesses to buy an SSL certificate. These are widely available from most domain registrars and web hosts. Some even provide them for free as an incentive to migrate.

As there are a lot of certificate types available it will depend on the website, the business requirements and the budget. Companies will need to choose the correct one for both the server and their requirements.

Once it is installed on the server, and the configuration changed for the website, then the server should be ready for the HTTPS switch over.

Update canonical tags

Canonicalisation is the name for the process of redirecting search engines from one or more of your brand's URLs to a single, accepted (or canonical) URL. These multiple pages are common, but can potentially cause problems with ranking – as the authority conferred upon content may be divided between multiple addresses, weakening the authority of each version and, therefore,decreasing the overall visibility of the page or meaning the wrong page may rank.

Common URL duplications

http://www.your_site.com http://your_site.com
http://www.your_site.co.uk http://www.your_site.com/index.html

Best practice

When optimising a site, the owner should always use the canonical tag (with few exceptions – such as for brands whose pages are repeated in multiple languages), ensuring all pages achieve the maximum Page Authority possible. All duplicate versions of a page should contain the same URL within the canonical tag – which should always be complete and identical, including the HTTP(S)://, or it may be ignored.

Some CMS's may need settings in the backend changed to support the migration to HTTPS. For example in WordPress a business would need to change the URL field so that the CMS knows what its canonical URL is. Failing this will land the user in a redirect loop and the site will fail to load.

Potential problems and things to remember

If a businesses website makes use of any 3rd party extensions then they will need to ensure that these work when they update to HTTPS and don't reference HTTP in any way. Any references to HTTP will cause the browser to display an error that the page contains insecure content. The page is then classed as not secure.

Update old redirects, sitemaps and robots.txt

Most websites will have some redirects already in place. These will need checking as chances are they will now be out of date. Take this opportunity to test all old redirects to make sure they still lead somewhere that is relevant and not a 404 page.

It is also crucial to make sure the html and xml sitemaps both have the URL's within them updated HTTP to HTTPS. This applies to the xml sitemap especially as this is the one that Google will primarily try to use for its crawling once the site goes live.

The robots.txt will contain an xml sitemap reference which will now likely be wrong as it will point to the HTTP version. This need changing to HTTPS to ensure that Google will read it.

Add the HTTPS version of the site to Google Search Console (GSC)

This is an important step because without it, Google will have to re-crawl the website from scratch a page at a time. With the xml sitemap in place, it can initially find and index the pages much more quickly.

Upload the disavow file

If a business had a disavow file on their HTTP version of Google Search Console, then they should download it and upload it to the HTTPS version. Ensuring that any domains that Google was previously ignoring are also ignored after the switch.

For these reasons it is essential that companies are constantly crawling their sites again and again. This is a foolproof way to make sure that they haven't missed any links that are still using HTTP, and that no other links are broken. Any insecure content can then be worked through.

Optional extras

Enable HSTS

HSTS (HTTP strict transport security) is a HTTP header which improves security by forcing the browser to always use HTTPS, and eliminates the need to do a server side check so your website loads faster. The implementation of HSTS is different depending on your server so bear that in mind.

Enable OCSP stapling

This enables a server to check if a security certificate is revoked instead of a browser, which keeps the browser from having to download or cross-reference with the issuing certificate authority.

Add HTTP/2 support

HTTP /2 is the next major revision in the HTTP protocol that is used by all browsers. HTTP /2 is significantly faster and optimised, but is only supported over HTTPS.

Go live!

You're ready to go live.

After going live, run some final tests to ensure nothing was missed in the final switch over. You should look at the following:

Analytics

In Google Analytics, and other analytics platforms that you use, make sure you update the default URL if one is required to ensure you are now tracking HTTPS properly.

Make sure you look at your analytics every day for a few days to make sure data is properly tracking.

Social share counts

When you change to HTTPS, your social URL's will change. This means that any social counts will disappear and some networks don't transfer it automatically.

Update other tools

If you are doing any A/B testing, heat maps or keyword tracking you may need to update to HTTPS URLs in those platforms as well.

Monitor everything during your migration

Check, double check, and triple check everything you are doing as you do it to make sure it's done properly.

GLOSSARY

For those who are continually reviewing their SEO and who are looking for the latest developments some of the terms in this piece will be second nature. We appreciate however that there is some depth in terms of technical language. With that in mind we've compiled this handy glossary.

Secure Sockets Layer (SSL) –

Secure Sockets Layer (SSL) was the most widely deployed cryptographic protocol to provide security over internet communications before it was succeeded by TLS (Transport Layer Security) in 1999. Despite the deprecation of the SSL protocol and the adoption of TLS in its place, most people still refer to this type of technology as 'SSL'.

SSL provides a secure channel between two machines or devices operating over the internet or an internal network. One common example is when SSL is used to secure communication between a web browser and a web server. This turns a website's address from HTTP to HTTPS, the 'S' standing for 'secure'.

Transport Layer Security (TLS) –

Transport Layer Security (TLS) – and its predecessor, Secure Sockets Layer (SSL), which is now prohibited from use – are cryptographic protocols that provide communications security over a computer network. Several versions of the protocols find widespread use in applications such as web browsing, email, internet faxing, instant messaging, and voice over IP (VoIP). Websites are able to use TLS to secure all communications between their servers and web browsers.

The TLS protocol aims primarily to provide privacy and data integrity between two communicating computer applications. When secured by TLS, connections between a client (e.g. a web browser) and a server (e.g. wikipedia.org) have one or more of the following properties:

The connection is private (or secure) because symmetric cryptography is used to encrypt the data transmitted. The keys for this symmetric encryption are generated uniquely for each connection and are based on a shared secret negotiated at the start of the session. The server and client negotiate the details of which encryption algorithm and cryptographic keys to use before the first byte of data is transmitted. The negotiation of a shared secret is both secure (the negotiated secret is unavailable to eavesdroppers and cannot be obtained, even by an attacker who places themselves in the middle of the connection) and reliable (no attacker can modify the communications during the negotiation without being detected).

The identity of the communicating parties can be authenticated using public-key cryptography. This authentication can be made optional, but is generally required for at least one of the parties (typically the server).

The connection ensures integrity because each message transmitted includes a message integrity check using a message authentication code to prevent undetected loss or alteration of the data during transmission.

Content Delivery Network (CDN) –

A content delivery network or content distribution network (CDN) is a geographically distributed network of proxy servers and their data centres. The goal is to distribute service spatially relative to end-users to provide high availability and high performance. CDNs serve a large portion of the Internet content today, including web objects (text, graphics and scripts), downloadable objects (media files, software, documents), applications (eCommerce, portals), live streaming media, on-demand streaming media, and social networks.

.htaccess –

A .htaccess (hypertext access) file is a directory-level configuration file supported by several web servers, used for configuration of site-access issues, such as URL redirection, URL shortening, Access-security control (for different webpages and files), and more.

A site could have more than one .htaccess file, and the files are placed inside the web tree (i.e. inside directories and their sub-directories), and hence their other name distributed configuration files.

.htaccess files act as a subset of the server's global configuration file (like httpd.conf) for the directory that they are in, or all sub-directories.

HTTP Strict Transport Security (HSTS) –

HTTP Strict Transport Security (HSTS) is a web security policy mechanism which helps to protect websites against protocol downgrade attacks and cookie hijacking. It allows web servers to declare that web browsers (or other complying user agents) should only interact with it using secure HTTPS connections, and never via the insecure HTTP protocol.

The HSTS Policy is communicated by the server to the user agent via an HTTPS response header field named Strict-Transport-Security. HSTS Policy specifies a period of time during which the user agent should only access the server in a secure fashion.

OSCP Stapling –

OCSP stapling, formally known as the TLS Certificate Status Request extension, is an alternative approach to the Online Certificate Status Protocol (OCSP) for checking the revocation status of X.509 digital certificates. It allows the presenter of a certificate to bear the resource cost involved in providing OCSP responses by appending (stapling) a time-stamped OCSP response.

Here at Click Consult we understand the technical side of the job and that the future of web based businesses rely on security. With that in mind we are constantly updating the way we do things and this eBook serves that purpose. For more information why not check out our other resources?

CLICKCONSULT®

EXPERTS IN SEARCH. SIMPLE.

ABOUT US

Located in North West England, Click Consult is a multi award-winning search marketing agency with a focus on organic (SEO) and paid search (PPC), with over 70 professionals employed and with a portfolio of over 60 clients from across the UK, Europe, Americas and Australia.

Click was named Best Digital Agency and Best Large eCommerce Agency in 2017, adding to its long list of other awards and accolades, and also ranks within both RAR's and Econsultancy's 'Top 100 Digital Agencies', and Prolific North's 'Top 50 Digital Agencies'.

ADDITIONAL RESOURCES

Blog

Click Consult regularly posts actionable insights on its blog – dealing with all aspects of search marketing. From technical SEO to PPC, content marketing and Analytics advice, the Click Consult blog has something for everyone of any ability.

eBooks

Click Consult produces in depth eBooks on all aspects of search marketing on a regular basis, dealing with the latest best practices, changes to standard practices, methods of improvement and more.

Infographics

Sometimes what you need is a quick point of reference about complex subjects. Thanks to its team of highly talented designers, Click Consult is also able to distil its vast experience into easy to understand visual content.

Follow us on Google+, Facebook, Twitter or LinkedIn, or to find out what Click Consult can do for you, call:

0845 205 0292

OUR ACCREDITATIONS

RAR RECOMMENDED

Bing Select Partner

PREMIER Google Partner

Specialized in: Search ads, Mobile ads, Video ads, Display ads, Shopping ads

CERTIFIED PARTNER

OUR AWARDS

NORTHERN eCOMMERCE AWARDS 2017 WINNER

THE DRUM SEARCH AWARDS AWARD WINNER 2017

UK AGENCY AWARDS 2017 WINNER

OUR TECHNOLOGIES

monitor TRAX rank TRAX feed TRAX link TRAX pro TRAX page TRAX