CLICK.CO.UK©

SEO OPTIMISATION

SCALABILITY FOR LARGE WEBSITES

CONTENTS

INTRODUCTION

WHAT IS SCALABILITY?

In broad terms, scalability refers to the ability of a system, model or function to withstand increasing workloads and its potential to be enlarged to accommodate growth. Insofar as it applies to search marketing, this covers a site's capability to function under unexpected or large-scale increases in traffic, as well as the implementation of SEO best practices on sites with large numbers of individual pages.

While scalability for large websites deals with many of the same issues as SEO for small sites, there are problems inherent in large-scale implementation that can cause unforeseen problems – these include site structure, linking strategies, content, sitemapping, canonicalisation and more.

As a result, while this eBook will assume some low-level familiarity with search marketing practices and expand upon them for use specifically with large sites.

WHY WORRY?

All sectors and all industries are operating online, if not entirely then increasingly. It is every brand's aim to become the top of their specific niche, and this means that all brands are essentially looking to increase traffic, increase conversions, increase sessions and session length, improve ranking positions and ranked keywords. All of this will place strain on hardware, software and SEO practices as sites increase in size and scope.

In order to ensure that there will not come a time when a site will face a tipping point where they are required to abandon a site, or lose traffic and revenue, it is necessary to have in place a specific scalability strategy – regularly reviewed to ensure that it grows with the brand. This can prevent revenue loss, and can improve ranking capabilities and even site security.

So, while dealing with scalability may not seem like an immediate concern for a well-prepared SME, there should always be a strategy in place to ensure that there never comes a point when the need for scalability comes as a surprise.

INFRASTRUCTURE

We won't go into a tremendous amount of detail here, as it is outside of the sphere in which people look to us for advice, but scalability of infrastructure should nevertheless be where any strategy begins.

By this, we don't mean that there should be a warehouse full of servers gathering dust while you reach a level at which they are necessary, nor that a brand should purchase unnecessary amounts of cloud server hosting – infrastructure scalability is about readiness for growth.

This means choosing hardware that can be easily augmented with additional capacity, or choosing a cloud host that has the potential to cover your predicted growth (and also possible growth).

Think about the times you've been frustrated trying to buy tickets for a popular event by slow or broken ticketing websites, or during seasonal sales on eCommerce sites which crash, causing complete basket (and therefore time) loss. Then think how that impacts possible return custom for your own brand.

For that reason, do not look for the cheapest option, or for infrastructure that will only suit short-term requirements. Success can come at any time – especially if a brand is operating a structured and creative search marketing strategy, therefore ensure that your hosting is capable of growing alongside your brand.

YOUR HOMEPAGE

As it seems logical, and we're reliably informed it's a good place, we're going to start at the beginning – or the beginning as it is for many of your consumers. Your homepage will change over time and as your brand grows, and as such it has many possible complications as traffic and demographics begin to grow and change.

PRIME REAL-ESTATE

For example, your homepage is likely to be (or should be) the most seen page on your site – meaning that the content on this page is also going to be the most seen. For many brands, the home page is a functional affair – with a short journey through a few major sections of the site, some simply with a call to action to begin a free trial or download software. Yet for large sites – such as eCommerce sites, news sites or any with a large number of unique pages and categories, the home page needs to accomplish an awful lot.

While there are easy decisions that can be made – site-wide best sellers or most downloaded can clearly feature in this place of prominence, as can top categories, but will the same homepage be visible to all consumers? As a brand's offering grow, generally speaking the site will grow along with it – which means that the specific interests of various consumers will begin to be more diffuse, meaning that generic front pages can represent a missed opportunity.

For this reason, it can benefit large sites to have areas of the homepage which are responsive to consumer habits. While there may be content which is core to the brand's positioning, or which represents its best chance to rank in search engine results pages (SERPs) for industry key terms, however, the chance to offer personalisation – not to mention the expectation of it – increases with the size of the site.

DYNAMIC HOMEPAGES

There was a good dissection of Optimizely's decision to move to a personalised home page over on their blog which deals with many of the reasons for their choice. While their 'average best' homepage had been optimised to the extent where it had a fantastic conversion rate, the conversions were sometimes of poor quality as leads for the sales team.

The same is true at larger scale. While simple, one-size-fits-all homepages may work well for small sites – with limited pages to navigate to – larger sites with far greater selections can become too busy if a brand attempts to put too much on there, or impossible to navigate if there is too little.

For this reason, it is often better for the brand to use data gathered (whether directly – through forms etc. or indirectly via cookies) to tailor content to specific demographics and consumer personas. This tailoring allows brands to feature demographic specific options, products or services – have navigation menus which enable consumers to better reach areas of the site which are relevant to them and therefore decrease the length of the journey from homepage to basket.

It is, of course, not necessary for every area of the page to be personalised – simply areas which will facilitate conversion. This means headers, navigation bars, and featured content. Much of the homepage may remain the same, but the capacity to personalise important areas can overcome many of the problems with them that a large number of pages and categories can cause – including scalability issues, as the available content for the personalised sections can be updated as the site grows.

The above is the Amazon homepage for a non-signed-in consumer – while the scrolling offers header remains the same for all consumers, the products below change dependent on the user's browsing and purchasing habits. The static sections are something to consider as a homepage that changes too often can fail to rank for key terms – therefore, even dynamic, personalised homepages should retain fixed sections of content featuring terms you wish the homepage to rank for.

STATIC HOMEPAGES

There are, of course, ways and means to ensure your homepage is usable without personalisation. The primary concern is, of course, delivering the right amount of information and content so that the page is useful to all demographics and so consumers can reach that any section of the site in as few steps as possible.

This is, in most cases, best achieved through well-ordered site structure and use of primary categories. As below, the homepage offers the most common primary categories clearly in the centre of the page, while the fixed menu-bar is still present at the top with easy access to more primary and sub-categories for anyone with a definite direction in mind for their browsing.

This also means that, provided your categories and sub-categories are well maintained, that you can maintain a steady and familiar presence even as your site (and therefore menu) grows and expands.

NAVIGATION

As stated above, navigation from your homepage is one of the most important issues to consider when considering scalability for your site. While we've covered the reasons why, we should probably give some information on how.
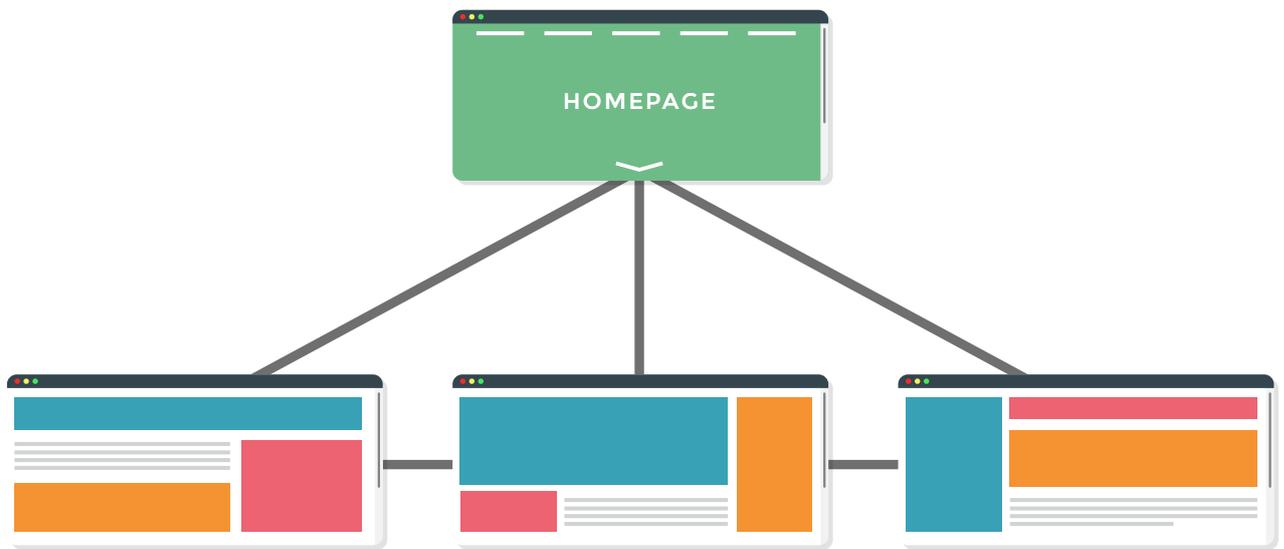
Whichever method you choose for the homepage of a large scale website, there are a few navigation issues you will definitely need to consider:

- **Consumer personas –** you will need to know what kinds of consumer use your site, their demographic, interests and traditional methods of site use. While this is something most brands will have when they implement any marketing strategy, there is a Click Consult eBook to help those who don't. These personas will inform areas of the site, products or services, navigation to which should take pride of place on your homepage – whether as personalised offers or standard category links.

- **Best sellers and offers –** While many of the offers will be driven by the consumer personas, there will obviously be some things which are more important for the brand's needs than others. Priority lists should be drawn up for products, goods or services which are important to the brand's goals so that a rotation of business driven navigational links and images can be added to those driven by the consumer.

- **Menu-bar –** Again, whichever way you decide to go, we would recommend a clearly formatted menu-bar linking to all of your main category pages (at least) and most popular sub-categories (at best). Even the best personalised homepage will occasionally fail to feature what the consumer is looking for and a menu-bar will should help them to find it easily.

YOUR CATEGORY PAGES

WHAT IS A CATEGORY?

By category, we're referring here to an umbrella term for a group of products or services – this could be as broad as 'men's fashion' and 'women's fashion' or as narrow as a specific electronics manufacturer. However broad you decide your categories should be, they should always be sufficiently descriptive to allow consumers within your targeted demographic to understand what to expect within them.

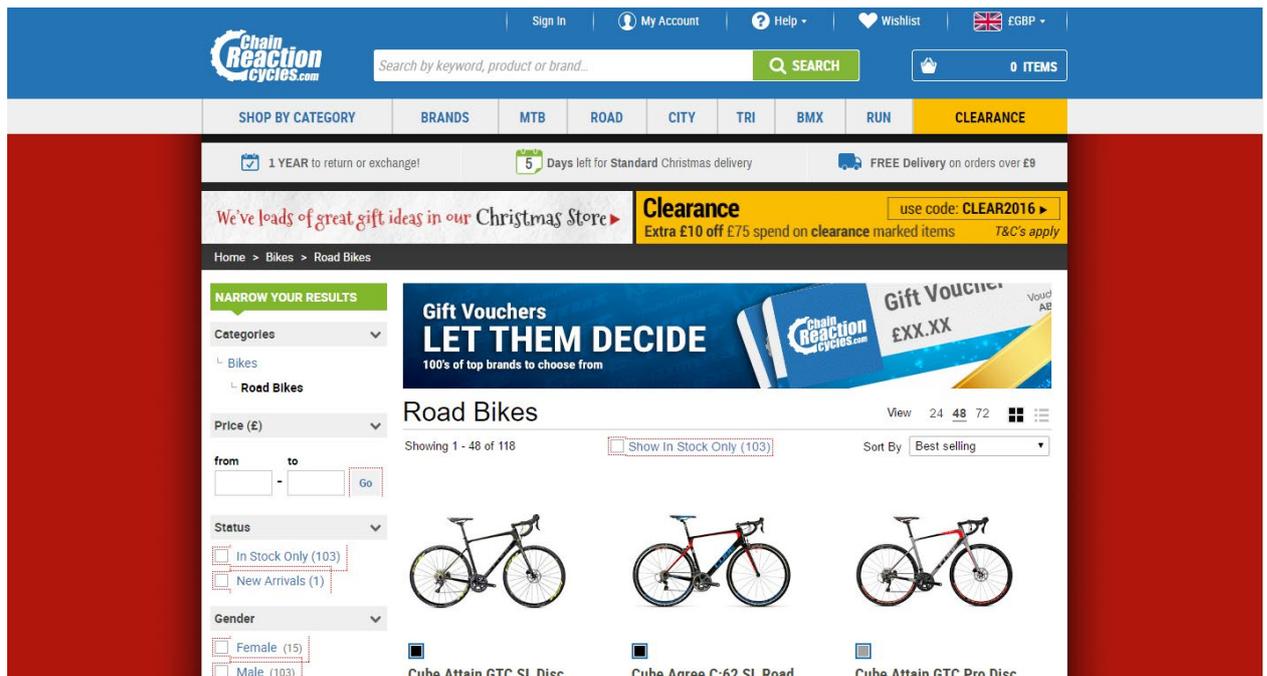Your category pages are, as above, nested in your sitemap directly beneath your homepage – they will follow the first forward slash after your site's URL. It is necessary for them to therefore be concise and descriptive – too long and they're difficult to remember, to non-descript and they can confuse consumers.

It is better for navigation purposes if these categories are kept to the minimum number necessary of distinct categories.

INTERNAL LINKING

As stated previously, your home page should always link to all of your primary categories. However, each category should also do this. This is generally achieved on the majority of sites by maintaining a fixed menu-bar on all sub-pages of the site. Again, with this as a fixture on all pages, there can be no sub-pages from which a consumer cannot navigate to any of the other main categories.

Following this, you can then choose one of two options – you can continue to sub-categorise pages, or you can add filters to the category. While, generally speaking, linking to sub-categories can be easier, for large websites a filter is going to be both the easiest to maintain and to navigate – it also has the benefit of being demarcated by a variable in the URL rather than a further forward slash and category name which is useful for ensuring there are no gaps or breaks in the flow of authority from inbound links.

HOW MANY IS TOO MANY SUB-CATEGORIES?

This is a difficult question to give a definitive answer to – there is no right or wrong way to approach the structure of a site provided there is clear, simple navigation and are no barrier to the flow of authority. However, for the most part it is better to consider the 'less is more' advice as best practice (there is more on ideal URL structure in Click Consult's eBook on Technical Optimisation). In short, however, ideally you want your product pages to be as close to the root domain as possible.

While this can be difficult to implement, and it may be tempting to – for example – sub-categorise science fiction below books and then authors below both, in reality, it is possible to limit sub-categories through use of variables:

URL BEST PRACTICE

Your URL has always been vital to your brand and, though domain name keywords are no longer as helpful as they were in determining SERPs rankings, an exact match domain can influence click-through rate. However, the important thing is to be memorable, easy to spell and short. This is the same, wherever possible, for all URL elements.

http://store.your_site.com/Item_Folder/Item_Page

Protocol://Subdomain.Root Domain/Subfolder/Page

When writing URLs, try to keep them simple and use descriptive keywords – if a subfolder contains blouses, call it "blouses". Also, wherever possible, attempt to keep content on the same subdomain – this will help with domain authority.

A good URL will always be as simple as possible. Though it may be tempting to subfolder each product for reasons of accuracy, the ideal is to have your product as close to your domain name as possible and to have it clearly described – not only is this easy to remember, but if someone chooses to type out a link, it is less likely they will misspell the URL and therefore deprive you of the link and link traffic.

PRODUCT PAGES

WHAT IS A PRODUCT PAGE?

The term 'product' essentially means any item, service or anything else on your site which is the endpoint of a consumer's navigation or search through your site. These are some of the most difficult pages to build – not, generally for technical reasons but for the amount of thought that needs to go into a page that is so close to conversion.

MULTIPLE PRODUCTS COLOURS/SIZES AND DUPLICATE CONTENT

According to Google's John Mueller, there are no direct penalties for duplicate content. However, clear indications that you have issues with duplicate content are if non-preferred URLs are appearing in searches, Google search console is showing multiple instances of title and description duplication, or if site crawlers are returning a greater number of page-crawls than you have pages. In addition to this, there are a number of free web resources which can trawl your site for duplicate content.

How to fix them
Do not use robot.txt to resolve duplication issues – this simply disguises the duplication, hiding it from the site crawlers without resolving the problem, as both pages will be indexed and the resultant problems stemming from duplication will persist.

Google understands that some duplication is unavoidable (such as with press releases, or third party product descriptions), and the key is to be consistent with your signals to Google and other search engines – ensure that you are using the desired URL everywhere (sitemap, canonical, href etc.), "rel canonical" is a good method of overcoming problems caused by multiple product pages (for size/colour etc.).

Avoid URL variations in your Content Management System, ensure you're making appropriate use of 301 redirects wherever possible and that your site is easy to read with a well-structured hierarchy. We would recommend extensive auditing for duplicate content both internally and external to your brand's site.

CANONICALISATION

WHAT IS CANONICALISATION?

Canonicalisation is the process through which it is possible to redirect search engines from multiple near identical URLs to a single, accepted (or canonical) URL. These multiple pages are common on most sites but, if left unaddressed, can cause problems with a site's position in SERPs, as the authority conferred upon any content on these pages may be divided, weakening the overall visibility of the page.

The canonical tag, which is used in the process is a meta tag (placed within the <header> area of a site. The tag, however, is not a redirect, so the users and search engines still arrive at the URL they expect, it simply means that the URL named in the tag is the one used for indexation with search engines.

Important!
When creating canonical tags on a website, ensure you make no errors in reproducing the canonical URL. A simple spelling mistake or slip of the typing finger can prevent the page from indexing entirely.

WHY USE CANONICALISATION?

Canonicalisation (along with 301 redirects, which can be read about here) can prevent dilution of page authority by promoting a specific variant of several identical or nearly identical URLs. By choosing a specific URL to feature as the homepage (or service page, or any other) of choice, it is possible to ensure that all authority earned from inbound links to the page is properly distributed throughout the site and that there are no duplicate or nearly duplicated pages ranking in SERPs.

HOW TO IMPLEMENT CANONICALISATION

When optimising, it is always a good idea to use the canonical tag, ensuring all pages achieve the maximum authority possible. All multiple versions of a page should contain the same URL within the canonical tag which should always be complete, including the HTTP://, or it may be ignored.

To set a canonical tag, use:

<link href=http://www.your_site.com/rel="canonical" />

This is a stand-alone tag, which should be placed between the <head></head> tags, and should refer to your preferred URL. This serves the double purpose of allowing you to control which of these a search engine sees, while ensuring your traffic is not split between duplicate content.

EXTERNAL LINKING STRATEGY

WHAT IS AN EXTERNAL LINKING STRATEGY?

An external linking strategy is simply the process of deciding which pages on your site to prioritise during any link building exercises or content marketing campaigns. For all sites, whether big or small, there will be pages that hold more value for a brand and which should be more heavily promoted in order to boost their visibility in SERPs.

The main difference is that, while there may be high earning, high priority products or services pages, there will likely to be too many to adequately targeted. Therefore, the strategy for large brands involves trying to target pages at least one up in the XML hierarchy so that authority can be spread across various pages and areas of a site.

WHICH PAGES SHOULD YOU TARGET AND WHY?

The target identification for external linking is an SEO practice that needs to be amended for large scale sites. While it is perfectly reasonable for a small site to seek to build links to each and every piece of content, looking to target improvement in SERPs for keywords across the site, there are obviously problems inherent in trying to accrue links for 10,000 individual products.

However, as authority flows from the top down, it is possible to increase the authority of pages across the site by identifying key pages for link building exercises. Therefore, while large brands should still begin by researching the best-selling products or most important pages on a site, the aim is then to use these products to earn links for important areas of your site. This can be done by tracking back through the previous strata of the sitemap and making sure that you choose a target (generally category) level for links so that each popular product covers as much of the site as necessary.

Once this is done, you can maybe look to make such popular products 'features' on these category pages to ensure people arriving at the page see what they are expecting to, while ensuring that authority is spread across the site.

XML SITEMAPPING FOR LARGE WEBSITES

Sitemaps are essentially a list of URLs and while they allow brands to have an organisational overview of the site, to search engines they're a lot more – they assist in the crawling of your site, facilitate its indexation and permit the proper flow of authority throughout the site.

The above is a visualisation of the top level structure of Wikipedia – it's possible to see how convoluted it's possible to get with a large website. As such, large sites should look to submit multiple sitemaps.

Because of the importance of sitemaps to SEO performance, it is imperative that all files are of exceptional quality and accuracy – without redirects, errors, or duplicates.

Due to number of pages on large sites, however, we recommend that sitemaps be segmented taxonomically – by categories, page types, services or similar – with product/service/item level pages featuring in their own sitemaps allowing for the monitoring of the performance of low level URLs comparative to the rest of the site.

This level of detailed organisation also permits the scalability of sites that are heading towards an increasingly large number of products. If a brand can have such a facility in place, there can be pages added, and pages removed with far less trouble and far more accuracy than if a single, more diffuse sitemap is in place which then has to be adapted as the brand grows.

CONCLUSION

Scalability is often reliant on best practice. While there are some areas of standard SEO practices that must be amended, the vast majority of issues with scalability are caused by insufficient preparation in the early stages of business planning.

In each area of a brand's digital presence, there should be room for progression and growth. Whether this is in hardware, software or simply in the practices in place for SEO strategy – brands must prepare for scalability at the earliest possible stage.

Allowing room for growth and room to manoeuvre can prevent issues further down the line.

Anyone that begins a business must believe in the vision that drives it and in the possibilities for profitability and success, yet too many brands forget to plan for (or don't know that they should plan for) the growth of their digital presence and the necessary changes that this can require.

As such, scalability shouldn't only be a consideration for brands which are already large, but for SMEs that might one day reach such a size.

GET IN TOUCH WITH US TODAY

ABOUT US

We are a multi award-winning search marketing agency working with brands & SMEs that want a tangible, measurable return from their online budget.

Our website features an online resources section, full of useful eBooks. We also write a blog, where you can find an abundance of search marketing tips. If you're feeling social, get in touch on Google+, Facebook, Twitter or LinkedIn.

To discuss how a bespoke strategy could amplify your business' search engine visibility, call us today on

0845 205 0292

Or visit our website.

WHAT OUR CLIENTS SAY

"Click, from the initial pitch, have delivered on their promise of providing data-driven solutions for particular SEM problems. I have found them both knowledgeable and effective in delivering results."
Fergal Lynch, Chill Insurance

"As business decisions go, this was probably one of our best. It is not often you find such a responsive and intuitive agency who understand what you are trying to achieve. They get it. They take feedback on the chin - we can be very direct sometimes, and they simply run with it and find solutions that work for us. In a market that over promises and under delivers we trust Click Consult to do a brilliant job. I highly recommend Click Consult. "
Eilish Adams, The Law House

OUR ACCREDITATIONS

RAR+ RECOMMENDED   Google Partner   Accredited Professional | bing ads   VWO CERTIFIED PARTNER

OUR TECHNOLOGIES

monitor **TRAX**   rank **TRAX**   feed **TRAX**   link **TRAX**   pro **TRAX**

CLICKCONSULT®

EXPERTS IN SEARCH. SIMPLE.