

22.2: How the Internet works

In its simplest form, the Internet is a collection of documents connected by hyperlinks.

A hyperlink is a virtual link from one document on the World Wide Web to another. It includes the Uniform Resource Locator (URL) of the linked-to document, which describes where on the Internet a document is. It is what you enter in the address bar of the browser, because it is the address of that document on the Internet.

A URL provides information to both browsers and people. URLs include domain names that translate to Internet Protocol (IP) addresses. Every website corresponds to an IP address, which is a structured series of dots and numbers indicating where it is physically located. In fact, every device on the network has an IP address.

When you enter a URL into the address bar of a browser, the Domain Name System (DNS) record indicates where the document you are linking to is.

Confused? Look at the domain name and IP address for Red & Yellow's website:

Domain name: www.redandyellow.co.za

IP address: 212.100.243.204

A domain name looks something like this: www.domainname.com.

But a lot more information can be included in this. URLs can carry the following information: subdomain.domain.tld/directory

Domain – the registered domain name of the website.

Subdomain – a domain that is part of a larger domain.

TLD – the top level domain, uppermost in the hierarchy of domain names.

Directory – a folder to organise content.

The TLD can indicate the country in which a domain is registered, and can also give information about the nature of the domain.

.com – the most common TLD.

.co.za, .co.uk, .com.au – these TLDs give country information.

.org – used by non-profit organisations.

.gov – used by governments.

.ac – used by academic institutions.

Domain names must be registered, and there is a fee for doing so.

A website, or any content on the Internet, is hosted on a server. A web server is a machine that serves web content, and the term often refers to the software (applications) and the hardware (machine) that serve the content.

Very simplistically, it works a little something like this:

1. Someone enters a URL in a browser.
2. This is translated into an IP address, which indicates where the content is located, or where the server for the content is.

Note

All of this happens in a fraction of a second!

3. The server then returns the content requested.
4. The person sees the website that they requested.

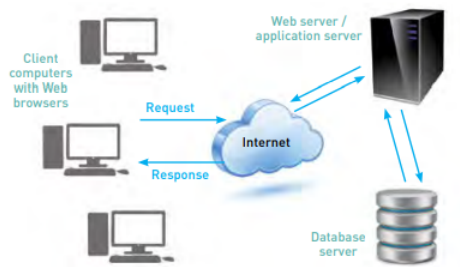


Figure 22.2.1: The process of serving a website *Adapted From Stokes, 2013*

Sometimes, the server is not able to fulfil the request meaning it cannot return the content requested, and instead it returns a status code. Two common status codes you will encounter in this book are below.

301: This is used to indicate that the content requested has moved permanently, and the new version of the content is returned instead. These 301 redirects are often used in search engine optimisation (SEO) or when a new website is launched to make sure that old links are redirected to the correct, new content.

404: This is returned when the content has not been found on the server, either because there was an error in the link, or because the content has been moved or deleted. Website owners can design a custom page for when a 404 error occurs, giving users useful information.



Figure 22.2.2: A fun custom 404 page from Lego *Adapted From Creative blog, 2017*

You can find a full list of status codes at www.w3.org/Protocols/rfc2616/rfc2616-sec10.html.

This information can be sent via Hypertext Transfer Protocol (HTTP), or HTTPS, which is a combination of HTTP with a secure way of transmitting information.

Note

How aware are you of security when browsing the web? Pay close attention to the sites that use secure protocols, what does this say about them?

HTTP makes it easy to request and transfer information. It's what makes our websites load, and allows us to connect with people on social networks. However, the information that is transferred is not transferred securely, meaning that it could be viewed by third parties. If this was the only way of sending information online, it would be a bad idea to bank online, or to purchase anything over the Internet.

This is why we use HTTPS to encrypt information when it is sensitive. In order to make use of HTTPS, the relevant website needs to get a security certificate, which ensures that various details have been verified by a trusted third party.

If you're unsure, look in the browser address bar to check whether the site you are on is HTTP or HTTPS. Most browsers will indicate a secure site with a little padlock in the address bar, or somewhere else in the browser, to make sure that you know you are in a secure site.



Figure 22.2.3: Indicators of a secure site *Adapted From Screenshot, Google, 2017*

This page titled [22.2: How the Internet works](#) is shared under a [CC BY-NC-SA 3.0](#) license and was authored, remixed, and/or curated by [Rob Stokes](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.