Case Study: Google

Submitted To :

Meskat Jahan

Lecturer,

Department of CSE

Comilla University

Submitted By:

Farah Nawar

Id : 11608001

Student ,

Department of CSE

Comilla University

Date of Submission : 19 January,2021

1. Introduction

The web creates new challenges for information retrieval. The amount of information on the web is growing rapidly, as well as the number of new users inexperienced in the art of web research. People are likely to surf the web using its link graph, often starting with high quality human maintained indices such as Yahoo! or with search engines. Human maintained lists cover popular topics effectively but are subjective, expensive to build and maintain, slow to improve, and cannot cover all esoteric topics. Automated search engines that rely on keyword matching usually return too many low quality matches. To make matters worse, some advertisers attempt to gain people's attention by taking measures meant to mislead automated search engines.

Google is a large-scale search engine which addresses many of the problems of existing systems. It makes especially heavy use of the additional structure present in hypertext to provide much higher quality search results. It is developed by Google LLC. It is the most used search engine on the World Wide Web across all platforms, with 92.16% market share as of December 2020, handling more than 5.4 billion searches each day. Google Search also provides many different options for customized searches, using symbols to include, exclude, specify or require certain search behavior, and offers specialized interactive experiences, such as flight status and package tracking, weather forecasts, currency, unit, and time conversions, word definitions, and more.

2. Background of the organization

Google's history began in 1995 when Larry Page met Sergey Brin. At the time, Larry Page was a Ph.D. student at Standford University, and Sergey was considering studying there. In 1996, the pair began work on a search engine called BackRub. The name comes from the algorithm-generated ranking for how many "back-links" a page has. (1)

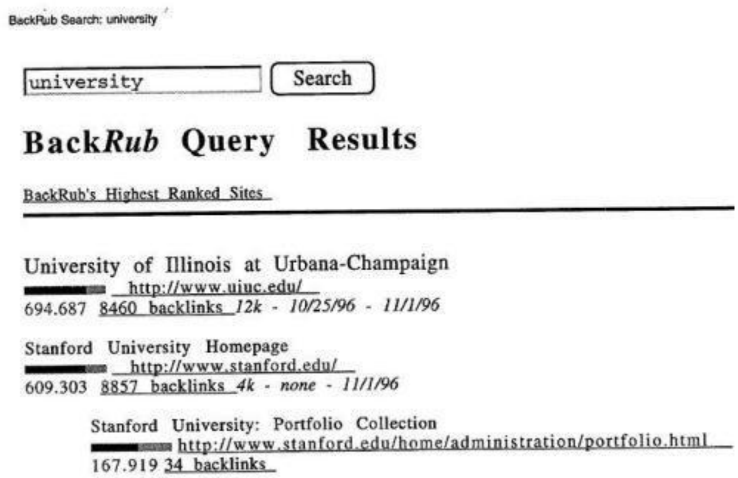

Fig.1. Query results of Backrub search engine

This engine worked on the Stanford servers for more than a year before it eventually clogged up the bandwidth and was forced to move. Google.com was registered on September 15th, 1997.

The name Google is a play on "googol," which is a mathematical term for the number 1 followed by 100 zeros. It is rumored that this reflects the founders' mission to organize the infinite amount of information on the internet.

Over the next few years, Google caught the attention of not only the academic community, butSilicon Valley investors as well. In August 1998, Sun co-founder Andy Bechtolsheim wrote Larry and Sergey a check for $100,000, and Google Inc. was officially born. With this investment, the newly incorporated team made the upgrade from the dorms to their first office: a garage in suburban Menlo Park, California, owned by Susan Wojcicki (employee #16 and now CEO of YouTube). (2)

Fig 2. Larry Page and Sergey Brin in their first office: Susan Wojcicki’s garage.

Fig 3.The garage where the Google was born in

In 1999, Google moved from its humble garage to new digs at 165 University Avenue, Palo Alto. At this time they were eight employees strong. About five months later, Kleiner Perkins Caufield and Byers and Sequoia Capital agreed to invest **$25 Million** in the company. These two venture capital firms were normally fierce rivals. Seeing Google's potential, however, they both took seats on the board of directors.

In June 2000, Google was recognized as the world's largest search engine. By 2002, Google earned several awards including Best Search Feature and Best Design awards. The company gained success by continuously enhancing its products and services. The company also launched a free email account, called Gmail (Google Milestones 2013).

n 2003, after outgrowing two other locations, the company leased an office complex from Silicon Graphics, at 1600 Amphitheatre Parkway in Mountain View, California. The complex became known as the Googleplex, a play on the word googolplex, the number one followed by a googol zeroes.

Figure 4.Google's headquarter in Mountain view,California

On August 19, 2004 the company has. The IPO earned Google USD $1.67 billion, which gave the company a total market capitalization of USD $23 billion (Elgin, 2004). (3)

Google has achieved great success in growing its internet-related products and services. In line with this, it has acquired several small entrepreneurial ventures like Keyhole Inc, YouTube, Double Click, Grand Central, Aardvark and On2 Technologies (Google Milestones 2013). In recent years, Google has become a significant player in the telecom industry with its development of the Android mobile system. The company is also increasing its hardware business through its partnerships with major electronics manufacturers.

3. Corporate Mission and Organizational Culture

Google's mission statement is

"to organize the world's information and make it universally accessible and useful,"

but also has an unofficial statement - *"Don't be evil"*. This motto was replaced in 2015 to *"Do the right thing"*.

Google is fundamentally built upon a culture of openness and sharing of ideas and opinions. This is primarily influenced by the company's beginnings as an internet startup company. In line with this principle, the company has encouraged its employees to ask questions directly to top executives about various company issues (Google 2013). Google is a dynamic company wherein everyone's ideas are respected and heard.

An important characteristic of learning organizations like Google is that they do not strictly dictate their employees' functions. Employees are instead provided with adequate flexibility to pursue their interest based on the premise that they are strategically aligned with the organization's goals. By doing this, employees play an active role in shaping the organization rather than passively following prescribed routines. Employees are encouraged to express ideas and challenge themselves with new targets, which consequently contribute towards an improved work environment. Such active employee participation is a paradigm shift from the traditional authoritarian management, which was deemed less potent at harnessing greater human potential. Learning organizations can 'create the results they truly desire and where they can learn to learn together for the betterment of the whole.

4. Design

The main design goal of the Google search engine was to improve the quality of web search engines. (4) These tasks are becoming increasingly difficult as the Web grows. However, hardware performance and cost have improved dramatically to partially offset the difficulty. There are, however, several notable exceptions to this progress such as disk seek time and operating system robustness.

In designing Google, they have considered both the rate of growth of the Web and technological changes. Google is designed to scale well to extremely large data sets. It makes efficient use of storage space to store the index. Its data structures are optimized for fast and efficient access Further, They designed Google in such a way that the cost to index and store text or HTML will eventually decline relative to the amount that will be available. This will result in favorable scaling properties for centralized systems like Google.

Generally there are three basic components of a search engine as listed below:

1. **Web Crawler**: It is also known as spider or bots**.** It is a software component that traverses the web to gather information. It gathers information from across hundreds of billions of webpages and organize it in the Search index.

2. **Database**: All the information on the web is stored in database. It consists of huge web resources. Google uses *Bigtable* for managing structured data that is designed to scale to a very large size: petabytes of data across thousands of commodity servers. (5) BigTable is built on *Google File System* **(***GFS*), which is used as a backing store for log and data files. GFS provides reliable storage for SSTables, a Google-proprietary file format used to persist table data.
3. **Search Interfaces:** This component is an interface between user and the database. It helps the user to search through the database.

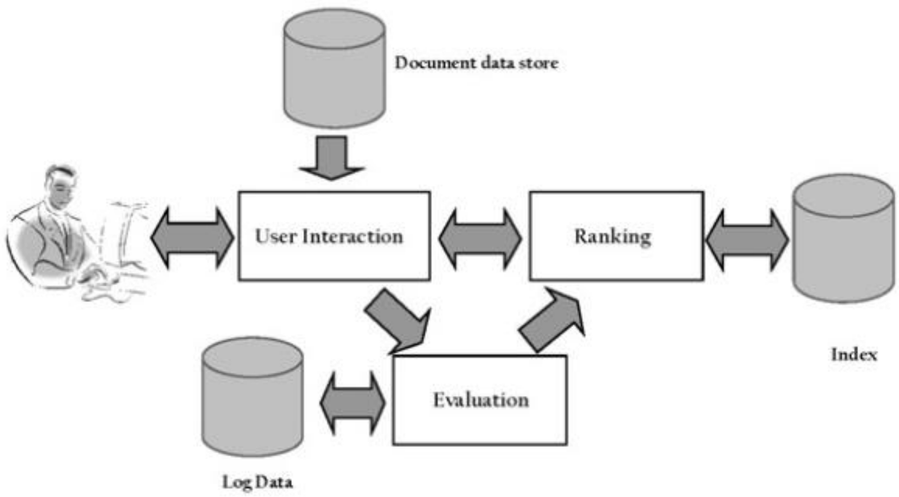

Fig 5.Design of Google search engine

Google refers to a huge database of internet resources such as web pages, newsgroups, programs, images etc. It helps to locate information on World Wide Web.User can search for any information by passing query in form of keywords or phrase. It then searches for relevant information in its database and return to the user.

4.1. Web Crawling

The crawling process begins with a list of web addresses from past crawls and sitemaps provided by website owners. As the crawlers visit these websites, they use links on those sites to discover other pages. The software pays special attention to new sites, changes to existing sites and dead links. Computer programs determine which sites to crawl, how often and how many pages to fetch from each site.

Google offer Search Console to give site owners granular choices about how Google crawls their site: they can provide detailed instructions about how to process pages on their sites, can request a recrawl or can opt out of crawling altogether using a file called “robots.txt”. Google never accepts payment to crawl a site more frequently — we provide the same tools to all websites to ensure the best possible results for the users.

When crawlers find a webpage, Google systems render the content of the page, just as a browser does.They take note of key signals from keywords to website freshness and keep track of it all in the Search index.

The Google Search index contains hundreds of billions of webpages and is well over 100,000,000 gigabytes in size. It's like the index in the back of a book — with an entry for every word seen on every webpage we index. When they index a webpage ,they add it to the entries for all of the words it contains. Google has its main crawler, Googlebot, which encompasses mobile and desktop crawling. But there are also several additional bots for Google, like Googlebot Images, Googlebot Videos, Googlebot News, and AdsBot.

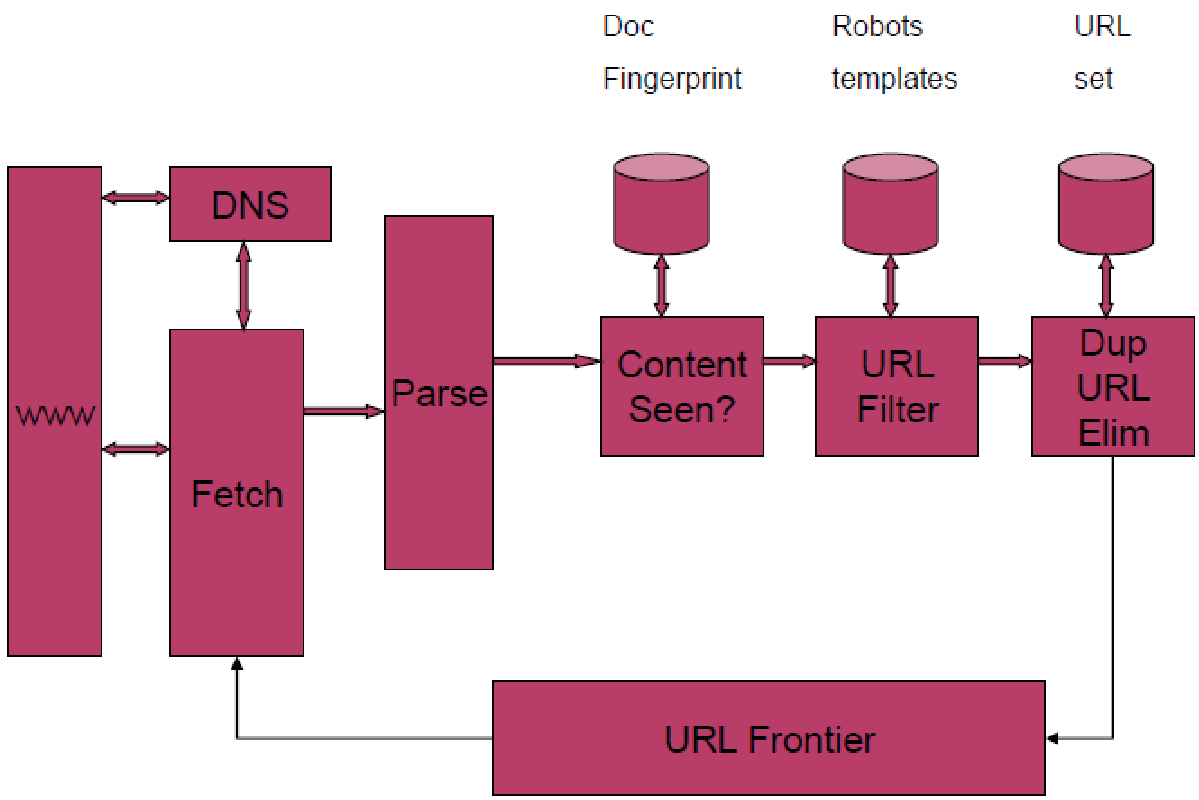

Fig.6. System Desgin of the web crawler

In the system design of the web Crawler:

- **URL Frontier:** Contains URLs yet to be fetches in the current crawl. At first, a seed set is stored in URL Frontier, and a crawler begins by taking a URL from the seed set.
- **DNS:** domain name service resolution. Look up IP address for domain names.
- **Fetch:** generally use the http protocol to fetch the URL.
- **Parse:** the page is parsed. Texts (images, videos, and etc.) and Links are extracted.
- **Content Seen?:** test whether a web page with the same content has already been seen at another URL. Need to develop a way to measure the fingerprint of a web page.

4.2. BigTable

Bigtable is a distributed storage system (built by Google) for managing structured data that is designed to scale to a very large size: petabytes of data across thousands of commodity servers. (6)

Some of the features of Big table are:

- Fast and extremely large-scale DBMS
- A sparse, distributed multi-dimensional sorted map, sharing characteristics of both row-oriented and column-oriented databases.
- Designed to scale into the petabyte range
- It works across hundreds or thousands of machines
- It is easy to add more machines to the system and automatically start taking advantage of those resources without any reconfiguration
- Each table has multiple dimensions (one of which is a field for time, allowing versioning)
- Tables are optimized for GFS (Google File System) by being split into multiple tablets - segments of the table as split along a row chosen such that the tablet will be ~200 megabytes in size.

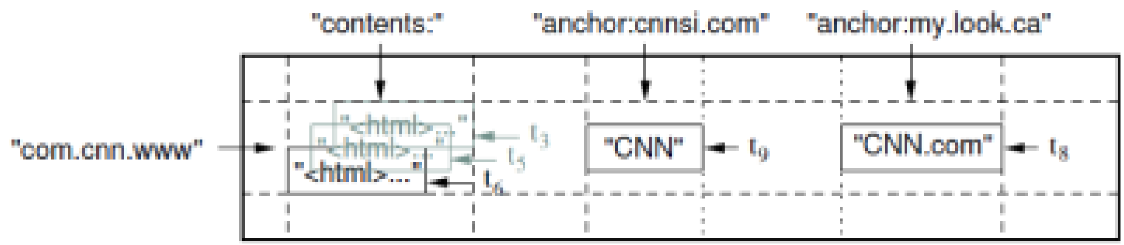

Fig.7.Architecture of Google Bigtable

Fig. shows slice of an example table that stores Web pages. The row name is a reversed URL. The contents column family contains the page contents, and the anchor column family contains the text of any anchors that reference the page. CNN's home page is referenced by both the Sports Illustrated and the MY-look home pages, so the row contains columns named anchor:cnnsi.com and anchor:my.look.ca. Each anchor cell has one version; the contents column has three versions, at timestamps t3, t5, and t6.

5. Architecture

In this section, we will give a high level overview of the architecture of the whole system . Most of Google is implemented in C or C++ for efficiency and can run in either Solaris or Linux.

In Google, the web crawling (downloading of web pages) is done by several distributed crawlers. There is a URLserver that sends lists of URLs to be fetched to the crawlers. The web pages that are fetched are then sent to the storeserver. The storeserver then compresses and stores the web pages into a repository.

Every web page has an associated ID number called a docID which is assigned whenever a new URL is parsed out of a web page. The indexing function is performed by the indexer and the sorter. The indexer performs a number of functions. It reads the repository, uncompresses the documents, and parses them. Each document is converted into a set of word occurrences called hits. The hits record the word, position in document, an approximation of font size, and capitalization. The indexer distributes these hits into a set of "barrels", creating a partially sorted forward index. The indexer performs another important function. It parses out all the links in every web page and stores important information about them in an anchors file. This file contains enough information to determine where each link points from and to, and the text of the link.

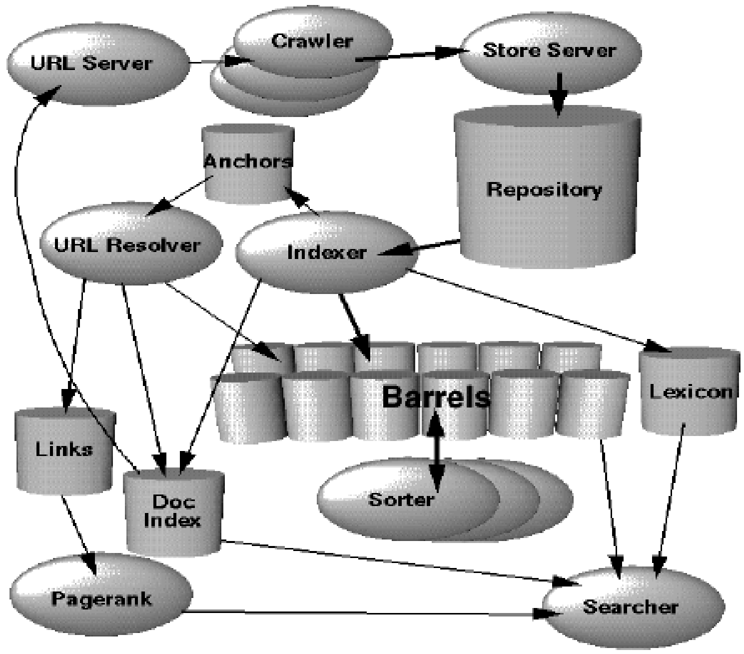

Figure8. High Level Google Architecture

The URLresolver reads the anchors file and converts relative URLs into absolute URLs and in turn into docIDs. It puts the anchor text into the forward index, associated with the docID that the anchor points to. It also generates a database of links which are pairs of docIDs. The links database is used to compute PageRanks for all the documents. The sorter takes the barrels, which are sorted by docID (this is a simplification, see Section 4.2.5), and resorts them by wordID to generate the inverted index. This is done in place so that little temporary space is needed for this operation. The sorter also produces a list of wordIDs and offsets into the inverted index. A program called DumpLexicon takes this list together with the lexicon produced by the indexer and generates a new lexicon to be used by the searcher. The searcher is run by a web server and

uses the lexicon built by DumpLexicon together with the inverted index and the PageRanks to answer queries.

6. Layout

Google launched their search tool in 1998 and their advertising platform Google AdWords debuted in 2000.

6.1.Google Search Interface

A search interface is the graphical user interface used by employees to search a data source. Google's search interface is very simple, neat and clean.

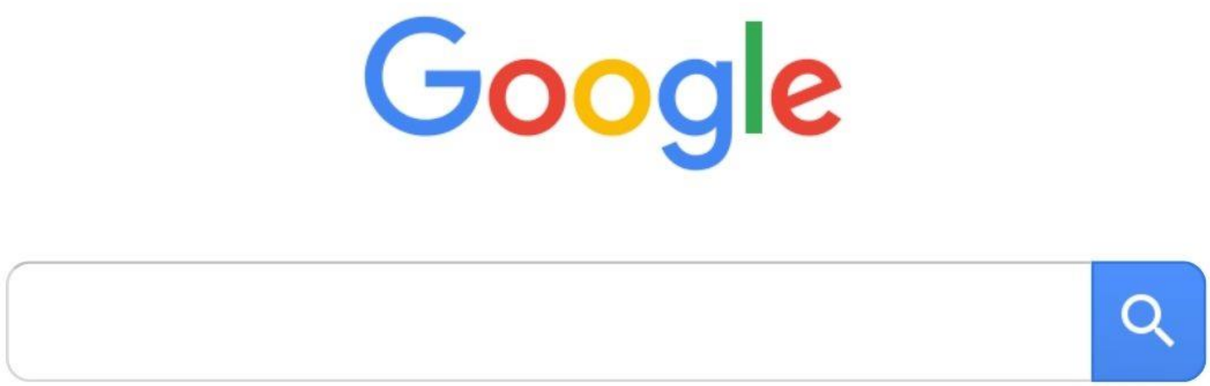

Fig9. Search interface of Google

6.2.Google Search Engine Result Page

Google is continuously iterating on not only their search algorithms but also the layout and display of how search results are formatted and presented to users.

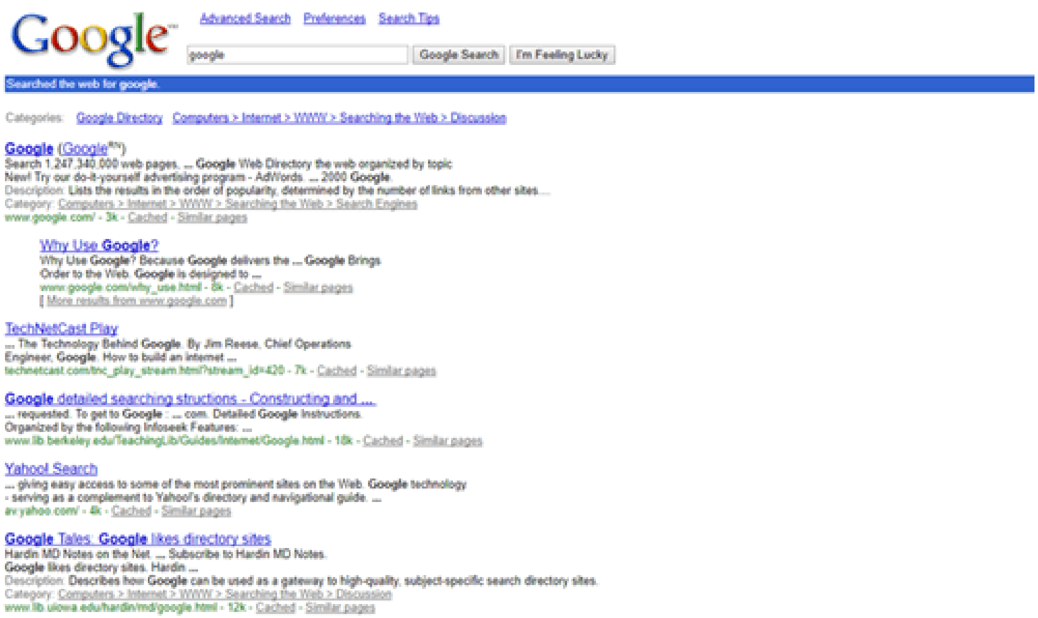

Fig 10.Google Search Engine Result Page in 2008

Google released a desktop SERP update on January 13, 2020 with the stated intention to have the desktop SERP more closely mirror the look of mobile search results. The essential change placed a company's flavicon and page URL above the traditional page title (blue link).

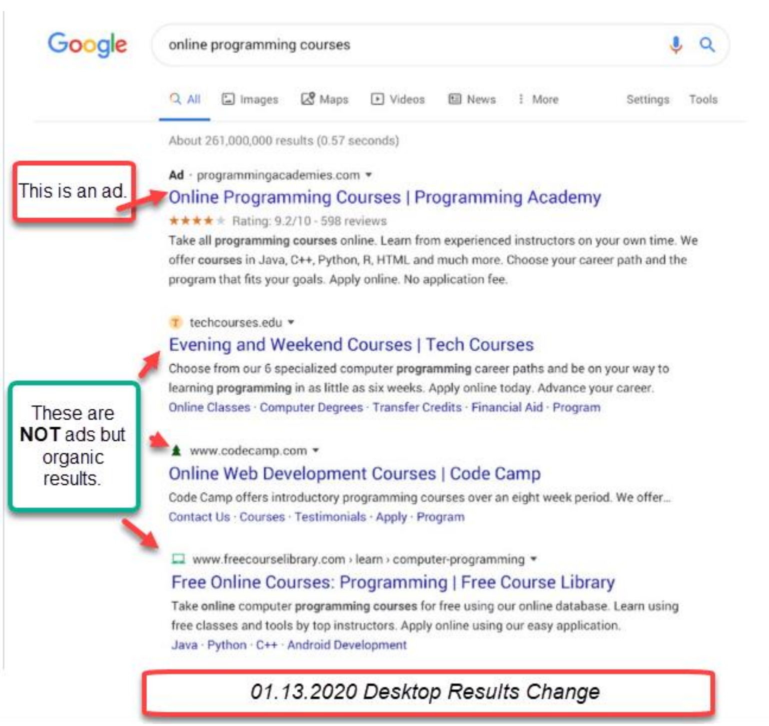

Fig 11.Google search engine result page(SERP) in 2020

6.3.Google Ads

Depending on the search query, Google will populate ad campaigns right at the top of a SERP. The more popular and competitive a phrase, the more ads we will see. On this SERP for the phrase, “call tracking software,” we see four ads before the organic results even begin. (7)

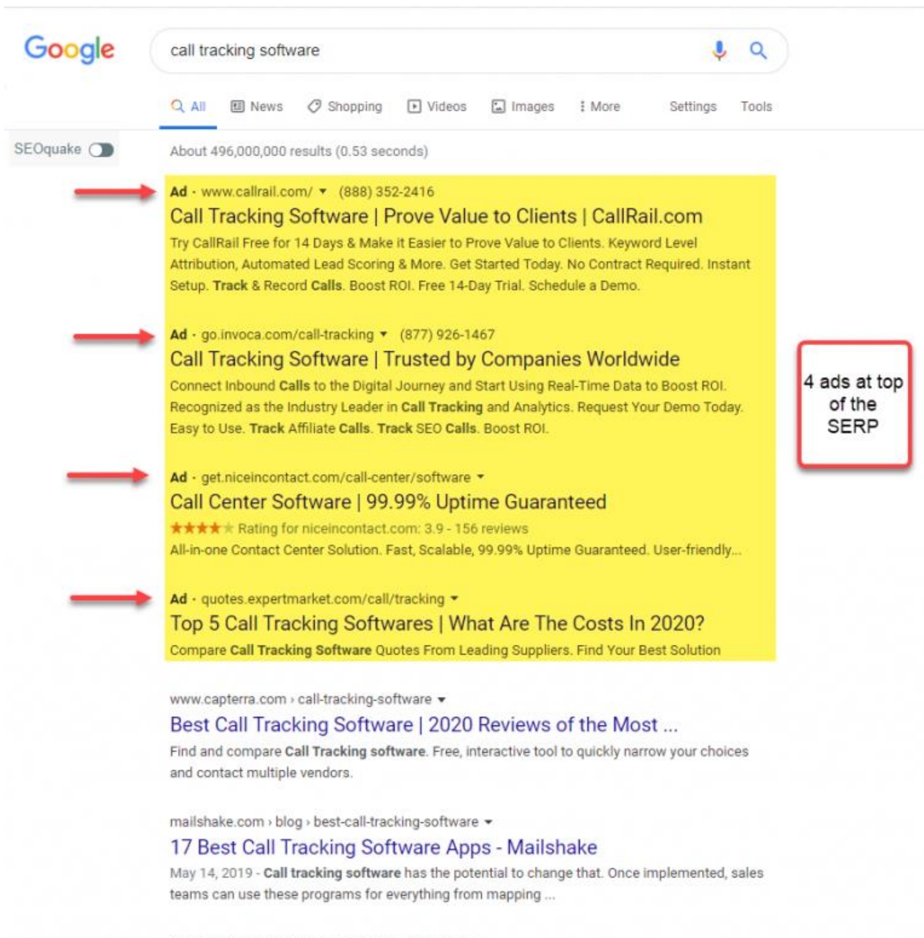

Fig 12.Google adsense result

7. Working Procedure

Google search engines work through three primary functions:

1) **Crawling:** The first step is finding out what pages exist on the web. There isn't a central registry of all web pages, so Google must constantly search for new pages and add them to its list of known pages. Some pages are known because Google has already visited them before. Other pages are discovered when Google follows a link from a known page to a new page. Still other pages are discovered when a website owner submits a list of pages (a *sitemap*) for Google to crawl. If you're using a managed web host, such as Wix or Blogger, they might tell Google to crawl any updated or new pages that you make.Once Google discovers a page URL, it visits, or *crawls*, the page to find out what's on it.Google renders the page and analyzes both the text and non-text content and overall visual layout to decide where it should appear in Search results. The better that Google can understand your site, the better we can match it to people who are looking for your content. (8)

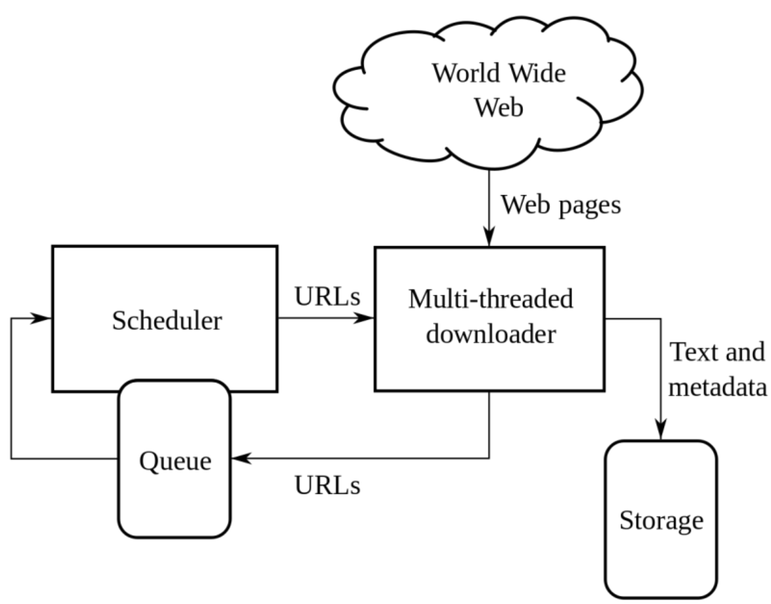

Fig 13. Functional diagram of web crawler

2) **Indexing:** After a page is discovered, Google tries to understand what the page is about. This process is called *indexing*. Google analyzes the content of the page, catalogs images and video files embedded on the page, and otherwise

tries to understand the page. This information is stored in the *Google index*, a huge database stored in many, many (many!) computers.

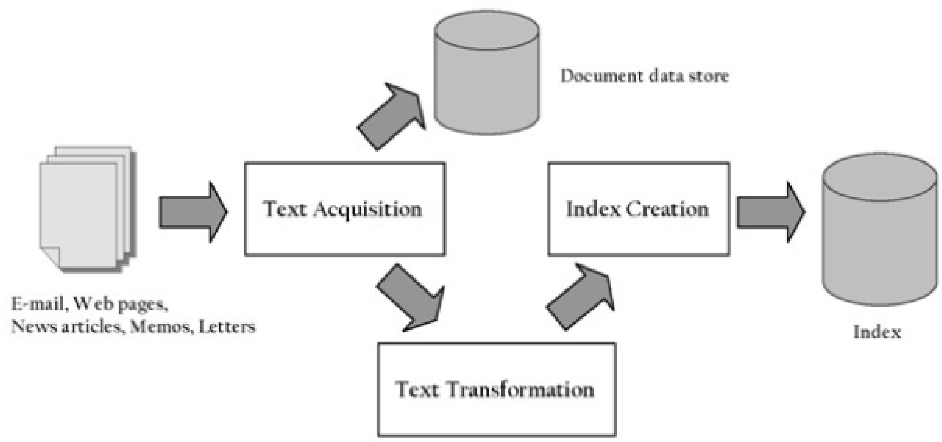

Fig 14.Indexing process

a) **Parsing :**Any parser which is designed to run on the entire Web must handle a huge array of possible errors. These range from typos in HTML tags to kilobytes of zeros in the middle of a tag, non-ASCII characters, HTML tags nested hundreds deep, and a great variety of other errors that challenge anyone's imagination to come up with equally creative ones. For maximum speed, instead of using YACC to generate a CFG parser, they use flex to generate a lexical analyzer which they outfit with its own stack. Developing this parser which runs at a reasonable speed and is very robust involved a fair amount of work.

b) **Indexing Documents into Barrels:** After each document is parsed, it is encoded into a number of barrels. Every word is converted into a wordID by using an in-memory hash table -- the lexicon. New additions to the lexicon hash table are logged to a file. Once the words are converted into wordID's, their occurrences in the current document are translated into hit lists and are written into the forward barrels. The main difficulty with parallelization of the indexing phase is that the lexicon needs to be shared. Instead of sharing the lexicon, we took the approach of writing a log of all the extra words that were not in a base lexicon, which we fixed at 14 million words. That way multiple indexers can run in parallel and then the small log file of extra words can be processed by one final indexer.

c) **Sorting :** In order to generate the inverted index, the sorter takes each of the forward barrels and sorts it by wordID to produce an inverted barrel for title and anchor hits and a full text inverted barrel. This process happens one barrel at a time, thus requiring little temporary storage. Also, we parallelize the sorting phase to use as many machines as we have simply by running multiple sorters, which can process different buckets at the same time. Since the barrels don't fit into main memory, the sorter further subdivides them into baskets which do fit into memory based on wordID and docID. Then the sorter, loads each basket into memory, sorts it and writes its contents into the short inverted barrel and the full inverted barrel.

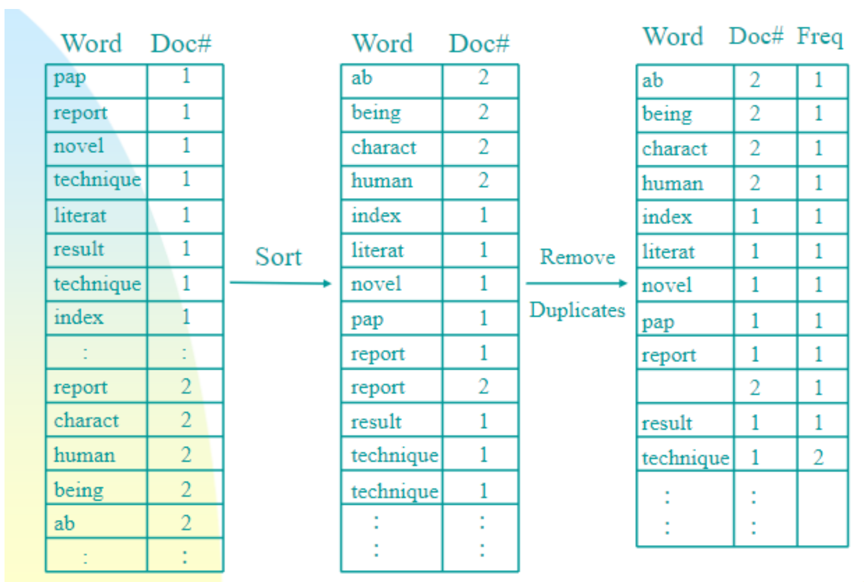

Fig 15.Sample Indexing

3) **Searching:** The goal of searching is to provide quality search results efficiently.

They do it with a ranking and feedback system. PageRank algorithm is used for ranking.

a) **Ranking:** Google ranking systems are designed to do just that: sort through hundreds of billions of webpages in our Search index to find the most relevant, useful results in a fraction of a second, and present them in a way that helps you find what you're looking for. These ranking systems are made up of not one, but a whole series of algorithms. To give one the most useful information, Search algorithms look at many factors, including the words of your query, relevance and usability of pages, expertise of sources, and your location and settings. The weight applied to each factor varies depending on the nature of your query—for example, the freshness of the content plays a bigger role in answering queries about current news topics than it does about dictionary definitions. (9)

b) The ranking function has many parameters like the type-weights and the type-prox-weights. Figuring out the right values for these parameters is something of a black art. In order to do this, we have a user feedback mechanism in the search engine. A trusted user may optionally evaluate all of the results that are returned. This feedback is saved. Then when we modify the ranking function, we can see the impact of this change on all previous searches which were ranked. Although far from perfect, this gives

8. Algorithm

The Google search engine has two important features that help it produce high precision results. It makes use of the link structure of the Web to calculate a quality ranking for each web page. This ranking algorithm is called *PageRank*.Google utilizes link to improve search results.

Google uses a machine-learning artificial intelligence system called *"RankBrain"* to help sort through its search results. (10)

Google has created maps containing as many as 518 million of these hyperlinks, a significant sample of the total. These maps allow rapid calculation of a web page's "PageRank", anobjective measure of its citation importance that corresponds well with people's subjective idea of importance. Because of this correspondence, PageRank is an excellent way to prioritize the results of web keyword searches. For most popular subjects, a simple text matching search that is restricted to web page titles performs admirably when PageRank prioritizes the results. For the type of full text searches in the main Google system, PageRank also helps a great deal.

8.1.**Description of PageRank Calculation**

Academic citation literature has been applied to the web, largely by counting citations or backlinks to a given page. (11)This gives some approximation of a page’s importance or quality. PageRank extends this idea by not counting links from all pages equally, and by normalizing by the number of links on a page.

PageRank is defined as follows:

We assume page A has pages T1...Tn which point to it (i.e., are citations). The parameter d is a damping factor which can be set between 0 and 1. We usually set d to 0.85. There are more details about d in the next section. Also C(A) is defined as the number of links going out of page A. The PageRank of a page A is given as follows:

$$PR(A) = (1-d) + d\,(PR(T1)/C(T1) + ... + PR(Tn)/C(Tn))$$

Note that the PageRanks form a probability distribution over web pages, so the sum of all web pages’ PageRanks will be one.

So basically,A page has a high PageRank R if

- There are many pages linking to it
- Or, if there are some pages with a high PageRank linking to it

PageRank or *PR(A)* can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web. Also, a PageRank for 26 million webpages can be computed in a few hours on a medium size workstation.

8.1.1. **Intuitive Justification**

PageRank can be thought of as a model of user behavior. We assume there is a "random surfer" who is given a web page at random and keeps clicking on links, never hitting "back" but eventually gets bored and starts on another random page. The probability that the random surfer visits a page is its PageRank. And, the *d* damping factor is the probability at each page the "random surfer" will get bored and request another random page. One important variation is to only add the damping factor *d* to a single page, or a group of pages. This allows for personalization and can make it nearly impossible to deliberately mislead the system in order to get a higher ranking. They also have several other extensions to PageRank, Another intuitive justification is that a page can have a high PageRank if there are many pages that point to it, or if there are some pages that point to it and have a high PageRank. Intuitively, pages that are well cited from many places around the web are worth looking at. Also, pages that have perhaps only one citation from something like the Yahoo! homepage are also generally worth looking at. If a page was not high quality, or was a broken link, it is quite likely that Yahoo's homepage would not link to it.

PageRank handles both these cases and everything in between by recursively propagating weights through the link structure of the web.

8.2.RankBrain Algorithm

RankBrain is considered the most useful search ranking signal, behind content and links. But prior to this, Google had only said publicly in 2015, October that RankBrain was used to process a "large fraction" of 15 percent of the searches it had never seen before. RankBrain is Google's name for a machine-learning artificial intelligence system that's used to help process its search results, as was reported by Bloomberg and also confirmed to us by Google.

RankBrain is part of Google's overall search "algorithm," a computer program that's used to sort through the billions of pages it knows about and find the ones deemed most relevant for particular queries. RankBrain is designed to help better interpret those queries and effectively translate them, behind the scenes in a way, to find the best pages for the searcher. it can see patterns between seemingly unconnected complex searches to understand how they're actually similar to each other. This learning, in turn, allows it to better understand future complex searches and whether they're related to particular topics. Most important, from what Google told us, it can then associate these groups of searches with results that it thinks searchers will like the most.

Google didn't provide examples of groups of searches or give details on how RankBrain guesses at what are the best pages. But the latter is probably because if it can translate an ambiguous search into something more specific, it can then bring back better answers.

9. Challenges and how to overcome them

Creating a search engine which scales even to today's web presents many challenges.

- Fast crawling technology is needed to gather the web documents and keep them up to date.
- Storage space must be used efficiently to store indices and, optionally, the documents themselves.
- The indexing system must process hundreds of gigabytes of data efficiently. Queries must be handled quickly, at a rate of hundreds to thousands per second.

These tasks are becoming increasingly difficult as the Web grows. However, hardware performance and cost have improved dramatically to partially offset the difficulty. There are, however, several notable exceptions to this progress such as disk seek time and operating system robustness.

- In designing Google, They have considered both the rate of growth of the Web and technological changes. Google is designed to scale well to extremely large data sets. It makes efficient use of storage space to store the index. Its data structures are optimized for fast and efficient access.The data structures that are used in Google are:

  - **BigFiles:** BigFiles are virtual files spanning multiple file systems and are addressable by 64 bit integers. The allocation among multiple file systems is handled automatically. The BigFiles package also handles allocation and deallocation of file descriptors, since the operating systems do not provide enough for our needs. BigFiles also support rudimentary compression options.

  - **Repository:** The repository contains the full HTML of every web page. Each page is compressed using zlib.The choice of compression technique is a tradeoff between speed and compression ratio.

- **Document Index:** The document index keeps information about each document. This design decision was driven by the desire to have a reasonably compact data structure, and the ability to fetch a record in one disk seek during a search

- **Lexicon:** In the current implementation we can keep the lexicon in memory on a machine with 256 MB of main memory. The current lexicon contains 14 million words

- They expect that the cost to index and store text or HTML will eventually decline relative to the amount that will be available.This will result in favorable scaling properties for centralized systems like Google.

10. Findings

Google's rise was largely due to a patented algorithm called PageRank which helps rank web pages that match a given search string. Google's unique and improving algorithm has made it one of the most popular search engines of all time. Other search engines continue to have a difficult time matching the relevancy algorithm Google has created by examining a number of factors such as social media, inbound links, fresh content, etc.

Google appeared on the search engine scene in 1996. Google was unique because it ranked pages according to citation notation, in which a mention of one site on a different website became a vote in that site's favor.

Google also began judging sites by authority. A website's authority, or trustworthiness, was determined by how many other websites were linking to it, and how reliable those outside linking sites were.

Google search history can be witnessed by taking a look at Google's homepage progressions over the years. It's remarkable to see how basic and primitive the now most popular search engine once was. No doubt Google has made revolutionary improvement in the field of web search engine.

11. Summary

Google is a large-scale search engine which addresses many of the problems of existing systems low quality matches, incapability to handle large amount of data etc . It makes especially heavy use of the additional structure present in hypertext to provide much higher quality search results. The order of search results returned by Google is based, in part, on a priority rank system called "PageRank". Google has started with a mission to organize the world's information and make it universally accessible and useful in 1997 by Larry Page, Sergey Brin, and Scott Hassan.Now It is the most used search engine on the World Wide Web across all platforms, with 92.16% market share as of December 2020, handling more than 5.4 billion searches each day.

12. Remarks

Google perfected the search engine, or at least greatly improved what it could be. It proved that by analyzing how people linked to web pages, you could get both highly relevant and comprehensive search results. We could search for both popular and obscure topics and receive great answers in response.

Google also thrived by being a “second-mover” player. The early search engines were largely prevented from making money directly off search because of outcry over paid results. These are commonplace and accepted today, but after an initial experiment by Open Text in 1996 sparked upset, other players shied away.

Over the years, Google eclipsed its competitors. Today, Google is top of mind when you think of search. In most countries, other than a few like China and Russia, it’s the most-used search service. It has earned that ultimate compliment for a business, having its name turned into a verb. To google is to search.

Works Cited

1. Almost Everything You Need to Know About Google's History. *Interesting Engineering.* [Online] June 28, 2020. https://interestingengineering.com/almost-everything-you-need-to-know-about-googles-history.

2. Our story. *Google.* [Online] https://about.google/our-story/.

3. Google Goes Public – Files for IPO. *Search Engine Journal.* [Online] April 29, 2014. https://www.searchenginejournal.com/google-goes-public-files-for-ipo/.

4. *The Anatomy of a Large-Scale Hypertextual.* **Page, Sergey Brin and Lawrence.** Stanford, CA 94305, USA : Stanford University.

5. Cloud Bigtable. *Google Cloud.* [Online] https://cloud.google.com/bigtable.

6. *Bigtable: A Distributed Storage System for Structured Data.* **Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach.** s.l. : Google Inc., 2006.

7. *Google advertising tools: Cashing in with AdSense, AdWords, and the Google APIs.* **Davis, Harold.** s.l. : O'Reilly Media, Inc., 2006.

8. How Google Search Works. *Google Search Central.* [Online]

9. How Search algorithms work. *Google Search.* [Online] https://www.google.com/search/howsearchworks/algorithms/.

10. *FAQ: All about the Google RankBrain algorithm.* **Sullivan, Danny.** s.l. : https://searchengineland. com/faq-all-about-the-new-google-rankbrain-algorithm-234440, 2016.

11. *The PageRank citation ranking: Bringing order to the web.* **Page, Lawrence and Brin, Sergey and Motwani, Rajeev and Winograd, Terry.** s.l. : tanford InfoLab, 1999.