

# Scalability and Efficiency Challenges in Commercial Web Search Engines

B. Barla Cambazoglu  
Yahoo! Labs  
Barcelona, Spain  
barla@yahoo-inc.com

Ricardo Baeza-Yates  
Yahoo! Labs  
Barcelona, Spain  
rbaeza@acm.org

## ABSTRACT

Commercial web search engines rely on very large compute infrastructures to be able to cope with the continuous growth of the Web and user bases. Achieving scalability and efficiency in such large-scale search engines requires making careful architectural design choices while devising algorithmic performance optimizations. Unfortunately, most details about the internal functioning of commercial web search engines remain undisclosed due to their financial value and the high level of competition in the search market. The main objective of this tutorial is to provide an overview of the fundamental scalability and efficiency challenges in commercial web search engines, bridging the existing gap between the industry and academia.

## Categories and Subject Descriptors

H.3.3 [Information Storage Systems]: Information Retrieval Systems

## General Terms

Algorithms, Design, Performance

## Keywords

Web search engines, crawling, indexing, query processing, caching, efficiency, scalability.

## 1. DESCRIPTION

Commercial web search engines need to process thousands of queries every second and provide responses to user queries within a few hundred milliseconds. As a consequence of these tight performance constraints, search engines construct and maintain very large compute infrastructures for crawling the Web, indexing discovered pages, and processing user queries [3]. The scalability and efficiency of these infrastructures require careful performance optimizations in every major component of the search engine [2].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGIR '13, July 28–August 1, 2013, Dublin, Ireland.  
ACM 978-1-4503-2034-4/13/07.

This tutorial aims to provide a fairly comprehensive overview of the scalability and efficiency challenges in large-scale web search engines [1]. In particular, the tutorial provides an in-depth architectural overview of a web search engine, mainly focusing on the web crawling, indexing, query processing, and caching components. The scalability and efficiency issues encountered in the above-mentioned components are presented at four different granularities: at the level of a single computer, a cluster of computers, a single data center, and a multi-center search engine. The tutorial also points at the open research problems and provides recommendations to researchers who are new to the field.

## 2. OBJECTIVES

The tutorial aims to achieve the following objectives:

1. To provide in-depth background on the basic architectural components in a web search engine.
2. To present the fundamental scalability and efficiency issues which have been often addressed in the information retrieval literature.
3. To shed some light into the techniques used in large-scale commercial search engines and bridge the gap between the industry and academia.
4. To identify the open research problems in the context of web search engine scalability and efficiency, promoting further research on the topic.

## 3. REFERENCES

- [1] B. B. Cambazoglu and R. Baeza-Yates. Scalability challenges in web search engines. In M. Melucci, R. Baeza-Yates, and W. B. Croft, editors, *Advanced Topics in Information Retrieval*, volume 33 of *The Information Retrieval Series*, pages 27–50. Springer Berlin Heidelberg, 2011.
- [2] A. Chowdhury and G. Pass. Operational requirements for scalable search systems. In *Proceedings of the 12th International Conference on Information and Knowledge Management*, pages 435–442, 2003.
- [3] J. Dean. Challenges in building large-scale information retrieval systems: invited talk. In *Proceedings of the 2nd ACM International Conference on Web Search and Data Mining*, page 1, 2009.