

# Construction of Green Information Retrieval System Based on Cloud Computing

Zhang Yan

Yun Yang Teachers' College  
Shi Yan, Hu Bei, 442200, China

**Abstract**—The development of information technology has greatly improved people's working efficiency, but also brings about environment pollution and resource consumption. The information retrieval functions also work in this way. Cloud computing technology can effectively reduce energy consumption, decrease environmental pollution and optimize the information retrieval system. This paper proposes to construct green retrieval system 4 key factors: standardization, sharable, reusable and green user behavior so as to provide reference for construction of green information retrieval system.

**Keywords**- cloud computing; information retrieval system

## I. ENVIRONMENT HARM BROUGHT BY TRADITIONAL INFORMATION RETRIEVAL

For a long time, the overall impression of the IT industry is "smokeless industry". The conventional wisdom says IT industry will not pollute the environment. In fact, is "smokeless industry", the influence of IT equipment on the environment of the IT industry is great, which exceeded the aviation industry. It is estimated that human because of the amount of carbon dioxide discharge increasing use of information technology equipment accounted for 2% of total CO<sub>2</sub> emissions in the world in order to maintain the normal operation of the global computer by the international agency for information technology research. Humans will increase about 35,000,000 tons of waste gas emissions to the atmosphere each year. Consumption is not only a large number of server power consumption, environmental pollution caused by the generation of these power energy is also worrying. These contaminants can cause serious damage to the urban environment. We should pay huge manpower and cost of capital to make up for the damage.

With the rapid development of communication technology and network, search engine has become the most commonly used tool for people, information retrieval tool which is convenient for a user to search for knowledge and information in the network or in digital library. However, these are rooted in information and communication technology information retrieval system on the device can cause a lot of environmental pollution and ecological destruction. In the environmental problems become increasingly prominent today, the extent we rely on information retrieval system is bigger and the negative impact on the environment will be greater. Therefore, it has the important practical significance to explore how to maintain the sustainable development and environment in the development of information retrieval systems.

## II. OPTIMIZATION OF INFORMATION RETRIEVAL SYSTEM BASED ON CLOUD COMPUTING

Through the sharing of computing and cyber source to reduce the economic and environmental costs, cloud computing is considered to be one of the most appropriate choice of realization of green IT. Cloud computing is a new computing as a service organization of resources allocation of computing resources and usage patterns. The type of the service can be divided into infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). According to different service levels, different types of services are related to different levels of resource sharing. Cloud computing can choose different ways to share software and hardware resources and information in order to achieve reasonable economic and environmental benefits.

We use the key words of "cloud computing" and "information retrieval" respectively in the ISI Web of Knowledge database and China National Knowledge Infrastructure (CNKI) database to search, results of research papers related to the earliest published on 2008, shows the related research in this field is at the starting stage, the related research papers in foreign countries total to 81. The number of the related research papers from china is 16. In addition, the related research is not confined to cloud computing and information retrieval, also contains the cloud computing and information management, information retrieval and sharing of resources etc. Basirat and Khand discussed how to use pattern recognition method of distributed cloud storage and cloud data retrieval; Bales discussed the use of mobile phone and other computing devices, especially smart mobile phone information access and access. An Junxiu put forward the data storage scheme of distributed cloud retrieval; Fu Yonggui put forward a new model of information retrieval for cloud computing platform for information retrieval characteristic.

These studies indicate that information retrieval researchers increasingly began to consider cloud computing as a new opportunity to their extended research field. But as discussed above, the current study of retrieval systems services lack the environmental impact. Specifically, it is the green information retrieval system and how to access the services in different application environments and the use of digital information in different forms in order to promote the sustainable development and the environment. However, the main challenge of green information retrieval system of greenhouse gases is facing the lack of an appropriate method to calculate the information retrieval system and the service

discharge. Information retrieval system and personal terminal is the main contribution to the energy consumption of internet. But cloud computing curing can reduce the information retrieval system can, in addition, the user's behavior (such as the choice of the type of terminal in information retrieval, retrieval time and retrieval behavior habit etc.) will also have an important effect on the reduction of energy consumption. Therefore, in order to build a green information retrieval system, we can ultimately achieve green goal through the use of cloud computing and change the habits of users.

### III. PRINCIPLES OF THE CONSTRUCTION OF GREEN INFORMATION RETRIEVAL SYSTEM

#### A. Standardization

Related research found, cloud computing technology superiority mainly manifests to create, organize and process standard in digital content. In addition, Hargreaves et al in the related research of the digital economy is innovation support knowledge intensive activities and put forward some suggestions, with regard to standard advice, also has the important reference for the development of the information retrieval system. Therefore, based on standardization, through the cloud architecture of the PaaS and SaaS layer, the remote access and computing technology is utilized to acquire the data cloud, will greatly facilitate information retrieval. At the same time, relative to the localized data center hosting, the digital content is transferred to the cloud will also significantly reduce greenhouse gas emissions.

However, the real advantage of cloud computing in this remote access is mainly reflected in the following aspects: first, the cloud data must be processed using standard metadata format digital information on a variety of sources, in order to create digital content can be retrieved; secondly, the use of standard model and tool enables developers to ignore the number format information. Develop standardized information retrieval system. Therefore, based on open standards, digital information created standard will greatly contribute to the development of information retrieval, including the data content identification, index and specific data used to create domain specific tool for mining use.

#### B. Sharable

In the past ten years, TREC (text retrieval conference) provides a good development opportunity for the sharing of resources. In addition, as a new information service mode of TREC, Crowdsourcing information retrieval has become a successful example of resource sharing. At the same time, the main advantage of cloud computing is embodied in sharing. Relative to the local server to the data independent retrieval, these emerging sharing technologies and sharing the search criteria play an important role in reducing greenhouse gas emissions.

At present, for the sharing of data and search terms, how to use open standards and technology to cooperate and exchange of different models is the focus on the problem.

#### C. Reusable

The reuse of the information retrieval tools and content will be a research focus in the field of information retrieval and digital library research in a long period. For example, the TREC study will be to improve information retrieval, has conducted the related research for a long time reuse retrieval tools and content. The cloud infrastructure is introduced to the relevant research will greatly promote the research progress, this is because a variety of research data set, and each index content and tools can be stored in the cloud, and convenient to use in a variety of development and applied research activities.

Obviously, the cloud computing infrastructure not only facilitates data sharing, but also is beneficial to the research on green retrieval. This is because the researchers no longer need to re-design and access to digital resources and the application has been completed, while also saving energy. In addition, the analysis of these existing indexes, log data, various data, can be re used for further study. It is also conducive to the construction of green retrieval system.

#### D. Green User Behavior

User's behavior can also be considered as a part of information retrieval systems. It determines whether the information retrieval system is green or not to a largely extent. The user's behaviors have a great influence on the energy consumption. Some users who are not using the information retrieval system still keep the information retrieval system available, consuming a large amount of electric energy. Some users use paper to print when the retrieval results needs no printing, consuming a large amount of wood.

Cloud computing can help the users to change behaviors. Users can easily access information through the cloud computing when using mobile electronic devices. The emergence of a large number of various mobile terminals such as iPad e-books is convenient for the user at any time to acquire knowledge. It also reduces paper printing and energy consumption. In addition, cloud computing can also connect a low energy consumption terminal with the cloud terminal to access data acquisition, which can significantly reduce greenhouse gas emissions.

### IV. CONCLUSION

This paper believes that green cloud computing and IT play key roles in reducing the impact of information and communication technologies on the environment. Standardization, sharing, reusability and user behavior of green can help us to build green information retrieval systems and services.

### REFERENCE

- [1] Gobinda Chowdhury. Building Environmentally Sustainable Information Services: A Green is Research Agenda [J]. Journal of the American Society for Information Science and Technology, 2012, 63(4):633—647.
- [2] Jenkin T A, Webster J, McShane L. An Agenda for Green Information Technology and Systems Research [J]. Information and Organization, 2011, 21(1):1—24.
- [3] Liu J, Zhao F, Liu X, He W. Challenges towards Elastic Power Management in Internet Data Centers [C]//Proceedings of the IEEE

- International Conference on Distributed Computer Systems Workshops, 2009: 65—72.
- [4] Cervone H F. An Overview of Virtual and Cloud Computing [J] .OCLC Systems and Services, 2010, 26(3): 162—165.