

***Abstract: Load Balance of the Original Pages Storage in the Distributed Crawling System***

Weizhe Zhang\*, Hongli Zhang, Jinfu Chen  
and Chuanliang Yu

*\*School of Computer Science and Technology, Harbin Institute of Technology,  
Harbin, China.  
wzzhang@hit.edu.cn*

**Abstract**

Load balancing plays a key role in making full and equal usage of the machines in the cluster system, and in shortening system response time. Storage and computation should be fully considered into load balancing. In order to improve the parallelism of crawlers' writing and reduce the response time of the crawlers' writing, this paper proposes a new method on the sensitive division of the channel data storage-compute-sensitive partition. Experiments have proven that this method is the optimal load partition strategy.