

## INTISARI

### IMPLEMENTASI PENCARIAN *FULL TEXT* PADA DATA BERITA ONLINE MENGGUNAKAN SISTEM *INDEXING APACHE SOLR*

Oleh :

Hendi Rusfandi

13/356756/PA/15747

Media Monitoring merupakan aktivitas untuk melakukan *monitoring* terhadap berita yang diterbitkan oleh media *online*. Kegiatan monitoring biasa dilakukan oleh perusahaan penyedia barang/jasa untuk mengetahui bagaimana reaksi pasar terhadap barang/jasa yang ditawarkannya. Untuk membantu kegiatan monitoring, dibangun sebuah aplikasi untuk membantu mengoleksi, mencari, *review*, dan *report* berita. Data berita *online* yang besar membuat beberapa proses pada aplikasi Media Monitoring terganggu. Salah satu masalah yang terjadi yaitu proses pencarian berita pada aplikasi yang lambat. Masalah tersebut terjadi karena tingginya transaksi *read and write* pada basis data aplikasi.

Untuk mengurangi beban kerja basis data, maka pencarian berita perlu dilakukan di luar basis data. Teknik yang dapat digunakan yaitu melakukan *indexing* data berita di luar dari basis data dan mencari berita menggunakan teknik pencarian *full text*. Pada penelitian ini, Apache Solr digunakan sebagai *search server*. Analisis kinerja pencarian *full text* antara MySQL dan Solr dibandingkan dengan melakukan pengujian menggunakan berbagai kata kunci *full text*. Analisis mengenai metode *delta indexing* dan *realtime indexing* pada Solr juga dilakukan dengan simulasi penambahan berita dan *update index* secara berkala serta melakukan pencarian *full text* untuk mengetahui metode *indexing* yang memenuhi kebutuhan aplikasi Media Monitoring.

Hasil pengujian pencarian berita diperoleh hasil bahwa pencarian *full text* pada Solr lebih cepat 2,5 kali lipat dibandingkan dengan MySQL. Metode *realtime indexing* dirasakan lebih cocok untuk diterapkan pada aplikasi Media Monitoring yang memiliki pertumbuhan data yang cepat, membutuhkan ketersediaan data yang *up-to-date*, dan pemakaian *resource hardware* yang optimal.

**Kata kunci :** *indexing*, pencarian *full text*, pencarian berita, MySQL *indexing*, Solr, *realtime indexing*, Media Monitoring

## ABSTRACT

### IMPLEMENTATION OF FULL TEXT SEARCH ON ONLINE NEWS DATA USING INDEXING SYSTEM OF APACHE SOLR

By :

Hendi Rusfandi

13/356756/PA/15747

Media Monitoring is an activity to monitor news published by online media. Monitoring activity is usually done by companies to know publics' reaction over offered goods or service. To assist the monitoring activity, an application is built to help collecting, searching, reviewing, and reporting the news. Big online news data makes some processes of the application disrupted. One of the problems that occur is news searching process on the application that is slowing down. This problem occurs because of high *read and write* transaction on the application's database.

To reduce the database workload, the news searching process should be done outside of the database. Techniques that can be used are to do news indexing outside the database and to search the news using full text search technique. In this research, Apache Solr is used as a search server. Performance analysis of full text search between MySQL and Solr by testing with various full text keyword. An analysis of delta indexing and real-time indexing method is also conducted by simulating the addition of news data to database, periodically update the index and doing full text searching to meet Media Monitoring application's needs.

The news search result produced a result that the full text searching on Solr is 2,5 times faster than MySQL. The real-time indexing method is more suitable to implement on the Media Monitoring application that has rapid data growth, requires up-to-date data availability, and optimal use of hardware resource.

**Keywords :** *indexing, full text search, news search, MySQL indexing, Solr, realtime indexing, Media Monitoring*