GettingNutchRunningWithSocksProxy

Running Nutch with SOCKS proxy

Created by Vinh Khuc Ngoc

Problem

Running Nutch with SOCKS proxy is necessary in some situations, but Apache Http Client 3.0 still doesn't support SOCKS.

Solution

Since SOCKS protocol is supported in JDK 1.4.2, there is a "quick and dirty" way to make Nutch work with SOCKS, that is adding properties **socksProxyHost** and **socksProxyPort** when we call JVM in bin/nutch script:

```
exec "$JAVA" -DsocksProxyHost=<host> -DsocksProxyPort=<port> $JAVA_HEAP_MAX $NUTCH_OPTS -classpath "$CLASSPATH"
$CLASS "$@"
```

Just keep waiting for Apache Http Client 4.0 with SOCKS support.

Here is the issue: http://issues.apache.org/jira/browse/HTTPCLIENT-334