EFFICIENT FILE SYNCHRONIZATION USING WEB SERVICES

Péter Dóbé, dobe@inf.bme.hu
Dr. Imre Szeberényi, szebi@iit.bme.hu
Budapest University of Technology and Economics
Department of Control Engineering and Information Technology

Especially in Grid computing, but perhaps in other environments as well, it may be necessary to update files that do not change substantially between two updates, from a distant location. If the files are large, transferring the whole file through the network each time would not be efficient. The rsync protocol has been created for such a task. Its essence is to transfer only the differences between the old and new version of the file.

The Saleve system developed at BME Department of Control Engineering and Information Technology supports solution of parameter study problems generating large files, in a Grid environment. It uses web services for data transfer, i.e. it sends data over SOAP. This paper presents a plugin for Saleve system, which does optimized file transfer with rsync protocol.

The paper introduces the Saleve system and the Saleve plugin developed for file transfer using rsync protocol. Besides the test implementation we show a generalized solution of web service oriented file synchronization. Finally we show the comparison of the non-optimized and optimized file transfer based on measured results.

The sample implementation written in C utilizes the gSOAP toolkit for creating the web service interface and for data transfer over SOAP, and the librsync library for performing the rsync algorithm. The ClusterGrid environment was the test environment.