

윈도우즈 자동화 스크립트를 이용한 멀티 세션 PDF 변환기 구현

강병준* · 김훈희* · 주상웅* · 심규철* · 강현진* · 김경환* · 정희경**

*주)커뮤 · **배재대학교 컴퓨터공학과

Implementation of Multi Session PDF Converter Using Automated Scripts for Windows

Byeong-Jun Kang* · Hun-Hee Kim* · Sang-Woong Joo* · Kyou-Chul Shim* · Hyun-Jin Kang*
· Kyung-Hwan Kim* · Hoe-Kyung Jeong**

*COMMU CO.,LTD · **Department of Computer Engineering, PaiChai University

E-mail : {junycom, hhkim, soklimt, kcshim, guswlsrkd, khkim}@commu.co.kr, hkjung@pcu.ac.kr

요 약

워드프로세서 파일(아래한글, MS-Office)에 입력된 대량의 데이터를 온·오프라인에서 PDF 파일로 변환 작업을 수행하면 현재의 PDF 변환 솔루션들은 하나의 원본 파일을 윈도우즈 가상 프린터 환경에서 PDF 변환 작업을 수행 한다. 이러한 PDF 변환 작업을 수행하면 원본 파일의 용량에 따라 PDF 변환 속도가 상이하고 대량의 PDF 변환 작업을 수행 할 경우 매우 느린 속도로 수행한다.

본 논문에서는 원본 파일의 PDF 변환 작업을 수행할 때 하나의 윈도우즈 서버에서 터미널 서버를 사용하여 터미널에 멀티 세션을 생성하여 동시에 PDF 변환 작업을 수행 할 수 있는 PDF 변환기에 대하여 연구하였다.

ABSTRACT

When you convert word processor files with large amounts of input data to PDF files in on/off line, current PDF conversion solutions do one of the source files to perform the conversion task in the Windows PDF virtual printer environment. Depending on the size of the original PDF file conversion speed to convert PDF to do a large amount of different and if you are performing a very slow speed. When you convert files to PDF in this sequence, depending on the capacity of the original file, PDF conversion rate is different, if automating a mass conversion to the PDF, the speed is very slow.

In this paper, the PDF conversion which can use the Windows Terminal Server from one Windows Server when performing a PDF conversion processing of the source file, to create a multi-session Terminal, to perform the PDF conversion process at the same time was studied.

키워드

PDF 변환, 멀티 세션, 멀티 세션 PDF 변환

I. 서 론

일반적으로 윈도우즈 서버에서 워드프로세서로 작성된 문서를 PDF 변환 작업을 수행하면 사용자가 원본 파일을 워드프로세서로 파일을 열어서 워드프로세서에서 제공하는 PDF 변환 메뉴를 사용하여 PDF 변환 작업을 수행하거나 또는 가상의 PDF 변환 프린터를 사용하여 PDF 파일로 인쇄

작업을 수행하여 PDF 변환 작업을 수행한다.

1개의 파일을 PDF 변환 작업을 수행 할 때는 큰 문제가 없이 변환 작업을 수행 할 수 있으나 온라인상에서 웹으로 원본 파일을 접수 받아 PDF 변환 작업을 수행하거나 또는 대량의 원본 파일을 PDF 변환 작업을 수행할 경우 많은 시간이 소요되기 때문에 PDF 변환 서버를 사용하여 PDF 변환 작업을 수행하고 있다[1].

하지만 현재 개발된 PDF 변환 서버의 경우 하나의 서버에서 워드프로세서 인스턴스를 1~4개 정도의 워드프로세서 인스턴스를 사용하여 변환 작업을 수행하여 그 속도가 매우 느리고 온라인 상에서 많은 사용자가 동시에 원본 파일을 업로드 할 경우 PDF 변환 작업에 많은 시간이 소요되어 사용자가 매우 불편하게 시스템 사용하고 있다[2]. 이런 문제점을 개선하기 위하여 본 논문에서는 윈도우 서버를 사용하여 윈도우 터미널에 설정된 세션을 동시에 접속하여 워드프로세서의 파일을 다수 실행하여 동시에 PDF 변환 작업을 수행할 수 있도록 하여 PDF 변환 서버의 성능을 향상시키는 방법을 연구하였다. 그림 1은 Virtual PDF Converter를 통하여 PDF 변환 작업을 수행하는 예 방법으로 사용자가 수동으로 조작하여 작업을 진행하는 방법이다.

본 논문에서는 PDF 변환 작업을 과정을 윈도우즈의 자동화 스크립트를 이용하여 PDF 변환 작업을 자동화하여 사용자가 조작하지 않아도 PDF 변환 작업이 수행 될 수 있도록 하였다.

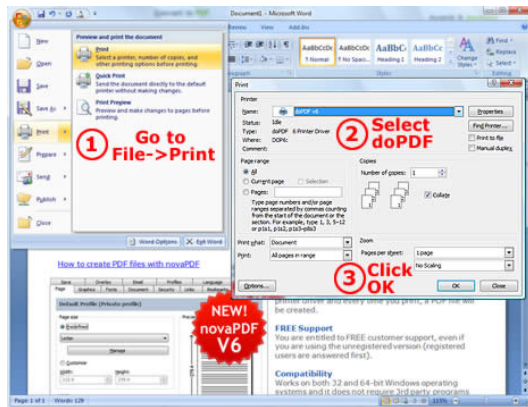


그림 1. Virtual PDF Converter의 실행 화면

II. 관련연구

2.1 멀티 윈도우 터미널

본 논문에서는 윈도우즈 서버의 RDP 프로토콜을 사용하여 윈도우즈 서버에 동시에 사용자 세션을 설정한 갯수 만큼 생성 할 수 있는 멀티 윈도우 터미널을 사용할 수 있도록 멀티 윈도우즈 터미널 프로그램을 개발하고 여러 개의 터미널이 동시에 실행 될 때 디스패치 서버에서 클라이언터에 작업을 할당할 수 있도록 클라이언트 프로그램을 개발하여 자동으로 실행할 수 있도록 구현한다[3].

그림 2는 멀티 윈도우즈 터미널을 윈도우즈 서버의 라이선스 방식에 따라 사용자가 필요한 만큼 구매가 가능하며 멀티 윈도우즈 터미널에서 그 개수만큼 접속할 수 있도록 설정하는 방식을 사용하여 필요한 만큼 사용자 세션을 생성 할 수 있다.

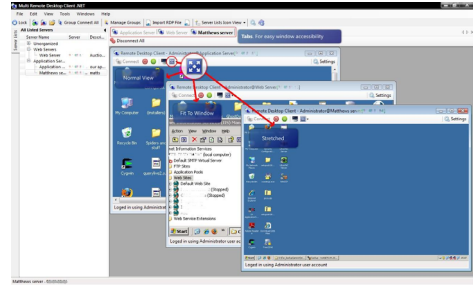


그림 2. 멀티 윈도우즈 터미널의 실행 화면

2.2 Virtual PDF Printer

본 논문에서는 가상의 PDF Printer 기능을 구현하여 다중 사용자가 동시에 프린터 출력을 할 수 있도록 구현하고 PDF Printer의 기능으로 PDF 암호화, 다중언어 지원, 페이지 사이즈 조정, 페이지 스케일링, PDF 파일 압축 및 비압축, subset 삽입, 폰트 저장, 디지털 사인, 북마크, 텍스트/이미지 워터마크 등의 기능을 구현하여 사용자가 다양한 옵션을 선택 할 수 있도록 하여 PDF 변환 작업을 수행 할 수 있도록 구현하였다.

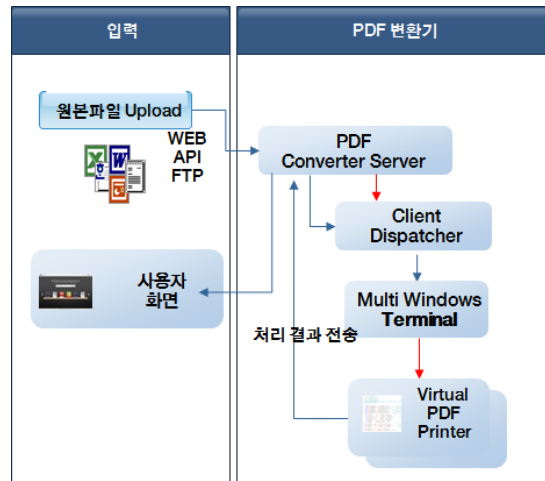


그림 3. PDF 변환 시스템 구성도

그림 3의 PDF 변환 시스템 구성도의 Virtual PDF Printer의 기능 구현에 가장 중요한 기능은 하나의 사용자가 프린터를 사용할 경우 다른 사용자의 작업 프린터 큐에 저장되어 이전 사용자의 작업이 끝나면 다음 작업을 수행하는 프린터의 고유 기능을 수정하여 서로 다른 사용자의 경우 동시에 프린터 큐를 사용할 수 있도록 기능을 구현하였다.

2.3 윈도우즈 자동화 스크립트

본 논문에서는 멀티 PDF 변환 기능을 사용하기 위하여 메인의 서버 기능의 어플리케이션과 PDF 변환 작업을 수행하는 클라이언트로 구성하였다. 서버는 사용자가 변환 작업을 요구하면 라운드로빈 방식으로 각각의 클라이언트에게 작업을 할당하고 할당된 작업이 종료 되면 PDF 변환

파일의 정보와 PDF 파일을 사용자에게 제공한다.

서버는 클라이언트의 헬스 체크를 주기적으로 수행하여 클라이언트 이상 유무를 확인하고 만약 클라이언트가 중지 되었거나 작업이 행이 걸리면 클라이언트를 초기화 하는 기능을 구현하여 지속적인 서비스가 가능하도록 구현하였다.

클라이언트는 서버로부터 PDF 변환 작업을 수신하면 클라이언트는 서버에서 제공하는 원본 파일의 확장자를 확인하여 해당 워드프로세서를 사용하여 원본 파일을 읽어 PDF 변환 작업을 수행 할 수 있다. 윈도우즈 자동화 스크립트는 해당 워드프로세서의 기능을 확인하여 Virtual PDF Printer로 원본 파일을 출력 할 수 있도록 스크립트를 수행한다. 만약 오류가 발생하면 오류 코드를 서버에 전송하여 사용자에게 알릴 수 있도록 하여야 한다. 클라이언트는 작업이 완료되면 서버에게 작업 완료 메시지를 전송하고 PDF 변환된 파일의 정보를 서버에 제공 한다. 서버는 사용자에게 작업의 이상유무와 PDF 변환된 파일을 제공하여 작업을 마무리 한다.

Virtual PDF Printer 기능으로 PDF 암호화, 다중언어 지원, 페이지 사이즈 조정, 페이지 스케일링, PDF 파일 압축 및 비압축, subset 삽입, 폰트 저장, 디지털 사인, 북마크, 텍스트/이미지 워터마크 등의 기능을 구현하여 사용자가 다양한 옵션을 선택 할 수 있도록 하여 PDF 변환 작업을 수행 할 수 있도록 구현하였다.

Virtual PDF Printer의 기능 구현에 가장 중요한 기능은 하나의 사용자가 프린터를 사용할 경우 다른 사용자의 작업 프린터 큐에 저장되어 이전 사용자의 작업이 끝나면 다음 작업을 수행하는 프린터의 고유 기능을 수정하여 서로 다른 사용자의 경우 동시에 프린터 큐를 사용할 수 있도록 하였다.

III. 결 론

온·오프라인 상에서 워드프로세서의 원본 파일을 PDF 변환 작업을 수행할 경우 현재 개발된 프로그램들은 동시에 하나의 작업을 수행하여 PDF 변환 작업이 매우 느리고 하나의 PDF 변환 작업에 문제가 발생할 경우 서버 전체가 중지되는 현상이 자주 발생하여 관리자가 서버의 이상 유무를 확인하여 서버를 초기화 하거나 다시 가동하여 사용하는 작업을 수행하고 있다.

본 논문에는 이런 단점을 보완하여 멀티 세션 PDF 변환 작업을 수행할 경우 윈도우즈 서버의 성능을 100% 사용할 수 있도록 설정할 수 있어서 동시에 PDF 변환 작업을 수행 할 수 있어서 그 속도로 매우 빠르고 PDF 변환 작업에 문제가 발생할 경우 서버 모듈에서 클라이언트 모듈을 초기화하는 방식으로 서버의 가용성을 향상 시킬 수 있다.

감사의 글

본 논문은 중소기업청에서 시행한 산학연 공동기술개발사업의 결과입니다.

참고문헌

- [1] <http://www.apache.org>
- [2] 전홍우, 정창후, 송사광, 최윤수, 최성필, 성원경, “PDF 문서로부터 의미 정보 인식 기반 XML 문서로의 변환,” 한국인터넷정보학회, Vol12, No1, pp.399-400, 2011.6
- [3] Akhtar, Waseem, “XSPARQL: Traveling between the XML and RDF worlds-and avoiding the XSLT pilgrimage.” The Semantic Web: Research and Applications. Springer Berlin Heidelberg, pp.432-447, 2008