

White Paper

## Server-Based PDF Creation: Basics

## Table of Contents

PDF Format.....	2
Description .....	2
Advantages of the File Format.....	2
PDF Creation: Desktop or Server Solution? .....	3
Advantages of Server Solutions.....	3
Requirements to the PDF Server.....	5
Alternative Ways of Server-Based PDF Production.....	5
Making of a conversion task .....	6
Processing of a conversion task.....	6
Distribution of the PDF Files.....	7
Example Application: PDF Render Center .....	7
The most important functions.....	8

## PDF FORMAT

Nowadays nearly every computer user knows about the PDF format. E-mail messages with an attachment in a form of a PDF file often displace the paper documents. For example, it is usual, that a person receives an invoice for a flight booking over the Internet within seconds by e-mail as a PDF document, or that telephone companies send the electronic versions of monthly bills as PDF files and sometimes even require an additional fee for the sending of the "original" paper bill.

### Description

The Portable Document Format (PDF) by Adobe® is a file format, with which reading of documents can occur regardless of the original application as well as from the hardware and the operating system used for creation of documents. In addition to the static design of pages a PDF document can also contain interactive elements, which are possible only in electronic form: PDF supports different types of so-called annotations as, for example, text comments, hyperlinks, modifications or selections, file attachments, sound data and video. A PDF document can define (within certain bounds) its own user interface while displaying. According to the settings, inputs from the keyboard or mouse can perform actions, which are exactly described by PDF objects. The PDF document can also contain interactive form fields, which can be filled by the user. Besides, export of field contents or import of these contents from other applications is possible. In addition to the design entries, the PDF documents often contain also the identification marks and details to the logical structure, which allow searching, modification or export of contents for further usage in another place. Finally, the scope of usage of the ready PDF file can be controlled over the security settings, which can be defined at the creation of the file.

### Advantages of the File Format

At the spreading of electronic documents and the accompanying replacement of traditional paper documents with them, the number of digital document formats grows. However, their functions and advantages also expand. Every person, organization or institution and every enterprise after the decision to perform - entire or partial - digital document management, should first make a decision on the data format, which would be used in the future. In this respect it is advised by many experts to include the PDF format in consideration. What are the advantages?

10 years ago talks on the openness of the PDF format and an extensively documented specification, and then on the standardization of PDF appeared. Other arguments for PDF were the flexibility of the technical requirements and the possible functionalities. All this is still true and important even today, but, in addition, PDF has matured meanwhile into a standard. The file format is now established as a worldwide standard, since its official assignment for the ISO standard by the *International Organization for Standardization* in July 2008.

The PDF format was already very popular without this "official initiation" and was well established, which favors its progressive spreading and forces its further propagation. Nowadays one can hardly imagine a computer without PDF software installed: at least for reading and printing nearly every computer user uses the PDF format. And this independently of the installed operating system. The number in software products for creation, modification and reading of PDF will further grow. Thereby the most important problem of digital documents can be solved - a trouble-free transferability: the contents of a document can be at least read on every computer because of the standardization of the format, and this also long time after its creation.

As files can be embedded in PDF documents in any format, it is easier to manage images of different types in a document workflow: it is much easier to administer and to keep only one file for a process than several files connected with each other. Besides, XML metadata can be used in a PDF document and can be connected with different objects in the document. This extends the possibilities for the integration of non-visual data in a PDF document.

A lot of attention was paid to the security of the document contents in the PDF format. At the work with PDF documents the encryption algorithm can be set by the author of the document and by the software developer, who integrates PDF into his applications, to protect data against unauthorized access. Besides, the PDF format allows integration of digital signatures right into the PDF file.

## **PDF CREATION: DESKTOP OR SERVER SOLUTION?**

Who currently looks around the software market for a good solution for converting of documents in PDF files, will come upon a relatively large quantity of products. Besides, the number of products with which the PDF production can occur only on the computer, where the PDF software itself is installed (desktop products) obviously forms the majority. As a rule they are cheaper than server solutions. With latter a central computer or central software, the server, takes over the conversion job for the decentralized clients. Server-based solutions are first of all more efficient at the production, processing and distribution of large amounts of PDF files for large user groups. Concerning the quality of the created PDF files (trustworthy representation of the source file) server solutions do not differ as a rule from the desktop versions. Their price is higher in compare with desktop pendants of the same manufacturer because of their special usage purpose.

### **Advantages of Server Solutions**

As it was mentioned, server solutions have some advantages at work with large amounts of documents, which are to be converted regularly and/or with a big user group. The essential reason for that is, on the one hand, the usually easier network-wide setup of the software. On the other hand, the functions refer to the central

administration of the PDF production and the **possibilities for comfortable distribution of the created files in a network** for a server solution.

For most server-based PDF products the **installation of software on every single personal computer is not necessary**. The PDF software should be installed merely on the server computer and then the access to the software for the workstations will be arranged by sharing it on the network. However, there are also solutions, for which PDF software or at least PDF printer driver should be installed on every personal computer.

Even more important than the comfortable installation is the ability of server solutions **to set and control the parameters of the PDF production centrally**. In this manner the PDF settings can be standardized for the users, for example, in processing rights, access protection of the created files or file size. Often the administrator can also define the user group profiles, where the PDF parameters can be arranged and saved according to the specific purpose of the PDF file (for example, use for the display in the Internet, for the printing preliminary stage, for the draft printing on paper, with or without a possibility of processing/commenting). The user should or can select only the suitable profile, without changing single PDF settings. Some software manufacturers choose another way to provide the predefined selections of PDF parameters: either an own PDF printer is set up for every selection, or different e-mail addresses or directories are arranged for incoming conversion tasks.

If for every selection of parameters an own printer will be set up, it will be quicker to find it in the well-known printing dialog of an application than a user profile in another dialog box.

The providers of server-based PDF software also partially differ according to whether a single user can change the PDF settings in his personal computer by himself or only apply the defaults and profiles set by the administrator.

If it is necessary, all documents can be also displayed according to the **unified document templates and layouts** (with a company logo, contact data or other defaults) thanks to the central conversion. Other advantages of server solutions are:

- Centralized update and license management
- For the **application** in existing or newly introduced **document management** systems (DMS) efficient server-based solutions usually offer useful functions (for example, for the distribution of documents) and interfaces (for example, web services) already included into the base packaging. To be able to control the cooperation with DMS in the optimal way, providers of PDF servers offer flexible methods of comprehensive adaptation of their product to the individual requirements of the customer.

## Requirements to the PDF Server

Basically the server-based solutions should fulfill the same requirements as the desktop solutions, namely, according to the respective tasks, they should contain an optimal mix of the following properties: **few conversion errors** and **high conversion speed**, **extensive functions** for the definition of the PDF settings, **low file size** and last but not least a **good representation quality**, also at the output on paper.

However, the following requirements to a good product arise from the basic technical advantages of the PDF server solutions: a release should require **no installation of software on every workstation** and the PDF settings should be **administrable** comfortably and **centrally**. Third, a **flexible and efficient distribution** of the created PDF files to the addressees is important (to the authors of the conversion task and other involved people).

Besides, the **offered price and license model** matter: flexible, modular products are preferable, where the customer pays only for the functions he really needs. In addition, the price should also depend on the number of the solution users in a network and on the data traffic amount (number of converted documents in a certain period).

Good server solutions provide **multithreading**, i.e. a synchronous processing of several operations within an application. At that the processor performance can be improved and the capacity of multiprocessor computers can be heightened. As a result the production and distribution of a large quantity of PDF files runs faster than on software without a multithreading ability.

From the technical point of view a server solution should consist of the following **components** in addition: a system service, which controls the synchronous processing of documents, a component(s) for the conversion of the documents in PDF as well as an administration program for the definition of the settings and for the control of conversion and distribution.

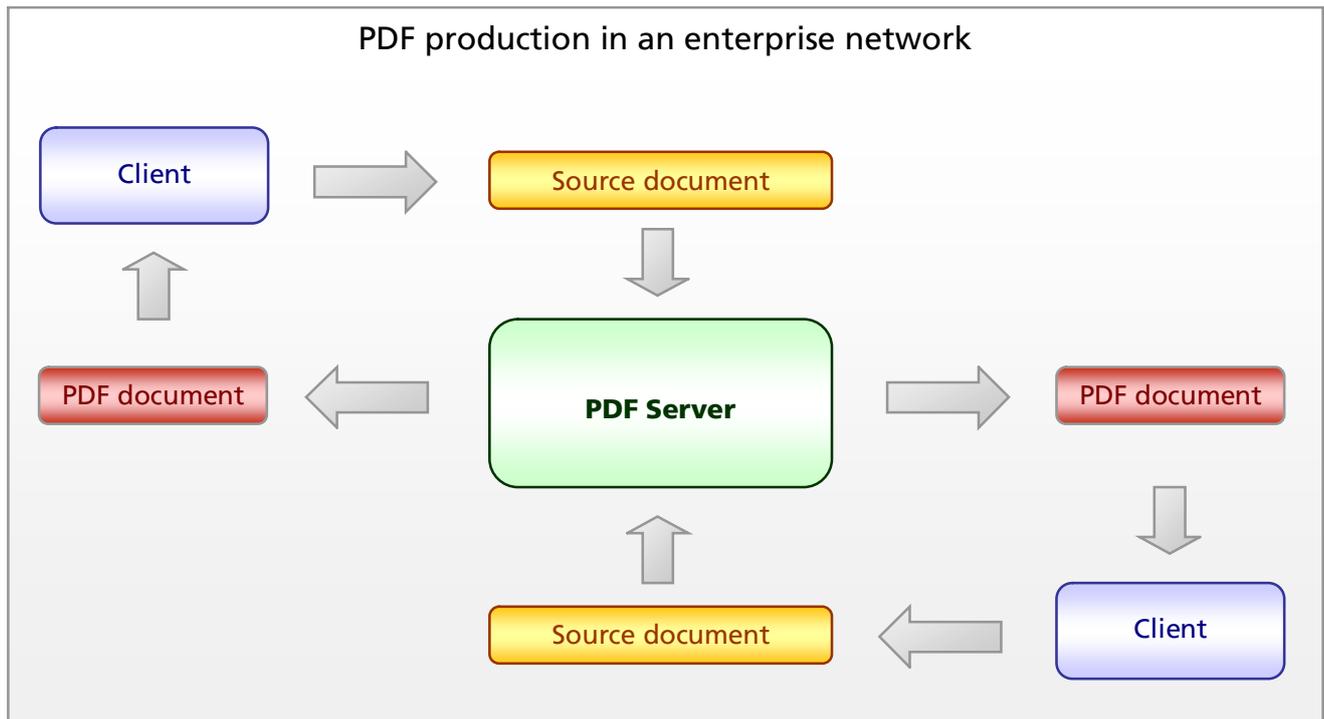
And finally the **system requirements** for the PDF server solutions should be mentioned. As a rule the programs can be run not only in the "real" server operating systems like Windows Server 2003 and 2008, but also in Windows XP (Professional) or Vista. In addition, there are also solutions for Linux or Apple Macintosh. Client-sided web-based solutions also exist; they can be run on Windows, Linux and Mac computers.

## Alternative Ways of Server-Based PDF Production

There are different methods of making a conversion task by a client as well as of its processing on a server:

## Making of a conversion task

To create a PDF file in a server solution, the user should act in most cases similarly as with the widespread desktop-based products: from an application, in which a PDF file should be created, it should be "printed" on a virtual printer from the personal computer. This PDF printer is installed at the installation of the server solution as an additional printer (driver) and is shared on the network afterwards. It displays the result of printing process not on paper, but as a PDF file. The printing or conversion task can be given by using a PDF printer driver for any applications and source formats, provided that the original application has a printing function.



The most popular server solutions use this way. However, there are also known products where the conversion tasks are sent as e-mails to a certain address and the source file is uploaded over a web browser or where the original file is put into a shared network folder and is automatically converted into a PDF file. Also combinations of these alternatives exist. At the production of PDF without a printer driver the number of convertible file formats, as a rule is limited to those, which the respective manufacturer makes available in his software.

## Processing of a conversion task

In the solutions, which are based on the use of a virtual printer, the first conversion step can occur on a personal computer. From the source file or from the data stream a (temporary) metafile is generated and sent to the server. Then the production of a PDF file occurs.

In this connection it is important, which way the solutions use to embed the character fonts: depending on whether they are already embedded on a personal computer or only on a server, all relevant character fonts should also be present on a server or not. If the imbedding occurs on a personal computer, the server does not need the used character fonts, at the imbedding on a server the same character fonts should be present there accordingly as on the personal computers.

PDF recognizes the following 14 standard character fonts: Times-Roman, Helvetica, Courier, Symbol, Times-Bold, Helvetica-Bold, Courier-Bold, ZapfDingbats, Times-Italic, Helvetica-Oblique, Courier-Oblique, Times-BoldItalic, Helvetica-BoldOblique, Courier-BoldOblique. Provided the used fonts cannot be embedded as they are not installed on a personal computer or on a server, the used fonts are replaced with similar standard fonts, but it leads as a rule to undesirable results - overlapping letters and/or wrong letter size.

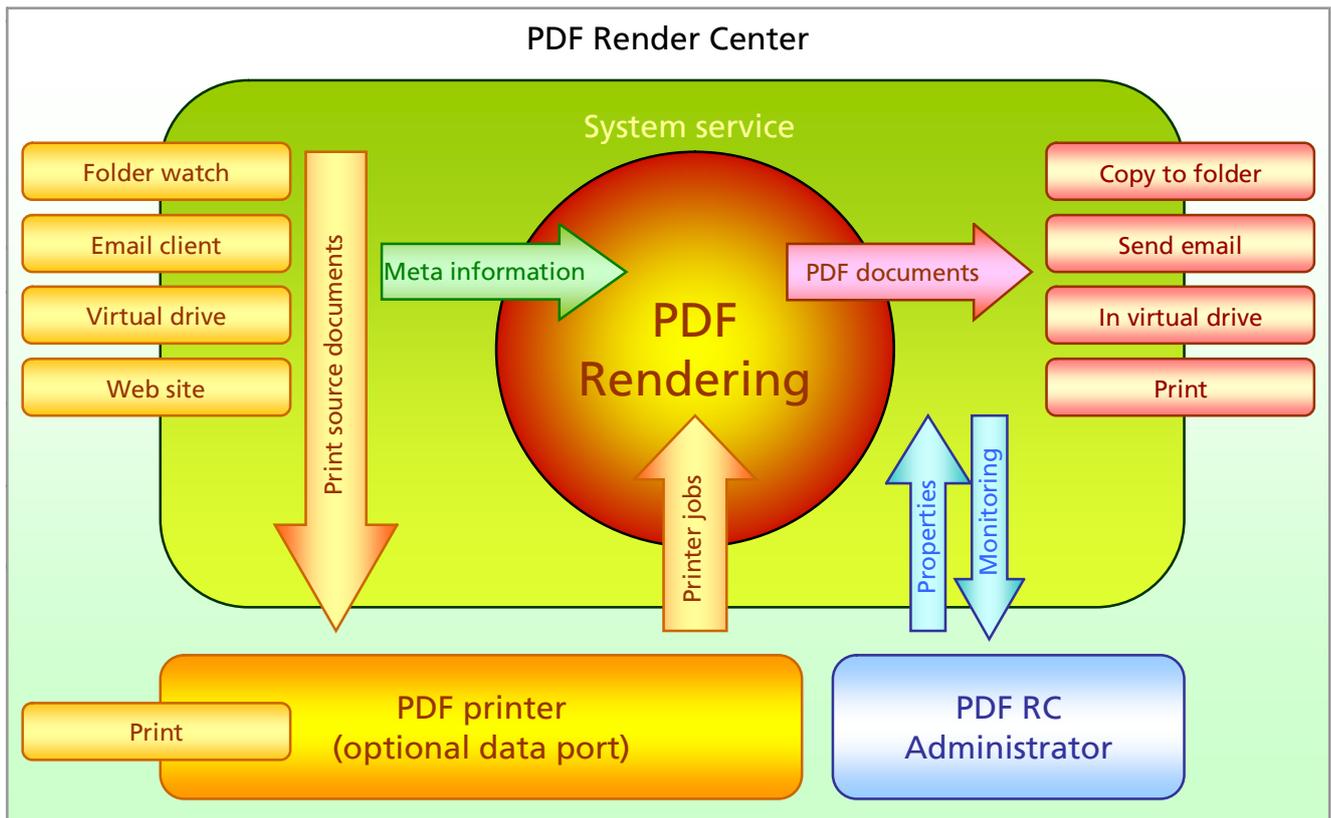
If a document is converted on a computer, where also the source file was created, the PDF production concerning character fonts occurs as a rule without difficulties, provided that all fonts are installed there locally and correctly. Therefore, a correct imbedding of the available fonts usually occurs easily for the desktop solutions. As for the server solutions, it occurs well on a personal computer, like in most PDF printer driver solutions. In the solutions, where the source file should be sent by e-mail to the server and uploaded to a directory, the character fonts should be also installed on a server. This can mean raised control expenditure, while on the personal computers there is a possibility to install own applications and/or character fonts.

## **DISTRIBUTION OF THE PDF FILES**

As in server solutions the production of PDF files from the point of view of the clients is "remote", the ways offered by the products matter for the distribution of the ready files. The following distribution options are possible: storage in a shared network folder, in a local directory on a personal computer, on virtual disk drives, sending by e-mail, web download or download from a FTP server. If the documents should be also printed on paper, it is possible over a network printer or local printers.

## **EXAMPLE APPLICATION: PDF RENDER CENTER**

A possible server solution is the [PDF Render Center](#) by soft Xpansion. A system service of Windows forms the core of the application; in addition a virtual printer (PDF printer) and an administration program are for the settings. The solution can be easily integrated into the existing document management systems. It is applicable in enterprises of all sizes and provides the avoidance of paper documents as well as the automatic administration of several thousand documents in a uniform format. PDF Render Center like all server-based solutions allows a more efficient processing of a large quantity of documents and at the same time minimizes or even removes costs for



- PDF and PDF/A production from any printable formats: for example, Microsoft Office format, formatted (RTF) and unformatted text, HTML files and websites, images (TIFF, JPEG, PNG, GIF, BMP)
- Import of Microsoft Office meta data: references, tables of contents (TOC), form fields and control elements, comments, from Word, Excel, PowerPoint documents in PDF documents
- Conversion of PDF in PDF/A-1a and 1b
- Export of PDF documents in ASCII-Text (ANSI or Unicode) or images (TIFF, JPEG, PNG), including multipage TIFFs, as well as CMYK

Basically PDF Render Center can convert any data formats into PDF. The only requirement is that on the server, where the PDF Render Center is installed, also the software should be set up, which can print this format on a PDF printer. So, for example, Microsoft Office should be installed on a server for the conversion of the Microsoft Office formats. The application-specific requirements are:

- Import of meta data from Microsoft Office documents is supported from Office 2000, but it is recommended to use a version starting from Office 2003
- Import of HTML files requires an Internet Explorer 6.0, but it is recommended to use version 7.0 or higher

In a network several communication variants between a personal computer and the PDF Render Center are possible, for example, Internet Explorer or Outlook. All variants are based on the standard software and require no additional software and no special EDP knowledge.

