



ENTERPRISE PDF RENDITION

The right approach to business-wide
PDF generation



WHITE PAPER

CONTENTS

The status quo in businesses and institutions	3
Enterprise rendition – what is it?	3
Technology / Requirements	4
The IT administrator’s perspective	5
The solution: Foxit Document Transformation Services	5
A look under the hood	5
Conclusion	8

THE STATUS QUO IN BUSINESSES AND INSTITUTIONS

Businesses often use a wide range of tools to convert files to PDF format or to create PDF files. But few have a clear picture of which solutions are installed in which department – or even at which individual employees' workstations. The mailroom converts scanned documents to PDF/A-2u, emails are turned into PDF/A-3b files and SAP journals become PDF 1.4 files... the list goes on and on, limited only by the size of the business itself. In addition to these planned scenarios, there are also unplanned ones. A typical example is the "Save as" function offered by an increasing number of software applications. As convenient as it is, it also results in this veritable menagerie of PDF files becoming larger and larger, as few employees take the time to consider exactly which boxes to check in the Settings menu. Last but not least, there are the employees who bring their own freeware into the business in order to create PDFs. This eventually leads to a more or less total loss of control over the ways in which businesses create PDF documents.

This white paper describes an approach to PDF creation which standardizes quality levels and adheres to all relevant compliance guidelines.

ENTERPRISE RENDITION – WHAT IS IT?

- How can a standardized high level of PDF quality be achieved across all systems?
- How can document-related guidelines (document policies), such as standardized storage of documents in PDF/A format, be implemented business-wide?
- How can regulatory requirements be met in areas such as process documentation, principles of standardized management and storage of books, diagrams and documents in digital form, or data access (GoBD guidelines)?
- How can an audit trail or a compliance check prove who processed a given document?
- How can all document-related processes be fully logged for auditing purposes?

The term "enterprise rendition" has been gaining prominence for some time now. Consider the following definition, originally sourced from the German-language Wikipedia: "Rendering describes the creation of a graphic from a sketch, a model or raw data such as geodata. Based on this, the term as used in the field of information technology describes the final creation of a (...) draft raw design which is rendered into a final image." Carrying this concept over into the world of document management, the term "enterprise rendition" describes the business-wide "rendering" of documents.

The goal of enterprise rendition is to display documents in a standardized fashion irrespective of the IT system used to generate them within the business, their original format, or their original source. The term encompasses the conversion, generation, processing and validation of PDFs. Other output formats are also theoretically possible.

Enterprise rendition enables controlled document creation. It is advisable to consolidate and centralize all document generation processes within a single shared service (“document conversion as a service”). This hands over control to the experts, lets users operate specific predefined conversion processes, and eliminates technical uncertainties. This – almost incidentally – results in the replacement of the existing menagerie of PDF creation and conversion tools with a single central component, reducing maintenance requirements and licensing costs. According to a study titled “Shared Services in Government” by the consulting firm A.T. Kearney, public institutions are achieving savings of over 20% in their IT systems and processing work through this approach. The result is a standardized data center architecture which controls the distribution of processes across the server infrastructure, ensuring that users get their processed documents back within a reasonable timeframe.

Such a system can theoretically be used across all departments in a business which convert, create or process documents: enterprise resource planning, SAP, archives, databases, input management, specialist applications, workflows, email clients, SharePoint, output management, document composing, CRM systems, ECM systems, EAI, DMS systems, BPM, big data, mobile applications... Consequently, the shared service can be addressed by the client, the server and by back-end or cloud systems.

TECHNOLOGY / REQUIREMENTS

It is advisable to use web services for enterprise rendition. They are easy to integrate into an existing IT infrastructure, since they are universally addressable and can be accessed using any computer or other end device such as a smartphone or tablet, regardless of the programming language used. In addition, they can be easily used on any platform thanks to a self-descriptive interface known as the Web Services Description Language (WSDL). However, it is important to take the amount of local reference data (the payload) into account.

The standardized data center architecture must be dynamically scalable, redistributing jobs (load balancing) and adding or removing resources as needed without any downtime. Naturally, it is also important to guarantee high operating stability without any outages.

A virtualized data center, sometimes also known as the “private cloud”, can be expanded if required using an elastic cloud service such as Amazon EC2. This means that even the highest of peak loads can be handled without any issues. Make sure you pay attention to data security!

In order for enterprise rendition to function smoothly, the IT department will set up a range of workflows – to convert digitized documents in the mailroom, for example, or for email archiving. These workflows are started by departments or devices as needed and can run in parallel with one another. An extensive tiered priority management system is essential for this to work, in order to ensure that nobody has to wait for their job to be completed. For example, if a large number of documents needs to be converted to PDF, the pages of these documents should be distributed across multiple instances. The system should also log who ordered which jobs, in order to correctly assign costs.

THE IT ADMINISTRATOR'S PERSPECTIVE

The IT administrator uses the web management interface to configure and monitor the system. This interface is included in existing standard software such as Cacti (to collect measuring data) or Nagios (for monitoring purposes). The interface guarantees robust, straightforward operation. If needed – perhaps in case of a network outage or corrupted files – the administrator can intervene to assist. The system itself, however, should also be able to monitor itself in order to guarantee performance and even load distribution.

THE SOLUTION: FOXIT DOCUMENT TRANSFORMATION SERVICES

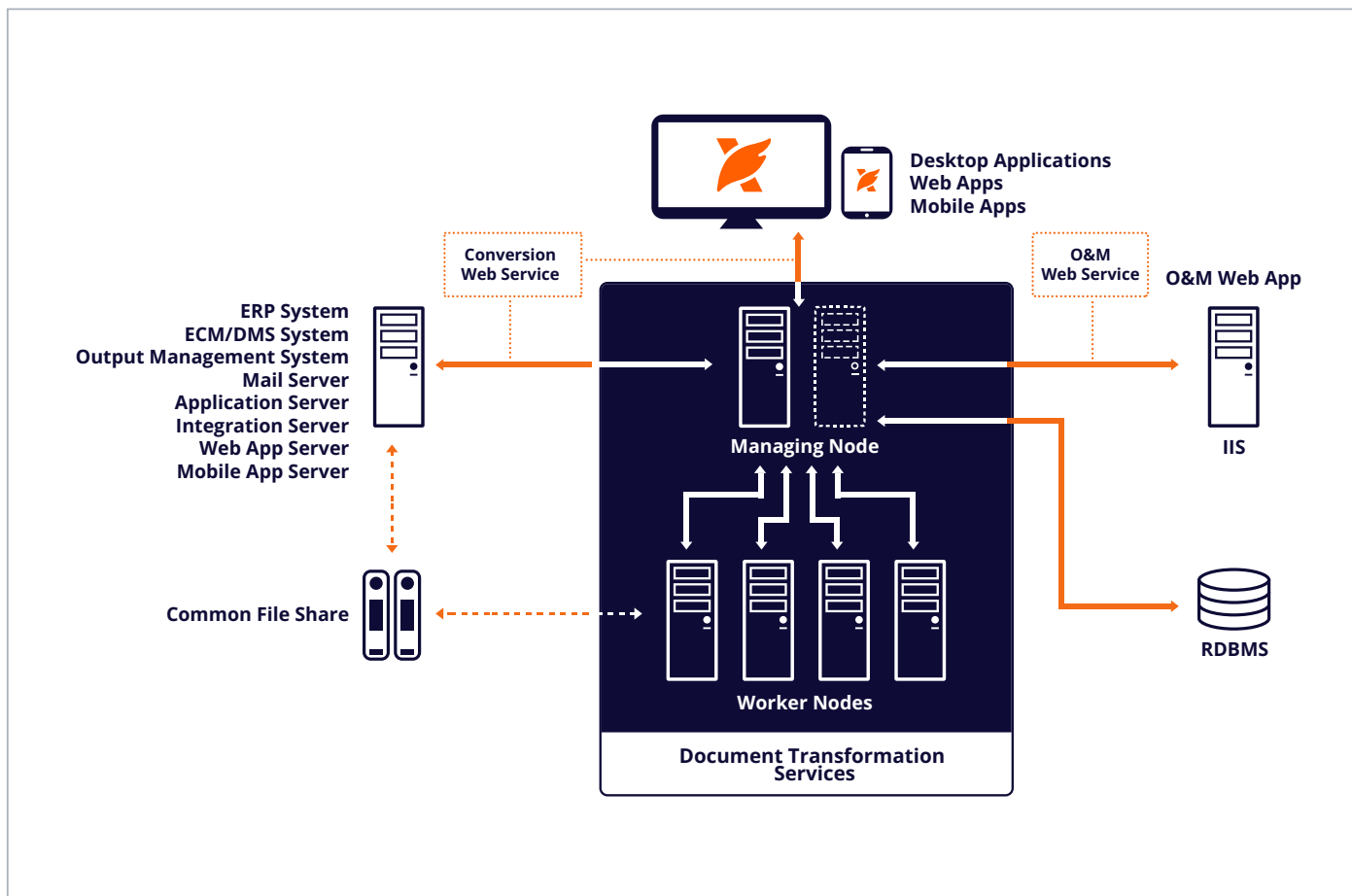
Foxit has responded to this need by developing a standalone product, the Document Transformation Services. This is a platform for implementing document conversion processes. Businesses can use it to ensure that PDFs are converted, generated, processed and validated at standardized quality levels. Other output formats are also possible. The process runs independently of any IT subsystems and of the source or format of the file to be processed. The software, which can be easily operated using a web management console, contains all conversion modules that may be required, all of which are based on Foxit technology. Additional modules can be easily integrated thanks to the solution's open interfaces. As many servers as required can be added for job distribution (load balancing) purposes in order to handle peak loads. All components can be operated redundantly.

A LOOK UNDER THE HOOD

A Foxit Document Transformation Services installation consists of a minimum of two nodes. One node can be considered the same as a (physical or virtual) server. The architecture of the Document Transformation Services distinguishes between two types of nodes: one or more managing nodes, and one or more worker nodes.

The managing node is responsible for providing the web conversion service and handles requests with their associated work items. It also maintains a connection with the database and takes care of all aspects of operations, monitoring and control.

A worker node is a “working server”. It has little awareness of its environment or of the Document Transformation Services array as a whole. Worker nodes are assigned their work items directly by the managing node, which is designed as an interface between the individual worker nodes and the outside world.



Each conversion job begins with a request sent to the conversion web service. A request contains the descriptions of the inputs and outputs, as well as the selected strategy and the properties supplied by the client. Each strategy defines which process to run, with which standard properties. These strategies are stored within the managing node and always relate to one assembly line type.

An assembly line type consists of one or more stations in a clearly defined sequence. It can contain both standard and custom stations. Each station specializes in one sub-function of the conversion process.

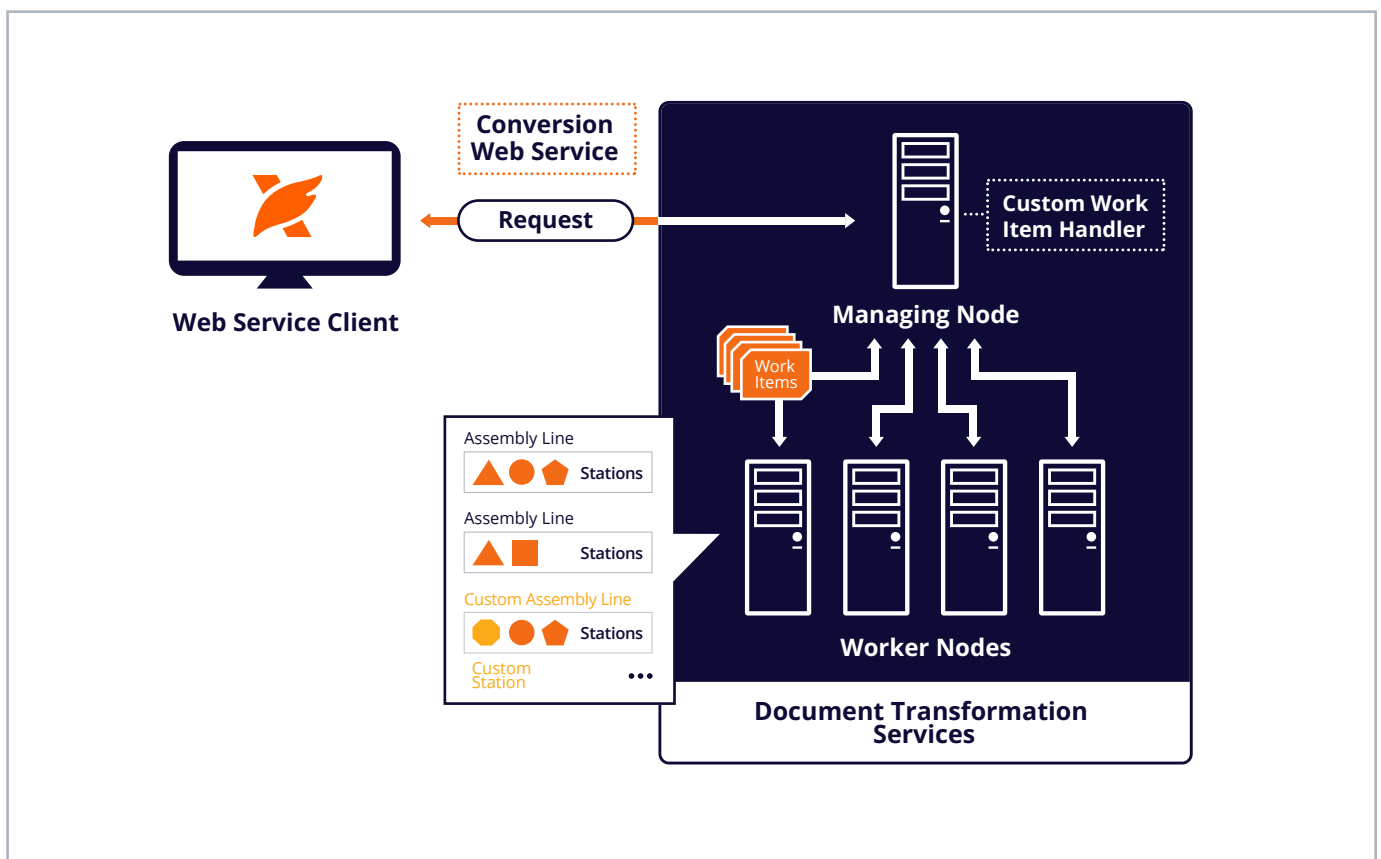
Identical functions may be required across multiple assembly line types; stations can be used to define these functions just once and in a single location. This increases output quality and avoids redundant code. Custom stations are generally built by customers or partners. They support or provide functions which are not available in the standard version of the Document Transformation Services.

Custom stations are made available to the Document Transformation Services in the form of a .Net assembly using the custom station plugin interface. In order for the stations to make use of them, they must first be integrated into a custom assembly line type. A request will then include the appropriate custom strategy (based on this custom assembly line type).

Custom stations implement the station interface and can therefore also be used together with standard stations in a custom assembly line type. This allows you to model highly efficient and powerful assembly lines without getting lost in the details.

The managing node receives the request and generates one or more work items in response. These work items are then processed in parallel by the worker nodes. At the end of the processing stage, the managing node then receives all work items again and returns the appropriate response. The managing node also has the option of executing external code using a plug in interface. This allows you to access the work item distribution options, for example.

A work item is created within a managing node and is generally sent to a worker node, which then returns it back to the managing node after processing. Work items are not sent back and forth between worker nodes. All work items can therefore be freely planned based on the current worker node availability and work item priority.



CONCLUSION:

The Foxit Document Transformation Services forms the basis for a stable, highly scalable document conversion process infrastructure which ensures consistently high quality. PDF generation and processing therefore becomes a centralized service. Businesses can implement document-related guidelines (document policies), meet regulatory requirements with corresponding process documentation, audit trails or compliance checks, and can access all data within a system in order to account for service costs with clients, branch offices or individual departments. In providing all of this, the Foxit Document Transformation Services replaces countless decentralized PDF export and conversion solutions.