A PPML/T Based Variable Data Printing Engine

Galo Gimenez and Luca Chiarabini
Hewlett Packard
Sant Cugat del Valles, Spain

Abstract

The introduction during the last years of a new generation of digital presses that allow the production of high quality, sophisticated, individualized documents, has catalyzed the investment in new variable data printing solutions. This paper describes a PPML/T based variable data printing engine and its application in the Retail Printing System (RPS), an end-to-end solution for real time targeted marketing (RTTM) that produces personalized point of sale materials that can be fulfilled in multiple channels, like high volume digital presses or flat panel displays. The paper discusses the state of the art PPML/T templates, the interfaces and internal architecture of the VDP engine designed to facilitate and drastically reduce the costs of the development of new VDP applications.

1 Introduction

The field of variable data printing (VDP) applications is on rapid expansion and new challenging applications are steadily emerging, like real time targeted marketing (RTTM). One consequence is the increased need of new document engineering, composition and production technology which can be more effective for producing flexible, large-scale and high quality, multi channel fulfillment applications.

Quite a few on-going attempts for establishing such technologies have appeared in last two years; to succeed it is necessary to develop a high quality document engineering framework that allows the integration with variable data from enterprise resource planning systems (ERP) or customer relationships management systems (CRM) to deliver personalized print materials on a cost-effective way. The personalized printing markup language PPML[4] and the PPML templates (PPML/T)[5] are in our view one of the most promising technologies to accomplish that.

To support our customers needs in terms of the degree of customization, quality of the resulting materials, productivity and cost it is essential to count with efficient and reliable fulfillment technologies like the last generation of digital presses and digital front ends (DFE); also a complete solution should support several fulfillment mechanisms and channels for the production of the personalized materials, like high volume presses, low-end printers, email, fax or video flat panel displays.

During the past year, an end to end solution for the retail market, the retail production system (RPS),[7] was developed by our team in the Hewlett-Packard Digital Publishing Solutions Lab. During the course of this development a VDP architecture that is efficient and can be adapted to a broad set of workflows has evolved.

2 The Retail Promotional Solution

The retail promotional solution (RPS), is a highly customizable end-to-end RTTM solution for the retail industry that integrates the business cycles with the business management and automating the production of high impact graphical customized Point Of Purchase (PoP) materials and flyers.

RPS helps retailers use digital assets, innovative printing technology and in-store floor space to: (a) increase sales through impulse buying, (b) improve the buying experience for their customers (c) increase revenue by renting communication opportunities to manufacturers and (d) differentiate from their competitors.

RPS integrates the printing material production cycles with the business management systems (ERP), facilitating campaign management and automating the production of high impact graphical customized point of purchase (PoP) materials. RPS allows retailers to centrally define and design the materials; the system will then automatically produce the customized documents for each store, taking into account its own promotions, prize zone, inventory, language, etc. The materials are transmitted, composed and produced on-demand by a distributed network of print stations connected to the system.

A Print Station (PS) can be deployed in the store, in the corporate offices or even in outsourced print service providers (PSP) facilities. Supported PS range from fashionable flat panel displays on the store to high volume production digital presses.

In order to fulfill the demanding requirements of the RPS system, a VDP engine based on PPML/T was developed; this system is described in the next sections.

3 PPML Templates

The personalized printing markup langue (PPML) is standard promoted by the PODi digital printing initiative, the purpose of which is to make much faster to process documents that have reusable content. PPML is an open

standard supported by several vendors of digital printing applications and devices.

PPML describes the layout of a document in terms of addressable objects, each of the objects can be represented using different formats, PPML itself does not specify a content format, it provides metadata about the content structure and layout, the content formats supported will depend on the consumer.

PPML templates or PPML/T enable very long runs of documents with personalized content be created. PPML/T is an XSLT script that can be applied to a stream of data producing as a result a PPML document. PPML/T documents contain the design and rules to apply to data in order to generate an actual personalized document.

PPML/T documents can be easily merged creating a new XSLT script that includes the original ones.

Because the PPML document engineering architecture is entirely XML[2] based all the content and structure can be generated, merged, manipulated and processed using standard tools.

4 Design Paradigms

4.1 Fundamental Features

An underlying design philosophy of the VDP engine was to develop a set of VDP facilities for application developers that isolate them of the details of PPML technologies while providing flexible and efficient support for different workflows and offering high level interfaces.

Fundamental features such middleware should provide are:

1. a business rules execution engine
2. document generation, merging and imposition according to destination needs
3. distribution and delivery to the digital front ends
4. support from digital front ends capable of generate output for a wide range devices

At the same time the internal architecture of the engine should be designed to be flexible enough to accommodate different workflows, business processing rules or digital asset managements systems.

All this needs to be provided within a high performance environment, reducing the usage of computing and networking resources while keeping up with the demanding requirements of the applications. Counting with such infrastructure should greatly reduce the developments and integrations costs of newer or existing applications.

4.2 Distributed Print Station Network

A print station (PS) represents a fulfillment facility, usually a DFE, where different types of devices (printers, presses or flat panel displays) can be attached, and different kinds of workflows can be executed.

Depending on the PS characteristics and location different formats and delivery protocols might be used; also the personalized materials themselves might have differences depending on the destination PS; for instance in a mass production material we might need to include cutting crop marks, while for a flat panel display we might want to include multimedia assets like video or an audio track. The middleware will take care of all those issues.

The diagram in Figure 1 shows how an application uses a network of PS for the fulfillment and delivery. Materials can be generated for instance locally at the destination by thin PS, in regional or main headquarters using clusters of midrange devices, or centralized in print service providers facilities that have digital presses. Depending on the application requirements different topologies provide the maximum return of investment for the customer.

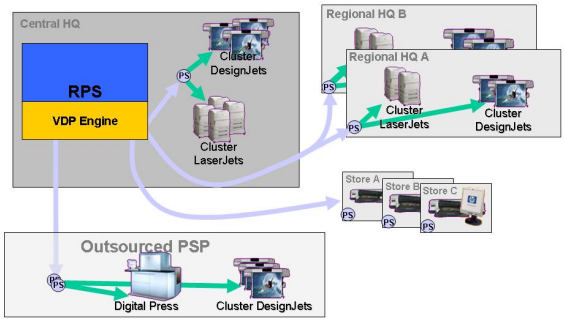

Figure 1 Distributed Print Stations

5 Middleware Architecture

The VDP middleware can be described as a pipeline with four layers as show in Figure 2, from the devices to the enterprise applications like ERPs or CRMs. Each stage in the pipeline is implemented by a component that provides a well defined set of interfaces, so applications can integrate with the lower layers. In this section the components and their main features are briefly discussed.

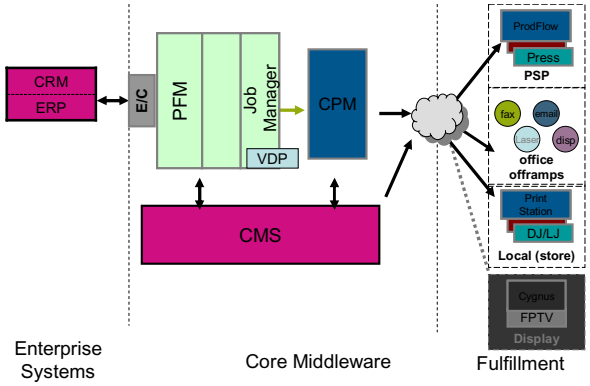

Figure 2. Architecture

5.1 Business Rules Execution

The Print Flow Manager (PFM) component is responsible of the business rules execution. This component

receives raw enterprise campaign data, in the form of XML, and processes it before submitting it to the job manufacturing component, the Job Manager.

The PFM can be configured with a set of predefined print flows; a print flow is a set of filters or flow actions that need to be processed over the stream of enterprise data. Providing different flows of filtering actions allows applications to apply different business rules depending on the type of promotional materials to generate. The flow actions executed for a personalized direct marketing campaign (PDM) can be different from the required for a centralized production of PoP materials. A print flow is implemented as a configurable SAX filter chain; SAX events are passed from one filter event handler to the other, making the process very efficient. The rules execution task might need to process millions of data records; therefore, the design must be able to process the stream of data as it comes from the enterprise systems. Each one of the filters acts on the data, and applies some business rule, for instance a filter can apply target language restrictions, providing translations to the target's language.

Filters can be developed in XSLT or Java languages, and appended to a print flow by system integrators, according to the application needs.

5.2 Job Manager

The Job Manager module provides support for the generation of VDP jobs. Within the job manager there are two primary tasks performed:

(1) Job Grouping

Job elements that are part of a campaign need to be grouped in chunks that we can fulfill together to support its manufacturability. There are business decisions to take into account, like having job elements that need to be delivered to different places, or have different priority levels become part of different jobs.

But there are also limitations of the print stations, related to the printing and imposition processes, like the maximum number of pages in a job, this will force some elements to be grouped in distinct chunks.

The system adopts a two level filtering approach to perform the grouping of job elements. The first level groups the elements that need to be part of different chunks, the second level groups the elements that will be part of different sections within a job, this is the case of a promotion where materials need to be delivered to different destinations be grouped in sections within the same job.

If needed, the grouping filters can be extended and configured by application integrators to implement the required logic for job manufacturing.

(2) Job Generation

This task is executed by the VDP Processor, and generates a PPML/T document that includes the job elements that are part of the same chunk. The processor will take into account the destination print station properties to generate and impose a document that merges each one of the PPML/T templates used in the individual job elements.

The process then adds the VDP data and assets and the result is a PPML/T document with its data section filled. The VDP processor relies in a core component, the PPML/T library, this library provides object addressability for PPML/T documents, and allows the processor to manipulate, merge and impose templates.

Job grouping and job generation are task intensive processes, that need to process from few elements for collateral on demand printing to millions of elements for personalized direct marketing applications, therefore it is important to count with scalable design.

5.3 Commercial Print Manager

The commercial print manager (CPM) module is responsible of the delivery of the jobs the print stations. For each print station a delivery service is defined; the delivery services are pluggable components that implement the mechanism for the transformation and delivery to the print stations of the PPML/T documents. The following are some of the delivery services provided for VDP applications:

(a) The remote printing protocol (RPP) delivery service is used for delivery to the HP Production Flow digital front ends that feed the HP Indigo presses. A job definition format (JDF)[1] ticket that describes the intent of the job is generated based on the characteristics of the elements in the job and the print station. The ticket, that includes references to the required assets to process the job are delivered using the remote printer protocol (RPP) which is based on JMF messages.

(b) A PDF[6] based delivery is provided to deliver jobs to thin print station. The PPML/T document is translated into PDF file which can be processed by devices with fewer resources.

Other delivery services include an export to folder delivery service or a service for the delivery jobs to flat panel display based print stations.

The Print Manager also provides application programming interfaces for (i) creation and setup the print station capabilities, such as the delivery service used, (ii) global print station queue management (resubmit, change priority, move or pause jobs) (iii) routing between print stations and (iv) job status and print station monitoring.

5.4 Print Stations and Production Flow

Fulfillment of high volumes is efficiently produced by DFE driven print stations like the HP Production Flow.[8] HP Production flow is a DFE that supports PPML/T workflows and can drive different devices like digital presses or clusters of laser or large format printers.

Production Flow uses deferred variable data printing and massive asset caching to greatly optimize network bandwidth. Intelligent auto-rotation and imposition/nesting is used to optimize printing and save paper.

Print Stations use encrypted communications and a pull mode transport protocol to communicate across firewalls, the RPP which is a job messaging format (JMF) based protocol where JDF tickets are exchanged.

The benefit is that PS may be deployed anywhere: in the retailer stores, corporate print centers or even outside the enterprise network, in outsourced PSP. The Print Station network can be arranged to fit the customers specific requirements to balance outsource vs. in-house, distribute-and-print vs. print-and-distribute, etc.

6 Future Research

The RPS is a multi-channel initiative, with a strong move into personalization and customization enabled by online technology. In the near future RPS will move beyond stoppers as customer's digital asset management capability expands, and will provide online catalogs with one-to-one personalization, through email, web and printed version output on digital presses.

Preparation, selection and composition of multiple delivery channels will be the next step in the VDP engine.

To accomplish that, newer extensions to the document description framework need to be developed. Although such framework should still benefit from current developed technologies like PPML, XSL/FO,[11] PDF or SVG,[9] more needs to be added, like the capability to relate different parts of the document, flowing objects between pages or include metadata to make template reengineering and multi channel fulfillment more efficient.[12]

References

1. JDF Specification, Release 1.1, CIP4, http://www.cip4.org
2. Tim Bray , “XML Specification”, W3C,February 18 1998
3. James Clark, “XSLT Transformations (XSLT)”, W3C, November 16 1998
4. ”PPML Functional Specification v 2.1”, The PPML Working Group. PODi., http://www.podi.org
5. ”PPML Template Methods and Workflows”, The PPML Working Group. PODI, http://www.podi.org
6. PDF Reference Third Edition, v1.5, Adobe Systems Incorporated.
7. Luca Chiarabini, “A PPML/T Variable Printing Solution” Proceeding DPP2003, Barcelona.
8. Jody Terrill, Loay Abu-Husein, James Sangroniz, “Rainbow MOLR Internal Development Overview” March 5 2003
9. Jon Ferraiolo, “Scalable Vector Graphics (SVG) 1.1 Specification”, W3C, January 14 2003
10. Andrew H. Watt, “Designing SVG Web Graphics”, New Riders
11. Dave Pawson, “XSL-FO”, O’Reilly and Associates, August 15, 2002
12. Fabio Gianetti “FOA: an XSL-FO Authoring tool”, HP Laboratories Bristol, May 25, 2001

Biography

Galo Gimenez is a senior software engineer at Hewlett Packard, Real Time Targeted Marketing Enterprise Solutions lab, where middleware software for variable data printing is being developed. He holds a BSc in Computer Sciencie by the Universitat Politecnica de Catalunya (UPC) and a MSc in Electrical and Computer Engineering by the University of California at Irvine(UCI).